به نام خداوند جان و خرد

نظریه احتمال پیشرفته

((مبتنی بر نظریه اندازه))

گردآورنده:

دکتر احمد پوردرویش

عضو هیأت علمی دانشگاه مازندران

زمستان ۱۳۹۷

صفحه اطلاعات فیپا

فهرست مطالب

پیشگفتار

از آنجا که کتاب جامعی به زبان فارسی برای دوره کارشناسی ارشد و دکتری آمار در دسترس نبود، این کتاب با بهره گیری از منابعی که ذکر شده در دست اقدام قرار گرفت.

کتاب حاضر محتوی ۱۷ فصل می‌باشد، که ۱۱ فصل اول منطبق بر سر فصل دوره کارشناسی ارشد وزارت علوم می‌باشد و ۶ فصل آخر کتاب منطبق بر سر فصل دوره دکتری آمار است.

این کتاب حاصل چندین دوره تدریس سالهای متوالی در دوره ارشد و دکتری در دانشگاه مازندران می‌باشد و عمدتاً اقتباس از کتاب های بات و پاتریک بیلینگزلی بوده و هر فصل شامل چند تمرین می‌باشد. تمرین‌های بیشتر از جمله، پاسخ سوالات چندین دوره ورودی دکتری دانشگاه‌های مختلف نیز تهیه شده که متعاقباً برای استفاده علاقمندان و مخصوصاً مدرسان این درس‌ها در نظر گرفته شده که جداگانه به چاپ می‌رسد.

با احتمال یک اشکالاتی در متن از نظر دور مانده‌اند که قصور اینجاب و ارسال نظرات ارزشمند صاحب‌نظران به ایمیل اینجانب به نشانی زیر موجب امتنان است

a.pourdarvish@umz.ac.ir

در پایان از همه دانشجویان تحصیلات تکمیلی اینجانب که در وارد نمودن برخی از این اشکالات همت گماشتند، تشکر می‌نمایم.

با آرزوی توفیق

دکتر احمد پوردرویش

فصل اول

مجموعه‌ها و کلاس پیشامدها

۱–۱ مقدمه

عموماً برآمدی دلخواه از یک آزمایش تصادفی را با ω نشان می‌دهیم و آن را یک نقطه نمونه می‌نامیم. مجموعه تمام برآمدهای ω را با Ω نمایش داده و آن را فضای نمونه گوییم. گردایه‌ای از برآمدهایω که مورد توجه هستند یک پیشامد نامیده می‌شود. پس یک پیشامد زیرمجموعه‌ای از Ω می‌باشد. (این تعریف چنان چه متعاقباً در بحث پیشرفته خواهد آمد زمانی برقرار است که میدان سیگما تعریف شده روی Ω، مجموعه توانی باشد.) پیشامدها را با حروف بزرگ لاتین A , B , C نشان خواهیم داد.

پیشامدی که تنها از یک نقطه نمونه‌ای همانند ω_0 تشکیل شود، بصورت $\{\omega_0\}$ نشان داده می‌شود که یک پیشامد تک عضوی است و به آن پیشامد ساده نیز می‌گویند.

زیرمجموعه‌ای از Ω که شامل هیچ نقطه‌ای نباشد، یعنی مجموعه $\emptyset$، پیشامد محال گوییم و همچنین تمام فضای نمونه، یعنی مجموعه Ω، را پیشامد حتمی گوییم.

۱–۲ جبر مجموعه‌ها

اعمال مهم مجموعه‌ها عبارت‌اند از:

متمم‌گیری: به هر مجموعه A می‌توانیم مجموعه دیگری مانند A^c، که شامل تمام نقاطی از Ω است که متعلق به A نیست، نسبت دهیم. A^c را متمم A می‌گوییم. بطور نمادی یعنی

$$A^c = \{\omega : \omega \notin A\}$$

واضح است که: $\Omega^c = \phi \quad , \phi^c = \Omega$

جزئیت: اگر تمام اعضای A متعلق به مجموعه B باشد می‌گوییم A یک زیرمجموعه B یا A جز B است. بطور نمادی یعنی

$$A \subset B \iff (\omega \in A \implies \omega \in B)$$

تساوی: اگر $A \subseteq B$ و $B \subseteq A$ آنگاه A و B را مساوی گوییم و می‌نویسیم: $A = B$.

پس برای اثبات تساوی دو مجموعه A و B باید ثابت کنیم:

$$(\omega \in A) \iff (\omega \in B)$$

اجتماع و اشتراک: اگر A و B دو مجموعه باشند، در آن صورت مجموعه تمام نقاط ω که به A یا B تعلق دارند را اجتماع A و B نامیده و آن را با $A \cup B$ نشان می‌دهیم. مجموعه تمام نقاط که به هر دو مجموعه A و B تعلق دارند را اشتراک A و B نامیده و آن را با $A \cap B$ نشان می‌دهیم. بطور نمادی

$$A \cup B = \{\omega : \omega \in A \ \ or \ \ \omega \in B\}$$

$$A \cap B = \{\omega : \omega \in A \ \ \& \ \ \omega \in B\}$$

نکته ۱-۱: مجموعه‌های A و B را دو مجموعه جدا از هم یا ناسازگار گوییم هرگاه $A \cap B = \phi$. در این حالت و فقط در این حالت $A \cup B$ را با $A + B$ نشان خواهیم داد، پس $A^c + A = \Omega$.

نکته ۱-۲: مجموعه $A \cap B^c + B \cap A^c = A\Delta B$ را تفاضل متقارن A و B و عملگر Δ را عملگر تفاضل متقارن گوییم.

لم ۱-۱: اگر گردایه $\{A_i, i = 1,2, \dots, n\}$ شامل n زیر مجموعه مفروض باشد، آنگاه گردایه‌ای مانند $\{B_i, i = 1,2, \dots, n\}$ از مجموعه‌های جدا از هم وجود دارد به قسمی که

$$\bigcup_{i=1}^{n} A_i = \sum_{i=1}^{n} B_i$$

اثبات: با استقرا ثابت می‌کنیم. می‌دانیم

$$A \cup B = A \cup (B - A)$$

بنابراین داریم:

$$A_1 \cup A_2 = A_1 + A_1^c A_2 = B_1 + B_2$$

که در آن B_1 و B_2 جدا از هم هستند. پس لم برای $n = 2$ درست است. فرض می‌کنیم برای $n = m$ برقرار باشد، حال برای $n = m + 1$ ثابت می‌کنیم

$$\bigcup_{i=1}^{m+1} A_i = \left(\bigcup_{i=1}^{m} A_i\right)\bigcup A_{m+1} = \left(\sum_{i=1}^{m} B_i\right)\bigcup A_{m+1}$$
$$= \sum_{i=1}^{m} B_i + \left(\sum_{i=1}^{m} B_i\right)^c A_{m+1}$$
$$= \sum_{i=1}^{m} B_i + B_{m+1}$$

که در آن B_{m+1} و B_i ،$i = 1,2,\dots,m$ جدا از هم هستند و در نتیجه B_{m+1} و $\sum_{i=1}^{m} B_i$ جدا از هم هستند، پس لم به ازای $n = m + 1$ صادق است و به موجب اصل استقرا به ازای هر n برقرار هست.

در ادامه یکی از نتایج لم ۱-۱ آورده شده است که بیان می‌کند چگونه از یک گردایه شمارا از مجموعه‌های دلخواه، گردایه‌ای از مجموعه‌های جدا از هم بسازیم به قسمی که اجتماع آنها برابر باشد.

نتیجه ۱-۱:

این نتیجه، تعمیم لم ۱-۱ برای تعداد شمارا مجموعه است.

$$\bigcup_{i=1}^{\infty} A_i = A_1 + A_1^c A_2 + A_1^c A_2^c A_3 + \cdots$$

۱–۳ دنباله ها

تعریف ۱-۱: فرض کنیم $\{A_n\}$ دنباله‌ای از مجموعه‌ها باشد، اگر به ازای هر n، $A_n \subset A_{n+1}$، آنگاه دنباله را صعودی یکنوا می‌گوییم و مجموعه $A = \bigcup_{n=1}^{\infty} A_n$ را حد دنباله نامیده و آن را با نماد $A_n \uparrow A$ نشان می‌دهیم. اگر به ازای هر n، $A_{n+1} \subset A_n$ باشد آنگاه دنباله $\{A_n\}$ را نزولی یکنوا گوئیم و مجموعه $A = \bigcap_{i=1}^{\infty} A_n$ را حد دنباله $\{A_n\}$ نامیده و آن را با نماد $A_n \downarrow A$ نشان می‌دهیم.

نکته ۱-۳: اگر $A_n \uparrow A$ آنگاه به موجب قاعده دمورگان نتیجه می‌شود

$$A_n^c \downarrow A^c$$

اثبات:

$$limA_n^c = \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m^c = (\bigcup_{m=1}^{\infty} A_m^c)\bigcap(\bigcup_{m=2}^{\infty} A_m^c)\bigcap \dots \bigcap(\bigcup_{m=n}^{\infty} A_m^c)\bigcap \dots$$

$$= \left(\bigcap_{m=1}^{\infty} A_m\right)^c \bigcap \left(\bigcap_{m=2}^{\infty} A_m\right)^c \bigcap \dots \bigcap \left(\bigcap_{m=n}^{\infty} A_m\right)^c \bigcap \dots$$

$$= \left(\left(\bigcap_{m=1}^{\infty} A_m\right) \bigcup \left(\bigcap_{m=2}^{\infty} A_m\right) \bigcup \dots \bigcup \left(\bigcap_{m=n}^{\infty} A_m\right) \dots \right)^c$$

$$= \left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m\right)^c = \left(limA_n\right)^c = A^c$$

تا اینجا حد دنباله‌ای از مجموعه‌های یکنوا را تعریف کردیم. حال آن را برای یک دنباله دلخواه تعریف می‌کنیم. برای هر دنباله $\{A_n\}$ تعریف می‌کنیم

$$C_n = \sup_{k\geq n} A_k = \bigcup_{k=n}^{\infty} A_k$$

$$B_n = \inf_{k\geq n} A_k = \bigcap_{k=n}^{\infty} A_k$$

در اینصورت $\{B_n\}$ یک دنباله صعودی یکنوا با حد زیر است:

$$B = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k = lim\ \inf_n A_n = \underline{lim}\, A_n$$

و $\{C_n\}$ یک دنباله نزولی یکنوا با حد زیر است:

$$C = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k = \lim \sup_n A_n = \overline{lim}\, A_n$$

واضح است که B مجموعه تمام نقاطی است که به کلیه A_n ها مگر تعداد متناهی از آنها تعلق دارند. لذا C مجموعه نقاطی است که به تعداد بی‌نهایت از A_n ها متعلق باشند. بنابراین اگر ω به تمام A_n ها مگر تعداد متناهی از آنها متعلق باشد به بی‌نهایت از آنها متعلق است در نتیجه:

$$\underline{lim}A_n \subseteq \overline{lim}A_n$$

۱-۳-۱ تعریف حد دنباله

تعریف ۱-۲: فرض کنیم $\{A_n\}$ دنباله‌ای از زیر مجموعه‌ها در Ω باشد در این صورت می‌گوییم دنباله $\{A_n\}$ دارای حد است هرگاه $\underline{lim}A_n = \overline{lim}A_n = A$. در این صورت می‌نویسیم $A_n \to A$ و می‌گوییم دنباله $\{A_n\}$ به A همگراست. باید توجه داشت که $\underline{\lim}A_n$ و $\overline{\lim}A_n$ همواره وجود دارند ولی حد $\{A_n\}$ ممکن است موجود نباشد.

مثال ۱-۱: اگر $\begin{cases} A_{2n} &= A \\ A_{2n+1} &= B \end{cases}$ آنگاه حدود بالایی و پایینی را بیابید. $(A \cap B = \phi)$

$$\forall\, n = 1 \implies \bigcap_{k\geq 1} A_k = A_1 \cap A_2 \cap A_3 \ldots = \phi$$

$$\forall\, n = 2 \implies \bigcap_{k\geq 2} A_k = A_2 \cap A_3 \cap A_4 \ldots = \phi$$

$$\vdots$$

$$\implies \bigcup_{n=1}^{\infty} \bigcap_{k\geq n} A_k = \underline{lim}A_n \quad = \lim\inf_n A_n = \phi$$

همچنین داریم:

$$\forall\, n = 1 \implies \bigcup_{k\geq 1} A_k = A_1 \cup A_2 \cup A_3 \ldots = A \cup B$$

$$\forall\, n = 2 \implies \bigcup_{k\geq 2} A_k = A_2 \cup A_3 \cup A_4 \ldots = A \cup B$$

$$\vdots$$

$$\implies \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k = \overline{lim}A_n \quad = \lim\sup_n A_n \quad = A \cup B$$

۱-۴ میدان، میدان مینیمال و افراز

فرض کنید Ω یک مجموعه ناتهی دلخواه باشد کلاس $\mathcal{F}$ از زیر مجموعه‌های Ω یک میدان نامیده می‌شود هرگاه هم

شامل Ω باشد و هم تحت اعمال متمم‌گیری و اجتماع‌های متناهی بسته باشد.

(۱) $\Omega \in \mathcal{F}$

(۲) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

(۳) $A_1, A_2, \dots, A_n \in \mathcal{F} \implies \bigcup_{i=1}^{n} A_i \in \mathcal{F}$

بدیهی است گردایه $\mathcal{F}$ تحت اشتراک متناهی نیز بسته است اگر

$$A_1, A_2, \dots, A_n \in \mathcal{F} \implies A_1^c, A_2^c, \dots, A_n^c \in \mathcal{F}$$
$$\implies \bigcup_{i=1}^{n} A_i^c \in \mathcal{F}$$
$$\implies \left(\bigcup_{i=1}^{n} A_i^c\right)^c \in \mathcal{F} \implies \bigcap_{i=1}^{n} A_i \in \mathcal{F}$$

نکته ۱-۴: هر میدان شامل مجموعه تهی و کل فضا Ω است. (با توجه به خواص ۱و۲)

نکته ۱-۵: گردایه که فقط شامل ϕ و Ω باشد یک میدان است. این کلاس کوچکترین میدان و زیرمجموعه هر میدان دیگر است و به آن یک میدان بدیهی یا تباهیده می‌گویند.

نکته ۱-۶: مجموعه توانی یعنی گردایه‌ی تمام زیر مجموعه‌های Ω بزرگ‌ترین میدان است.

مثال ۱-۲: کلاس $\{A, A^c, \Omega, \phi\}$ یک میدان است که زیر مجموعه هر میدان شامل A خواهد بود. لذا کوچک‌ترین میدان شامل A است.

مثال ۱-۳: فرض کنید A_3, A_2, A_1 به قسمی باشند که:

$$A_i \cap_{i \neq j} A_j = \phi \quad , \sum_{i=1}^{3} A_i = \Omega , \forall i, j = ۱,۲,۳$$

پس Ω توسط سه مجموعه دو به دو جدا از هم تقسیم شده است. برای تعیین میدانی که شامل گردایه $\{A_1, A_2, A_3\}$ باشد کافی است به گردایه فوق مجموعه‌های $A_1 + A_3$, $A_1 + A_2$, $A_2 + A_3$, Ω , ϕ را اضافه کنیم. چون $A_2 + A_3 = A_1^c , A_1 + A_3 = A_2^c , A_1 + A_2 = A_3^c$

پس کلاس حاصل عبارت است از:

$$\{\phi , A_1, A_2, A_3, A_1 + A_2, A_2 + A_3, A_1 + A_3 \ \ , \Omega \}$$

که تحت متمم‌گیری و اجتماع متناهی بسته بوده و در نتیجه یک میدان است. همچنین کوچک‌ترین میدان شامل A_3, A_2, A_1 است.

تعریف ۱-۳: کلاس دلخواه $\mathcal{A}$ از زیر مجموعه‌های Ω را در نظر بگیرید. کوچکترین میدان شامل $\mathcal{A}$ را میدان مینیمال شامل $\mathcal{A}$ یا میدان تولید شده توسط $\mathcal{A}$ می‌نامیم و این میدان جزء هر میدان دیگر شامل $\mathcal{A}$ است.

قضیه ۱-۱: اشتراک هر تعداد دلخواهی از میدان‌ها، یک میدان است.

اثبات: فرض می‌کنیم $\mathcal{F} = \bigcap_{\theta \in I} F_\theta$ که I دلخواه بوده و F_θ میدان است،

۱) $\Omega \in F_\theta \ , \forall\theta \implies \Omega \in \bigcap_\theta F_\theta$

۲) $A \in \mathcal{F}, \forall\theta \implies A \in F_\theta \implies A^c \in F_\theta \implies A^c \in \bigcap_\theta F_\theta$

۳) $A_i \in \mathcal{F}; \forall\theta, i \implies A_i \in F_\theta \implies \bigcup_{i=1}^{n} A_i \in F_\theta$

$$\implies \bigcup_{i=1}^{n} A_i \in \bigcap_{\theta} F_\theta = \mathcal{F}$$

نکته ۱-۷: اجتماع میدان‌ها لزوماً یک میدان نیست. به عنوان مثال A وB را طوری در نظر بگیرید که هیچ کدام زیر مجموعه هم نباشند، مثلاً $\{A, A^c, \Omega, \phi\}$ و $\{B, B^c, \Omega, \phi\}$ میدان هستند ولی اجتماع آنها یعنی $\{A, A^c, B, B^c, \Omega, \phi\}$ میدان نمی‌باشد.

تعریف ۱-۴: فرض کنید $\{A_i\}, i = 1,2, \dots, n$ گردایه‌ای از مجموعه‌های جدا از هم باشد به طوری که $\sum_{i=1}^{n} A_i = \Omega$ و $\{A_i\}$ دو به دو جدا از هم هستند. در این صورت $\{A_i\}$ را یک افراز Ω گوییم. بنابرین داریم

$$A_1^c = A_2 + A_3 + \cdots + A_n \quad , (A_1 + A_2)^c = A_3 + A_4 + \cdots + A_n \ \ , \dots$$

و الی آخر، پس کلاس

$$\{\phi, A_1, A_2, A_3, \dots, A_n, A_1 + A_2, \dots, A_{n-1} + A_n, \dots, A_1 + A_2 + \cdots + A_n = \Omega\}$$

که شامل ϕ،Ω، اجتماع A_i ها یک به یک، دو به دو و غیره بوده و در نتیجه تحت اجتماع متناهی بسته است. تحت متمم نیز بسته بوده و بنابراین یک میدان خواهد بود. در واقع یک میدان مینیمال شامل مجموعه‌های $A_1, A_2, \dots A_n$ می‌باشد که تعداد مجموعه‌های آن برابر است با:

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n}=2^n$$

تعریف ۱-5: افراز $\acute{\mathcal{A}}$ را ظریف‌تر از افراز $\mathcal{A}$ گوییم هرگاه میدان مینیمال $\mathcal{F}(\acute{\mathcal{A}})$ که شامل $\acute{\mathcal{A}}$ است، میدان مینیمال $\mathcal{F}(\mathcal{A})$ که شامل $\mathcal{A}$ است را در برگیرد. در این صورت $\mathcal{A}$ را درشت‌تر از $\acute{\mathcal{A}}$ گوییم.

مثلاً $\{A, A^c\}$ درشت‌تر از $\{A, BA^c, (B \cup A)^c\}$ می‌باشد ولی افراز $\{AB, AB^c, BA^c, (B \cup A)^c\}$ ظریف‌تر از هر دوی آنها است. برای اثبات می‌توان میدان مینیمال متناظر با هر یک از افرازها را بدست آورده وسپس با مقایسه آنها به نتیجه مطلوب دست یافت.

۱–۵ سیگما میدان و میدان بورل

گردایه ناتهی $\mathcal{F}$ از زیر مجموعه‌های Ω را یک سیگما میدان می‌گوییم اگر اولاً دارای خواص میدان باشد و ثانیاً نسبت به اتحاد و یا اشتراک شمارا بسته باشد:

۱) $\Omega \in \mathcal{F}$

۲) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

۳) $A_1, A_2, \ldots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

اگر گردایه‌ای نسبت به اتحادهای شمارا بسته باشد آنگاه حتماً نسبت به اتحادهای متناهی نیز بسته است.

توجه کنید:

۱. یک میدان لزومی ندارد که یک سیگما میدان باشد. به همین دلیل یک میدان را گاهی اوقات یک سیگما میدان به طور متناهی جمع‌پذیر می‌نامند.

۲. یک میدان تحت عملگرهای متناهی و یک سیگما میدان تحت عملگرهای شمارا بسته است.

۳. اگر میدان $\mathcal{F}$ فقط شامل تعداد متناهی مجموعه باشد آن یک سیگما میدان خواهد بود ولی یک میدان شامل تعداد نامتناهی مجموعه ممکن است یک سیگما میدان نباشد.

قضیه ۱-۲: اشتراک هر تعداد سیگما میدان یک سیگما میدان است.

اثبات: فرض می‌کنیم $\mathcal{F} = \bigcap_{\theta\in I} \mathcal{F}_\theta$ که I دلخواه بوده و $\mathcal{F}_\theta$سیگما میدان است،

۱) $\Omega \in \mathcal{F}_\theta, \forall\theta \in I \implies \Omega \in \bigcap_{\theta\in I} \mathcal{F}_\theta = \mathcal{F}$

۲) $A \in \mathcal{F} \Rightarrow A \in \mathcal{F}_{\theta}\ ,\forall \theta \in I \ \Rightarrow\ A^c \in \mathcal{F}_{\theta}\ \ ,\forall \theta \in I$

$$\Rightarrow \bigcup_{i} A_i \in \bigcap_{\theta \in I} \mathcal{F}_{\theta} = \mathcal{F}$$

نکته ۱-۸: اجتماع سیگما میدان‌ها لزوماً سیگما میدان نیست.

نکته ۱-۹: $\sigma(\mathcal{A})$ یا کوچکترین سیگما میدان شامل $\mathcal{A}$، دارای سه ویژگی زیر است:

الف) $\mathcal{A} \subset \sigma(\mathcal{A})$

ب)$\sigma(\mathcal{A})$ یک سیگما میدان می باشد.

ج) اگر $\mathcal{A} \subset \mathcal{G}$ و $\mathcal{G}$ یک سیگما میدان باشد، آنگاه $\sigma(\mathcal{A}) \subset \mathcal{G}$.

قسمت (ج) بیان می‌کند که $\sigma(\mathcal{A})$ کوچکترین سیگما میدان حاوی $\mathcal{A}$ و یا اشتراک تمام سیگما میدان‌های شامل $\mathcal{A}$ است و $\mathcal{G}$ یکی از سیگما میدان‌های شامل $\mathcal{A}$ است اما لزوماً کوچکترین آنها نیست.

توجه کنید:

_ اگر $\mathcal{F}$ یک سیگما میدان باشد آنگاه $\sigma(\mathcal{F}) = \mathcal{F}$.

_ اگر$\mathcal{A} = \{\phi\}$ یا $\mathcal{A} = \{\Omega\}$ آنگاه $\sigma(\mathcal{A}) = \{\phi\ ,\Omega\}$.

_ اگر $\mathcal{A} \subset \acute{\mathcal{A}}$ آنگاه $\sigma(\mathcal{A}) \subset \sigma(\acute{\mathcal{A}})$.

تعریف ۱-۶: کلاس C شامل تمام بازه‌های $x \in R$, $(-\infty\ , x)$ را به عنوان زیر مجموعه‌هایی از خط حقیقی R در نظر می‌گیریم. این کلاس تحت اشتراک متناهی بسته است ولی تحت متمم و اشتراک شمارا بسته نیست. فرض کنید $\sigma(C) = \mathcal{B}$ سیگما میدان شامل C باشد. در این صورت $\mathcal{B}$ شامل بازه‌هایی به صورت $[x\ ,\infty)$ است که مکمل مجموعه‌های $(-\infty\ , x)$ می‌باشند. همچنین شامل بازه‌هایی بصورت زیر است:

$$(-\infty, a] = \bigcap_{n=1}^{\infty}\left(-\infty, a+\frac{1}{n}\right)$$

$$(a, \infty) = (-\infty, a]^c$$

$$(a, b) = (-\infty, b) \cap (a, \infty) \quad , a < b$$

کلاس $\mathcal{B}$ را میدان بورل از زیر مجموعه‌های خط حقیقی گوییم و مجموعه‌های متعلق به $\mathcal{B}$ را مجموعه‌های بورل می‌نامیم.

لم ۱-۲: اگر C_1 کلاس تمام بازه‌هایی به شکل $(a,b) \quad , a<b$ باشد که در آن $a,b \in R$ دلخواه‌اند، آنگاه:

$$\sigma(C_1) = \mathcal{B}$$

اثبات: به ازای هر $a,b \in R$ داریم $(a,b) \in \mathcal{B}$. پس $C_1 \subset \mathcal{B}$. بنابراین طبق تعریف سیگما میدان مینیمال داریم $\sigma(C_1) \subseteq \mathcal{B}$.

حال باید ثابت کنیم $\mathcal{B} \subseteq \sigma(C_1)$. توجه کنید که داریم $\sigma(C_1)$ یک سیگما میدان از بازه‌های باز (a,b) است، بنابراین اجتماع شمارا از این بازه‌ها را دربردارد یعنی

$$\bigcup_{n=1}^{\infty}(-n,x) \in \sigma(C_1), \forall x \implies (-\infty,x) \in \sigma(C_1) \implies C \subset \sigma(C_1)$$

که در آن C همان گردایه معین شده در تعریف فوق می‌باشد. پس $\sigma(C) \subseteq \sigma(C_1)$ از طرفی می‌دانیم که $\sigma(C) = \mathcal{B}$ ، بنابراین باتوجه به رابطه فوق:

$$\sigma(C) = \sigma(C_1) = \mathcal{B}$$

به همین ترتیب می‌توان اثبات نمود که میدان بورل یک میدان مینیمال است که شامل هر یک ازگردایه‌های زیر می‌باشد:

$$C_2 = \{(-\infty,x], x \in R\}$$

$$C_3 = \{(a,b], a<b; a,b \in R\}$$

$$C_4 = \{[a,b], a<b; a,b \in R\}$$

$$C_5 = \{[a,b), a<b; a,b \in R\}$$

$$C_6 = \{(x,\infty), x \in R\}$$

پس میدان $\mathcal{B}$ شامل کلیه زیر مجموعه‌های R به صورت فوق و متمم، اجتماع و اشتراک شمارای آنها می‌باشد. ولی گردایه $\mathcal{B}$ شامل کلیه زیر مجموعه‌های R نیست.

مثال ۱-۴: اگر $x \in R$ مجموعه تک عضوی $\{x\}$ یک مجموعه بورل است، زیرا

$$\{x\} = (-\infty,x] \cap [x,\infty)$$

بنابراین چون هر زیر مجموعه شمارا از R اجتماعی شمارا از مجموعه‌های تک عضوی است، یک مجموعه بورل می‌باشد. پس $\{0,1,2,\ldots\}$ یک مجموعه بورل است ولی اجتماع ناشمارا از اینگونه مجموعه‌های تک عضوی، ممکن است یک مجموعه بورل نباشد.

تعریف ۱-۶: میدان $\mathcal{F}$ را میدان یکنوا گوییم اگر تحت اعمال یکنوا بسته باشد. اگر $\{A_n\}$ دنباله‌ای یکنوا از مجموعه‌های $\mathcal{F}$ باشد آنگاه داریم:

$$lim\, A_n \in \mathcal{F}$$

یعنی:

$$A_n \in \mathcal{F} \quad \text{و} \quad A_n \uparrow A \implies A \in \mathcal{F}$$

$$A_n \in \mathcal{F} \quad \text{و} \quad A_n \downarrow A \implies A \in \mathcal{F}$$

قضیه ۱-۳: هر سیگما میدان یک میدان یکنواست و برعکس.

اثبات: فرض کنید $\mathcal{F}$ یک سیگما میدان باشد و $A_n \in \mathcal{F}$.

اگر $A_n \uparrow A$ آنگاه $A = \bigcup_n A_n$ یک اجتماع شمارا از مجموعه‌های $\mathcal{F}$ است، بنابراین $A_n \in \mathcal{F}$. همینطور اگر $A_n \downarrow A$ آنگاه $A = \bigcap_n A_n$ اشتراک شمارا از مجموعه‌های $\mathcal{F}$ است، بنابراین $A \in \mathcal{F}$. در نتیجه $\mathcal{F}$ یک میدان یکنواست.

برعکس، فرض کنید $\mathcal{F}$ یک میدان یکنوا باشد، و $A_1, A_2, \ldots$ مجموعه‌هایی متعلق به $\mathcal{F}$ باشند. در این صورت $\bigcup_{k=1}^{n} A_k$ و $\bigcap_{k=1}^{n} A_k$ به $\mathcal{F}$ متعلق‌اند زیرا $\mathcal{F}$ یک میدان است و هم چنین دنباله‌هایی یکنوا هستند و حدود آنها یعنی $\bigcap_{k=1}^{\infty} A_k = \lim_{n\to\infty} \bigcap_{k=1}^{n} A_k$ به $\mathcal{F}$ متعلق هستند. پس $\mathcal{F}$ یک سیگما میدان است.

$$\implies A^c \in \cap_{\theta \in I} F_\theta = \mathcal{F}$$

$$\forall i\, , A_i \in \mathcal{F} \implies A_i \in F_\theta \quad , \forall \theta, i \implies \bigcup_i A_i \in F_\theta \ \ , \forall \theta, i$$

۱-۶ کلاس پیشامدها

Ω مجموعه‌ی تمام برآمدهای E است لذا اگر A یک پیشامد باشد A^C نیز یک پیشامد خواهد بود.

اگر A و B دو پیشامد باشند، $A \cup B$ بیانگر پیشامدی است که در آن حداقل یکی از پیشامدهای A یا B رخ می‌دهد، و $A \cap B$ پیشامدی است که در آن هر دو پیشامد A و B با هم رخ می‌دهند. پس $\mathcal{A}$ تحت اجتماع و اشتراک متناهی بسته است و یک میدان خواهد بود. اگر $A_1, A_2, \ldots, A_n$ دنباله‌ای از پیشامدها باشند، $\bigcup_{k=1}^{n} A_k$ پیشامدی است که نشان می‌دهد حداقل یکی از پیشامدهای A_i رخ می‌دهند و $\bigcap_{k=1}^{n} A_k$ پیشامدی است که در آن تمام پیشامدهای

A_i با هم رخ می‌دهند. چون اجتماع و اشتراک را می‌توان برای حالت n بی‌نهایت نیز تعمیم داد، اگر E دارای بی‌نهایت برآمد باشد طبیعتاً یک سیگما میدان خواهد بود. بنابراین، به هر آزمایش E می‌توانیم یک فضای تمام برآمدها، Ω و یک سیگما میدان از پیشامدها $\mathcal{A}$، اختصاص دهیم.

۱–۷ مسائل

۱- فرض کنید H مجموعه‌ای باشد که خارج $\mathcal{F}$ قرار می‌گیرد. از آنجایی که $\mathcal{F}$ یک میدان است نشان دهید که میدان تولید شده به وسیله $\mathcal{F} \cup \{H\}$ شامل مجموعه‌هایی از فرم زیر است:

$$(H \cap A) \cup (H^c \cap B) \quad , \quad A, B \in \mathcal{F}$$

۲- روابط زیر را ثابت کنید:

الف) $\lim_n sup(A_n \cap B_n) = \lim_n sup(A_n) \cap \lim_n sup(B_n)$

ب) $\lim_n sup(A_n \cup B_n) = \lim_n sup(A_n) \cup \lim_n sup(B_n)$

۳- اگر $A_n = \begin{cases} A & , n = 1,3,5, \ldots \\ B & , n = 2,4,6, \ldots \end{cases}$ حدود بالایی و پایینی را بیابید.

۴- اگر $A_n \longrightarrow A$ نشان دهید $A_n^c \longrightarrow A^c$

فصل دوم

متغیرهای تصادفی

۲–۱ توابع وتوابع معکوس

فرض می‌کنیم Ω فضای نمونه با نقاط نمونه‌ی ω باشد. گاهی مقادیری مانند $X(\omega)$ یا $f(\omega)$ به جای خود ω مورد توجه است.

۲–۱–۱ تابع نقطه‌ای و تابع مجموعه‌ای

تعریف ۲-۱: اگر شناسه‌های f نقاطی از فضای Ω باشند، یک تابع نقطه‌ای داریم و اگر شناسه‌های یک تابع مجموعه‌هایی از گردایه‌ای معین باشند، در آن صورت یک تابع مجموعه‌ای خواهیم داشت.

مثال ۲-۱: اگر $X(\omega) = \omega^{۲}$ باشد، در این صورت $X(\omega)$ تابعی یک به یک نیست، لذا داریم:

$$\begin{cases} \omega_{۱} = \ \ ۲ \\ \omega_{۲} = -۲ \end{cases} \Rightarrow X\left(\omega_{۱}\right) = ۴ = X\left(\omega_{۲}\right) \nRightarrow \ \omega_{۱} = \ \omega_{۲}$$

تعریف ۲-۲: تابعی مانند X بر فضای Ω به توی فضای Ω' را که به هر $\omega \in \Omega$ یک نقطه منحصر به فرد در Ω' نسبت دهد با $X(\omega)$ نشان می‌دهیم. $X(\omega)$ نقش ω تحت تابع X است، که آن را گاهی مقدار X در ω نیز می‌گویند. X همچنین یک تناظر بین نقاط Ω و نقاط Ω' برقرار می‌سازد. Ω را حوزه‌ی تعریف یا دامنه X ، Ω' را حوزه‌ی مقادیر آن و مجموعه $\Omega'' = \{X(\omega): \omega \in \Omega\}$ را که زیر مجموعه‌ای از Ω' است برد تابع X می‌گوییم. اگر $\Omega'' = \Omega'$ باشد می‌گوییم X تابعی از Ω به روی Ω' است، در غیر این صورت X تابعی از Ω به توی Ω' است.

مثال ۲-۲: فرض کنید $X(\omega) = \omega^{۲}$ و داریم:

$$\Omega'' = \left\{0,۱,\ ۴,\ \dots\right\}\ ,\Omega' = \left\{0,۱,۲,\ \dots\right\},\Omega = \left\{0,\ \pm۱,\ \pm۲\ ,\ \dots\right\}$$

بنابراین،

$$\begin{cases} X:\Omega \xrightarrow{\text{به توی}} \Omega' \\ X:\Omega \xrightarrow{\text{به روی}} \Omega'' \end{cases} \Rightarrow \Omega''=\{0,۱,۴,\ldots\}$$ برد تابع

و اگر $X(\omega)=\omega+2$ آنگاه X نگاشتی از Ω به روی Ω' است.

مثال۲-۳: اگر $\Omega=(-\infty<\omega<\infty)$ و $\Omega'=(0<\omega'<\infty)$ خط حقیقی باشند، در این صورت تابع $X(\omega)=e^{\omega}$ تابعی یک به یک از $\Omega \xrightarrow{\text{به روی}} \Omega'$ است.

۲-۱-۲ تابع معکوس

در حالت کلی اگر $B'\subset\Omega'$، مجموعه تمام نقاط Ω را به قسمی که $X(\omega)\in B'$، نقش معکوس B' به وسیله X می‌گوییم و آن را با $X^{-۱}(B')$ نمایش می‌دهیم:

$$X^{-۱}(B')=\{\omega: X(\omega)\in B'\},\quad B'\subset\Omega' \qquad (۱\text{-}۲)$$

در واقع در اینجا ما به هر تابع نقطه‌ایX، یک تابع مجموعه‌ای $X^{-۱}$ وابسته کردیم که در آن، $\mathfrak{B}'$ حوزه تعریف گردایه‌ای از زیر مجموعه‌های Ω' است و $\mathfrak{B}$ حوزه مقادیر گردایه‌ای از زیر مجموعه‌های Ω می‌باشد. $X^{-۱}$ را نقش معکوس (نگاشت معکوس) X می‌گوییم و می‌نویسیم:

$$X^{-۱}(\mathfrak{B}')=\{X^{-۱}(B'): B'\in\mathfrak{B}'\}$$

بدیهی است که:

$$X^{-۱}(\Omega')=\{\omega: X(\omega)\in\Omega'\}=\Omega$$

مثال ۲ -۴: تابع $X(\omega)=\omega^{۲}$ از R به R را در نظر بگیرید. فرض کنید B' مجموعه تمام نقاط واقع در بازه $(۱,۲)$ باشد، در این صورت:

$$X^{-۱}(B')=\left\{\omega: X(\omega)=\omega^{۲}\in B'\right\},\qquad B'\subset R$$

$$X^{-۱}(B')=\left(۱,\sqrt{2}\right)\cup\left(-\sqrt{2},-۱\right)$$

اگر دامنه و حوزه مقادیر X یکسان و برابر $(0,\infty)$ باشد، آنگاه

$$X^{-۱}(B')=(۱,\sqrt{2}).$$

تعریف۲ -۳: تابع حقیقی I_A را بر روی Ω به صورت زیر تعریف می‌کنیم:

$$I_A(\omega) = \begin{cases} ۱ & \omega \in A \\ 0 & \omega \in A' \end{cases} \quad (۲-۲)$$

تابع I_A را تابع نشانگر یا تابع شاخص مجموعه A می‌گوییم. برد تابع I_A مجموعه $\{۰, ۱\}$ می‌باشد. اگر $B \subset R$، آنگاه:

$$I_A^{-۱}(\omega) = \begin{cases} \phi & 0, ۱ \notin B \\ A & 0 \notin B, ۱ \in B \\ A^c & 0 \in B,\ ۱ \notin B \\ \Omega & 0,۱ \in B \end{cases}$$

بنابراین،

$$\sigma(A) = I_A^{-۱}(\mathfrak{B}) = \{\phi, A, A^c, \Omega\}$$

واضح است که، cI_A مقدار c را بر A و صفر را بر A^c اختیار می‌کند. در نتیجه:

$$(cI_A^{-۱})(\mathfrak{B}) = I_A^{-۱}(\mathfrak{B}) = \sigma(A)$$

قضیه ۲-۱: نگاشت معکوس دارای خواص زیر است:

الف) $X^{-۱}(A^c) = \left(X^{-۱}(A)\right)^c$

ب) $X^{-۱}(\cap A_\alpha) = \cap X^{-۱}(A_\alpha)$

ج) $X^{-۱}(\cup A_\alpha) = \cup X^{-۱}(A_\alpha)$

د) اگر $B \subset C \subset \Omega'$ آنگاه $X^{-۱}(B) \subset X^{-۱}(C)$

اثبات:

الف)

$$\omega \in X^{-۱}(A^c) \Leftrightarrow X(\omega) \in A^c \Leftrightarrow X(\omega) \notin A \Leftrightarrow \omega \notin X^{-۱}(A) \Leftrightarrow \omega \in \left[X^{-۱}(A)\right]^c$$

ب)

$$\begin{aligned} \omega \in X^{-۱}(\cap_\alpha A_\alpha) &\Leftrightarrow X(\omega) \in \cap_\alpha A_\alpha \\ &\Leftrightarrow X(\omega) \in A_\alpha, \quad \forall \alpha \end{aligned}$$

$$\Leftrightarrow \omega \in X^{-۱}(A_\alpha), \forall\alpha$$

$$\Leftrightarrow \omega \in \cap_\alpha X^{-۱}(A_\alpha)$$

ج) اثبات این قسمت مشابه قسمت قبل می باشد.

د)

$$X^{-۱}(B) = \{\omega: X(\omega) \in B\} \subset \{\omega: X(\omega) \in C\} = X^{-۱}(C) \Rightarrow X^{-۱}(B) \subset X^{-۱}(C)$$

توجه کنید اجتماع و اشتراک‌های فوق لزومی ندارد شمارا باشد.

قضیه ۲-۲: اگر $\mathcal{A}$ گردایه‌ای از زیر مجموعه‌های Ω و یک سیگما میدان باشد آن‌گاه گردایه‌ی $\mathfrak{B}$ شامل همه مجموعه‌هایی که نقش معکوس آنها به $\mathcal{A}$ متعلق است، نیز یک سیگما میدان می‌باشد.

اثبات: نشان می‌دهیم که $\mathfrak{B} = [B: X^{-۱}(B) \in \mathcal{A}]$ یک سیگما میدان است:

۱) $\Omega' \in \mathfrak{B}$

$\Omega \in \mathcal{A}, X(\Omega) = \Omega'$ یا $X^{-۱}(\Omega') = \Omega \in \mathcal{A} \quad \Rightarrow \Omega' \in \mathfrak{B}.$

۲) $B \in \mathfrak{B} \Rightarrow B^c \in \mathfrak{B}$

$$B \in \mathfrak{B} \Rightarrow X^{-۱}(B) \in \mathcal{A} \Rightarrow \left[X^{-۱}(B)\right]^c \in \mathcal{A} \quad \Rightarrow X^{-۱}(B^c) \in \mathcal{A} \quad \Rightarrow B^c \in \mathfrak{B}.$$

۳) $B_۱, B_۲, \ldots \in \mathfrak{B} \Rightarrow \bigcup_i B_i \in \mathfrak{B}$

$$B_۱, B_۲, \ldots \in \mathfrak{B} \Rightarrow X^{-۱}\left(B_۱\right), X^{-۱}\left(B_۲\right), \ldots \in \mathcal{A}$$

$$\Rightarrow \bigcup_i X^{-۱}(B_i) \in \mathcal{A}$$

$$\Rightarrow X^{-۱}\left(\bigcup_i B_i\right) \in \mathcal{A} \Rightarrow \bigcup_i \in \mathfrak{B}.$$

قضیه ۲-۳: اگر $\mathfrak{B}$ یک میدان (یا سیگما میدان) از زیر مجموعه‌های Ω' باشد، $X^{-۱}(B)$ یک میدان (یا سیگما میدان) از زیر مجموعه‌های Ω' خواهد بود.

اثبات: $\mathcal{A}$ را گردایه نقش‌های معکوس اعضای $\mathfrak{B}$ در نظر می‌گیریم و ثابت می‌کنیم که $\mathcal{A}$ یک سیگما میدان است.

$1)\ \Omega \in \mathrm{A}$

$\Omega' \in \mathfrak{B} \Rightarrow \mathrm{X}^{-۱}(\Omega') = \Omega \in \mathrm{A}.$

$2)\ \mathrm{X}^{-۱}(B) \in \mathcal{A} \Rightarrow \left[\mathrm{X}^{-۱}(B)\right]^{c} \in \mathrm{A}$

$\mathrm{X}^{-۱}(B) \in \mathcal{A} \Rightarrow B \in \mathfrak{B} \Rightarrow B^{c} \in \mathfrak{B} \Rightarrow \mathrm{X}^{-۱}(B^{c}) \in \mathrm{A}.$

باید نشان دهیم:

$3)\ \mathrm{X}^{-۱}\left(B_{۱}\right), \mathrm{X}^{-۱}\left(B_{۲}\right), \ldots \in \mathrm{A} \Rightarrow \bigcup_{i} \mathrm{X}^{-۱}(B_{\mathrm{i}}) \in \mathrm{A}.$

$$
\begin{aligned}
&B_1, B_2, \ldots \in \mathfrak{B} \Rightarrow \bigcup B_i \in \mathfrak{B} \\
&B \in \mathfrak{B} \\
&X^{-1}(B) \in \mathcal{A} \\
&\Rightarrow X^{-1}(\bigcup B) \in \mathcal{A} \\
&\Rightarrow \bigcup (X^{-1}(B)) \in \mathcal{A}
\end{aligned}
$$

۲–۱–۳ تابع اندازه پذیر، تابع بورل، سیگما میدان القا شده

تعریف ۲-۴: اگر X یک خط حقیقی بر Ω و $\mathfrak{B}$ میدان بورل باشد،

به موجب قضیه ۲-۳ رابطه زیر یک سیگما میدان است که آن را *سیگما میدان القا شده* توسط X می‌گوییم.

$$\mathrm{X}^{-۱}(\mathfrak{B}) = \left\{\ \mathrm{X}^{-۱}(B) : B \in \mathfrak{B}\right\}$$

به عنوان مثال $I_A^{-۱}(\mathfrak{B}) = \{\Phi, A, A^c, \Omega\}$ سیگما میدان القا شده توسط I_A است.

تعریف ۲-۵: فرض کنید X یک تابع حقیقی بر Ω به توی R هستند. فرض کنید $\mathfrak{B}$ میدان بورل از زیر مجموعه‌های R و $\mathcal{A}$ یک سیگما میدان از زیر مجموعه‌های Ω باشد. اگر برای تمام مجموعه‌های بورل $B \in \mathfrak{B}$، $\mathrm{X}^{-۱}(B) \in \mathcal{A}$ باشد، X را تابع اندازه‌پذیر نسبت به $\mathcal{A}$ می‌گوییم.

تعریف ۲-۶: اگر Ω خط حقیقی R باشد و X نسبت به میدان بورل $\mathfrak{B}$ بر R اندازه‌پذیر باشد، آنگاه X را یک تابع بورل می‌نامیم. پس، X یک تابع بورل است اگر به ازای هر مجموعه بورل $\mathfrak{B}$ در فضای مقادیر، و میدان بورل $\mathfrak{B}$ در دامنه R، داشته باشیم $\mathrm{X}^{-۱}(B) \in \mathfrak{B}$.

۲-۲ متغیرهای تصادفی

تعریف ۲-۷: تابع $X:\Omega \to R$ را متغیر تصادفی نسبت به سیگما میدان $\mathcal{F}$ ، از زیر مجموعه‌های Ω می‌گوییم هرگاه به ازای هر مجموعه بورل B از سیگما میدان بورل $\mathfrak{B}$ داشته باشیم:

$$\mathrm{X}^{-1}(\mathfrak{B}) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

به عبارت دیگر متغیر تصادفی تابعی است که تصویر معکوس هر مجموعه بورل توسط آن، پیشامد است.

تعریف ۲-۸: سیگما میدان القا شده به وسیله یک تابع ثابت، سیگما میدان تباهیده است. بنابراین یک تابع ثابت نسبت به هر سیگما میدان از زیرمجموعه‌های Ω اندازه‌پذیر است.

لم ۲-۱: X یک متغیر تصادفی است اگر و تنها اگر برای هر گردایه C از زیر مجموعه‌های R که $\mathfrak{B}$ را تولید می‌کند داشته باشیم $\mathrm{X}^{-1}(\mathrm{C}) \subset \mathcal{A}$.

اثبات: باید ثابت کنیم:

$$\mathrm{X}^{-1}(\mathrm{C}) \subset \mathcal{A} \Leftrightarrow \mathrm{X}^{-1}(\mathfrak{B}) \subset \mathcal{A}$$

چون $C \subset \mathfrak{B}$ و $\mathrm{X}^{-1}(\mathfrak{B}) \subset \mathcal{A}$ پس،

$$\mathrm{X}^{-1}(\mathrm{C}) \subset \mathcal{A}$$

برای اثبات عکس آن، با توجه به اینکه $\mathcal{A}$ یک سیگما میدان است و

$$\mathrm{X}^{-1}(\mathrm{C}) \subset \mathcal{A} \Rightarrow \sigma\left(\mathrm{X}^{-1}(\mathrm{C})\right) \subset \mathcal{A}$$

$$\Rightarrow \mathrm{X}^{-1}\left(\sigma(\mathrm{C})\right) \subset \mathcal{A}$$

$$\Rightarrow \mathrm{X}^{-1}(\mathfrak{B}) \subset \mathcal{A}$$

بنابراین لم ثابت می‌گردد.

تعریف ۲-۹: فرض کنید $\mathrm{X}:\Omega \to \Omega'$, $X':\ \Omega' \to \Omega''$ آنگاه $X'\big(\mathrm{X}(\omega)\big)$ تابعی از Ω به Ω'' است که با $X'X$ یا $X'(\mathrm{X})$ نشان داده می‌شود و آن را تابع ترکیبی یا تابع مرکب می‌نامند.

تعریف ۲-۱۰: نقش معکوس $X'X$ تابعی است بر زیر مجموعه‌های Ω'' به زیر مجموعه‌های Ω به قسمی که به ازای هر $\mathrm{B} \subset \Omega''$ داشته باشیم:

$$(X'X)^{-۱}(B) = \{\omega: X'(\mathrm{X}(\omega))\epsilon B\}$$

$$= \left\{\omega: X(\omega)\epsilon X'^{-۱}(\mathrm{B})\right\}$$

$$= X^{-۱}\left(X'^{-۱}(B)\right)$$

$$= (X'X)^{-۱} = X^{-۱}X'^{-۱}$$

۲–۲–۱ متغیر تصادفی بردارهای k بعدی

در ابتدا ذکر می‌کنیم که فضای اقلیدسی دو بعدی $R^{۲}$ در صفحه میدان بورل $\mathfrak{R}^{۲}$، توسط مستطیل‌هایی به شکل زیر تولید می‌شود:

$$\{(x, y): a < x \leq b, c < y \leq d\}$$

یا حاصل‌ضرب مجموعه‌هایی به شکل زیر است :

$$\mathfrak{R}_{۱} \times \mathfrak{R}_{۲} = \left\{(x, y): x \in \mathfrak{R}_{۲}, \mathrm{y} \in \mathfrak{R}_{۲}\right\}, \mathfrak{R}_{۱}, \mathfrak{R}_{۲} \in \mathfrak{R}$$

لم ۲-۲: هر تابع بورل از تابع اندازه‌پذیر X نسبت به $\mathcal{A}$ تابعی اندازه‌پذیر نسبت به $\mathcal{A}$ بوده و زیر سیگما میدان القا شده بوسیله آن با زیر سیگما میدان القا شده بوسیله X برابر است.

اثبات: اگر $f(X)$ یک تابع بورل از X، و B یک مجموعه بورل باشد،

$$\{f(X) \in B\} = \left\{X \in f^{-۱}(B)\right\}$$

چون B یک مجموعه بورل باشد $f^{-۱}(B)$ نیز یک مجموعه بورل مانند B'' است و در نتیجه:

$$\left\{X \in f^{-۱}(B)\right\} = \{\omega: \mathrm{X}(\omega)\epsilon B'\} \subset \mathrm{X}^{-۱}(\mathfrak{B}) \subset \mathcal{A}$$

پس $f(X)$ نسبت به $\mathcal{A}$ اندازه‌پذیر است و آن یک زیر سیگما میدان $\mathrm{X}^{-۱}(B)$ که همان سیگما میدان القا شده بوسیله X است را القا می‌کند.

قضیه ۲-۴: $z^{-۱}(\mathfrak{R}^{۲})$ کوچکترین سیگما میدان است که X و Y نسبت به آن اندازه‌پذیرند.

اثبات: از نظر نمادی باید ثابت کنیم که:

$$z^{-۱}\left(\mathfrak{R}^{۲}\right) = \sigma\left\{(X^{-۱}(\mathfrak{R}) \cup (Y^{-۱}(\mathfrak{R})\right\}$$

مستطیل زیر را در نظر می‌گیریم:

$$B = \{(x,y): a < X < b, c < Y < d\}\ a,b,c,d \in R$$

$$z^{-1}(B) = \{\omega: a < X(\omega) < b, c < Y(\omega) < d\}$$
$$= \{\omega: a < X(\omega) < b\} \cap \{\omega: c < Y(\omega) < d\}$$
$$= \left(X^{-1}(a,b)\right) \cap (Y^{-1}(c,d))$$

با فرض $c = -\infty, d = \infty$ داریم $Y^{-1}(c,d) = Y^{-1}(R) = \Omega$

بنابراین،

$$\left\{X^{-1}(a,b): a < b,\ a,b \in R\right\} \subset \left\{Z^{-1}(B): \text{است یک مستطیل } B\right\} \tag{۳-۲}$$

ولی،

$$\sigma\left\{Z^{-1}(B): \text{است یک مستطیل } B\right\} = Z^{-1}\left(\mathfrak{R}^{2}\right)$$

و

$$\sigma\left\{X^{-1}(a,b):\quad a,b \in R\right\} = X^{-1}(\mathfrak{R})$$

که اولی سیگما میدان القا شده بوسیله X و دومی سیگما میدان القا شده بوسیله Z است. پس با توجه به رابطه ۲-۳ داریم

$$X^{-1}(\mathfrak{R}) \subset Z^{-1}\left(\mathfrak{R}^{2}\right) \text{ و } Y^{-1}(\mathfrak{R}) \subset Z^{-1}\left(\mathfrak{R}^{2}\right)$$
$$\Rightarrow \left\{X^{-1}(\mathfrak{R}) \cup Y^{-1}(\mathfrak{R})\right\} \subset Z^{-1}\left(\mathfrak{R}^{2}\right)$$

(۴-۲)

$$\text{چون } Z^{-1}\left(\mathfrak{R}^{2}\right) \text{ سیگما میدان است}$$
$$\Rightarrow \sigma\left\{X^{-1}(\mathfrak{R}) \cup Y^{-1}(\mathfrak{R})\right\} \subset Z^{-1}\left(\mathfrak{R}^{2}\right)$$

برای اثبات رابطه عکس ملاحظه کنید که

$$\left\{Z^{-1}(B)\text{: یک مستطیل است } B\right\} \subset \sigma\left\{X^{-1}(\Re) \cup Y^{-1}(\Re)\right\}$$

$$\left\{Z^{-1}(B)\text{: یک مستطیل است } B\right\} \subset \sigma\left\{X^{-1}(\Re) \cup Y^{-1}(\Re)\right\}$$

$$\Rightarrow Z^{-1}(\Re^{2}) \subset \sigma\left\{X^{-1}(\Re) \cup Y^{-1}(\Re)\right\} \qquad (۵-۲)$$

با توجه به روابط (۲-۴) و (۲-۵) اثبات قضیه کامل می‌شود.

از قضیه ۲-۴ نتیجه می‌شود که (X, Y) یک متغیر تصادفی برداری است اگر و فقط اگر X, Y متغیرهای تصادفی باشند.

قضیه ۲-۵: هر تابع بورل از متغیر تصادفی برداری (X, Y) یک متغیر تصادفی است.

اثبات از لم ۲-۲ نتیجه می‌شود. زیرا یک بردار تصادفی (X, Y) یک سیگما میدان القا می‌کند که یک زیر سیگما میدان از سیگما میدان القا شده به وسیله (X, Y) است.

از قضیه ۲-۵ نتیجه می‌شود که اگر (X, Y) یک متغیر تصادفی برداری باشد، آنگاه $(X - Y)$ ، $(X + Y)$، $XY, \frac{X}{Y}$ و.... که توابع بورل از (X, Y) هستند نیز متغیرهای تصادفی خواهند بود.

مثال ۲- ۵: برای به دست آوردن سیگما میدان القا شده به وسیله (I_A, I_B) ملاحظه کنید،

$$I_{\mathrm{A}}^{-1}(\Re) = \{\Phi, A, A^c, \Omega\}$$

$$I_{\mathrm{B}}^{-1}(\Re) = \{\Phi, B, B^c, \Omega\}$$

$$I_{\mathrm{A}}^{-1}(\Re) \cup I_{\mathrm{B}}^{-1}(\Re) = \{\Phi, A, A^c, B, B^c, \Omega\}$$

$$\sigma(I_{\mathrm{A}}^{-1}(\Re) \cup I_{\mathrm{B}}^{-1}(\Re) = \sigma\{AB, A^cB, AB^c, A^cB^c\} = \sigma\{A, B\}$$

که همان سیگما میدان القا شده به وسیله (I_A, I_B) است.

مجموعه$\{AB, A^cB, AB^c, A^cB^c\}$یک افراز Ω است و آن را افراز القا شده به وسیله (I_A, I_B) گوییم.

مثال ۲-۶: تابع X رابه صورت زیر در نظر بگیرید:

$$X(\omega) = \begin{cases} C_0 & \omega \in A_0 \\ C_1 & \omega \in A_1 \\ C_2 & \omega \in A_2 \end{cases}$$

آنگاه $\{A_0, A_۱, A_۲\}$ یک افراز Ω است پس X نسبت به $\mathcal{A}$ اندازه‌پذیر خواهد بود اگر و فقط اگر:

$\{A_0, A_۱, A_۲\} \in \mathcal{A}$

$$X^{-۱}(B) = \begin{cases} \emptyset & \text{اگر } B \text{ شامل } C_0 \text{ یا } C_۱ \text{ یا } C_۲ \text{ نباشد} \\ A_0 & \text{اگر } B \text{ شامل } C_0 \text{ باشد} \\ A_۱ & \text{اگر } B \text{ شامل } C_۱ \text{ باشد} \\ A_۲ & \text{اگر } B \text{ شامل } C_۲ \text{ باشد} \\ A_. + A_۱ & \text{اگر } B \text{ شامل } C_0 \text{ و } C_۱ \text{ باشد} \\ A_. + A_۲ & \text{اگر } B \text{ شامل } C_0 \text{ و } C_۲ \text{ باشد} \\ A_۱ + A_۲ & \text{اگر } B \text{ شامل } C_۱ \text{ و } C_۲ \text{ باشد} \\ \Omega & \text{اگر } B \text{ شامل } C_0 \text{ و } C_۱ \text{ و } C_۲ \text{ باشد} \end{cases}$$

$$X^{-۱}(\mathfrak{B}) = \{\emptyset, A_0, A_۱, A_۲, A_0 + A_۱, A_. + A_۲, A_۱ + A_۲, \Omega\}$$

که یک سیگما میدان مینیمال شامل افراز $\{A_0, A_۱, A_۲\}$ است.

۲-۲-۲ نگاشت‌ها به R^k

تعریف ۲-۱۱: متغیر تصادفی برداری K بعدی، $Z = (X_1, X_2, \dots, X_k)$ نگاشتی از Ω به R^k، فضای اقلیدسی K بعدی، است به قسمی که $Z^{-۱}(B)$ که در آن B یک مستطیل K بعدی به صورت زیر است یک پیشامد است،

$$B = \{(X_1, X_2, \dots, X_k):\ a_i < X_i < b_i \quad , a_i, b_i \in R \quad , i = ۱,\dots,K\} \qquad (۲-۶)$$

اگر $\mathfrak{B}_k$ میدان بورل در R^k باشد، یعنی سیگما میدان شامل تمام مستطیل‌های K بعدی به صورت (۲-۶)، آنگاه $Z^{-۱}(\mathfrak{B})$ همان سیگما میدان القا شده به وسیله Z است.

اگر $X_1, X_2, \dots, X_k$ متغیرهای تصادفی باشند، $\sum X_i$ ، $\sum X_i^2$ ، $\text{Max}(X_i)$ ، $\text{Min}(X_i)$ و غیره که توابعی بورل از $(X_1, X_2, \dots, X_k)$ هستند، متغیرهای تصادفی خواهند بود.

۲-۲-۳ انواع متغیرهای تصادفی

تعریف ۲-۱۲: اگر متغیر تصادفی X تعداد متناهی یا شمارا مقدار متمایز مانند $X_۱, X_۲, \dots$ اختیار کند، آن را یک *متغیر تصادفی گسسته* گوییم.

تعریف ۲-۱۳: هرگاه مقادیر ممکن یک متغیر تصادفی غیر قابل شمارش باشد آن را یک متغیر تصادفی پیوسته می‌نامیم.

تعریف ۲-۱۴: متغیرهای تصادفی گسسته و پیوسته متغیرهای دیگری نیز وجود دارد که آمیخته‌ای از متغیرهای تصادفی پیوسته وگسسته می‌باشند. چنین متغیرهایی را متغیرهای آمیخته می نامیم.

تعریف ۲-۱۵: یک متغیر تصادفی مختلط تابعی است از Ω به توی صفحه مختلط به قسمی که جزء حقیقی و موهومی آن متغیرهای تصادفی هستند. پس $Z = X + iY$ یک متغیر تصادفی است اگر و فقط اگر (X, Y) یک بردار تصادفی باشد.

تعریف ۲-۱۶: اگر یک متغیر تصادفی بتواند مقادیر $\pm\infty$ را اختیار کند آن را متغیر تصادفی توسعه یافته گوییم. یک متغیر تصادفی توسعه یافته نگاشتی از Ω به $\mathrm{R} = [-\infty, +\infty]$ است.

۲-۳ حدود متغیرهای تصادفی

تعریف ۲-۱۷: اگر $(\Omega, \mathcal{F})$ فضای اندازه‌پذیر و $\{X_\mathrm{n}\}_{\mathrm{n}\geq ۱}$ دنباله‌ای از متغیرهای تصادفی در $\mathcal{F}$ باشد. آنگاه:

$$\lim_{n\to\infty} \inf X_\mathrm{n} \equiv \underline{\lim} X_\mathrm{n} \equiv \bigcup_{\mathrm{k}=۱}^{\infty} \bigcap_{\mathrm{n}\geq ۱} X_\mathrm{n}$$

$$\lim_{n\to\infty} \sup X_\mathrm{n} \equiv \overline{\lim} X_\mathrm{n} \equiv \bigcap_{\mathrm{k}=۱}^{\infty} \bigcup_{\mathrm{n}\geq ۱} X_\mathrm{n}$$

اگر A مجموعه‌ای همگرا باشد در آن صورت A^c مجموعه تمام ω هایی است که در آنها $\{X_\mathrm{n}\}$ همگرا نیست.

مجموعه تمام ω هایی که به ازای آنها داریم

$$\underline{\lim} X_\mathrm{n}(\omega) = \overline{\lim} X_\mathrm{n}(\omega) = X(\omega) = \pm\infty$$

مجموعه واگرایی است. اگر Ω مجموعه همگرایی باشد دنباله $\{X_\mathrm{n}\}$ را همه جا همگرا و اگر Φ مجموعه همگرایی باشد دنباله $\{X_\mathrm{n}\}$ را هیچ جا همگرا گوییم.

۲-۳-۱ سیگما میدان القا شده به وسیله دنباله‌ای از متغیرهای تصادفی

سیگما میدان مینیمال C_1 که شامل $\cup \mathcal{A}_n$ است، سیگما میدان القا شده به وسیله دنباله نامتناهی $Z=(X_1, X_2, \dots)$ نامیده می‌شود، R^∞، مجموعه‌ی تمام دنباله‌های نامتناهی از اعداد حقیقی، فضای مقادیر Z می‌باشد.

تعریف ۲-۱۸: سیگما میدان القا شده به وسیله $X_1, \dots, X_n$ را به صورت سیگما میدان مینیمال $\mathcal{A}_n$ تعریف می‌کنیم که شامل سیگما میدان القا شده به وسیله هر یک از متغیر های $X_1, \dots, X_n$ است یعنی:

$$\mathcal{A}_n = \sigma\left[\bigcup_{i=1}^{n} X_i^{-1}(\mathfrak{B})\right]$$

به وضوح $\{\mathcal{A}_n\}$ یک دنباله صعودی یکنوا از سیگما میدان‌های با حد $\cup_{i=1}^{\infty} \mathcal{A}_n$ می‌باشد.

میدان بورل $\mathfrak{B}_\infty$ از زیر مجموعه‌های R^∞ به صورت سیگما میدان مینیمال شامل تمام مجموعه‌های استوانه‌ای با قاعده‌ای با بعد متناهی به صورت زیر تعریف می‌شود:

$$B = \left\{(x_1, \dots, x_n), a_{ij} < x_{ij} < b_{ij} \quad , j = 1,\dots,K \quad -\infty < x_{ij} < \infty\,, i = k+1 \ ,k+2,\dots\right\}$$

که در آن $(i_1, i_2, \dots)$ یک جایگشت $\left(1,2,3, \dots\right)$ بوده، $b_i \in R, a_i \in R$ و k دلخواه ولی متناهی است. بدیهی است که،

$$Z^{-1}(B) = \left\{\omega:\ a_{ij} < x_{ij}(\omega) < b_{ij} \quad , j = 1,\dots,K \ -\infty < x_{ij}(\omega) < \infty\,, i = k+1, k+2,\dots\right\}$$

$$= \bigcap_{j=1}^{k}\{\omega: \quad a_{ij} < x_{ij}(\omega) < b_{ij}\}$$

چون به ازای هر مقدار از $Z^{-1}(B), k$ متعلق به $\sigma\left\{\cup_{i=1}^{k} X_i^{-1}(\mathfrak{B})\right\}$ است پس به ازای مقداری از

$n = 1,2,\dots$ داریم $Z^{-1}(B) \in \mathcal{A}_n$.

در نتیجه $Z^{-1}(B) \in \cup_{n=1}^{\infty} \mathcal{A}_n$ و بنابراین متعلق به C_1 است. چون C_1 سیگما میدان است، سیگما میدان مینیمال شامل مجموعه‌هایی به شکل، $\{Z^{-1}(B)$: مجموعه استوانه‌ای $B\}$ جزئی از C_1 خواهد بود، در نتیجه خواهیم داشت:

$$Z^{-1}(\mathfrak{B}_\infty) \subset C_1 \quad (*)$$

به عکس می‌توان دید:

$$\mathcal{A}_n \subset Z^{-1}(\mathfrak{B}_\infty) \ \ \forall n$$

$$\Rightarrow \bigcup \mathcal{A}_n \subset Z^{-1}(\mathfrak{B}_\infty)$$

$$\Rightarrow C_1 \subset Z^{-1}(\mathfrak{B}_\infty) \quad (**)$$

$$(*)(**) \qquad \Rightarrow C_1 = Z^{-1}(\mathfrak{B}_\infty)$$

در نتیجه مشاهده می‌کنید که سیگما میدان القا شده به وسیله یک دنباله نامتناهی از متغیرهای تصادفی سیگما میدان مینیمال، شامل سیگما میدان‌های القا شده به وسیله هر یک از آنها می‌باشد.

قضیه ۲-۶: توابع Y_n، Z_n، $\underline{\lim}X_n$ و $\overline{\lim}X_n$ متغیرهای تصادفی توسعه‌یافته هستند. $Lim\ X_n$ (هنگامی که موجود است) نیز یک متغیرهای تصادفی توسعه‌یافته است.

اثبات: تابع Y_n را در نظر می‌گیریم. واضح است که آن دارای مقادیرحقیقی $-\infty$ می‌باشد. به ازای هر $x \in R$ داریم،

$$\{Y_n < x\} = \left\{\omega : \text{برای } k \geq n\text{، حداقل یک } X_n(\omega) \text{ کمتر از } x \text{ است}\right\}$$

$$= \bigcup_{k \geq n}^{\infty} \{X_k < x\}$$

چون X_k ها متغیر تصادفی هستند، $\{X_k < x\} \in \mathcal{A}$. بنابراین به ازای هر $x \in R$ داریم $\{Y_n < x\} \in \mathcal{A}$. در این صورت، Y_n به ازای هر n یک متغیر تصادفی توسعه یافته است.

همینطور، Z_n دارای مقادیر حقیقی یا $+\infty$ است و به ازای هر $x \in R$ و هر n،

$$\{Z_n > x\} = \bigcup_{k \geq n}^{\infty} \{X_k > x\} \in \mathcal{A}$$

پس Z_n یک متغیر تصادفی توسعه یافته است. ولی سوپریمم و اینفیمم متغیرهای تصادفی توسعه یافته خود متغیرهای تصادفی توسعه یافته‌اند در نتیجه $\sup Y_n$ و $\inf Z_n$ متغیرهای تصادفی توسعه یافته هستند. و این یعنی $\underline{\lim}X_n$ و $\overline{\lim}X_n$ متغیرهای تصادفی توسعه یافته‌اند. اگر $Lim\ X_n$ بر Ω وجود داشته باشد، آنگاه آن برابر مثلا $\underline{\lim}X_n$ خواهد بود و در نتیجه یک متغیر تصادفی توسعه یافته است.

۲-۳-۲ توابع ساده و مقدماتی

تعریف ۲-۱۹: یک ترکیب خطی از تعداد متناهی توابع نشانگر روی مجموعه‌های جدا از هم A_i، به صورت $\sum_{\mathrm{i}=۱}^{\mathrm{n}} x_\mathrm{i} \mathrm{I}_{\mathrm{A_i}}$ را *تابع ساده* گوییم.

تعریف ۲-۲۰: یک ترکیب خطی از تعداد شمارا توابع نشانگر روی مجموعه‌های جدا از هم را توابع مقدماتی گوییم هرگاه:

$$X = \sum_{\mathrm{i}=۱}^{\infty} x_\mathrm{i} \mathrm{I}_{\mathrm{A_i}}$$

همچنین $\{A_i\}$ یک افراز از Ω است.

قضیه ۲-۷: شرط لازم و کافی برای آنکه تابع ساده $f = \sum_{i=۱}^{n} a_i\, I_{A_i}$ اندازه‌پذیر باشد آن است که A_iها، $i = 1,2, \dots, n$ اندازه‌پذیر باشند.

اثبات: اگر A_iها اندازه‌پذیر باشند، I_{A_i}ها نیز اندازه‌پذیرند و لذا $f = \sum_{i=۱}^{n} a_i\, I_{A_i}$ اندازه‌پذیر خواهد بود.

حال فرض کنیم تابع ساده $f = \sum_{i=۱}^{n} a_i\, I_{A_i}$ اندازه‌پذیر باشد،

اگر A_iها جدا از هم باشند از آنجایی که $A_i = f^{-۱}(a_i)$، به راحتی ثابت می‌شود که A_iها اندازه‌پذیرند، اما اگر A_iها جدا از هم نباشند، دقت می‌کنیم که:

$$f(x) = \begin{cases} a_i & x \in A_i - \bigcup_{j \neq ۱} A_\mathrm{j} \\ a_i + a_j & x \in A_i \cap A_j - \bigcup_{\ell \neq i,j} A_\ell \\ \vdots & \vdots \\ a_۱ + a_۲ + \cdots + a_n & x \in A_۱ \cap \ldots \cap A_\mathrm{n} \end{cases}$$

از آنجایی که f اندازه‌پذیر فرض شده‌است، هر یک از مجموعه‌های سمت راست برابری بالا اندازه‌پذیرند. چون هر یک از A_i ها را می‌توان به کمک این مجموعه‌ها نوشت، نتیجه می‌شود که A_i ها اندازه‌پذیراند. به عنوان مثال:

$$A_۱ = (A_۱ - \bigcup_{i=2}^{n} A_\mathrm{i}) \cup (A_۱ \cap A_۲ - \bigcup_{i=3}^{n} A_\mathrm{i}) \cup \dots \cup (A_۱ \cap A_۲ \dots \cap A_\mathrm{n})$$

۲–۳–۳ جزء مثبت و جزء منفی

تعریف ۲–۲۱: جزء مثبت و جزء منفی تابع X، به صورت زیر تعریف می شوند.

$$X^- = \begin{cases} -X(\omega) & X(\omega) < 0 \\ \cdot & X(\omega) \geq 0 \end{cases}$$

$$X^+ = \begin{cases} X(\omega) & X(\omega) \geq 0 \\ \cdot & X(\omega) < 0 \end{cases}$$

در واقع به هر متغیر تصادفی مانند X می‌توانیم دو متغیر تصادفی نامنفی که به طور معادل به صورت زیر تعریف می‌شوند وابسته کنیم؛

$$X^+ = XI_{\{X\geq 0\}}$$

$$X^- = -XI_{\{X<0\}}$$

اگر X یک متغیر تصادفی باشد$\{\omega : X(\omega) \geq 0\}$, $\{\omega : X(\omega) < 0\}$ پیشامد هستند و در نتیجه توابع نشانگر آنها متغیر تصادفی می‌باشند.

پس X^+ که حاصلضرب دو متغیر تصادفی است، یک متغیر تصادفی خواهد بود. همین طور X^- نیز یک متغیر تصادفی است چون اینها نامنفی‌اند، دنباله‌های صعودی یکنوا از توابع ساده وجود دارند که به سمت آنها همگرا باشند. توجه کنید که چون

$$|X| = X^+ + X^- \quad , X = X^+ - X^-$$

$|X|$ و X توابع خطی از X^+ و X^- هستند.

۲–۴ مسائل

1- فرض کنید f, g دو تابع بورل بر $(\Omega, \mathcal{A})$ باشند به ازای $A \in \mathcal{A}$ تعریف کنید:

$$h(\omega) = \begin{cases} f(\omega) & \omega \in A \\ g(\omega) & \omega \in A^c \end{cases}$$

نشان دهید که h یک تابع بورل است.

2- نشان دهید تابع f اندازه‌پذیر است اگر وفقط اگر f^+ و f^- اندازه‌پذیر باشند.

3- به وسیله یک مثال نقض نشان دهید که اگر $|f|$ نسبت به $\mathcal{A}$ اندازه‌پذیر باشد لازم نیست f نسبت به $\mathcal{A}$ اندازه‌پذیر باشد.

4- جزء های مثبت و منفی تابع $f: R \to R$ را پیدا کنید:

$$f(x) = ax^2 + bx + c$$

فصل سوم

فضای احتمال

۳–۱ فضای نمونه و پیشامد ها

فرض می‌کنیم Z مجموعه‌ای شامل تمام برآمدهای ممکن از یک آزمایش تصادفی باشد آنگاه این مجموعه را فضای نمونه می‌نامیم. اگر $\mathcal{F}$ ، سیگما میدانی روی Z باشد و $A \in \mathcal{F}$ می‌گوییم A نسبت به $\mathcal{F}$ یک پیشامد است اگر A عضو $\mathcal{F}$ باشد. از این تعریف چنین بر می‌آید که اگر سیگما میدانی را عوض کنیم پیشامدها نیز تغییر می‌کنند.

مثال ۳-۱: فرض کنیم $\Omega = \{۱,۲,۳,۴\}$ ، $\mathcal{F}_۱ = \{\phi, \Omega, \{۱,۲\}, \{۳,۴\}\}$ در این صورت $\{۱,۲\}$ نسبت به $\mathcal{F}_۱$ پیشامد است. ϕ و Ω همواره پیشامداند و با تغییر سیگما میدان پیشامد بودن این دو مجموعه تغییر نمی‌کند.

تعریف ۳-۱: فرض می‌کنیم Ω فضای نمونه،$\mathcal{F}$ سیگما میدانی روی آن باشد در این صورت:

الف)ϕ و Ω پیشامدند.

ب) متمم هر پیشامدی، پیشامد است.

ج)اجتماع و اشتراک شمارایی از پیشامدها، پیشامد است.

د) تفاضل دو پیشامد، می‌تواند پیشامد باشد.

و) اگر دنباله‌ای از پیشامدها حد داشته باشد، حد آن نیز پیشامد است.

۳–۲ تعریف احتمال

تعریف ۳-۲: وقتی یک آزمایش صورت می‌گیرد یکی از چند برآمد ممکن نتیجه شود. ممکن است قادر نباشیم بطور دقیق برآمد مربوطه را پیش‌بینی کنیم ولی می‌توانیم با نسبت دادن یک اندازه کمی به شانس وقوع یک برآمد خاص، درباره برآمدهای ممکن اظهار نظر نماییم، پس به هرپیشامد A که شامل یک یا چند برآمد آزمایش است ممکن است

کمیتی عددی بنام "احتمال A " نسبت دهیم که آن را با $P(A)$ نشان می دهیم که اندازه‌ی "شانس" وقوع آن پیشامد را اندازه خواهد گرفت.

۳–۲–۱ روش های مختلف تعیین احتمال پیشامدها

الف) روش کلاسیک

این تعریف احتمال فقط در حالتی قابل اجرا است که آزمایش شامل تعداد متناهی، مثلاً n برآمد $W_1, \dots, W_n$ باشد (تمام برآمدها هم شانس هستند).

ب) روش آماری (فراوانی نسبی)

فرض کنید A مجموعه‌ای از برآمدهای یک آزمایش باشد، ممکن است آزمایش را به تعداد دفعات زیادی تکرار کرد. اگر n_A تعداد برآمدهای A در n تکرار باشد آن را فراوانی A و $\frac{n_A}{n}$ را فراوانی نسبی می‌گویند. مشاهده می‌شود که وقتی n بزرگ می‌شود فراوانی نسبی به صورت پایداری به سمت مقدار معینی میل می‌کند. این مقدار پایدار را احتمال A می‌نامیم و به صورت زیر تعریف می‌کنیم.

$$P(A) = \lim_{n\to\infty} \frac{n_A}{n} \qquad (۱-۳)$$

ج) روش ذهنی

وقتی آزمایشی قابل تکرار نباشد این روش را می‌توان برای نسبت دادن احتمال‌ها به پیشامدها پذیرفت.

۳–۳ اندازه احتمال

قبلاً در احتمال مقدماتی برای محاسبه احتمال از نسبت $\frac{\text{تعدادحالات مساعد}}{\text{تعداد حالات ممکن}}$ استفاده می کردیم. استفاده از این دستور این محدودیت را پیش می‌آورد که فضای نمونه باید متناهی باشد، در غیر این صورت خارج قسمت مذکور یا تعریف نشده است یا برابر صفر است. بنابراین باید مکانیزمی برای تعریف احتمال پیشامدها طراحی شود که این مشکل را نداشته باشد.

۳–۳–۱ اصول احتمال

فرض می‌کنیم Ω فضای نمونه و $\mathcal{F}$ سیگما میدانی از آن باشد در این صورت تابع مجموعه‌ای $P:\mathcal{F} \to [0,1]$ را اندازه احتمال گوییم هر گاه اصول زیر در مورد P صادق باشد.

الف) $P(A) \geq 0$

ب) $P(\Omega) = ۱$

ج) اگر A_i ها پیشامدهای دوبه دو مجزا از هم باشند، یعنی به ازای هر $i \neq j$ ، $A_i \cap A_j = \phi$ آنگاه:

$$P(\cup_i A_i) = \sum_i P(A_i)$$

بطوریکه؛

(جمع‌پذیری شمارا) $$P(\cup_{i=۱}^{\infty} A_i) = \sum_{i=۱}^{\infty} P(A_i)$$

(جمع‌پذیری متناهی) $$P(\cup_{i=۱}^{n} A_i) = \sum_{i=۱}^{n} P(A_i)$$

۳–۴ فضای احتمال

تعریف۳–۳: فرض کنید Ω فضای نمونه، $\mathcal{F}$ سیگما میدانی از زیر مجموعه‌های Ω و $P: \mathcal{F} \to [0,۱]$ یک اندازه احتمال باشد در این صورت ۳ تایی مرتب $(\Omega, \mathcal{F}, P)$ را فضای احتمال می‌نامیم. فضای احتمال مشخص می‌کند که اولاً پیشامدها به عنوان زیر مجموعه، از چه مجموعه‌ای در نظر گرفته می‌شوند (Ω) و ثانیا کدام یک از این مجموعه‌ها پیشامد هستند و کدام یک پیشامد نیستند (اگر عضو $\mathcal{F}$ باشند پیشامداند و اگر عضو $\mathcal{F}$ نباشند پیشامد نیستند) و سرانجام احتمال پیشامدها چگونه محاسبه می‌شوند (با استفاده از تابع P).

۳–۴–۱ ویژگی‌های اندازه احتمال

قضیه ۳–۱: فضای احتمال $(\Omega, \mathcal{F}, P)$ رادر نظر می‌گیریم در این صورت:

الف) $P(\phi) = 0$

ب) $P(A) + P(A^c) = ۱$

ج) اگر A و B دو پیشامد باشند که $A \subset B$ آنگاه

$$P(B - A) = P(B) - P(A) \qquad \& \qquad P(A) \leq P(B)$$

د) برای هر دو پیشامد دلخواه، A و B داریم

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

اثبات:

الف) $$\Omega = \Omega \cup \phi,\ \ ۱ = P(\Omega) = P(\Omega \cup \phi) = P(\Omega) + P(\phi)$$

$$۱ - P(\phi) = ۱ \rightarrow P(\phi) = 0$$

ب) $$\Omega = \mathrm{A} \cup A^c \rightarrow\ \ ۱ = P(\Omega) = P(\mathrm{A} \cup A^c) = P(\mathrm{A}) + P(A^c)$$

ج) با استفاده از شرط $A \subset B$ داریم

$$B = A \cup (B - A)$$

که A و $B - A$ جدا از هم هستند، لذا $P(B) = P(A) + P(B - A)$. از این برابری بنابر نامنفی بودن اندازه احتمال نتیجه می‌شود که

$$P(B - A) = P(B) + P(A) \text{ و } P(A) \leq P(B)$$

د) برای دو پیشامد دلخواه، A و B داریم $A \cup B = A \cup (B - (A \cap B))$

در این برابری دو پیشامد A و $B - (A \cap B)$ جدا از هم هستند لذا:

$$P(A \cup B) = P(A) + P(B - (A \cap B))$$

اما $(A \cap B) \subset B$ لذا بنابر قسمت (ج) داریم

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

نکته ۳-۱: اگر $P(A) = ۱$ ، آنگاه A را پیشامد تقریباً حتمی می‌گوییم، در این صورت مکمل آن A^c دارای احتمال صفر است. A^c را پیشامد پوچ یا تباهیده گوییم. اگر پدیده‌ای با احتمال یک رخ دهد گوییم آن پدیده، قریب به یقین رخ می‌دهد.

قضیه۳-۲: برای پیشامدهای دلخواه A_i، می‌توان گفت

$$P(\cup_{i=۱}^{n} A_i) \leq \sum_{i=۱}^{n} P(A_i).$$

که این نامساوی را نامساوی بول می‌نامند.

اثبات: فرض کنید:

$$\begin{cases} B_1 = A_1 \\ B_2 = A - A_1 \\ \vdots \\ B_k = A_k \cap (A_1^c \cap A_2^c \cap \cdots \cap A_{k-1}^c) \end{cases} \tag{۳-۲}$$

لذا B_k ها مجزا هستند و خواهیم داشت $\cup_{k=1}^{n} B_k = \cup_{k=1}^{n} A_k$ در نتیجه:

$$P(\cup_{k=1}^{n} A_k) = P(\cup_{k=1}^{n} B_k) = \sum_{k=1}^{n} P(B_k) \le \sum_{k=1}^{n} P(A_k)$$

۳– ۴–۲ فضای احتمال گسسته

تعریف۳-۴: در یک فضای نمونه‌ای Ω، اگر کلاس پیشامدهای $\mathcal{F}$ به وسیله افزار شمارایی از زیر مجموعه‌های Ω تولید شده باشد و همچنین P یک تابع احتمال در آن باشد آنگاه سه‌تایی $(\Omega, \mathcal{F}, P)$ را فضای احتمال گسسته می‌گوییم.

اگر $\mathcal{F} = \sigma(A_i)$ باشد که در آن $\{A_i\}$ یک افزار Ω است هر پیشامد A اجتماع شمارایی از $\{A_i\}$ ها است. بنابراین احتمال A را می‌توان با معلوم بودن احتمال $\mathrm{P_i}$ از A_i معین کرد.

۳– ۴–۳ فضای احتمال شمارا

تعریف ۳-۵: در مسائل واقعی و کاربردی معمولاً فرض می‌کنیم Ω شمارا و متناهی است و $\mathcal{F} = \sigma(\Omega)$.

اگر $\Omega = \{a_1, a_2, \dots\}$ در این صورت چون به ازای هر i خواهیم داشت $\{a_i\} \in \mathcal{F}$، بنابراین اگر P اندازه احتمال روی $\mathcal{F}$ باشد، $P(\{a_i\})$ تعریف شده است و مقدار آن را برابر P_i فرض می‌کنیم. همچنین Ω را به صورت

$\Omega = \cup \{a_i\}$ که اجتماع شمارایی از مجموعه‌های دوبه‌دو مجزا است، می‌توان نوشت:

$$1 = P(\Omega) = p(\cup \{a_i\}) = \sum_i P(\{a_i\}) = \sum_i p_i$$

بنابراین P_i در دو شرط زیر صدق می‌کند.

الف) به ازای هر $0 \le P_i \le 1$

ب) $\sum P_i = 1$

با داشتن P_i ها می‌توانیم مقدار P راروی هر مجموعه دلخواه مانند $A = \{a_{i1}, a_{i2}, \dots\}$ بدست آوریم.

ملاحظه می‌کنیم که ... $A = \{a_{i1}\} \cup \{a_{i2}\} \cup$ چون مجموعه‌های سمت راست جدا از هم می‌باشند پس:

$P(A) = \sum_j P\big(a_{ij}\big) = \sum_j p_{ij}$، یعنی برای محاسبه $P(A)$ کافی است احتمال تک تک اعضای A را با هم جمع کنیم. بنابراین در حالتی که Ω شمارا است و هر تک عضوی از Ω پیشامد است برای تعریف اندازه احتمال، کافی است مقدار اندازه احتمال را روی پیشامدهای تک عضوی بدانیم. این مقادیر لازم است در دو شرط (الف) و (ب) صدق کند.

۳-۴-۴ فضای احتمال متناهی

تعریف ۳-۶: فرض می‌کنیم $\Omega = \{a_۱, \dots, a_n\}$ متناهی و تعداد اعضای آن برابر n باشد، در این صورت با فرض $p_i = p(\{a_i\})$، $i = ۱,۲, \dots, n$ داریم:

الف) $0 \le p_i \le ۱$

ب) $\sum_{i=1}^{n} p_i = 1$

حال در این مجموعه متناهی فرض می‌کنیم p_i ها برابر باشند یعنی $p_۱ = \cdots = p_n$ چون لازم است مجموع آنها برابر یک شود لازم است هر یک از P_i ها برابر $\frac{1}{n}$ باشند. در این صورت اگر A مجموعه‌ای با k عضو و به صورت $A = \{a_{i_۱}, \dots a_{ik}\}$ باشد، آنگاه برای $۱ \le k \le n$ و $A = \{a_{i_۱}, \dots a_{ik}\}$ داریم:

$$P(A) = P_{i_۱} + \cdots + P_{ik} = \frac{۱}{n} + \cdots + \frac{۱}{n} = \frac{k}{n} = \frac{\neq (A)}{\neq (\Omega)}$$

یعنی فقط در حالت خاص که Ω متناهی است و P_i ها برابرند، احتمال هر پیشامد برابر با خارج قسمت تعداد اعضا A بر تعداد اعضای Ω است.

مثال ۳-۲: فرض کنید n سکه یک بار یا یک سکه n بار پرتاب شوند، چون هر یک از مولفه‌ها ممکن است شیر (H) یا خط (T) باشند، تعداد حالت‌های ممکن برابر $N = ۲^{n}$ می‌باشد. اگر تمام مولفه‌ها هم‌شانس باشند، یعنی $P_i = (\frac{۱}{۲})^n$ می توانیم احتمال پیشامدهایی به صورت زیر را محاسبه کنیم:

$$\mathrm{A} = (\text{خط}\, n-k, \text{شیر}\, k)$$

$$\mathrm{B} = (\text{حداقل یک شیر})$$

تعداد برآمدهای مساعد با A برابر $\binom{n}{k}$ و تعداد برآمد های مساعد با B برابر تمام n تایی‌ها به جز نقطه $(TT \dots .T)$ می باشد. بنابراین،

$$P(A) = \frac{\binom{n}{k}}{۲^{n}} \quad , P(B) = ۱ - \frac{1}{۲^{n}}$$

فرض کنید احتمال نقطه نمونه‌ای $w_{i=}\{\text{HTH}, \dots.\}$ را که شامل k شیر و $n-k$ خط است به صورت زیر تعریف کنیم:

$$P_i = \mathrm{p}^{\mathrm{k}}\mathrm{q}^{\mathrm{n-k}} \qquad q = ۱\text{-}\mathrm{p} \quad , \qquad i = ۱,۲,\dots,\mathrm{N}$$

در این صورت دیگر در حالت هم شانسی نخواهیم بود و داریم،

$$p(A) = \binom{n}{k} p^k q^{n-k} \qquad\qquad p(B) = ۱ - q^n$$

۳–۵ فضای احتمالی کلی

تاکنون فضای احتمال $(\Omega, \mathcal{F}, P)$ را معرفی کردیم و هر عضو $\mathcal{F}$ را پیشامد خواندیم. اینک ثابت می‌کنیم اندازه احتمال دارای ویژگی‌هایی است که یادآور خاصیت پیوستگی است.

قضیه۳-۳: فرض می‌کنیم$\{A_n\}$، دنباله‌ای از پیشامدهای صعودی باشد در این صورت:

$$p(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$$

اثبات:

اگر $A_{\mathrm{n}} \uparrow A$ آنگاه $\lim\limits_{n\to\infty} A_n = \cup_{n=۱}^{\infty} A_n$. حال سعی می‌کنیم $\cup_{n=۱}^{\infty} A_n$ را به پیشامدهای جدا از هم افراز کنیم. برای این منظور پیشامدهای زیر را تعریف می‌کنیم:

$$\begin{cases} A_0 = \emptyset \\ B_۱ = A_۱ \\ B_۲ = A_۲ - A_۱ \\ \quad \vdots \\ B_{\mathrm{n}} = (A_{\mathrm{n}} - A_{\mathrm{n}\text{-}۱}) \end{cases}$$

به ازای هر $i \neq j$ داریم،

$$A_n = \cup_{i=۱}^{n} B_i \text{ و } \cup_{n=۱}^{\infty} B_n = \cup_{n=۱}^{\infty} A_n \text{ و } B_i \cap B_j = \emptyset$$

در نتیجه

$$p\left(\lim_{n\to\infty} A_n\right) = p\left(\bigcup_{n=۱}^{\infty} A_n\right) = p\left(\bigcup_{n=۱}^{\infty} B_n\right)$$

$$= \sum_{n=۱}^{\infty} p(B_n) = \sum_{n=۱}^{\infty} p(A_n - A_{n-1}) = \lim_{N\to\infty} \sum_{n=۱}^{N} p(A_n - A_{n-1})$$

$$= \lim_{N\to\infty} (p(A_n) - p(A_0)) = \lim_{N\to\infty} p(A_N)$$

مشابه این قضیه برای حالت نزولی نیز برقرار است.

نتیجه۳-۱: وقتی $n \longrightarrow \infty$،

$$A_n \longrightarrow A \quad \Rightarrow \quad P(A_n) \longrightarrow p(A)$$

اثبات:

$$\bigcap_{K\geq n} A_k \subset A_n \subset \bigcup_{K\geq n} A_k \qquad (۳\text{-}۳)$$

لذا داریم:

$$p\left(\bigcap_{K\geq n} A_k\right) \leq p(A_n) \leq p\left(\bigcup_{K\geq n} A_k\right)$$

می‌دانیم وقتی A_n حد داشته باشد، خواهیم داشت

$\overline{\lim} A_n = \underline{\lim} A_n = \lim A_n = A$ که نتیجه می‌دهد:

$$\bigcup_{K\geq n} A_k \downarrow \overline{\lim} A_n = A \quad , \quad \bigcap_{K\geq n} A_k \uparrow \underline{\lim} A_n = A$$

همچنین با استفاده از قضیه پیوستگی داریم:

$$p\left(\bigcup_{K\geq n} A_k\right) \downarrow p(A) \quad , \quad p\left(\bigcap_{K\geq n} A_k\right) \uparrow p(A)$$

در نتیجه با استفاده از رابطه (۳-۳) داریم: $p(A_n) \longrightarrow p(A)$

۳-۵-۱ توسیع اندازه احتمال

تعریف۳-۷: فرض کنید P بر میدان معین $\mathcal{F}$ تعریف شده باشد. P^* را بر $\sigma(\mathcal{F})$ طوری تعریف می‌کنیم که بر مجموعه‌های $\mathcal{F}$ ، P^* با P برابر باشند. در این صورت، P^* را توسیع P بر $\sigma(\mathcal{F})$ و P را تحدید P^* بر $\mathcal{F}$ می‌نامند. پس P^* دارای قلمرو وسیع‌تر از P بوده و بر قلمرو مشترک با هم برابرند.

قضیه۳-۴: (توسیع کارا تئودوری)؛ احتمال تعریف شده بر میدانی مانند $\mathcal{F}$ را می‌توان بطور منحصر بفرد بر میدان سیگمای مینیمال شامل $\mathcal{F}$ تعمیم داد.

اثبات: برای هر $A \subset \Omega$، فرض کنید $A_j \in \mathcal{G}$ و $A \subset \bigcup A_j$. در این صورت، $\{A_j\}$ را یک پوشش A گویند. برای هر A ، تابع مجموعه‌ای زیر را تعریف می‌کنیم

$$P^*(\mathrm{A}) = \inf \sum_{j=1}^{\infty} P(A_j)$$

اینفیمم روی تمام پوشش های A اختیار شده است. به آسانی دیده می‌شود که:

i) $P^*(\emptyset) = 0$

ii) $P^*(\mathrm{A}) \leq P^*(\mathrm{B}) \qquad A \subset B$

iii) $P^*\left(\sum_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} P^*\left(A_j\right)$ (زیر سیگما – جمع پذیری)

هر تابع P^* را که در ۳ شرط بالا صدق کند توسیع خارجی گویند.

مجموعه‌ی A را P^* اندازه‌پذیر گویند اگر به ازای هر $D \subset \Omega$

$$P^*(\mathrm{D}) = P^*(A \cap D) + P^*(A^c \cap D)$$

با استفاده از خاصیت زیر سیگما-جمع‌پذیری P^* می توان ثابت کرد که، گردایه A^* از تمام مجموعه‌های P^* اندازه‌پذیر، تشکیل یک سیگما میدان خواهد داد و P^* بر A^* یک اندازه است به طوری که برای،

$$\forall A \in \mathcal{G},\ \ \forall D \subset \Omega\,, \forall \varepsilon > 0 \qquad \exists \{A_n\}_{\mathrm{n}=1}^{\infty}\ A_n \in \mathcal{G}, D \subset \bigcup A_n$$

بنابراین:

$$P^*(\mathrm{D}) + \varepsilon \geq \sum_{j=1}^{\infty} P\left(A_j\right) = \sum_{j=1}^{\infty} P\left(\mathrm{A} \cap A_j\right) + \sum_{j=1}^{\infty} P\left(A^c \cap A_j\right)$$

$$\geq P^*(A \cap D) + P^*(A^c \cap D)$$

چون ε اختیاری است؛

$$P^*(\mathrm{D}) \geq P^*(A \cap D) + P^*(A^c \cap D)$$

همچنین،

$$\begin{aligned} P^*(\mathrm{D}) &= P^*\big(D \cap (A \cup A^c)\big) = P^*[(D \cap A) \cup (D \cap A^c)] \\ &\leq P^*(A \cap D) + P^*(A^c \cap D) \end{aligned}$$

$$\Rightarrow P^*(D) = P^*(A \cap D) + P^*(A^c \cap D)$$

بنابراین A مجموعه‌ای، P^* اندازه‌پذیر است و به A^* تعلق دارد پس $\mathcal{G} \subseteq A^*$.

$\mathcal{G} \subseteq A^* \to \sigma(g)\ ,\ g$ میدان سیگمای مینیمال شامل $\Rightarrow \sigma(\mathcal{G}) \subseteq A^*$

P^* یک توسیع P بر $\sigma(\mathcal{G})$ است. برای اثبات منحصر به فرد بودن نیاز به مفاهیم زیر داریم.

تعریف ۳-۸: (Π سیستم)؛ گردایه P از زیر مجموعه‌های Ω یک Π سیستم است اگر نسبت به اشتراک متناهی بسته باشد.

تعریف ۳-۹: (λ سیستم)؛ گردایه f را یک λ سیستم گوییم هرگاه Ω را شامل شود و تحت مکمل اجتماع شمارش پذیر بسته باشد.

نتیجه ۳-۲:

الف) یک گردایه که هم Π سیستم و هم λ سیستم باشد، یک سیگما میدان است.

ب) اگر p یک Π سیستم باشد و f یک λ سیستم باشد و $P \subset f$،

آنگاه: $\sigma\ (P) \subset f$

چون یک میدان همواره یک Π سیستم خواهد بود، لذا یکتایی قضیه توسیع نتیجه مستقیم قضیه بالا خواهد بود.

لم ۳-۱: تابع احتمال تعریف شده روی تمام بازه‌هایی به صورت $(a,b]$ بطور یکتا یک توسیع بر میدان مینیمال شامل تمام بازه‌ها تعریف می‌کند.

اثبات: فرض کنید $\mathfrak{B}_\mathrm{t}$ کلاس تمام بازه‌های به صورت $(a,b]$ و $\mathfrak{B}$. میدان مینیمال شامل $\mathfrak{B}_\mathrm{t}$ باشد. این میدان مینیمال شامل اجتماع متناهی مجموعه‌های $\mathfrak{B}_\mathrm{t}$ و مکمل مجموعه‌های $\mathfrak{B}_\mathrm{t}$ است، یعنی مجموعه‌هایی به صورت $(-\infty,b] \cup (b,\infty)$. فرض کنید $\mathrm{A} \in \mathfrak{B}_\mathrm{t}$ و $A = \sum_k A_k$ ، $A_k \in \mathfrak{B}_\mathrm{t}$ ، در این صورت،

$P(A) = \sum_k P(A_k)$ و اگر A نیز برابر $\sum_{j=1}^{n} A'_j$ ، $A'_j \in \mathfrak{B}_\mathrm{t}$ باشد آنگاه $P(A) = \sum_{j=1}^{n} P\big(A'_j\big)$.

باید ثابت کنیم:

$$\sum_k P(A_k) = \sum_j P(A'_j)$$

از طرفی داریم:

$$A_k = A \cap A_k = \sum_j A_k \cap A'_j$$

$$A'_j = A \cap A'_j = \sum_k A_k \cap A'_j$$

چون P بر $\mathfrak{B}_t$ جمع پذیر است در نتیجه:

$$\sum_j P(A'_j) = \sum_j P(\sum_k A_k \cap A'_j) = \sum_j \sum_k P(A_k \cap A'_j)$$

$$= \sum_k P(\sum_j (A'_j \cap A_k)) = \sum_k P(A_k)$$

بنابراین چون $p(A^c) = ۱\text{-p(A)}$ به ازای هر $A \in \mathfrak{B}_t$ قضیه ثابت می شود.

مثال ۳-۳: فرض‌کنید $\Omega = \{\omega: 0 \leq \omega \leq ۱\}$. دراین صورت Ω فضای ناشمارا می‌باشد. فرض کنید سیگما میدان پیشامدها میدان بورل محدود شده به زیر مجموعه‌های $[0,1]$ باشد.

پس، $\mathcal{A} = \{B \cap [0,1]; B \in \mathfrak{B}\}$ کوچکترین سیگما میدان تولید شده بوسیله گردایه زیرمجموعه‌هایی به صورت $[0, x]$، $۰ \leq x \leq ۱$، می‌باشد. تعریف می‌کنیم، $P[0,x) = x$,$0 \leq x \leq ۱$. بنابراین، $P(\Omega) = ۱$. این تابع احتمال دارای تعمیم منحصر بفردی بر $\mathcal{A}$ است که آن را توزیع یکنواخت یا مستطیلی بر $[0,1]$ گوییم.

حال فرض کنید $\Omega = \{\omega: 0 \leq \omega \leq a\}$ و $\mathcal{A}$ میدان بورل تولید شده بوسیله زیر مجموعه‌های $[0, a]$ باشد. تابع احتمال را به صورت زیر تعریف می‌کنیم:

$$P\{[0,x)\} = \frac{x}{a} \qquad (0 < x \leq a)$$

داریم $P(\Omega) = ۱$، و P را می توان بر $\mathcal{A}$ بطور منحصر بفردی تعمیم داد. P را توزیع مستطیلی بر $[0, a]$ گوییم.

مثال ۳-۴: فرض کنید $\Omega = R^۲ = \{(x, y): -\infty < x, y < \infty\}$ و $\mathcal{A}$ میدان بورل زیر مجموعه‌های $R^۲$ باشد. مستطیل‌های به صورت زیر را در نظر می‌گیریم:

$$\{(\mathrm{x},\mathrm{y}): -\infty < x \le \mathrm{a}, -\infty < y \le b\} \quad a,b \in R$$

چون این مستطیل‌ها $\mathfrak{B}_{۲}$ را پدید می‌آورند، و تحت اشتراک‌های متناهی بسته‌اند، با تعریف احتمال‌ها روی این مستطیل‌ها می‌توان آنها را به تمام مجموعه‌های بورل $R^{۲}$ تعمیم داد. فرض کنید $\mathrm{f}(\mathrm{x},\mathrm{y})$ تابع نامنفی از (x,y) باشد به طوری که انتگرال آن بر $R^{۲}$ واحد باشد. با تعریف

$$F(a,b) = P\{-\infty < x \le a, -\infty < y \le b\} = \int_{-\infty}^{a}\int_{-\infty}^{b} f(\mathrm{x},\mathrm{y})\, dxdy$$

یک توزیع احتمال بر $R^{۲}$ را با تابع چگالی احتمال $\mathrm{f}(\mathrm{x},\mathrm{y})$ بدست می‌آوریم. اگر $\Omega = R^{\mathrm{k}}$ و $\mathfrak{B}_{\mathrm{k}}$ میدان بورل در R^{k} باشد، P را می‌توان بوسیله تابع $f(\mathrm{x}_{۱},\mathrm{x}_{۲},\dots,\mathrm{x}_{\mathrm{k}})$ تعریف کرد بطوری که $f(\mathrm{x}_{۱},\mathrm{x}_{۲},\dots,\mathrm{x}_{\mathrm{k}}) \ge 0$

و

$$\int \dots . \int_{R^k} f(\mathrm{x}_{۱},\dots,\mathrm{x}_{\mathrm{k}})\ d\mathrm{x}_{۱} \dots\dots d\mathrm{x}_{\mathrm{k}} = ۱$$

در این صورت،

$$F\left(a_{۱},\dots,a_{\mathrm{k}}\right) =$$

$$p\{-\infty < \mathrm{x}_{۱} \le a_{۱}, -\infty < \mathrm{x}_{۲} \le a_{۲}, \dots -\infty < \mathrm{x}_{\mathrm{k}} \le a_{\mathrm{k}}\}$$

$$= \int_{-\infty}^{a_{۱}}\int_{-\infty}^{a_{۲}} \dots . . \int_{-\infty}^{a_{\mathrm{k}}} f(\mathrm{x}_{۱},\dots,\mathrm{x}_{\mathrm{k}})\, d\mathrm{x}_{\mathrm{k}} \dots\dots d\mathrm{x}_{۱}$$

تابع $f(\mathrm{x}_{۱},\dots,\mathrm{x}_{\mathrm{k}})$ را تابع چگالی بر R^{k} گوییم.

۳–۶ فضای احتمال القا شده

فرض کنید X یک متغییر تصادفی برای مجموعه‌های جدا از هم بر $(\Omega,\mathcal{A},P)$ باشد به طوری که برای هر $B \in \mathfrak{B}$ تعریف می‌کنیم:

$$P_X(B) = P\{\omega: X(\omega) \in B\} = P\{X^{-۱}(B)\}$$

در این صورت $P_X(B) \ge 0$ و

$$P_X(\sum_j \mathrm{B}_j) = \mathrm{p}(\mathrm{x}^{-۱}(\sum_j \mathrm{B}_j)) = \mathrm{p}(\sum_j \mathrm{x}^{-۱}(\mathrm{B}_j)) = \sum_j \mathrm{p}(\mathrm{x}^{-۱}(\mathrm{B}_j)) = \sum_j \mathrm{p}_X(\mathrm{B}_j)$$

چون $P_X(\Omega) = 1$ ، P_X بر $(\mathrm{R}, \mathfrak{B})$ یک اندازه احتمال است. پس $(\mathrm{R}, \mathfrak{B}, P_X)$ یک فضای احتمال جدید به‌نام فضای القا شده بوسیله‌ی X بر فضای بردش می‌باشد. P_X را توزیع احتمال X گویند.

مثال ۳-۵: اگر g هر تابع بورل از X باشد آنگاه به ازای هر $\mathrm{B} \in \mathfrak{B}$خواهیم داشت.

$$P\{\mathrm{g}(\mathrm{X}(\omega)) \in \mathrm{B}\} = P\{X(\omega) \in \mathrm{g}^{-۱}(\mathrm{B})\} = P_X\{\mathrm{g}^{-۱}(B)\}$$

چون g یک تابع بورل است، $\mathrm{g}^{-۱}(B)$ یک مجموعه بورل است. بنابراین توزیع احتمال X به صورت منحصربفردی توزیع احتمال هر تابع بورل مانند $\mathrm{g}(\mathrm{x})$ را معین می‌کند.

۳-۷ توزیع توابع بورل از متغیرهای تصادفی

اگر $y = \mathrm{g}(X)$ یک تابع یک مقداری باشد (اگر g صعودی باشد) خواهیم داشت

$$P(\mathrm{g}(\mathrm{x}) \leq \mathrm{y}) = \mathrm{P}(\mathrm{x} \leq \mathrm{g}^{-۱}(y))$$

اگر $f(X)$ تابع چگالی احتمال X و $\mathrm{g}^{-۱}$ موجود باشد و g و $\mathrm{g}^{-۱}$ مشتق‌پذیر باشند با استفاده از تبدیل متغیرها داریم:

$$P(\mathrm{g}(\mathrm{x}_۱) \leq \mathrm{g}(\mathrm{x}) \leq \mathrm{g}(\mathrm{x}_۲)) = \int_{\mathrm{x}_۱}^{\mathrm{x}_۲} f(x)dx = \int_{y_۱}^{y_۲} f(h(y))h(y)dy$$

که در آن $y_۱ = \mathrm{g}(\mathrm{x}_۱)$ ، $y_۲ = \mathrm{g}(\mathrm{x}_۲)$ ، $\mathrm{g}^{-۱}(\mathrm{y}) = \mathrm{x} = \mathrm{h}(\mathrm{y})$

با استفاده از این روش گاهی می‌توان تابع توزیع $\mathrm{g}(\mathrm{x})$ را بدست آوریم.

مثال ۳-۶: اگر X دارای توزیع نرمال با تابع چگالی $\frac{1}{\sqrt{2\pi}} e^{\frac{-\mathrm{x}^2}{۲}}$ $-\infty < \mathrm{x} < \infty$ باشد و داشته باشیم

$\log \mathrm{y} = \mathrm{x}$ در این صورت y دارای توزیع نرمال لگاریتمی با چگالی احتمال زیر است:

$$f_y(y) = \frac{d\mathrm{x}}{dy} f_{\mathrm{x}}(\log y) = \frac{۱}{e^x \sqrt{۲\pi}} e^{\frac{-(\log y)^۲}{۲}} = \frac{۱}{y\sqrt{۲\pi}} e^{\frac{-(\log y)^۲}{۲}} 0 < \mathrm{y} < \infty$$

۳-۸ اندازه های دیگر

۳-۸-۱ اندازه احتمال تعمیم یافته

تعریف۳-۱۰: احتمال، یک تابع مجموعه‌ای استاندارد شده نامنفی و جمع‌پذیر سیگما بر $(\Omega,\mathcal{A})$ است. بدون شرط استاندارد شدن، اگر یک تابع مجموعه‌ای نامنفی جمع‌پذیر داشته باشیم، آن را یک اندازه می‌نامیم و با μ نشان می‌دهیم. اگر $\mu(\Omega)$ مخالف صفر و متناهی باشد $\frac{\mu(.)}{\mu(\Omega)}$ را به صورت یک تابع مجموعه‌ای بر $(\Omega,\mathcal{A})$ در نظر می‌گیریم. دیده می‌شود که این تابع در تمام خواص تعریف (۳-۳-۱) صدق کرده و در نتیجه می‌تواند به عنوان یک اندازه احتمال در نظر گرفته شود، در نتیجه استاندارد بودن را می‌توان با تقسیم تابع اندازه بر $\mu(\Omega)$ بدست آورد. در نتیجه اندازه‌های متناهی را می‌توان به عنوان تعمیم اندازه‌های احتمال در نظر گرفت.

۳-۸-۲ اندازه احتمال شرطی

تعریف۳-۱۱: فرض کنید B زیر مجموعه‌ای اختیاری ولی ثابت و متعلق به $\mathcal{A}$ باشد. گردایه پیشامدهای زیر را در نظر می‌گیریم:

$$B\cap\mathcal{A}=\{B\cap A,\ \ A\in\mathcal{A}\}=\mathcal{A}_B$$

این گردایه یک سیگما میدان از زیر مجموعه‌های B است. پس $(B,\mathcal{A}_B)$ یک فضای اندازه پذیر است. اندازه احتمال P که بر این فضا تعریف شده است، استاندارد نمی‌باشد زیرا در حالت کلی $P(B)$ نمی‌تواند برابر واحد باشد. ولی می‌توان آن را با تعریف اندازه P^* بر $\mathcal{A}_B$ ، اگر $P(B)\geq o$ ، به صورت زیر استاندارد کرد.

$$P^*(A\cap B)=\frac{P(A\cap B)}{P(B)}$$

در این صورت $(\mathrm{B},\mathcal{A}_{\mathrm{B}},\mathrm{P}^*)$ یک فضای احتمال است. قلمرو P^* گردایه پیشامدهایی است که زیرمجموعه‌های B هستند. به ازای هر پیشامد مفروض B ، P^* را می‌توان به عنوان تابع مجموعه‌ای تعریف شده بر $\mathcal{A}$ در نظر گرفت. در این صورت، آن را اندازه احتمال شرطی با فرض B گوییم و به صورت P_{B} یا $\mathrm{P}(.|\mathrm{B})$ نشان می‌دهیم. در نتیجه،

$$P_B(A)=P^*(A\cap B)=\frac{P(A\cap B)}{P(B)}=P(A|B)$$

بدیهی است که P_B یک اندازه احتمال بر $(\Omega,\mathcal{A})$ است زیرا

$$\begin{cases} P_B(A) \geq ۰ \\ P_B(\Omega) = ۱ \\ P_B(\bigcup_j A_j) = P^*(B \cap \bigcup_j A_j) = \dfrac{P(B \cap \cup_j A_j)}{P(B)} = \dfrac{\sum_j P(B \cap A_j)}{P(B)} = \sum_j P_B(A_j) \end{cases}$$

در نتیجه $(\Omega, \mathcal{A}, P_B)$ یک فضای احتمال است.

مثال ۳-۷: فرض کنید

$$P(X_N = \mathrm{x} | N = n) = \frac{۱}{n + ۱} \qquad \mathrm{x} = 0,۱,۲,\ldots,\mathrm{n}$$

$$P(N = n) = \frac{e^{-\lambda}\lambda^n}{n!}$$

احتمال غیر شرطی $P\{X_N = \mathrm{x}\}$ را محاسبه کنید.

$$P(X_N = \mathrm{x} | N = n) = \frac{۱}{n + ۱} \qquad x = 0,۱,\ldots,\mathrm{n}$$

$$P(N = n) = \frac{e^{-\lambda}\lambda^n}{n!} \qquad n = 0,۱,۲,\ldots$$

$$P\{X_N = \mathrm{x}\} = \sum_{n=0}^{\infty} P(N = n) P(X_N = \mathrm{x} | N = n) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \frac{۱}{n + ۱}$$

$$= \sum_{n=۰}^{\infty} \frac{e^{-\lambda}\lambda^n}{(n + ۱)!}$$

$$= \lambda^{-۱} \sum_{n=۰}^{\infty} \frac{e^{-\lambda}\lambda^{n+1}}{(n + ۱)!}$$

$$= \lambda^{-۱} \left[\sum_{m=1}^{\infty} \frac{e^{-\lambda}\lambda^m}{m!} \right]$$

$$= \lambda^{-۱} \left(۱\text{-}e^{-\lambda} \right)$$

مثال ۳-۸: سکه‌ای را دو مرتبه پرتاب می‌کنیم. با فرض اینکه هر ۴ عضو فضای نمونه آزمایش

$S = \{(H,H),(H,T),(T,H),(T,T)\}$ هم شانس باشند، احتمال اینکه در هر دو پرتاب شیر ظاهر شود به شرط اینکه پرتاب اول شیر آمده باشد چقدر است؟

$$E = \{(H,H)\} \rightarrow \text{هر دو پرتاب شیر}$$

$$F = \{(H,H)(H,T)\} \rightarrow \text{پرتاب اول شیر}$$

$$p(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(\{H,H\})}{P\{(H,H),(H,T)\}} = \frac{\frac{۱}{۴}}{\frac{۲}{۴}} = \frac{۱}{۲}$$

۳-۸-۳ اندازه شمارنده

تعریف۳-۱۲: فرض کنید Ω خط حقیقی R و$\mathcal{A}$گردایه تمام مجموعه‌های بورل در R باشد تابع مجموعه‌ای μ را به صورت زیر تعریف می‌کنیم:

$\mu(B)$: تعداد اعداد صحیح (مثبت، منفی یا صفر) متعلق به B

در این صورت،

$$\mu(\emptyset) = 0 \quad , \quad \mu(\textstyle\sum_i B_i) = \sum_i \mu(B_i) \quad , \quad B_i \in \mathfrak{B}$$

پس μ یک اندازه بر $\mathcal{A}$ است که آن را اندازه شمارنده می‌گوییم. اندازه شمارنده را ممکن است بر $(R^k, \mathfrak{B}_k)$ به صورت زیر تعریف کرد.

برای هر $B_k \in \mathfrak{B}_k$ ، تعداد نقاط$(X_۱, \dots \dots, X_k)$ با مختصات صحیح (مثبت،منفی یا صفر) درB_k :$\mu(B_k)$

۳-۸-۴ اندازه لبگ

تعریف۳-۱۳: خط حقیقی R و گردایه$\mathcal{G}$ شامل تمام بازه‌های کراندار به صورت $(a,b], a < b$ را در نظر می‌گیریم تابع مجموعه‌ای μ را بر اعضای $\mathcal{G}$ به قسمی تعریف می‌کنیم که:

$$\mu(a,b] = b - a \qquad , \mu(\emptyset) = 0 \qquad , \mu(\cup E_i) = \sum_i \mu(E_i)$$

که در آن .. ... $E_۱, E_2,$ مجموعه‌های جدا از هم $\mathcal{G}$ هستند. پس، μ یک اندازه بر $\mathcal{G}$ است. چون $\mathcal{G}$ تحت اشتراک‌های متناهی بسته است، با استفاده از لم (۳-۱) می‌توان μ را به طور منحصر بفردی به تمام مجموعه‌های بورل تعمیم داد.

چون $R = \sum_{n=-\infty}^{\infty}(n, n+1]$ و $\mu(n, n+1] = 1$، μ یک اندازه سیگما متناهی است.

چون$\{x\} = \lim(x - \frac{۱}{n}, x]$ و $\mu(x - \frac{۱}{n}, x] = \frac{۱}{n}$، با توجه به خاصیت پیوستگی μ، برای تمام یکتایی‌های $\{x\}$ داریم $\mu(\{x\}) = 0$. مجموعه‌ای که اندازه μ آن صفر است مجموعه پوچ بر حسب μ نامیده می‌شود.

مثال ۳-۹: (مجموعه کانتور)؛ بازهٔ $[0,1]$ را به زیر بازهٔ $[0, \frac{۱}{۳}], (\frac{۱}{۳}, \frac{۲}{۳}), [\frac{۲}{۳}, ۱]$ افراز کرده و بازهٔ باز میانی را از آن حذف می‌کنیم. مجدداً هر یک از دو بازهٔ باقی‌مانده را به سه قسمت مساوی تقسیم کرده و بازهٔ میانی آنها را حذف می‌کنیم. مجدداً بازه‌های باقی مانده را (که البته بازه‌های بسته هستند) به سه قسمت مساوی تقسیم می‌کنیم و جزء میانی آنها را حذف می‌کنیم. این روش را بی‌نهایت مرتبه ادامه می دهیم. آنچه که از بازهٔ $[0,1]$ باقی می‌ماند، مجموعهٔ کانتور است. معرفی صوری مجموعهٔ کانتور به شرح زیر است.

فرض کنیم:

$E_{۱,۱}: \quad [0,1]$ بازهٔ میانی بازهٔ

$E_{۲,۲}, E_{۲,۱}: \quad [0,1] - E_{۱۱}$ بازه‌های میانی مجموعهٔ

$E_{n,۱} \dots . E_{n,n}: \quad [0,1] - \cup_{m=۱}^{n-۱} \cup_{k=۱}^{۲^m-۱} E_{m,k}$ بازه‌های میانی مجموعهٔ

باشند، در این صورت مجموعهٔ کانتور را با نماد C به صورت زیر تعریف می‌شود

$$c = [0,1] - \bigcup_{m=۱}^{\infty} \bigcup_{k=۱}^{۲^m-۱} E_{m,k}$$

۳-۸-۵ اندازه علامت‌دار

تعریف ۳-۱۴: اگر یک تابع مجموعه‌ای جمع‌پذیر سیگما مانند v که بر $\mathcal{A}$ تعریف شده، بتواند مقادیر مثبت و منفی را اختیار کند ولی حداکثر یکی از مقادیر $\pm\infty$ را بگیرد و در شرط $v(\emptyset) = 0$ صدق کند آن را اندازه علامت‌دار می‌نامیم. مثلا اگر $v = \mu_1 - \mu_2$ که در آن μ_1 و μ_2 اندازه‌های تعریف شده بر $\mathcal{A}$ باشند بطوریکه حداقل یکی از آنها متناهی است، آنگاه v یک اندازه علامت‌دار خواهد بود.

۳-۹ قانون احتمال کل و دستور بیز

۳-۹-۱ قانون احتمال کل

قضیه ۳-۵: فرض کنیم $\{A_n\}$ دنباله‌ای از پیشامدهای دو به دو ناسازگار (به ازای هر $i \neq j A_i \cap A_j = \emptyset$) و اجتماع آنها برابر Ω باشد در این صورت به ازای هر پیشامد دلخواه A داریم:

$$P(A) = \sum_i P(A_i)P(A|A_i)$$

اثبات: از آنجا که مجموعه پیشامدهای $A_i \cap A$ به علت ناسازگاری A_i ها، ناسازگارند و اجتماع آنها برابر A است داریم:

$$P(A) = P(\bigcup_i (A \cap A_i)) = \sum_i P(A \cap A_i) = \sum_i P(A_i)P(A|A_i)$$

۳-۹-۲ دستور بیز

قضیه۳-۶: فرض کنیم پیشامدهای... $A_1, A_2,$ افرازی از فضای نمونه باشند (دو به دو ناسازگار و اجتماع آنها فضای نمونه Ω است). به ازای هر پیشامد مانند A داریم :

$$P(A_i|A) = \frac{P(A_i)P(A|A_i)}{\sum_j P(A_j)P(A|A_j)}$$

اثبات: اثبات این قضیه با استفاده از تعریف احتمال شرطی و دستور احتمال کل به آسانی حاصل می‌شود.

$$P(A_i|A) = \frac{P(A_i \cap A)}{P(A)} = \frac{P(A_i)P(A|A_i)}{\sum_j P(A_j)P(A|A_j)}$$

۳–۱۰مسائل

۱- فرض کنید $A_1, A_2, \dots, A_n$ پیشامدهای دلخواه باشند و قرار دهید:

$$I_k = \bigcap_i (A_{i_1} \cup A_{i_2} \cup A_{i_3} \dots \cup A_{i_k}$$

$$U_k = \bigcup_i (A_{i_1} \cap A_{i_2} \cap A_{i_3} \dots \cap A_{i_k}$$

۲- نشان دهید $U_k = I_{n+k+1}$.

۳- نشان دهید.

$$P(A \cap B) \leq min[P(A), P(B)]$$

۴-اگر $\Omega = \mathrm{R}\ , A_{2n} = \{x; -n<x<n\}\ , A_{2n+1} = \{x; 0<x<\frac{1}{n}\}$ باشدآنگاه $\lim inf A_n\ , \lim supA_n$ را بدست آورید.

۵ -اگر $A_n = \begin{cases} \left(-\frac{1}{n}, 1\right] & n \text{ فرد} \\ \left(-1, \frac{1}{n}\right] & n \text{ زوج} \end{cases}$ باشد. $\lim inf A_n\ , \lim supA_n$ را پیدا کنید.

۶- ثابت کنید $F = \{(X, \infty), X \in R\}$ سیگما میدان نیست.

۷- اگر

$$\Omega = \{a, b, c, d\}$$

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{a\}\}$$

$$\mathcal{F}_2 = \{\{b, c, d\}, \{a\}, \emptyset, \Omega\}$$

باشد آیا سیگما میدان‌های $\mathcal{F}_1 \mathcal{F}_2$ یک میدان می باشند.

فصل چهارم

توابع توزیع

۴–۱ تابع توزیع یک متغیر تصادفی

تعریف ۴–۱: در آمار توصیفی عموماً فراوانی نسبی را به عنوان برآورد احتمال یک پیشامد در نظر می‌گیرند یعنی اگر در n آزمایش k بار مقدار X کمتر یا مساوی x مشاهده شود $\frac{k}{n}$ یک برآورد برای عبارت زیر است:

$$F_n(X) = P\{X \leq x\} \tag{۴-۱}$$

که همان تابع توزیع تجربی است که در انتهای فصل مورد بحث قرار می‌گیرد. حال اگر X گسسته باشد و m تعداد دفعاتی باشد که X مقدار x را اختیار کرده است، $P\{X = x\} = \frac{m}{n}$ خواهد بود. پس توابع توزیع، اساس نظریه‌ی آمار و احتمال را تشکیل می‌دهند.

۴–۱–۱ چند خاصیت توابع توزیع

اگر X یک متغیر تصادفی بر فضای احتمال (Ω, A, P) باشد برای $x \in R$ تعریف می‌کنیم:

$i)\ P(X \leq x) = F_X(x)$

$ii)\ P_X(a,b] = P\{a < X \leq b\} = F_X(b) - F_X(a)$

که در آن $F_X(x)$ تابع توزیع تجمعی متغیر تصادفی X است. اگر داشته باشیم $F_X(-\infty) = 0$ هر یک از خواص فوق دیگری را نتیجه می‌دهند.

اثبات:

$ii) \leftrightarrow\ P_X(a,b] = F_X(b) - F_X(a)$

$$a = -\infty\ , b = x\ \Rightarrow P_X(-\infty, x) = F_X(x) - F_X(-\infty) = F_X(x) \leftrightarrow (i)$$

نکته ۴–۱: ممکن است دو متغیر تصادفی مختلف، تابع توزیع یکسانی داشته باشند.

مثال ۴–۱: فرض کنید X و Y دو متغیر تصادفی باشند و $Y = -X$,

$$F_X(x)=\begin{cases}0 & x<-1\\ \frac{1}{2} & -1\le x<1\\ 1 & x\ge 1\end{cases} \quad \text{و} \quad F_Y(y)=\begin{cases}0 & y<-1\\ \frac{1}{2} & -1\le y<1\\ 1 & y\ge 1\end{cases}$$

اما $X \neq Y$ است.

مثال ۴-۲: اگر X پیشامد آمدن خط و Y پیشامد آمدن شیر در پرتاب یک سکه باشد، آنگاه این دو متغیر تصادفی دارای احتمالات مساوی و توابع توزیع یکسان هستند اما با هم برابر نیستند.

مثال ۴-۳: اگر $X=\mathrm{I}(A)$ و $P(A)=P$ آنگاه

$$F_X(x)=\begin{cases}0 & x<0\\ 1-P & 0\le x<1\\ 1 & x\ge 1\end{cases}$$

همانطور که مشخص است تابع توزیع گسسته بالا تابع توزیع دو جمله‌ای است که در نقطه 0 جهشی به اندازه‌ی $(1-P)$ و در نقطه1 جهشی به اندازه 1 دارد.

تعریف ۴-۲: اگر $F(x)$ در $X=c$ جهشی به اندازه واحد داشته باشد متغیر تصادفی X را تباهیده گویند.

$$F_X(x)=\begin{cases}0 & x<c\\ 1 & x\ge c\end{cases} \tag{۴-۲}$$

تذکر ۴-۱: اگر X متغیر حقیقی باشد آنگاه

$$F_X(+\infty)-F_X(-\infty)=P(-\infty<X<+\infty)=P\big(X^{-1}(R)\big)=P(\Omega)=1$$

اما اگر X متغیر حقیقی توسعه یافته باشد آنگاه:

$$\begin{aligned}P\big(X^{-1}(\bar{R})\big)&=p\{(-\infty<X<+\infty)\cup(X=+\infty)\cup(X=-\infty)\}\\ &=P\big(X^{-1}(R)\big)+P(X=+\infty)+P(X=-\infty)\end{aligned}$$

در این‌صورت باید مشخص کنیم که :

$$F_X(+\infty)=P\big(X^{-1}(\bar{R})\big)=1 \quad \text{یا} \quad P\big(X^{-1}(\bar{R})\big)=1$$

$$F_X(-\infty)=0 \quad \text{یا} \quad P(X=-\infty)=0$$

قضیه ۴-۱: اگر $F(x)$ تابع توزیع متغیر تصادفی X باشد آنگاه $F(x)$ تابعی نانزولی، از راست پیوسته و $F_X(-\infty) = 0$، $F_X(\infty) = 1$ است. بر عکس، هر تابع F با خواص فوق، تابع توزیع یک متغیر تصادفی بر یک فضای احتمال است.

اثبات :

(i) غیر نزولی بودن: فرض کنیم $x < x'$، آنگاه داریم:

$$A = \{X \leq x\}$$

$$B = \{X \leq x'\}$$

$$\Longrightarrow A \subseteq B \Rightarrow P(A) \leq P(B) \Rightarrow P(X \leq x) \leq P(X \leq x') \Rightarrow F_X(x) \leq F_X(x')$$

(ii) از راست پیوستگی:

دنباله $\{x'_n\}$ و $x'_n \downarrow x$ را در نظر می‌گیریم. از آنجایی که اگر $x'_n \downarrow x$ آنگاه $\{x < X \leq x'_n\} \to \phi$. در نتیجه داریم:

$$\lim_n \{x < X < x'_n\} = \bigcap_n \{x < X < x'_n\} = \phi$$

$$\Longrightarrow P(\lim\{x < X < x'_n\}) = P(\phi) = 0$$

چون مجموعه‌های $\{x < X \leq x'_n\}$ نزولی‌اند پس می‌توان جای احتمال و $\lim$ را عوض کرد در نتیجه داریم:

$$\lim P\{x < X < x'_n\} = 0 \Rightarrow \lim\{P(X \leq x'_n) - P(X \leq x)\} = 0$$

$$\Rightarrow \lim F_X(x'_n) = F_x(x) \Rightarrow F_X(x^+) = F_X(x)$$

چون این حکم برای هر دنباله $\{x'_n\}$ برقرار است پس $F(x)$ از راست پیوسته است.

(iii) فرض کنید $n = 1,2,\ldots$ $A_n = \{X \leq n\}$ در این صورت:

$$A_n \subseteq A_{n+1}$$

$$A_n \uparrow \{X < \infty\} = X^{-1}(R) = \Omega$$

$$n \to \infty \implies P(A_n) = P(\Omega) = 1$$

بنابراین $F_X(\infty) = 1$

(iv) حال اگر $n = 1,2,\ldots$ $B_n = \{X \leq -n\}$ آنگاه

$$B_{n+1} \subseteq B_n$$

$$B_n \downarrow \phi \quad n \to \infty \quad \Longrightarrow \quad P(B_n) \to 0 \quad \Longrightarrow \quad F_X(-\infty) = 0$$

اثبات عکس قضیه:

از آنجایی که هر تابع یکنوای نانزولی F روی R با $F_X(-\infty) = 0$ و $F_X(\infty) = 1$ با یک اندازه لبگ استیلجس متناظر است که می‌توان آن را به عنوان اندازه احتمال P بر (R, B) در نظر گرفت.

تذکر ۴-۲: تابع توزیع یک متغیر تصادفی گسسته با جهش‌ها صعود می‌کند.

۴-۱-۲ تعریف دیگر تابع توزیع

$$P(X < x) = F'_X(x) \tag{۴-۳}$$

این تابع توزیع تابعی غیرنزولی، از چپ پیوسته، $F'_X(+\infty) = 1$ و $F'_X(-\infty) = 0$ است.

اثبات: فقط از چپ پیوستگی را ثابت می‌کنیم، سایر خواص آن مانند F است.

فرض کنیم $x_n \uparrow x$ یعنی $x_n \subset x_{n+1} \subset \cdots$ آنگاه

$$A = \{x_n \le X < x\} \Rightarrow P\{X < x\} - P\{X < x_n\} = P\{x_n \le X < x\} \to 0$$

$$\Rightarrow P(X < x) - \lim P(X < x_n) = 0$$

$$\Rightarrow P(X < x) = \lim F'_X(x_n) = F'_X(x^-)$$

در نتیجه F' از چپ پیوسته است.

نکته ۴-۲: اگر در یک نقطه انفصال F مانند x ، $F(x)$ را به‌صورت $\left\{\frac{F(x^+)+F(x^-)}{2}\right\}$ تعریف کنیم، آنگاه تابع F نه از راست و نه از چپ پیوسته خواهد بود.

اثبات:

فرض خلف: اگر F در این نقطه هم از چپ و هم از راست پیوسته باشد آنگاه F در x پیوسته بوده و در نتیجه x دیگر نقطه‌ی انفصال نخواهد بود، پس فرض خلف باطل است.

تعریف ۴-۳: گوییم تابع توزیع $F(x)$ در $X = x$ دارای جهش است هرگاه $P(X = x) = c$.

نکته ۴-۳: هرگاه $P(X = x_0) = 0$ باشد، آنگاه تابع $F(x)$ در x_0 پیوسته خواهد بود.

اثبات:

$$P(X = x_0) = P\{X \le x_0\} - P\{X < x_0\} = 0$$

$$\Longrightarrow F(x_0) - F({x_0}^-) = 0 \Rightarrow F(x_0) = F({x_0}^-)$$

پس از چپ پیوسته است و می‌دانیم F از راست نیز پیوسته است پس F در x_0 پیوسته است.

نکته ۴-۴: توابع F و F' در تمام نقاط پیوستگی مساوی خواهند بود و در نقطه انفصال تابع $F(x)$ یا $F'(x)$ مانند x_* داریم:

$$F_X(x_*) = F_X(x_*^+) \qquad , \qquad F_X'(x_*) = F_X(x_*^-)$$

تعریف ۴-۴: اگر X فقط تعداد متناهی مقدار مانند $x_1, x_2, \dots, x_n$ را اختیار کند آنگاه $F(x)$ در هر یک از این نقاط جهش دارد و $F(x)$ فقط با این جهش‌ها صعود می‌کند، این قبیل توابع $F(x)$ را که فقط با این جهش‌ها صعود می‌کنند، توابع پله‌ای می‌نامند. تعداد جهش‌ها (تعداد x_i ها) ممکن است متناهی یا شمارا باشد.

نکته ۴-۵: از آنجا که تابع F یکنوا می‌باشد، می‌توان نوشت:

$$F(-\infty) = \lim_{x \to -\infty} F(X)$$

$$F(+\infty) = \lim_{x \to +\infty} F(X) \qquad (۴-۴)$$

مثال ۴-۴: اگر X متغیر تصادفی هندسی باشد چون X فقط مقادیر $0,1,2, \dots$ را اختیار می‌کند $F_X(x)$ در این نقاط دارای جهش بوده و یک تابع پله‌ای است. اندازه‌ی جهش درنقاط X برابر $\theta^x(1-\theta)$, $(0 < \theta < 1)$ است. همین طور تابع توزیع متغیر تصادفی پواسون، دوجمله‌ای و دو جمله‌ای منفی توابع پله‌ای هستند. توزیع متغیرهای تصادفی نرمال، گاما و مستطیلی پیوسته‌اند.

۴-۱-۳ تکیه گاه F

مجموعه مقادیر X را که به ازای آنها $F(x)$ صعود می‌کند، تکیه گاه F می‌نامند. اگر X یک متغیر تصادفی با تابع توزیع F و A تکیه گاه F باشد داریم

$$P\{X \in A^c\} = 0$$

مثلا در توزیع پواسن مجموعه‌ی اعداد صحیح نامنفی و در توزیع گاما مجموعه‌ی اعداد حقیقی نامنفی، تکیه‌گاه F هستند.

۴-۱-۴ تابع توزیع کلی

اگر $F(x)$ تابعی نامنفی، غیر نزولی و از راست پیوسته باشد که $F(-\infty) = 0$ و $F(+\infty) = c$ (c ثابت و مخالف صفر باشد) آنگاه F را تابع توزیع کلی می‌گویند.

اگر یک تابع توزیع کلی داشته باشیم با تقسیم آن بر $F(+\infty)$ تابع توزیع برای یک متغیر تصادفی بدست می‌آید.

۴-۲ تجزیه ی توابع توزیع

آمیزه توابع توزیع:

اگر $F_i(x)$ برای هر $i = 1,2,\dots,n$ تابع توزیع یک متغیر تصادفی باشد و $\alpha_i \geq 0$, $\sum_{i=1}^{n} \alpha_i = 1$ آنگاه

$$G(x) = \sum_{i=1}^{n} \alpha_i F_i(x)$$

در این صورت G نیز تابع توزیع یک متغیر تصادفی است و به آن آمیزه‌ای از توابع توزیع $F_i(x)$ با وزن‌های α_i گویند.

تذکر ۴-۳: اگر $\sum_{i=1}^{n} \alpha_i \neq 1$ آنگاه G یک تابع توزیع کلی خواهد بود.

تذکر ۴-۴: این تابع نقش مهمی در مطالعه‌ی استنباط‌های بیزی دارد و $\{\alpha_i\}$ ها توزیع پیشین نامیده می‌شوند.

قضیه ۴-۲ (قضیه‌ی تجزیه‌ی جردن):

i) نقاط انفصال هر تابع توزیع F یک مجموعه شماراست.

ii)علاوه بر این، $F = F_c + F_d$ که در آن F_c پیوسته و F_d تابع پله‌ای است و این تجزیه منحصر به فرد است.

اثبات:

قسمت i): بازه $(l, l+1)$ $\quad l = 0, \pm1, \pm2, \dots$ را در نظر می‌گیریم. فرض کنید که $x_1 < x_2 < \cdots < x_n$ نقاط انفصال $F(x)$ در $(l, l+1)$ با جهش‌های بزرگتر از $\frac{1}{m}$ باشند. چون

$$F(l) \leq F(x_1^-) < F(x_1) \leq F(x_2^-) < F(x_2) \leq \cdots \leq F(x_n^-) < F(x_n) \leq F(l+1)$$

با فرض $P(x_k) = F(x_k) - F(x_k^-)$ داریم

$$\frac{n}{m} < \sum_{k=1}^{n} P(X = x_k) = \sum_{k=1}^{n} [F(x_k) - F(\mathrm{x_k^-})]$$

$$= F(x_n) - F(\mathrm{x_1^-}) \leq F(l+1) - F(l)$$

$$\Rightarrow F(l+1) - F(l) > \frac{n}{m} = \frac{1}{m} \times \text{(تعداد نقاط انفصال)}$$

پس تعداد نقاط انفصال در بازه $(l, l+1)$ که اندازه جهش آنها بزرگتر از $\frac{1}{m}$ است، نمی‌تواند بزرگ‌تر از $m[F(l+1) - F(l)]$ باشد، بنابراین وقتی $n \to \infty$ تعداد نقاط انفصال در $(l, l+1)$ شماراست و چون تعداد این بازه‌ها شماراست قسمت اول قضیه ثابت می‌شود.

قسمت (ii): اما $F_d(x) = \sum_{x_k \leq x} P(x_k) \quad , \quad x \in R$ تابعی نامنفی، نانزولی و از راست پیوسته است. این خصوصیات برای $F_c(x) = F(x) - F_d(x)$ هم برقرار است علاوه بر این $F_c(x)$ از چپ هم پیوسته است، زیرا برای $x < x'$

$$0 \leq F_c(x') - F_c(x) = F_c(x') - F(x) - \sum_{x < x_k \leq x'} P(x_k)$$

$$= F(\mathrm{x'^-}) - F(x) - \sum_{x < x_k \leq x'} P(x_k)$$

وقتی $x \uparrow x'$، داریم $F(x) = F(\mathrm{x'^-})$ و عبارت فوق به سمت صفر میل می‌کند. در نتیجه:

$$\lim_{x \to x'} F_c(x) = F_c(x')$$

پس $F_c(x)$ از راست نیز پیوسته است.

این تجزیه منحصر به فرد است زیرا در غیر این صورت اگر

$$F(x) = F_c + F_d \text{ و } F(x) = F'_c + F'_d$$

$$\Longrightarrow F_c + F_d = F'_c + F'_d$$

$$\Longrightarrow F_c - F'_c = F'_d - F_d$$

و این غیر ممکن است زیرا سمت چپ پیوسته وسمت راست گسسته است مگر اینکه دو طرف صفر شوند. یعنی:

$$F_c = F'_c$$

$$F_d = F'_d$$

نتیجه ۴-۱: یک تابع توزیع کلی را می‌توان به دو تابع توزیع کلی به طور یکتا تجزیه کرد به طوری که یکی تابعی پیوسته و دیگری تابعی پله‌ای است.

با توجه به قضیه تجزیه لبگ می‌توان به طور یکتا قسمت پیوسته را به یک جزء مشتق‌پذیر و یک جزء منفرد تجزیه کرد.

با نرم کردن مناسب این سه جزء می‌توان هر یک را به عنوان تابع توزیع متغیر تصادفی در نظر گرفت و تابع توزیع تجزیه شده‌ی اولیه را به عنوان آمیزه‌ای از این سه نوع تابع توزیع در نظر گرفت.

۴-۳ توابع توزیع متغیرهای تصادفی برداری

۴-۳-۱ حالت دومتغیره

فرض کنید (X, Y) یک متغیر تصادفی دو بعدی باشد، تابع توزیع توأم آن به صورت زیر تعریف می‌شود:

$$F: R^2 \to [0,1]$$

$$F(x,y) = P\{X \le x, Y \le y\}$$

$$= P\{[X \le x] \cap [Y \le y]\} \quad (۵-۴)$$

خواص توابع توزیع دو متغیره:

1) $F(x,+\infty) = P\{[X \le x] \cap \mathrm{R}\} = P(X \le x) = F_X(x) \quad ,$

$$F(+\infty, y) = F_Y(y)$$

2) $P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = P\{[x_1 < X \le x_2] \cap [y_1 < Y \le y_2]\}$

با تعریف پیشامدها به‌صورت زیر:

$$\{x_1 < X \le x_2\} = \{X \le x_2\} - \{X \le x_1\} \quad , \quad \{y_1 < Y \le y_2\} = \{Y \le y_2\} - \{Y \le y_1\}$$

$$\begin{aligned}[x_1 < X \le x_2] \cap [y_1 < Y \le y_2] \\ &= \{X \le x_2, Y \le y_2\} - \{X \le x_2, Y \le y_1\} - \{X \le x_1, Y \le y_2\} \\ &+ \{X \le x_1, Y \le y_1\}\end{aligned}$$

آنگاه داریم

$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

3) $F(+\infty, +\infty) = 1$

4) $F(-\infty, y) = 0 = F(x, -\infty) \qquad \forall x, y \in R$

نکته ۴-۶: اگر مشتق $\frac{\partial^2}{\partial X \partial Y} = f(x,y)$ تقریباً در تمام نقاط وجود داشته باشد و F دارای جهش نباشد آنگاه $f(x,y) \ge 0$ و آن را تابع چگالی احتمال توأم (X,Y) گوییم، در این‌صورت داریم:

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, du\, dv$$

۴-۳-۲ تابع توزیع دو بعدی گسسته

اگر (X,Y) مقادیر شمارا $i,j = 1,2,\dots$ (x_i, y_j) را با احتمال‌های

$$P_{ij} = P\{X = x_i, Y = y_j\}$$

اختیار کنند آنگاه

$$F(x,y) = \sum_{X \le x_i, Y \le y_j} P_{ij}$$

پس $F(x,y)$ در نقاط (x_i, y_j) دارای جهش است و مشخص است که $\sum_{\mathrm{i}} \sum_{\mathrm{j}} P_{ij} = 1$.

توزیع حاشیه‌ای X برابر $P_{i.} = \sum_j P_{ij} = P(X = X_i)$

و توزیع حاشیه‌ای Y برابر $P_{.j} = \sum_i P_{ij} = P(Y = Y_j)$ است.

۴-۳-۳ حالت K متغیره

فرض کنید $(X_1, X_2, \dots, X_k)$ یک متغیر تصادفی برداری K بعدی بر Ω به توی R باشد، در این صورت مجموعه-هایی به صورت زیر اندازه‌پذیر هستند.

$$\{-\infty < X_1 \le x_1, \dots, -\infty < X_k \le x_k\} = \bigcap_{i=1}^{k} \{\infty < X_i \le x_i\} \qquad (۶-۴)$$

احتمال عبارت بالا را با $F(x_1, x_2, \dots, x_k)$ نشان می‌دهیم که تابع توزیع توأم $(X_1, X_2, \dots, X_k)$ است. وقتی یکی از مجموعه‌های بالا تهی باشد، اشتراک فوق تهی می‌شود پس داریم:

$$0 = F(-\infty, x_2, \dots, x_k) = F(x_1, -\infty, x_3, \dots, x_k) = \cdots = F(x_1, \dots, x_{k-1}, -\infty)$$

$$1 = F(+\infty, +\infty, \dots, +\infty)$$

$$F_{X_i,X_j}(x_i, x_j) = F\big(+\infty, \dots, +\infty, X_i, +\infty, \dots, +\infty, X_j, +\infty, \dots, +\infty\big) \quad , \qquad \forall\, i \le j$$

۴-۴ قضیه تطابق

۴-۴-۱ تابع توزیع بر یک زیر مجموعه‌ی فشرده‌ی R

یک تابع توزیع دلخواه به صورت تابعی نانزولی متناهی که از راست پیوسته باشد تعریف می‌شود. پس:

$$F(x^-) = \lim_{x_n \uparrow x} F(x_n) = \sup_{x_n < x} F(x_n)$$

$$F(x^+) = \lim_{x_n \downarrow x} F(x_n) = \inf_{x_n > x} F(x_n) \qquad (۴-۷)$$

موجودند و کران‌هایشان عبارتند از $F(-\infty) \ , \ F(+\infty)$.

تعریف ۴-۵: مجموعه D را در R چگال می‌گوییم اگر هر نقطه R را بتوان به عنوان یک نقطه حدی D در نظر گرفت. مثلا مجموعه تمام اعداد گویا در اعداد حقیقی چگالند. زیر مجموعه چگال را مجموعه جداساز گوییم. اعداد گویا یک زیر مجموعه جداساز در R است.

لم ۴-۱: فرض کنید F_D تابعی نانزولی متناهی باشد که بر D تعریف شده است تابع $F(x)$ را بر R به صورت زیر تعریف می‌کنیم:

$$(*) \qquad F(x) = \inf_{x_n \ge x} F_D(x_n) \qquad ; \quad x_n \in D \ , \ \ x \in R$$

ثابت می‌کنیم F تابع توزیع یک متغیر تصادفی است.

اثبات: F توسعه یافته $F(x)$ بر R است و بنا به تعریف $x_n > x \ , \ x_n \downarrow x \ ,$ $\exists\{x_n\} \subset D$، $F_D(x_n)$ تابعی یکنوا نانزولی و از پایین کراندار است.

$$\Longrightarrow F(x) = \lim_{n \to \infty} F_D(x_n) \quad , \qquad \forall\, x \in R$$

در D توابع F و F_D بر هم منطبق هستند و F_D تابعی نانزولی است. در نتیجه F نانزولی و از راست هم طبق رابطه (*) پیوسته است.

اگر $F_D(-\infty) = F(-\infty) = 0$ و $F_D(\infty) = F(\infty) = 1$ آنگاه F تابع توزیع یک متغیر تصادفی خواهد بود.

قضیه ۴-۳ (قضیه تطابق): هر اندازه احتمال P بر (R, B) یک تابع توزیع F را بطور یکتا معین می‌کند و این تناظر به صورت زیر است:

$$P(-\infty, x] = F(x) \tag{۸-۴}$$

$$F(\infty) = 1 , F(-\infty) = 0$$

بر عکس، هر تابع توزیع یک متغیر تصادفی، یک اندازه‌ی احتمال P را بطور یکتا بر (R, B) معین می‌کند و تناظر مربوطه همان $P(-\infty, X] = F(X)$ است.

اثبات: اول: فرض می‌کنیم $a < b$ آنگاه داریم:

$$P(a, b] = F(b) - F(a) \geq 0 \Rightarrow F(b) \geq F(a)$$

پس F غیر نزولی است.

حال فرض می‌کنیم $\{X_n\}$ یک دنباله باشد به طوری که

$x_1 > x_2 > \cdots > x_n$ و $x_n \downarrow x$

آنگاه

$$F(x_n) - F(x) = P(x, x_n) \downarrow 0 \Rightarrow F(x_n) \downarrow F(x)$$
$$\Rightarrow \lim_{x_n \to x^+} F(x_n) = F(x) \Rightarrow F(x^+) = F(x)$$

پس F از راست پیوسته است و طبق فرض چون $F(+\infty) = 1 \ , \ F(-\infty) = 0$ پس بنابر عکس قضیه (۴-۱)، F تابع توزیع یک متغیر تصادفی است.

دوم: P را اندازه‌ی لبگ و $\Omega = (0,1)$ در نظر می‌گیریم. دو حالت را بررسی می‌کنیم:

1) F پیوسته و اکیداً صعودی باشد.

2) F تابعی پله‌ای باشد یا تابعی اکیداً صعودی نباشد.

حالت ۱) اگر F پیوسته و اکیداً صعودی باشد آنگاه F یک نگاشت یک به یک از R به $(0,1)$ می‌باشد. فرض کنید φ تصویر معکوس F باشد.

$$\varphi = F^{-1} : (0,1) \to R$$

برای $0 < \omega < 1$ فرض کنید داشته باشیم $X(\omega) = \varphi(\omega)$. چون φ صعودی است، به طور یقین X یک متغیر تصادفی است.

اگر $0 < u < 1$ آنگاه

$$\varphi(u) \leq x \Leftrightarrow F^{-1}(u) \leq x \Leftrightarrow F\left(F^{-1}(u)\right) \leq F(x) \Leftrightarrow u \leq F(x)$$

چون P یک اندازه‌ی لبگ است، پس:

$$P(X \leq x) = P(\omega \in (0,1), \varphi(\omega) \leq x) = P\left(0 < \omega < 1, \omega \leq F(x)\right)$$

$$= P\left(o < \omega < F(x)\right) = F(x) - 0$$

$$= F(x)$$

$$\Longrightarrow P(-\infty, x] = F(x)$$

حالت ۲) اگر F ناپیوسته باشد یا اکیداً صعودی نباشد، تعریف می‌کنیم:

$$\varphi(u) = \inf\left(x: u \leq F(x)\right)$$

چون F غیرنزولی است، $[x: u \leq F(x)]$ فاصله‌ای است که به $+\infty$ میل می‌کند.

چون F از راست پیوسته است، این فاصله از چپ بسته است یعنی برای $0 < u < 1$

$$[x: u \leq F(x)] = I_{[\varphi(u), +\infty]}(x)$$

بنابراین:

$$\Longrightarrow \varphi(u) \leq x \iff u \leq F(x)$$

و بقیه را مشابه حالت ۱ ادامه می‌دهیم و اثبات کامل می‌شود.

۴-۴-۲ تابع توزیع تجربی

فرض کنید $x_{(1)}, x_{(2)}, \dots, x_{(n)}$، n مقدارمشاهده شده یک متغیر تصادفی باشند که به ترتیب صعودی مرتب شده‌اند و آنها را آماره‌های ترتیبی گوییم. تعریف می‌کنیم:

$$F_n(x) = \begin{cases} 0 & x < x_{(1)} \\ \frac{k}{n} & x_k \leq x < x_{k+1} \qquad , \quad k = 1,2, \dots, n-1 \\ 1 & x \geq x_{(n)} \end{cases} \qquad (۴-۹)$$

آنگاه $F_n(x)$ تابع توزیع تجربی X نامیده می‌شود. این تابع پله‌ای بوده و در هر مقدار مشاهده شده جهشی به اندازه $\frac{1}{n}$ دارد.

اگر $F_n(x)$ به $F(x)$ همگرا باشد، آنگاه $F_n(x)$ تقریبی برای $F(x)$ است و $\sup|F_n(x) - F(x)|$ اندازه‌ای از این تقریب را نشان می‌دهد.

در ادامه نشان می‌دهیم این اندازه، خواص یک فاصله یا متر را دارد.

1) $\sup|F_n(x) - F(x)| \geq 0$

2) $\sup|F(x) - F_n(x)| = \sup|F_n(x) - F(x)|$

3) $\sup|F_n(x) - F(y)| \ \leq \sup\{|F_n(x) - F(x)| + |F(x) - F(y)|\}$

$$\leq \sup\{|F_n(x) - F(x)|\} + \sup\{|F(x) - F(y)|\}$$

به فاصله‌ی فوق فاصله‌ی کولموگروف-اسمیرنف بین F_n و F گویند. این فاصله را می‌توان برای اثبات اینکه آیا الگوی نظری رضایت‌بخش است و با مقادیر مشاهده شده توافق دارد، بکار برد.

قضیه ۴-۴: فرض کنید F و G بر بازه‌ی $[a, b]$ دو تابع غیرنزولی و از راست پیوسته باشند. اگر F و G در $(a, b]$ نقاط ناپیوسته‌ی مشترک نداشته باشند، آنگاه خواهیم داشت:

$$\int_{(a,b]} G(x)\, dF(x) = \Delta_{ba}\, FG - \int_{(a,b]} F(x)\, dG(x)$$

$$= F(b)G(b) - F(a)G(a) - \int_{(a,b]} F(x)\, dG(x) \qquad (۱۰-۴)$$

نکته ۴-۷: در قضیه‌ی فوق اگر F و G نقاط ناپیوسته‌ی مشترک داشته باشند، داریم:

$$\int_{(a,b]} G(x)\, dF(x) = \Delta_{ba}\, FG - \int_{(a,b]} F(x)\, dG(x)$$

$$+ \sum_{z} \big(F(z) - F(z^-)\big)\big(G(z) - G(z^-)\big)$$

که در آن Z مجموعه نقاط ناپیوستگی مشترک F و G است.

مثال ۴-۵: فرض کنید F تابع توزیع متغیر تصادفی X و G تابع توزیع متغیر تصادفی Y باشد، ثابت کنید:

$$E\big(F(Y)\big)+E\big(G(X)\big)=1$$

حل:

$$E\big(F(Y)\big)+E\big(G(X)\big)=\int F(y)dG(y)+\int G(x)dF(x)=\Delta FG$$
$$=F(+\infty)G(+\infty)-F(-\infty)G(-\infty)=1$$

مثال ۴-۶: فرض کنید $F_X(x)=x\mathrm{I}_{[0,\frac{1}{2})}+\mathrm{I}_{[\frac{1}{2},+\infty)}$، مقدار $\int F\,dF$ را بدست آورید.

$$\int F\,dF=\Delta FF-\int F\,dF+\sum_z\big(F(z)-F(z^-)\big)^2$$

$$\Rightarrow 2\int F\,dF=\Delta FF+\sum_z\big(F(z)-F(z^-)\big)^2$$

$$\Rightarrow \int F\,dF=\frac{1}{2}\left[\Delta FF+\sum_z\big(F(z)-F(z^-)\big)^2\right]$$

به عنوان مثال اگر $z=\frac{1}{2}$، آنگاه

$$\int F\,dF=\frac{1}{2}\left[1+\left(1-\frac{1}{2}\right)^2\right]=\frac{5}{8}$$

۴-۵ مسائل

۱- اگر یک تابع مانند F دارای خواص $F(+\infty)=1$و $F(-\infty)=0$ باشد آنرا تابع توزیع یک متغیر تصادفی گویند. در آمار برای برازش یک الگو به یک پدیده‌ی فیزیکی، از میان خانواده‌ای از توابع توزیع، تابعی که مناسب‌تر است را انتخاب می‌کنیم. با تغییر پارامترهای مشخص کننده یک توزیع معین، خانواده‌ای از توابع توزیع بدست می‌آیند.

الف) خانواده نرمال:

$$F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt \quad , \quad -\infty<\mu<+\infty \ ,\sigma>0$$

ب) خانواده نمائی:

$$F(x) = \int_0^x \frac{1}{\theta} e^{-\frac{t}{\theta}}\, dt \quad , \qquad \theta > 0$$

ج) خانواده کوشی:

$$F(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{dt}{1+(t-\mu)^2} \quad , -\infty < \mu < +\infty$$

د) خانواده گاما:

$$F(x) = \int_o^x \frac{1}{\Gamma(p)\theta^p} e^{-\frac{t}{\theta}} t^{p-1}\, dt \quad , \;\; p > 0 \;\; , \theta > 0$$

ه) توابع توزیع گسسته مانند دوجمله‌ای، پواسن، هندسی و دوجمله‌ای منفی دارای پارامترهای λ , p و غیره می‌باشند که با تغییر آنها، خانواده‌ای از توابع توزیع بدست می‌آیند.

۲- دو تابع توزیع را معادل می‌گویند هرگاه آنها تنها در یک مقدار ثابت با هم متفاوت باشند. نشان دهید که این رابطه یک رابطه‌ی هم ارزی است. توجه کنید که یک توزیع احتمال منحصر به فرد وجود دارد که با یک کلاس هم ارزی متناظر است.

۳- تابع $F: R \to R$ را به صورت زیر در نظر می‌گیریم:

$$F(x) = \begin{cases} 0 & x < 0 \\ \alpha + k e^{-\frac{x^2}{2}} & x \geq 0 \end{cases}$$

مقادیر $k \;\; , \;\; \alpha$ را به قسمی تعیین کنید که $F(x)$ تابع توزیع باشد.

۴- اگر F تابع توزیع یک متغیر تصادفی باشد، مطلوبست تابع توزیع:

$$i)\; Y = 3X^2 + 2 \qquad\qquad , \qquad\qquad ii)\; Y = F(x)$$

۵- اگر X متغیر تصادفی نرمال استاندارد باشد، نشان دهید که:

$$P\{a < X^2 < b\} = 2P\{-\sqrt{b} < X < -\sqrt{a}\}$$

۶- پیچش دو تابع $F_1(x)$ و $F_2(x)$ به صورت زیر تعریف شده و با $F = F_1 * F_2$ نشان داده می‌شود.

$$F(z) = \int F_1(z-y)dF_2(y) = \int F_2(z-x)dF_1(x) \qquad\qquad (z \in R)$$

ثابت کنید که:

الف: اگر F_1 و F_2 توابع توزیع متغیرهای تصادفی باشند آنگاه F نیز تابع توزیع است.

ب: اگر به ازای $z < 0$ ، F_1 و F_2 صفر شوند آنگاه

$$F(z) = \int_0^z F_1(z-y)dF_2(y)$$

۷- فرض کنید:

$$F(x,y) = \begin{cases} 0 & x<0\,, y<0\ \ , x+y<1 \\ 1 & o.w \end{cases}$$

نشان دهید که $F(x,y)$ یک تابع توزیع نیست.

۸- نشان دهیدکه

$$f(x,y) = \frac{1}{4}\{1 + xy(x^2 - y^2)\}\ \ , (|x| < 1, |y| < 1)$$

تابع چگالی یک متغیر تصادفی دوبعدی است. توابع توزیع کناری را بدست آورید.

فصل پنجم

امید ریاضی و گشتاورها

یکی از مفاهیم بسیار مفید در مسائل مربوط به بخش متغیرهای تصادفی امید ریاضی می‌باشد. در این بخش امید ریاضی متغیرهای تصادفی را در چند مرحله تعریف می‌کنیم. ابتدا امید ریاضی یک متغیر تصادفی ساده و سپس متغیر تصادفی نامنفی و در نهایت امید ریاضی یک متغیر تصادفی دلخواه رابا ذکر یک سری از ویژگی‌های آنها بیان می‌کنیم. درانتها نگاهی به گشتاورها و نامساوی‌های مهم خواهیم داشت.

۵–۱ امید ریاضی متغیرهای تصادفی

یادآوری (فصل ۲)

تعریف متغیر تصادفی: اگر Ω فضای نمونه ای و A سیگما میدان پیشامدهای مربوط به آزمایش باشد. هر تابع حقیقی که بر Ω تعریف شده باشد و نسبت به A اندازه‌پذیر باشد را یک متغییر تصادفی می‌نامیم.

۵–۱–۱ امید ریاضی متغیرتصادفی ســاده

متغیر تصادفی ساده، متغیری تصادفی است که مجموعهٔ مقادیر آن متناهی باشد.

فرض کنید X متغیرتصادفی ساده برفضای احتمال $(\Omega, \mathcal{F}, P)$ با مجموعهٔ مقادیر $\{x_1, x_2, \dots, x_n\}$ باشد. برای $1 \le K \le n$، تعریف می‌کنیم:

$$A_k = \{\omega \in \Omega: X(\omega) = x_k\} = X^{-1}(x_k) = (X = x_k)$$

دراین‌صورت A_k ها افرازی از Ω تشکیل می‌دهند و X را می‌توانیم به صورت زیر بنویسیم:

$$X = \sum_{k=1}^{n} x_k \quad I_{A_k} \quad , \quad A_k \in A \quad , \quad \bigcup_{k=1}^{n} A_k = \Omega$$

اگر x متغیر تصادفی ساده بصورت بالا باشد، آنگاه امید ریاضی X که آن را با نماد $E(X)$ نشان می‌دهیم، بصورت زیر تعریف می‌کنیم:

$$E(X)=\sum_{k=1}^{n} x_k\, P(\omega\in A_k)=\sum_{k=1}^{n} x_k\, P(X=x_k)\quad ;\quad X(\omega)=\begin{cases} x_1 & \omega\in A_1\\ x_2 & \omega\in A_2\\ . & .\\ . & .\\ . & .\\ x_n & \omega\in A_n\end{cases}$$

که آن را انتگرالX نسبت به اندازهٔ P نیز می‌گویند و آن را به صورت$\int X$ یا $\int X\,dp$ یا $\int_\Omega X\,dp$ نیز نمایش می‌دهند.

اگر $X=I_A$ باشد، داریم:

$$X=I_A=\begin{cases}1 & \omega\in A\\ 0 & \omega\in A^c\end{cases}$$

$$\Rightarrow E(X)=p(\omega\in A)=P(A)$$

۵–۱–۲ خاصیت های مهم امید ریاضی متغیرهای تصادفی ساده

۱ - خاصیت خطی:

اگر X و Y متغیرهای تصادفی ساده باشند، آنگاه

$$E(aX\pm bY)=aE(X)\pm bE(Y)$$

که در آن a و b ثابت‌های دلخواه می‌باشند.

اثبات: فرض کنید $X=\sum_{j=1}^{n} x_j\, I_{A_j}$ و $Y=\sum_{k=1}^{n} y_k\, I_{B_k}$ که در آن $\{A_j\}$ و$\{B_k\}$ افزارهایی از Ω هستند.

$$aX\pm bY=\sum_{j=1}^{n}\sum_{k=1}^{n}\left(ax_j\pm by_k\right)\mathrm{I}(A_jB_k)$$

بنابر تعریف امید داریم:

$$E(aX\pm bY)=\sum_{j=1}^{n}\sum_{k=1}^{n}\left(ax_j\pm by_k\right)\mathrm{P}(A_jB_k)$$

$$=a\sum_{j=1}^{n} x_j\sum_{k=1}^{n} P(A_jB_k)\pm b\sum_{k=1}^{n} y_k\sum_{j=1}^{n} P(A_jB_k)$$

عبارت بالا با استفاده از روابط زیر ساده سازی میشود:

$$\sum_{k=1}^{n} P(A_j B_k) = P\left(\sum_{k=1}^{n} A_j B_k\right) = P(A_j)$$

$$\sum_{j=1}^{n} P(A_j B_k) = P\left(\sum_{k=1}^{n} A_j B_k\right) = P(B_k)$$

داریم:

$$= a\sum_{j=1}^{n} x_j\, P(A_j) \pm b\sum_{k=1}^{n} y_k\, P(B_k)$$

$$= aE(X) \pm bE(Y)$$

۲ - حفظ مقیاس:

$$E(cX) = cE(X)$$

اثبات:

$$\mathrm{cX} = \mathrm{c}\sum_{k=1}^{n} x_k\, I_{A_k} = \sum_{k=1}^{n} (cx_k)\, I_{A_k}$$

بنابر تعریف امید

$$E(\mathrm{cX}) = \sum_{k=1}^{n} cx_k \mathrm{P}(A_k) = c\sum_{k=1}^{n} x_k \mathrm{P}(A_k) = cE(X)$$

۳ - نامنفی بودن:

$$\mathrm{X} \geq 0 \quad \Rightarrow \quad \mathrm{E(X)} \geq 0$$

اثبات:

اگر $X = \sum_{k=1}^{n} x_k\, I_{A_k} \geq 0$، آنگاه به ازای تمام مقادیر k، $x_k > 0$ می‌باشد. بنابراین داریم

$$E(X) = \sum_{k=1}^{n} x_k \mathrm{P}(A_k) \geq 0$$

حتی اگر به ازای بعضی مقادیر k که برای آنها $P(A_k) = 0$، $x_k \geq 0$ است. پس برای اینکه $E(X)$ نامنفی باشد کافی است که X بر مجموعه‌هایی که احتمال غیر صفر دارند نامنفی باشد. یعنی $x \geq 0$، $a.s$.

$$X \geq 0 \quad a.s \quad \Rightarrow P\{\omega: x(\omega) \geq 0\} = 1$$

یا

$$P\{\omega: x(\omega) < 0\} = 0$$

اگر $X = \sum_{k=1}^{n} x_k I_{A_k} \geq 0$ و $X \geq 0$، $a.s$ آنگاه $x_k \geq 0$ است. اگر $P(A_k) \geq 0$ و $x_k < 0$، نتیجه می‌دهد که $P(A_k) = 0$ می‌باشد، چون:

$$E(X) = \sum_k x_k P(A_k) = \sum_{x_k \geq 0} x_k P(A_k) + \sum_{x_k < 0} x_k P(A_k)$$

جملهٔ اول سمت راست نامنفی و جملهٔ دوم صفر می‌باشد. پس خواهیم داشت $E(X) \geq 0$.

نتیجه ۵-۱: اگر $X \geq Y$، $a.s$ آنگاه $E(X) \geq E(Y)$.

$$X \geq Y \rightarrow X - Y \geq 0 \Rightarrow E[X - Y] \geq 0 \Rightarrow E(X) - E(Y) \geq 0$$
$$\Rightarrow E(X) \geq E(Y)$$

نتیجه ۵-۲: اگر $X = 0$ آنگاه $E(X) = 0$.

$$X = 0 \ a.s \Rightarrow P\{\omega: X(\omega) \neq 0\} = 0$$
$$\Rightarrow \quad P(\{\omega: X(\omega) < 0\} \cup \{\omega: X(\omega) > 0\}) = 0$$
$$\Rightarrow P(\{\omega: X(\omega) < 0\}) + (\{\omega: X(\omega) > 0\}) = 0$$
$$\Rightarrow P(\{\omega: X(\omega) < 0\}) = 0 \ , \ P(\{\omega: X(\omega) > 0\}) = 0$$
$$\begin{cases} X(\omega) \geq 0 \Rightarrow E\big(X(\omega)\big) \geq 0 \\ X(\omega) \leq 0 \Rightarrow E\big(X(\omega)\big) \leq 0 \end{cases}$$
$$\Rightarrow E\big(X(\omega)\big) = 0$$

نمونه‌ای دیگر از امید ریاضی متغیر تصادفی برگرفته ازکتاب احتمال پیشرفته[1]

امید ریاضی متغیر تصادفی گسسته:

[1] . Billingsley - برگرفته از کتاب احتمال

امید برای متغیر تصادفی گسسته دارای محاسبات ساده‌تری می‌باشد.

فرض کنید X متغیر تصادفی گسسته باشد که مقادیر ... $x_1, x_2,$ با احتمال $P(X=x_k)= P_k$ می پذیرد. برای k≥1 و g بورل داریم:

$$E_g(\mathbf{X}) = \sum_{k=1}^{\infty} g(x_k)P(X = x_k) = \sum_{k=1}^{\infty} p_k\, g(x_k)$$

$$\sum_{k=1}^{\infty} P_k\, |g(x_k)| < \infty$$

۵–۱–۳ امید متغیر تصادفی نامنفـی

فرض کنید که X یک متغیر تصادفی نامنفی باشد، ولی لزوماً ساده نباشد، می‌دانیم که دنباله‌ای از توابع ساده نامنفی صعودی، مانند $\{X_n\}$ یافت می‌شوند که به X همگرا هستند. چون $X_{n+1} \geq X_n \geq 0$ پس (طبق نتیجه ۵-۱)

$E(X_{n+1}) \geq E(X_n) \geq 0$، در نتیجه $\{E(X_n)\}$ نیز یکنوا و صعودی است. لذا به سمت $\lim_{n\to\infty} E(X_n)$ میل می‌کند که می‌تواند متناهی یا $+\infty$ باشد.

قضیه ۵-۱: فرض کنید X نامنفی و $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی نامنفی ساده و صعودی باشد که به X میل می‌کند، در این صورت برای هر متغیر تصادفی ساده و نامنفی مانند X که $Y \leq X$ و $a.s\ \ n \to \infty$ داریم:

$$\lim_{n\to\infty} E(X_n) \geq E(Y)$$

اثبات: فرض کنید $\varepsilon > o$ داده شده باشد. به ازای هر n مجموعه A_n را به صورت زیر تعریف می‌کنیم:

$$A_n = \{\omega: X_n(\omega) > Y(\omega) - \varepsilon\} = \{X_n > Y - \varepsilon\}$$

بنابر صعودی بودن $\{X_n\}$ نتیجه می‌شود که A_n نیز صعودی است و

$$A_n = \{X_n > Y - \varepsilon\} \uparrow \Omega$$

با استفاده از خاصیت خطی امید ریاضی داریم:

$$E(X_n) = \int X_n dp = \int X_n\, I_{A_n} dp + \int X_n\, I_{A_n^c} dp$$

$$\geq \int X_n\, I_{A_n}\ \ dp \quad (X_n \geq 0)$$

می دانیم: $\qquad X_n > Y - \varepsilon$

$$\geq \int (Y-\varepsilon)\, I_{A_n} dp$$

$$= \int Y\, I_{A_n} dp - \varepsilon \int I_{A_n}\, dp$$

$$= \int Y\, dp - \int \mathrm{Y}\; I_{A_n^c} dp - \varepsilon\, P(A_n)$$

فرض می‌کنیم$\{M = \max Y(\omega) < \infty\}$ درنتیجه:

$$E(X_n) \geq E(Y) - M\, P(A_n^c) - \varepsilon\, P(A_n)$$

دیدیم که$A_n \uparrow \Omega$ ولذا $A_n^c \downarrow \emptyset$. درنتیجه $P(A_n) \uparrow 1$ و $P(A_n^c) \downarrow 0$

بنابراین:

$$\lim E(X_n) \geq E(Y) - \varepsilon$$

چون نابرابری اخیر به ازای هر $\varepsilon > 0$ برقرار است، بنابراین:

$$\lim_{n\to\infty} E(X_n) \geq E(Y)$$

قضیه ۵-۲: فرض کنید$X_n \uparrow X$, $Y_n \uparrow X$ به طوری کهX_n و Y_n متغیرهای تصادفی ساده نامنفی باشند، دراین صورت داریم

$$\lim_{n\to\infty} E(X_n) = \lim_{n\to\infty} E(Y_n)$$

اثبات: باتوجه به اینکه $Y_n \leq X = \lim_{n\to\infty} X_n$ و باتوجه به قضیه قبل داریم

$$\lim_{n\to\infty} E(Y_n) \leq \lim_{n\to\infty} E(X_n) \qquad (۵-۱)$$

و نیز با درنظر گرفتن $X_n \leq X = \lim_{n\to\infty} Y_n$:

$$\lim_{n\to\infty} E(X_n) \leq \lim_{n\to\infty} E(Y_n) \qquad (۵-۲)$$

در نتیجه بنابر (۵-۱) و (۵-۲) :

$$\lim_{n\to\infty} E(X_n) = \lim_{n\to\infty} E(Y_n)$$

۵–۱–۴ امید متغیر تصادفی دلخواه

برای هرمتغیر تصادفی مانند X دو متغیر تصادفی نامنفی مانند X^+ و X^- به صورت زیر تعریف می‌شود:

$$X^+ = X \ I_{(X \geq 0)} \quad , \quad X^- = -X \ I_{(X<0)} \ \Longrightarrow \ X = X^+ - X^-$$

بنابراین چون X^+ و X^- نامنفی هستند پس امید ریاضی آن‌ها تعریف می‌شود، برای هر متغیر دلخواه امید ریاضی به صورت زیر است:

$$E(X) = E(X^+) - E(X^-)$$

به شرطی که سمت راست به صورت $\infty - \infty$ نباشد، $E(X)$ وجود دارد. بدیهی است که $E(X)$ متناهی خواهد بود اگر و فقط اگر $E(X^+)$ و $E(X^-)$ هر دو متناهی باشد، X انتگرال‌پذیر است اگر و فقط اگر X^+ و X^- انتگرال‌پذیر باشند.

با توجه به فضای احتمال $(R, \mathfrak{B}, P_X)$، $E(X)$ را می‌توان به صورت زیر تعریف کرد:

اگر X مقادیر... $x_1, x_2, \dots x_n$ را با احتمال ... $P_1, P_2, \dots P_n$ اختیار کند، آنگاه:

$$E(X) = \sum_k x_k \ P(A_k)$$

$$= \sum_k x_k \ P(X = x_k)$$

(۵-۳)

و اگر P_X اندازه لبگ استیلجس باشد:

$$F_X(x) = P_x(-\infty, x) = \int_{-\infty}^{x^+} dF(x)$$

آنگاه $E(X) = \int_{-\infty}^{+\infty} x \ dF(x)$ را انتگرال لبگ استیلجس بر R گوییم. اگر F_X دارای مشتق $f(x)$ باشد، آنگاه:

$$E(X) = \int_{-\infty}^{0} x \ f(x) dx + \int_{0}^{+\infty} x \ f(x) dx \qquad (۵\text{-}۴)$$

$$= \int_{-\infty}^{+\infty} x \ dF(x)$$

مثال ۵- ۱: فرض کنید X یک متغیر تصادفی کوشی با تابع چگالی احتمال $f(x)=\frac{1}{\pi(1+x^2)}$ باشد داریم:

$$E(X^+)=\frac{1}{\pi}\int_0^\infty \frac{x}{1+x^2}dx=\frac{1}{2\pi}\ln(1+x^2)\Big|_0^\infty=\infty$$

و

$$E(X^-)=\frac{1}{\pi}\int_{-\infty}^0 \frac{x}{1+x^2}dx=\infty$$

$$\Longrightarrow E(X^+)=E(X^-)=\infty$$

بنابراین $E(X)$ وجود ندارد. ولی حد انتگرال لبگ موجود و برابر صفر خواهد بود.

$$\frac{1}{\pi}\int_0^\infty \frac{x}{1+x^2}dx=\lim_{A\to\infty}\int_{-A}^{A}\frac{x\,dx}{\pi(1+x^2)}$$

مثال ۵-۲: فرض کنید X مقادیر $\{0,\pm1,\pm2,\dots\}$ را با احتمال‌های زیر اختیار کند

$$P(X=0)=0$$

$$P\{X=k\}=P\{X=-k\}=\frac{3}{\pi^2k^2}\quad,\quad k=1,2,\dots$$

به این ترتیب X یک متغیر تصادفی می‌باشد که (طبق رابطه (۵-۳)):

$$E(X^+)=\sum_{k=1}^{\infty}\frac{3\,k}{\pi^2k^2}=\frac{3}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k}=\infty$$

$$E(X^-)=\sum_{k=-\infty}^{0}\frac{3\,k}{\pi^2k^2}=\frac{3}{\pi^2}\sum_{k=-\infty}^{0}\frac{1}{k}=\infty$$

$$\Longrightarrow E(X)=\infty-\infty$$

حال اگر X مقادیر $\{1,2,\dots\}$ را با احتمال‌های $P(X=k)=\frac{6}{\pi^2k^2}$ اختیار کند، آنگاه:

$$E(X^+)=\infty$$

$$E(X^-)=0$$

$$\Longrightarrow E(X)=\infty-0=\infty$$

یعنی $E(X)$ وجود دارد، گرچه مقدار آن ∞ است.

۵-۱-۵ خواص امید ریاضی متغیر تصادفی دلخواه

در بخش قبل چند خاصیت از متغیرتصادفی ساده را مورد بررسی قرار دادیم. در این جا به بررسی چند خاصیت دیگر یک متغیرتصادفی دلخواه و نامنفی می‌پردازیم. ابتدا یک لم را بررسی می‌کنیم.

لم۵-۱: اگر $X \geq 0$ و انتگرال‌پذیر باشد، X می‌تواند حداکثر بر یک مجموعه با احتمال صفر ∞ شود.

اثبات: فرض می‌کنیم که این‌طور نباشد. یعنی مجموعه‌ای باشد که X بر آن ∞ شود. یعنی:

$$X = X\, \mathrm{I}_{\{X<\infty\}} + X\, \mathrm{I}_{\{X=\infty\}}$$

و $X\, \mathrm{I}_{\{X<\infty\}} = \lim X_n$ که در آن X_n ساده است و داریم:

$$X_n \uparrow X\, \mathrm{I}_{\{X<\infty\}}$$

$$X = \lim(X_n + \infty \times \mathrm{I}_{\{X=\infty\}})$$

جملهٔ اخیر به ازای $P(X=\infty) > 0$ بی‌نهایت می‌شود. و چون X انتگرال‌پذیر است، بنابراین $E(X) < \infty$ نتیجه می‌دهدکه $E(X^+)$ و $E(X^-)$ محدود هستند. داریم: $\left[X^- < \infty \text{ و} X^+ < \infty\right], a.s$،

بنابراین $[X < \infty], a.s$.

قضیه ۵-۳: فرض کنید $E(X)$، $E(Y)$ و $E(X+Y)$ وجود داشته باشند،

الف) $X \geq 0 \;\; a.s \implies E(X) \geq 0$

ب) $E(cX) = cE(X)$

پ) $E(X+Y) = E(X) + E(Y)$

ت) $X \geq Y \;\; a.s \implies E(X) \geq E(Y)$

ج) $X = 0 \;\; a.s \implies E(X) = 0$

ح) X انتگرال‌پذیر است اگر وتنها اگر $|X|$ انتگرال‌پذیر باشد.

د) اگر Y انتگرال‌پذیر باشد و $|X| \leq Y$ آنگاه X انتگرال‌پذیر است.

اثبات:

الف) اگر $X \geq 0$ ،$a.s$ آنگاه دنباله‌ای از توابع صعودی ساده مانند $\{X_n\}$ وجود دارد که $X_n \uparrow X$. از خاصیت نامنفی بودن متغیر تصادفی ساده نتیجه می‌شود که $0 \geq E(X_n)$. چون X_n صعودی (غیر نزولی) است، پس $E(X_n)$ نیز صعودی یا غیرنزولی است. چون $E(X) = \lim E(X_n)$ در نتیجه $E(X) \geq 0$.

ب) فرض کنید $X \geq 0$، در این صورت دنباله‌ای از توابع ساده مانند $\{X_n\}$ یافت می‌شوند که $0 \leq X_n \uparrow X$. فرض کنید که $c > 0$ باشد، در این صورت $0 \leq cX_n \uparrow cX$ و به موجب خاصیت امید ریاضی داریم

$$E(cX_n) = cE(X_n) \uparrow cE(X)$$

به عبارتی دیگر:

$$E(cX) = \lim_{n\to\infty} E(cX_n) = \lim_{n\to\infty} cE(X_n) = c\lim E(X_n) = cE(X)$$

حال اگر X متغیر تصادفی دلخواه باشد، داریم:

$$cX = cX^+ - cX^-$$

با امید گرفتن از طرفین داریم:

$$\begin{aligned} E(cX) &= E(cX^+) - E(cX^-) \\ &= c[E(X^+) - E(X^-)] \\ &= cE(X) \end{aligned}$$

پ) برای اثبات $E(X+Y) = E(X) + E(Y)$ ابتدا حالتی را بررسی می‌کنیم که $\{X_n\}$ و $\{Y_n\}$ هایی که $0 \leq X, Y$ هستند. در این صورت دنباله‌هایی از متغیر تصادفی ساده مانند $\{X_n + Y_n\}$ هستند که:

$$0 \leq X_n \uparrow X \;\;,\quad 0 \leq Y_n \uparrow Y \;\;,\quad 0 \leq X_n + Y_n \uparrow X + Y$$

$$\begin{aligned} E(X+Y) &= \lim E(X_n + Y_n) \\ &= \lim E(X_n) + \lim E(Y_n) \\ &= E(X) + E(Y) \end{aligned}$$

حال اگر X متغیر تصادفی دلخواه باشد آنگاه برای ادامه اثبات Ω را بر مجموعه‌های $A_1, A_2, \dots, A_6$ به صورت زیر افراز می‌کنیم:

$$A_1 = \{X \geq 0 \;,\quad Y < 0 \;,\quad X + Y \geq 0\}$$

$$A_2 = \{X < 0 \;,\quad Y \geq 0 \;,\quad X + Y \geq 0\}$$

$$A_3 = \{X \geq 0 \ , \quad Y \geq 0 \quad , \quad X + Y \geq 0\}$$

$$A_4 = \{X < 0 \ , \quad Y \geq 0 \quad , \quad X + Y < 0\}$$

$$A_5 = \{X \geq 0 \ , \quad Y < 0 \quad , \quad X + Y < 0\}$$

$$A_6 = \{X < 0 \ , \quad Y < 0 \quad , \quad X + Y < 0\}$$

فرض می‌کنیم که حداقل یکی از متغیرهای تصادفی مثلاً Y، انتگرال‌پذیر باشد و نیز $E(Y) < \infty$ باشد.

برای افراز A_1 داریم: $X + Y \geq 0 \ , -Y > 0$. پس خاصیت جمعی برای توابع نامنفی برقرار است یعنی:

$$E\left\{(X+Y)\mathrm{I}_{(A_1)}\right\} + E\left\{(-Y)\mathrm{I}_{(A_1)}\right\} = E(X\mathrm{I}_{(A_1)})$$

در نتیجه با افزودن مقدار متناهی $E\left\{(Y)\mathrm{I}_{(A_1)}\right\}$ داریم

$$E\left\{(X+Y)\mathrm{I}_{(A_1)}\right\} = E\left\{(X)\mathrm{I}_{(A_1)}\right\} + E(Y\,\mathrm{I}_{(A_1)})$$

برای A_2 نیز داریم $X + Y \geq 0 \ , -X > 0$. پس مانند قبل داریم

$$E\left\{(X+Y)\mathrm{I}_{(A_2)}\right\} + E\left\{(-X)\mathrm{I}_{(A_2)}\right\} = E\left(Y\,\mathrm{I}_{(A_2)}\right) < \infty$$

در نتیجه داریم $E\left(Y\,\mathrm{I}_{(A_2)}\right) < \infty$. بعد با افزودن مقدار متناهی $E\left(X\mathrm{I}_{(A_2)}\right)$ به طرفین به نتیجه مطلوب می‌رسیم.

برای A_3 که در آن $\sum_{i=1}^{n} X_i Y_i$، $Y \geq 0$ و $X \geq 0$ می‌باشد در قبل بررسی کردیم.

برای افراز A_4 نیز داریم: $-(X+Y) \geq 0 \ , Y \geq 0$. بنابراین

$$E\left\{-(X+Y)\mathrm{I}_{(A_4)}\right\} + E\left\{(Y)\mathrm{I}_{(A_4)}\right\} = E(-X\mathrm{I}_{(A_4)})$$

با کم کردن عبارت $E(Y\mathrm{I}_{(A_4)})$ از طرفین و تغیر علامت طرفین، به نتیجه مطلوب می‌رسیم.

برای افراز A_5 نیز $-(X+Y) \geq 0 \ , X \geq 0$ پس داریم

$$E\left\{-(X+Y)\mathrm{I}_{(A_5)}\right\} + E\left\{(X)\mathrm{I}_{(A_5)}\right\} = E\left(-Y\mathrm{I}_{(A_5)}\right) < \infty$$

بنابراین $E\left(Y\mathrm{I}_{(A_5)}\right) < \infty$. در نتیجه با کم کردن $E\left(X\mathrm{I}_{(A_5)}\right) < \infty$ از طرفین داریم

$$E\left\{-(X+Y)\mathrm{I}_{(A_5)}\right\} = -E\left\{(X)\mathrm{I}_{(A_5)}\right\} - E(Y\,\mathrm{I}_{(A_5)})$$

$$E\left\{(X+Y)\mathrm{I}_{(A_5)}\right\} = E\left\{X\,\mathrm{I}_{(A_5)}\right\} + E\left(Y\,\mathrm{I}_{(A_5)}\right)$$

برای A_6 نیز به همین ترتیب برقرار می‌باشد. پس در حالت کلی داریم:

$$E\left\{(X+Y)\mathrm{I}_{(A_i)}\right\} = E\left\{X\,\mathrm{I}_{(A_i)}\right\} + E\left(Y\,\mathrm{I}_{(A_i)}\right) \qquad (۵-۵)$$

و

$$E\left\{X\left(\mathrm{I}_{(A_1)} + \mathrm{I}_{(A_3)} + \mathrm{I}_{(A_5)}\right)\right\} = E(X^+)$$

$$E\left\{Y\left(\mathrm{I}_{(A_2)} + \mathrm{I}_{(A_3)} + \mathrm{I}_{(A_4)}\right)\right\} = -E(Y^-)$$

در نتیجه:

$$\sum_{i=1}^{6} E(Y\,(\mathrm{I}_{(A_i)})) = E(Y)$$

و بطور مشابه داریم

$$\sum_{i=1}^{6} E((Y+X)\,(\mathrm{I}_{(A_i)})) = E(Y+X)$$

و با توجه به رابطه (۵-۵) داریم $E(X+Y)=E(X)+E(Y)$ و از رابطه $E(cX)=cE(X)$ نتیجه می‌شود:

$$E(aX+bY) = aE(X)+bE(Y)$$

ت) $X \geq Y \;\; a.s \;\; \Rightarrow E(X) \geq E(Y)$

$$\begin{aligned} X \geq Y \;\; a.s \;\; &\Rightarrow X - Y \geq 0 \\ &\Rightarrow E(X-Y) \geq 0 \\ &\Rightarrow E(X) - E(Y) \geq 0 \\ &\Rightarrow E(X) \geq E(Y) \end{aligned}$$

ج) $X = 0 \;\; a.s \;\Rightarrow E(X) = 0$

$$X = 0 \;\; a.s \;\Rightarrow X \geq 0 \;\; a.s \quad , \quad X \leq 0 \;\; a.s$$

$$\Rightarrow E(X) \geq 0 \quad , \quad E(-X) \geq 0 \;\Rightarrow E(X) = 0$$

ح) X انتگرال‌پذیر است اگر وتنها اگر $|X|$ انتگرال پذیر باشد.

اگر X انتگرال‌پذیر باشد آنگاه $E(X^-)$ و $E(X^+)$ متناهی می‌باشند. در نتیجه

$$E(|X|) = E(X^+ + X^-) = E(X^+) + E(X^-) < \infty$$

پس $|X|$ نیز انتگرال‌پذیر می‌باشد. برعکس اگر $|X|$ انتگرال‌پذیر باشد، چون $|X| \geq X$ پس داریم

$$X \leq |X| \Rightarrow E(X) \leq E(|X|) < \infty$$

درنتیجه X نیز انتگرال‌پذیر خواهد شد.

د) اگر Y انتگرال‌پذیر باشد و $|X| \leq Y$ آنگاه X انتگرال‌پذیر است.

$$|X| \leq Y \Rightarrow E(|X|) \leq E(Y) < \infty$$

بنابراین $|X|$ انتگرال‌پذیر است در نتیجه X انتگرال‌پذیر است.

۵-۱-۶ خواص دیگر از امیدها

۱- اگر $a \leq X \leq b$ آنگاه $a \leq E(X) \leq b$.

اثبات: با توجه به خاصیت بند (ت) داریم :

$$X \geq Y \ \ a.s \ \ \Rightarrow E(X) \geq E(Y)$$

بنابراین می‌توان به نتیجه فوق یعنی $a \leq E(X) \leq b$ رسید.

۲- $E(|X|) \geq |E(X)|$.

اثبات: (طبق نامساوی مثلثی)

$$\begin{aligned}|\mathrm{E(X)}| &= |\mathrm{E}(X^+) - \mathrm{E}(X^-)| \leq |\mathrm{E}(X^+)| + |\mathrm{E}(X^-)| \\ &= \mathrm{E}(X^+) + \mathrm{E}(X^-) \\ &= \mathrm{E}(|\mathrm{X}|)\end{aligned}$$

۳- اگر X فقط مقادیر صحیح مثبت را با احتمال‌های $P(X = i) = P_i$ اختیار کند، آنگاه:

$$E(X) = \sum_{i=1}^{\infty} P(X \geq i)$$

اثبات:

$$\mathrm{E(X)} = \sum_{i=1}^{\infty} i\, P_i$$

$$= P_1 + 2P_2 + 3P_3 + \cdots$$

$$= \big(P(X \geq 1) - P(X \geq 2)\big) + 2\big(P(X \geq 2) - P(X \geq 3)\big) + \cdots$$

$$= P(X \geq 1) + P(X \geq 2) + \cdots$$

$$= \sum_{i=1}^{\infty} P(X \geq i)$$

۴- اگر $f(x)$ تابع چگالی متغیر تصادفی X بر R باشد و $X \geq 0$ آنگاه داریم

$$E(X) = \int_0^{\infty} x\, f(x)dx = \int_0^{\infty} (1 - F(x))dx$$

اثبات: بنابر انتگرال‌گیری جزء به جزء داریم:

$$\begin{cases} dx = dv & \Rightarrow x = v \\ 1 - F(x) = u & \Rightarrow -f(x)d(x) = d(u) \end{cases}$$

$$\int_0^{\infty} (1 - F(x))dx = \big(1 - F(x)\big)(x)\Big|_0^{\infty} + \int_0^{\infty} x\, f(x)dx$$

$$= 0 + \int_0^{\infty} x\, f(x)dx$$

حال برای جزء مثبت و منفی داریم: $X^+ \geq 0$ و $X^- \geq 0$

$$E(X^+) = \int_0^{\infty} (1 - F(x))dx$$

$$E(X^-) = \int_{-\infty}^{0} (F(x))dx$$

$$\Longrightarrow E(X) = E(X^+) - E(X^-)$$

$$= \int_0^{\infty} (1 - F(x)dx - \int_{-\infty}^{0} (F(x))dx$$

$$= \int_0^\infty (1 - F(x)dx - \int_0^\infty \big(F(-x)\big)dx$$

$$= \int_0^\infty \big(1 - F(x) - F(-x)\big)dx$$

۵-۱-۷ امید ریاضی در حالت دو متغیره

حال در حالت دو متغیره، اگر X و Y دو متغیرتصادفی و $g(X,Y)$ تابعی بورل از (X,Y) باشد آنگاه:

$$E\big(g(X,Y)\big) = \int_\Omega g(X,Y)dp$$

اگر $(R^2, \mathcal{R}_2, P_{XY})$ فضای احتمال القاء شده بوسیله X و Y باشد عبارت بالا به صورت زیر در می‌آید:

$$E\big(g(X,Y)\big) = \int_{R_2} g(X,Y)dP_{XY}$$

اگر F تابع توزیع توأم (X,Y) باشد آنگاه داریم

$$E\big(g(X,Y)\big) = \int_{R_2} g(x,y)dF(X,Y) = \int_{R_2} g(x,y)\, f(x,y)dxdy$$

که در آن $f(x,y)$ تابع چگالی (X,Y) در R^2 می‌باشد.

۵-۱-۸ امید ریاضی متغیــرهای تصادفــی مختلط

اگر Z به صورت $Z = X + iY$ نوشته شود، به طوری که X و Y متغیرهای تصادفی و $i = \sqrt{-1}$ باشد، آنگاه Z یک متغیر تصادفی مختلط می‌باشد. اگر به صورت قطبی $Z(\omega) = P(\omega)e^{i\,\theta(\omega)}$ که در آن $0 \le P(\omega)$ و $0 \le \theta(\omega) \le 2\pi$ نوشته شود، θ و P متغیر تصادفی خواهند بود.

بنا به تعریف داریم

$$Z = e^{iuX} = \cos uX + i \sin uX$$

امید عبارت فوق را می‌گیریم

$$E(Z) = E(e^{iuX}) = E(\cos uX) + iE(\sin uX)$$

$$= \int (\cos ux)dF(x) + i \int (\sin ux)\, dF(x)$$

$$= \int e^{iuX}\, dF(x) = \varphi_X(u)$$

که در آن $\varphi_X(u)$ تابع مشخصه متغیر تصادفی X می‌باشد و نقش عمده‌ای در مطالعه خواص همگرایی توابع توزیع دارد. از آنجایی که تشخیص هم‌توزیعی از طریق تابع توزیع کاری دشوار است معمولاً از تابع مشخصه استفاده می‌کنند. بر طبق قضیه یکتایی، دو متغیر تصادفی هم‌توزیع‌اند، اگر و فقط اگر تابع مشخصه آنها برابر باشد.

لم ۵-۲: اگر Z یک متغیر تصادفی مختلط باشد آنگاه:

$$E(|Z|) \geq |E(Z)|$$

اثبات: فرض کنید $Z = \rho e^{i\alpha}$ که در آن ρ و α تصادفی هستند و $E(Z) = re^{it}$ زیرا:

$$E(Z) = E(X + iY) \;\; = E(X) + iE(Y) = M + iN = re^{it}$$

دراین صورت داریم

$$|E(Z)| = |r|\left|e^{it}\right| = |r||\cos t + i\sin t|$$

$$= |r|\sqrt{cos^2 t + sin^2 t} = |r|.1 = |r|$$

بنابراین $|E(Z)| = |r|$. در نتیجه:

$$|E(Z)| = |r| = e^{-it}E(\rho e^{i\alpha}) = E\left(\left|\rho e^{i(\alpha - t)}\right|\right)$$

$$= E(|\rho\cos(\alpha - t)) + iE(\rho\sin(\alpha - t)|)$$

$$= E(|\rho\cos(\alpha - t)| \leq E(|\rho|) = E(|Z|)$$

لم ۵-۳: اگر Z مختلط باشد آنگاه Z انتگرال‌پذیر است اگر و فقط اگر $|Z|$ انتگرال‌پذیر باشد.

اثبات: اگر $|Z|$ انتگرال‌پذیر باشد، با توجه به لم قبل $|E(Z)| \leq E(|Z|) < \infty$ در نتیجه Z نیز انتگرال‌پذیر خواهد بود، برعکس، اگر Z انتگرال‌پذیر باشد،

$$|Z| = \sqrt{X^2 + Y^2} \leq |X| + |Y|$$

چون Z انتگرال‌پذیر است بنابراین X و Y نیز انتگرال‌پذیرند و چون X و Y انتگرال‌پذیرند، با توجه به قسمت ح قضیه (۵-۳)، آنگاه $|X|$ و $|Y|$ انتگرال‌پذیرخواهند بود. بنابراین از

$$|Z| = \sqrt{X^2 + Y^2} \leq |X| + |Y|$$

خواهیم داشت:

$$E(|Z|) < \infty$$

یعنی $|Z|$ نیز انتگرال‌پذیر است.

۵–۲ گشتاورها

$E(X-a)^k$ را گشتاور مرتبه‌ی kام X حول a و $E(X^k)$ را گشتاور مرتبه‌ی kام X حول مبدأ یا به اختصارگشتاور kام می‌گوییم. که ممکن است، موجود باشد یا نباشد.

$E|X-a|^r$ را گشتاور مطلق r ام X حول a و $E|X^r|$ را گشتاور مطلق مرتبه‌ی rام گوییم. اگر متغیر تصادفی $a.s$ صفر باشد این گشتاورها صفر می‌شوند.

$E(X-E(X))^2$ گشتاور مرتبهٔ دوم X حول میانگین گویند که همان واریانس X می‌باشد. درواقع پراکندگی مقادیرمختلف X را حول میانگین اندازه‌گیری می‌کند.

۵–۲–۱ تابع مولد گشتاورها

اگر گشتاورهای تمام مراتب X وجود داشته باشد و متناهی باشد آنگاه $M(\theta)$ را تابع مولد گشتاور X می‌نامند،

در واقع

$$M(\theta) = 1 + E(X)\theta + E(X^2)\frac{\theta^2}{2!} + \cdots + E(X^n)\frac{\theta^n}{n!} + \cdots$$

$$\Longrightarrow M(\theta) = \sum_{n=0}^{\infty} E(X^n)\frac{\theta^n}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{\theta^n}{n!}\int_{-\infty}^{\infty} X^n \; dF(X)$$

$$= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{\theta^n X^n}{n!} dF(X)$$

$$= \int_{-\infty}^{\infty} e^{\theta X} dF(X) = E(e^{\theta X})$$

می‌توان نتیجه گرفت:

$$\mu'_n = E(X^n) = \frac{d^n M(\theta)}{d^n\theta}\Big|_{\theta = 0}$$

مثال ۵-۳: اگر X گسسته و دارای توزیع دو جمله‌ای منفی زیر باشد، آنگاه تابع مولد گشتاور آن را بیابید.

$$P(X=x) = \binom{r+x-1}{x} p^r(-q)^x \ , \quad x = 0,1,2,\ldots$$

حل:

$$M(\theta) = \sum_{n=0}^{\infty} \binom{r+x-1}{x} p^r(-q)^x e^{\theta X}$$

$$= \sum_{x=0}^{\infty} \binom{-r}{x} p^r(-qe^{\theta})^x = (\frac{p}{1-qe^{\theta}})^r$$

همچنین می‌توان نتیجه گرفت:

$$E(X) = r\frac{q}{p} \qquad , \qquad Var(X) = \frac{rq}{p^2}$$

۵–۳ نامساوی ها

لم ۵-۴: اگر $E(|X|^r) < \infty$ آنگاه به ازای $r' \le r$, $E\left(|X|^{r'}\right) < \infty$ و به ازای $k < r$, $E(X^k)$ موجود و متناهی است.

اثبات: به ازای $0 < r' \le r$ و مقدار مفروض ω داریم

$$|X(\omega)|^r \le 1 \Rightarrow |X(\omega)|^{r'} \le 1$$

$$|X(\omega)|^r > 1 \Rightarrow |X(\omega)|^{r'} \le |X(\omega)|^r$$

بنابراین از رابطه فوق می‌توان نتیجه گرفت:

$$|X(\omega)|^{r'} \le 1 + |X(\omega)|^r$$

ازطرفین رابطه فوق امید می‌گیریم

$$E|\mathrm{X}(\omega)|^{r'} \leq 1 + E|\mathrm{X}(\omega)|^{r}$$

با استفاده از این موضوع که X^k انتگرال‌پذیر است اگر و فقط اگر $|X^k|$ انتگرال‌پذیر باشد و همچنین با توجه به این‌که:

$$E|X(\omega)| < \infty \Rightarrow E|X(\omega)|^{r'} < \infty \quad , \quad |X|^k = |X^k|$$

می‌توان نتیجه گرفت:

$$E|X|^k < \infty \Rightarrow E|X^k| < \infty \Rightarrow E(X)^k < \infty$$

نتیجه ۵-۳: اگر گشتاور مرتبهٔ معینی متناهی باشد، تمام گشتاورهای مراتب پایین‌تر نیز متناهی خواهند بود. که عکس این لم برقرار نمی‌باشد.

قضیه یانگ

فرض کنید که h یک تابع پیوسته و اکیداً صعودی باشد و $h(0)=0$ و $h(\infty)=\infty$ ، حال فرض کنید $g=h^{-1}$. سپس برای هر $a>0$ و $b>0$ داریم

$$ab \leq \int_0^a h(t)\,dt + \int_0^b g(t)\,dt$$

قضیه هولدر

از تعریف نامساوی یانگ استفاده می‌کنیم و داریم

$$h(t) = t^{p-1}$$

$$g(s) = s^{1/(p-1)} = \mathrm{s}^{(1/\mathrm{p})/(\mathrm{p}-1/\mathrm{p})} = \mathrm{s}^{(\mathrm{q}-1/\mathrm{q})/(1/\mathrm{q})} = s^{q-1}$$

$$\left.\begin{aligned} \int_0^a h(t)dt = \frac{a^p}{p} \\ \int_0^b g(s)ds = \frac{b^q}{q} \end{aligned}\right\} => ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

که نامساوی اخیر همان نامساوی یانگ می‌باشد.

$$\begin{cases} p = r \\ q = s \end{cases}$$

که در عبارت فوق $b = \frac{|Y|}{E^{\frac{1}{s}}(|Y|^s)}$ و $a = \frac{|X|}{E^{\frac{1}{r}}(|X|^r)}$ قرار می دهیم.

در نتیجه

$$\frac{1}{r}.\frac{|X|}{E^{\frac{1}{r}}(|X|^r)} + \frac{1}{s}.\frac{|Y|}{E^{\frac{1}{s}}(|Y|^s)} \geq \frac{XY}{E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)}$$

با گرفتن امید ریاضی از طرفین داریم

$$\frac{1}{r} + \frac{1}{s} \geq \frac{E|XY|}{E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)}$$

اگر در قضیه هولدر قرار دهیم $\frac{1}{r} + \frac{1}{s} = 1$ آنگاه به نامساوی هولدر می رسیم.

نامساوی هولدر

$$\mathrm{E}(|\mathrm{XY}|) \leq E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)$$

که در آن $\frac{1}{r} + \frac{1}{s} = 1$ ، $s\,, r > 1$ می‌باشد.

اثبات: تابع $\phi(p) = \frac{1}{r}p^r + \frac{1}{s}p^{-s}\ ,\ o < p < \infty$ را در نظر بگیرید. این تابع به ازای $p = 1$ مینیمم می‌شود. بنابراین

$$\phi(p) = \frac{1}{r}p^r + \frac{1}{s}p^{-s} \geq 1$$

باشد. درنتیجه:

$$p = \frac{a^{\frac{1}{s}}}{b^{\frac{1}{r}}}$$

فرض می کنیم که

$$\frac{1}{r}\frac{a^{\frac{r}{s}}}{b^{\frac{r}{r}}} + \frac{1}{s}\frac{a^{\frac{-s}{s}}}{b^{\frac{-s}{r}}} \geq 1$$

$$\frac{1}{br}a^{\frac{r}{s}} + \frac{1}{as}b^{\frac{s}{r}} \geq 1$$

$$\frac{as\ a^{\frac{r}{s}} + br\ b^{\frac{s}{r}}}{ab\ rs} \geq 1$$

$$\frac{s\,a^{\frac{r+s}{s}}+r\,b^{\frac{s+r}{r}}}{rs}\geq ab$$

$$r^{-1}a^{\frac{r+s}{s}}+s^{-1}b^{\frac{s+r}{r}}\geq ab$$

$$r^{-1}a^{r\frac{r+s}{sr}}+s^{-1}b^{s\frac{s+r}{sr}}\geq ab$$

$$r^{-1}a^{r}+s^{-1}b^{s}\geq ab$$

درعبارت فوق $a=\frac{|X|}{E^{\frac{1}{r}}(|X|^r)}$ و $b=\frac{|Y|}{E^{\frac{1}{s}}(|Y|^s)}$ قرار می‌دهیم.

در نتیجه:

$$\frac{1}{r}.\frac{|X|^r}{E(|X|^r)}+\frac{1}{s}.\frac{|Y|^s}{E(|Y|^s)}\geq\frac{|XY|}{E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)}$$

با گرفتن امید ریاضی از طرفین داریم:

$$\frac{1}{r}+\frac{1}{s}\geq\frac{E|XY|}{E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)}$$

$$E|XY|\leq E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)$$

نتیجه ۵-۴: در نامساوی هولدر اگر $s=r=2$ باشد نامساوی هولدر به شوارتز تبدیل می‌شود.

$$E|XY|\leq\sqrt{E(|X|^2).E(|Y|^2)}$$

۵-۳-۱ نامساوی مینکووسکی

$$E^{\frac{1}{r}}(|X+Y|^r)\leq E^{\frac{1}{r}}(|X|^r)+E^{\frac{1}{r}}(|Y|^r)$$

اثبات:

$$E^{\frac{1}{r}}(|X+Y|^r)=E^{\frac{1}{r}}(|X+Y|^{r-1}.|X+Y|)$$

$$\leq E^{\frac{1}{r}}\big(|X+Y|^{r-1}.(|X|+|Y|)\big)$$

$$=E^{\frac{1}{r}}(|X||X+Y|^{r-1})+E^{\frac{1}{r}}(|Y||X+Y|^{r-1})$$

با استفاده از نامساوی هولدر داشتیم:

$$E|XY| \leq E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|Y|^s)$$

بنابراین:

$$\begin{cases} E^{\frac{1}{r}}(|X||X+Y|^{r-1}) \leq E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|X+Y|)^{s(r-1)} \\ E^{\frac{1}{r}}(|Y||X+Y|^{r-1}) \leq E^{\frac{1}{r}}(|Y|^r).E^{\frac{1}{s}}(|X+Y|)^{s(r-1)} \end{cases}$$

و همچنین

$$\left[\frac{1}{r}+\frac{1}{s}=1 \Rightarrow \frac{r+s}{rs}=1 \Rightarrow r+s=rs \Rightarrow r=s(r-1)\right]$$

بنابراین:

$$E(|X+Y|^r) \leq E^{\frac{1}{r}}(|X|^r).E^{\frac{1}{s}}(|X+Y|)^r \ + E^{\frac{1}{r}}(|Y|^r).E^{\frac{1}{s}}(|X+Y|)^r$$

$$\leq \left(E^{\frac{1}{r}}(|X|^r)+E^{\frac{1}{r}}(|Y|^r)\right)E^{\frac{1}{s}}(|X+Y|)^r$$

$$\Longrightarrow E^{1-\frac{1}{s}}(|X+Y|^r) \leq \left(E^{\frac{1}{r}}(|X|^r)+E^{\frac{1}{r}}(|Y|^r)\right)$$

$$E^{\frac{1}{r}}(|X+Y|^r) \leq E^{\frac{1}{r}}(|X|^r)+E^{\frac{1}{r}}(|Y|^r).$$

۵-۳-۲ نامساوی جنسن

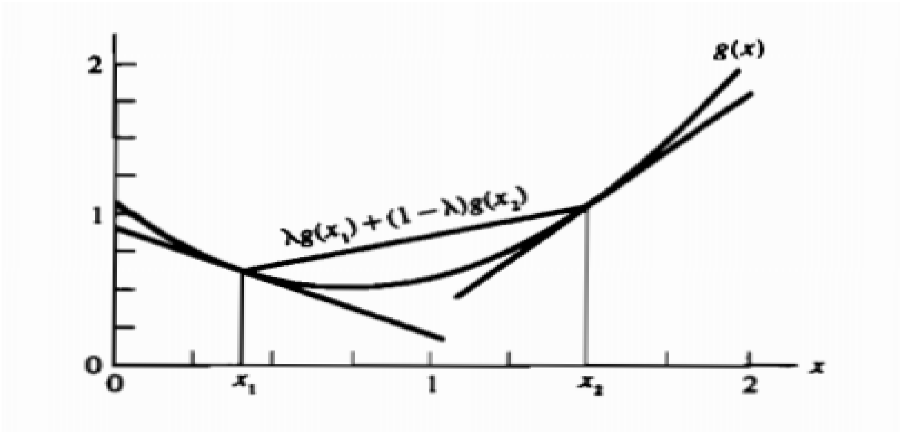
قبل از مطرح کردن نابرابری جنسن تعریف تابع محدب و برخی از ویژگی‌های آنرا بیان می‌کنیم

تابع محدب

داشته باشیم $0\leq a\leq 1$ که a و هر y و x را محدب گوییم هر گاه به ازای هر R →R :

$$\varphi(ax+(1-a)y) \leq a\varphi(x)+(1-a)\varphi(y)$$

تعبیر هندسی: برای هر x و y از R $(x \leq y)$, خطی که دو نقطه $(y,\varphi(y))$, $(x,\varphi(x))$ را به هم وصل می‌کند. بالای نمودار تابع φ روی بازه (x,y) قرار گیرد.



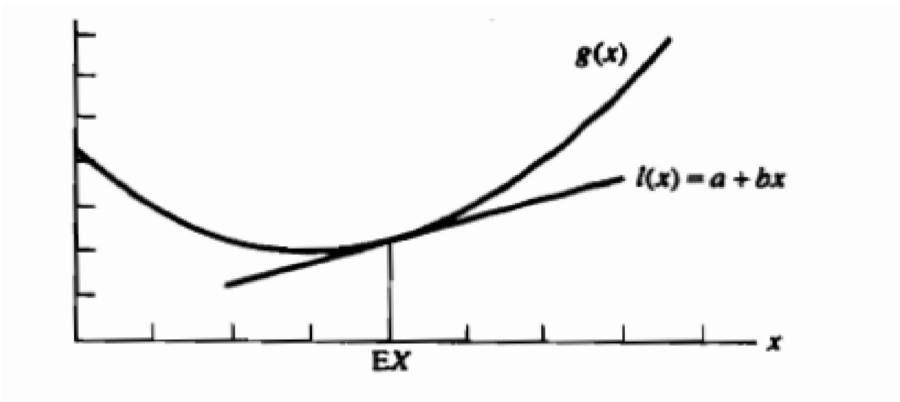
یعنی محدب بودن تابع φ هم ارز محدب بودن مجموعه زیر در R^2 است.

$$A=\{(x,y)/y \geq \varphi(x)\}$$

(مجموعه‌ای محدب است که هر دو نقطه ای از ان را اگر به وسیله پاره خطی به هم وصل کنیم ان پاره خط تماماً در داخل مجموعه قرار گیرد.)

ویژگی مجموعه محدب

هر خطی که در صفحه رسم شود نمودار آنها را در دو نقطه قطع می‌کند. به هر حال Ø اگر تابعی محدب باشد و A نقطه دلخواه روی آن باشد آنگاه خطی وجود دارد که از نقطه A می‌گذرد و تمام نقاط بالای خط قرار می‌گیرد مانند:



و از این ویژگی در اثبات قضیه جنسن استفاده می‌کنیم.

اگر $(x,\varphi(x))$ مختصات نقطه A باشد در این صورت خطی به معادله $y(x) = ax + b$ وجود دارد

الف) $y(x) = ax + b$ (خط از نقطه A می گذرد)

ب) به ازای $y(x) \geq ax + b$ (نمودار تابع بالای خط قرار میگیرد)

اثبات قضیه

مطابق توضیحات بالا، نقطه A به مختصات $(E(x), \varphi E(y))$ را روی نمودار $Y+\varphi(x)$ در نظر می گیریم.

فرض می‌کنیم که $y = ax + b$ خطی باشد که از نقطه A گذشته و تمام این نمودار بالای خط قرار می‌گیرد.

$$\varphi\big(E(x)\big) = a\,E(x) + b$$

به ازای هر $\varphi(x) \geq ax + b$

مجموعه‌های محدب و توابع محدب ویژگی‌های جالب وکاربردهای فراوانی دارند. از جمله ویژگی‌های آنها که در بیان ویژگی‌های تابع محدب نیز گفته شد این است که، هر خطی که درصفحه رسم شود آنها را حداکثر در دو نقطه قطع می‌کند. بعلاوه اگر f تابع محدبی باشد و A نقطه‌ای دلخواه روی آن، خطی وجود دارد که از نقطه A می‌گذرد و تمام منحنی بالای خط قرار می‌گیرد. که از این ویژگی در اثبات نابرابری جنسن استفاده می‌کنیم، که این نابرابری به صورت زیر است.

اگر $g(X)$ تابعی محدب باشد و $E(X) < \infty$ آنگاه:

$$g(E(X)) \leq E(g(X))$$

اثبات: فرض می‌کنیم که $L(X) = a + bX$, $X_0 = E(X)$ باشد:

$$L(X) = a + bX \Rightarrow \begin{cases} L(X) \leq g(X) \\ L(X_0) = g(X_0) \end{cases}$$

از رابطه $L(X) \leq g(X)$ امید می‌گیریم:

$$E(L(X)) \leq E\big(g(X)\big)$$

$$E(a + bX) \leq E\big(g(X)\big)$$

$$a + bE(X) \leq E\big(g(X)\big)$$

$$L(E(X)) \leq E\big(g(X)\big)$$

طبق فرض $X_0 = E(X)$ بنابراین داریم:

$$L(X_0) = g(X_0\,) \quad \Rightarrow \quad L\big(E(X)\big) = E\big(g(X)\big)$$

در نتیجه:

$$g\big(E(X)\big) \leq E\big(g(X)\big).$$

نامساوی $\boldsymbol{C_r}$

$$E(|X+Y|^r) \leq C_r E(|X|^r) + C_r E(|Y|^r)$$

که درآن C_r به صورت زیر تعریف می‌شود:

$$C_r = \begin{cases} 1 & r < 1 \\ 2^{r-1} & r \geq 1 \end{cases}$$

اثبات: تابع زیر را برای $a > 0 \ , \ b > 0$ و $r < 1$ تعریف می‌کنیم:

$$(\frac{a}{a+b})^r + (\frac{b}{a+b})^r \geq 1$$

$$a^r + b^r > (a+b)^r$$

بنابراین به ازای تمام مقادیر $r < 1 \ , \ \omega$ داریم

$$|X(\omega)|^r + |Y(\omega)|^r \geq (|X(\omega)| + |Y(\omega)|)^r$$

$$\geq (|X(\omega) + Y(\omega)|)^r$$

با گرفتن امید از طرفین رابطهٔ فوق به نامساوی C_r برای $r < 1$ می‌رسیم.

حال برای $r \geq 1$ داریم

تابع $\phi(p) = p^r + (1-p)^r$ را برای $0 \leq p \leq 1 \ , \ r \geq 1$ در نظر می‌گیریم.

$\phi(p)$ در $p = \frac{1}{2}$ دارای مینیمم است، پس به ازای $a > 0 \ , \ b > 0$ و $1-p = \frac{b}{a+b}$ ، $p = \frac{a}{a+b}$ ،

در نتیجه: $\phi(p) \geq 2^{-(r-1)}$

$$(\frac{a}{a+b})^r + (\frac{b}{a+b})^r \geq 2^{-(r-1)}$$

$$2^{(r-1)}(a^r + b^r) \geq (a+b)^r$$

بنابراین به ازای تمام مقادیر $r \geq 1 \ , \ \omega$ داریم:

$$2^{r-1}(|X(\omega)|^r + |Y(\omega)|^r) \geq (|X(\omega)| + |Y(\omega)|)^r$$

$$\geq (|X(\omega) + Y(\omega)|)^r$$

اگر از طرفین امید ریاضی بگیریم، نابرابری C_r حاصل می‌شود.

نتیجه ۵-۵: اگر گشتاورهای مطلق مرتبهٔ r ام X و Y موجود و متناهی باشند، گشتاورهای مطلق مرتبه‌ی r ام،

$X+Y$ نیز موجود است. هیچ محدودیتی روی توزیع X و Yوجود ندارد. به عبارتی دیگر کلاس متغیرهای تصادفی که گشتاور مرتبه‌ی sام آنها متناهی است نسبت به عمل جمع بسته است.

نامساوی اساسی

فرض کنید کهX یک متغیر تصادفی دلخواه و g بر R یک تابع نامنفی و بورل باشد. اگرg زوج و بر بازه‌ی$[0,\infty)$ نانزولی باشد آنگاه

$$\frac{E\big(\mathrm{g}(X)\big)-\mathrm{g}(a)}{a.s\quad \sup \mathrm{g}(X)} \le P(|X|\ge a) \le \frac{E\big(\mathrm{g}(X)\big)}{\mathrm{g}(a)} \quad , \quad \forall\, a>0$$

(در اینجا منظور از $a.s\quad \sup \mathrm{g}(X)$ عبارت است از کوچک‌ترین مقدار ثابتی که تقریباً تمام مقادیر $\mathrm{g}(X)$ زیر آن قرار دارند.)

اثبات: چون g بر R تابعی بورل است، نتیجه می‌شود که g بر Ω تابعی اندازه‌پذیر است که یک متغیر تصادفی می‌باشد. چون g نامنفی است، انتگرال آن وجود دارد. یعنی:

$$E[\mathrm{g}(X)] = \int_A \mathrm{g}(X) + \int_{A^c} \mathrm{g}(\mathrm{X})$$

که در آن$A=\{|X|\ge a\}$ و چون g نانزولی و زوج است داریم

$$\mathrm{g}(\mathrm{a}) \le \mathrm{g}(X) \le a.s\ \sup \mathrm{g}(X)$$

ازطرفین رابطه فوق، انتگرال نسبت به A می‌گیریم:

$$\int_A \mathrm{g}(a)dp \le \int_A \mathrm{g}(\mathrm{X}) \le \int_A \sup \mathrm{g}(X)\ a.s$$

$$\mathrm{g}(a).p(A) \le \int_A \mathrm{g}(\mathrm{X}) \le a.s\ \sup g(X)\,.p(A) \qquad (۶-۵)$$

بر روی ناحیهٔ A^c داریم

$$0 \le \mathrm{g}(X) \le \mathrm{g}(a)$$

با انتگرال‌گیری از رابطه فوق نسبت به A^c داریم

$$0 \le \int_{A^c} \mathrm{g}(\mathrm{X}) \le \mathrm{g}(a).p(A^c) \le \mathrm{g}(a) \qquad (۷-۵)$$

و با جمع‌کردن رابطه (۵-۶) و (۵-۷) به نتیجه زیر می‌رسیم:

$$g(a).p(A) \leq \int_A g(X) + \int_{A^c} g(X) \leq a.s \ \ \sup g(X)\,.\,p(A) + g(a)$$

$$g(a).p(A) \leq E\big(g(X)\big) \leq a.s \ \ \sup g(X)\,.\,p(A) + g(a)$$

از سمت راست به رابطه زیر می‌رسیم:

$$E\big(g(X)\big) - g(a) \leq a.s \ \ \sup g(X)\,.\,p(A)$$

$$\frac{E\big(g(X)\big) - g(a)}{a.s \ \ \sup g(X)} \leq p(A)$$

از سمت چپ به رابطه زیر می‌رسیم:

$$p(|X| \geq a) = p(A) \leq \frac{E\big(g(X)\big)}{g(a)}$$

در نتیجه:

$$\frac{E\big(g(X)\big) - g(a)}{a.s \ \ \sup g(X)} \leq P(|X| \geq a) \leq \frac{E\big(g(X)\big)}{g(a)} \quad , \ \ \forall\, a > 0$$

نتیجه ۵-۶: فرض کنید $g(X) = |X|^r$ و $r > 0$ آنگاه:

$$\frac{E(|X|)^r - |a|}{a.s \ \ \sup g(X)} \leq P(|X| \geq a) \leq \frac{E(|X|)^r}{|a|}$$

به سمت راست رابطه اخیر، نامساوی مارکف می‌گویند. به ازای $r = 2$ نامساوی مارکف به چپیشف تبدیل می‌شود. سمت چپ نامساوی بدیهی است اگر g تابعی کراندار باشد. زیرا اگر $g(X)$ کراندار نباشد $\sup g(X)$ نامتناهی است و سمت چپ صفر می‌شود.

۵–۴ مسائل

۱- اگر X متغیر تصادفی با میانگین m و واریانس متناهی σ^2 باشد آنگاه ثابت کنید:

$$P[X - m \geq \alpha] \leq \frac{\sigma^2}{\sigma^2 + \alpha^2}$$

۲- ثابت کنید:

$$E(X) = E(X|X > o)P(X > o) + E(X|X \leq 0)P(X \leq 0)$$

۳-اگر $E(|X|) < \infty$، $\sum_n P(A_n) < \infty$ روابط زیر را اثبات کنید:

الف) $\lim\limits_{n\to\infty} \int_{A_n} X\, dp = 0$

ب) $\lim\limits_{n\to\infty} \int_{|X|>n} X\, dp = 0$

۴- نابرابری دوتایی جنسن را برای تابع f که یک تابع حقیقی محدب روی مجموعه‌های محدب است ثابت کنید.

$$f[E(X), E(Y)] \leq E[f(X, Y)]$$

فصل ششم

همگرایی متغیرهای تصادفی

در این فصل ابتدا انواع مختلف همگرایی‌ها را تعریف و شرط کوشی را برای این همگرایی‌ها بیان و اثبات می‌کنیم. در ادامه قضایای همگرایی برای امید ریاضی را مطرح و اثبات می‌کنیم. این قضایا در واقع شرایطی را که می‌توان جای حد و انتگرال را عوض کرد بیان می‌کنند، و در انتها قضیه فوبینی را که یکی از قضایای مهم در نظریه اندازه و احتمال است تشریح می‌نماییم.

۶-۱ همگرایی نقطه به نقطه

گوییم دنباله ای از متغیرهای تصادفی X_n در Ω به متغیر تصادفی X همگراست اگر

$$\forall\ \omega \in \Omega\ \Rightarrow\ \overline{lim} X_n(\omega) = \underline{lim} X_n(\omega) = X(\omega) < \infty.$$

تعریف۶-۱: تعریف ریاضی همگرایی نقطه به نقطه

$$\forall\ \varepsilon > 0, \forall\omega \in \Omega\ \equiv n(\varepsilon,\omega);\ \ n >\ n(\varepsilon,\omega)\ \ \rightarrow |X_n(\omega) - X(\omega)| < \varepsilon$$

و می‌نویسیم

$$X_n \overset{p.w}{\rightarrow} X$$

نکته۶-۱: اگر $C \subset \Omega$ باشد و به ازای $\omega \in C$ دنباله $X_n(\omega)$ به $X(\omega)$ همگرا باشد آنگاه مجموعه C را مجموعه همگرایی دنباله $X_n(\omega)$ گوییم

$$C = \{\omega : X_n(\omega) \rightarrow X(\omega)\}$$

$$= \bigcap_{\varepsilon>0} \bigcup_{n} \bigcap_{m} \{\ \omega : |X_{n+m}(\omega) - X_n(\omega)| < \varepsilon\} \qquad (1-6)$$

اگر $C = \Omega$ آنگاه دنباله $X_n(\omega)$ را همه جا همگرا گوییم و اگر $C = \Phi$ باشد آن را هیچ جا همگرا گوییم.

تذکر۶-۱: اگر $C \in \mathcal{A}$ آنگاه $\lim X_n$ یک متغیر تصادفی توسعه یافته است.

تذکر۶-۲: همان طور که در تعریف مجموعه همگرایی (۶-۱) می‌بینیم C با اعمال شمارا روی مجموعه‌های اندازه-پذیر به دست می‌آید پس $C \in \mathcal{A}$ (یعنی اندازه‌پذیر است)

تعریف۶-۲: (دنباله کوشی) دنباله $\{P_n\}$ در فضای متری X را یک دنباله کوشی نامند هرگاه

$$\forall\, \varepsilon > 0 \quad \exists N\,;\ \ if\ \ n \geq N\,, m \geq N \ \ \to d(P_n, P_m) < \varepsilon$$

تعریف۶-۳: تعریف ریاضی متغیرهای تصادفی کوشی

$$\forall\, \omega \in \Omega;\ m, n \to \infty \to X_m(\omega) - X_n(\omega) \to .$$

تعریف۶-۴: محک کوشی برای همگرایی (معیار همگرایی کوشی)

$$\forall\, \omega \in \Omega\,; X_n(\omega) \xrightarrow{p.w} X(\omega) < \infty \ \Leftrightarrow m, n \to \infty \ \ ; X_m(\omega) - X_n(\omega) \to .$$

یا

$$X_{m+n} - X_n \to 0$$

اگر دنباله‌ای از متغیرهای تصادفی به مفهوم کوشی همگرا باشند، گفته می‌شود که به طور دو جانبه همگراست و C' مجموعه همگرایی دو جانبه به صورت زیر تعریف می‌شود

$$C' = \{\omega : X_m(\omega) - X_n(\omega) \to .\}$$

$$\bigcap_k \bigcup_n \bigcap_m \left\{ \omega : |X_{n+m}(\omega) - X_n(\omega)| < \frac{1}{k} \right\}$$

۶-۲ همگرایی در احتمال

دنباله‌ای از متغیرهای تصادفی مانند $\{X_n\}$ را در احتمال به X همگرا می‌گوییم و آن را با نماد $X_n \xrightarrow{P} X$ نشان می‌دهیم اگر

$$\forall\, \varepsilon > 0 \ \ ; \ \lim_{n\to\infty} P\{|X_n - X| < \varepsilon\} = 1$$

یا

$$\forall\, \varepsilon > 0 \ \ ; \ \lim_{n\to\infty} P\{|X_n - X| \geq \varepsilon\} = 0$$

مفهوم‌های زیر نقش مهمی در آمار ایفا می‌کنند که همگرایی در احتمال است.

الف) سازگاری برآوردگرها

$$\forall\ \varepsilon > 0\ ;\ \lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1$$

یا

$$\forall\ \varepsilon > 0\ ;\ \lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| \geq \varepsilon\} = 0$$

ب) قانون ضعیف اعداد بزرگ

$$\forall\ \varepsilon > 0\ ;\ \lim_{n\to\infty} P\{|\bar{X}_n - \mu| < \varepsilon\} = 1$$

یا

$$\forall\ \varepsilon > 0\ ;\ \lim_{n\to\infty} P\{|\bar{X}_n - \mu| \geq \varepsilon\} = 0$$

نکته۶-۲: در واقع سازگاری برآوردگرها و قانون ضعیف اعداد بزرگ حالت خاصی از همگرایی در احتمال هستند که دنباله متغیرها به یک مقدار ثابت میل می‌کنند.

تعریف۶-۵: دو متغیر تصادفی X و X' را معادل گوییم اگر

$$as \qquad X = X'$$

مجموعه زیر، مجموعه تمام ω هایی است که در آنها X و X' برابر نیستند

$$\{\omega: X(\omega) \neq X'(\omega)\} = \bigcap_k \left\{|X - X'| \geq \frac{1}{k}\right\}$$

پس Xو X' معادل هستند اگر $P\{X \neq X'\} = 0$ یا $P\left\{|X - X'| \geq \frac{1}{k}\right\} = 0$.

لم۶-۱: اگر $X_n \xrightarrow{P} X$ و $X_n \xrightarrow{P} X'$ آنگاه X و X' معادل هستند. به عبارت دیگر یک دنباله از متغیرهای تصادفی نمی‌توانند در احتمال به دو متغیر تصادفی متفاوت همگرا باشند.

اثبات: داریم

$$P(|X - X'| \geq \varepsilon) = P(|X - X_n + X_n - X'| \geq \varepsilon)$$
$$\leq P\left\{|X_n - X| \geq \frac{\varepsilon}{2}\right\} + P\left\{|X_n - X'| > \frac{\varepsilon}{2}\right\}$$

$$\lim_{n\to\infty} P(|X-X'|>\varepsilon) \le \lim_{n\to\infty} P\left(|X_n-X'|>\frac{\varepsilon}{2}\right)+\lim_{n\to\infty} P\left(|X_n-X'|>\frac{\varepsilon}{2}\right)$$

$$\lim_{n\to\infty} P(|X-X'|>\varepsilon)=0 \Rightarrow P(X=X')=1 \Rightarrow as\ X=X'$$

لم۶-۲: شرط کافی برای همگرایی در احتمال: اگر $E(|X_n|^r) \longrightarrow 0$ آنگاه $X_n \xrightarrow{P} 0$.

برهان: برای اثبات از نامساوی مارکـف استفاده می‌کنیم. داریم

$$P\{g(X)\ge\varepsilon\}\le\frac{E\{g(X)\}}{\varepsilon}$$

در نامساوی مارکف به جای $g(X)$ تابـع $|X_n|^r$ را قـرار می‌دهیم. پس

$$P(|X_n|^r\ge a^r)\le\frac{E(|X_n|^r)}{a^r}\Rightarrow P(|X_n|\ge a)\le\frac{E(|X_n|^r)}{a^r}$$

$$\lim_{n\to\infty}P(|X_n|\ge a)\le\lim_{n\to\infty}\frac{E(|X_n|^r)}{a^r}=0\Rightarrow\lim_{n\to\infty}P(|X_n|\ge a)=0\quad\Rightarrow X_n\xrightarrow{P}0$$

نتیجه۶-۱: اگر $Var(X_n)\longrightarrow 0$ آنگاه $X_n\xrightarrow{P}c$ (c عددی ثابت است).

برهان: از نامساوی چیبیشف داریم

$$P(|\bar{X}_n-\mu|>k)\le\frac{Var(\bar{X}_n)}{k^2}=0$$

در نتیجه

$$\lim_{n\to\infty}P(|\bar{X}_n-\mu|>k)\le\lim_{n\to\infty}\frac{Var(\bar{X}_n)}{k^2}=0$$

و $\lim_{n\to\infty}P(|\bar{X}_n-\mu|>k)=0$، بنابراین $\boldsymbol{\bar{X}_n\to\mu}$.

۶-۲-۱ معیار همگرایی در احتمال

لم۶-۳: شرط لازم و کافی برای همگرایی در احتمال به صورت زیر می‌باشد

$$n\to\infty\,,X_n\xrightarrow{P}0\ \Leftrightarrow E\left(\frac{|X_n|}{1+|X_n|}\right)\to 0$$

برهان: فرض می‌کنیم $E\left(\frac{|X_n|}{1+|X_n|}\right) \to 0$ ثابت می‌کنیم $X_n \xrightarrow{P} 0$. برای اثبات از نامساوی اساسی استفـاده می‌کنیم. در نامساوی اساسی $g(X)$ را $\frac{|X_n|}{1+|X_n|}$ می‌گیریم.

$$g(X_n) = \frac{|X_n|}{1+|X_n|}$$

همواره داریم

$$\frac{|X|}{1+|X|} \leq 1 \ \Rightarrow \text{a.}s \qquad \sup g(X) = 1$$

$$\frac{E\left(\frac{|X_n|}{1+|X_n|}\right) - g(\varepsilon)}{a.s \ \ \sup g(X) = 1} \leq \ P\{|X_n| \geq \varepsilon\} \leq \frac{E\left(\frac{|X_n|}{1+|X_n|}\right)}{g(\varepsilon)} \qquad (2-6)$$

چون فرض کردیم

$$E\left(\frac{|X_n|}{1+|X_n|}\right) \to 0$$

پس سمت راست نامساوی (۶-۲)، به سمت صفر میل می‌کند. یعنی:

$$\lim_{n\to\infty} P(|X_n| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{E\left(\frac{|X_n|}{1+|X_n|}\right)}{g(\varepsilon)} = 0 \Rightarrow \lim_{n\to\infty} P(|X_n| \geq \varepsilon) = 0 \quad \Rightarrow X_n \xrightarrow{P} 0$$

اثبات عکس: حال فرض می‌کنیم $X_n \xrightarrow{P} 0$. ثابت می‌کنیم $E\left(\frac{|X_n|}{1+|X_n|}\right) \to 0$.

برای اثبات از طرف چپ نامساوی (۶-۲)، استفاده می‌کنیم

$$E\left(\frac{|X_n|}{1+|X_n|}\right) - g(\varepsilon) \leq P\{|X_n| \geq \varepsilon\}$$

$$\lim_{n\to\infty} E\left(\frac{|X_n|}{1+|X_n|}\right) - g(\varepsilon) \leq \lim_{n\to\infty} P\{|X_n| \geq \varepsilon\} = 0$$

$$\lim_{n\to\infty} E\left(\frac{|X_n|}{1+|X_n|}\right) - g(\varepsilon) = 0$$

$$g(\varepsilon) = \frac{\varepsilon}{1+\varepsilon}$$

چون ε اختیاری است می‌توان $\varepsilon = 0$ پس $g(\varepsilon) = 0$ خواهد بود (همچنین $\frac{|X_n|}{1+|X_n|}$ نامنفی است و بنابراین $E\left(\frac{|X_n|}{1+|X_n|}\right) \geq 0$) داریم

$$\lim_{n\to\infty} E\left[\frac{|X_n|}{1+|X_n|}\right] = 0$$

لم۶-۴: برای هر $X_n \xrightarrow{P} X$ $m,n > 0$ آنگاه $X_n - X_m \xrightarrow{P} 0$.

در واقع این لم نشان می‌دهد که همگرایی در احتمال همگرایی دو جانبه را نتیجه می‌دهد.

داریم

$$P(|X_n - X_m| > \varepsilon) = P(|X_n - X + X - X_m| > \varepsilon)$$
$$\leq P\left(|X_n - X| > \frac{\varepsilon}{2}\right) + P\left(|X_m - X| > \frac{\varepsilon}{2}\right)$$
$$\lim_{n\to\infty} P(|X_n - X_m| > \varepsilon) \leq \lim_{n\to\infty} P\left(|X_n - X| > \frac{\varepsilon}{2}\right) + \lim_{n\to\infty} P\left(|X_m - X| > \frac{\varepsilon}{2}\right)$$
$$\lim_{n\to\infty} P(|X_n - X_m| > \varepsilon) = 0 \implies X_n - X_m \xrightarrow{P} 0$$

نکته۶-۳: مفهوم همگرایی در احتمال حالت خاصی از همگرایی در یک اندازه معین است. در واقع $\{f_n\}$ ها در اندازه μ به f همگراست اگر $n \to \infty$ آنگاه

$$\mu\{|f_n - f| > \varepsilon\} \to 0 \implies f_n \xrightarrow{\mu} f$$

اگر μ را اندازه احتمال P بگیریم همگرایی در احتمال را نتیجه می‌دهد.

مثال۶-۱: فرض کنید $\Omega = [0,1]$ و اندازه P ، اندازه لبگ بر $[0,1]$ باشد. دنباله X_n ها را به صورت زیر در نظر بگیرید.

$$X_n(\omega) = \begin{cases} 1 & \omega \in [0,\frac{1}{n}] \\ 0 & \omega \in [\frac{1}{n},1] \end{cases}$$

ثابت می‌کنیم دنباله X_n ها در احتمال به صفر همگرا هستند.

در واقع $X_n = I(A_n)$ خواهد بود که $A_n = \left\{\omega : \omega \in [0,\frac{1}{n}]\right\}$. چون $P(A_n) = \frac{1}{n}$ ، (P اندازه لبگ است) پس وقتی $n \to \infty$

$$P(A_n) = \frac{1}{n} \to 0$$

از طرفی

$$P\{|X_n| > \varepsilon\} = P\{X_n > \varepsilon\} + P\{X_n < -\varepsilon\}$$
$$= P\{X_n = 1 = I(A_n)\}$$
$$= P\{I(A_n) = 1\} = P(A_n) \to 0$$

در نتیجه

$$P\{|X_n| > \varepsilon\} \to 0 \Rightarrow X_n \xrightarrow{P} 0$$

مثال۶-۲: فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی باشد به طوری که

$$\begin{cases} P(X_n = 0) = 1 - \dfrac{1}{n^2} \\ P(X_n = n) = \dfrac{1}{n^2} \end{cases}$$

نشان دهید که $X_n \xrightarrow{P} 0$.

اثبات:

$$P(|X_n - 0| > \varepsilon) = P(|X_n| > \varepsilon) = P(X_n > \varepsilon) + P(X_n < -\varepsilon) = P(X_n > \varepsilon) + 0$$

فقط برای $n > \varepsilon$ ، $P(X_n = n) = \frac{1}{n^2}$ می شود ، پس

$$= \frac{1}{n^2}$$

در نتیجه $\lim_{n\to\infty} P(|X_n - 0| > \varepsilon) = 0$ ، و $X_n \xrightarrow{P} 0$.

قضیه۶-۱: اگر $f(x)$ تابعی پیوسته باشد و $X_n \xrightarrow{P} X$ آنگاه

$$f(X_n) \xrightarrow{P} f(X)$$

برهان: X متغیر تصادفی است برای $\varepsilon > 0$ ثابت k را طوری در نظر می‌گیریم که $P\{|X_n| > k\} < \frac{\varepsilon}{2}$.

f روی R پیوسته است پس روی $[-k, k]$ پیوسته یکنواخت است. (زیرا تابع پیوسته روی بازه بسته فشرده. پیوسته یکنواخت است.)

$$\forall\varepsilon > 0\,, \exists\delta(\varepsilon,k); |X_n - X| < \delta \;\Rightarrow |f(X_n) - f(X)| < \varepsilon$$

با فرض $\delta = \frac{\varepsilon}{2}$ می‌توان نوشت

$$\{|f(X_n) - f(X)| < \varepsilon\} \supset \{|X_n - X| < \delta\}$$

$$\Rightarrow P\{|f(X_n) - f(X)| < \varepsilon\} \geq P\{|X_n - X| < \delta\}$$

$$\Rightarrow \lim_{n\to\infty} P\{|f(X_n) - f(X)| < \varepsilon\} \geq \lim_{n\to\infty} P\{|X_n - X| < \delta\} = 1$$

$$\lim_{n\to\infty} P\{|f(X_n) - f(X)| < \delta\} = 1$$

$$f(X_n) \xrightarrow{P} f(X)$$

قضیه۶-۲: (قضیه اسلوتسکی) عمل همگرایی در احتمال تحت اعمال حسابی بسته است.

فرض کنید $X_n \xrightarrow{P} X$ و $Y_n \xrightarrow{P} Y$ آنگاه

الف) $aX_n \xrightarrow{P} aX$ ،

ب) $X_n + Y_n \xrightarrow{P} X + Y$ ،

ج) $X_nY_n \xrightarrow{P} XY$ ،

د) $\frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y}$.

اثبات:

الف) اگر $a = 0$ باشد نتیجه بدیهی است. فرض می‌کنیم $a \neq 0$ باشد داریم

$$P\{|aX_n - aX| \geq \varepsilon\} = P\{|a||X_n - X| \geq \varepsilon\} = P\left\{|X_n - X| \geq \frac{\varepsilon}{|a|}\right\}$$

$$\Rightarrow \lim_{n\to\infty} P\{|aX_n - aX| \geq \varepsilon\} = \lim_{n\to\infty} P\left\{|X_n - X| \geq \frac{\varepsilon}{|a|}\right\} = 0$$

$$\lim_{n\to\infty} P\{|aX_n - aX| \geq \varepsilon\} = 0 \Rightarrow aX_n \xrightarrow{P} aX$$

ب) داریم

$$P\{|(X_n+Y_n)-(X+Y)|\geq \varepsilon\} = P\{|(X_n-X)+(Y_n-Y)|\geq \varepsilon\}$$
$$\leq P\{|(X_n-X)|+|Y_n-Y|\geq \varepsilon\}$$
$$\leq P\left\{|(X_n-X)|\geq \frac{\varepsilon}{2}\right\}+P\left\{|(Y_n-Y)|\geq \frac{\varepsilon}{2}\right\}$$
$$\Longrightarrow \lim_{n\to\infty} P\{|(X_n+Y_n)-(X+Y)|\geq \varepsilon\}$$
$$\leq \lim_{n\to\infty} P\left\{|(X_n-X)|\geq \frac{\varepsilon}{2}\right\}+\lim_{n\to\infty} P\left\{|(Y_n-Y)|\geq \frac{\varepsilon}{2}\right\}$$
$$\lim_{n\to\infty} P\{|(X_n+Y_n)-(X+Y)|\geq \varepsilon\} = 0 \;\Longrightarrow\; X_n+Y_n \xrightarrow{P} X+Y$$

ج) همواره داریم

$$X_nY_n-XY=(X_n-X)(Y_n-Y)+X(Y_n-Y)+Y(X_n-X)$$
$$\Longrightarrow P\{|X_nY_n-XY|>\varepsilon\}=P\{|(X_n-X)(Y_n-Y)+X(Y_n-Y)+Y(X_n-X)|>\varepsilon\}$$

$$\leq P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3}\right\}+P\left\{|X(Y_n-Y)|>\frac{\varepsilon}{3}\right\}$$
$$+P\left\{|Y(X_n-X)|>\frac{\varepsilon}{3}\right\}$$

باید ثابت کنیم توابع احتمال رابطه بالایی به صفر میل می‌کند اولین تابع از سمت چپ را در نظر می‌گیریم داریم

$$P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3}\right\}$$
$$=P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3},|X_n-X|>\frac{\varepsilon}{3}\right\}$$
$$+P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3},|X_n-X|<\frac{\varepsilon}{3}\right\}$$
$$\leq P\left\{|X_n-X|>\frac{\varepsilon}{3}\right\}+P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3},|X_n-X|<\frac{\varepsilon}{3}\right\}$$

بنا به فرض $P\left\{|X_n-X|>\frac{\varepsilon}{3}\right\}$ به سمت صفر میل می‌کند $X_n \xrightarrow{P} X$ و

$$P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3},|X_n-X|<\frac{\varepsilon}{3}\right\}$$

چون دو پیشامد هیچ اشتراکی ندارند $P(\Phi)=0$ پس داریم

$$\lim_{n\to\infty} P\left\{|(X_n-X)(Y_n-Y)|>\frac{\varepsilon}{3}\right\}=0$$

به همین ترتیب می‌توان ثابت کرد که دو تابع احتمال نیز به سمت صفر میل می‌کنند و در نتیجه داریم

$$\lim_{n\to\infty} P\{|\, X_nY_n - XY| > \varepsilon\} \le 0 \implies \lim_{n\to\infty} P\{|\, X_nY_n - XY| > \varepsilon\} = 0 \implies X_nY_n \xrightarrow{P} XY$$

د) برای اثبات از قضیه ۶-۱ استفاده می‌کنیم. فرض کنید $f(x) = \frac{1}{x}$ داریم

$$\text{طبق فرض}\; Y_n \xrightarrow{P} Y \implies f(y_n) \xrightarrow{P} f(y) \implies \frac{1}{Y_n} \xrightarrow{P} \frac{1}{Y}$$

$$\text{طبق قسمت (ج)} \qquad X_n\frac{1}{Y_n} \xrightarrow{P} X\frac{1}{Y} \implies \frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y}$$

۶-۳ همگرایی در میانگین مرتبه r

تعریف۶-۶: دنباله‌ای از متغیرهای تصادفی مانند $\{X_n\}$ را در میانگین مرتبه r به X همگرا گوییم

اگر وقتی $n \to \infty$، داشته باشیم

$$\mathrm{E}\{|X_n - X|^r\} \to 0 \implies X_n \xrightarrow{r} X$$

به ازای $r = 2$ آن را همگرایی در مربع میانگین و به ازای $r = 1$ آن را همگرایی در میانگین مرتبه اول گوییم.

لم۶-۵: اگر $X_n \xrightarrow{r} X$ آنگاه $\mathrm{E}\{|X_n|^r\} \to \mathrm{E}\{|X|^r\}$.

اثبات:

حالت اول $r \le 1$:

$$C_{\mathrm{r}}\ \text{نامساوی}: \mathrm{E}(|X+Y|^r) \le C_{\mathrm{r}}\,\mathrm{E}(|X|^r) + C_{\mathrm{r}}\,\mathrm{E}(|Y|^r)$$

در نامساوی C_{r} به جای X و Y به ترتیب $(X_n - X)$ و X قرار می‌دهیم

$$C_{\mathrm{r}} = \begin{cases} 1 & r \le 1 \\ 2^{r-1} & r \ge 1 \end{cases} \qquad (۶-۷)$$

$$\implies \mathrm{E}(|X_n|^r) \le \mathrm{E}(|X_n - X|^r) + \mathrm{E}(|X|^r)$$

$$\mathrm{E}(|X_n|^r) - \mathrm{E}(|X|^r) \le \mathrm{E}(|X_n - X|^r) \qquad (۶-۸)$$

چون طبق فرض $X_n \xrightarrow{r} X$ یعنی $\mathrm{E}\,(|X_n - X|^r) \to 0$ پس

$$\mathrm{E}\,(|X_n|^r) \to \mathrm{E}\,(|X|^r)$$

حالت دوم $r > 1$:

$$\mathrm{E}^{\frac{1}{r}}(|X+Y|^r) \le \mathrm{E}^{\frac{1}{r}}(|X|^r) + \mathrm{E}^{\frac{1}{r}}(|Y|^r)$$: نامساوی مینکووسکی(۶-۹)

در نامساوی مینکووسکی به جای X و Y به ترتیب $(X_n - X)$ و X قرار می‌دهیم.

$$\Longrightarrow \mathrm{E}^{\frac{1}{r}}(|X_n|^r) \le \mathrm{E}^{\frac{1}{r}}(|X_n - X|^r) + \mathrm{E}^{\frac{1}{r}}(|X|^r)$$

$$\mathrm{E}^{\frac{1}{r}}(|X_n|^r) - \mathrm{E}^{\frac{1}{r}}(|X|^r) \le \mathrm{E}^{\frac{1}{r}}(|X_n - X|^r) \to 0$$

$$\Longrightarrow \mathrm{E}\,(|X_n|^r) \to \mathrm{E}\,(|X|^r)$$

لم ۶-۶: اگر $X_n \xrightarrow{r} X$ همگرا باشد آنگاه $X_n \xrightarrow{P} X$.

و وقتی X_n ها $a.s$ کراندار باشند آنگاه به ازای تمام r ها وقتی $X_n \xrightarrow{P} X$ آنگاه $X_n \xrightarrow{r} X$.

اثبات: فرض می کنیم $X_n \xrightarrow{r} X$ همگرا باشند ثابت می‌کنیم

$$X_n \xrightarrow{P} X$$

برای اثبات از نامساوی مارکف استفاده می‌کنیم.

$$P\{g(X) \ge \varepsilon\} \le \frac{E\{g(X)\}}{\varepsilon}$$: نامساوی مارکف

$$g(X) = |X_n - X|^r\ ;\ P\{|X_n - X|^r \ge \varepsilon^r\} \le \frac{\mathrm{E}(|X_n - X|^r)}{\varepsilon^r}$$

$$\Longrightarrow P\{|X_n - X| \ge \varepsilon\} \le \frac{\mathrm{E}(|X_n - X|^r)}{\varepsilon^r} \xrightarrow{\text{طبق فرض}} 0$$

$$\Longrightarrow P\{|X_n - X| \ge \varepsilon\} \to 0 \Longrightarrow X_n \xrightarrow{P} X$$

برعکس فرض می‌کنیم X_n ها کراندار باشد و $X_n \xrightarrow{P} X$ همگرا باشند ثابت می‌کنیم $X_n \xrightarrow{r} X$.

اگر X_n ها $a.s$ کراندار باشند و $X_n \xrightarrow{P} X$ آنگاه X نیز $a.s$ کراندار است و در نتیجه $X_n - X$، $a.s$ کراندار خواهد بود. با استفاده از نامساوی اساسی داریم

$$\frac{E\{g(X)\}-g(a)}{a.s \quad sup\, g(X)} \leq P\{|X| \geq a\} \leq \frac{E\{g(X)\}}{g(a)}$$: نامساوی اساسی

با قرار دادن $g(X) = |x|^r$ و $X = (X_n - X)$ داریم

$$\frac{\mathrm{E}(|X_n - X|^r) - a^{\mathrm{r}}}{\sup g(X)} \leq P\{|X_n - X| \geq a\} \xrightarrow{\text{طبق فرض}} 0$$

$$\mathrm{E}(|X_n - X|^r) - a^{\mathrm{r}} \leq 0$$

چون $|X_n - X|^r \geq 0$ پس $\mathrm{E}(|X_n - X|^r) \geq 0$. پس به ازای $a = 0$ نیز داریم: $E(|X_n - X|^r) \leq 0$

آنگاه خواهیم داشت:

$$\mathrm{E}(|X_n - X|^r) \to 0$$

نتیجه۶-۲: همگرایی در میانگین مرتبه r قویتر از همگرایی در احتمال است و هردو معادل هستند اگر متغیرهای تصادفی مربوطه $a.s.$ کراندار باشند.

قضیه ۶-۳: دنباله متغیرهای تصادفی $\{X_n\}$ در احتمال همگرا به X است اگر و تنها اگر $E\left(\frac{|X_n-X|^r}{1+|X_n-X|^r}\right)$ همگرا به صفر باشد.یعنی

$$X_n \xrightarrow{P} X \Leftrightarrow E\left(\frac{|X_n - X|^r}{1 + |X_n - X|^r}\right) \to 0$$

اثبات: فرض می کنیم $X_n \xrightarrow{P} X$ ثابت می‌کنیم

$$E\left(\frac{|X_n - X|^r}{1 + |X_n - X|^r}\right) \to 0$$

در نامساوی اساسی به ازای $r > 0$، $g(x) = \frac{|X|^r}{1+|X|^r}$

و به جای X ، $X_n - X$ قرار می‌دهیم، داریم

$$\frac{E\left(\frac{|X_n-X|^r}{1+|X_n-X|^r}\right) - \frac{a^r}{1+a^r}}{a.s \;\; \sup g(X)} \leq P\{|X_n - X| \geq a\} \leq \frac{E\left(\frac{|X_n-X|^r}{1+|X_n-X|^r}\right)}{\frac{a^r}{1+a^r}}$$

ابتدا سمت چپ نامساوی فوق را در نظر می‌گیریم، از آن جایی که $X_n \xrightarrow{P} X$ پس

$$P\{|X_n - X| \geq \varepsilon\} \to 0$$

و چون $\frac{|X_n-X|^r}{1+|X_n-X|^r}$ مثبت است، پس امید آن نیز مثبت است. به ازاء $a = 0$ داریم: $0 \leq E(\frac{|X_n-X|^r}{1+|X_n-X|^r})$

آنگاه خواهیم داشت:

$$E\left(\frac{|X_n - X|^r}{1 + |X_n - X|^r}\right) \to 0$$

برای اثبات عکس، سمت راست نامساوی فوق را در نظر می‌گیریم.

$$E\left(\frac{|X_n - X|^r}{1 + |X_n - X|^r}\right) \to 0 \implies P\{|X_n - X| \geq a\} \leq 0 \implies P\{|X_n - X| \geq a\} \to 0$$

$$\implies X_n \xrightarrow{P} X$$

۶–۴ قضایای همگرایی برای امید ریاضی

۶–۴–۱ قضیه همگرایی یکنوا (MCT)

اگر $X \geq 0$ ↑ X_n، آنگاه $0 \leq E(X)$ ↑ $E(X_n)$.

برهان: ممکن است $E(X)$ متناهی یا نامتناهی باشد. چون X_n ها به X صعود می‌کنند

پس $X_n \leq X_{n+1}$ و $E(X_n) \leq E(X_{n+1})$. لذا دنباله $\{E(X_n)\}$ ها به طور یکنوا صعودی است و می‌دانیم که هر دنباله یکنوا دارای حد است، پس دنباله $\{E(X_n)\}$ ها نیز حد دارد. ثابت می‌کنیم حد این دنباله برابر $E(X)$ است.

برای هر یک از متغیرهای تصادفی $X_k \geq 0$ می‌توان دنباله‌ای از متغیرهای تصادفی نامنفی ساده مانند $\{X_{k_m}\}$ را وابسته کرد که به طور یکنوا به X_k متقارب باشد، یعنی $X_{k_m} \uparrow X_k$

$$\begin{gathered} 0, X_{11}, X_{12}, \dots, X_{1n} \quad \uparrow X_1 \\ \vdots \\ 0, X_{n1}, X_{n2}, \dots, X_{nn} \quad \uparrow X_n \end{gathered}$$

فرض کنید $Y_n = \max_{i \leq n, j \leq n} (X_{ij})$ (چون X_{ij} ها توابع ساده هستند، Y_n ماکزیمم X_{ij} ها هستند پس Y_n ها نیز توابع ساده هستند.)

$$X_{kn} \leq Y_n = \max_{i\leq n, j\leq n}\left(X_{ij}\right) \leq X_n$$

$$\Longrightarrow E(X_{kn}) \leq E(Y_n) \leq \mathrm{E}(X_n)$$

$$\text{(۶-۱۱)} \Longrightarrow \lim_n E(X_{kn}) \leq \lim_n E(Y_n) \leq \lim_n E(X_n)$$

چون X_{kn} ها ساده هستند جای امید و حد را می توانیم عوض کنیم پس داریم:

$$\text{(۶-۱۲)} \lim_n E(X_{kn}) = E(\lim_n X_{kn}) = E(X_k)$$

رابطه (۶-۱۲) را در رابطه (۶-۱۱) قرار می‌دهیم.

$$E(X_k) \leq \lim_n E(Y_n) \leq \lim_n E(X_n)$$

$$\Longrightarrow \lim_k E(X_k) \leq \lim_n E(Y_n) \leq \lim_n E(X_n)$$

از طرفی چون $\lim_n E(X_n) = \lim_k E(X_k)$، پس

$$\text{(۶-۱۳)} \lim_n E(X_n) \leq \lim_n E(Y_n) \leq \lim_n E(X_n)$$

$$X_{kn} \leq Y_n \leq X_n$$

$$\lim_n X_{kn} = X_k \leq \lim_n Y_n \leq \lim_n X_n = X$$

$$\lim_k X_k = X \leq \lim_n Y_n \leq X$$

$$\text{(۶-۱۴)} \lim_n Y_n = X$$

چون Y_n ها نیز توابع ساده هستند،جای امیدوحدرا می توانیم عوض کنیم و با توجه به $\lim_n E(Y_n) = E\left(\lim_n Y_n\right)$، و عبارت (۶-۱۴)، $E\left(\lim_n Y_n\right) = E(X)$ (※) . پس با برابری رابطه (※)، رابطه (۶-۱۳) را به صورت زیر خواهیم داشت:

$$\lim_n E(X_n) \leq E(X) \leq \lim_n E(X_n)$$

$$\Longrightarrow \lim_n E(X_n) = E(X)$$

نتیجه ۶-۳: اگر $0 \leq X \downarrow X_n$، آنگاه $E(X) \downarrow E(X_n)$.

برهان: فرض کنید $Z_n = (-X_n)$ نامنفی و به طور یکنوا صعودی است پس بنا به قضیه MCT داریم

$$Z_n \uparrow Z \overset{MCT}{\Longrightarrow} E(Z_n) \uparrow E(Z)$$

$$\Longrightarrow E(-X_n) \uparrow E(-X)$$

$$\Longrightarrow -E(X_n) \uparrow -E(X)$$

$$\Longrightarrow E(X_n) \downarrow E(X)$$

نتیجه ۶-۴: اگر $X_n \geq 0$ آنگاه

$$\mathrm{E}\left(\sum_{i=1}^{\infty} \mathrm{X_i}\right) = \sum_{i=1}^{\infty} \mathrm{E(X_i)}$$

اثبات:

$$\mathrm{E}\left(\sum_{i=1}^{\infty} \mathrm{X_i}\right)$$

$$= E\left(\lim_{n} \sum_{i=1}^{\mathrm{n}} \mathrm{X_i}\right)$$

$$= \lim_{n} \mathrm{E}\left(\sum_{i=1}^{\mathrm{n}} \mathrm{X_i}\right) = \lim_{n} \sum_{i=1}^{\mathrm{n}} \mathrm{E(X_i)} = \sum_{i=1}^{\infty} \mathrm{E(X_i)}$$

نکته۶-۴: می‌توان صورت قضیه همگرایی یکنوا را از این منظر نیز نگاه کرد.

اگر $0 \leq X_n \uparrow X$ ، آنگاه $0 \leq E(X_n) \uparrow E(X)$.

$$\lim_{n} X_n = X$$

$$E\left(\lim_{n} X_n\right) = E(X)$$

چون $X_n \geq 0$ ، پس می توان جای حد و انتگرال را عوض کرد.

$$\lim_{n} E(X_n) = E(X)$$

به عبارت دیگر قضیه MCT یک شرط برای امکان تعویض حد و امید را مطرح می‌کند و می‌گوید که اگر دنباله X_n ها نامنفی و یکنوا باشد می‌توان جای حد و امید ریاضی را عوض کرد.

قضیه۶-۴: قضیه فاتو (لم فاتو)

یک نتیجه مهم از قضیه همگرایی یکنوا، لم فاتو است که به صورت زیر بیان و اثبات می‌کنیم.

الف) اگر $Y \le X_n$ و Y انتگرال پذیر باشد آنگاه

$$E\left(\underline{\lim} X_n\right) \le \underline{\lim} E(X_n)$$

ب) اگر $X_n \le Z$ و Z انتگرال پذیر باشد آنگاه

$$\overline{\lim} E(X_n) \le E\left(\overline{\lim} X_n\right)$$

ج) اگر $Y \le X_n \le Z$ و Y, Z انتگرال‌پذیر باشند و $X_n \xrightarrow{a.s} \mathrm{X}$ آنگاه $\lim\limits_{n} E(X_n) = E(X)$.

باید توجه کرد که وقتی $Y \le X_n$ باشد $\underline{\lim} X_n$ موجود و $a.s$ متناهی است ولی $\overline{\lim} X_n$ ممکن است متناهی نباشد و همچنین وقتی $X_n \le Z$، $\overline{\lim} X_n$ موجود و متناهی خواهد بود ولی $\underline{\lim} X_n$ ممکن است متناهی نباشد.

اثبات:

الف) فرض می کنیم $Y_n = \inf\limits_{k \ge n} X_k$. Y_n ها دنباله‌های صعودی هستند که همگرا به $\lim \inf\limits_{k \ge n} X_k$ می‌باشند، در نتیجه

$$\lim Y_n = \underline{\lim}\, X_n \qquad (۶\text{-}۱۵)$$

فرض می‌کنیم $Y_n \ge 0$ بنا به قضیه MCT داریم

$$E(\,Y_n) \uparrow E\left(\underline{\lim}\, X_n\right) \qquad (۶\text{-}۱۶)$$

از طرفی $X_n \ge Y_n$ و $E(X_n) \ge E(Y_n)$، در نتیجه

$$\underline{\lim}\, E(X_n) \ge \underline{\lim}\, E(Y_n) = \lim E(Y_n) = E\left(\underline{\lim}\, X_n\right)$$

$$\Longrightarrow \underline{\lim}\, E(X_n) \ge E(\underline{\lim}\, X_n)$$

حال حالت کلی را در نظر می‌گیریم یعنی حالتی که Y_n لزوماً نامنفی نباشد. بنا به فرض $X_n \ge Y$، پس

$Z_n = X_n - Y \ge 0$. چون Z_n ها نامنفی هستند بنا به قسمت فوق که برای نا منفی‌ها اثبات کردیم داریم

$$\underline{\lim}\, E(Z_n) \ge E\left(\underline{\lim}\, Z_n\right) \qquad (۶\text{-}۱۷)$$

به طوری که

$$\underline{\lim}\, Z_n = \underline{\lim} X_n - Y \qquad (۶\text{-}۱۸)$$

در نتیجه

$$E\left(\underline{\lim}\, Z_n\right) = E\left(\underline{\lim} X_n - Y\right) = E\left(\underline{\lim} X_n\right) - E(Y) \quad (۱۹-۶)$$

و

$$\underline{\lim}\, E(Z_n) = \underline{\lim}\, E(X_n) - E(Y) \quad (۲۰-۶)$$

با جایگذاری عبارات (۶-۱۹) و (۶-۲۰) در (۶-۱۷) نتیجه می‌شود

$$\underline{\lim}\, E(X_n) \geq \ E(\underline{\lim}\, X_n).$$

ب) فرض می کنیم $X_n \leq Z$، درنتیجه $\mathrm{Z_n} = Z - X_n \geq 0$. چون $Z_n \geq 0$ پس بنا به قسمت ۱ داریم

$$\underline{\lim}\, E(Z_n) \geq \ E\left(\underline{\lim}\, Z_n\right)$$

از طرفی

$$\mathrm{E}(Z_\mathrm{n}) = \mathrm{E(Z)} - \mathrm{E}(X_n) \quad (۲۱-۶)$$

$$\underline{\lim}\, Z_n = Z - \ \overline{\lim} X_n \quad (۲۲-۶)$$

$$\underline{\lim}\, E(Z_n) = \underline{\lim}\left(\mathrm{E(Z)} - \mathrm{E}(X_n)\right) = E(Z) - \overline{\lim}\, E(X_n) \quad (۲۳-۶)$$

$$E\left(\underline{\lim}\, Z_n\right) = E\left(Z - \overline{\lim}\ (X_n)\right) = E(Z) - E\left(\overline{\lim}\, X_n\right) \quad (۲۴-۶)$$

در نتیجه

$$E(Z) - \overline{\lim}\, E(X_n) \geq E(Z) - E(\overline{\lim}\, X_n)$$
$$\Longrightarrow \overline{\lim}\, E(X_n) \leq E(\overline{\lim}\, X_n)$$

پس در حالت کلی داریم

$$E(\underline{\lim}\, X_n) \leq \underline{\lim}\, E(X_n) \leq \overline{\lim}\, E(X_n) \leq E(\overline{\lim}\, X_n) \quad (۲۵-۶)$$

ج) با فرض این که $X_n \xrightarrow{a.s} \mathrm{X}$، داریم $\lim\limits_n X_n = X$ و $\underline{\lim}\, X_n = \overline{\lim}\, X_n = \lim\limits_n X_n = X$. با استفاده از رابطه (۶-۲۵)، خواهیم داشت

$$E(X) \leq \underline{\lim}\, E(X_n) \leq \overline{\lim}\, E(X_n) \leq E(X)$$
$$\Longrightarrow \lim E(X_n) = E(X)$$

نتیجه ۶-۵: اگر $|X_n| \leq Y$ و Y انتگرال‌پذیر باشد داریم

$$\text{اگر } X_n \xrightarrow{a.s} \mathrm{X} \Longrightarrow E(X_n) \to E(X)$$

اثبات: با توجه به قسمت (ج) لم فاتو انجام می‌شود. چون $-Y \leq X_n \leq Y$ و $-Y, Y$ انتگرال‌پذیر هستند، شرایط قسمت (ج) لم فاتو برقرار است.

قضیه ۶-۵: (**قضیه همگرایی مغلوب**)؛ در نتیجه (۶-۵)، اگر شرط $X_n \xrightarrow{a.s} \mathrm{X}$ را با شرط $X_n \xrightarrow{P} \mathrm{X}$ عوض کنیم در این صورت این نتیجه، قضیه ای را که به قضیه همگرایی مغلوب معروف است، نتیجه می دهد.

پس اگر $|X_n| \leq Y$ و Y انتگرال پذیر باشد و اگر $X_n \xrightarrow{P} \mathrm{X}$ آنگاه $E(X_n) \to E(X)$.

اثبات: ابتدا قضیه را برای $X = 0$ یعنی وقتی که $X_n \xrightarrow{P} 0$ میل می‌کند ثابت می‌کنیم.

دنباله $E(X_n)$ و $\underline{\lim}\, E(X_n)$ را در نظر بگیرید (چون $-Y \leq X_n \leq Y$ است، پس $\underline{\lim}\, E(X_n)$ موجود است).

زیر دنباله‌ای مانند $X_{n'} \xrightarrow{P} 0$ وجود دارد به قسمی که

$$E(X_{n'}) \to \underline{\lim}\, E(X_n)$$

بنا به قضیه ای از قبل، یک زیر دنباله مانند $\{X_{n''}\}$ از $\{X_n\}$ وجود دارد به قسمی که $X_{n''} \xrightarrow{a.s} 0$. بنا به نتیجه ای از قبل و فرض $|X_n| < Y$ نتیجه می شود $E(X_{n''}) \to 0$.

پس دو زیر دنباله همگرا از یک زیر دنباله همگرا وجود دارد پس داریم

$$\underline{\lim}\, E(X_n) = \lim E(X_{n''}) = 0 \Longrightarrow \underline{\lim}\, E(X_n) = 0 \qquad (۶-۲۶)$$

به همین ترتیب می‌توان زیر دنباله‌ای از X_n ها انتخاب کرد که امید ریاضی آن به $\overline{\lim}\, E(X_n)$ همگرا باشد و مانند قسمت بالا

$$\overline{\lim}\, E(X_n) = 0 \qquad (۶-۲۷)$$

با استفاده از (۶-۲۶) و (۶-۲۷) داریم

$$\underline{\lim}\, E(X_n) = \overline{\lim}\, E(X_n) = 0 \Longrightarrow \lim E(X_n) = 0$$

حال فرض می‌کنیم $X_n \xrightarrow{P} \mathrm{X}$ داریم

$$X_n \xrightarrow{P} \mathrm{X} \Longrightarrow X_n - \mathrm{X} \xrightarrow{P} 0$$

چون $|X_n| < Y$، پس

$$-Y \leq X_n \leq Y \Rightarrow -X - Y \leq X_n - \mathrm{X} \leq Y - X$$

$$\Rightarrow -2Y \leq -X - Y \leq X_n - \mathrm{X} \leq \mathrm{Y} - \mathrm{X} \leq \mathrm{Y} + \mathrm{Y} = 2\mathrm{Y} \Rightarrow |X_n - \mathrm{X}| \leq 2\mathrm{Y}$$

و چون $X_n - \mathrm{X} \xrightarrow{P} 0$ و $|X_n - \mathrm{X}| \leq 2\mathrm{Y}$ شرایط حالت قبل برقرار است و داریم

$$\lim E(X_n - \mathrm{X}) = 0 \Rightarrow \lim E(X_n) = \mathrm{E}(\mathrm{X})$$

۶–۵ قضیه فوبینی

در این بخش به بسط قضیه‌ای می‌پردازیم که طی آن قادر خواهیم بود ترتیب را در انتگرال دوگانه عوض کنیم. این قضیه ابزار نیرومندی در اثبات قضایا و حل مسائل است.

فرض کنید (X, χ) و (Y, Υ) دو فضای اندازه‌پذیر باشند و همچنین μ و υ دو اندازه تعریف شده روی این فضاها باشد. فضای حاصل ضربی (حاصل ضرب دکارتی) به صورت

$$\mathrm{X} \times Y = \{(\omega_1, \omega_2);\ \ \omega_1 \in X\ \ , \omega_2 \in Y\}$$

می باشد. در فضای حاصل ضربی $\mathrm{X} \times Y$ مستطیل‌های اندازه‌پذیر را به صورت زیر تعریف می‌کنیم:

مجموعه تمام زوج‌های $\{(\omega_1, \omega_2);\ \ \omega_1 \in A_1\ \ , \omega_2 \in A_2\}$ مستطیل $A_1 \times A_2$ را تشکیل می‌دهند.

اگر $A_1 \in \chi$ و $A_2 \in \Upsilon$ آنگاه $A_1 \times A_2$ یک مستطیل اندازه‌پذیر خواهد بود.

۶–۵–۱ سیگما میدان حاصل ضربی

سیگما میدان مینیمال شامل گردایه $\{A_1 \times A_2;\ \ A_1 \in \chi\ \ , A_2 \in \Upsilon\}$ را سیگمای میدان حاصل‌ضربی گویند.

$$\chi \times \Upsilon = \sigma\{A_1 \times A_2;\ \ A_1 \in \chi\ \ , A_2 \in \Upsilon\}$$

تعریف ۶–۷: فرض کنیم E عضوی از $\chi \times \Upsilon$ باشد، دراین صورت به ازای $x\ \in \chi$ ، زیر مجموعه E_x از Υ را به صورت زیر تعریف می‌کنیم

$$E_{\mathrm{x}} = \{\mathrm{y} \in \Upsilon\,; (x\,, y)\ \in\ E\} \subset \Upsilon$$

E_x را مقطع مجموعه E در نقطه x می‌نامیم. به همین ترتیب می‌توانیم مقطع E در نقطه $y\ \in\ \Upsilon$ را نیز تعریف کنیم.

$$E_{\mathrm{y}} = \{\mathrm{x}\ \in \chi\,; (x\,, y)\ \in\ E\} \subset \chi$$

در واقع E_x بخشی از E است که توسط x تعیین می‌شود و E_y بخشی از E است که توسط y تعیین می‌شود.

لم ۶-۷: اگر $E \in \chi \times \Upsilon$ باشد آنگاه برای هر x مجموعه E_x نسبت به Υ اندازه‌پذیر است و برای هر y مجموعه E_y نسبت به χ اندازه پذیر است.

اثبات: برای x ثابتی فرض کنید T_x یک نگاشت به صورت زیر باشد

$$T_{\mathrm{x}}: \mathrm{y} \to \mathrm{X} \times Y$$

با ضابطه

$$T_{\mathrm{x}}(\mathrm{y}) = (\mathrm{x}, \mathrm{y})$$

اگر $E = A \times B$ یک مستطیل اندازه‌پذیر باشد آنگاه داریم

$$T_{\mathrm{x}}{}^{-1}(\mathrm{E}) = \{\mathrm{y} : (x, y) \in E = A \times B\} = \begin{cases} B & x \in A \\ \phi & x \notin A \end{cases} \Longrightarrow T_{\mathrm{x}}{}^{-1}(\mathrm{E}) \in \Upsilon$$

بنابراین با استفاده از قضیه خوانده شده داریم

$$T_{\mathrm{x}}{}^{-1}(\mathrm{E}) \in \Upsilon \Longrightarrow \text{خواهد شد } \frac{\Upsilon}{\chi \times \Upsilon} \text{ اندازه پذیر } T_{\mathrm{x}}$$

تعریف ۶-۸: فرض کنید (Υ, Υ) و (X, χ) دو فضای اندازه پذیر باشند و $(X \times Y, \chi \times \Upsilon)$ فضای اندازه‌پذیر حاصل از ضرب این دو فضا باشد. فرض کنید $f: X \times Y \to R$ تابعی دلخواه باشد، در این صورت به ازاء هر $x \in \mathrm{X}$ تابع $f(x,.): Y \to R$ با ضابطه $f(x,.) = f(x,y), \forall\, y \in Y$ و برای هر $y \in \mathrm{Y}$ تابع $f(.,y): X \to R$ با ضابطه $f(.,y) = f(x,y), \forall\, x \in X$ را در نظر می‌گیریم.

توابع $f(x,.)$ و $f(.,y)$ را مقاطع تابع f به ترتیب در نقاط x و y می‌نامیم. در واقع $f(x,.)$ بخشی از f است که توسط x تعیین می‌شود و $f(.,y)$ بخشی از f است که توسط y تعیین می‌شود.

لم۶-۸: اگر f اندازه‌پذیر $(\chi \times \Upsilon)$ باشد آنگاه برای هر x ثابتی تابع $f(x,.)$ اندازه‌پذیر Υ است و برای هر y ثابتی تابع $f(.,y)$ اندازه‌پذیر χ است.

اثبات: اگر f اندازه پذیر $\frac{\chi \times \Upsilon}{\mathrm{R}}$ باشد آنگاه با استفاده از قضیه خوانده شده f_{T_x} اندازه‌پذیر $\frac{\Upsilon}{\mathrm{R}}$ خواهد بود. بنابراین $f(x,.) = f_{T_x}(.)$ اندازه‌پذیر y خواهد بود.

۲-۵-۶ اندازه حاصل ضربی (Product Measure)

حال فرض کنید $(X, \chi\, , \mu)$ و (Y, Υ, υ) دو فضای اندازه باشند و فضای اندازه‌پذیر $(X \times Y\, , \chi \times \Upsilon)$ از حاصل‌ضرب آنها حاصل شده باشد. فرض کنیم $E\ \in \chi \times \Upsilon$ و E_x، E_y به ترتیب مقاطع E در نقطه $x\ \in X$ و $y\ \in Y$ باشد در لم (۶-۷)، ثابت کردیم که E_x، E_y به ترتیب نسبت به Υ, χ اندازه‌پذیر هستند بنابراین می‌توان روی E_x،E_y اندازه‌های زیر را تعریف کرد.

$$\upsilon[y : (x, y) \in\ E] =\ I_A(x)\upsilon(B)$$

$$\mu[x : (x, y) \in\ E] =\ I_B(y)\mu(A)$$

و همچنین اندازه‌های متناهی $\Pi'(E)\, , \Pi"(E)$ را به صورت زیر در نظر می‌گیریم

$$\Pi'(E) = \int_X \upsilon[y : (x, y) \in\ E]\mu(dx), \qquad E \in\ \chi \times \Upsilon$$

$$\Pi"(E) = \int_Y \mu[x : (x, y) \in\ E]\upsilon\,(dy), \qquad E \in\ \chi \times \Upsilon$$

ثابت می‌کنیم $\Pi"(A \times B) = \Pi'(A \times B) = \mu(A).\upsilon(B)$.

$$\Pi'(A \times B) = \int_X \upsilon[y : (x, y) \in\ A \times B]\mu(dx)$$

$$= \int I_A(x)\upsilon(B)\mu\,(dx) =\ \upsilon(B) \underbrace{\int I_A(x)\mu\,(dx)}_{\mu(A)} = \upsilon(B).\mu(A)$$

$$\Pi"(A \times B) = \int_Y \mu[x : (x, y) \in\ A \times B]\upsilon\,(dy)$$

$$= \int I_B(y)\mu(A)\upsilon(dy) = \mu(A) \underbrace{\int I_B(y)\upsilon(dy)}_{\upsilon(B)} = \mu(A).\upsilon(B)$$

$$\Longrightarrow \Pi'(A \times B) = \Pi"(A \times B) = \mu(A).\upsilon(B)$$

وقتی δ یک تابع اندازه‌پذیر نامنفی باشد اندازه υ را به صورت زیر تعریف می‌کنیم

$$\upsilon(A) = \int_A \delta\ d\mu$$

$$\int_A f\ d\mu = \int I_A f\ d\mu$$

قضیه۶-۶: (**قضیه فوبینی**)؛ اگر $(X, \chi\,, \mu)$ و (Y, Υ, υ) دو فضای اندازه‌پذیر سیگما متناهی باشند و $\prod(E) = \Pi'(E) = \Pi''(E)$ به عنوان اندازه سیگما متناهی روی سیگما میدان $\chi \times \Upsilon$ تعریف شود. داریم

الف)

$$\iint_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_X \left[\int_Y f(x,y)\upsilon(dy)\right] \mu(dx)$$

ب)

$$\iint_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_Y \left[\int_X f(x,y)\mu(dx)\right] \upsilon(dy)$$

اثبات: فرض می‌کنیم $f = I_A$، در این صورت

$$\int_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_{X\times Y} \mathrm{I_A}\Pi\big(d(x\,,y)\big) = \Pi(A)$$

$$\int_X \left[\int_Y f(x,y)\upsilon(dy)\right] \mu(dx) = \underbrace{\int_X \left[\int_Y \mathrm{I_A}\,\upsilon(dy)\right] \mu(dx)}_{\upsilon(A)} = \overbrace{\int_X \upsilon(A)\,\mu(dx)}^{\text{بنا به تعریف}\,\Pi'(A)} = \Pi'(A)$$

$$\int_Y \left[\int_X f(x,y)\mu(dx)\right] \upsilon(dy) = \underbrace{\int_Y \left[\int_X \mathrm{I_A}\,\mu(dx)\right] \upsilon(dy)}_{\mu(A)} = \int_Y \mu(A)\,\upsilon(dy) = \Pi''(A)$$

از آن جایی که ثابت کردیم $\Pi(A) = \Pi'(A) = \Pi''(A)$ پس داریم

$$\int_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_X \int_Y f(x,y)\,\upsilon(dy)\;\mu(dx)$$

$$= \int_Y \int_X f(x,y)\,\mu(dx)\,\upsilon(dy)$$

حال فرض می‌کنیم $f = \sum a_i I_{A_i}$ توابع ساده غیر منفی باشند، به ازاء $(x,y) \epsilon A_i$،

قرار می دهیم $x = A_{i1}$ و $y = A_{i2}$. داریم

$$\int_X \int_Y f(x,y)\upsilon(dy)\,\mu(dx) = \int_X \int_Y \sum a_i I_{A_i}\,\upsilon(dy)\,\mu(dx) = \sum \int_X \int_Y a_i I_{A_i}\upsilon(dy)\,\mu(dx)$$

$$= \sum \int_X a_i \left[\underbrace{\int_Y I_{A_i}\upsilon(dy)}_{\upsilon[y:(x,y)\epsilon A_i]}\right]\mu(dx) = \sum \int_x a_i \upsilon(A_{i2})\mu(dx)$$

$$= \sum a_i \upsilon(A_{i2}) \underbrace{\int_x \mu(dx)}_{\mu(A_{i1})}$$

$$= \sum a_i \upsilon(A_{i2})\ \mu(\, A_{i1}) \qquad (۶-۲۸)$$

از طرفی دیگر

$$\int_Y \int_X f(x,y)\ \mu(dx)\ \upsilon(dy) = \int_Y \int_X \sum a_i I_{A_i}\ \ \mu(dx)\ \upsilon(dy)$$

$$= \ \sum \int_Y \int_X a_i I_{A_i}\ \mu(dx)\ \upsilon(dy) = \sum \int_Y a_i \underbrace{\left[\int_X I_{A_i}\ \mu(dx)\right]}_{\mu[x:(x,y)\epsilon A_i]} \upsilon(dy)$$

$$= \sum \int_Y a_i\ \mu(A_{i1})\,\upsilon(dy) = \ \sum a_i\ \mu(A_{i1}) \underbrace{\int_Y \upsilon(dy)}_{\upsilon(A_{i2})}$$

$$= \sum a_i\ \mu(A_{i1})\ \upsilon(A_{i2}) \qquad (۶-۲۹)$$

از طرفی دیگر

$$\int_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_{X\times Y} \sum a_i I_{A_i}\,\Pi\big(d(x\,,y)\big)$$

$$= \sum a_i \int_{X\times Y} I_{A_i}\Pi\big(d(x\,,y)\big) \qquad (۶-۳۰)$$

$$= \sum a_i\ \mu(A_{i1})\,\upsilon(A_{i2})$$

پس از (۶-۲۸)، (۶-۲۹) و (۶-۳۰) داریم

$$\int_{X\times Y} f(x,y)\Pi\big(d(x\,,y)\big) = \int_X \int_Y f(x,y)\ \upsilon(dy)\ \ \mu(dx) \quad = \int_Y \int_X f(x,y)\ \mu(dx)\ \upsilon(dy)$$

حال فرض کنید $f \geq 0$ تابعی غیر منفی باشد آنگاه بنا به قضیه خوانده شده توابع ساده غیر منفی وجود دارد که $f_n \uparrow f$ پس بنا به قضیه MCT، $[\int f_n\, d\, \mu] \uparrow [\int f d\, \mu]$. در نتیجه

$$\left[\int_Y f_n \upsilon(dy)\right] \uparrow \left[\int_Y f\ \upsilon(dy)\right]$$

و

$$\left[\int_X f_n\ \mu(dx)\right] \uparrow \left[\int_X f\ \mu(dx)\right]$$

و همین طور

$$\lim_{n\to\infty} \int_Y f_n \upsilon(dy) = \int_Y f\ \upsilon(dy)$$

و

$$\lim_{n\to\infty} \int_X f_n\ \mu(dx) = \int_X f\ \mu(dx)$$

در نتیجه

$$\int_X \int_Y f(x,y)\ \upsilon(dy)\ \ \mu(dx) = \int_X \lim_{n\to\infty} \int_Y f_n \upsilon\ (dy)\, \mu(dx) = \lim_{n\to\infty} \int_X \int_Y f_n(x,y)\ \upsilon(dy)\ \ \mu(dx)$$

و

$$\int_Y \int_X f(x,y)\ \ \mu(dx) \upsilon(dy) = \int_Y \lim_{n\to\infty} \int_X f_n \mu\ (dx) \upsilon(dy) = \lim_{n\to\infty} \int_Y \int_X f_n(x,y)\ \ \mu(dx) \upsilon(dy)$$

از طرفی چون f_n توابع ساده غیر منفی هستند و در قسمت قبل برای توابع ساده ثابت کردیم

$$\int_Y \int_X f\ \ \mu(dx) \upsilon(dy) = \int_X \int_Y f\ \upsilon(dy)\ \ \mu(dx)$$

پس برای توابع $f \geq 0$ هم رابطه برقرار است.

$$\int_{X\times Y} f(x,y) \Pi\big(d(x\ ,y)\big) = \lim_{n\to\infty} \int_{X\times Y} f_n(x,y) \Pi\big(d(x\ ,y)\big)$$

حال فرض می‌کنیم f یک تابع کلی باشد. $f = f^+ - f^-$ و $|f| = |f^+| + |f^-|$. در نتیجه

$$\int_X \int_Y f^+ \upsilon(dy)\ \mu(dx) = \int_Y \int_X f^+\ \mu(dx)\upsilon(dy)$$
$$= \int_{X\times Y} f^+\Pi\big(d(x\,,y)\big)$$

و

$$\int_X \int_Y f^-\upsilon(dy)\,\mu(dx) = \int_Y \int_X f^-\ \mu(dx)\upsilon(dy)\ = \int_{X\times Y} f^-\Pi\big(d(x\,,y)\big)$$

$$\int_X \int_Y f^+\,\upsilon(dy)\mu(dx) - \int_X \int_Y f^-\,\upsilon(dy)\mu(dx)$$
$$= \int_Y \int_X f^+\ \mu(dx)\upsilon(dy) - \int_Y \int_X f^-\ \mu(dx)\upsilon(dy)$$
$$= \int_{X\times Y} f^+\Pi\big(d(x\,,y)\big) - \int_{X\times Y} f^-\Pi\big(d(x\,,y)\big)$$
$$= \int_{X\times Y} (f^+ - f^-)\Pi(d(X,Y))$$

$$\int_X \int_Y f\ \upsilon(dy)\mu(dx)\ = \int_Y \int_X f\ \ \mu(dx)\upsilon(dy) = \int_{X\times Y} f\ \Pi\big(d(x\,,y)\big)$$

۶-۶ مسائل

۱) فرض کنید فضای اندازه $(R\,,\mathcal{R}\,,\lambda)$ که λ اندازه لبگ است. متغیرهای X_n را به صورت زیر در نظر بگیرید:

$$X \equiv 0\ \ ,X_n(\omega) = \begin{cases} \mathrm{n}^2 & 0 < \omega < \dfrac{1}{\mathrm{n}} \\ 0 & \mathrm{o.w} \end{cases}$$

آنگاه نشان دهید:

$$X_n(\omega) \longrightarrow X(\omega),\ \ \forall\omega\epsilon\Omega$$

۲) فرض کنید:

$$I = \int_{-\infty}^{\infty} e^{\text{-x^2}}\ dx$$

آنگاه جواب انتگرال I را پیدا کنید.

۳) جواب انتگرال زیر را پیدا کنید.

$$\int_0^\infty \frac{\sin x}{x} dx$$

۴) اگر $X_1, \dots, X_n$ متغیرهای تصادفی با توزیع $U(0,\theta)$ باشند توزیع حدی $Y_n = X_{(n)}$ را بیابید

فرض کنید $\{X_n\}$دنباله‌ای از متغیرهای تصادفی باشد به طوری که:

$$p\ (X_n = 0) = 1 - \frac{1}{n}\ , \qquad p\ (X_n = 1) = \frac{1}{n} n = 1,2, \dots$$

۵) نشان دهید $X_n \xrightarrow{L^2} 0$. اگر $X_1, \dots, X_n$ یک نمونه تصادفی از توزیع یکنواخت $U(0,\theta)$ باشد آنگاه توزیع حدی $Z_n = nX_{(1)}$ را پیدا کنید.

۶) فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی با شرایط زیر باشد:

$$Var(X_n) = \sigma_n^2\ , E(X_n) = m_n$$

ثابت کنید وقتی $\sigma_n^2 \to 0$، $n \to \infty$ آنگاه $X_m - X_n$ در احتمال به صفر همگراست.

۷) اگر $X_n \xrightarrow{a.s} X$ و $Y_n \xrightarrow{a.s} Y$ نشان دهید: $X_n + Y_n \xrightarrow{a.s} X + Y$.

۸) اگر $\{X_n, Y_n\}$ دنباله‌ای از زوج‌های متغیرهای تصادفی باشند و $Y_n \xrightarrow{l} Y$ و $|X_n - Y_n| \xrightarrow{l} 0$ آنگاه:

$X_n \xrightarrow{l} Y$.

۹) فرض کنید g تابعی پیوسته باشد و $X_n - Y_n \xrightarrow{p} 0$ اگر $Y_n \xrightarrow{l} Y$ نشان دهید که

$g(X_n) - g(Y_n) \xrightarrow{p} 0$.

۱۰) نشان دهید اگر $X_n \xrightarrow{p} X$ آنگاه از $supE(|X_n|^r) = c < \infty$ نتیجه می‌شود که

$$X_n \xrightarrow{\mathrm{r}'} \mathrm{X}, \forall r' \le r$$

۱۱) نشان دهید که اگر $X_n \epsilon l_r$ و $X_n \xrightarrow{r} X$ آنگاه $X_m - X_n \xrightarrow{r} 0$ و برعکس.

۱۲) اگر $X_n \xrightarrow{r} X$ آنگاه $X_n \xrightarrow{r'} X\ , r' < r$

۱۳) فرض کنید $X_n \epsilon l_r$ ثابت کنید $X_n \xrightarrow{r} X$ اگر و فقط اگر $X_n \xrightarrow{p} X$ و

$$E(|X_n|^r) \to E(|X|^r) < \infty$$

۱۴) اگر $X_n - {X_n}' \xrightarrow{p} 0$ و $F_n(x) \to F(x)$ نشان دهید که برای همان مقادیر x

$.F_n'(x) \to F(x)$

$$X_n - X_n' \xrightarrow{p} 0 \;\Rightarrow\; X_n \xrightarrow{p} X_n' \;\Rightarrow\; \left.\begin{matrix} F_n(x) \to F_n'(x) \\ F_n(x) \to F(x) \end{matrix}\right\} \;\Rightarrow\; F_n'(x) \to F(x).$$

۱۵) دنباله متغیرهای تصادفی $\{X_n\}$ را به طور تقریبا یکنواخت $(a.u)$ به X همگرا گوییم هر گاه برای هر

$\varepsilon > 0$ یک مجموعه اندازه پذیر وجود داشته باشد به طوری که $P(A) < \varepsilon$ و به طور یکنواخت

$$X_n(\omega) = X(\omega), \qquad \forall\ \omega \epsilon A^c$$

۱۶) نشان دهید

$$X_n \xrightarrow{a.u} X \;\Rightarrow\; X_n \xrightarrow{a.s} X \;\Rightarrow\; X_n \xrightarrow{p} X$$

۱۷) آیا $X_n \xrightarrow{a.s} X$ نتیجه می‌دهد که $X_n \xrightarrow{a.u} X$؟

۱۸) فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشند به قسمی که

$$P[X_n = 0] = 1 - n^{-2} \qquad P[X_n = e^n] = n^{-2}$$

وضعیت دنباله را از جهت همگرایی $a.s$ و همگرایی در میانگین مرتبه r بررسی کنید.

۱۹) فرض کنید $\{X_n\}$دنباله‌ای از متغیرهای تصادفی مستقل باشند به قسمی که

$$P[X_n = 0] = 1 - n^{-1} \qquad P\left[X_n = n^{\frac{1}{2r}}\right] = n^{-1}$$

وضعیت دنباله را از جهت همگرایی $a.s$ و همگرایی در میانگین مرتبه r بررسی کنید.

فصل هفتم

تابع مشخصه

تابع مشخصه یکی از مباحث مهم در آمار می‌باشد. همان‌طور که از قبل می‌دانید تابع مولد گشتاور همواره وجود ندارد . در صورتی‌که تابع مشخصه همواره وجود دارد و همچنین مشخصا با توجه به خواصی که برای تابع توزیع بیان شده است از این رو نتیجه میگیریم که تناظر یک به‌یک بین تابع مشخصه و تابع توزیع برقرار است.

۷–۱ تابع مشخصه

فرض کنید F تابع توزیع متغیر تصادفی X با دو خاصیت: $F(-\infty) = 0$ و $F(+\infty) = 1$ باشد

تابع مشخصه یک F تابع مختلط مقادیر با رابطه

$$\varphi_F(u) = \int_R e^{iux} dF(x)$$

$$u \in R \qquad , \qquad i = \sqrt{-1}$$

تعریف می‌شود

ازآنجایی‌که $e^{ix} = cosx + isinx$ خواهیم داشت

$$\Phi_F(u) = \int_R cosuxdF(x) + i\int_R sinuxdF(X)$$

مثال ۷-۱:

-اگر X دارای چگالی نرمال استاندارد باشد داریم :

$$X \sim N(0,1)$$

$$\Phi_F(u) = \int_R \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} e^{iux} dx = \int_R \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(x-iu)^2 - \frac{u^2}{2}} dx$$

$$= \int_R \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(x-iu)^2} e^{-\frac{u^2}{2}} dx = e^{-\frac{u^2}{2}} \int_R \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(x-iu)^2} dx = e^{-\frac{u^2}{2}}$$

-اگر X داری توزیع نمایی باشد داریم :

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\Phi_F(u) = \int_R e^{iux} dF(x) = \int_0^\infty e^{-x} e^{iux} dx = \int_0^\infty e^{(iu-1)x} dx$$

$$= \int_0^\infty e^{-(1-iu)x} dx = \frac{-1}{1-iu} e^{-(1-iu)x} \Big|_0^\infty$$

$$= \frac{-1}{1-iu} [e^{-\infty} - e^0] = \frac{1}{1-iu} = (1-iu)^{-1}$$

-اگر X داری توزیع دوجمله‌ای باشد داریم:

$$X \sim B(n,p)$$

$$q = 1-p$$

$$\Rightarrow p_x(x) = \binom{n}{x} p^x q^{n-x} \qquad x = 0,1,\dots,n$$

$$\varphi_x(u) = \sum_{x=0}^n e^{iux} p_x(x) = \sum_{x=0}^n e^{iux} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^n \binom{n}{x} (pe^{iu})^x q^{n-x} = (pe^{iu} + q)^n$$

-اگر X داری توزیع پواسن باشد داریم :

$$X \sim P(\lambda)$$

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!} \qquad x=0,1,2,\dots$$

$$\Rightarrow \Phi_x(u)=\sum_{x=0}^{\infty} e^{iux}\frac{e^{-\lambda}\lambda^{x}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^{iu})^{x}}{x!}$$

$$=e^{-\lambda}e^{\lambda e^{iu}}=e^{\lambda(e^{iu}-1)}$$

-اگر X در C تباهیده باشد خواهیم داشت :

$$P(X=x)=\begin{cases}1 & x=c\\ 0 & o.w\end{cases}$$

$$\Rightarrow \Phi_x(u)=\sum_{x} e^{iux}p(x)=e^{iuc}$$

نکته ۷-۱: اگر Φ تابع مشخصه F باشد تابع مشخصه $G(x)=F(x)+c$ نیز می‌باشد که c ثابت است. به این دلیل

$$\Phi_G(u)=\int_R e^{iux}dG(x)=\int_R e^{iux}dF(x)=\Phi_F(u)=\Phi$$

$$dG(x)=d[F(x)+c]$$

تعریف۷-۱: اگر $0<\alpha<1$ ؛ $F_2(x), F_1(x)$ توابع توزیع باشند و F بصورت زیر تعریف شود

$$F(x)=\alpha F_1(x)+(1-\alpha)F_2(x)$$

در این صورت تابع مشخصه F بصورت زیر خواهد بود

$$F(x)=\alpha F_1(x)+(1-\alpha)F_2(x)$$

$$dF(x)=\alpha dF_1(x)+(1-\alpha)dF_2(x)$$

$$\Phi_F(u)=\int_R e^{iux}dF(x)=\int_R e^{iux}[\alpha dF_1(x)+(1-\alpha)dF_2(x)]$$

$$=\alpha\int_R e^{iux}\,dF_1(x)+(1-\alpha)\int_R e^{iux}\,dF_2(x)$$

$$= \alpha\Phi_{F_1}(u) + (1-\alpha)\Phi_{F_2}(u)$$

در حالت کلی اگر $F(x) = \sum_i \ \alpha_i F_i(x)$ باشد، که $\sum_i \ \alpha_i = 1$ آنگاه :

$$dF(x) = \sum_i \alpha_i \, dF_i(x)$$

و بنابراین خواهیم داشت :

$$\Phi_F(u) = \sum_i \alpha_i \, \Phi_{F_i}(u)$$

مثال۷-۲: با فرض اینکه F_1 توزیع تباهیده در صفر و F_2 توزیع نمایی باشد داریم :

$$F(x) = pF_1(x) + (1-p)F_2(x)$$

$$\Rightarrow \Phi_F(u) = p\Phi_{F_1}(u) + (1-p)\Phi_{F_2}(u)$$

$$= p + (1-p)(1-iu)^{-1}$$

۷-۱-۱ تابع مشخصه دومتغیره

فرض کنید X و Y دو متغیر تصادفی یا یک متغیر تصادفی دوبعدی باشند. در این‌صورت داریم

$$\Phi_{X,Y}(u,v) = E(e^{iuX+ivY})$$

$$\Phi_{X,Y}(u,0) = E(e^{iuX}) = \Phi_X(u)$$

$$\Phi_{X,Y}(0,v) = E(e^{ivY}) = \Phi_Y(v)$$

مثال۷-۳: فرض کنید چگالی احتمال (X,Y) بصورت زیر باشد

$$P(-1,-1) = \frac{1}{6} \quad ,P(-1,1) = \frac{1}{6} \quad ,P(1,-1) = \frac{1}{3} \quad ,P(1,1) = \frac{1}{3}$$

بنابراین، تابع مشخصه (X,Y) به صورت زیر در می‌آید

$$\Phi_{X,Y}(u,v) = \sum_x \sum_y e^{iux+ivy} P_{X,Y}(x,y)$$

$$= \frac{1}{6}e^{-iu-iv} + \frac{1}{6}e^{-iu+iv} + \frac{1}{3}e^{iu-iv} + \frac{1}{3}e^{iu+iv}$$

$$= \frac{1}{3}cosv(3cosu + i\sin u)$$

تعریف۷-۲: تابع مشخصه بردار تصادفی $(X_1, X_2, \dots, X_P)$ برابر است با

$$\Phi_{X_1,X_2,\dots,X_P}(u_1, u_2, \dots, u_p) = E(e^{iu_1X_1+iu_2X_2+\cdots+iu_pX_p}$$

مثال ۷-۴: تابع مشخصه توزیع کوشی استاندارد بصورت زیر است:

$$f(x) = \frac{1}{\pi(1+x^2)} \Rightarrow \Phi_x(u) = e^{-|u|}$$

و تابع مشخصه توزیع کوشی مضربی از تابع چگالی توزیع لاپلاس است. در اینصورت به زوج توزیع کوشی و لاپلاس، زوج‌های مزدوج گفته می‌شود.

۷-۲ چند خاصیت ساده

قضیه ۷-۱: اگر Φ تابع مشخصه یک متغیر تصادفی کلی مانند F باشد، آنگاه:

الف) Φ پیوسته است.

ب) $|\Phi_x(u)| \leq \Phi(0)$

ج) $\Phi_x(u)$ مزدوج مختلط $\overline{\Phi}(u)$ است و $\Phi_x(-u) = \overline{\Phi}(u)$

د) اگر $\Phi_x(u)$ تابع مشخصه X باشد، تابع مشخصه $a + bX$ برابر $e^{iua}\Phi(ub)$ است و $\overline{\Phi}$ تابع مشخصه $-X$ است.

هـ) Φ حقیقی است اگر و فقط اگر X نسبت به مبدأ قرینه باشد.

و) اگر $\Phi_n(u)$ ها توابع مشخصه باشند، با شرط $\sum_n \lambda_n = 1, \lambda_n \geq 0$ نتیجه می‌دهد $\sum \lambda_n \Phi_n(u)$ نیز تابع مشخصه است.

ی) با فرض اینکه $E(X^t)$ وجود داشته باشد $\varphi(t)$ دارای بسط مک لورن $\varphi(t) = \sum_{k=0}^{t} \mu_k^1 \frac{(it)^k}{k!} + O((it)^k)$ می‌باشد که در آن $O(x)$ کمیتی را نشان دهد که هنگامی که $x \to 0$ میل می کند داریم $O(x)/x \to 0$

ن) تابع مولد کومولان یا تابع مشخصه دوم $\Psi_{(t)} = Log\varphi(t)$ دارای بسط $\Psi_{(t)} = \sum_{j=0}^{t} K_j \frac{(it)^k}{j!} + o(t^k)$ است که در آن K_j کومولان مرتبه j ام X نامیده می‌شود.

اثبات: الف) کافی است با توجه به تعریف پیوسته بودن نشان دهیم که

$$\left|u - u'\right| \to 0$$

$$\left|\Phi(u) - \Phi(u')\right| \to 0$$

در نتیجه خواهیم داشت :

$$\left|\Phi(u) - \Phi(u')\right| = \left|\int e^{iuX} dF(x) - \int e^{iu'X} dF(x)\right|$$

$$\leq \int \left|e^{iuX}\left(1 - e^{i(u'-u)x}\right)\right| dF(x)$$

$$= \int \left|e^{iuX}\right|\left|1 - e^{i(u'-u)x}\right| dF(x)$$

$$= \int \left|1 - e^{i(u'-u)x}\right| dF(x)$$

و چون فرض کردیم $\left|u - u'\right| \to 0$ به صفر میل می کند، در نتیجه $e^{i(u'-u)} \to e^0 = 1$ میل می کند، در نهایت داریم :

$$\left|1 - e^{i(u'-u)x}\right| \to 0$$

$$\left|\Phi(u) - \Phi(u')\right| \to 0$$

در نتیجه Φ پیوسته است.

ب)

$$|\Phi(u)| = \left|\int e^{iuX} dF(x)\right| \leq \int \left|e^{iuX}\right| dF(x)$$

$$= \int dF(x) = F(+\infty) - F(-\infty) = 1 - 0 = 1$$

که قسمت (ب) را نتیجه می‌دهد.

ج)

$$\Phi(-u) = \int e^{-iuX} dF(x)$$

$$= \int \overline{(e^{iuX})} dF(x)$$

$$= \overline{\int e^{iuX} dF(x)} = \overline{\Phi(u\ )}$$

د)

$$\Phi_{a+bX}(u) = \int e^{iu(a+bx)} dF(x) = \int e^{iua+iubx} dF(x)$$

$$= e^{iua} \int e^{i(ub)x} dF(x) = e^{iua} \Phi_x(ub)$$

$$\overline{\Phi(u)} = \Phi(-u) \Rightarrow a = 0\,; b = -1$$

$$\overline{\Phi_x(u)} = \Phi_{-x}(u)$$

$$X = -X$$

هـ) ابتدا فرض می‌کنیم که X نسبت به مبدأ متقارن است. در نتیجه داریم

$$X = -X$$

$$\Phi_x(u) = \Phi_{-x}(u) = \overline{\Phi_x(u)}$$

در نتیجه ، Φ حقیقی است.

حال، عکس قضیه. فرض می‌کنیم Φ حقیقی باشد. در نتیجه داریم

$$\Phi_x(u) = \overline{\Phi_x(u)} = \Phi_{-x}(u)$$

$$X = -X$$

در نتیجه ، X نسبت به مبدأ متقارن است.

و) برای این قسمت کافی است F را بصورت زیر در نظر بگیریم

$$F(x) = \sum_n \lambda_n F_n(x) \Rightarrow dF(x) = \sum_n \lambda_n\, dF_n(x)$$

$$\Rightarrow \Phi_F(u) = \int e^{iux} dF(x) = \int e^{iuX}(\sum_n \lambda_n \, dF_n(x))$$

$$= \sum_n \lambda_n \int e^{iux} dF_n(x) = \sum_n \lambda_n \, \Phi_n(u)$$

که Φ_n ها تابع مشخصه هستند. در نتیجه Φ_F تابع مشخصه است.

نکته ۷-۲: اگر $u = 0$ باشد در نتیجه داریم $\Phi(u) = \Phi(0) = 1$ ولی عکس این حالت برقرار نیست، یعنی اگر $\Phi(u) = 1$ باشد، نمی‌توان نتیجه گرفت $u = 0$ است. با توجه به نکات گفته شده همواره $|\Phi(u)| \leq 1$

به عنوان مثال در توزیع دوجمله ای داریم :

$$X \sim B(n, p)$$

$$\Phi(u) = (pe^{iu} + q)^n$$

$$if \ u = 2n\pi \ \Rightarrow e^{iu} = 1$$

$$\Phi(u) = \Phi(2n\pi) = (p + q)^n = 1 \, ; u \neq 0$$

۷-۲-۱ چند نامساوی

لم۷-۱: برای هر تابع مشخصه مانند Φ داریم

الف) $Re(1 - \Phi(u)) \geq \frac{1}{4} Re(1 - \Phi(2u))$

ب)$|\Phi(u) - \Phi(u + h)|^2 \leq 2(1 - Re\Phi(h))$

ج)$\int_{|x| \leq u^{-1}} x^2 dF(x) \leq \frac{3}{u^2}(1 - Re\Phi(u))$

د) با استفاده از نامساوی جنسن- هولدر $1 + \Phi(2U) \geq 2[\Phi(U)]^2$

که در آنها $Re\Phi$ جزء حقیقی Φ است.

اثبات: الف)

$$Re\big(1 - \Phi(2u)\big) = 1 - Re\Phi(2u) = 1 - Re \int e^{2iux} dF(x)$$

$$= 1 - \int cos2uxdF(x)$$

$$= \int dF(x) - \int cos2uxdF(x)$$

$$= \int (1 - cos\ 2ux)dF(x)\ = 2\int (1 - \cos^2 ux)\ dF(x)$$

$$= 2\int (1 - \cos ux)\ (1 + \cos ux)dF(x)$$

$$\leq 4\int (1 - \cos ux)dF(x) = 4Re\big(1 - \Phi(u)\big)$$

ب)

$$|\Phi(u) - \Phi(u+h)|^2 = \left|\int e^{iux}dF(x) - \int e^{i(u+h)x}dF(x)\right|^2$$

$$= \left|\int e^{iux} - e^{i(u+h)x}dF(x)\right|^2$$

$$= \left|\int e^{iux}(1 - e^{ihx})\ dF(x)\right|^2$$

$$\leq [\int \left|e^{iux}\right| \left|1 - e^{ihx}\right|dF(x)]^2$$

$$= [\int \left|1 - e^{ihx}\right|dF(x)]^2$$

$$\leq \int \left|1 - e^{ihx}\right|dF(x)$$

$$= 2\int (1 - coshx)dF(x))$$

$$= 2(1 - Re\Phi(h))$$

ج)

$$\big(1 - Re\Phi(u)\big) = 1 - Re\int e^{iux}dF(x)$$

$$= 1 - Re\int (cosux + isinux)dF(x)$$

$$= 1 - \int cosuxdF(x) = \int (1 - cosux)dF(x)$$

$$\geq \int_R \frac{(ux)^2}{2}(1-\frac{(ux)^2}{12})dF(x)$$

$$\geq \int_{|x|\leq u^{-1}} \frac{(ux)^2}{2}(1-\frac{(ux)^2}{12})dF(x)$$

$$\geq \int_{|x|\leq u^{-1}} \frac{(ux)^2}{2}\left(\frac{11}{12}\right)dF(x)$$

$$\geq \frac{u^2}{3}\int_{|x|\leq u^{-1}} x^2 dF(x)$$

$$\Rightarrow 1-Re\Phi(u) \geq \frac{u^2}{3}\int_{|x|\leq u^{-1}} x^2 dF(x)$$

$$\int_{|x|\leq u^{-1}} x^2 dF(x) \leq \frac{3}{u^2}(1-Re\Phi(u))$$

د)

$$(Ef)^2 \leq E(f)^2 \quad ; g(x) = |x|^p$$

$$2[\Phi(U)]^2 = 2[E[e^{iux}]]^2 \leq 2E(e^{i(2u)x})$$

$$= E\left(e^{i(2u)x}\right) + E\left(e^{i(2u)x}\right) \leq 1 + \Phi(2u)$$

تعریف ۷-۳: تابع g را معین نامنفی گویند اگر برای هر مجموعه متناهی از عناصر $u_1,u_2,...,u_n$ و $v_1,v_2,...,v_n$ و هر تابع حقیقی یا مختلط مانند h داشته باشیم:

$$\sum_u \sum_v g(u-v)h(u)\bar{h}(v) \geq 0$$

بنابراین ، تابع مشخصه، معین نامنفی است.

لم ۷-۲: فرض کنید $(k=1,2,...,n)$ u_k, v_k مجموعه هایی دلخواه از اعداد حقیقی باشند. در این‌صورت برای هر تابع h (حقیقی یا مختلط) که بر این مجموعه تعریف شده باشد، داریم:

$$\sum_u \sum_v \Phi(u-v)h(u)\bar{h}(v) \geq 0$$

اثبات:

$$\sum_u \sum_v \Phi(u-v)h(u)\bar{h}(v) \geq 0$$

$$\sum_u \sum_v \int e^{i(u-v)x} dF(x)\, h(u)\bar{h}(v) \geq 0$$

$$\int \sum_u \sum_v e^{iux-ivx}\, h(u)\bar{h}(v) dF(x)$$

$$= \int (\sum_u e^{iux})(\sum_v e^{-ivx})\, dF(x)$$

$$= \int (\sum_u e^{iux})\overline{(\sum_v e^{ivx})}\, dF(x)$$

$$= \int \left|\sum_u e^{iux}\right|^2 dF(x) \geq 0 = \sum_u \sum_v \mathrm{g}(u-v)h(u)\bar{h}(v) \geq 0$$

نکته۷-۳: در قضیه زیر دیده می‌شود تابع مشخصه علاوه بر مشخص‌کردن یک توزیع وسیله‌ای برای مطالعه رفتار مجانبی دنباله‌ای از متغیرهای تصادفی می‌باشد.

قضیه ۷-۲: (**پیوستگی برای توابع مشخصه لوی-کرامر**)؛ فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی و $\{F_n(x)\}$ و $\{\varphi_n(t)\}$ دنباله‌های توابع توزیع و توابع مشخصه متناظرشان باشد. شرط لازم و کافی برای اینکه دنباله توابع توزیع $F_n(x)$ به تابع توزیع $F(x)$ در هر نقطه پیوستگی F میل کند این است که وقتی $n \to 0$ برای هر t داشته باشیم $\varphi_n(t) \to \varphi(t)$ و $\varphi(t)$ در $t = 0$ پیوسته باشد. هرگاه این رابطه برقرار باشد تابع مشخصه حدی $\varphi(t)$ تابع مشخصه تابع توزیع حدی F می‌باشد.

۷-۳ فرمول وارون

قضیه ۷-۳: (**فرمول وارون**)؛ اگر a و $b\left(a<b\right)$ نقاط پیوستگی F باشند. آنگاه

$$F(b)-F(a)=F(a,b)=\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}\Phi(u)du$$

اثبات: تابع زیر انتگرال را در نظر می‌گیریم

$$g(u)=\frac{e^{-iua}-e^{-iub}}{iu}\Phi(u)$$

چون تابع پیوسته است، مقدار تابع در $u=0$ تعریف شده است.

$$g(0)=(b-a)\Phi(0)$$

حال به اثبات قضیه می پردازیم :

$$I_u=\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}\Phi(u)du$$

$$=\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}[\int e^{iux}dF(x)]du$$

$$=\frac{1}{2\pi}\int\int_{-u}^{u}\frac{e^{iu(x-a)}-e^{iu(x-b)}}{iu}dudF(x)$$

$$=\frac{1}{2\pi}\int j_u\,F(x)\ ;\ j_u=\int_{-u}^{u}\frac{e^{iu(x-a)}-e^{iu(x-b)}}{iu}du$$

$$j_u=\int_{-u}^{u}\frac{\cos u(x-a)-\cos u(x-b)+i[\sin u(x-a)-\sin u(x-b)]}{iu}du$$

$$=2\int_{0}^{u}\frac{\sin u(x-a)-\sin u(x-b)}{u}du$$

می‌دانیم

$$\int_{0}^{\infty}\frac{\sin ax}{x}dx=(signa)\frac{\pi}{2}\qquad ;\ signa=\begin{cases}-1 & a<0\\ 0 & a=0\\ 1 & a>0\end{cases}$$

در نتیجه اگر از طرفین تساوی حد بگیریم داریم :

$$\lim_{u\to\infty} 2\int_{0}^{u} \frac{sinu(x-a)-\mathrm{sinu}(x-b)}{u}du$$

$$=2\int_{0}^{\infty} \frac{sinu(x-a)}{u}du-2\int_{0}^{\infty} \frac{\mathrm{sinu}(x-b)}{u}du$$

$$=2\left(sign(x-a)\tfrac{\pi}{2}-sign(x-b)\tfrac{\pi}{2}\right)$$

در نهایت:

$$F(a,b)=\lim_{u\to\infty} I_u=\lim_{u\to\infty}\int j_u dF(x)$$

$$= \frac{1}{2\pi}\int 2[sign(x-a)\frac{\pi}{2}-sign(x-b)\frac{\pi}{2}]dF(x)$$

$$=\frac{1}{\pi}[\int_{\{a\}}\frac{\pi}{2}dF(x)+\int_{a}^{b}\pi\, dF(x)+\int_{\{b\}}\frac{\pi}{2}dF(x)]$$

$$=\tfrac{1}{\pi}\pi\int_a^b dF(x)=F(b)-F(a)=F(a,b)$$

که نتیجه مشاهده می‌شود.

نتیجه ۷-۱:

$$\lim_{u\to\infty} I_u=\frac{1}{2}[F(b+0)-F(b-0)]-\frac{1}{2}[F(a+0)-F(a-0)]$$

اثبات:

$$\lim_{u\to\infty} I_u=\int_{\{a\}} dF(x)+\int_{a}^{b} dF(x))+\int_{\{b\}} dF(x)$$

$$=\frac{1}{2}[F(a+0)-F(a-0)]+[F(b-0)-F(a+0)]+\frac{1}{2}[F(b+0)-F(b-0)]$$

$$=\frac{1}{2}F(a+0)-\frac{1}{2}F(a-0)+F(b-0)-F(a+0)+\frac{1}{2}F(b+0)-\frac{1}{2}F(b-0)$$

$$=\frac{1}{2}[F(b+0)+F(b-0)]-\frac{1}{2}[F(a+0)-F(a-0)]$$

نتیجه ۷-۲: تابع توزیع یک متغیر تصادفی F و تابع مشخصه آن یکدیگر را معین می‌کنند.

اثبات: اگر F تابع توزیع باشد، طبق تعریف تابع مشخصه داریم:

$$\Phi(u)=\int e^{iux}dF(x)$$

بنابراین F، تابع مشخصه را معین می‌کند.

حال اگر تابع مشخصه را داشته باشیم

$$F(b)-F(a)=F(a,b)=\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}\Phi(u)du$$

حال می‌خواهیم نشان دهیم که Φ ،F را بطور یکتا معین می‌کند.

فرض می‌کنیم تابع مشخصه دو توزیع F و F' با هم برابر باشد.

$$\Phi_F(u)=\Phi_{F'}(u)$$

$$F(a,b)=\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}\Phi_F(u)du$$

$$=\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iua}-e^{-iub}}{iu}\Phi_{F'}(u)du=F'(a,b)$$

$$\Rightarrow F(a,b)=F'(a,b)\ \Rightarrow F(b)-F(a)=F'(b)-F'(a)$$

$$\Rightarrow F(b)-F'(b)=F(a)-F'(a)$$

اگر b را تغییر دهیم، برای مقدار ثابت a داریم

$$b=\infty\ \Rightarrow\ F(b)-F'(b)=F(\infty)-F'(\infty)=1-1=0=F(a)-F'(a)$$

چون $F(a)-F'(a)$ مقداری ثابت است و در اینجا صفر شد. پس در همه جا صفر است. بنابراین

$$F(b)-F'(b)=0$$

$$\forall b\qquad F(b)=F'(b)$$

نتیجه ۷-۳: اگر F در X مشتق‌پذیر باشد، $F'(x) = f(x)$ آنگاه

$$f(x) = \lim_{h\to 0}\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\left(\frac{1-e^{iuh}}{iuh}\right)e^{-iux}\Phi(u)du$$

اگر و فقط اگر حد سمت راست عبارت فوق وجود داشته باشد. اگر Φ بر R انتگرال‌پذیر باشد و $F(X)$ برای تمام مقادیر وجود داشته باشد، آنگاه

$$f(x) = \frac{1}{2\pi}\int_{R} e^{-iux}\,\Phi(u)du$$

اثبات: چون F در X مشتق‌پذیر است، داریم

$$F'(x) = \lim_{h\to 0}\frac{F(x+h)-F(x)}{h}$$

$$= \lim_{h\to 0}\frac{\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iux}-e^{-iu(x+h)}}{iu}\Phi(u)du}{h}$$

$$= \lim_{h\to 0}\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iux}-e^{-iu(x+h)}}{iuh}\Phi(u)du$$

$$= \lim_{h\to 0}\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iux}(1-e^{-iuh})}{iuh}\Phi(u)du$$

زمانی وجود دارد که حد سمت راست وجود داشته باشد.

اگر Φ انتگرال‌پذیر باشد، $|\Phi|$ نیز انتگرال‌پذیر است. در نتیجه داریم:

$$\left|\int_{-u}^{u}\frac{(1-e^{-iuh})}{iuh}\Phi(u)e^{-iux}du\right| \le \int_{-u}^{u}|\Phi(u)|\,du \le \int_{R}|\Phi(u)|\,du < \infty$$

لذا می‌توانیم جای حد و انتگرال را عوض کنیم.

$$f(x) = \lim_{h\to 0}\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-iux}(1-e^{-iuh})}{iuh}\Phi(u)du$$

$$= \lim_{h\to 0}\frac{1}{2\pi}\int_R \frac{e^{-iux}(1-e^{-iuh})}{iuh}\,\Phi(u)du$$

$$= \frac{1}{2\pi}\int_R e^{-iux}[\lim_{h\to 0}\frac{1-e^{-iuh}}{iuh}]\,\Phi(u)du$$

$$= \frac{1}{2\pi}\int_R e^{-iux}\,\Phi(u)du$$

$$\Rightarrow f(x) = \frac{1}{2\pi}\int_R e^{-iux}\,\Phi(u)du$$

نتیجه ۷-۴: تحت شرایط نتیجه ۷-۳، $F(X)$ پیوسته است. اگر $f(x) \geq 0$ و $F(x) = \int_{-\infty}^{x} f(x)dt$ آنگاه $f(x)$ پیوسته است. در نتیجه $F(X)$ از چپ پیوسته است.

بنابراین کافی است ثابت کنیم f پیوسته است.

$$|f(x)-f(x-\delta)| = \left|\frac{1}{2\pi}\int_R e^{-iux} - e^{-iu(x-\delta)}\,\Phi(u)du\right|$$

$$= \left|\frac{1}{2\pi}\int_R e^{-iux}(1-e^{iu\delta})\,\Phi(u)du\right|$$

$$\leq \frac{1}{2\pi}\int_R \left|1-e^{iu\delta}\right|\Phi(u)du$$

اگر $\delta \to 0$ آنگاه $e^{iu\delta} \to 1$ آنگاه $\left|1-e^{iu\delta}\right| \to 0$

$$|f(x)-f(x-\delta)| \to 0$$

در نتیجه f از چپ پیوسته است و از آنجایی که تابع توزیع از راست پیوسته است لذا f از راست پیوسته است بنابراین f پیوسته است و در نتیجه F پیوسته است.

نتیجه ۷-۵: به ازای هر $x \in R$:

$$F(x+0)-F(x-0)=\lim_{u\to\infty}\frac{1}{2u}\int_{-u}^{u} e^{-iux}\,\Phi(u)du$$

اثبات:

$$\left|\frac{1}{2u}\int_{-u}^{u} e^{-iux}\int_{R} e^{iut}dF(t)\,du\right|=\left|\frac{1}{2u}\int_{-u}^{u}\int_{R} e^{iu(t-x)}dF(t)du\right|$$

$$\leq\frac{1}{2u}\int_{-u}^{u}\int_{R}\left|e^{iu(t-x)}\right|dF(t)du=1$$

$$\Rightarrow\left|\frac{1}{2u}\int_{-u}^{u} e^{-iux}\int_{R} e^{iut}dF(t)\,du\right|\leq 1$$

پس با تغییر کران‌های انتگرال‌گیری داریم

$$\lim_{t\to x}\int_{-\infty}^{+\infty}\frac{sinu(t-x)}{u(t-x)}dF(t)=\int_{-\infty}^{+\infty}dF(t)=F(x+0)-F(x-0)$$

مثال ۷-۵: با استفاده از تابع مشخصه توزیع کوشی، تابع چگالی مربوطه رابیابید.

$$\Phi(u)=e^{-|u|}$$

$$f(x)=\frac{1}{2\pi}\int_{R} e^{-iux}\,e^{-|u|}du$$

$$=\frac{2}{2\pi}\int_{0}^{\infty} 2\cos uxe^{-u}du=\frac{1}{\pi(1+x^2)}$$

۷-۳-۱ فرمول وارون برای توزیع های مشبکه ای

اگر X یک متغیر تصادفی گسسته با مقادیر $0,\pm\lambda,\pm2\lambda,...$ باشد آنگاه X دارای توزیع حسابی با مولد λ می‌باشد.

اگر مبدأ را a قرار دهیم ، یعنی X مقادیر $a, a\pm\lambda, a\pm2\lambda,...$ را بگیرد، گوییم X دارای توزیع مشبکه‌ای است.

فرض کنید X متغیر حسابی با مولد $\lambda=1$ باشد. در این‌صورت داریم:

$$p\{X=x\}=p_t$$

$$\Phi(u)=\sum_{t=-\infty}^{\infty} p_t e^{iut}$$

$$p_x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(u)e^{-iux}du$$

اثبات:

$$\int_{-\pi}^{\pi}\Phi(u)e^{-iux}du=\int_{-\pi}^{\pi}\sum_t p_t e^{iut}\,e^{-iux}du$$

$$=\sum_t\int_{-\pi}^{\pi}p_t e^{iu(t-x)}\,du$$

$$=p_x\int_{-\pi}^{\pi}du+\sum_{t\neq x}p_t\int_{-\pi}^{\pi}e^{iu(t-x)}\,du \qquad =2\pi p_x$$

$$\Rightarrow p_x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(u)e^{-iux}du$$

مثال ۷-۶: با استفاده از تابع مشخصه توزیع هندسی، تابع چگالی مربوطه را بیابید.

$$\Phi(u)=p(1-qe^{iu})^{-1}$$

$$p_x=\frac{1}{2\pi}\int_{-\pi}^{\pi}p(1-qe^{iu})^{-1}e^{-iux}du=\frac{p}{2\pi}\int_{-\pi}^{\pi}\frac{1}{1-qe^{iu}}e^{-iux}du$$

$$=\frac{p}{2\pi}\int_{-\pi}^{\pi}\sum_{n=0}^{\infty}(qe^{iu})^n\,e^{-iux}du=\frac{p}{2\pi}\sum_{n=0}^{\infty}q^n\int_{-\pi}^{\pi}e^{iu(n-x)}\,du$$

اگر $n=x$ آنگاه $\int_{-\pi}^{\pi}e^{iu(n-x)}\,du=2\pi$

$$p_x=pq^x$$

مثال ۷-۷: با استفاده از تابع مشخصه توزیع دوجمله‌ای، تابع چگالی مربوطه را بیابید.

$$\Phi(u) = (q + pe^{iu})^n$$

$$p_x = \frac{1}{2\pi}\int_{-\pi}^{\pi}(q + pe^{iu})^n\, e^{-iux}du$$

$$= \binom{n}{x} p^x q^{n-x} \;\; ; x = 0,1, \dots, n$$

لم۷-۳: اگر γ_r گشتاور مطلق مرتبه r ام $F(X)$ متناهی باشد، تابع مشخصه آن r مرتبه مشتق‌پذیر است و مشتق r ام برابر است با

$$\Phi^{(r)}(u) = \int (ix)^r e^{iux} dF(x)$$

به عکس، اگر $\Phi^{(r)}(0)$ موجود و متناهی باشد، آنگاه برای $r' < r \;\; , \;\; \gamma_{r'} < \infty$

اثبات: ابتدا نشان می‌دهیم که وقتی که γ_r متناهی باشد، مقدار $\Phi^{(r)}(u)$ موجود است.

$$\left|\Phi^{(r)}(u)\right| = \left|\int (ix)^r e^{iux} dF(x)\right| \le \int |x|^r \left|e^{iux}\right| dF(x)$$

$$= \int |x|^r dF(x) = \gamma_r < \infty$$

حال نشان می‌دهیم که فرمول مشتق مرتبه r ام $ch.fn$ به همان شکل بالاست.

$$\Phi'(u) = \lim_{h\to 0}\frac{\Phi(u+h) - \Phi(u)}{h}$$

$$= \lim_{h\to 0}\frac{\int \mathrm{e}^{i(u+h)x}\, dF(x) - \int \mathrm{e}^{iux} dF(x)}{h} = \lim_{h\to 0}\int \frac{\mathrm{e}^{iux}(e^{ihx} - 1)}{h} dF(x)$$

$$= \int \mathrm{e}^{iux} \lim_{h\to 0}\left(\frac{e^{ihx} - 1}{h}\right) dF(x) = \int \mathrm{e}^{iux}\; ix\; dF(x)$$

$$= \int \; (\, ix\,)\, \mathrm{e}^{iux} dF(x)$$

$$\Phi''(u) = \lim_{h\to 0}\frac{\Phi'(u+h) - \Phi'(u)}{h}$$

$$= \lim_{h\to 0}\int (ix)\, \mathrm{e}^{iux}\,\frac{e^{ihx}-1}{h}\,dF(x)$$

$$= \int\ (\ ix\ )\ \mathrm{e}^{iux}(\lim_{h\to 0}\frac{e^{ihx}-1}{h})dF(x)$$

$$= \int (\ ix\ )^2\ \mathrm{e}^{iux}dF(x)$$

و به راحتی از استقراء ثابت می‌شود

$$\Phi^{(n)}(u) = \int (ix)^n e^{iux} dF(x)$$

حال، **عکس لم**:

فرض می‌کنیم $\Phi^{(r)}(0)$ موجود باشد، در این صورت داریم

$$\Phi'(0) = \left|\lim_{h\to 0}\frac{\Phi(h)-\Phi(-h)}{2h}\right|$$

$$= \lim_{h\to 0}\left|\int \frac{e^{ihx}-e^{-ihx}}{2h}\,dF(x)\right|$$

$$= \lim_{h\to 0}\left|\int \frac{2i\ sin\ hx}{2h}\,dF(x)\right|$$

$$= \lim_{h\to 0}\left|\int \frac{i\ sin\ hx}{h}\,dF(x)\right|$$

$$= \lim_{h\to 0}\left|\int \frac{sinhx}{h}\,ixdF(x)\right|$$

به راحتی با استقراء ثابت می‌شود که

$$\Phi^{(r)}(0) = \left|\lim_{h\to 0}\int \left(\frac{sinhx}{hx}\right)^r (ix)^r\, dF(x)\right|$$

اگر r زوج باشد، چون $\Phi^{(r)}(0)$ وجود دارد، لذا γ_r نیزموجود است. اما اگر r فرد باشد، می‌دانیم $r-1$ زوج است و حالت فوق تا $r-1$ برقرار است، بنابراین نتیجه مورد نظر برقرار است.

۷–۴ سری تیلور برای $\Phi(u)$

قضیه ۷–۴: اگر γ_r متناهی باشد، آنگاه

$$\Phi(u)=\sum_{s=0}^{r-1}\frac{\mu_s' iu^s}{s!}+u^r\int_0^1\frac{(1-t)^{(r-1)}}{(r-1)!}\Phi^{(r)}(tu)dt$$

اثبات: با استفاده از انتگرال‌گیری جزء به جزء داریم

$$(iux)^r\int_0^1\frac{(1-t)^{(r-1)}}{(r-1)!}e^{iutx}dt=e^{iux}-\sum_{s=0}^{r-1}\frac{(iux)^s}{s!}$$

حال، از عبارت بالا نسبت به $F(X)$ انتگرال می‌گیریم.

$$\int_R(iux)^r\int_0^1\frac{(1-t)^{(r-1)}}{(r-1)!}e^{iutx}dtdF(x)$$

$$=\int_R e^{iux}dF(x)-\int_R\sum_{s=0}^{r-1}\frac{(iux)^s}{s!}dF(x)$$

$$=\int_R(iux)^r\int_0^1\frac{(1-t)^{(r-1)}}{(r-1)!}e^{iutx}dtdF(x)$$

$$=\int_0^1\int_R\frac{(1-t)^{(r-1)}}{(r-1)!}e^{iutx}dtdF(x)$$

$$=\int_0^1u^r\left[\int_R(\,ix\,)^r\,\mathrm{e}^{iutx}dF(x)\right]\frac{(1-t)^{(r-1)}}{(r-1)!}dt$$

$$=\int_0^1u^r\Phi^{(r)}(tu)\frac{(1-t)^{(r-1)}}{(r-1)!}dt$$

$$=u^r\int_0^1\frac{(1-t)^{(r-1)}}{(r-1)!}\Phi^{(r)}(tu)dt$$

$$= \int_R e^{iux} dF(x) - \int_R \sum_{s=0}^{r-1} \frac{(iux)^s}{s!} dF(x)$$

$$= \Phi(u) - \sum_{s=0}^{r-1} \frac{(iu)^s \mu_s'}{s!}$$

$$\Rightarrow \Phi(u) - \sum_{s=0}^{r-1} \frac{(iu)^s \mu_s'}{s!} = u^r \int_0^1 \frac{(1-t)^{(r-1)}}{(r-1)!} \Phi^{(r)}(tu) dt$$

$$\Rightarrow \Phi(u) = \sum_{s=0}^{r-1} \frac{(iu)^s \mu_s'}{s!} + u^r \int_0^1 \frac{(1-t)^{(r-1)}}{(r-1)!} \Phi^{(r)}(tu) dt$$

و در نتیجه قضیه اثبات می‌شود.

نتیجه۷-۶: اگر γ_r متناهی باشد، خواهیم داشت

$$\Phi(u) = \sum_{s=0}^{r-1} \mu_s' \frac{(iu)^s}{s!} + \rho \mu_r' \frac{(iu)^r}{r!}$$

که در آن $|\rho| \leq 1$ است و وقتی که u به سمت صفر میل کند، ρ به سمت یک میل می‌کند. علاوه بر این اگر تمام گشتاورها متناهی باشند، داریم:

$$\Phi(u) = \sum_{s=0}^{r-1} \frac{(iu)^s}{s!} \mu_s'$$

اثبات: در اثبات قضیه بالا عبارت زیر را داشتیم:

$$e^{iux} = \sum_{s=0}^{r-1} \frac{(iux)^s}{s!} + (iux)^r \int_0^1 \frac{(1-t)^{(r-1)}}{(r-1)!} e^{iutx} dt$$

$$\left| \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} e^{iutx} dt \right| \leq \int_0^1 \left| \frac{(1-t)^{r-1}}{(r-1)!} \right| \left| e^{iutx} \right| dt$$

$$= \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} dt = \int_0^1 \frac{u^{r-1}}{(r-1)!} du = \frac{1}{r} \frac{1}{(r-1)} u^r \Big|_{-1}^{0} = \frac{1}{r!}$$

$$\Rightarrow (u \longrightarrow 0)\ \Rightarrow \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!}dt \longrightarrow \frac{1}{r!}$$

درنتیجه خواهیم داشت:

$$\int_0^1 \frac{(1-t)^{r-1}}{(r-1)!}dt = \frac{\rho}{r!} \quad ; \quad |\rho| \le 1 \quad ; \quad \begin{cases} u \to 0 \\ \rho \to 1 \end{cases}$$

بنابراین خواهیم داشت:

$$e^{iux} = \sum_{s=0}^{r-1} \frac{(iux)^s}{s!} + \frac{(iux)^r}{r!}\rho$$

حال، اگر از طرفین عبارت فوق نسبت به $F(X)$ انتگرال بگیریم، خواهیم داشت:

$$\Phi(u) = \sum_{s=0}^{r-1} \mu_s^{'} \frac{(iu)^s}{s!} + \rho\mu_r^{'} \frac{(iu)^r}{r!}$$

این در حالتی است که γ_r متناهی باشد. اگر به ازای هر γ ای ، γ_r متناهی باشد، داریم

$$r \to \infty \ \Rightarrow\ \Phi(u) = \sum_{s=0}^{r-1} \frac{(iu)^s}{s!}\mu_s^{'}$$

نتیجه۷-۷: اگر $\gamma_{r+\delta} < \infty$ ، $0 < \delta \le 1$ آنگاه

$$\rho\mu_r^{'} \frac{(iu)^r}{r!} = \mu^{'} \frac{(iu)^r}{r!} + 2^{1-\delta}\theta\gamma_{r+\delta} \frac{|u|^{\delta+r}2^{1-\delta}\theta\gamma_{r+\delta}}{(\delta+1)(\delta+2)\dots(\delta+r)}$$

اثبات:

در اثبات قضایا و نتایج فوق از معادلات زیر استفاده کردیم.

$$\left|e^{ia}-1\right| \le |a| \le 2\left|\frac{a}{2}\right|^{\delta} \qquad \left|\frac{a}{2}\right| < 1$$

$$\left|e^{ia}-1\right| \le 2 \le 2\left|\frac{a}{2}\right|\delta \qquad \left|\frac{a}{2}\right| \ge 1$$

$$e^{iux} - \sum_{s=0}^{r-1} \frac{(iux)^s}{s!} = (iux)^r \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!}e^{iutx}dt$$

$$= e^{iux} - \sum_{s=0}^{r-1} \frac{(iux)^s}{s!} = \rho \frac{(iux)^r}{r!}$$

$$\rho \frac{(iux)^r}{r!} = (iux)^r \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} e^{iutx} dt$$

$$= (iux)^r \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} (e^{iutx} - 1) dt + \frac{(iux)^r}{r!}$$

$$\left| \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} (e^{iutx} - 1) dt \right| \leq \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} \left|(e^{iutx} - 1)\right| dt$$

$$\leq \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} 2 \left|\frac{utx}{2}\right|^{\delta} dt$$

$$\Rightarrow \left| \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} (e^{iutx} - 1) dt \right| \leq \frac{2^{1-\delta} |ux|^{\delta} \theta}{(\delta+1)(\delta+2)\dots(\delta+r)}$$

$$\rho \frac{(iux)^r}{r!} = (iux)^r \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} (e^{iutx} - 1) dt + \frac{(iux)^r}{r!}$$

$$= \frac{2^{1-\delta} |ux|^{\delta} \theta |iux|^r}{(\delta+1)(\delta+2)\dots(\delta+r)} + \frac{(iux)^r}{r!}$$

در نتیجه با انتگرال‌گیری از تساوی فوق نسبت به $F(X)$ خواهیم داشت:

$$\rho \mu_r^{'} \frac{(iu)^r}{r!} = \mu^{'} \frac{(iu)^r}{r!} + \frac{|u|^{\delta+r} 2^{1-\delta} \theta \gamma_{r+\delta}}{(\delta+1)(\delta+2)\dots(\delta+r)}$$

قضیه ۷-۵: اگر سری $\sum_{r=0}^{\infty} \frac{\gamma_r u^r}{r!}$ به ازای مقداری مانند $u_0 > 0$ همگرا باشد، آنگاه دنباله گشتاورهای $\{\mu_r^{'}\}$ توزیع F را بطور یکتا معین می‌سازد.

اثبات: اگر $\sum_{r=0}^{\infty} \frac{\gamma_r u^r}{r!}$ همگرا باشد آنگاه برای u_0 داریم

$$|h| < u_0 \quad , \quad n \to \infty$$

$$\frac{\gamma_n |h|^n}{n!} \underset{n\to\infty}{\to} 0$$

بنابراین در بسط تیلور تابع مشخصه، قدر مطلق جمله باقیمانده به سمت صفر میل می‌کند و داریم

$$\Phi(u) = \sum_r \frac{iu^r}{r!}\mu'_r$$

چون Φ تابع F را بطور یکتا معین می‌کند و Φ از طریق $\{\mu'_r\}$ معین می‌شود، $\{\mu'_r\}$ نیز توزیع F را بطور یکتا معین می‌کند.

نتیجه۷-۸: اگر

$$\overline{\lim}\left(\frac{{\gamma_n}^{\frac{1}{n}}}{n}\right) = \overline{\lim}\left(\frac{{\mu'_r}^{\frac{1}{n}}}{n}\right) = 1 < \infty$$

آنگاه مجموعه گشتاورها ، توزیع را بطور یکتا معین می‌کنند.

اثبات: شعاع همگرایی سری $\sum_{r=0}^{\infty}\frac{\gamma_r u^r}{r!}$ بصورت زیر است.

$$\rho = [\overline{\lim}\left(\frac{{\gamma_n}^{\frac{1}{n}}}{n!}\right)]^{-1}$$

برای یافتن شعاع همگرایی از آزمون ریشه استفاده می‌کنیم. و با استفاده از تقریب استرلینگ شعاع همگرایی به صورت زیر بدست می‌آید.

تقریب استرلینگ برابر است با $(n!)^{\frac{1}{n}} \approx \frac{n}{e}$

در نتیجه برای شعاع همگرایی خواهیم داشت

$$[\overline{\lim}\left(\frac{{\gamma_n}^{\frac{1}{n}}}{n!}\right)]^{-1} = \overline{\lim}\left(\frac{{\gamma_n}^{\frac{1}{n}}}{\frac{n}{e}}\right) = e\overline{\lim}(\frac{{\gamma_n}^{\frac{1}{n}}}{n}) < \infty$$

با توجه به نامساوی لیاپانف داریم:

$${\gamma_{2n-1}}^{\frac{1}{2n-1}} \le {\gamma_{2n}}^{\frac{1}{2n}} \le {\gamma_{2n+1}}^{\frac{1}{2n+1}}$$

$$\overline{\lim}\left(\frac{\gamma_{2n-1}{}^{\frac{1}{2n-1}}}{2n-1}\right) \le \overline{\lim}\left(\frac{\gamma_{2n}{}^{\frac{1}{2n}}}{2n}\right) \le \overline{\lim}\left(\frac{\gamma_{2n+1}{}^{\frac{1}{2n+1}}}{2n+1}\right)$$

$$\Rightarrow \overline{\lim}\left(\frac{\mu_{2n}{}^{\frac{1}{2n}}}{2n}\right) = \overline{\lim}\left(\frac{\gamma_n{}^{\frac{1}{n}}}{n}\right) = 1$$

در نتیجه شعاع همگرایی دو سری $\sum_{r=0}^{\infty} \frac{\gamma_r u^r}{r!}$ و $\sum_{r=0}^{\infty} \frac{\mu_r u^r}{r!}$ با هم برابر است و چون $l < \infty$ است، در نتیجه سری اول همگراست و بنابراین، سری دوم نیز همگرا می‌باشد.

۵–۷ قضیه بکنر

تابع $\boldsymbol{g}$ بر $\boldsymbol{R}$ یک تابع مشخصه است اگر و فقط اگر معین نامنفی و پیوسته باشد.

اثبات: g تابع مشخصه است. از قبل ثابت کردیم که تابع مشخصه معین نامنفی و پیوسته است. پس، یک طرف حکم اثبات شد.

حال، طرف دیگر حکم: فرض می‌کنیم g یک تابع معین نامنفی و پیوسته باشد. باید نشان دهیم که تابع مشخصه است.

تعریف معین نامنفی:

$$\sum_u \sum_v \mathrm{g}(u-v)h(u)\bar{h}(v) \ge 0$$

چون g پیوسته است، داریم:

$$\int\int g(u-v)h(u)\bar{h}(v)dudv \ge 0$$

به ازای هر تابع (حقیقی یا مختلطh) تابع h را به صورت زیر در نظر می‌گیریم

$$h(u) = e^{-iux-u^2\sigma^2}$$

$$= \int\int g(u-v)h(u)\bar{h}(v)dudv = 0$$

$$= \int\int g(u-v)e^{-iux-u^2\sigma^2}\overline{\left(e^{-iux-u^2\sigma^2}\right)}dudv$$

$$= \int\int g(u-v)\, e^{-iux-u^2\sigma^2} e^{ivx-u^2\sigma^2}dudv$$

حال با استفاده از تغییر متغیر بدست می‌آوریم:

$$u-v=t$$

$$u+v=s$$

$$\Rightarrow\ u^2+v^2=\frac{s^2+t^2}{2}$$

$$= \int\int g(t)e^{-itx-(\frac{s^2+t^2}{2})\sigma^2}\frac{1}{2}dtds$$

$$= \int g(t)e^{-itx-\frac{t^2\sigma^2}{2}}\frac{1}{2}[\int e^{-\frac{s^2\sigma^2}{2}}ds]dt$$

$$= \sqrt{\frac{\pi}{2}}\frac{1}{\sigma}\int g(t)e^{-itx-\frac{t^2\sigma^2}{2}}dt \geq 0$$

$$\Rightarrow f(x,\sigma) = \int \frac{1}{2\pi}g(t)e^{-itx-\frac{t^2\sigma^2}{2}}dt \geq 0$$

$$\int f(x,\sigma) = \int\int \frac{1}{2\pi}g(t)e^{-itx-\frac{t^2\sigma^2}{2}}dt\ \geq 0$$

$$= \lim_{\beta\to 0}\int\int \frac{1}{2\pi}g(t)e^{-itx-\frac{t^2\sigma^2}{2}}e^{\frac{-x^2\beta^2}{2}}$$

$$= \lim_{\beta\to 0}\int \frac{g(t)}{\sqrt{2\pi}}e^{-\frac{t^2u^2}{2}}[\int \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2\beta^2}{2}-itx}dx]\,dt$$

$$= \lim_{\beta\to 0}\int \frac{g(t)}{\sqrt{2\pi}}e^{\frac{-t^2u^2}{2}}(e^{\frac{-t^2}{2}}\frac{1}{\beta})dt$$

$$= \lim_{\beta\to 0}\int \frac{g(ub)}{\sqrt{2\pi}}e^{\frac{-t^2\beta^2\sigma^2}{2}-\frac{u^2}{2}}du$$

$$= \int (\lim_{\beta\to 0}\frac{g(ub)}{\sqrt{2\pi}}e^{\frac{-t^2\beta^2\sigma^2}{2}-\frac{u^2}{2}})du$$

$$= \int \frac{g(0)}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du = g(0) \geq 0$$

$$f(x,\sigma) \geq 0 \;\; \Rightarrow \;\; \int f(x,\sigma)dx \geq 0 \;\; \Rightarrow \;\; g(0) \geq 0$$

پس $f(x,\sigma)$ می‌تواند تابع چگالی یک تابع توزیع کلی باشد.

می‌دانیم که تعریف تابع مشخصه به صورت زیر است

$$\Phi(t) = \int f(x,\sigma)e^{iut}dx$$

$$\begin{aligned}
\int f(x,\sigma)e^{iut}dx &= \lim_{\beta\to 0}\int f(x,\sigma)e^{iut-\frac{x^2\beta^2}{2}}dx \\
&= \lim_{\beta\to 0}\int \frac{g(t+u\beta)}{\sqrt{2\pi}}e^{-\frac{(t+u\beta)^2\sigma^2}{2}-\frac{u^2}{2}}du \\
&= \int \frac{g(t)}{\sqrt{2\pi}}e^{-\frac{t^2\sigma^2}{2}-\frac{u^2}{2}}du \\
&= g(t)e^{-\frac{t^2\sigma^2}{2}}
\end{aligned}$$

بنابراین برای هر σ ، $g(t)e^{-\frac{t^2\sigma^2}{2}}$ تابع مشخصه برای $f(x,\sigma)$ است. پس، طبق خاصیت پیوستگی داریم:

$$\lim_{\sigma\to 0} g(t)e^{-\frac{t^2\sigma^2}{2}} = g(0)$$

پس $g(t)$ نیز یک تابع مشخصه است. اگر $g(0)$ برابر با یک باشد، آنگاه تابع مشخصه یک متغیر تصادفی خواهد بود.

۷–۶ مسائل

۱- با فرض اینکه X دارای تابع احتمال زیر باشد، تابع مشخصه X را بدست آورید.

$$P_X(x) = \begin{cases} \frac{1}{3} & x = 0 \\ \frac{1}{3} & x = 1 \\ \frac{1}{3} & x = 2 \\ 0 & x \notin R_X \end{cases}$$

۲- از تابع مشخصه‌ای که درسوال قبل بدست آمده استفاده کنید و واریانس X را بدست آورید.

فصل هشتم

همگرایی

در این فصل $F(\,.\,)$ تابعی یکنواخت، نانزولی و از راست پیوسته است بطوری که $F(\infty)-F(-\infty)<\infty$ یعنی $F(X)$ یک تابع توزیع کلی است. در این فصل به معرفی همگرایی ضعیف و همگرایی کامل می‌پردازیم. همچنین قضایایی را در ارتباط با انتگرال توابع مشخصه و گشتاورها بیان می‌کنیم. لازم به ذکر است که همگرایی در توزیع، نقش عمده‌ای در آمار دارد. مباحثی مانند نظریه نمونه‌های بزرگ، که در آن قضیه حد مرکزی نقشی کلیدی دارد، یکی از این موارد است. خواص برآوردگرها در نمونه‌های بزرگ، کارایی نسبی در استنباط آماری ناپارامتری نیز با استفاده از همگرایی ضعیف توزیع‌ها بدست می‌آید.

۸–۱ همگرایی ضعیف و کامل توابع توزیع

تعریف ۸-۱: (**همگرایی ضعیف**)؛ دنباله‌ای از توابع توزیع F_n را در نظر بگیرید. گوییم دنبال F_n به‌طور ضعیف به یک تابع توزیع مانند F همگراست اگر در تمام نقاط پیوستگی F ، F_n به‌طور یکنواخت به F همگرا باشد به عبارت دیگر

$$F_n \overset{w}{\rightarrow} F \;\Leftrightarrow\; F_n(x)\rightarrow F(x) \qquad \forall\, x\in C(F)$$

یادآوری لم ۴-۱: فرض کنید F_D تابعی نانزولی متناهی باشد که بر D تعریف شده است تابع $F(x)$ را بر R به صورت زیر تعریف می‌کنیم:

$$F(x)=\inf_{x_n\geq x} F_D(x_n) \qquad ;\quad x_n\in D\,,\;\; x\in R$$

نتیجه لم ۴-۱: اگر دو تابع F_1 و F_2 بر مجموعه‌ی چگال D در R با یکدیگر برابر باشند روی کلیه نقاط R با هم برابرند.

لم ۸-۱: اگر $F_n \overset{w}{\rightarrow} F$ آنگاه F یکتاست.

اثبات: فرض کنیم F_n به دو تابع توزیع F و F' همگرا باشد، بنابراین

$$F_n \overset{w}{\to} F \iff F_n(x) \to F(x) \qquad \forall x \in C(F)$$

$$F_n \overset{w}{\to} F' \iff F_n(x) \to F'(x) \qquad \forall x \in C(F')$$

طبق خاصیت یکتایی حد توابع، $\forall\, x \in C(F) \cap C(F')$ داریم $F(x) = F'(x)$.

از طرفی طبق قضیه تجزیه جردن تعداد نقاط انفصال یک تابع توزیع شماراست. اما مجموعه‌ی نقاط $C(F) \cap C(F')$ یک مجموعه پیوسته است پس شامل تمام اعضای R به غیر از تعداد شمارایی از نقاط آن هست که یک مجموعه چگال در R است. با توجه به نتیجه لم ۴-۱، F و F' همه جا بر روی R با یکدیگر برابرند.

لم ۸-۲: دنباله‌ی F_n از توابع توزیع به‌طور ضعیف به F همگراست اگر و تنها اگر بر مجموعه چگال D در R همگرا باشد.

اثبات: ($\Leftarrow$) ابتدا نشان می‌دهیم که اگر $F_n \overset{w}{\to} F$ آنگاه بر یک مجموعه چگال D در R همگراست.

$$F_n \overset{w}{\to} F \iff F_n(x) \to F(x) \qquad \forall\, x \in C(F)$$

چون $C(F)$ یک مجموعه چگال در R است با انتخاب $D = C(F)$ حکم ثابت می‌شود.

حال برای اثبات ($\Rightarrow$) نشان می‌دهیم که اگر F_n بر مجموعه چگال D در R همگرا باشد آنگاه F_n به طور ضعیف همگراست.

فرض کنیم $F_D(x)$ حد $F_n(x)$ برای هر $x \in D$ باشد. $F_D(x)$ تابعی غیر نزولی است زیرا بر طبق خاصیت غیر نزولی بودن توابع توزیع داریم

$$x_1 \le x_2 \Longrightarrow F_n(x_1) \le F_n(x_2)$$

$$\Longrightarrow \lim_{n\to\infty} F_n(x_1) \le \lim_{n\to\infty} F_n(x_2)$$

$$\Rightarrow F_D(x_1) \le F_D(x_2)$$

همچنین چون دنباله توابع توزیع F_n ها متناهی هستند بنابراین حد آن یعنی F_D نیز متناهی است. بنابراین F_D تابعی غیر نزولی و متناهی است و بر اساس لم ۴-۱ اگر تعریف کنیم

$$F(x) = \lim_{x^*\downarrow x} F_D(x^*) = \lim_{x^*\uparrow x} F_D(x^*) \qquad \forall x \in C(F), \quad \forall x^* \in D$$

$F(x)$ یک تابع توزیع است و اکنون کافیست نشان دهیم که $F_n(x)$ به طور ضعیف به $F(x)$ همگراست. به ازای هر $x' < x < x''$، $x'', x' \in D$ به دلیل غیر نزولی بودن F_n داریم

$$F_n(x') \le F_n(x) \le F_n(x'') \ \Rightarrow F_n(x') \le \inf_{x'<x<x''} F_n(x) \le \sup_{x'<x<x''} F_n(x) \le F_n(x'')$$

با میل دادن $n \to \infty$ خواهیم داشت:

$$F_D(x') \le \lim_{n\to\infty} \inf F_n(x) \le \lim_{n\to\infty} \sup F_n(x) \le F_D(x'')$$

$$\Rightarrow F_D(x') \le \lim_{n\to\infty} F_n(x) \le F_D(x'')$$

اگر x یک نقطه پیوستگی F باشد و در D قرار بگیرد و چون $x'' \downarrow x$ و $x' \uparrow x$ خواهیم داشت:

$$\lim_{x'\uparrow x} F_D(x') = F(x)$$

$$\lim_{x''\downarrow x} F_D(x'') = F(x)$$

با توجه به دو رابطه بالا و قضیه فشردگی(ساندویچ) خواهیم داشت:

$$\lim_{n\to\infty} F_n(x) = F(x)$$

اثبات لم به روش دیگر:

دنباله $\{x_k\}$ و $\{y_k\}$ در D وجود دارد که وقتی $k \to \infty$ ، $y_k \downarrow x$, $x_k \uparrow x$ بنابراین برای $k, n \in N$ داریم:

$$x_k < x < y_k$$

چون F یک تابع نانزولی است

$$F_n(x_k) < F_n(x) < F_n(y_k)$$

چون $\{x_k\}$ و $\{y_k\}$ دو دنباله در D هستند پس طبق فرض به x همگرا هستند.

$$F(x_k) = \lim_{n\to\infty} F_n(x_k)$$

$$F(y_k) = \lim_{n\to\infty} F_n(y_k)$$

می‌دانیم

$$\lim_{n\to\infty} \inf F_n(x) \le \lim_{n\to\infty} F_n(x) \le \lim_{n\to\infty} \sup F_n(x)$$

از روابط بالا داریم:

$$F(x_k) \le \lim_{n\to\infty} F_n(x) \le F(y_k)$$

$$\lim_{k\to\infty} F_n(x_k) \le \lim_{n\to\infty} F_n(x) \le \lim_{k\to\infty} F_n(y_k)$$

از طرفی $x \in C(F)$ و $x_k \uparrow x\,, y_k \downarrow x$ داریم:

$$\lim_{k\to\infty} F(x_k) \le F(x) \le \lim_{k\to\infty} F(y_k)$$

پس

$$\lim_{n\to\infty} F_n(x) = F(x)$$

قضیه ۸-۱: (قضیه همگرایی هلی یا فشردگی ضعیف)؛ اگر F_n دنباله‌ای نامتناهی از توابع توزیع باشد آنگاه زیر دنباله‌ای از آن موجود است که به طور ضعیف به F همگراست.

اثبات: بر اساس لم ۸-۲ کافیست ثابت کنیم زیر دنباله ای از F_n بر مجموعه چگال D در R همگراست. مجموعه $D = \{x_n\}$ را یک مجموعه جداساز در R در نظر می گیریم. چون F_n ها کراندار هستند پس دنباله $F_n(x_1)$ یک دنباله نامتناهی و کراندار است و زیر دنباله‌ای از آن مانند $F_{n1}(x_1)$ وجود دارد که کراندار است و می‌دانیم که هر دنباله کراندار در R^K دارای زیر دنباله همگراست. دنباله $F_{n1}(x_2)$ نیز دنباله‌ای نامتناهی و کراندار است و بنابراین زیر دنباله‌ای مانند $F_{n2}(x_2)$ وجود دارد که کراندار است. این دنباله در $x = x_1, x_2$ نیز همگراست. و به همین شکل زیر دنباله $F_{nk}(x_k)$ در $x = x_1, x_2, \dots, x_k$ همگراست.

$$F_{11}(x_1), F_{21}(x_1), \dots, F_{n1}(x_1) \longrightarrow F_1(x_1)$$

$$F_{12}(x_2), F_{22}(x_2), \dots, F_{n2}(x_2) \longrightarrow F_2(x_2)$$

$$\vdots$$

$$F_{1n}(x_n), F_{2n}(x_n), \dots, F_{nn}(x_n) \longrightarrow F_n(x_n)$$

دنباله قطری F_{nn} بر D همگراست و بنابراین $F_n \xrightarrow{w} F$.

مثال ۸-۱: دنباله $F_n(x)$ را به صورت زیر داریم

$$F_n(x) = \begin{cases} 0 & x < 0 \\ nx & 0 \le x < \dfrac{1}{n} \\ 1 & x \ge \dfrac{1}{n} \end{cases}$$

در نظر بگیرید. حد تابع $F_n(x)$ برابر است با:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0 \end{cases}$$

می‌دانیم که $F(x)$ در نقطه صفر پیوسته نیست. بنابراین $C(F) = R - \{0\}$ برای هر $x \neq 0$ ، $F_n \xrightarrow{w} F$.

مثال ۸-۲: اگر $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل از هم با توزیع $U(0,\theta)$ باشند، اگر $X_{(n)} = Y_n$ بزرگترین آماره‌ی ترتیبی باشد آنگاه:

$$F_{Y_n}(y) = F_n(y) = (F_X(y))^n = (\frac{y}{\theta})^n$$

بنابراین

$$F_n(y) = \begin{cases} 0 & y \le 0 \\ (\frac{y}{\theta})^n & 0 < y < \theta \\ 1 & y \ge \theta \end{cases}$$

در نتیجه:

$$\lim_{n\to\infty} F_n\,(y) = \begin{cases} 0 & y < \theta \\ 1 & y \ge \theta \end{cases}$$

می‌بینیم که $F(x)$ در نقطه θ پیوسته نیست. بنابراین $C(F) = R - \{\theta\}$ پس برای هر $x \neq \theta$ ، $F_n \xrightarrow{w} F$.

لم ۸-۳: (**هلی – برای**)؛ اگر $F_n \xrightarrow{w} F$ ، برای $a, b \in C(F)$ ، به طوری که $a < b$ است و g یک تابع کراندار بر $[a,b]$ و پیوسته بر R می‌باشد آنگاه داریم:

$$\int_a^b g\, dF_n \to \int_a^b g\, dF$$

اثبات: اگر $\{x_1, x_2, \dots, x_m\}$ را یک افراز از $[a,b]$ در نظر بگیریم و g_m را به صورت تابع ساده زیر تعریف کنیم

$$g_m = \sum_{i=1}^{m-1} g(x_i) I[x_i, x_{i+1}]$$

چون g_m بر $[a,b]$ پیوسته و کراندار است زمانی که $m \to \infty$ داریم

$$g_m(x) \to g(x)$$

بنابر خواص انتگرال خواهیم داشت:

$$\int_a^b g_m dF_n \to \int_a^b g\, dF_n \;\; \Longrightarrow \left|\int_a^b g_m dF_n - \int_a^b g\, dF_n\right| \longrightarrow 0 \qquad (۱\text{-}۸)$$

$$\int_a^b g_m dF \to \int_a^b g\, dF \quad \Longrightarrow \left|\int_a^b g_m dF - \int_a^b g\, dF\right| \longrightarrow 0$$

از اینکه $F_n \xrightarrow{w} F$ و $n \to \infty$ داریم

$$F_n(x_{i+1}) - F_n(x_i) \to F(x_{i+1}) - F(x_i)$$

از این رو خواهیم داشت:

$$\int_a^b g_m\, dF_n = \sum_{i=1}^{m-1} g(x_i)\left(F_n(x_{i+1}) - F_n(x_i)\right) = \sum_{i=1}^{m-1} g(x_i)\left(F(x_{i+1}) - F(x_i)\right)$$

(۲-۸)

$$= \int_a^b g_m\, dF$$

بنابراین داریم:

$$\left|\int_a^b g_m\; dF_n - \int_a^b g_m\; dF\right| \longrightarrow 0$$

اکنون داریم:

$$\left|\int_a^b g\, dF_n - \int_a^b g\; dF\right|$$

$$= \left|\int_a^b g\; dF_n - \int_a^b g_m\; dF_n + \int_a^b g_m\, dF_n - \int_a^b g_m\, dF + \int_a^b g_m\, dF\; - \int_a^b g\, dF\right|$$

$$\leq \left|\int_a^b g\, dF_n - \int_a^b g_m\, dF_n\right| + \left|\int_a^b g_m\, dF_n - \int_a^b g_m\, dF\right| + \left|\int_a^b g_m\, dF - \int_a^b g\, dF\right|$$

با استفاده از روابط (۸-۱) و (۸-۲) داریم

$$\left|\int_a^b g\, dF_n - \int_a^b g\; dF\right| \to 0 \Longrightarrow \int_a^b g\, dF_n \longrightarrow \int_a^b g\, dF$$

تعمیم لم ۸-۳: اگر $F_n \overset{w}{\rightarrow} F$ ، برای تابع کراندار و پیوسته‌ی g خواهیم داشت:

$$\int_{-\infty}^{+\infty} g\,dF_n \to \int_{-\infty}^{+\infty} g\,dF$$

اثبات: چون g کراندار است بنابراین $|g| \le c$ ، با استفاده از لم ۸-۳ برای $a,b \in C(F)$ داریم:

$$\left|\int_{-\infty}^{+\infty} g\,dF_n - \int_{-\infty}^{+\infty} g\,dF\right|$$

$$= \left|\int_{-\infty}^{+\infty} g\,dF_n - \int_a^b g\,dF_n + \int_a^b g\,dF_n - \int_a^b g\,dF + \int_a^b g\,dF - \int_{-\infty}^{+\infty} g\,dF\right|$$

$$\le \left|\int_{-\infty}^{+\infty} g\,dF_n - \int_a^b g\,dF_n\right| + \left|\int_a^b g\,dF_n - \int_a^b g\,dF\right| + \left|\int_{-\infty}^{+\infty} g\,dF - \int_a^b g\,dF\right|$$

$$= I + II + III$$

بنابراین

$$I = \left|\int_{-\infty}^{+\infty} g\,dF_n - \int_a^b g\,dF_n\right| = \left|\int_{-\infty}^{a} g\,dF_n + \int_b^{+\infty} g\,dF_n\right|$$

$$\le \left|\int_{-\infty}^{a} g\,dF_n\right| + \left|\int_b^{\infty} g\,dF_n\right|$$

$$\le \int_{-\infty}^{a} |g|\,dF_n + \int_b^{+\infty} |g|\,dF_n \le \int_{-\infty}^{a} c\,\,dF_n + \int_b^{+\infty} c\,\,dF_n$$

$$= c\big(F_n(a) - F_n(-\infty)\big) + c\big(F_n(\infty) - F_n(b)\big)$$

$$= c\big(F_n(a) - F_n(-\infty) + F_n(\infty) - F_n(b)\big)$$

وقتی که $a \to -\infty$، $b \to \infty$، مقدار I به سمت صفر میل می‌کند. به همین ترتیب برای III نیز ثابت می‌شود که وقتی $a \to -\infty$ ، $b \to \infty$ ، مقدارآن به سمت صفر میل می‌کند. مقدار II نیز با استفاده از لم ۸-۳ به سمت صفر میل می‌کند.

از این رو $\left|\int_{-\infty}^{+\infty} g\,dF_n - \int_{-\infty}^{+\infty} g\,dF\right| \to 0$ و خواهیم داشت:

$$\int_{-\infty}^{+\infty} g dF_n \to \int_{-\infty}^{+\infty} g dF$$

حالت کلی لم هلی- برای: اگر p_n و ... ,p_1, p_2 اندازه‌های احتمال روی مجموعه‌ی S باشند، آنگاه:

$$p_n \overset{w}{\rightarrow} p \quad \Leftrightarrow \quad \int g dp_n \rightarrow \int g dp$$

که g تابع پیوسته و کراندار روی S است.

لم ۸-۴: اگر $\lim_n \int g dp_n = \int g dp$ برای همه g ها بطور یکنواخت کراندار و پیوسته باشد داریم

$$p_n \overset{w}{\rightarrow} p$$

اثبات: اگر $P(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$ و $P_n(A) = \begin{cases} 1 & x_n \in A \\ 0 & x_n \notin A \end{cases}$ و $x_n \rightarrow x$ آنگاه:

$$\int g dp_n = g(x_n) \rightarrow g(x) = \int g dp \quad \Longrightarrow \quad p_n \overset{w}{\rightarrow} p$$

که در آن روابط فوق با استفاده از روش انتگرال گیری جز به جز به آسانی بدست می آید و اگر $x_n \nrightarrow x$ می‌توان ثابت کرد که رابطه فوق برقرار نیست .

تعریف۸-۲: (**همگرایی کامل**): دنباله توابع F_n را به طور کامل به F همگرا گوییم اگر

$$F_n \overset{w}{\rightarrow} F \qquad \text{و} \qquad VarF_n \rightarrow VarF$$

که در آن $VarF = F(+\infty) - F(-\infty)$، این نوع همگرایی را به صورت $F_n \overset{c}{\rightarrow} F$ نشان می‌دهند. در واقع

$$VarF_n \rightarrow VarF \quad \equiv \quad F_n(\infty) \rightarrow F(\infty) \ \text{و} \ F_n(-\infty) \rightarrow F(-\infty)$$

مثال ۸-۳: درمثال ۸-۱ ثابت کردیم که همگرایی ضعیف برقرار است برای بررسی همگرایی کامل داریم

$$F_n(-\infty) = 0 \rightarrow F(-\infty) = 0$$

$$F_n(+\infty) = 1 \rightarrow F(+\infty) = 1$$

پس نتیجه می‌گیریم که $VarF_n \rightarrow VarF$. پس به طور کامل همگراست.

با بیان مثالی نشان می‌دهیم که همگرایی کامل قوی‌تر از همگرایی ضعیف است.

مثال۸-۴: فرض کنید F_0 تابع توزیع یک متغیر تصادفی مانند X باشد که در نقطه صفر تباهیده است. یعنی

$$F_0 = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

تعریف می‌کنیم $F_n(x) = F_0(x+n)$ ، $\forall x \in R$ در اینصورت $\forall x$ زمانیکه $n \to \infty$ داریم:

$F_n(x) \to F_0(\infty) = 1$، پس به ازای تمام مقادیر x تعریف می‌کنیم $F(x) \equiv 1$ ، پس $F(x)$ در تمام نقاط پیوسته است و به ازای تمام مقادیر x ، $F_n \xrightarrow{W} F$ ، ولی

$$F(\infty) = \lim_{x\to\infty} F(x) = 1 = \lim_{x\to-\infty} F(x) = F(-\infty)$$

$$\lim_{n\to\infty} F_n(-\infty) = \lim_{n\to\infty} F_0(-\infty) = 0$$

پس $F_n \nrightarrow^{c} \mathrm{F}$.

قضیه ۸-۲: (**قضیه هلی- برای**)؛ اگر g تابعی پیوسته و کراندار در R باشد آنگاه

$$F_n \xrightarrow{C} F \qquad \Leftrightarrow \qquad \int_{-\infty}^{+\infty} g dF_n \to \int_{-\infty}^{+\infty} g dF$$

اثبات: چون $F_n \xrightarrow{C} F$ بنا بر تعریف $F_n \xrightarrow{w} F$ و بر اساس تعمیم لم هلی- برای

$$\int_{-\infty}^{+\infty} g dF_n \to \int_{-\infty}^{+\infty} g dF$$

برعکس فرض کنیم $\int_{-\infty}^{+\infty} g dF_n \to \int_{-\infty}^{+\infty} g dF$، چون g تابعی پیوسته و کراندار است داریم

$$F_n(x) \to F(x) \qquad \forall x \in C(F)$$

و این یعنی $F_n \xrightarrow{w} F$. اکنون باید نشان دهیم $VarF_n \to VarF$.

تابع $g(x) = 1$ را که تابعی پیوسته و کراندار است در نظر می‌گیریم خواهیم داشت:

$$\int_{-\infty}^{+\infty} dF_n \longrightarrow \int_{-\infty}^{+\infty} dF \Longrightarrow F_n(\infty) - F_n(-\infty) \to F(\infty) - F(-\infty) \Longrightarrow VarF_n \to VarF$$

و بنابراین $F_n \xrightarrow{c} F$.

تعریف همگرایی ضعیف با استفاده از اندازه: همگرایی ضعیف عبارت است از همگرایی دنباله اندازه‌های P_n به P که P_n به وسیله F_n بر R القا شده است یعنی اگر $B = (a, b]$ داریم:

$$P_n(B) = \big(F_n(b) - F_n(a)\big) \longrightarrow \big(F(b) - F(a)\big) \quad \Longrightarrow \qquad P_n(B) \longrightarrow P(B)$$

۸–۲ قضیه ی شفه برای توابع توزیع با استفاده از چگالی ها

اگر $f_n(x)$ تابع چگالی $F_n(x)$ باشد آنگاه $f_n(x) \geq 0$ و $\int_{-\infty}^{+\infty} f_n(x)\mathrm{dx} = \mathrm{c}$ و اگر برای $x \in R$ ، $f_n(x) \to f(\,x)$، که در آن $\int_{-\infty}^{+\infty} f_n(x)\mathrm{dx} = \mathrm{c}$، آنگاه گوییم $F_n(x) \xrightarrow{c} F(x)$. این قضیه با توجه به قضیه زیر متعلق به شفه اثبات می‌شود.

قضیه ۸–۳: فرض کنید $f_n(x)$ توابع اندازه‌پذیر نامنفی باشند و $f_n(x) \to f(x)$، اگر داشته باشیم: $\int_{-\infty}^{+\infty} f_n(x)dx = \int_{-\infty}^{+\infty} f(x)dx = c$ آنگاه به‌ازای تمام x ها ، $F_n(x) \xrightarrow{c} F(x)$ و رابطه زیر را خواهیم داشت:

$$\sup_{B\in\beta}\left|\int_B f_n(t)dt - \int_B f(t)dt\right| = \frac{1}{2}\int_{-\infty}^{+\infty}|f_n(t) - f(t)|\mathrm{dt} \to 0$$

اثبات: Bرا مجموعه $(-\infty, x]$ اختیار می‌کنیم. چون

$$|F_n(x) - F(x)| = \left|\int_B f_n(t)dt - \int_B f(t)dt\right|$$

$$\leq \sup_{B\in\beta}\left|\int_B f_n(t)dt - \int_B f(t)dt\right|$$

کافیست نشان دهیم:

$$\sup_{B\in\beta}\left|\int_B f_n(t)dt - \int_B f(t)dt\right| \to 0$$

$$\int_{-\infty}^{+\infty} f_n(x)dx - \int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty}\big(f_n(x) - f(x)\big)dx$$

$$= \int_B \big(f_n(t) - f(t)\big)dt + \int_{B^c}\big(f_n(t) - f(t)\big)dt = c - c = 0$$

$$\Longrightarrow \int_B \big(f_n(t) - f(t)\big)dt = -\int_{B^c}\big(f_n(t) - f(t)\big)dt$$

بنابراین

$$2\left|\int_{B}\ \big(f_n(t)-f(t)\big)dt\right|$$

$$=\left|\int_{B}\ \big(f_n(t)-f(t)\big)dt\right|+\left|\int_{B^c}\big(f_n(t)-f(t)\big)dt\right|$$

$$=\left|\int_{-\infty}^{+\infty}\big(f_n(t)-f(t)\big)dt\right|\le\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt$$

$$\Rightarrow\left|\int_{B}\ \big(f_n(t)-f(t)\big)dt\right|\le\frac{1}{2}\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt$$

چون این رابطه به ازای هر B برقرار است بنابراین

$$\sup_{B\in\beta}\left|\int_{B}\ \big(f_n(t)-f(t)\big)dt\right|\ \le\frac{1}{2}\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt\qquad(8-3)$$

اکنون اگر $B_0=\{f_n\ge f\}$ داریم

$$\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt=\int_{B_0}(f_n(t)-f(t))dt-\int_{B_0^c}(f_n(t)-f(t))dt$$

$$=2\int_{B_0}(f_n(t)-f(t))dt$$

چون برای مجموعه B_0 رابطه (۸-۳)، به صورت تساوی می‌باشد. بنابراین برای هر مجموعه B خواهیم داشت:

$$\sup_{B\in\beta}\left|\int_{B}\ \big(f_n(t)-f(t)\big)dt\right|=\frac{1}{2}\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt$$

حال کافیست نشان دهیم $\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt\to 0$

$$\int_{-\infty}^{+\infty}\left|\big(f_n(t)-f(t)\big)\right|dt=\int_{B_0}(f_n(t)-f(t))dt\ -\int_{B_0^c}(f_n(t)-f(t))dt$$

$$=2\int_{B_0}(f_n(t)-f(t))dt$$

و چون $f_n(x)\to f(x)$ خواهیم داشت $\int_{B_0}(f_n(t)-f(t))dt\to 0$ بنابراین

$$\int_{-\infty}^{+\infty}|f_n(t)-f(t)|\mathrm{dt}\to 0$$

و حکم ثابت شد.

مثال ۸-۵: اگر $X\sim t(n)$ باشد زمانیکه $n \to \infty$ تابع‌چگالی به سمت توزیع نرمال استاندارد میل می‌کند.

نتیجه ۸-۱: قضیه شفه برای توابع احتمال نیز برقرار است به این صورت که اگر

$$P_n(x) \to P(x) \text{ و } P_n(x) \geq 0$$

$$\sum_x P_n(x) = \sum_x P(x) = c$$

آنگاه

$$\sum_{i\leq x} P_n(i) \to \sum_{i\leq x} P(i)$$

مثال ۸-۶:

$$p\left(X_n = -\frac{2n+3}{\mathrm{n}}\right) = \frac{4\mathrm{n}-3}{12\mathrm{n}} \to p(X=-2) = \frac{1}{3}$$

$$p\left(X_n = -\frac{3n+1}{\mathrm{n}}\right) = \frac{8\mathrm{n}+3}{12\mathrm{n}} \to p(X=3) = \frac{2}{3}$$

طبق نتیجه ۸-۱ چون $P_n(i) \to P(i)$ و $\sum P_n(i) = \sum P(i) = 1$ پس $F_n \xrightarrow{c} F$.

۸-۳ همگرایی توابع توزیع و توابع مشخصه

تعریف ۸-۳: (**انتگرال تابع مشخصه**)؛ انتگرال تابع مشخصه F ، $\hat{\phi}$، به صورت زیر تعریف می‌شود

$$\hat{\phi}(u) = \int_0^u \phi(v)dv = \int_0^u \int_{-\infty}^{+\infty} e^{ivx}\, dF(x)dv$$

چون $|\phi(v)| \leq 1$ بنابراین $\hat{\phi}(u)$ یک تابع کراندار است پس می‌توان ترتیب انتگرال‌گیری را عوض کرد.

$$\hat{\phi}(u) = \int_0^u \phi(v)dv = \int_{-\infty}^{+\infty} \int_0^u e^{ivx} dv dF(x)$$

و چون $\int_0^u e^{ivx}\, dv = \frac{e^{iux}-1}{ix}$، از این رو

$$\hat{\phi}(u) = \int_{-\infty}^{+\infty} \frac{e^{iux} - 1}{ix} dF(x)$$

حال نشان می‌دهیم که یک تناظر یک به یک بین $\phi(u)$ و $\hat{\phi}(u)$ وجود دارد. فرض کنیم دو تابع مشخصه $\phi_1(u)$ و $\phi_2(u)$ وجود دارند

$$\hat{\phi}(u) = \int_0^u \phi_1(u) dv,$$

$$\hat{\phi}(u) = \int_0^u \phi_2(u) dv \ \Longrightarrow \int_0^u \phi_1(u) dv \ = \ \int_0^u \phi_2(u) dv$$

با مشتق‌گیری نسبت به u از رابطه فوق خواهیم داشت:

$$\phi_1(u) = \phi_2(u)$$

از طرفی چون بین F و ϕ یک رابطه‌ی یک به یک وجود دارد. بنابراین F و $\hat{\phi}$ یکدیگر را بطور یکتا مشخص می‌کند.

قضیه زیر روشی برای مشخص‌سازی همگرایی ضعیف است.

قضیه ۸-۴: اگر $F_n \overset{w}{\to} F$ آنگاه $\hat{\phi}_n \longrightarrow \hat{\phi}$ ، بر عکس اگر $\hat{\phi}_n \longrightarrow \hat{\psi}$ آنگاه یک تابع توزیع مانند F وجود دارد به طوری که $F_n \overset{w}{\to} F$ و $\hat{\phi}_F = \hat{\psi}$.

اثبات: تابع $g(x) = \frac{e^{iux}-1}{ix} = \frac{\cos(ux)+i\sin(ux)-1}{ix}$ تابعی کراندار است زیرا

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} \frac{cos(\,ux) + i\,sin(ux) - 1}{ix}$$

با استفاده از قاعده هوپیتال خواهیم داشت:

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} \frac{-u\sin(ux) + iu\cos(ux)}{i} = \frac{iu}{\mathrm{i}} = u$$

همچنین

$$\lim_{x\to\pm\infty} g(x) = \lim_{x\to\pm\infty} \frac{cos(\,ux) + i\,sin(ux) - 1}{ix}$$

چون صورت کسر مقداری متناهی است بنابراین حد فوق برابر صفر خواهد شد. از این رو $g(x)$ تابعی کراندار است.

با استفاده از تعمیم لم (هلی- برای) داریم:

$$\int_R \frac{e^{iux}-1}{ix}dF_n(x) \to \int_R \frac{e^{iux}-1}{ix}dF(x)$$

$$\Longrightarrow \hat{\phi}_n(u) \longrightarrow \hat{\phi}(u) \qquad (۴-۸)$$

برعکس فرض کنید $\hat{\phi}_n \longrightarrow \hat{\psi}$ و F_n دنباله توابع متناظر با $\hat{\phi}_n$ باشد بنابر قضیه همگرایی هلی زیر دنباله‌ای از F_n ، مانند $F_{n'}$ وجود دارد که $F_{n'} \xrightarrow{w} F$ بنابراین طبق رابطه (۸ – ۴) داریم $\hat{\phi}_{n'}(u) \longrightarrow \hat{\phi}_F(u)$.

از طرفی $\hat{\phi}_n(u) \longrightarrow \hat{\psi}(u)$ می‌دانیم هر زیر دنباله‌ی یک دنباله همگرا، همگراست. بنابراین $\hat{\phi}_F(u) = \hat{\psi}(u)$ همچنین فرض کنیم زیر دنباله دیگری مانند $F_{n''}$ موجود باشد که $F_{n''} \xrightarrow{w} \bar{F}$ طبق آنچه گفته شد داریم

$$\hat{\phi}_{\bar{F}}(u) = \hat{\psi}(u) \text{ و } \hat{\phi}_{n''}(u) \longrightarrow \hat{\phi}_{\bar{F}}(u)$$

چون $\hat{\psi}$ توزیع F را به طور یکتا مشخص می‌کند پس $F = \bar{F}$. چون این مطلب برای هر زیر دنباله همگرا از F_n برقرار است از این رو $F_n \xrightarrow{w} F$ با

$$\hat{\phi}_F = \hat{\psi}$$

نتیجه ۸-۲: اگر $\phi_n \to \psi$ آنگاه $F_n \xrightarrow{w} F$ با $\phi_F = \psi$.

اثبات: چون ϕ_n و ψ کراندار و اندازه‌پذیر لبگ هستند طبق نتیجه قضیه همگرایی مغلوب داریم:

$$\phi_n \to \psi \Longrightarrow \int_0^u \phi_n(v)dv = \int_0^u \psi(v)dv \Longrightarrow \hat{\phi}_n \longrightarrow \hat{\psi}$$

چون رابطه بالا وجود دارد ϕ و $\hat{\phi}$ خواهیم داشت واز طرفی رابطه‌ای یک به یک بین $\hat{\phi}_F = \hat{\psi}$ و $F_n \xrightarrow{w} F$ وجود دارد بنابراین طبق قضیه ۸-۴،

$$\phi_F = \psi$$

قضیه ۸-۵:(قضیه پیوستگی لوی)؛ اگر $F_n \xrightarrow{c} F$ آنگاه $\phi_n \to \phi$، به عکس اگر $\phi_n \to \psi$ که در آن ψ در $u = 0$ پیوسته است آنگاه $F_n \xrightarrow{c} F$ و $\phi_F = \psi$.

اثبات: اگر $F_n \xrightarrow{c} F$ آنگاه $F_n \xrightarrow{w} F$ و $VarF_n \to VarF$.

تابع g را به صورت $g(x) = e^{iux}$ در نظر می‌گیریم چون $|g(x)| \leq 1$ و همه جا پیوسته است با استفاده از قضیه (هلی– برای) داریم

$$\int_R e^{iux}\,dF_n(x) \to \int_R e^{iux}\,dF(x) \quad \Longrightarrow \phi_n(u) \longrightarrow \phi(u)$$

برعکس فرض کنید $\phi_n \to \psi$ در این صورت طبق نتیجه ۸-۲ داریم $F_n \overset{w}{\to} F$ و $\phi_F = \psi$ اکنون کافیست نشان دهیم $VarF_n \to VarF$

$$VarF_n = F_n(\infty) - F_n(-\infty) = \int_{-\infty}^{+\infty} d\,F_n = \phi_n(0) \to \psi(0)$$

$$VarF = F(\infty) - F(-\infty) = \int_{-\infty}^{+\infty} d\,F = \phi(0)$$

باید نشان دهیم $\phi(0) = \psi(0)$

چون $\hat{\phi}_n$ و ϕ_n یکدیگر را بطور یکتا مشخص می‌کنند داریم

$$\phi_n \to \psi \Longrightarrow \hat{\phi}_n \to \hat{\psi}$$

پس طبق عکس قضیه ۸-۴ داریم $F_n \overset{w}{\to} F$، حالا طبق رفت قضیه ۸-۴، $\hat{\phi}_n \to \phi$ چون می‌دانیم برای هر $u > 0$،

$$\hat{\psi}(u) = \hat{\phi}(u) \Longrightarrow \int_0^u \psi(v)dv = \int_0^u \phi(v)dv \Longrightarrow \frac{1}{u}\int_0^u \psi(v)dv = \frac{1}{u}\int_0^u \phi(v)dv$$

$$\Longrightarrow \lim_{u\to 0}\frac{1}{u}\int_0^u \psi(v)dv = \lim_{u\to 0}\frac{1}{u}\int_0^u \phi(v)dv$$

با استفاده از روابط هم ارزی در انتگرال و پیوستگی ψ در صفر نتیجه می‌شود که $\phi(0) = \psi(0)$ و بنابراین $F_n \overset{c}{\to} F$.

مثال۸-۷: فرض کنید X_n بر $(-n, n)$ به طور یکنواخت توزیع شده باشد.

$$f_n(x) = \frac{1}{2n}$$

$$F_n(x) = \frac{x+n}{2n}$$

تابع مشخصه توزیع $F_n(x)$ برابر است با

$$\phi_n(u) = \int_R \frac{e^{iux}}{2n} dx = \frac{1}{2n} \frac{e^{inu} - e^{-inu}}{iu} = \frac{1}{2n} \frac{2i\,sin(nu)}{iu} = \frac{sin(nu)}{nu}$$

$$\phi(u) = \lim_{n\to\infty} \phi_n(u) = \begin{cases} 1 & u = 0 \\ 0 & u \neq 0 \end{cases}$$

در حالت $u = 0$ از روابط هم ارزی انتگرال استفاده می‌کنیم. و در حالت $u \neq 0$،

$$\lim_n \left(0 \times \text{کراندار}\right) = 0$$

چون تابع $\phi(u)$ در $u = 0$ پیوسته نیست از این رو دنباله‌ی F_n به تابع توزیع یک متغیر تصادفی همگرا نیست.

۸–۴ همگرایی گشتاورها

لم ۸-۵: فرض کنید $F_n \xrightarrow{c} F\,(X_n \xrightarrow{d} X)$ و به ازای عددی مانند $p > 0$ داشته باشیم

$\sup_n E(|{X_n}^p|) = M$ که M ثابت است. در این صورت برای هر $r < p$ خواهیم داشت

$$E(|X_n|^r) \to E(|X|^r)$$

اثبات: تابع f را به صورت زیر تعریف می‌کنیم

$$f_A(x) = \begin{cases} |x|^r & |x| \leq A \\ A^r & |x| > A \end{cases}$$

تابع $f_A(x)$ بر R کراندار و بر $[-A, A]$ پیوسته است بنابراین طبق لم هلی-برای داریم

$$\int_R f_A(x) dF_n(x) \to \int_R f_A(x)\ dF(x) \tag{۸-۵}$$

$$\int_R |x|^r dF_n(x) \to \int_R |x|^r\ dF(x) \tag{۸-۶}$$

اگر $A \to \infty$،

$$\int_R |f_A(x_n) - |x_n|^r| dF_n(x)$$

$$= \int_{|x_n|\le A} |f_A(x_n) - |x_n|^r|\, dF_n(x) + \int_{|x_n|>A} |f_A(x_n) - |x_n|^r|\, dF_n(x)$$

$$= \int_{|x_n|\le A} ||x_n|^r - |x_n|^r|\, dF_n(x) + \int_{|x_n|>A} |A^r - |x_n|^r|\; dF_n(x)$$

$$= \int_{|x_n|>A} |A^r - |x_n|^r|\, dF_n(x)$$

برای $|x_n| > A$ داریم

$$|A^r - |x_n|^r| \le |A^r| + |x_n|^r \le |x_n|^r + |x_n|^r = 2|x_n|^r$$

در نتیجه

$$\int_{|x_n|>A} |A^r - |x_n|^r|\; dF_n(x) \le 2\int_{|x_n|>A} |x_n|^{\mathrm{r}}\; dF_n(x) \quad = 2E(|x_n|^{\mathrm{r}}\, I_{(|x_n|>A)})$$

از طرفی داریم

$$|x_n| > A \Longrightarrow \frac{|x_n|}{A} > 1 \Longrightarrow (\frac{|x_n|}{A})^{p-r} > 1 \Rightarrow \frac{|x_n|^p}{A^{p-r}} > |x_n|^{\mathrm{r}}$$

از طرفین امید می‌گیریم

$$2E\left(|X_n|^{\mathrm{r}}\, I_{(|x_n|>A)}\right) \le 2E\left(\frac{|X_n|^p}{A^{p-r}}\right) = \frac{2E(|X_n|^p)}{A^{p-r}} \le \frac{2M}{A^{p-r}}$$

از این رو

$$\int_{|x_n|>A} |A^r - |x_n|^r|\, dF_n(x) \le \frac{2M}{A^{p-r}}$$

زمانی که $A \to \infty$ مقدار $\frac{2M}{A^{p-r}} \to 0$ به سمت صفر میل می‌کند. پس $f_{\mathrm{A}}(x_n) \to |x_n|^r$

حال طبق رابطه (۸-۵) و(۸-۶) داریم

$$E(|X|^r) = \lim_{A\to\infty} \int_{-\infty}^{+\infty} f_A(x) dF = \lim_{A\to\infty} \lim_{n\to\infty} \int_{-\infty}^{+\infty} f_A(x)\, dF_n$$

$$= \lim_{n\to\infty} \lim_{A\to\infty} \int_{-\infty}^{+\infty} f_A(x)\, dF_n = \lim_{n\to\infty} E(|X_n|^r)$$

قضیه ۸-۶: اگر F_n دنباله‌ای از توابع توزیع باشد به قسمی که دنباله گشتاورهای متناظر با آن، γ'_{r_n} به γ'_r که به ازای تمام r ها متناهی است، همگرا باشد و γ'_{r_n} تابع توزیع متناظر با γ'_r یعنی F را به طور یکتا معین کند، آنگاه $F_n \xrightarrow{c} F$.

اثبات: بنا بر قضیه همگرایی هلی زیر دنباله F_{n_k} از F_n موجود است به طوری که $F_{n_k} \xrightarrow{w} F$ حال نشان می‌دهیم که $F_{n_k} \xrightarrow{c} F$. پس کافی است ثابت کنیم $VarF_{n_k} \to VarF$

از نامساوی چبیشف داریم

$$VarF_{n_k} = \lim_{A\to\infty}(F_{n_k}(A) - F_{n_k}(-A)) = \lim_{A\to\infty}\int_{-A}^{A} dF_{n_k}(x) \quad \geq \lim_{A\to\infty}(1 - \frac{\gamma'_2(n_k)}{A^2}) = 1$$

همچنین

$$\begin{aligned} VarF &= \lim_{A\to\infty}(F(A) - F(-A)) \\ &= \lim_{A\to\infty}\lim_{k\to\infty}\left(F_{n_k}(A) - F_{n_k}(-A)\right) \\ &= \lim_{k\to\infty}\lim_{A\to\infty}\left(F_{n_k}(A) - F_{n_k}(-A)\right) \\ &= \lim_{k\to\infty} VarF_{n_k} = 1 \end{aligned}$$

پس $F_{n_k} \xrightarrow{c} F$ و از اینکه $\gamma'_{r_n} \to \gamma'_r$ نتیجه می‌شود که $\gamma'_{r_{n_k}} \to \gamma'_r$. چون γ'_r کراندار است از لم(۸-۵)، نتیجه می‌شود که به ازای هر $s < r$ ، $\gamma'_{s_{n_k}} \to \gamma'_s$ و چون γ'_r به طور یکتا F را مشخص می‌کند از این رو هر زیر دنباله همگرا F_n به طور کامل به F همگراست بنابراین $F_n \xrightarrow{c} F$.

همگرایی ضعیف در احتمال:

اگر چه مفهوم همگرایی ضعیف تابع توزیع روی خط حقیقی یا فضای اقلیدسی است ولی مفهوم همگرایی ضعیف روی اندازه احتمال می‌تواند روی همه فضاهای متریک بیان شود و این دلیل ترجیح همگرایی ضعیف روی اندازه‌ی احتمال نسبت به همگرایی ضعیف تابع توزیع است.

تعریف ۸-۴: دنباله $(\mu_n)_{n>1}$ که روی $(R, \mathfrak{B}(R))$ تعریف می‌شود بطور ضعیف به سمت μ همگراست $(\mu_n \xrightarrow{w} \mu)$ اگر $\int f d\mu_n \to \int f d\mu$ برای تمام fهای پیوسته و کراندار روی R.

لم ۸-۶: اگر μ_n بطور ضعیف به μ همگرا باشد و F_n تابع توزیع مرتبط با μ_n و F تابع توزیع مرتبط با μ باشد آنگاه $F_n \xrightarrow{w} F$.

اثبات: x نقطه‌ای از تابع پیوسته F است. تعریف می‌کنیم

$$\psi(t) = \begin{cases} 1 & t < x - \frac{1}{k} \\ -k(t-x) & x - \frac{1}{k} \leq t < x \\ 0 & t \geq x \end{cases}$$

و $\varphi_k(t) = \psi(t-x)$ پس

$$I_{(-\infty, x-\frac{1}{k})} \leq \psi_k \leq I_{(-\infty,x)} \leq \varphi_k \leq I_{(-\infty, x+\frac{1}{k})}$$

همچنین φ_k و ψ_k توابع پیوسته کراندار هستند و داریم

$$F\left(x - \frac{1}{k}\right) \leq \int \psi_k d\mu = \lim_{n\to\infty} \int \psi_k d\mu_n \leq \lim_{n\to\infty} inf\, F(x)$$

و

$$\lim_{n\to\infty} sup\, F_n(x) \leq \lim_{n\to\infty} \int \varphi_k d\mu_n = \int \varphi_k d\mu \leq F(x + \frac{1}{k})$$

تابع F در x پیوسته است و

$$\lim_{n\to\infty} inf\, F_n\,(x) = \lim_{n\to\infty} sup\, F_n\,(x)$$

در نتیجه

$$\lim_{n\to\infty} F_n(x) = F(x)$$

همگرایی ضعیف در فضای متریک:

اگر $(\mu_n)_{n>1}$ و μ اندازه‌های احتمال روی فضای متریک (S,d) باشد و اگر $\int f d\mu_n \to \int f d\mu$ برای همه f های پیوسته و کران‌دار روی S باشد آنگاه (μ_n) بطور ضعیف به μ همگراست.

۸-۵ مسائل

۱- تحقیق کنید آیا دنباله‌های زیر از توابع توزیع بطور کامل به یک تابع توزیع همگرا می‌باشند.

الف)

$$F_n(x) = \begin{cases} 0 & x < \frac{1}{\sqrt{n}} \\ \frac{2}{3}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \leq x < 1 + \frac{1}{\sqrt{n}} \\ 1 & x \geq 1 + \frac{1}{\sqrt{n}} \end{cases}$$

ب)

$$F_n(x) = \begin{cases} 0 & x < \frac{1}{n} \\ \frac{2}{3}\frac{1}{n} & \frac{1}{n} \leq x < 1 + \frac{1}{n} \\ 1 & x \geq 1 + \frac{1}{n} \end{cases}$$

۲- آیا دنباله‌های زیر از متغیرهای تصادفی $\{X_n\}$ در توزیع به یک متغیر تصادفی مانند X همگرا هستند.

الف)

$$X_n \to N(\frac{1}{n} - 1, \frac{1}{n^2})$$

ب)

$$X_{\mathrm{n}} \to C(\frac{1}{n} - 1, \frac{1}{n^2})$$

ج)

$$F_n(x) = \begin{cases} 0 & x < 0 \\ x^n & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

۳- میانه تابع توزیع $F(x)$ مقداری مانند m از x است که به ازای آن نامساوی زیر برقرار باشند،

$$F(m-0) \leq \frac{1}{2} \leq F(m)$$

الف) اگر $F_n \xrightarrow{w} F$ ثابت کنید m_n میانه F_n به m میانه F همگراست .

ب) اگر F مطلقاً پیوسته باشد ثابت کنید $E|X-c|$ وقتی کمترین مقدار را دارد که c برابر میانه F باشد.

۴- اگر $X_n \sim B(N_n, p_n)$ باشد وقتی $n \to \infty$، $N_n \to \infty$ و $p_n \to \infty$ و $p_n N_n \to \lambda$ و $\lambda \in (0, \infty)$

آنگاه $X_n \xrightarrow{d} X$ که X دارای توزیع پواسن با پارامتر λ است.

۵- $X_n \sim P(\lambda_n)$ که در آن $\lambda_n \in (0, \infty)$، اگر $Y_n = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}}$ را تعریف کنیم اگر $\lambda_n \to \infty$ وقتی $n \to \infty$ آنگاه $Y \xrightarrow{d} N(0,1)$

فصل نهم

استقلال

استقلال یکی از فرضیات اساسی در نظریه‌ی احتمال است. نتایج حاصل از پرتاب‌های متوالی یک سکه یا تاس مستقل هستند. به عبارت دیگر رخداد نتیجه اولین پرتاب سکه یا تاس تأثیری در نتیجه پرتاب دوم یا سایر پرتاب‌ها ندارد. کسب نتایج موفقیت‌آمیز در کارت بازی وقتی که کارت‌ها را بدون جایگذاری انتخاب می‌کنیم، از هم مستقل نیستند. چون کارتی که قبلاً انتخاب شده دوباره انتخاب نمی‌شود. فضای احتمال، همان فضای اندازه است، که در آن اندازه‌ی کل فضا برابر با یک است. اما نظریه احتمال صرفاً زیرمجموعه‌ای از نظریه اندازه نیست. وجه تمایز نظریه احتمال و نظریه اندازه، مفهوم استقلال است.

توجه کنید که استقلال ویژگی ذاتی پیشامدها نیست، بلکه به انتخاب اندازه احتمال بستگی دارد. یعنی ممکن است دو پیشامد نسبت به اندازه‌ی احتمال P مستقل باشند ولی نسبت به اندازه‌ی احتمال Q مستقل نباشند.

۹–۱ استقلال

تعریف ۹-۱: اگر (Ω, F, P) یک فضای احتمال و $\{B_1, B_2, \dots, B_n\} \subset F$ مجموعه‌ای متناهی از پیشامدها باشد:

الف) $B_1, B_2, \dots, B_n$ نسبت به P مستقل هستند اگر برای تمام $1 < k \leq n$ داشته باشیم

$$\mathrm{P}\left(\bigcap_{j=1}^{k} B_j\right) = \prod_{j=1}^{k} P(B_j)$$

ب) $B_1, B_2, \dots, B_n$ را دو به دو نسبت به P مستقل می‌گوییم اگر

برای تمام $\mathrm{i} \neq \mathrm{j}$ $\quad P(B_i \cap B_j) = P(B_i)P(B_j)$

شایان ذکر است که مجموعه‌ای از پیشامدها ممکن است که نسبت به یک اندازه احتمال مثل P مستقل باشند ولی نسبت به اندازه احتمال دیگری مثل P' مستقل نباشند.

مثال ۹-۱: اگر

$$\Omega = \{1, 2, 3\}$$

$$F = \{\Phi, \Omega, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}$$

$$P_1 = 0, 1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}$$

$$P_2 = 0, 1, \frac{1}{2}, \frac{1}{2}, 0, 1, \frac{1}{2}, \frac{1}{2}$$

P_1 و P_2 دو اندازه احتمال هستند که مقادیر متناظر با پیشامدها موجود در F برای هر یک از آنها مشخص شده است. واضح است که

$$P_1 = (\{1,2\} \cap \{2,3\}) = P_1(\{2\}) = \frac{1}{3} \neq \frac{2}{3} \times \frac{2}{3}$$

پس دو پیشامد $\{1,2\}$ و $\{2,3\}$ با اندازه احتمال P_1 مستقل نیستند. ولی تحت اندازه‌ی احتمال P_2 داریم

$$P_2 = (\{1,2\} \cap \{2,3\}) = P_1(\{2\}) = \frac{1}{2} = 1 \times \frac{1}{2}$$

بنابراین دو پیشامد مذکور تحت این اندازه احتمال مستقل‌اند.

مثال ۹-۲: فرض کنید X و Y دو متغیر تصادفی مستقل باشند که مقادیر ۱+ و۱- را با احتمال $\frac{1}{2}$ ختیار می‌کنند. اگر $Z = XY$ باشد، داریم

X	-1	1
$P(X=x)$	$1/2$	$1/2$

Z	-1	1
$P(Z=z)$	$1/2$	$1/2$

Y	-1	1
$P(Y=y)$	$1/2$	$1/2$

در این مثال X, Y, Z دو به دو مستقل‌اند ولی توأماً مستقل نیستند. برای بررسی استقلال متغیرهای X, Z داریم

$$P(X = 1, Z = 1) = P(X = 1)P(Z = 1)$$

$$P(X = 1, Z = 1) = P(X = 1, XY = 1) = P(X = 1, Y = 1)$$

$$= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = P(X = 1)P(Z = 1)$$

$$P(X = -1, Z = 1) = P(X = -1)P(Z = 1)$$

$$P(X = -1, Z = 1) = P(X = -1, XY = 1) = P(X = -1, Y = -1)$$

$$= \frac{1}{4} = P(X = -1)P(Z = 1)$$

همچنین

$$P(X = 1, Z = -1) = P(X = 1)P(Z = -1)$$

$$P(X = 1, Z = -1) = P(X = 1, y = -1) = P(X = 1)P(Y = -1) = \frac{1}{4}$$

پس X, Y, Z دو به دو مستقل‌اند. ولی متقابلاً مستقل نیستند، زیرا در سمت چپ رابطه داریم

$$P(\mathrm{X} = 1, y = 1, z = 1) = P(\mathrm{X} = 1)P(Y = 1)P(Z = 1) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

$$P(\mathrm{X} = 1, Y = 1, Z = 1) = P(\mathrm{X} = 1, Y = 1) = \frac{1}{4} \neq \frac{1}{8}$$

قضیه ۹-۱: شرط لازم و کافی برای آنکه دو متغیر تصادفی X و Y مستقل باشند آن است که به ازای هر دو عدد حقیقی u و v داشته باشیم:

$$P(X \in (-\infty, u], Y \in (-\infty, v]) = P(X \in (-\infty, u])P(Y \in (-\infty, v]) \qquad (۹\text{-}۱)$$

کافی است در تعریف ۹-۱، مجموعه‌های بورل را به صورت $(-\infty, t]$ در نظر بگیریم. در حقیقت قضیه ۹-۱ برابری (۹-۱) را با استقلال X و Y هم ارز می‌داند.

اثبات: اگر X و Y مستقل باشند برابری $F_{X,Y} = F_X(u)F_Y(v)$ بدیهی است. برای اثبات طرف دیگر فرض می‌کنیم پیشامدهای A و B به صورت زیر تعریف شده است

$$A = \{(-\infty, x], x \in R\}, \qquad B = \{(-\infty, y], y \in R\}$$

اگر

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \qquad (۹\text{-}۲)$$

برقرار باشد می‌خواهیم ثابت کنیم این برابری به ازای هر $A \in \mathcal{B}(R)$ و هر $B \in \mathcal{B}(R)$ برقرار است.

الف) می‌توان نشان داد که رابطه‌ی (۹-۲)، برای هر بازه‌ی دلخواه برقرار است.

ب) با استفاده از قسمت الف ثابت می‌شود که رابطه‌ی (۹-۲)، برای هر اجتماع متناهی و دو به دو جدا از هم، از اعضا A و B نیز برقرار است. (گردایه اجتماع‌های متناهی از اعضای A یک میدان است، این میدان را با G نشان می‌دهیم.)

ج) فرض کنیم

$$A = X^{-1}(G) \qquad B = Y^{-1}(G)$$

$$F_1 = X^{-1}(\mathcal{B}(R)) \;\; F_2 = Y^{-1}(\mathcal{B}(R))$$

در این صورت A و B میدان و F_1 و F_2 سیگما میدان هستند. بنابراین

$$F_1 = \sigma(A) \qquad F_2 = \sigma(B)$$

آنچه در قسمت (ب) ثابت شد، اثبات استقلال دو میدان $B = Y^{-1}(G)$ و $A = X^{-1}(G)$ است. مجدداً ثابت می‌شود که F_1 و F_2 نیز مستقل‌اند. استقلال F_1و F_2 به معنای استقلال X و Y است.

تعریف ۹-۲: فرض کنید (Ω, F, P) یک فضای احتمال باشد، مجموعه‌ای از پیشامدها

$\{B_\alpha, \alpha \in A\} \subset F$ نسبت به P مستقل‌اند اگر به ازای هر زیرکلاس متناهی داشته باشیم:

$$P\left(\bigcap_{i=1}^{k} B_{\alpha_i}\right) = \prod_{i=1}^{k} P\left(B_{\alpha_i}\right), \quad \{\alpha_1, \dots, \alpha_k\} \subset A, \quad 1 < K < \infty.$$

قضیه ۹-۲: فرض کنید $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و $h_1, \dots, h_n$ توابع اندازه‌پذیر باشند.

آنگاه $h_1(X_1), \dots, h_n(X_n)$ مستقل هستند.

اثبات: فرض کنید $A_1, \dots, A_n$ مجموعه‌های بورل اندازه‌پذیر باشند، داریم

$$P\left(\bigcap_{k=1}^{n} \{h_k(X_k) \in A_k\}\right) = P\left(\bigcap_{k=1}^{n} \{X_k \in h_k^{-1}(A_k)\}\right)$$

$$= \prod_{k=1}^{\boldsymbol{n}} P\left(X_k \in h_k^{-1}(A_k)\right)$$

$$= \prod_{k=1}^{n} P(h_k(X_k) \in A_k)$$

تعریف ۹-۳: فرض کنید (Ω, F, P) فضای احتمال و$\{X_\alpha\!: \alpha \in A\}$ مجموعه‌ای از متغیرهای تصادفی باشد. خانواده‌ای از سیگما میدان‌های $\{\sigma(X_\alpha)\!: \alpha \in A\}$ به صورت زیر تعریف می‌شود

$$\sigma(X_\alpha) = \{X_\alpha^{-1}(B) \colon B \in \mathfrak{B}(R)\}$$

۹–۲ کلاس‌های مستقل

تعریف ۹-۴: دو کلاس C_1 و C_2 را مستقل می‌گوییم هرگاه برای هر $A_1 \in C_1$ و $A_2 \in C_2$ و A_1، A_2 مستقل باشند. کلاس‌های $\{C_1, \dots, C_n\}$ را مستقل (متقابلا مستقل) می‌گوییم اگر برای هر $A_1 \in C_1, \dots, A_n \in C_n$ کلاس $\{A_1, \dots, A_n\}$ شامل پیشامدهای مستقل (متقابلاً مستقل) باشد.

نکته ۹-۱: فرض می‌کنیم $1 \le n_1 < n_2 < \cdots < n_k = n$ و f_1 یک تابع بورل اندازه‌پذیر از n_1 متغیر و f_2 از $n_2 - n_1$ متغیر و به همین ترتیب f_k از $n_k - n_{k-1}$ متغیر باشد.

اگر $\{X_j, 1 \le j \le n\}$ متغیرهای تصادفی مستقل باشند، آنگاه k متغیرتصادفی

$$f_1(X_1, \dots, X_{n_1}), f_2(X_{n_1+1}, \dots, X_{n_2}), \dots, f_k(X_{n_{k-1}+1}, \dots, X_{n_k})$$

نیز مستقل‌اند.

از استقلال A, B می‌توان استقلال $A,\ B^c$ و $A^c,\ B$ و $A^c,\ B^c$ را نیز نتیجه گرفت.

به عنوان مثال برای استقلال A, B^c داریم:

$$A = A \cap \Omega$$
$$= A \cap (B \cup B^C) = (A \cap B) \cup (A \cap B^c) \Rightarrow P(A) = P(A \cap B) + P(A \cap B^C)$$

بنابراین

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)\big(1 - P(B)\big)$$
$$= P(A)P(B^c)$$

نتیجه ۹-۱: پیشامدهای پوچ و حتمی مستقل‌اند.

اثبات: فرض کنیم N پیشامد پوچ و N^c پیشامد حتمی باشد.

$$P(N)P(N^c) = 0 \times 1 = 0$$
$$P(N \cap N^c) = P(\Phi \cap \Omega) = P(\Phi) = 0$$

بنابراین می‌توان گفت که این دو پیشامد از هم مستقل هستند.

به طور مشابه می‌توان ثابت کرد که هر پیشامد دلخواه از پیشامدهای پوچ یا حتمی مستقل است. واضح است که یک پیشامد پوچ از خودش مستقل است، این مطلب برای یک پیشامد حتمی نیز صادق است.

می‌دانیم پیشامدهای پوچ و حتمی از هر پیشامد دلخواه مستقل هستند. پس اگر به دو کلاس مستقل Φ و Ω دو پیشامد مستقل A و B را اضافه کنیم آن دو کلاس مورد نظر یعنی $\{A, \Phi, \Omega\}$ و $\{B, \Phi, \Omega\}$ نیز مستقل‌اند. واضح است که دو کلاس $\{A, \Phi, \Omega, \mathrm{A^c}\}$ و $\{B, \Phi, \Omega, \mathrm{B^c}\}$ نیز مستقل‌اند.

قضیه ۹-۳: اگر کلاس‌های مستقل $\{C_t : t \in T\}$ تحت اشتراک متناهی بسته باشند، سیگما میدان‌های مینیمال بر این کلاس‌ها مستقل‌اند.

اثبات: اگر به کلاس‌های مستقل Φ و Ω مجموعه‌های معینی را اضافه کنیم بر استقلال آن‌ها تأثیری ندارد. مثل اضافه کردن A, B^c، A^c, B و A^c, B^c. در ادامه نشان می‌دهیم اگر بر این کلاس‌ها اجتماع شمارایی از مجموعه‌ها را اضافه کنیم بر استقلال آن‌ها تأثیر گذار نیست.

در واقع به بررسی این موضوع می‌پردازیم که اگر به جای مجموعه‌ی $A_{t_{1i}}$، مجموعه‌ی $\sum_{i=1}^{\infty} A_{t_{1i}}$ را قرار دهیم استقلال پیشامدها تغییری می‌کند یا خیر.

$$P\left(\left(\bigcup_{i=1}^{\infty} A_{t_{1i}}\right) A_{t_2} \dots A_{t_n}\right)$$

$$= P\left(\left(\sum_{i=1}^{\infty} A_{t_{1i}}\right) A_{t_2} \dots A_{t_n}\right)$$

$$= P\left(\left(A_{t_{11}} + A_{t_{12}} + \cdots\right) \times A_{t_2} \times \cdots \times A_{t_n}\right)$$

$$= P\left(\left(A_{t_{11}} A_{t_2} \dots A_{t_n}\right) + \left(A_{t_{12}} A_{t_2} \dots A_{t_n}\right) + \cdots\right)$$

$$= \sum_{i=1}^{\infty} P\left(A_{t_{1i}} A_{t_2} \dots A_{t_n}\right)$$

$$= \sum_{i=1}^{\infty} P\left(A_{t_{1i}}\right) P\left(A_{t_2}\right) \cdots P\left(A_{t_n}\right)$$

$$= \sum_{i=1}^{\infty} \left(P\left(A_{t_{1i}}\right)\right) P\left(A_{t_2}\right) \cdots P(A_{t_n})$$

همچنین داریم

اگر $A_{t_1}, A'_{t_1} \in C_{t_1}$ و $A_{t_n} \in C_{t_n}$ به قسمی که $A'_{t_1} \subset A_{t_1}$ چون $\{A'_{t_1}, \dots, A_{t_n}\}$ و $\{A_{t_1}, \dots, A_{t_n}\}$ کلاسهای تواما مستقل‌اند.

$$\begin{aligned} P\left(\left(A_{t_1} - A'_{t_1}\right)A_{t_2} \dots A_{t_n}\right) &= P\left(A_{t_1}, \dots, A_{t_n}\right) - P\left(A'_{t_1}, \dots, A_{t_n}\right) \\ &= P\left(A_{t_1}\right) \cdots P\left(A_{t_n}\right) - P\left(A'_{t_1}\right) \cdots P\left(A_{t_n}\right) \\ &= \left(P\left(A_{t_1}\right) - P\left(A'_{t_1}\right)\right)\left(P\left(A_{t_2}\right) \cdots P\left(A_{t_n}\right)\right) \\ &= P\left(A_{t_1} - A'_{t_1}\right)P\left(A_{t_2}\right) \cdots P\left(A_{t_n}\right). \end{aligned}$$

قضیه ۹-۴: فرض کنید $s_t, u_t \in R$ برای $t = t_1, \dots t_n \in T$ مجموعه های بورل هستند که در آن t_i ها دلخواه و n عدد طبیعی متناهی است. اگر متغیرهای تصادفی $t \in T,\ X_t$ تواما مستقل باشند تعاریف زیر معادلند:

الف) $P\left(\bigcap_{k=1}^{n}\left[x_{t_k} \in s_{t_k}\right]\right) = \prod_{k=1}^{n} P\left(x_{t_k} \in s_{t_k}\right)$

ب) $F_{x_{t_1} \dots x_{t_n}}\left(x_{t_1}, \dots, x_{t_n}\right) = F_{x_{t_1}}\left(x_{t_1}\right) \cdots F_{x_{t_n}}\left(x_{t_n}\right)$

ج) $\phi_{x_{t_1}, \dots, x_{t_n}}\left(u_{t_1}, \dots, u_{t_n}\right) = \phi_{x_{t_1}}\left(u_{t_1}\right) \dots \phi_{x_{t_n}}\left(u_{t_n}\right)$

اثبات: الف ← ب:

فرض کنیم $s_{t_k} = (-\infty, x_{t_k}]$ باشد بنابراین:

$$\begin{aligned} P\left(\bigcap_{k=1}^{n}(X_{i_k} \in s_{t_k})\right) &= P\left(X_{t_1} \in \left(-\infty, x_{t_1}\right], \dots, X_{t_n} \in \left(-\infty, x_{t_n}\right]\right) \\ &= P\left(X_{t_1} \in \left(-\infty, x_{t_1}\right]\right) \dots P\left(X_{t_n} \in \left(-\infty, X_{t_n}\right]\right) \\ &\Rightarrow P\left(X_{t_1} \le x_{t_1}, \dots, X_{t_n} \le x_{t_n}\right) \end{aligned}$$

$$\Rightarrow F_{X_{t_1}, \dots, X_{t_n}}\left(x_{t_1}, \dots, x_{t_n}\right) = F_{X_{t_1}}\left(x_{t_1}\right) \cdots F_{X_{t_n}}\left(x_{t_n}\right)$$

از ب ←الف:

$$F_{X_{t_1}, \dots, X_{t_n}} = F_{X_{t_1}} \cdots F_{X_{t_n}}$$

$$= P(X_{t_1} \in (-\infty, x_{t_1}], \dots, X_{t_n} \in (-\infty, x_{t_n}])$$

$$\Rightarrow P\left(X_{t_1} \in s_{t_1}, \dots, X_{t_n} \in s_{t_n}\right) = P\left(X_{t_1} \in s_{t_1}\right) \cdots P\left(X_{t_n} \in s_{t_n}\right)$$

$$P\left(\bigcap_{k=1}^{n}(X_{i_k} \in s_{t_k})\right) = \prod_{k=1}^{n} P(X_{t_k} \in s_{t_k})$$

از ج ← ب:

برای اثبات از این نکته استفاده می‌کنیم که اگر $\Phi(u)$ بر R^* مطلقاً انتگرال‌پذیر باشد تابع توزیع متناظر آن مطلقاً پیوسته است. بنابراین

$$f_{x_t}(\boldsymbol{x}) = \frac{1}{(2\pi)^n}\int_{R^n} e^{-iu_t' x_t}\,\phi_{x_{t_1},\dots x_{t_n}}\left(u_{t_1},\dots,u_{t_n}\right)du_{t_1}\cdots du_{t_n}$$

$$= \frac{1}{(2\pi)^n}\int_{R^n} e^{-i\Sigma u_{t_i}x_{t_i}}\,\phi_{x_{t_1}}\left(u_{t_1}\right),\dots,\phi_{x_{t_n}}\left(u_{t_n}\right)du_{t_1}\cdots du_{t_n}$$

$$= \frac{1}{2\pi}\int \phi_{x_{t_1}}\left(u_{t_1}\right)e^{-iu_{t_1}x_{t_1}}du_{t_1}\cdots \frac{1}{2\pi}\int \phi_{x_{t_n}}\left(u_{t_n}\right)e^{-iu_{t_n}x_{t_n}}du_{t_n}$$

از ب ←ج:

$$\phi_{x_t}(\boldsymbol{u}_t) = E\left(e^{iu_t x_t}\right) = \int_{R^n} e^{iu_t' x_t} f_{x_{t_1},\dots,x_{t_n}}d\boldsymbol{u}_t$$

$$= \int \cdots \int e^{iu_{t_1}x_{t_1}}\cdots e^{iu_{t_n}x_{t_n}} f_{x_{t_1}}\cdots f_{x_{t_n}}du_{t_1}\cdots du_{t_n}$$

$$= \int e^{iu_{t_1}x_{t_1}} f_{x_{t_1}}du_{t_1}\cdots \int e^{iu_{t_n}x_{t_n}} f_{x_{t_n}}du_{t_n}$$

$$= \phi_{x_{t_n}}\left(u_{t_n}\right)\cdots\phi_{x_{t_1}}\left(u_{t_1}\right)$$

بنابراین اثبات کامل است.

تعریف ۹-۵: کلاس P از زیر مجموعه‌های Ω یک π سیستم است اگر

$$A,B \in P \Rightarrow \quad A\cap B \in P$$

تعریف ۹-۶: کلاس Y یک λ سیستم است اگر

۱) $\Omega \in Y;$

۲) $A \in Y \ \Rightarrow A^c \in Y;$

۳) $A_1, A_2 \ldots \in Y$ ، برای $m \neq n$ $A_m \cap A_n = \phi \quad \Rightarrow \cup_n A_n \in Y$

نتیجه ۹-۲: اگر مجموعه‌ی $C_t = \{X^{-1}(-\infty, X_t],\ t\in T\}$ به ازای T های مختلف مستقل باشند، آنگاه $X^{-1}(B_t)$ که $B_t \in \mathcal{B}$، نیز مستقل‌اند.

$$(-\infty, v] \cap (-\infty, u] = (-\infty, \min(u\ , v)] \in (-\infty, x_t]$$

کلاس $C_t = \{X^{-1}(-\infty, X_t],\ t \in T\}$ تحت اشتراک متناهی بسته است.

بنابراین طبق قضیه ۳-۹ می‌توان گفت $\sigma\ (C_t)$ ها، نیز مستقل‌اند.

$$\sigma(C_t) = \sigma(X^{-1}(-\infty, X_t]) = X^{-1}(\sigma(-\infty, X_t]) = X^{-1}(B_t)$$

قضیه ۹-۵: فرض کنید

$$A_{11}, A_{12}, \dots$$

$$A_{21}, A_{22}, \dots$$

$$\vdots$$

دنباله‌ای از پیشامدهای مستقل باشد. تعداد پیشامدهای هر سطر می‌تواند متناهی یا نامتناهی باشد، در مورد تعداد سطرها نیز این مطلب برقرار است. اگر F_i سیگما میدان تولید شده توسط i امین سطر باشد، $F_1, F_2, \dots$ مستقل هستند.

اثبات: اگر A_i کلاس تمام اشتراک‌های متناهی از i امین سطر باشد، A_i یک π سیستم است. فرض کنید I مجموعه‌ای متناهی از اندیس‌ها باشد و به ازای هر $i \in I$، J_i نیز مجموعه‌ای متناهی از اندیس‌ها باشد. به ازای هر

$C_i = \cap_{j \in J_i} A_{ij}$ را در نظر می‌گیریم.

$$P(\cap_i C_i) = P\big(\cap_i \cap_j A_{ij}\big) = \prod_i \prod_j P\big(A_{ij}\big) = \prod_i P\big(\cap_j A_{ij}\big) = \prod_i P(C_i)$$

این نتیجه می‌دهد که $A_1, A_2, \dots$ مستقل‌اند.

قضیه ۹-۶: اگر $A_1, \dots, A_n$ مستقل باشند و هر A_i یک π سیستم باشد، آنگاه $\sigma(A_1), \dots, \sigma(A_n)$ نیز مستقل هستند.

اثبات: فرض کنید B_i کلاس A_i باشد که به Ω اضافه شده است. هر B_i یک π سیستم است و با توجه به فرض استقلال داریم

$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n), \qquad B_i \in \mathcal{B}_i$$

برای مجموعه‌های ثابت $B_2, \dots, B_n$ که در $\mathcal{B}_2, \dots, \mathcal{B}_n$ قرار می‌گیرند Y را کلاسی از مجموعه‌های B_1 در نظر می‌گیریم که در رابطه بالا صدق می‌کنند. بنابراین Y یک λ سیستم شامل π سیستم B_1 است و بنابراین شامل $\sigma(A_1) = \sigma(B_1)$ است.

بنابراین $P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n)$ برقرار است، اگر $B_1, B_2, \dots, B_n$ به ترتیب در $\sigma(A_1), A_2, \dots, A_n$ قرار بگیرد که این به این معناست که $\sigma(A_1), A_2, \dots, A_n$ مستقل‌اند. به وضوح اگر اندیس ۱ توسط هر یک از اندیس‌های ۲ و ۳ و ... و n جایگذاری شود، از استقلال $\sigma(A_1), A_2, \dots A_n$ می‌توان استقلال $\sigma(A_1), \sigma(A_2), \dots, A_n$ را نتیجه گرفت بنابراین $\sigma(A_1), \ \dots, \sigma(A_n)$ مستقل‌اند.

نتیجه۹-۳: مجموعه $\{X_\alpha : \alpha \in A\}$ از متغیرهای تصادفی (Ω, F, P) نسبت به P مستقل است اگر و تنها اگر به ازای هر $\{\alpha_1, \dots, \alpha_k\} \subset A$ و به ازای هر $x_1, \dots, x_k$ در R داشته باشیم:

$$F_{\alpha_1,\dots,\alpha_k}(x_1, \dots, x_k) = P(X_{\alpha_i} \le x_i, i = 1, \dots, k)$$

$$= \prod_{i=1}^{k} P\left(X_{\alpha_i} \le x_i\right) = \prod_{i=1}^{n} F_{\alpha_i}(x_i)$$

اگر $C_\alpha = \{X_\alpha^{-1}(-\infty, X_t]:\ x \in R\}, \alpha \in R$ با استفاده از قضیه (۹-۶)، اثبات کامل است.

لم ۹-۱: زیر کلاس‌های، کلاس‌های مستقل، مستقل‌اند.

اثبات: فرض کنید $\{C_t : t \in T\}$ کلاس‌های مستقل باشند و $\{C'_{t'} : t' \in T'\}$

را به عنوان زیر کلاسی از $\{C_t : t \in T\}$ در نظر می‌گیریم.

$T' \subset T$ چون $\{\boldsymbol{t_1}, \dots, \boldsymbol{t_n}\} \in \boldsymbol{T}, \{\boldsymbol{t'_1}, \dots, \boldsymbol{t'_n}\} \in \boldsymbol{T'}$ بنابراین می‌توان گفت $\{\boldsymbol{t'_1}, \dots, \boldsymbol{t'_n}\} \in \boldsymbol{T}$ و این یعنی

$$A_{t_i} \in C_{t'_i} \xrightarrow{C'_{t_i} \subset C_{t_i}} A_{t_i'} \in C_{t_i} \Rightarrow P\left(\cap A_{t'_i}\right) = \prod P\left(A_{t'_i}\right).$$

۹–۳ لم بورل کانتلی

قضیه ۹-۷: اگر (Ω, F, P) فضای احتمال و $\{A_n\}_{n\ge1}$ دنباله‌ای از پیشامدها در F باشد.

الف) لم اول بورل کانتلی: اگر $\sum_{n=1}^{\infty} P(A_n) < \infty$ آنگاه

$$P\left(\overline{\lim} A_n\right) = 0$$

ب) لم دوم بورل کانتلی: اگر $\{A_n\}_{n\ge1}$ دنباله‌ای مستقل از پیشامدها باشند و $\sum_{n=1}^{\infty} P(A_n) = \infty$ آنگاه

$$P\left(\overline{\lim} A_n\right) = 1$$

اثبات قسمت الف:

$$P(\overline{\lim} A_n) = P\left(\lim_{n\to\infty} \cup_{j\geq n} A_j\right)$$
$$= P\left(\cup_{n=1}^{\infty} \cap_{j=n}^{\infty} A_j\right) \left(\text{از پیوستگی } P\right)$$
$$\leq \quad P\left(\cup_{j=n}^{\infty} A_j\right) (\text{ از زیر جمعی بودن })$$
$$\leq \quad \sum_{j=n}^{\infty} P(A_j) \to \ 0$$
$$= 0$$

به دلیل اینکه همگرایی $\sum_n P(A_n)$ نتیجه می‌دهد که $\sum_{j=n}^{\infty} P(A_j) \to 0$ وقتی که $n \to \infty$.

اثبات قسمت ب: برای اثبات این قسمت کافی است نشان دهیم که $P(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) = 0$ و در واقع اثبات اینکه به ازای تمام مقادیر n، $P(\cap_{k=n}^{\infty} A_k^c)$کفایت می‌کند. از آنجایی که$1 - x \leq e^{-x}$ است داریم

$$P\left(\cap_{k=n}^{n+j} A_k^c\right) = \prod_{k=n}^{n+j}\left(1 - P(A_k)\right) \leq e^{-\sum_{k=n}^{n+j} P(A_k)}$$

و چون $\sum_k P(A_k)$ واگراست، عبارت سمت راست به سمت صفر میل می‌کند وقتی که $n \to \infty$، و از این رو

$$P(\cap_{k=n}^{\infty} A_k^c) = \lim_{\mathrm{j}} P\left(\cap_{k=n}^{n+j} A_k^c\right) = 0$$

و اثبات کامل است.

مثال ۹-۳: اگر $\{X_n, n \geq 1\}$ متغیرهای تصادفی برنولی باشند به طوری که

$$P(X_n = 1) = p_n = 1 - P(X_n = 0)$$

نشان می‌دهیم که $P\left(\lim_{\mathrm{n}\to\infty} X_n = \ 0\right) = 1$ اگر $\sum_n p_n < \infty$ باشد.

طبق لم اول بورل کانتلی اگر $\sum_n p_n = \sum_n P(X_n = 1) < \infty$ آنگاه با احتمال وقوع نامتناهی بار $(i.o.)$ پیشامد $A_n = [X_n = 1]$ برابر با صفر است.

با متمم‌گیری از طرفین داریم

$$1 = P\left(\lim_{\mathrm{n}\to\infty} \sup[X_n = 1]^c\right) = P\left(\lim_{\mathrm{n}\to\infty} \inf[X_n = 0]\right) = 1$$

از طرفی چون X_n یک متغیر برنولی است و طبق لم اول بورل کانتلی $P([X_n = 1]\ i.o.) = 0$ پس می‌توان گفت

$$P([X_n = 0]\ i.o.) = P\left(\lim_{n\to\infty} \sup\, [X_n = 0]\right) = 1$$

واضح است که $P(\lim_{n\to\infty} X_n = 0) = 1$ است.

نتیجه۹-۴:

$$\sum P(A_n^c) < \infty \Rightarrow P\big(\underline{\lim} A_n\big) = 1$$

اثبات: از قضیه (۹-۷) داریم

$$\sum P(A_n^c) < \infty \Rightarrow P\big(\overline{\lim} A_n^c\big) = 0 \Rightarrow P(\cap\cup\, A_n^c) = 0$$

$$P\big(\cap\cup\, A_n\big)^c = 0 \Rightarrow 1 - P(\cup\cap\, A_n) = 0 \Rightarrow P(\cup\cap\, A_n) = 1$$

$$\Rightarrow P\big(\underline{\lim} A_n\big) = 1$$

نتیجه۹-۵:

$$\sum P(A_n^c) < \infty, \sum P\big(A_n\big) = \infty \Rightarrow P(\lim A_n) = 1$$

نتیجه۹-۶:

$$P\big(\overline{\lim} A_n\big) = 0 \Rightarrow \sum P\big(A_n\big) < \infty$$

۹–۴ قوانین صفر و یک

تعریف ۹-۷: برای دنباله‌ای از پیشامدهای $A_1, \dots, A_n$ از فضای احتمال (Ω, F, P)، سیگما میدان $\sigma(A_n, A_{n+1}, \dots)$ را در نظر بگیرید تعریف می‌کنیم $T = \cap_{n=1}^{\infty} \sigma(A_n, A_{n+1}, \dots)$، T را یک *سیگما میدان دُمی* (دنباله ای) می‌نامیم که به وسیله A_n ها مشخص می‌شود. این سیگما میدان شامل پیشامدهایی است که فقط به رفتار حدی دنباله‌ی پیشامدها وابسته است. به هر $A \in T$ یک *پیشامد دمی* گفته می‌شود. به متغیر تصادفی که نسبت به T اندازه‌پذیر باشد، *متغیر تصادفی دمی* می‌گویند. پیشامدهای دمی توسط رفتار دنباله $\{X_n\}_{n\geq 1}$ برای n بزرگ تعیین می‌شوند و اگر هر یک از زیر کلاس‌های $\{X_n\}_{n\geq 1}$ حذف شوند و یا توسط مجموعه‌ای از متغیرهای تصادفی جایگزین شوند، بدون تغییر باقی می‌مانند. پیشامدهای $\left\{\lim_{n\to\infty} X_n = x\right\}$ و

$\{\limsup_{n\to\infty} X_n < x\}$ به T تعلق دارند.

۹-۴-۱ معیار ۱-۰ بورل

قضیه۹-۸: اگر $\{A_n\}$ دنباله‌ای از پیشامدهای مستقل باشد، در این صورت $I(A_n)$ دنباله‌ای از متغیرهای مستقل خواهند بود. بنابراین $\lim I(A_n)$ و $\underline{\lim}\, I(A_n)$, $\overline{\lim}\, I(A_n)$ در صورت وجود توابع دنباله‌ای‌اند. در نتیجه تقریباً همه جا $(a.s.)$ تباهیده هستند، چون:

$$\overline{\lim}\, I(A_n) = I(\overline{\lim} A_n)$$

$$\underline{\lim}\, I(A_n) = I(\underline{\lim} A_n)$$

پس $I(\overline{\lim} A_n)$ به طور $a.s.$ برابر با صفر یا یک است به عبارت دیگر $P\left(\underline{\lim}\,(A_n)\right)$ برابر صفر و یا یک است. همین‌طور $P(\overline{\lim}(A_n))$ برابر صفر و یا یک است. اگر $A = \lim A_n$ باشد $P(A)$ برابر صفر یا یک خواهد بود.

نتیجه ۹-۷: اگر x_n ها مستقل باشند و $x_n \xrightarrow{a.s.} 0$ آنگاه برای هر مقدار متناهی $c > 0$ داریم

$$\sum P(|x_n| \geq c) < \infty$$

اثبات: اگر x_n ها متغیرهای تصادفی مستقل باشند $A_n = \{|x_n| \geq c\}$ نیز پیشامدهای مستقل‌اند چون

$x_n \xrightarrow{a.s.} 0$ اگر و فقط به ازای هر $c > 0$ وقتی $n \to \infty$ داشته باشیم:

$$P(\cup_{n=1}^{\infty} (|x_n| \geq c)) = P(\sup A_n) \to 0$$

خواهیم داشت

$$P(\overline{\lim}(A_n)) = 0$$

بنا به قضیه (۹-۷)، $\sum P(A_n) < \infty$ و بنابراین نتیجه (۹-۷)، برقرار خواهد بود.

نتیجه۹-۸: اگر $\{x_n\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشد یا $a.s.$ همگرا و یا $a.s.$ واگرا است.

نتیجه‌ی فوق برای $(b_n \uparrow \infty)$,$\frac{s_n}{b_n}$ که در آن $s_n = \sum_{k=1}^{n} x_k$ نیز برقرار است. علاوه بر این حدود دنباله‌های $\{x_n\}$ و $\{\frac{s_n}{b_n}\}$، نیز $a.s.$ تباهیده‌اند.

اثبات: می‌دانیم $\overline{\lim}x_n$ ، $\lim x_n$ و $\underline{\lim}x_n$ و مجموعه همگرایی $\{X_n\}$ یعنی جایی که $\overline{\lim}x_n = \underline{\lim}x_n$ است، نسبت به سیگما میدان دنباله $\{x_n\}$ اندازه‌پذیرند یعنی $\lim, \overline{\lim}$ و $\underline{\lim}$ متعلق به سیگما میدان دنباله $\{x_n\}$ هستند. از آنجایی که x_nها مستقل‌اند. طبق قضیه قبلی پیشامدهای موجود در سیگما میدان دنباله‌ای $a.s.$ تباهیده‌اند و چون $\lim, \overline{\lim}$ و $\underline{\lim}$ متعلق به سیگما میدان دنباله‌ای‌اند، لذا آن‌ها هم $a.s.$ تباهیده‌اند چون $\underline{\lim} = \overline{\lim}$ به طور $a.s.$ تباهیده‌اند. پس تقریباً برای تمام w ها $\overline{\lim}x_n = \underline{\lim}x_n$ یا $\overline{\lim}x_n \neq \underline{\lim}x_n$ که در حالت اول همگرایی $a.s.$ وجود دارد و $\lim x_n$ وجود دارد و در حالت دوم هیچ همگرایی وجود ندارد و $\lim x_n$ در هیچ نقطه‌ای موجود نیست. برای $\frac{s_n}{b_n}$ باید ثابت کنیم که حد $\frac{s_n}{b_n}$ به طور $a.s.$ تباهیده است.

چون $\frac{s_n}{b_n}$ همگراست لذا حد بالا و پایین آن برابر است. بنابراین

$$\frac{s_n}{b_n} = \frac{x_1 + \cdots + x_m}{b_n} + \frac{x_{m+1} + \cdots + x_n}{b_n}$$

$(b_n \uparrow \infty)$

$$\underline{\lim}\frac{s_n}{b_n} = \underline{\lim}\frac{x_{m+1} + \cdots + x_n}{b_n}$$

تابع اخیر نسبت به $\mathcal{B}(X_{m+1}, \dots)$ اندازه‌پذیراست چون $X_{m+1}, \dots$ در $\mathcal{B}(X_{m+1}, \dots)$ است و در نتیجه $\lim \frac{s_n}{b_n}$ نسبت به سیگما میدان دنباله اندازه‌پذیر است و در نتیجه $a.s.$ تباهیده است. بنابراین

$$\underline{\lim}\frac{s_n}{b_n} = \sup\inf\frac{s_n}{b_n} = \sup\inf\frac{\sum_{k=m+1}^{n} x_k}{b_n}$$

۹-۴-۲ قانون ۱-۰ کلموگروف

قضیه۹-۹: اگر $A_1, A_2, \dots$ دنباله‌ای از پیشامدهای مستقل باشد، به ازای هر $A \in T$، 1یا $P(A) = 0$ است.

اثبات: می‌دانیم اگر

$$A_{11}, A_{12}, \dots$$

$$A_{21}, A_{22}, \dots$$

$$\vdots$$

پیشامدهای مستقل باشند، آنگاه $F_1, F_2, \dots$ مستقل‌اند به طوری‌که F_i، سیگما میدان تولید شده توسط سطر iام است. و اگر $A_1, \dots, A_n$ مستقل از هم و $\pi-$سیستم باشند، $\sigma(A_1), \dots, \sigma(A_n)$ نیز از هم مستقل‌اند. بنابراین $\sigma(A_n, A_{n+1}, \dots)$ و $\sigma(A_1), \dots, \sigma(A_{n-1})$ نیز از هم مستقل‌اند.

اگر $A \in T$ باشد، آنگاه $A \in \sigma(A_n, A_{n+1}, \dots)$ بنابراین A، $A_1, \dots A_{n-1}$ مستقل‌اند. از آنجایی که استقلال یک مجموعه از پیشامدها توسط استقلال هر زیر مجموعه غیرمتناهی تعریف می‌شود. دنباله $A, A_1, A_2, \dots$ مستقل است. بنابراین $\sigma(A)$ و $\sigma(A_1, A_2, \dots)$ از هم مستقل‌اند. از آنجایی که

$$A \in T \subset \sigma(A_1, A_2, \dots)$$

$$A \in \sigma(A)$$

$$A \in \sigma(A_1, A_2, \dots)$$

بنابراین A از خودش مستقل است. پس می‌توان نوشت

$$P(A \cap A) = P(A)P(A)$$

$$\Rightarrow P(A) = P^2(A)$$

$$\Rightarrow P^2(A) - P(A) = 0$$

واضح است که $P(A)$ باید برابر با صفر یا یک باشد.

نکته ۹-۲: در لم بورل کانتلی پیشامد رخداد، تعداد نامتناهی از A_nها مورد نظر است. بنابراین با رخ دادن تعداد متناهی از پیشامدهای مورد نظر نمی‌توان در مورد وقوع تعداد نامتناهی تا از آن‌ها بحث کرد. بنا بر قضیه 0-1 کلموگروف احتمال وقوع تعداد نامتناهی از A_nها، وقتی A_nها مستقل باشند برابر با ۰ یا ۱ است.

۹-۵ خواص ضرب

قضیه ۹-۱۰: اگر $X_1, \dots, X_n$ مستقل باشند آن‌گاه

$$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i) \qquad (۹-۴)$$

مشروط بر آنکه دو طرف موجود باشند.

اثبات:

الف) رابطه (۹-۴)، را به ازای $n=2$ اثبات می‌کنیم سپس با استفاده از استقراء روی n و استفاده از این خاصیت که استقلال X_{n+1} و $(X_1,\dots,X_n)$ استقلال X_{n+1} و $X_k,\ k=1,2,\dots,n$ را نتیجه می‌دهد می‌توانیم رابطه‌ی (۹-۴)، را برای هر n دلخواه ثابت کنیم.

ب) اگر X و Y دو تابع ساده باشند، داریم

$$X=\sum_{j=1}^{m} X_j\, I(A_j)\quad,\quad Y=\sum_{k=1}^{n} Y_k I(B_k)$$

که در آن $\{A_1,\dots,A_m\}$، $\{B_1,\dots,B_n\}$ دو افراز مستقل Ω هستند آن‌گاه

$$E(XY)=E\left(\sum_{j=1}^{m}\sum_{k=1}^{n} X_j Y_k I(A_j B_k)\right)$$

$$=\sum_{j=1}^{m}\sum_{k=1}^{n} x_j y_k P(A_j B_k)=\sum_{j}\sum_{k} x_j y_k P(A_j)P(B_k)$$

$$=\left(\sum_{j} x_j P(A_j)\right)\left(\sum_{k} y_k P(B_k)\right)=E(X)E(Y)$$

ج) فرض کنید $x,y\geq 0$ در اینصورت دنباله‌ای از توابع ساده مانند $\{X_n\},\{Y_n\}$ وجود دارند به قسمی که $0\leq X_n \uparrow X$ و $0\leq Y_n \uparrow Y$ در نتیجه $0\leq X_nY_n \uparrow XY$. اگر X و Y مستقل باشند، آنگاه $\{X\in B\}$ و $\{Y\in C\}$ نیز به ازای هر دو مجموعه بورل B و C مستقل هستند. بنابراین X_n و Y_n مستقل‌اند.

از (ب) نتیجه می‌شود

$$E(X_nY_n)=E(X_n)E(Y_n)$$

چون

$$E(X_n)\uparrow E(X_n), E(Y_n)\uparrow E(Y_n),$$

$$E(X_nY_n)\uparrow E(X_n)E(Y_n)$$

E(XnYn) ↑ E(XY)

و

$$E(X_nY_n) = E(X_n)E(Y_n) \uparrow E(X)E(Y)$$

داریم

$$E(XY) = E(X)E(Y)$$

د) فرض کنید X و Y دلخواه باشند، در این صورت

$$\begin{aligned} XY &= (X^+ - X^-)(Y^+ - Y^-) \\ &= X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^- \end{aligned}$$

چون X^+ و X^- توابعی بورل از X هستند، (Y^-, Y^+, X^+, X^-) از هم مستقل خواهند بود. بنابراین

$$\begin{aligned} E(XY) &= E(X^+)E(Y^+) - E(X^+)E(Y^-) - E(X^-)E(Y^+) + E(X^-)E(Y^-) \\ &= (E(X^+) - E(X^-))(E(Y^+) - E(Y^-)) \\ &= E(X)E(Y) \end{aligned}$$

تعریف ۹-۸: (**ناسازگاری**)؛ دو پیشامد A, B را ناسازگار گوییم، هرگاه $P(A \cap B) = 0$. ناسازگاری دو پیشامد دلیلی بر استقلال آنها نیست.

توجه کنید در صورتی دو پیشامد ناسازگار A, B، مستقل‌اند که یکی از آن‌ها پوچ یا حتمی باشند.

اگر A پوچ باشد

$$P(A \cap B) = 0, \qquad P(A) = 0 \Rightarrow P(A \cap B) = P(A)P(B)$$

اگر A حتمی باشد

$$P(A \cap B) = 0 \Rightarrow P(B \cap \Omega) = P(B) = 0 \Rightarrow 0 = 1 \times 0 \Rightarrow 0 = 0$$

تعریف ۹-۹: اگر $X_1 \sim F_1, X_2 \sim F_2$ از هم مستقل باشند آن‌گاه تابع توزیع $X_1 + X_2$ عبارتست از پیچش F_1 و F_2 که بصورت زیر تعریف می‌کنیم.

$$\begin{aligned} F(X) &= \int F_1(x-y)dF_2(y), \qquad x \in R \\ &= \int F_2(x-y)dF_1(y), \qquad x \in R \end{aligned}$$

مثال ۹-۴: اگر x_1, x_2 متغیرهای مستقل نرمال باشند از تابع مشخصه $x_1 + x_2$ معلوم می‌شود که $x_1 + x_2$ نرمال است.

اگر x_1, x_2 متغیرهای مستقل نرمال باشند، $x_1 + x_2$ نیز نرمال است

$$\phi_{X_1+X_2}(u) = E\left(e^{iu(X_1+X_2)}\right) = \phi_{X_1}(u)\phi_{X_2}(u)$$

$$\phi_X(t) = M_X(it) = e^{i\mu t+\frac{\sigma^2(it)^2}{2}} = e^{i\mu t-\frac{\sigma^2t^2}{2}}$$

$$\Rightarrow \phi_{X_1+X_2}(u) = e^{i\mu_1 u-\frac{\sigma_1^2u^2}{2}}e^{i\mu_2 u-\frac{\sigma_2^2u^2}{2}} = e^{iu(\mu_1+\mu_2)-\frac{u^2}{2}(\sigma_1^2+\sigma_2^2)}$$

$$\Rightarrow x_1 + x_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

برای سهولت می‌توان از تابع مشخصه‌ی مجموع متغیرهای تصادفی استفاده کرد. فرض کنید $y = \sum x_i$ باشد که X_iها مستقل و هم توزیع‌اند در آن صورت **:**

$$\phi_Y(u) = E(e^{iuY}) = E(e^{iu\Sigma x_i}) = (\phi_X(u))^n$$

$$\sqrt[n]{\phi_Y(u)} = \phi_X(u)$$

به طور کلی کلاس توابع توزیع‌ها تحت پیچش بسته است هرگاه ریشه n ام تابع مشخصه مجموع متغیرها، به ازای هر n دلخواه یک تابع مشخصه باشد. با استفاده از تابع مشخصه می‌توان توزیع را محاسبه کرد. پس می‌توان گفت که کلاس توابع توزیع‌ها تحت پیچش بسته است هرگاه ریشه n ام تابع توزیع مجموع متغیرها یک تابع توزیع باشد. به طور کل X, Y به یک نوع توزیع تعلق دارند اگر یکی ترکیب خطی از دیگری باشد.

اگر $ax + by$ به ازای هر a و b نرمال باشد داریم:

$$\phi_{x,y}(u, v) = E(e^{iux+ivy}) = E(e^{iu'ax+iu'by})$$

$$E\left(e^{iu'(aX+bY)}\right) = \phi_{aX+bY}(u')$$

$$(v = bu', u = au')$$

مثال ۹-۵: اگر x, y متغیرهای مستقل پواسن باشند آن‌گاه

$$\phi_{ax}(t) = E(e^{itax}) = \phi_x(at) = e^{\lambda(e^{iat}-1)}$$

$ax + by$ از نوع پواسون است ولی پواسون نیست. یعنی تحت پیچش بسته است.

اگر x_iها متغیرهای مستقل نرمال استاندارد باشند، آنگاه تابع چگالی احتمال $\sum_{i=1}^{n} X_i^2$ کای دو با n درجه آزادی است.

نکته ۹-۳: اگر X, Y مستقل باشند ، هر تابع دلخواه از X, Y نیز از هم مستقل‌اند.

نتیجه ۹-۹: اگر $X_1, \dots, X_n$ مستقل باشند آن‌گاه تابع مشخصه مجموع متغیرهای تصادفی برابر با حاصل‌ضرب توابع مشخصه x_kهاست. به عبارت دیگر

$$\phi_{\sum x_k}(u) = \prod_{k=1}^{n} \phi_{x_k}(u)$$

عکس این حالت برقرار نیست.

اثبات:

$$\phi_{\sum x_k}(u) = E(e^{iu\sum x_k}) = E\left(e^{iu(x_1+\cdots+x_n)}\right) = E(\prod_{k=1}^{n} e^{iux_k})$$

$$= \prod_{k=1}^{n} E\left(e^{iux_k}\right) = \prod_{k=1}^{n} \phi_{x_k}(u)$$

یعنی اگر تابع مشخصه مجموع متغیرهای تصادفی برابر با حاصل‌ضرب توابع مشخصه باشد، دلیلی ندارد که x_kها مستقل باشند.

مثال ۹-۶: اگر X یک متغیر تصادفی با توزیع کوشی و تابع مشخصه $e^{-|u|}$ باشد، در این صورت تابع مشخصه متغیر cX به صورت $e^{-c|u|}$ است و تابع مشخصه $(1+c)X$ به صورت زیراست:

$$\phi_X(u) = e^{-|u|}$$

$$\phi_{cX}(u) = E(e^{iucx}) = \phi_X(cu) = e^{-c|u|}$$

$$\phi_{(c+1)X}(u) = E\left(e^{iu(c+1)x}\right) = \phi_X((c+1)u) = e^{-(c+1)|u|}$$

با استفاده از روابط بالا می‌توان نوشت

$$\phi_{X+cX}(u) = \phi_X(u)\phi_{cX}(u)$$

بنابراین دلیلی ندارد که X و cX مستقل باشند.

چون X و cX سیگما میدان غیر تباهیده یکسانی تولید می‌کنند لذا مستقل نیستند. اگر قرار باشدکه X و cX مستقل باشد باید میدان سیگمای القاء شده توسط آن‌ها مستقل باشد ولی در اینجا X و cX سیگما میدان یکسانی تولید می‌کنند و چون این سیگما میدان شامل پیشامدهای مشابه است، لذا پیشامد دلخواه A نمی‌تواند از خودش مستقل باشد. در نتیجه X, cX مستقل نیستند.

$$X = \sum X_k I(A_k) = X_1 I(A_1) + \cdots + X_n I(A_n)$$

$$X = \begin{cases} X_1 & w \in A_1 \\ \vdots & \vdots \\ X_n & w \in A_n \end{cases} \Rightarrow X^{-1}(B) = \begin{cases} \phi & x_1, \dots, x_n \notin \mathcal{B} \\ A_i & x_i \in \mathcal{B} \\ A_i + A_j & x_i, x_j \in \mathcal{B} \\ \vdots & \vdots \\ \Omega & \forall\, x_i \in \mathcal{B} \end{cases}$$

$$CX = CX_1 I(A_1) + \cdots + CX_n I(A_n)$$

$$X = \begin{cases} CX_1 & w \in A_1 \\ \vdots & \vdots \\ CX_n & w \in A_n \end{cases} \Rightarrow (CX)^{-1}(B) = \begin{cases} \phi & Cx_1, \dots, Cx_n \notin \mathcal{B} \\ A_i & Cx_i \in \mathcal{B} \\ A_i + A_j & Cx_i, Cx_j \in \mathcal{B} \\ \vdots & \vdots \\ \Omega & \forall\, Cx_i \in \mathcal{B} \end{cases}$$

$$CX^{-1}(B) = (CX)^{-1}(B)$$

۹–۶ مسائل

۱) اگر X و Y متغیرهای مستقل گاما باشند،نشان دهید $X + Y$ و $\frac{X}{Y}$ به طور مستقل توزیع شده‌اند.

۲) شرایطی که تحت آن توزیع دوجمله‌ای تحت عمل پیچش بسته است را بیان کنید.

$$X_1 \sim B(n_1, P_1), \qquad X_2 \sim B(n_2, P_2)$$

۳) ثابت کنید اگر X و Y مستقل باشند $g(x)$ و $h(y)$ نیز مستقل‌اند. اگر X, Y با توزیع‌های $F(y), G(x)$ مستقل باشند نشان دهید احتمال آنکه Y کوچکتر از X باشد برابر است با $\int F(x) dF(x)$.

۴) اگر $A_1, \dots, A_n$ پیشامدهای متقابلاً مستقل باشند ثابت کنید

$$P(\cup A_k) = P(A_1) + \big(1 - P(A_1)\big)P(A_2) + \cdots$$

۵) اگر $P(X=i,\ Y=j) = P_{ij}$ که در آن

$$P_{-1,0} = P_{1,0} = \frac{3}{16}, P_{0,0} = \frac{1}{8}, P_{-1,1} = \frac{1}{16}, P_{0,1} = \frac{3}{8}, P_{1,1} = \frac{1}{16}$$

۶) نشان دهید که X, Y ناهمبسته هستند ولی مستقل نیستند.

۷) تابع چگالی توأم سه متغیر به صورت زیر است

$$f(x_1,x_2,x_3) = \begin{cases} \frac{1}{2} + 4x_1x_2x_3 & 0 \le x_1,x_2,x_3 \le 1 \\ 0 & \text{در غیر این صورت} \end{cases}$$

آیا x_1, x_2, x_3دو به دو مستقل هستند؟ توأماً چطور؟

۸) یک سکه سالم را سه بار پرتاب می‌کنیم، پیشامدهای زیر را در نظر بگیرید

$$B_{uv} = \{\omega: d_v(\omega) = d_u(w)\}$$

(به این معنی که نتیجه پرتاب سکه در بار uام و vام با هم برابر است.) اگر فرض کنیم که

$$A_1 = B_{12} \qquad A_2 = B_{13} \qquad A_3 = B_{23}$$

آنگاه ثابت کنید A_1, A_2, A_3 دو به دو مستقل‌اند؟ آیا متغیرها متقابلاً مستقل‌اند؟

۹) فرض کنید $\{x_n\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم توزیع باشند و $E(|X_1|) = \infty$ و نشان دهید: $E(X) = \int_0^\infty P(X \ge t)dt$

الف) $\sum P(|X_n| \ge an) = \infty \quad \forall\, a$

ب) $\sup \frac{|X_n|}{n} = \infty \quad (a.s.)$

فصل دهم

قانون اعداد بزرگ

در این فصل به مبحث قانون اعداد بزرگ می‌پردازیم. اما در ابتدا نیازمند یادگیری مفاهیمی هم‌چون تمرکز، برش، متغیرهای به‌طور دنباله‌ای معادل و معادل همگرایی می‌باشیم. در ادامه همگرایی در احتمال S_n را بررسی کرده و سپس نامساوی کلموگروف را تحت عنوان قضیه مطرح می‌کنیم و بعد از آن همگرایی a.s یک سری را بررسی می‌کنیم. هم‌چنین با مفهوم جدید دنباله متقارن شده روبرو می‌شویم و در پایان قانون قوی اعداد بزرگ را که بیانگر رفتار متغیرهای تصادفی مستقل و میانگین آن‌هاست، مورد بحث قرار می‌دهیم.

۱۰–۱ همگرایی یک سری از متغیرهای تصادفی مستقل

۱۰–۱–۱ تعاریف

تعریف۱۰-۱: X در C مرکزی شده است اگر به جای X، عبارت $X - C$ را قرار دهیم و اگر به جای X، $X - E(X)$ قرار دهیم، می‌گوییم X در میانگین خود مرکزی شده است. همان طور که می‌بینیم مرکزی کردن همان تغییر مبدا است که تأثیری در واریانس نخواهد داشت

$$Var(X) = Var(X - C)$$

تعریف۱۰-۲: X در C بریده شده است، اگر به جای X متغیر X^c را که به صورت زیر تعریف می‌شود، قرار دهیم

$$X^c = \begin{cases} X & |X| < C \\ 0 & |X| \geq C \end{cases} (0 < C < \infty).$$

در واقع به جای متغیر تصادفی دلخواه X با یک متغیر تصادفی X^c سر و کار داریم که a.s کراندار است و همچنین تمام گشتاورهای X^c موجود و متناهی هستند.

به عنوان مثال گشتاور مرتبه اول و دوم X^c را به دست می‌آوریم که متناهی است

$$E(X^c) = \int_{|X|<C} X^c \, \mathrm{dF} + \int_{|X|\geq C} X^c \, \mathrm{dF}$$

$$= E\big(XI(|X| < C)\big) + 0 \le C$$

$$E(X^c)^2 = E\big(XI(|X| < C)\big)^2$$

$$= E\big(X^2 I(|X| < C)\big) = \int_{|X|<C} X^2 \, \mathrm{dF} \le C^2$$

تعریف ۱۰-۳: دو دنباله $\{X_n\}$, $\{X'_n\}$ از متغیرهای تصادفی را به طور دنباله‌ای معادل گوییم اگر تنها در تعداد متناهی جمله با هم متفاوت باشند. به عبارت دیگر

$$\forall \omega \in \Omega, \;\; \exists\, n(\omega) < \infty \;\; s.t \;\; \forall\, n \ge n(\omega),$$

دو دنباله $X_n(\omega), X'_n(\omega)$ یکسان باشند به این صورت که

$$\exists n(\omega), \forall\, n \ge n(\omega) \; s.t. \; P\{X_n(\omega) = X'_n(\omega)\} = 1$$

$$\exists n(\omega), \forall\, n \ge n(\omega) \; s.t. \; P\{X_n(\omega) \ne X'_n(\omega)\} = 0$$

$$P\left\{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}(X_k \ne X'_k)\right\} = 0$$

یا

$$P\{\overline{lim}\; (X_k \ne X'_k)\} = 0$$

به طور معادل، $P\{\underline{lim}\; (X_k = X'_k)\} = 1$.

نتیجه ۱۰-۱: با توجه به لم بورل-کانتلی داریم

اگر $\sum P(X_n \ne X'_n) < \infty$ باشد، آنگاه $P\{\overline{lim}(X_n \ne X'_n)\} = 0$

بنابراین می‌توان لم را به این صورت نیز بیان کرد:

اگر $\{X_n\}$, $\{X'_n\}$ به طور دنباله‌ای معادل باشند آنگاه $P\{\overline{lim}(X_n \ne X'_n)\} = 0$

لم ۱۰-۱: اگر $S_n = \sum X_k$, $S'_n = \sum X'_k$ در نظر بگیریم و $\sum_n P(X_n \ne X'_n) < \infty$، آنگاه $\{X_n\}$, $\{X'_n\}$ به طور دنباله‌ای معادل می‌باشند و $\{S_n\}, \{S'_n\}$ معادل همگرایی هستند. علاوه بر این $\left\{\frac{S_n}{b_n}\right\}, \left\{\frac{S'_n}{b_n}\right\}$ وقتی $b_n \uparrow \infty$ در پیشامدهای یکسان به سمت حد یکسان همگرا خواهند بود.

اثبات: به موجب لم بورل کانتلی $\{X_n\}$, $\{X'_n\}$ به طور دنباله‌ای معادل هستند. بنابراین به ازای هر ω ی معین

(ثابت $= (S_n(\omega) - S'_n(\omega)$)، اگر $\lim S_n(\omega)$ به ازای هر مقدار مفروض ω موجود باشد $\lim S'_n(\omega)$ نیز وجود دارد. پس مجموعه‌های همگرایی S_n , S'_n به طور a.s یکسان هستند. علاوه بر این وقتی $n \to \infty$ آنگاه:

$$\frac{S_n(\omega) - S'_n(\omega)}{b_n} = \frac{\text{ثابت}}{b_n} \to 0$$

پس $\left\{\frac{S_n}{b_n}\right\}, \left\{\frac{S'_n}{b_n}\right\}$ به طور a.s به سمت یک حد همگرا هستند.

تا این‌جا استقلال X_nها جزء مفروضات نبوده اما از این به بعد باید این شرط را نیز اضافه کنیم.

۱۰-۱-۲ همگرایی در احتمال S_n

لم۱۰-۲: فرض کنید $\sigma_k^2 = \sigma^2{}_{(X_k)} = Var\ X_k$ و $\sigma_n^2 = \sum_{k=1}^n \sigma_k^2$، در این صورت

$$\frac{\sum_{k=1}^n \sigma_k^2 - \varepsilon^2}{sup[S_n - E(S_n)]^2} \le P\{|S_n - E(S_n)| > \varepsilon\} \le \frac{\sum_{k=1}^n \sigma_k^2}{\varepsilon^2}$$

اثبات: بنا بر نامساوی اساسی داریم

$$\frac{E\big(g(X) - g(a)\big)}{\sup g(X)} \le P\{|X| \ge a\} \le \frac{E\big(g(X)\big)}{g(a)}$$

حال اگر تغییرات زیر را در نامساوی فوق اعمال کنیم داریم

فرض کنیم داشته باشیم

$$\begin{cases} a = \varepsilon \\ g(X) = X^2 \\ X = S_n - E(S_n) \end{cases}$$

و همچنین داشته باشیم

$$E\big(g(X)\big) = E(X^2) = E\big(S_n - E(S_n)\big)^2 = Var\ S_n$$

در نتیجه:

$$\frac{\sum_{k=1}^n \sigma_k^2 - \varepsilon^2}{sup[S_n - E(S_n)]^2} \le P\{|S_n - E(S_n)| > \varepsilon\} \le \frac{\sum_{k=1}^n \sigma_k^2}{\varepsilon^2}$$

نتیجه۱۰-۲: اگر $\sum \sigma_k^2 < \infty$ آنگاه $\sum[X_k - E(X_k)]$ در احتمال همگرا است.

اثبات: فرض کنید $Z_n = \sum_{k=1}^{n}[X_k - E(X_k)] = S_n - E(S_n)$ و همچنین

$$Z_{n+m} = \sum_{k=1}^{n+m}[X_k - E(X_k)]$$

$$= \sum_{k=1}^{m}[X_k - E(X_k)] + \sum_{k=m+1}^{n+m}[X_k - E(X_k)]$$

اگر بتوان نشان داد وقتی $m, n \to \infty$ آنگاه $P\{|Z_{n+m} - Z_m| \geq \varepsilon\} \to 0$ می‌توان نتیجه گرفت Z_n در احتمال همگراست. با توجه به

$$Var(Z_{n+m} - Z_m) = Var\,(Z_{n+m}) + \,Var\,(Z_m) - 2cov(Z_{n+m}, Z_m)$$

$$= Var(Z_{n+m}) + Var(Z_m) \; - 2cov\left(Z_m + \sum_{k=m+1}^{n+m}[X_k - E(X_k)]\,, Z_m\right)$$

$$= Var(Z_{n+m}) + Var(Z_m) - 2Var(Z_m) - 2cov\left(\sum_{k=m+1}^{n+m}[X_k - E(X_k)]\,, Z_m\right)$$

$$= Var(Z_{n+m}) - Var(Z_m)$$

$$= Var\left(Z_m + \sum_{k=m+1}^{n+m}[X_k - E(X_k)]\right) - Var(Z_m)$$

$$= Var\left(\sum_{k=m+1}^{n+m}[X_k - E(X_k)]\right) + Var(Z_m) - Var(Z_m)$$

$$= Var\left(\sum_{k=m+1}^{n+m}[X_k - E(X_k)]\right) = \sum_{k=m+1}^{n+m}\sigma_k^2$$

و به موجب لم (۱۰-۲)، می‌توان نوشت

$$P\{|Z_{n+m} - Z_m| \geq \varepsilon\} \leq \frac{\sum_{k=m+1}^{n+m}\sigma_k^2}{\varepsilon^2} \to 0$$

وقتی $\sum \sigma_k^2 < \infty$ و $m, n \to \infty$ در نتیجه $\sum_{k=m+1}^{n+m}\sigma_k^2 \to 0$ پس Z_n در احتمال همگرا است

نتیجه۱۰-۳: اگر $\sum_{k=1}^{n}\frac{\sigma_k^2}{b_n^2} \to 0$ آنگاه $\frac{[S_n - E(S_n)]}{b_n} \to 0$

اثبات: اگر به جای ε مقدار εb_n را در لم (۱۰-۲)، جایگزین کنیم، داریم

$$P\{|Z_n| > \varepsilon b_n\} = P\{|S_n - E(S_n)| > \varepsilon b_n\} \leq \frac{\sum \sigma_k^2}{\varepsilon^2 b_n^2} \xrightarrow{b_n \uparrow \infty} 0$$

$$\Rightarrow P\{|S_n - E(S_n)| > \varepsilon b_n\} \to 0 \Rightarrow \frac{S_n - E(S_n)}{b_n} \xrightarrow{p} 0$$

۱۰-۲ نامساوی کلموگروف

قضیه۱۰-۱: اگر X_k ها مستقل و انتگرال‌پذیر باشند و $|X_k| \leq c < \infty$ آنگاه

$$\forall \varepsilon > 0, \quad 1 - \frac{(\varepsilon + 2c)^2}{\sum \sigma_k^2} \leq P\left\{\max_{k \leq n} |S_k - E(S_k)| \geq \varepsilon\right\} \leq \frac{\sum_{k=1}^{n} \sigma_k^2}{\varepsilon^2}$$

اثبات: ابتدا باید دقت کرد که اگر یکی از واریانس‌ها نامتناهی شود آنگاه $\sum \sigma_k^2$ هم نامتناهی می‌شود و سمت راست نامساوی بدیهی می‌باشد. هم‌چنین اگر c نامتناهی باشد آن‌گاه سمت چپ نامساوی $-\infty$ می‌شود بنابراین نامساوی بی‌معنی می‌شود پس لازم است فرض کنیم c و تمام واریانس‌ها متناهی هستند.

اگر X کراندار با کران c باشد سپس $|X - E(X)|$ دارای کران $2c$ است. پس در اثبات قضیه از متغیرهایی استفاده می‌کنیم که درمیانگین خود مرکزی شده‌اند و دارای کران $2c$ می‌باشند. مجموعه A_k ها که نزولی می‌باشند را به صورت زیر تعریف می‌کنیم

$$A_k = \left\{\max_{j \leq k} |S_j| < \varepsilon\right\} = \{|S_1| < \varepsilon, \ldots, |S_k| < \varepsilon\}$$

نزولی‌اند، فرض می‌کنیم $A_0 = \Omega$، بنابراین $A_k^c = \left\{\max_{j \leq k} |S_j| \geq \varepsilon\right\}$ و

$$\begin{aligned} B_k &= A_{k-1} - A_k = A_{k-1} \cap A_k^c \\ &= \left\{\max_{j \leq k-1} |S_j| < \varepsilon, \max_{j \leq k} |S_j| \geq \varepsilon\right\} \\ &= \{|S_1| < \varepsilon, \ldots, |S_{k-1}| < \varepsilon, |S_k| \geq \varepsilon\} \end{aligned}$$

تساوی بالا زمانی وجود خواهد داشت که $S_k = \max_{j \leq k} |S_j|$، در غیر این‌صورت B_k تهی خواهد شد و B_k ها جدا از همدیگر هستند.

$$B_1 = A_0 - A_1 = A_0 \cap A_1^c = \Omega \cap A_1^c = A_1^c$$

$$B_2 = A_1 - A_2 = A_1 \cap A_2^c$$

$$\vdots$$

$$B_k = A_{k-1} - A_k = A_{k-1} \cap A_k^c$$

$$\Rightarrow B_k \cap B_{k-1} = (A_{k-1} \cap A_k^c) \cap (A_{k-2} \cap A_{k-1}^c) = \emptyset$$

و همچنین

$$\sum_{k=1}^{n} B_k = A_0 - A_1 + A_1 - A_2 + \cdots + A_{n-1} - A_n$$

$$= A_0 - A_n = \Omega - A_n = A_n^c = \left\{\max_{j \le n} |S_j| \ge \varepsilon\right\}$$

B_k ها به وسیله متغیرهای تصادفی $(X_1, X_2, \dots, X_k)$ معین می‌شوند پس B_k ها از هر متغیر بر پایه $(X_{k+1}, X_{k+2}, \dots, X_n)$ استقلال دارند (طبق فرض مساله می‌گوییم مستقل از یکدیگرند)، بنابراین

$$E\big(S_k I(B_k)(S_n - S_k)\big) = E\big(S_k I(B_k)\big) E(S_n - S_k) = 0$$

با توجه به این‌که X ها را طوری فرض کردیم که در میانگین خود مرکزی شده باشد و از قبل هم می‌دانیم X زمانی در میانگین خود مرکزی می‌شود که $E(X) = 0$ باشد در نتیجه عبارت فوق برقرار است.

$$\int_{B_k} S_n^2 = \mathrm{E}\big(S_n \mathrm{I}(B_k)\big)^2 = \mathrm{E}\big(S_k \mathrm{I}(B_k) + (S_n - S_k)\mathrm{I}(B_k)\big)^2$$

از طرفی می‌دانیم که $\mathrm{I}^2(B_k) = \mathrm{I}(B_k)$ و همواره $|S_k| \ge \varepsilon$ در نتیجه

$$\int_{B_k} S_n^2 = \mathrm{E}\big(S_k^2 \mathrm{I}(B_k)\big) + \mathrm{E}\big((S_n - S_k)^2 \mathrm{I}(B_k)\big)$$

$$+2E[S_k \mathrm{I}(B_k)(S_n - S_k)] \ge E\big(S_k^2 \mathrm{I}(B_k)\big) \ge \varepsilon^2 p(B_k)$$

از طرفی

$$\sum P(B_k) E(S_n^2) = \int_{A_n} S_n^2 + \int_{A_n^c} S_n^2 \ge \int_{A_n^c} S_n^2 = \int_{\sum B_k} S_n^2$$

$$= \sum_k \int_{B_k} S_n^2 \ge \varepsilon^2 \sum P(B_k) = \varepsilon^2 P(A_n^c)$$

$$= \varepsilon^2 P\left\{\max_{j\ge n}\left|S_j\right| \ge \varepsilon\right\}$$

$$\Rightarrow P(A_n^c) = P\left\{\max_{j\ge n}\left|S_j\right| \ge \varepsilon\right\} \le E\left(\frac{S_n^2}{\varepsilon^2}\right)$$

و با توجه به

$$E\left(\frac{S_n^2}{\varepsilon^2}\right) = \frac{var(S^2) + \left(E(S^2)\right)^2}{\varepsilon^2} = \frac{var(S_n^2)}{\varepsilon^2} = \frac{\sum \sigma_k^2}{\varepsilon^2}$$

و

$$P\left\{\max_{k\ge n}|S_k| \ge \varepsilon\right\} \le \frac{\sum \sigma_k^2}{\varepsilon^2}$$

بنابراین سمت راست نامساوی اثبات شد.

حال برای اثبات سمت چپ داریم

$$S_{k-1}\mathrm{I}(A_{k-1}) + X_k\mathrm{I}(A_{k-1}) = S_k\mathrm{I}(A_{k-1})$$

$$= S_k\mathrm{I}(A_k) + S_k\mathrm{I}(B_k) \qquad (۱-۱۰)$$

اما همان‌طور که می‌دانیم (S_{k-1}, A_{k-1}) مستقل از X_k است. بنابراین داریم

$$E\{[S_{k-1}\mathrm{I}(A_{k-1})][X_k\ \mathrm{I}(A_{k-1})]\} = \mathrm{E}[S_{k-1}\ \mathrm{I}(A_{k-1})]E(X_k) = 0$$

و همچنین داریم

$$A_k \cap B_k = A_k \cap A_{k-1} \cap A_k^c = \emptyset$$

$$E\{[S_k\ \mathrm{I}(A_k)][S_k\ \mathrm{I}(B_k)]\} = \mathrm{E}[S_k\ \mathrm{I}(A_kB_k)] = 0$$

چون $|S_{k-1}| < \varepsilon,$ و X_k دارای کران $2c$ است. بنابراین

$$|S_k\mathrm{I}(B_k)| = |S_{k-1}\mathrm{I}(B_k) + X_k\mathrm{I}(B_k)| \le |\mathrm{\varepsilon} + 2\mathrm{c}|\mathrm{I}(B_k)$$

با توجه به رابطه (۱۰-۱)، داریم

$$E[S_{k-1}\mathrm{I}(A_{k-1})]^2 + \mathrm{E}[X_k\mathrm{I}(A_{k-1})]^2 = E[S_k\mathrm{I}(A_k)]^2 + \mathrm{E}[S_k\mathrm{I}(B_k)]^2$$

$$\le E[S_k\mathrm{I}(A_k)]^2 + |\mathrm{\varepsilon} + 2\mathrm{c}|^2 p(B_k) \qquad (۲-۱۰)$$

با جمع روی مقادیر k از ۱ تا n داریم

$$\sum_{k=1}^{n} E[S_{k-1}\mathrm{I}(A_{k-1})]^2 + \sum_{k=1}^{n} \mathrm{E}[X_k \mathrm{I}(A_{k-1})]^2$$
$$\leq \sum_{k=1}^{n} E[S_k \mathrm{I}(A_k)]^2 + \sum_{k=1}^{n} |\varepsilon + 2\mathrm{c}|^2 P(B_k)$$

بنابراین

$$\sum_{k=1}^{n} \mathrm{E}(X_k^2)\, p(A_{k-1}) \leq \sum_{k=1}^{n} E[S_k \mathrm{I}(A_k)]^2 + \sum_{k=1}^{n} |\varepsilon + 2\mathrm{c}|^2 P(B_k)$$

با توجه به اینکه $\forall k,\ P(A_k) < P(A_{k-1})$ داریم

$$\sum_{k=1}^{n} \mathrm{E}(X_k^2)\, P(A_k) \leq \sum_{k=1}^{n} \mathrm{E}(X_k^2)\, P(A_{k-1})$$
$$\leq \sum_{k=1}^{n} E[S_k \mathrm{I}(A_k)]^2 + \sum_{k=1}^{n} (\varepsilon + 2\mathrm{c})^2 P(B_k)$$

$$\sum_{k=1}^{n} \mathrm{E}(X_k^2)\, P(A_k) \leq \sum_{k=1}^{n} E[S_k\ \mathrm{I}(A_k)]^2 + \sum_{k=1}^{n} (\varepsilon + 2\mathrm{c})^2 P(B_k)$$

$$\sum_{k=1}^{n} \sigma_k^2\, P(A_n) \leq \sum_{k=1}^{n} E(S_k^2)\, P(A_k) + (\varepsilon + 2\mathrm{c})^2 P(\mathrm{A}_n^c)$$
$$\leq \varepsilon^2 P(A_n) + (\varepsilon + 2\mathrm{c})^2 P(\mathrm{A}_n^c) \leq (\varepsilon + 2\mathrm{c})^2$$

بنابراین

$$P(A_n) \leq \frac{(\varepsilon + 2\mathrm{c})^2}{\sum \sigma_k^2}$$

و اثبات کامل است.

۱۰-۲-۱ همگرایی تقریبا حتمی($a.s$) یک سری

قضیه۱۰-۲: اگر $\sum \sigma_k^2 < \infty$، آنگاه $\sum_n [X_n - E(X_n)]$ به طور تقریباً حتمی (a.s) همگراست، و اگر X ها a.s کراندار باشد، آن‌گاه عکس قضیه نیز برقرار است.

اثبات: اگر $\sum \sigma_k^2 < \infty$، آنگاه وقتی $\mathrm{m}, \mathrm{n} \to \infty$ ،$\sum_{m+1}^{n+m} \sigma_k^2 \to 0,$ حال Y_n را به صورت زیر تعریف می‌کنیم.

$$Y_n = \sum_{k=1}^{n} [X_k - E(X_k)]$$

بنابراین $Y_{n+m} - Y_m = \sum_{k=m+1}^{n+m} [X_k - E(X_k)]$.

حال برای اثبات همگرایی بطور تقریباً حتمی (a.s) بودن همگرایی Y_n کافی است نشان دهیم وقتی $m \to \infty$، آنگاه

$$P\{\cup_{k=1}^{\infty} [|Y_{m+k} - Y_m| \geq \varepsilon]\} \to 0.$$

با توجه به قضیه ۱-۱۰، $\sum \sigma_n^2 < \infty$، وقتی $m, n \to \infty$ داریم

$$P\left\{\bigcup_{k=1}^{n} [|Y_{m+k} - Y_m| \geq \varepsilon]\right\} = P\left\{\max_{1 \leq k \leq n} |Y_{m+k} - Y_m| \geq \varepsilon\right\}$$

$$\leq \frac{\sum_{m+1}^{n+m} \sigma_k^2}{\varepsilon^2} \to 0$$

حال برای اثبات عکس قضیه، فرض کرده‌ایم که X_n ها بطور یکنواخت دارای کرانی مانند c باشند.

پس $|X_n - E(X_n)|$ دارای کران $2c$ می‌باشد. با استفاده از قضیه (۱-۱۰)، داریم

$$1 - \frac{(\varepsilon + 2c)^2}{\sum \sigma_k^2} \leq P\left\{\max_{k \leq n} |Y_{m+k} - Y_m| \geq \varepsilon\right\} \to 0$$

$$\frac{(\varepsilon + 2c)^2}{\sum \sigma_k^2} \geq 1 \Rightarrow \sum \sigma_k^2 \leq (\varepsilon + 2c)^2 < \infty$$

و اثبات کامل می‌شود.

مثال۱۰-۱: اگر $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشند به طوری که

$$P\left(X_n = \frac{1}{n}\right) = \frac{1}{2} = P\left(X_n = \frac{-1}{n}\right), \quad n = 1,2, \ldots$$

آنگاه

$$E(X_n) = 0, \qquad \operatorname{var}(X_n) = \sigma_n^2 = \frac{1}{n^2}\,, \qquad \sum \sigma_n^2 = \sum \frac{1}{n^2} < \infty$$

پس می‌توان نتیجه گرفت که $\sum [X_n - E(X_n)] = \sum X_n$ به طور تقریبا حتمی (a.s) همگراست.

۱۰–۲–۲ معیار همگرایی تقریبا حتمی ($a.s$)

قصد داریم مجموعه شرایطی برای همگرایی تقریباً حتمی (a.s) تعیین کنیم. برای این کار به ایده‌ی متقارن‌سازی نیازمندیم.

تعریف ۱۰-۴: (**دنباله متقارن شده**)؛ اگر $\{X_n\}$ دنباله‌ای مستقل از متغیرهای تصادفی باشد و $\{X_n'\}$ دنباله‌ای دیگر از متغیرهای تصادفی مستقل از $\{X_n\}$ که با $\{X_n\}$ هم‌توزیع باشند. آنگاه $\{X_n^s\}$ به صورت زیر تعریف می‌شود

$$X_n^s = X_n - X_n'$$

چون X_n , X_n' هم‌توزیع هستند میانگین و واریانس یکسانی دارند و می‌توان نتیجه گرفت

$$E(X_n^s) = 0, \;\; Var(X_n^s) = Var(X_n) + Var(X_n') = 2\sigma_n^2$$

و اگر $|X_n| \le C$ آنگاه $|X_n^s| \le 2C$.

قضیه۱۰-۳: اگر X_n ها به طور یکنواخت کراندار و $\sum X_n$ به طور تقریباً حتمی (a.s) همگرا باشد آنگاه $\sum E(X_n)$, $\sum \sigma_n^2$ همگرا هستند.

اثبات: فرض می‌کنیم $\{X_n'\}$ و $\{X_n\}$ در شرایط متقارن‌سازی صدق کنند و هم‌چنین $\sum X_n'$ به طور تقریباً حتمی (a.s) همگرا باشد و $X_n^s = X_n - X_n'$ و $\sum X_n^s = \sum X_n - \sum X_n'$ بنابراین $\{X_n^s\}$ و $\sum X_n^s$ به طور a.s همگرا است. چون $\{X_n^s\}$ کراندار است و

$$\sum [X_n^s - E(X_n^s)] = \sum X_n^s < \infty$$

پس با توجه به عکس قضیه (۱۰-۲)، و کرانداری $\{X_n^s\}$ و همگرایی $\sum X_n^s$ داریم

$$\sum Var\, X_n^s < \infty \Rightarrow \sum 2\sigma_n^2 < \infty \Rightarrow \sum \sigma_n^2 < \infty$$

حال با توجه به این‌که $\sum \sigma_n^2 < \infty$، می‌توان نتیجه گرفت $\sum \sigma_n^2$ به طور a.s همگراست. طبق فرض قضیه $\sum_n [X_n - E(X_n)]$ نیز به طور a.s همگرا است، بنابراین

$$\sum_n E(X_n) = \sum X_n - \sum \big(X_n - E(X_n)\big)$$

هر دو عبارت سمت راست به طور تقریباً حتمی (a.s) همگرا هستند بنابراین $\sum E(X_n)$ همگرا است.

قضیه۱۰-۴. (قضیه سه سری کلموگروف) : فرض کنید X^c بریده شده X در $C > 0$ باشد. آنگاه سری $\sum X_n$ از متغیرهای تصادفی مستقل به طور تقریباً حتمی (a.s) به یک متغیر تصادفی همگراست اگر و فقط اگر به ازاء مقدار معین $C > 0$ سه سری زیر همگرا باشند.

$$(i) \sum P(|X_n| \geq C)$$

$$(ii) \sum \sigma^2{}_{(X_n^c)}$$

$$(iii) \sum E\,(X_n^c)$$

اثبات: ابتدا فرض می‌کنیم سه سری فوق همگرا باشند، می‌خواهیم ثابت کنیم $\sum X_n$ به طور تقریباً حتمی (a.s) همگراست. اگر (i) برقرار باشد، بنابراین $\sum P(|X_n| \geq C) = \sum P(X_n \neq X_n^c) < \infty$، در این صورت بنا بر لم(۱۰-۱)، $\sum X_n^c, \sum X_n$ معادل همگرایی هستند یعنی اگر $\sum X_n^c$ همگرا باشد، می‌توان نتیجه گرفت $\sum X_n$ نیز همگراست، بنابراین کافی است نشان دهیم که $\sum X_n^c$، تقریباً حتمی (a.s) همگراست.

اگر رابطه (ii) همگرا باشد، چون X_n^c به طور یکنواخت کراندار است در نتیجه بنا بر قضیه (۱۰-۲)،

$\sum\left(X_n^c - E(X_n^c)\right)$ به طور تقریباً حتمی (a.s) همگراست.

حال اگر (iii) همگرا باشد، آنگاه داریم

$$\sum E(X_n^c) < \infty,$$

$$\sum\left(X_n^c - E(X_n^c)\right) < \infty$$

$$\Rightarrow \sum X_n^c = \sum\left(X_n^c - E(X_n^c)\right) + \sum E(X_n^c) < \infty,$$

$\sum X_n^c$ به طور تقریباً حتمی (a.s) همگراست پس متقابلاً $\sum X_n$ نیز به طور a.s همگرا است.

اثبات عکس قضیه: بنابر فرض داریم $\sum X_n$ به طور تقریباً حتمی (a.s) همگرا است، در این صورت $X_n \xrightarrow{a.s} 0$ و بنابر نتیجه (۹-۷)، داریم اگر $X_n \xrightarrow{a.s} 0$، آنگاه $\sum P(|X_n| \geq C) < \infty$.

بنابراین سری (i) برای هر $C > 0$ همگرا می‌باشد. از همگرایی (i) نتیجه می‌شود $\sum X_n^c, \sum X_n$، معادل همگرایی هستند پس $\sum X_n^c$ نیز به طور تقریباً حتمی (a.s) همگرا است و چون X_n^c به طور یکنواخت کراندار است طبق قضیه (۱۰-۲)، نتیجه می‌گیریم $\sum \sigma^2{}_{(X_n^c)}$ همگرا است، پس (ii) نیز برقرار است.

حال طبق قضیه (۱۰-۲)، طبق متناهی بودن $\sum \sigma^2$ نتیجه می‌گیریم که به طور a.s، $\sum\left(X_n^c - E(X_n^c)\right)$ همگرا است بنابراین داریم

$$\sum E(X_n^c) = \sum X_n^c - \sum\left(X_n^c - E(X_n^c)\right) < \infty$$

۱۰–۳ پایداری سری های متغیرهای تصادفی مستقل

دنباله $\{S_n\}$ حاصل‌جمع‌های جزئی متغیرهای تصادفی مستقل $\{X_n\}$ را در احتمال یا a.s پایدار گوییم اگر برای دنباله‌هایی مانند $\{b_n\}$, $\{a_n\}$ وقتی $n \to \infty$ دنباله $\frac{S_n - a_n}{b_n}$ در احتمال یا a.s همگرا به صفر شود.

مثال۱۰-۲: اگر $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی مستقل برنولی با پارامتر p باشد و $S_n = \sum_{k=1}^{n} X_k$ آنگاه $E(S_n) = np$ و $Var(S_n) = npq$ و با استفاده از نامساوی چبیشف وقتی $n \to \infty$ نتیجه می‌شود،

$$P\left\{\left|\frac{S_n - np}{n}\right| > \varepsilon\right\} < \frac{pq}{n\varepsilon^2} \to 0$$

در نتیجه

$$\left|\frac{S_n - np}{n}\right| \xrightarrow{p} 0$$

بنابراین S_n پایدار است و این خاصیت پایداری را قانون اعداد بزرگ برنولی گویند.

توجه کنید که طبق قانون اعداد بزرگ برنولی

$$\left|\frac{S_n - np}{n}\right| \xrightarrow{p} 0 \Longrightarrow \frac{S_n}{n} \xrightarrow{p} p$$

از طرفی طبق نتیجه فصل قبل $\frac{S_n}{n} \xrightarrow{a.s} p$، که این خاصیت را قانون قوی اعداد بزرگ بورل گویند. به طور کلی دنباله $\{X_n\}$ از قانون قوی اعداد بزرگ پیروی می‌کند اگر

$$\frac{S_n - E(S_n)}{n} \xrightarrow{a.s} 0$$

لم۱۰-۳: لم کرونکر: اگر $\sum_{k=1}^{n} X_k = S_n \to S$ ، $b_n \uparrow \infty$، آنگاه $\frac{1}{b_n}\sum_{k=1}^{n} b_k X_k \to 0$

اثبات: فرض می‌کنیم $b_0 = 0, b_k - b_{k-1} = a_{k-1}$، در این صورت $\sum_{k=0}^{n-1} a_k = b_n$ پس داریم

$$\frac{1}{b_n}\sum_{k=1}^{n} b_k X_k = \frac{1}{b_n}\sum_{k=1}^{n} b_k\,(S_k - S_{k-1}) = \frac{1}{b_n}\left[\sum_{k=1}^{n} b_k\, S_k - \sum_{k=1}^{n} b_k\, S_{k-1}\right]$$

$$= \frac{1}{b_n}[b_1S_1 + \cdots + b_nS_n - b_1S_0 - b_2S_1 - \cdots - b_nS_{n-1}]$$

$$= \frac{1}{b_n}[-b_1S_0 + (b_1 - b_2)S_1 + (b_2 - b_3)S_2 + \cdots\ (b_{n-1} - \ b_n)S_{n-1} + b_nS_n]$$

$$= \frac{1}{b_n}[-a_0S_0 - a_1S_1 - \cdots - a_{n-1}S_{n-1}] + S_n$$

$$= \frac{-1}{b_n}\sum_{k=0}^{n-1} a_kS_k + S_n \qquad (3-10)$$

رابطه سمت راست را باز می‌کنیم

$$\frac{1}{b_n}\sum_{k=0}^{n-1} a_kS_k \ = \frac{1}{b_n}\sum_{k=0}^{n-1} a_kS + \ \frac{1}{b_n}\sum_{k=0}^{n-1} a_k(S_k - S)$$

$$= \frac{1}{b_n}\ S\sum_{k=0}^{n-1} a_k + \ \frac{1}{b_n}\sum_{k=0}^{n-1} a_k(S_k - S)$$

$$= \ \frac{b_n}{b_n}\ S + \frac{1}{b_n}\sum_{k=0}^{n-1} a_k(S_k - S)$$

$$= S + \frac{1}{b_n}\sum_{k=0}^{n-1} a_k(S_k - S)$$

$$\left|\sum_{k=1}^{\infty}\frac{a_k}{b_n}(S_k - S)\right| \le \sum_{k=1}^{N}\left|\frac{a_k}{b_n}\right||(S_k - S)| + \sum_{k=N+1}^{\infty}\left|\frac{a_k}{b_n}\right||(S_k - S)| \qquad (4-10)$$

از طرفی داریم، طبق فرض S_n به S همگراست بنابراین

$$\forall\varepsilon > 0 \text{ و } \exists k > N, \qquad s.t,\ \ |S_k - S| \le \frac{\varepsilon}{C}$$

پس رابطه دوم در سمت راست (۱۰-۴)، حداکثر ε می‌شود و وقتی $n \to \infty$ آنگاه $\frac{a_k}{b_n} \to 0$، بنابراین

$$\sum_{k=1}^{N}\left|\frac{a_k}{b_n}\right||(S_k-S)|\to 0$$

پس به طور کلی داریم

$$|\sum_{k=1}^{\infty}\frac{a_k}{b_n}(S_k-S)|\leq \varepsilon$$

در نتیجه $\sum_{k=1}^{\infty}\frac{a_k}{b_n}(S_k-S)\to 0$ و

$$\frac{1}{b_n}\sum_{k=0}^{n-1}a_kS_k = \ S+\sum_{k=0}^{n-1}\frac{a_k}{b_n}(S_k-S)\to S$$

حال با جایگذاری در رابطه (۱۰-۳)، داریم:

$$\frac{1}{b_n}\sum_{k=1}^{n}b_kX_k = \ \frac{-1}{b_n}\sum_{k=1}^{n-1}a_kS_k+S_n\to -S+S=0$$

نتیجه۱۰-۴: اگر $\sum_{k=0}^{n-1}a_k=b_n\uparrow\infty$, $S_n\to S$ ، آنگاه

$$\frac{1}{b_n}\sum_{k=0}^{n-1}a_kS_k\to S$$

نتیجه۱۰-۵: اگر $S_n\to S$ آنگاه $\frac{1}{n}\sum_{k=0}^{n-1}S_k\to S$

قضیه۱۰-۵: اگر X_n ها انتگرال‌پذیر و مستقل باشند و $\sum\left(\frac{\sigma_n^2}{b_n^2}\right)<\infty,\ b_n\uparrow\infty$ آنگاه

$$\frac{S_n-E(S_n)}{b_n}\xrightarrow{a.s}0$$

اثبات: اگر $\sum\left(\frac{\sigma_n^2}{b_n^2}\right)<\infty$ طبق قضیه (۱۰-۲)، $\sum\frac{X_k-E(X_k)}{b_k}$ ، به طور a.s همگرا است، پس برای تقریبا تمام ω ها سری $\sum\frac{X_k(\omega)-E(X_k)}{b_k}$ همگرا می‌باشد.

حال اگر قرار دهیم $X_k=\sum\frac{X_k(\omega)-E(X_k)}{b_k}$ بنا بر لم کرونکر خواهیم داشت

$$\frac{1}{b_n}\sum_{k=1}^{n}b_kX_k=\frac{1}{b_n}\sum_{k=1}^{n}\left(X_k(\omega)-E(X_k)\right)\to 0$$

$$\Rightarrow \frac{1}{b_n}\sum_{k=1}^{n}\left(X_k(\omega)-E(X_k)\right)\xrightarrow{a.s}0$$

$$\Rightarrow \frac{S_n-\sum_{k=1}^{n}E(X_k)}{b_n}=\frac{S_n-E(S_n)}{b_n}\xrightarrow{a.s}0$$

قضیه۱۰-۶: قانون قوی اعداد بزرگ حالت i.i.d (قانون قوی اعداد بزرگ کلموگروف):

اگر X_i ها متغیرهای تصادفی مستقل و هم‌توزیع باشند آنگاه $\frac{S_n}{n}\xrightarrow{a.s}c$ و c عددی متناهی است اگر و تنها اگر $E|X|<\infty$. در این صورت $c=E|X|$.

اثبات: ابتدا فرض می‌کنیم $\frac{S_n}{n}\xrightarrow{a.s}c<\infty$، می‌خواهیم ثابت کنیم $E|X|<\infty$.

تعریف می‌کنیم

$$A_n=\{|X|\geq n\},\qquad (n=0,1,2,\dots),\quad A_0=\Omega$$

و دنباله A_n نزولی است، لذا $A_{n+1}\subset A_n$ در نتیجه $\cup A_i=A_0$، هم‌چنین تعریف می‌کنیم

$$B_n=A_n-A_{n+1},\quad \cup B_n=\cup A_n$$

$$\sum_{n=0}^{\infty}B_n=B_0+B_1+\cdots=A_0-A_1+A_1-A_2+\cdots=A_0=\Omega$$

با انتگرال‌گیری نسبت B_n ازطرفین $n\leq|X|<n+1$

$$n\,P(B_n)\leq\int_{B_n}|X|\leq(n+1)P(B_n)$$

با جمع روی n داریم

$$\sum_{n=1}^{\infty}n\,P(B_n)\leq E|X|\leq\sum_{n=1}^{\infty}(n+1)P(B_n)$$

$$\sum_{n=0}^{\infty}n\,P(B_n)=\sum_{n=0}^{\infty}n\,P(A_n-A_{n+1})=0+P(A_1-A_2)+2P(A_2-A_3)+\cdots$$

$$=P(A_1)+P(A_2)+\cdots=\sum_{n=1}^{\infty}P(A_n)$$

و هم‌چنین

$$\sum_{n=1}^{\infty} P(A_n) \leq E|X| \leq \sum_{n=1}^{\infty} P(A_n) + \sum_{n=1}^{\infty} P(B_n)$$
$$= \sum_{n=1}^{\infty} P(A_n) + P(\Omega) = \sum_{n=1}^{\infty} P(A_n) + 1 \qquad (5-10)$$

و طبق فرض داریم $\frac{S_n}{n} \xrightarrow{a.s} c < \infty$ بنابراین

$$\frac{X_n}{n} = \frac{S_n}{n} - \left(\frac{n-1}{n}\right)\frac{S_{n-1}}{n-1} \xrightarrow{a.s} 0$$

بنا بر نتیجه (۹-۵)، که اگر $X_n \to 0$ آنگاه $\sum P(|X_n| \geq n) < \infty$. بنابراین $\sum_{n=1}^{\infty} P(A_n) < \infty$،

در نتیجه $E|X| < \infty$

حال برای اثبات عکس قضیه داریم، اگر $E|X| < \infty$ از رابطه (۱۰-۵)، نتیجه می‌شود $\sum_{n=1}^{\infty} P(A_n) < \infty$ و فرض که

$$X_k^k = \begin{cases} X_k & |X_k| < k \\ 0 & |X_k| \geq k \end{cases}$$

که همان مفهوم برش است و $\overline{S_n} = \sum_{n=0}^{\infty} X_k^k$ چون X_n ها هم‌توزیع اند

$$\sum_{k=1}^{\infty} P(X_k \neq X_k^k) = \sum_{k=1}^{\infty} P(|X_k| \geq k) = \sum_{k=1}^{\infty} P(A_k) < \infty$$

بنابراین $\frac{S_n}{n}$ و $\frac{\overline{S_n}}{n}$ معادل همگرایی هستند و کافی است ثابت کنیم

$$\frac{\overline{S_n}}{n} \xrightarrow{a.s} E(X)$$

حال چون $X_n^n \to X$ ، $|X_n^n| \leq X$ و X انتگرال‌پذیر است بنا بر قضیه همگرایی مغلوب وقتی $n \to \infty$ داریم $E(X_n^n) \to E(X)$.

و بنا بر نتیجه (۱۰-۵)، $E\left(\frac{\overline{S_n}}{n}\right) = \frac{1}{n}\sum E(X_k^k) \to E(X)$

پس کافی است ثابت کنیم که $\frac{\overline{S_n} - E(\overline{S_n})}{n} \xrightarrow{a.s} 0$.

با توجه به قضیه (۱۰-۵)، کافی است ثابت کنیم که $\sum_{k=1}^{\infty} \frac{\sigma^2\left(X_k^k\right)}{k^2} < \infty$، از طرفی

$$\sigma^2(X_k^k) \le E(X_k^k)^2 = \int_{-k}^{k} X^2 \, dF(X)$$

و داریم

$$\sum_{k=j}^{\infty} \frac{1}{k^2} \le \sum_{k=j+1}^{\infty} \frac{1}{k(k-1)} = \frac{1}{j}$$

$$\sum_{k=1}^{\infty} \frac{\sigma^2\left(X_k^k\right)}{k^2} \le \sum_{k=1}^{\infty} \frac{E\left(X_k^k\right)^2}{k^2}$$

$$\le \sum_{k=1}^{\infty} \frac{E\left(X^2 \mathrm{I}(|X| < k)\right)}{k^2} = \sum_{k=1}^{\infty} \frac{\int_{-k}^{k} X^2 \, dF(X)}{k^2}$$

$$\sum_{k=1}^{\infty} \frac{\int_{-k}^{k} X^2 \, dF(X)}{k^2} = \sum_{k=1}^{\infty} \frac{\left[\int_{-k}^{0} X^2 \, dF(X) + \int_{0}^{k} X^2 \, dF(X)\right]}{k^2}$$

$$= \sum_{k=1}^{\infty} \sum_{j=1}^{k} \frac{\left[\int_{j-1}^{j} X^2 \, dF(X) + \int_{-j}^{1-j} X^2 \, dF(X)\right]}{k^2}$$

$$= \sum_{j=1}^{k} \sum_{k=j}^{\infty} \frac{\left[\int_{j-1}^{j} X^2 \, dF(X) + \int_{-j}^{1-j} X^2 \, dF(X)\right]}{k^2}$$

$$= \sum_{j=1}^{\infty} \left[\int_{j-1}^{j} X^2 \, dF(X) + \int_{-j}^{1-j} X^2 \, dF(X)\right] \sum_{k=j}^{\infty} \frac{1}{k^2}$$

$$\le \sum_{j=1}^{\infty} \left[\int_{j-1}^{j} X^2 \, dF(X) + \int_{-j}^{1-j} X^2 \, dF(X)\right] \frac{1}{j}$$

$$= \sum_{j=1}^{\infty} \int X^2 \, I(j-1 < |X| < j) \, dF(X) \frac{1}{j}$$

از آنجایی که $|X| < j$ $\quad X^2 = |X||X| < j|X|,$

$$\le \sum_{j=1}^{\infty} j \int |X| \, I(j-1 < |X| < j) \, dF(X) \frac{1}{j}$$

$$= \sum_{j=1}^{\infty} \int |X| \, I(j-1 < |X| < j) \, dF(X) = E|X| < \infty$$

و اثبات کامل است.

تعمیم قضیه۱۰-۶: فرض کنید که $\{X_n\}$ دنباله‌ای $i.i.d$ است. در این صورت

۱) اگر $E(|X|^r) < \infty$ آنگاه

$$n^{-1/r}\sum X_k \to 0 \quad (a.s), \qquad (0 < r < 1)$$

و بالعکس،

۲)اگر $E(|X|^r) < \infty$ آنگاه

$$n^{-1/r}\sum \left(X_k - E(X_k)\right) \to 0 \quad (a.s), \qquad (1 \le r < 2)$$

وبالعکس.

۱۰-۴ قانون ضعیف اعداد بزرگ

۱۰-۴-۱ قانون ضعیف اعداد بزرگ حالت*i.i.d*

همان طور که می‌دانیم همگرایی در احتمال را می‌توان از همگرایی با احتمال یک نتیجه گرفت. اما برای اینکه به‌طور مستقیم به قانون ضعیف اعداد بزرگ دست پیدا کنیم، داریم

$$P[|n^{-1}S_n - E(X)| \ge \varepsilon] \le \frac{Var(S_n)}{n^2\varepsilon^2} = \frac{n\,Var(X)}{n^2\varepsilon^2}$$

۱) اگر $X_1, X_2, \ldots$ به قسمی باشند که

$$\frac{Var(\sum X_k)}{n^2} \to 0$$

آنگاه وقتی $n \to \infty$ می‌توان ثابت کرد

$$P[|\bar{X}_n - E(\bar{X}_n)| \ge \varepsilon] \to 0$$

نکته ۱۰-۱: قانون ضعیف اعداد بزرگ وقتی $n \to \infty$ برای حالتی که متغیرها هم‌توزیع نباشند، به صورت زیر است

$$P\left[\left|\frac{S_n - E(S_n)}{n}\right| > \varepsilon\right] \le \frac{Var(S_n)}{n^2\varepsilon^2} \to 0$$

یعنی اگر انحراف معیار $\frac{S_n}{n}$ به سمت صفر میل کند، آن‌گاه قانون ضعیف برقرار خواهد بود و با انتخاب متفاوت $\{a_n\}$ و $\{b_n\}$ حدود متفاوتی برای $\frac{S_n - a_n}{b_n}$ می‌توان به دست آورد.

مثلاً با انتخاب $a_n = E(S_n)$ و $b_n = \sqrt{Var(S_n)}$ ثابت می‌شود که $\frac{S_n - a_n}{b_n}$ به نرمال همگرا است که به آن قضیه حد مرکزی گویند. تحت شرایط معین می‌توان a_n , b_n را به قسمی انتخاب کرد که

$$P\left(\limsup\left(\frac{S_n - a_n}{b_n}\right) = 1\right) = 1 = P\left(\liminf\left(\frac{S_n - a_n}{b_n}\right) = -1\right)$$

برای متغیرهای تصادفی برنولی این رابطه به ازای

$$b_n = \sqrt{npq(2\log\log n)}, \qquad a_n = np$$

برقرار است که در آن $\log$ در مبنای دو در نظر گرفته شده است. در این صورت نتیجه را قانون لگاریتم‌های مکرر خین‌چین می‌نامند.

۱۰-۵ مسائل

۱) به ازای مقادیر زیر نشان دهید $\sum \sigma_n^2$ ممکن است نامتناهی باشد ولی $\sum X_n$ ممکن است a.s همگرا باشد.

$$P(X_n = -n) = \frac{1}{2n^2}, \qquad P(X_n = n) = \frac{1}{2n^2}, \qquad P(X_n = 0) = 1 - \frac{1}{n^2}$$

۲) همگرایی در احتمال معادل است با:

$$E\left\{\frac{(X_n - X)^2}{1 + (X_n - X)^2}\right\} \to 0$$

اگر و فقط اگر:

$$E\left\{\frac{\left[\sum\left(X_k - E(X_k)\right)\right]^2}{n^2 + \left[\sum\left(X_k - E(X_k)\right)\right]^2}\right\} \to 0$$

برای متغیرهای تصادفی مستقل اگر و فقط اگر سه شرط زیر برقرار باشند:

الف) $\sum_{k=1}^{n} P(|X_k| > n) \to 0$

ب) $\frac{1}{n}\sum_{k=1}^{n} E(X_k^n) \to 0$

ج) $\frac{1}{n^2}\sum_{k=1}^{n} \sigma^2(X_k^n) \to 0$

که در آن X_k^n بریده شده متغیر X_k در n می باشد

تعیین کنید آیا قانون ضعیف اعداد بزرگ برای دنباله‌های متغیرهای تصادفی مستقل زیر برقرار است:

الف) $p(X_n = 2^n) = \frac{1}{2} = p(X_n = -2^n)$

ب) $p(X_n = n) = \frac{1}{2}n^{-\lambda} = p(X_n = -n), \quad p(X_n = 0) = 1 - n^{-\lambda}$

۳) اگر X_2 و X_1 به قسمی باشند که (حتی فرض نمی‌کنیم دوبه دو مستقل باشند) ، وقتی و

$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n}(X_k)$ آنگاه با استفاده از نابرابری چبیشف می‌توان ثابت کرد

$$p(|\bar{X}_n - E(\bar{X}_n)| \geq \varepsilon) \to 0$$

۴) اگر دنباله‌ای از متغیرهای دوبه دو مستقل به قسمی باشد که

$$Var(X_i) \leq c, \quad E(X_i) = a, \quad (i = 1,2, \dots)$$

آنگاه به ازای هر عدد ثابت وقتی

$$p(|\bar{X}_n - a| < \varepsilon) \to 1, \quad \bar{X}_n = \frac{1}{n}\sum_{k=1}^{n}(X_k)$$

فصل یازدهم

قضیه حد مرکزی

۱۱–۱ حالت برنولی

فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی برنولی باشد که مقدار یک را با احتمال p و صفر را با احتمال $q = p - ۱$ اختیار کند آنگاه

$$E(X_k) = p, \qquad var(X_k) = pq, \tag{۱۱-۱}$$

$$S_n = \sum_{k=1}^{n} X_k \to E(S_n) = np, \quad , \quad var(S_n) = npq \tag{۱۱-۲}$$

در این صورت $S_n^* = \frac{\mathrm{S_n - np}}{\sqrt{\mathrm{npq}}}$ یک متغیر تصادفی استانداردشده با میانگین صفر و واریانس یک می‌باشد. دموآور ثابت نمود که وقتی n به سمت بی‌نهایت میل می‌کند آنگاه

$$p(S_n^* \le x) \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} dy = \varphi(x), \quad -\infty < x < \infty \quad (3 - 11)$$

یعنی $S_n^* \xrightarrow{d} N(0,1)$.

۱۱– ۲ CLT به عنوان تعمیمی برای L. L. N

اگر $X_k = \frac{S_n}{n}$ $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n}$, $\sigma(S_n) = Sd(S_n) = s_n$ آنگاه

$$\frac{\bar{X}_n - E(\bar{X}_n)}{\sigma(\bar{X}_n)} = \frac{S_n - E(S_n)}{s_n} = S_n^*$$

یک متغیر استاندارد با میانگین صفر و واریانس واحد است. اگر X_k ها مستقل و هم‌توزیع باشند، آنگاه،

$var(S_n) = n\sigma^2$ که در آن $\sigma^2 = var(X_k)$ و در این‌صورت خواهیم داشت

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad (۴-۱۱)$$

و اگر بپذیریم $p(S_n^* \le x) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy = \varphi(x)$ آنگاه

$$p(|S_n^*| \le x) \to \frac{1}{\sqrt{2\pi}}\int_{-x}^{x} \mathrm{e}^{-\frac{y^2}{2}}\,dy = \varphi(x) - \varphi(-x) \qquad (۵-۱۱)$$

و یا به‌طور معادل خواهیم داشت

$$p\left(|\bar{X}_n - \mu| \le x\frac{\sigma}{\sqrt{n}}\right) \to \frac{1}{\sqrt{2\pi}}\int_{-x}^{x} \mathrm{e}^{-\frac{y^2}{2}}\,dy$$
$$= \varphi(x) - \varphi(-x) \qquad (۶-۱۱)$$

حال با فرض این‌که $x\frac{\sigma}{\sqrt{n}} = \varepsilon$ یعنی $x = \frac{\varepsilon\sqrt{n}}{\sigma}$ ، وقتی که $n \to \infty$، داریم

$$p(|\bar{X}_n - \mu| \le \varepsilon) \to 1 \qquad (۷-۱۱)$$

یعنی $\bar{X}_n \xrightarrow{p} \mu$، که همان قانون ضعیف اعداد بزرگ است، بنابراین CLT کران احتمالی برای $|\bar{X}_n - \mu|$ معین می‌کند در صورتی‌که قانون اعداد بزرگ فقط مقدار حدی آن را نشان می‌دهد.

۱۱– ۳ حالت I. I. D

۱۱–۳–۱ واریانس متناهی

فرض کنید $\{X_k\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع $(i.i.\mathrm{d})$ باشد و هم‌چنین داشته باشیم

$$E(X_k) = \mu, \qquad \sigma(X_n) = \sigma, \qquad \sigma(S_n) = \sqrt{n}\sigma$$

می‌خواهیم ثابت کنیم که $S_n^* = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k)}{\sqrt{n}\sigma}$ دارای توزیع نرمال استاندارد می‌باشد.

قضیه۱۱-۱: (لیندبرگ-لوی)، فرض کنید $\{X_k\}$ دنباله‌ای از متغیرهای تصادفی $(i.i.\mathrm{d})$ باشند به قسمی که $E(X_k) = \mu < \infty, \quad \sigma(X_k) = \sigma < \infty$،

قرار دهید $S_\mathrm{n} = \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{X}_k$ در این صورت وقتی $n \to \infty$ داریم

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(\cdot,1)$$

اثبات: طبق قضیه پیوستگی توابع مشخصه می‌دانیم $X_n \to X \Leftrightarrow \varphi_n \to \varphi$.

بنابراین کافی است ثابت کنیم که تابع مشخصه S_n^* به تابع مشخصه توزیع نرمال استاندارد یعنی $exp\left\{-\frac{u^۲}{۲}\right\}$ همگرا است، با در نظر گرفتن $Y_k = \frac{X_k-\mu}{\sigma}$ داریم

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}}\sum_{k=1}^n\left(\frac{X_k-\mu}{\sigma}\right) = \frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k \qquad (۸\text{-}۱۱)$$

از طرفی می‌دانیم که $E(Y_k) = \cdot,\ \mathrm{var}(Y_k) = ۱$

$$\varphi_{S_n^*}(t) = \varphi_{\frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k}(t) = \varphi_{\sum_{k=1}^n Y_k}\left(\frac{t}{\sqrt{n}}\right)$$

$$= \prod_{k=1}^n \varphi_{Y_k}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n \qquad (۹\text{-}۱۱)$$

می دانیم اگر $E(|X|^r) < \infty$ آنگاه:

$$\varphi(t) = ۱ + \sum_{k=۱}^r \frac{(it)^k}{k!} E(X^k) + O(|t|^r)$$

که در آن $O(\cdot)$ تابعی است از t که در هر بازه متناهی $(-a\,,a)$ دارای قدرمطلق کراندار است و وقتی $t \to \cdot$، به صفر همگرا است.

حال چون $E(Y_k^2) < \infty$ ، برای $r = 2$ داریم

$$\varphi_{S_n^*}(t) = \left[1 + \sum_{k=1}^{2}\frac{\left(\frac{it}{\sqrt{n}}\right)^k}{k!}E(X^k) + O\left(\left|\frac{t}{\sqrt{n}}\right|^2\right)\right]^n$$

$$= \left[1 - \frac{t^2}{2n} + O\left(\frac{t^2}{n}\right)\right]^n \to \left[\exp\left(\frac{-t^2}{2}\right)\right]^n \qquad (۱۰\text{-}۱۱)$$

بنابراین حکم ثابت شد.

مثال۱۱-۱: فرض کنید ... $X_۱, X_2,$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع $(i.i.\mathrm{d})$ نمایی با میانگین λ باشند و $Y_n = \sum_{k=1}^n X_k$، می‌دانیم که Y_n دارای توزیع گاما با میانگین $n\lambda$ و واریانس $n\lambda^2$ می‌باشد، با این فرض $\frac{Y_n - n\lambda}{\lambda\sqrt{n}}$ دارای توزیع مجانبی نرمال استاندارد می‌باشد.

۱۱-۳-۲ حالتی که n تصادفی است

قضیه ۱۱-۲: فرض کنید $\{X_k\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع با میانگین صفر و واریانس یک باشد. هم‌چنین با فرض این‌که $\{N_k\}$ دنباله‌ای از متغیرهای تصادفی با مقادیر صحیح و مثبت باشد، به قسمی که $\frac{N_k}{k}\xrightarrow{p} c\,,\ (0<c<\infty)$ در این صورت $\frac{s_{N_K}}{\sqrt{N_K}}$ در توزیع به یک متغیر نرمال استاندارد میل می‌کند وقتی که $s_{N_K}=\sum_{i=1}^{N_k}X_i$ می باشد.

اثبات: با استفاده از امید ریاضی شرطی داریم

$$E(S_{N_k})=E(\sum_{i=1}^{N_k}X_i)=E(E(\sum_{i=1}^{N_k}X_i|N_k))$$

$$=E\left(\sum_{i=1}^{N_k}E(X_i)\right)$$

$$=E\Big((N_k)\left(\cdot\right)\Big)=E\left(\cdot\right)=\cdot \tag{۱۱-۱۱}$$

و همچنین برای واریانس داریم

$$var(S_{N_k})=var\left(\sum_{i=1}^{N_k}X_i\right)=E(var(\sum_{i=1}^{N_k}X_i|N_k))+var(E(\sum_{i=1}^{N_k}X_i|N_k))$$

$$=E\left(\sum_{i=1}^{N_k}var(X_i)\right)$$

$$=E(N_k)=N_k \tag{۱۲-۱۱}$$

$$\frac{S_{N_k}-E(S_{N_k})}{\sigma(S_{N_k})}=\frac{S_{N_k}}{\sqrt{N_k}}\to N\left(\cdot,1\right)$$

۱۱–۴ توزیع های متغیر

۱۱–۴–۱ صورت لیاپانفσ_k^2

در این بخش فرض می‌کنیم که $\{X_k\}$ها دنباله‌ای از متغیرهای تصادفی مستقل باشند به‌طوری‌که

$var(\mathrm{X}_k) = \sigma_k^2 < \infty$، در این صورت:

$$var(S_n) = \sum_{k=1}^{n} \sigma_k^2 = s_n^2$$

که آن را متناهی فرض کرده‌ایم تحت شرایط فوق فرض می‌کنیم گشتاورهای مطلق مرتبه (۲+δ)ام متغیرهای X_k متناهی هستند.

قضیه۱۱-۳: (لیاپانف)؛ اگر برای مقدار معینی از δ، که $0 < \delta \leq 1$ وقتی $n \to \infty$ و داشته باشیم

$$\frac{\sum_{k=1}^{n} E(|X_k|^{2+\delta})}{S_n^{2+\delta}} \to 0 \qquad (۱۴\text{-}۱۱)$$

آنگاه $\frac{S_n}{\mathrm{s_n}} \xrightarrow{\mathrm{d}} \mathrm{N}(\cdot,1)$ که اینجا $\mathrm{s_n}=\sum_{k=1}^{n}\sigma_k^2$, $S_n = \sum_{k=1}^{n} X_k$

اثبات: می‌دانیم که X_k انتگرال‌پذیر است اگر و فقط اگر $|X_k|$ انتگرال‌پذیر باشد یا به طور معادل

$$E(X_k) < \infty \Leftrightarrow E(|X_k|) < \infty$$

طبق فرض $E(X_k) = \cdot < \infty$ بنابراین $E(|X_k|) < \infty$ و چون $\sigma^2{}_k > \cdot$ پس X_k غیرتباهیده است. همچنین از رابطه (۱۱-۱۴)، نتیجه می‌شود وقتی $n \to \infty$ آنگاه $S_n \to \infty$

طبق نامساوی لیاپانف $[E(|X_k|^r)]^{\frac{1}{r}}$ برحسب r صعودی است و چون $0 < \delta \leq 1$ پس $۲ + \delta > ۲$ ، بنابراین داریم

$$\sigma_k = \mathrm{E}^{\frac{۱}{۲}}(|X_k|^{۲}) \leq \mathrm{E}^{\frac{۱}{۲+\delta}}(|X_k|^{۲+\delta}) \Rightarrow \sigma_k{}^{۲+\delta} \leq \mathrm{E}(|X_k|^{۲+\delta})$$

در نتیجه

$$Max_{k\leq n}(\frac{\sigma_k}{\mathrm{S_n}})^{۲+\delta} \leq Max_{k\leq n}\frac{\mathrm{E}|X_k|^{۲+\delta}}{\mathrm{S_n}{}^{۲+\delta}} \leq \frac{۱}{\mathrm{S_n}{}^{۲+\delta}}\sum \mathrm{E}(|X_k|^{۲+\delta})$$

اگر $\varphi_k(t)$ تابع مشخصه X_k باشد، تابع مشخصه $\frac{S_n}{s_n}$ به صورت زیر می‌باشد:

$$\varphi_{\frac{S_n}{s_n}}(t) = \varphi_{S_n}\left(\frac{t}{s_n}\right) = \prod_{k=1}^{n} \varphi_k\left(\frac{t}{s_n}\right) \qquad (۱۱-۱۵)$$

از طرفی می‌دانیم که اگر گشتاورهای مطلق مرتبه $(r+\delta)$ ام متناهی باشند و $0 < \delta \leq 1$ آنگاه:

$$\varphi_x(t) = \sum_{s=0}^{r-1} \frac{(it)^s}{s!} E(X^s) + \frac{(it)^r}{r!} E(X^r)000 \ + ۲^{۱-\delta} \theta \frac{|t|^{r+\delta} E\left(|X|^{r+\delta}\right)}{\left(۱+\delta\right)\left(۲+\delta\right)\ldots\left(r+\delta\right)}$$

پس برای $r = ۲$ خواهیم داشت

$$\varphi_k\left(\frac{t}{s_n}\right) = ۱ - \frac{t^۲}{۲}\frac{\sigma_k{}^۲}{s_n{}^۲} + \frac{\theta}{\left(۱+\delta\right)\left(۲+\delta\right)} ۲^{۱-\delta} \frac{E(|X_k|^{۲+\delta})}{s_n{}^{۲+\delta}} |t|^{۲+\delta}$$

که به ازای $k \leq n$ بطور یکنواخت به سمت یک میل می‌کند، از طرفی برای مقادیر بزرگ n داریم

$$\log\left(1+z\right) = \int \frac{1}{1+z} dz$$

$$= \int \sum_{k=0}^{\infty} (-z)^k dz = \int \left(۱ - Z + Z^۲ - Z^۳ + \cdots\right) dz$$

$$= z - \frac{z^۲}{۲} + \frac{z^۳}{۳} - \cdots = z\left(۱ - \frac{z}{۲} + \frac{z^۲}{۳} - \cdots\right) = z\left(1 + O\left(1\right)\right)$$

بنابراین

$$log\varphi_k\left(\frac{t}{S_n}\right) = -\frac{t^۲}{۲}\frac{\sigma_k^2}{S_n^2}\left(۱ + O\left(۱\right)\right) + ۲\theta \ \frac{E(|X_k|^{۲+\delta})}{S_n^{2+\delta}} |t|^{۲+\delta}$$

که در آن $O(1)$, θ بستگی به k ندارند. بنابراین، وقتی $n \to \infty$ داریم

$$\sum_{k=۱}^{n} log\varphi_k\left(\frac{t}{S_n}\right) = -\frac{t^۲}{۲}\left(۱ + O\left(۱\right)\right) + ۲\theta \ |t|^{۲+\delta} \frac{\sum_{k=۱}^{n} E(|X_k|^{۲+\delta})}{S_n^{2+\delta}} \to -\frac{t^۲}{۲}$$

$$\prod_{k=1}^{n} \varphi_k(\frac{t}{s_n}) \to \exp\{-\frac{t^۲}{۲}\}$$

نتیجه۱۱-۱: اگر شرط لیاپانف به ازای مقدار معیّنی از $\delta > 0$ صادق باشد، قضیه (۱۱-۳)، برقرار خواهد ماند.

اثبات: نشان خواهیم داد که اگر رابطه (۱۱-۱۴)، به ازای مقدار معیّنی از $\delta \geq 1$ صادق باشد، آنگاه به ازای $\delta = 1$ نیز برقرار است. فرض می‌کنیم Y یک متغیر تصادفی با تابع توزیع $\frac{1}{n}\sum_{K=1}^{n} F_k$ باشد.

$$E(|Y|^r) = \frac{1}{n}\sum_{k=1}^{n} \int |Y|^r \ dF_k = \frac{1}{n}\sum_{k=1}^{n} E(|X_k|^r) \qquad (۱۷\text{-}۱۱)$$

$logE(|Y|^r)$ تابعی محدب از r، ($r >۰$) است،

با توجه به تعریف تابع محدب و با انتخاب $\lambda = \frac{\delta-1}{\delta}$ $\quad X_1 = 2, \quad X_2 = 2+\delta,$ داریم

با استفاده از نامساوی جنسن

$$logE\left(|Y|^۳\right) = logE\left\{|Y|^{۲\left(\frac{\delta-۱}{\delta}\right)+\frac{\delta+۲}{\delta}}\right\} \leq \frac{\delta-۱}{\delta} logE\left(|Y|^۲\right) + \frac{۱}{\delta} logE(|Y|^{۲+\delta})$$

زیرا $3 = \frac{\delta+۲}{\delta} + \frac{\delta-۱}{\delta}۲$ ولی $E\left(|Y|^۲\right) = \frac{S_n^2}{n}$. با لگاریتم معکوس‌گیری و حذف n از طرفین داریم

$$\sum_{k=۱}^{n} E(|X_k|^۳) \leq (S_n^2)^{\frac{\delta-۱}{\delta}} \sum_{k=۱}^{n} \left(E\left(|X_k|^{\delta+۲}\right)\right)^{\frac{۱}{\delta}}$$

بنابراین

$$\frac{1}{S_n^3}\sum_{k=۱}^{n} E(|X_k|^۳) \leq \frac{1}{(s_n)^{\frac{\delta+2}{\delta}}} \sum_{k=1}^{n} \left(E\left(|X_k|^{\delta+۲}\right)\right)^{\frac{۱}{\delta}} \qquad (18-11)$$

سمت چپ نامساوی فوق همان شرط لیاپانف به ازای $\delta = 1$ است پس شرط لیاپانف به ازای $\delta \geq 1$ همان نتیجه را برای $\delta = 1$ به‌دست می‌دهد. پس می‌توان فرض کرد $0 < \delta \leq 1$.

شرط لازم و کافی قضیه قبل در قضیه بعدی آمده است.

۱۱-۴-۲ قضیه لیندبرگ- لوی- فلر

$$\frac{S_n}{s_n} \to N(۰,۱) \quad , \quad Max_{k\leq n}(\frac{\sigma_k}{s_n}) \to ۰$$

اگر و فقط اگر به ازای هر $۰ < \varepsilon$

$$g_n(\varepsilon) = \frac{1}{S_n^2}\sum_{k=1}^{n}\int_{|x|>\varepsilon s_n} X^۲ \, dF_k \to 0 \qquad (19-11)$$

اثبات شرط کافی: ابتدا نشان می‌دهیم $Max_{k\leq n}(\frac{\sigma_k}{s_n}) \to 0$

$$\sigma^۲ = \int_{|x|<\varepsilon s_n} X^۲ \, dF_k + \int_{|x|>\varepsilon s_n} X^۲ \, dF_k \leq \varepsilon^۲ {s_n}^۲ + \int_{|x|>\varepsilon s_n} X^۲ \, dF_k$$

$$\Rightarrow Max_{k\leq n}\left(\frac{{\sigma_k}^۲}{{s_n}^۲}\right) \leq \varepsilon^۲ + \frac{۱}{{s_n}^۲}\sum\int_{|x|>\varepsilon s_n} X^۲ \, dF_k \Rightarrow Max_{k\leq n}(\frac{\sigma_k}{s_n}) \to ۰$$

همچنین با استفاده از رابطه (۱۱-۱۹)، می‌توان نتیجه گرفت

$$\limsup \ \max_{k\leq n}(\frac{\sigma_k}{s_n}) \leq \varepsilon$$

وقتی که$n \to \infty$ که در آن ε دلخواه می‌باشد.

دنباله زیر را تعریف می‌کنیم

$$X_{nk} = \begin{cases} X_k & |X_k| \leq \varepsilon_n s_n \\ 0 & |X_k| > \varepsilon_n s_n \end{cases}$$

که در آن $۰ \downarrow \varepsilon$ به قسمی که $۰ \to {\varepsilon_n}^{-۲} g_n(\varepsilon)$، با فرض $s_{nn} = \sum_{k=۱}^{n} X_{nk}$ داریم

$$p\left\{\frac{s_{nn}}{n} \neq \frac{s_n}{n}\right\} \leq \sum_{k=۱}^{n} p\{X_{nk} \neq X_K\} = \sum_{k=۱}^{n}\int_{|x|>\varepsilon_n s_n} X^۲ \, dF_k$$

$$\leq (۱/{\varepsilon_n}^۲ \, {s_n}^۲)\sum_{K=۱}^{n}\int_{|X|\geq\varepsilon_n s_n} X^۲ dF_K = \frac{g_n(\varepsilon_n)}{{\varepsilon_n}^۲} \to ۰$$

در نتیجه وقتی که $n \to \infty$، $\frac{s_{nn}}{n}$ و $\frac{S_n}{n}$ در قانون دارای یک حد یکسان می‌باشند.

حال کافی است، نشان دهیم که$\frac{s_{nm}}{s_n} \xrightarrow{d} N(۰,۱)$.

$$S_{nn}^2 = \sum_{K=۱}^{n} var(X_{nk}) \le \sum_{K=۱}^{n} \left(E|X_{nk}|^۲\right) = \sum_{K=۱}^{n} \int_{|X| \le \varepsilon_n s_n} X^۲ dF_K$$

$$= S_n^2 - \sum_{k=۱}^{n} \int_{|X| > \varepsilon_n s_n} X^۲ \ \mathrm{d}F_K$$

$$\frac{\mathrm{s_{nn}}}{s_n} \to 1$$

ازطرفی می‌دانیم که

$$|X_{nk} - E(X_{nk})| \le |X_{nk}| + |E(X_{nk})| \le \varepsilon_n s_n + E(|X_{nk}|) \le \varepsilon_n s_n + \varepsilon_n s_n = ۲\varepsilon_n s_n$$

بنابراین

$$\frac{۱}{S_{nn}^3} \sum_{k=۱}^{n} E(|X_{nk} - E(X_{nk})|)^۳ \le \frac{۲\varepsilon_n s_n}{S_{nn}^3} \sum_{k=۱}^{n} E\left(X_{nk} - E(X_{nk})\right)^۲$$

$$\le \frac{۲\varepsilon_n s_n}{\mathrm{s_{nn}}} \to ۰ \qquad (۱۱-۲۰)$$

پس شرط کافی قضیه ثابت شد. اثبات فوق توسط لیند برگ ارائه شده است.

اثبات شرط لازم: ابتدا فرض می‌کنیم که $\frac{S_n}{s_n} \xrightarrow{d} N(۰,۱)$ هرگاه $n \to \infty$ آنگاه نتیجه می‌دهد

$$\varphi_{\frac{S_n}{s_n}}(t) = \varphi_{S_n}\left(\frac{t}{s_n}\right) = \prod_{k=۱}^{\mathrm{n}} \varphi_k\left(\frac{t}{s_n}\right) \to \exp\left\{-\frac{t^۲}{۲}\right\} \quad \to A$$

$$\Rightarrow \prod_{k=۱}^{\mathrm{n}} \ln \varphi_k\left(\frac{t}{s_n}\right) \to -\frac{t^۲}{۲}$$

$$\xrightarrow{\text{از طرفی داریم}} \sum_{k=0}^{n} \varphi(t) = 1 - \frac{t^۲}{۲} \Rightarrow \frac{t^۲}{۲} = 1 - \sum_{k=0}^{n} \varphi(t)$$

$$\Rightarrow \frac{t^{۲}}{۲} = \sum_{k=1}^{n} 1 - \varphi(t)$$

یا

$$-\sum_{k=1}^{n} \varphi(t) - 1$$

$$\xrightarrow{\text{از } A \text{ داریم}} \sum_{k=1}^{n} ln\, \varphi_k\left(\frac{t}{s_n}\right) - \sum_{k=1}^{n}\left[ln\, \varphi_k\left(\frac{t}{s_n}\right) - 1\right] \to 0$$

$$\xrightarrow{\text{با جایگذاری در } 1 \text{ و } 2 \text{ داریم}} \sum\left(\varphi\left(\frac{t}{s_n}\right) - 1\right) + \frac{t^{۲}}{۲} \to 0$$

حال اگر $Z = \varphi\left(\frac{t}{s_n}\right) - 1$ قرار دهیم برای nهای بزرگ $|z| \le 1/2$ که داریم با توجه به بسط

$$|e^z - 1 - z| \le z^2$$

$$\to \left|e^{it} - 1 - it\right| < t/2 \Rightarrow \left(\varphi_n(t/s_n) - 1\right) =$$

$$\longrightarrow \left|Ee^{itX_{n,k}} - 1 - itEX_{n,k}\right| \le E\left|e^{itX_{n,k}} - 1 - itX_{n,k}\right| \le \frac{1}{2}t^2 EX_{n,k}^2$$

و طبق تعریف داریم $EX_{n,k}^2 = \sigma^2$

$$\longrightarrow\ Max\left(\varphi\left(t/s_n\right) - 1\right) \le \frac{1}{2}t^2 \underbrace{Max\ \sigma^2}_{\text{یک می شود}} \Rightarrow \sum_{k=1}^{n}\left|\varphi_{n,k}(t/s_n) - 1\right| \le \frac{t^2}{2}$$

ازطرفی ثابت می‌شود وقتی $n \to \infty$ داریم

$$\Rightarrow \sum_{k=1}^{n}\left|\varphi\left(t/s_n\right) - 1\right|^2 \le \frac{t^4}{4} g(\varepsilon) \to 0$$

حال طبق بسط داریم

$$|e^z - 1| < e^{|z|} - 1$$

با قرار دادن $z = \varphi_k\left(\frac{t}{s_n}\right) - ۱$، آنگاه برای nهای بزرگ $|z| \le \frac{۱}{۲}$ که نتیجه می‌دهد

$$\left|e^{z} - ۱ - \mathrm{z}\right| \leq \mathrm{z}^{۲}$$

حال با استفاده از تعریف Z داریم

$$Z = \ \varphi_n\left({}^{t}/_{s_n}\right) - 1$$

$$\begin{aligned}
\rightarrow \left|e^{\varphi_k\left({}^{t}/_{s_n}\right)-1} - 1\right| &< \left|e^{\varphi_k\left({}^{t}/_{s_n}\right)-1} - \varphi_k\left({}^{t}/_{s_n}\right) + \varphi_k\left({}^{t}/_{s_n}\right) - 1\right| \\
&\leq \left|e^{\varphi_k\left({}^{t}/_{s_n}\right)-1} - \varphi_k\left({}^{t}/_{s_n}\right)\right| + \left|\varphi_k\left({}^{t}/_{s_n}\right) - 1\right| \\
&\leq \left|\varphi_k\left({}^{t}/_{s_n}\right) - 1\right|^2 + \left|\varphi_k\left({}^{t}/_{s_n}\right) - 1\right|
\end{aligned}$$

از طرفی با توجه به اینکه $|z| \leq \frac{1}{2}$ پس در نامساوی آخر داریم

$$\left|\varphi_k\left({}^{t}/_{s_n}\right) - 1\right|^2 + \left|\varphi_k\left({}^{t}/_{s_n}\right) - 1\right| \ \leq \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

با استفاده از (۱۱-۱۹)، داریم

$$\left|\prod_{k=1}^{n} \varphi_k\left(\frac{t}{s_n}\right) - \prod_{k=1}^{n} \exp\left\{\varphi_k\left(\frac{t}{s_n}\right) - 1\right\}\right| \leq \sum_{k=1}^{n} \left|\varphi_k\left(\frac{t}{s_n}\right) - \exp\left\{\varphi_k\left(\frac{t}{s_n}\right) - 1\right\}\right|$$

$$= \sum_{k=1}^{n} \left|\exp\left\{\varphi_k\left(\frac{t}{s_n}\right) - 1\right\} - 1 - \left\{\varphi_k\left(\frac{t}{s_n}\right) - 1\right\}\right| \leq \sum_{k=1}^{n} \left|1 - \varphi_k\left(\frac{t}{s_n}\right)\right|^2 \rightarrow 0$$

با استفاده از نامساوی مثلثی و اولین عبارت در اثبات لزوم، وقتی$n \rightarrow \infty$ داریم

$$\prod_{k=1}^{n} \exp\left\{\varphi_k\left(\frac{t}{s_n}\right) - 1\right\} \rightarrow \exp(-\frac{t^2}{2n}) \rightarrow 1$$

که مشابه است با:

$$\exp\left\{\sum_{k=1}^{n}\left(\varphi_k\left(\frac{t}{s_n}\right) - 1 + \frac{{\sigma_k}^2 t^2}{2{s_n}^2}\right)\right\} \rightarrow 1$$

قسمت حقیقی سمت چپ باید به یک همگرا شود، بنابراین توان قسمت نمایی باید به صفر میل کند.

$$\sum_{k=1}^{\mathrm{n}}\left(\mathrm{E}\cos\left(\frac{\mathrm{X}_k \mathrm{t}}{\mathrm{s}_n}\right) - 1 + \frac{{\sigma_k}^2 t^2}{2{s_n}^2}\right) \rightarrow 0$$

چون تابع زیر انتگرال نامنفی است $\cos x - ۱ + \frac{x^۲}{۲} \geq ۰$، امید ریاضی محدود شده به هر زیرمجموعه آن نیز وقتی $n \to \infty$ به صفر همگرا است به‌طوری‌که

$$E\left(\sum_{k=۱}^{n}\left(\mathrm{E}\cos\left(\frac{\mathrm{X}_k \mathrm{t}}{\mathrm{s}_n}\right) - ۱ + \frac{\sigma_k^2 t^۲}{۲S_n^2}\right) I\{|\mathrm{X}_k| > \varepsilon s_n\}\right) \to ۰$$

که مشابه رابطه زیر می‌باشد

$$\frac{t^۲}{۲} g_n(\varepsilon) - \sum_{k=۱}^{n} E\left(۱ - \cos\left(\frac{\mathrm{X}_k \mathrm{t}}{\mathrm{s}_n}\right)\right) I\{|\mathrm{X}_k| > \varepsilon s_n\} \to ۰$$

در نهایت چون $\left|۱ - \cos(\mathrm{y})\right| \leq ۲$ برای هر y و با استفاده از نامساوی مارکف داریم

$$\sum_{k=۱}^{\mathrm{n}} \mathrm{E}\left(۱ - \cos\left(\frac{\mathrm{X}_k \mathrm{t}}{\mathrm{s}_n}\right)\right) I\{|\mathrm{X}_k| > \varepsilon s_n\} \leq \sum_{k=۱}^{\mathrm{n}} ۲\mathrm{p}(|\mathrm{X}_k| > \varepsilon s_n) \leq ۲\sum_{k=۱}^{\mathrm{n}} \frac{{\sigma_k}^۲}{(\varepsilon \mathrm{s}_n)^۲} = \frac{۲}{\varepsilon^۲}$$

به‌طوری‌که

$$۰ \leq \lim_{n\to\infty} \sup \mathrm{g_n}(\varepsilon) \leq \frac{۴}{(\mathrm{t}\varepsilon)^۲}$$

که با انتخاب t به قدر کافی بزرگ می‌توان آن را به اندازه دلخواه کوچک نمود، در نتیجه قضیه اثبات می‌شود.

نتیجه ۱۱-۲: شرط لیاپانف شرط لیندبرگ را نتیجه می‌دهد.

اثبات:

$$\frac{۱}{S_n^2}\sum_{k=۱}^{n}\int_{|x|>\varepsilon s_n} X^۲ F_K \leq \frac{۱}{\varepsilon^{\delta} S_n^{2+\delta}}\sum_{k=۱}^{n}\int_{|x|>\varepsilon s_n} |X|^{۲+\delta} dF_K$$

$$\leq \frac{\left[\sum_{k=1}^{n} E\left(|X_k|^{2+\delta}\right)\right]}{\varepsilon^{\delta} S_n^{2+\delta}} \qquad (21-11)$$

بنابراین اگر شرط لیاپانف برقرار باشد، شرط لیندبرگ هم برقرار خواهد بود.

مثال ۱۱-۲: فرض کنید $\mathrm{X}_1, \mathrm{X}_2, \ldots$ مستقل باشند و داشته باشیم

$$p(X_K = a_K) = p(X_K = -a_K) = \frac{۱}{۲} \quad , \ a_K > ۰ \qquad (\mathrm{k} = ۱,۲, \dots)$$

$$\sigma_n = a_K \quad , \quad s_n = \sqrt{\sum_{k=۱}^{n} {a_K}^{۲}} \quad , \quad E(X_K) = ۰$$

حال فرض می‌کنیم $m_\mathrm{n} = \mathrm{Max}_{\mathrm{k \le n}} a_K$، از طرفی می‌دانیم که

$$\sigma_k^{2+\delta} = \sigma_k^2 . \sigma_k^\delta = \mathrm{E}\left({X_K}^{۲}\right) . \sigma_k^\delta \le \mathrm{E}\left({X_K}^{۲}\right) . m_n^\delta = m_n^\delta . \sigma_k^2$$

$$\sum_{k=۱}^{n} a_k^{2+\delta} \ge Max(a_k^{2+\delta}) = m_n^{2+\delta}$$

$$\left(\frac{m_\mathrm{n}}{s_n}\right)^{۲+\delta} \le \frac{\sum_{k=۱}^{n} a_k^{2+\delta}}{s_n^{2+\delta}} = \frac{۱}{s_n^{2+\delta}} \sum_{k=۱}^{n} E\left(|X_K|^{۲+\delta}\right)$$

$$\le \frac{m_n^\delta \sum_{k=۱}^{n} E\left(|X_K|^{۲}\right)}{s_n^{2+\delta}} = \frac{m_n^\delta}{s_n^\delta} = \left(\frac{m_\mathrm{n}}{s_n}\right)^{\delta}$$

$$\Rightarrow \left(\frac{m_\mathrm{n}}{s_n}\right)^{۲+\delta} \le \left(\frac{m_\mathrm{n}}{s_n}\right)^{\delta} \Rightarrow \frac{m_\mathrm{n}}{s_n} \longrightarrow ۰$$

(توضیح: در نهایت ما دیدیم که $1 > \left(\frac{m_n}{s_n}\right)^{\delta}$ بنابراین چون این عبارت مثبت بود و کسر کوچکتر از واحد می‌باشد

بنابراین زمانی که $n \to \infty$ میل میکند عبارت $\frac{m_n}{s_n}$ به سمت صفر میل می کند

$$0 \le \frac{m_n}{s_n} = \frac{Max_{k \le n} a_K}{\sqrt{\sum_{k=۱}^{n} {a_K}^{۲}}} < 1 \qquad \text{وقتی} \ n \to \infty$$

که شرط لازم و کافی برای برقراری $C.L.T$ می‌باشد.

اگر $a_k = k^\lambda (\lambda > 0)$، $m_\mathrm{n} = n^\lambda$، آنگاه $\sqrt{\sum_{k=۱}^{\mathrm{n}} k^{۲\lambda}} = \frac{n^{\lambda + \frac{۱}{۲}}}{\sqrt{۲\lambda + ۱}} \approx s_\mathrm{n}$.

پس اگر $\frac{m_{\mathrm{n}}}{s_n} \longrightarrow ۰$ ، شرایطCLT برقرار می‌باشد. اگر $\lambda < ۰$ باشد، این شرط به ازای $\lambda < ۰$، $m_{\mathrm{n}} = 1$، برقرار است.

و اگر $\frac{m_{\mathrm{n}}}{s_n} \longrightarrow ۱$ و $s_{\mathrm{n}} \to \infty$ به عبارت دیگر $\sum_{k=۱}^{\mathrm{n}} k^{۲\lambda} \to ۰$ یا اگر $\lambda \geq -\frac{۱}{۲}$ به ازای تمام مقادیر

$\lambda \geq -\frac{۱}{۲}$ ، CLT برقرار است.

۱۱–۵ حد مرکزی چند متغیره

۱۱–۵–۱ تابع مشخصه نرمال p متغیره

بردار میانگین بردار تصادفی p متغیره $X = (X_1 , \dots , X_p)$ که ماتریس واریانس-کوواریانس آن، عبارت است از:

$$\sum = \left(\sigma_{ij}\right) = (cov(X_i, X_j))$$

گوییم بردار X دارای توزیع نرمال p متغیره است، اگر تابع چگالی آن به صورت زیر باشد

$$N_{\mathrm{p}}\left(\mu , \sum\right) = (۲\pi)^{-\frac{\mathrm{p}}{۲}} |\Sigma|^{-\frac{۱}{۲}} \exp\{-\frac{۱}{۲}(X-\mu)'\Sigma^{-1}(X-\mu)\}$$

می‌توان نشان داد که تابع مشخصه آن عبارت است از

$$\varphi(u) = \mathrm{E}\{\exp(\mathrm{i}u'\mathrm{X})\} = \exp\{\mathrm{i}u'\mu - \frac{۱}{۲}u'\Sigma u\}$$

که درآن u یک بردار p بعدی از مقادیر حقیقی می‌باشد. تابع مشخصه یک بردار نرمال با میانگین صفر و ماتریس کوواریانس $\Sigma = I$ عبارت است از

$$\varphi(u) = \exp\left\{\frac{-1}{2}u'u\right\} = \exp\{-\frac{1}{2}\sum_{j=1}^{p} u_j^2\}$$

که همان حاصل‌ضرب توابع مشخصه p متغیره نرمال استاندارد مستقل می‌باشد. چون تابع مشخصه تابع توزیع را به‌طور یکتا معین می‌کند، متغیرهای نرمال ناهمبسته، مستقل هستند.

در حالت خاص اگر $l'X = \sum_{i=1}^{p} l_i X_i$، تابعی حقیقی از $\mathrm{X_i}$ ها باشد، تابع مشخصه آن

عبارت است از

$$E\{exp(\mathrm{i}\acute{u}l'\mathrm{X})\} = E\{\exp i(lu)'\, X\}$$

$$= \exp\{i(lu)'\mu - \tfrac{1}{2}(lu)'\Sigma(lu)\}$$

$$=\exp\{i(lu)'\mu - \tfrac{1}{2}(lu)'\textstyle\sum(lu)\} = \exp\{i(lu)'\mu - \tfrac{1}{2}(lu)'\textstyle\sum(lu)\}$$

$$= \exp\{iu(l'\mu) - \tfrac{1}{2}(u'(l'\textstyle\sum l)\,u) \implies l'X \sim N(l'\mu, l'\textstyle\sum l) \qquad (۲۲\text{-}۱۱)$$

وبالعکس، اگر هر ترکیب خطی از مؤلفه‌های بردار X دارای توزیع نرمال باشد، X دارای توزیع نرمال چندمتغیره p است.

۱۱–۵–۲ $C.L.T$ برای حالت p متغیره

دنباله‌ای از بردارهای p بعدی $\{X_K\}$، $X_k = (X_{k1}, \ldots, X_{kp})$ را درنظر بگیرید و فرض کنید $\mathrm{F_k}$ توزیع p بعدی X_kو F_{lk} تابع توزیع یک بعدی $\sum_{i=1}^{\mathrm{p}} l_{\mathrm{i}}\mathrm{X_{ki}}$، $(k = ۱,۲,..)$و F_{lk} به ازای تمام l ها به‌طور ضعیف به F_l همگرا باشد.

از آن جایی که $\{F_l: l \in R^p\}$، تابع توزیع بردار p بعدی $X = (X_1, \ldots, X_p)$ و $\{\mathrm{F}_{lk}\}$ تابع توزیعی از $\mathrm{F_k}$، $\{X_K\}$ را معین می‌کنند، آن‌گاه $\mathrm{F_k} \xrightarrow{\mathrm{w}} \mathrm{F}$ اگر و فقط اگر به ازای هر l، $\mathrm{F}_{lk} \to \mathrm{F}_l$.

طبق این خاصیت می‌توان مسأله تعیین توزیع حدی دنباله‌ای از متغیرهای تصادفی برداری را به مسأله دنباله‌ای از متغیرهای تصادفی یک بعدی تبدیل نمود.

با فرض $Y_{lk} = l'X_K = \sum_{i=1}^{p} l_i X_{Ki}$داریم

$$E(l'X_K) = l'E(X_K) = \textstyle\sum_{i=1}^{p} l_i \mu_{ki} = \mu_{lk}$$

$$\mathrm{Var}(l'X_K) = \textstyle\sum_i \sum_j l_i l_j \sigma_{ij}(k) = \sigma_l(k)$$

از طرفی بردار میانگین $E(X_K) = \mu_K = (\mu_{k1}, \ldots, \mu_{kp})$و $\sum_k = (\sigma_{ij}(k))$ بردار ماتریس کوواریانس X_K می‌باشند. اگر $\{X_K\}$ دنباله‌ای از بردارهای i.i.d باشد آن‌گاه y_{lk} دنباله‌ای از متغیرهای تصادفی i.i.d خواهند بود. اگر y_{lk} از $\mathrm{C.L.T}$ پیروی کند یعنی مجموع‌های جزئی استاندارد شده آن دارای توزیع مجانبی نرمال استاندارد باشد، می‌گوئیم که دنباله $\{X_K\}$ از قانون C.L.T، پیروی می‌کند. در این حالت مجموع جزیی استاندارد شده (دارای بردار میانگین صفر و ماتریس واریانس-کوواریانس همانی) متغیرهای تصادفی برداری $\{X_K\}$ دارای توزیع مجانبی نرمال p متغیره خواهد بود.

حال صورت لیندبرگ- لوی C.L.T را برای متغیرهای تصادفی برداری به صورت زیر بیان می‌کنیم.

قضیه۱۱- ۴: اگر $\{X_K\}$ بردارهای i.i.d با بردار میانگین صفر و ماتریس کوواریانس $\sum$ باشد آنگاه تابع توزیع $\frac{S_n}{\sqrt{n}} = \frac{\sum X_k}{\sqrt{n}}$ به سمت توزیع نرمال p متغیره با میانگین صفر و ماتریس کواریانس $\sum$ میل می‌کند.

اثبات: فرض می‌کنیم به ازای یک l خاص

$$, \quad Y_{lk} = I'X_k = \sum_{i=1}^{p} l_i X_{ki}$$

$$E(Y_{lk}) = \sum l_i\, E(X_{ki}) = 0, \qquad var(Y_{lk}) = l'\Sigma l$$

باشد، پس $Y_{lk}, \quad k = 1,2, \ldots$، دنباله‌ای از متغیرهای تصادفی i.i.d است و به موجب قضیه (۱۱-۱) لیندبرگ- لوی $\frac{\sum Y_{lk}}{\sqrt{n}}$ دارای توزیع مجانبی $N(0, l'\Sigma l)$ می‌باشد.

۱۱–۶ مسائل

۱) در هر یک از حالات زیر تعیین کنید آیا C.L.T برای دنباله‌های متغیرهای تصادفی مستقل $\{X_n\}$, $(n = 1,2, \ldots)$، برقرار است.

الف) X_n دارای توزیع نرمال با واریانس واحد و $E(x_{2n}) = -۱\ ,\ E\left(x_{2n+1}\right) = ۱$

ب) X_n دارای توزیع پواسون با میانگین $n\lambda$

پ) X_n دارای توزیع نرمال با $var(X_n) = n^2,\ \ E(X_n)$

فصل دوازدهم

قضیه رادون – نیکودایم

اگر f تابعی نامنفی در فضای اندازه $(\Omega,\mathcal{F},\mu)$ باشد، آنگاه $\upsilon(A)=\int_A f\,d\mu$ اندازه دیگری روی میدان $\mathcal{F}$ تعریف می‌کند، به طوری که υ دارای چگالی f نسبت به μ می‌باشد. برای هر مجموعه A متعلق به $\mathcal{F}$، $\mu(A)=0$ الزاماً $\upsilon(A)=0$ را نتیجه می‌دهد. هدف این بخش آن است که نشان دهد عکس این رابطه نیز برقرار است. یعنی اگر υ و μ هر دو اندازه‌های σ- متناهی روی $\mathcal{F}$ باشند، آنگاه υ دارای چگالی مانند f نسبت به μ $\upsilon(A)=\int_A f\,d\mu$ است.

۱۲–۱ توابع مجموعه ای جمع پذیر

در سراسر این بخش دوتایی $(\Omega,\mathcal{F})$ یک فضای اندازه پذیر می‌باشد. نیز فرض بر این است که تمام مجموعه ها در $\mathcal{F}$ قرار دارند.

تعریف ۱۲-۱: یک تابع مجموعه‌ای جمع پذیر، تابعی است مانند φ از میدان $\mathcal{F}$ به مقادیر حقیقی ، به طوری‌که

$$\varphi(\cup_n A_n)=\sum_n \varphi(A_n) \qquad (۱-۱۲)$$

هرگاه $A_1,A_2,\dots,A_n$ دنباله‌ای متناهی یا نامتناهی از مجموعه‌های مجزا باشند. فرق عمده یک تابع مجموعه ای جمع پذیر با اندازه این است که، $\varphi(A)$ می تواند منفی باشد، اما متناهی است. بدین معنا که سمت راست رابطه (۱-۱۲) باید قطعاً همگرا باشد و نمی تواند مقادیر $+\infty$ یا $-\infty$ را اختیار کند. توجه شود که $\varphi(\emptyset)=0$ است

مثال ۱۲-۱: اگر μ_1 و μ_2 دو اندازه متناهی باشند آنگاه تفاضلشان یعنی $\varphi(A)=\mu_1(A)-\mu_2(A)$ یک تابع مجموعه‌ای جمع پذیر است. زیرا

$$\varphi\left(\bigcup A_n\right)=\mu_1\left(\bigcup A_n\right)-\mu_2\left(\bigcup A_n\right)=\sum_n \mu_1(A_n)-\sum_n \mu_2(A_n)$$
$$=\sum_n[\mu_1(A_n)-\mu_2(A_n)]=\sum_n \varphi(A_n)\,.$$

یک حالت خاص آن زمانی است که، $\varphi(A) = \int_A f\, d\mu$، که f انتگرال پذیر است. (اما لزوماً نامنفی نیست).

اثبات قضیه اصلی این فصل (قضیه رادون- نیکودایم) نیاز به ارائه حقایقی درباره توابع مجموعه‌ای جمع پذیر دارد؛ اگرچه صورت این قضیه تنها، بر اساس اندازه ها بیان شده است.

لم ۱۲-۱: اگر $E_U \uparrow E$ یا $E_U \downarrow E$ ، یعنی اگر E_U دنباله‌ای باشد که به E صعود یا نزول کند، آنگاه

$$\varphi(E_U) \to \varphi(E).$$

اثبات: اگر $E_U \uparrow E$ می توان نوشت:

$$\varphi(E) = \varphi\left(\bigcup_{U=1}^{\infty} E_U\right) = \varphi\left[E_1 \bigcup_{U=1}^{\infty}(E_{U+1} - E_U)\right] = \varphi(E_1) + \sum_{U=1}^{\infty} \varphi(E_{U+1} - E_U)$$
$$= \lim_{c\to+\infty}\left[\varphi(E_1) + \sum_{U=1}^{c-1} \varphi(E_{U+1} - E_U)\right] = \lim_{c\to+\infty} \varphi(E_c)$$

همچنین اگر $E_U \downarrow E$ ، آنگاه متمم آن $E_U^c \uparrow E^c$ و در نتیجه:

$$\varphi(E_U) = \varphi(\Omega) - \varphi(E_U^c) \to \varphi(\Omega) - \varphi(E^c) = \varphi(E)$$

مشابه این نتیجه در مورد اندازه‌ها نیز وجود دارد، اما اثبات آن متفاوت است. توجه شود که در اینجا نیازی نیست که حدها یکنوا باشند.

۱۲-۲ تجزیه هان

قضیه ۱۲-۱ (تجزیه هان): برای هر تابع مجموعه‌ای جمع پذیر φ، مجموعه‌های مجزای A^+ و A^- وجود دارند به طوری که $A^+ \cup A^- = \Omega$ و برای هر مجموعه $E \in A^+$، $\varphi(E) \geq 0$ و برای هر $E \in A^-$، $\varphi(E) \leq 0$.

هر مجموعه A را مثبت گویند اگر برای هر $E \subset A$، داشته باشیم $\varphi(E) \geq 0$، و منفی گویند اگر برای هر $E \subset A$، داشته باشیم $\varphi(E) \leq 0$. در واقع این قضیه Ω را به وسیله ی یک تابع مجموعه ای جمع پذیر به دو جزء مثبت و منفی تقسیم می کند. این قضیه را تجزیه هان گویند.

اگر $\varphi(A) = \int_A f d\mu$ باشد، $A^+ = [f \geq 0]$ و $A^- = [f < 0]$ در نظر گرفته شود. حال اگر $\varphi^+(A) = \varphi\left(A \cap A^+\right)$ و $\varphi^-(A) = -\varphi(A \cap A^-)$ تعریف کنیم، آنگاه φ^+ و φ^- اندازه‌های متناهی هستند که آن‌ها را تغییرات بالایی و پایینی φ می نامند، بطوری که:

$$\varphi(A) = \varphi^+(A) - \varphi^-(A) \qquad (۱۲-۲)$$

رابطه (۱۲-۲) را تجزیه جردن گویند، که در واقع تابع مجموعه ای φ را به صورت تفاضل دو اندازه متناهی با تکیه گاه های مجزا نشان می دهد. $|\varphi|$ را نیز تغییرات کل در نقطه A می گویند و برابراست با:

$$|\varphi(A)| = \varphi^{+}(A) + \varphi^{-}(A).$$

۱۲–۳ منفرد بودن و پیوستگی مطلق

تعریف ۱۲-۲: دو اندازه μ و ν روی فضای اندازه‌پذیر (Ω, F) را تواماً منفرد گویند هرگاه دارای تکیه‌گاه های جدا باشند، یعنی مجموعه‌های s_μ و s_ν وجود داشته باشند به طوری‌که:

$$\begin{cases} \mu(\Omega - s_\mu) = 0 \\ \nu(\Omega - s_\nu) = 0 \end{cases}, s_\mu \cap s_\nu = \emptyset \qquad (۳-۱۲)$$

در این صورت ν نسبت به μ و μ نسبت به ν منفرد می باشند. توجه شود که اگر یکی از اندازه‌ها صفر باشد، هر اندازه بطور خودکار نسبت به دیگری منفرد خواهند بود.

بر مبنای قضیه ای (یادآوری ۱۲-۱) یک اندازه متناهی روی R با تابع توزیع F نسبت به اندازه لبگ منفرد است، اگر و فقط اگر $F'(x) = 0$ به جز روی مجموعه‌ای با اندازه لبگ صفر.

(**یادآوری ۱۲-۱:** فرض کنید

$$\mu(\mathrm{a}, \mathrm{b}] = \mathrm{F}(\mathrm{b}) - \mathrm{F}(\mathrm{a})$$

شرط لازم و کافی برای آنکه دو اندازه μ و ν دارای تکیه گاه های مجزا باشند، آن است که $F'(x) = 0$ به جز روی مجموعه‌ای با اندازه لبگ صفر.)

تعریف ۱۲-۳: اندازه ν نسبت به اندازه μ مطلقاً پیوسته گفته می‌شود، اگر برای $A \in \mathcal{F}$، $\mu(A) = 0$ ایجاب کند $\nu(A) = 0$. در این صورت ν تحت سلطه μ بوده و آن را با نماد $\nu \ll \mu$ نشان می دهیم. اگر $\nu \ll \mu$ و هم‌چنین $\mu \ll \nu$ آنگاه دو اندازه معادلند و می‌نویسیم $\nu \equiv \mu$.

نکته: در واقع یک ایده $\varepsilon - \delta$ مربوط به تعریف پیوستگی مطلق وجود دارد. فرض کنید

$$\forall \varepsilon > 0\ \exists \delta > 0 \ \ni \ (\mu(A) < \delta \ \Rightarrow \nu(A) < \varepsilon)$$

آنگاه $\nu \ll \mu$ می باشد؛ زیرا اگر $\mu(A) = 0$ باشد، در نتیجه به ازای هرε، $\nu(A) < \varepsilon$ است و بنابراین $\nu(A) = 0$ است یعنی داریم $\nu \ll \mu$. اما عکس این رابطه، یعنی از $\nu \ll \mu$ به ایده $\varepsilon - \delta$ زمانی برقرار است که ν متناهی باشد. این ادعا را به کمک برهان خلف به صورت زیر تشریح میکنیم.

اثبات: (برهان خلف) فرض کنید ν متناهی و $\mu \gg \nu$ باشد، اما رابطه $\varepsilon - \delta$ برقرار نباشد. آنگاه برای برخی εها دنباله ای از مجموعه‌های A_n وجود دارد به طوری که $\mu(A_n) < n^{-2}$ و $\nu(A_n) \geq \varepsilon$. اگر $\limsup_n A_n = A$ آنگاه بنا به لم بورل کانتلی خواهیم داشت $\mu(A) = 0$،

$$\sum \mu(A_n) < \sum \frac{1}{n^2} < \infty \Rightarrow \mu\left(\limsup_n A_n\right) = \mu(A) = 0$$

اما (بنابر لم فاتو) $\nu(A) > \varepsilon > 0$ است. زیرا:

$$\nu(A_n) \geq \varepsilon \Rightarrow \limsup_n \nu(A_n) \geq \varepsilon \Rightarrow \nu\left(\limsup_n A_n\right) \geq \limsup_n \nu(A_n) \geq \varepsilon \Rightarrow \nu(A) \geq \varepsilon \geq 0$$

که با فرض اولیه مان یعنی $\mu \gg \nu$ در تناقض است ($\mu(A) = 0 \not\Rightarrow \nu(A) = 0$)، پس حضور شرط متناهی بودن ν نیز لازم می باشد. بنابراین، رابطه $\varepsilon - \delta$ برقرار است هرگاه ν متناهی و $\mu \gg \nu$ باشد.

یادآوری ۱۲-۲: لم بورل کانتلی: فرض کنید A_n ها دنباله ای از پیشامد ها از فضای احتمال $(\Omega, \mathcal{F}, \mathrm{P})$ با شند در این صورت

الف) اگر $\sum_{n=1}^{\infty} P(A_n) < \infty$ باشد، آنگاه $P\left(\limsup_n A_n\right) = 0$،

ب) اگر $\sum_{n=1}^{\infty} P(A_n) = \infty$ یعنی سری واگرا باشد و A_nها مستقل باشند، آنگاه $P\left(\limsup_n A_n\right) = 1$.

۱۲–۴ قضیه اصلی

اگر $\nu(A) = \int_A f d\mu$ باشد، آنگاه مسلماً $\mu \gg \nu$ است. قضیه رادون - نیکودایم عکس این مطلب را بیان می‌کند

قضیه ۱۲-۲ (رادون – نیکودایم):

اگر μ و ν دو اندازه σ-متناهی باشند به‌طوریکه $\mu \gg \nu$ باشد، آنگاه یک تابع نامنفی (چگالی) f وجود دارد، بطوری که برای همه $A \in \mathcal{F}$ داریم $\nu(A) = \int_A f d\mu$. برای دو چگالی f و g که دارای این خاصیت باشند، داریم $\mu[f \neq g] = 0$.

f را مشتق رادون نیکودایم ν نسبت به μ گویند و آن را با نماد $\frac{d_\nu}{d_\mu}$ نمایش می دهند.

$$\nu(A) = \int_A d\nu = \int_A f d\mu \Rightarrow f = \frac{d\nu}{d\mu}$$

ابتدا یکتایی این چگالی را ثابت می کنیم و سپس بحث وجودی آن اثبات شود.

لم ۱۲-۲: اگر μ و ν اندازه‌های متناهی بوده و توأماً منفرد نباشند، آنگاه یک مجموعه ای مانند A و مقدار مثبتی مانند ε وجود دارد به طوری‌که $\mu(A) > 0$ و برای هر $E \subset A$، $\varepsilon\mu(E) \leq \nu(E)$.

اثبات: فرض کنید تجزیه هان برای تابع مجموعه‌ای $\varphi = \nu - n^{-1}\mu$ ، به صورت $A_n^+ \cup A_n^-$ باشد. اگر قرار دهیم $M = \bigcup_n A_n^+$ ، بنابراین $M^c = \bigcap_n A_n^-$. از آنجایی که M^c جزء منفی مجموعه می‌باشد، طبق تعریف جزء منفی داریم $\varphi(M^c) \leq 0$است. بنابراین:

$$\varphi(M^c) = \nu(M^c) - n^{-1}\mu(M^c) \leq 0 \Rightarrow \nu(M^c) \leq n^{-1}\mu(M^c)$$

چون این رابطه برای هر n برقرار است، پس برای nهای بزرگ داریم $\nu(M^c) = 0$ودر نتیجه M تکیه گاه ν است. زیرا:

$$\lim_n \nu(M^c) \leq \lim_n n^{-1}\mu(M^c) = 0 \Rightarrow \nu(M^c) = 0$$

از طرفی چون μ و ν توأماً منفرد نمی‌باشد، پس M^c نمی‌تواند تکیه‌گاه μ باشد. پس M فقط تکیه گاه ν نبوده و بنابراین $\mu(M) > 0$، پس داریم:

$$\mu(M) > 0 \Rightarrow \exists n \ni \mu(A_n^+) > 0$$

حال اگر A_n^+ را برابر A و ε را n^{-1}در نظر بگیریم، در نتیجه $\nu(E) \geq \varepsilon\mu(A)$، زیرا:

$$(\nu - n^{-1}\mu)(A_n^+) > 0 \Rightarrow \nu(A_n^+) > \frac{\mu}{n}(A_n^+) \Rightarrow \nu(A) > \varepsilon\mu(A).$$

اثبات قضیه ۱۲-۲:

یکتایی f: فرض کنیم g و f دو تابع اندازه‌پذیر نسبت به $\mathcal{F}$ بوده به طوری‌که

$$\varphi(A) = \int_A f d\mu = \int_A g d\mu$$

آنگاه

$$\forall A \in \mathcal{F}: \int_A (f-g)\, d\mu = 0 \Rightarrow f - g = 0 a.s$$

در غیر این‌صورت $A:\{|f-g| > \varepsilon\}$ به طوری‌که $\mu(\text{A})>۰$ و لذا

$$\int_A (f-g) d\mu > \varepsilon\mu(A) > 0$$

که این متناقض با $f - \text{g} = ۰$ است. لذا f منحصر به فرد است.

وجودی f:

در قضیه رادون نیکودایم ν و μ هر دو σ- متناهی هستند. در این صورت یک افراز شمارا از Ω ، مانند A_n وجود خواهد داشت که در آن $\mu(A_n)$ و $\nu(A_n)$ هر دو متناهی‌اند. اگر

$$\mu_n(A) = \mu(A \cap A_n), \quad \nu_n(A) = \nu(A \cap A_n)$$

آنگاه $\nu << \mu$ نتیجه می‌دهد $\nu_n << \mu_n$ و بنابراین چگالی ای مانند f_n وجود دارد که $\nu_n(A) = \int_A f_n d\mu_n$.
از آنجایی که μ_n دارای چگالی I_{A_n} نسبت به μ است، پس

$$\nu(A) = \sum_n \nu_n(A) = \sum_n \int_A f_n d\mu_n = \sum_n \int_A f_n I_{A_n} d\mu = \int_A \sum_n f_n I_{A_n} d\mu = \int_A f d\mu$$

پس ν دارای چگالی $f = \sum_n f_n I_{A_n}$ نسبت به اندازه μ است.
با توجه به مطالب بیان شده کافی است قضیه را برای μ و ν متناهی اثبات کنیم.
فرض کنید μ و ν اندازه‌های متناهی و $\nu << \mu$ باشد. $\mathcal{G}$ را نیز کلاسی از توابع نامنفی g در نظر بگیریم به‌طوری که برای هر E داشته باشیم $\int_E g d\mu \leq \nu(E)$. اگر g و g' در کلاس $\mathcal{G}$ قرار داشته باشند، ماکسیمم آن‌ها نیز در $\mathcal{G}$ قرار دارد؛ زیرا

$$\begin{aligned}\int_E \max(g, g')\, d\mu &= \int_{E\cap(g\geq g')} g d\mu + \int_{E\cap(g<g')} g' d\mu \\ &\leq \nu\left(E \cap (g \geq g')\right) + \nu\left(E \cap (g < g')\right) = \nu(E)\end{aligned}$$

بنابراین $\mathcal{G}$کلاسی است که تحت عمل ماکسیمم گیری متناهی بسته است. حال فرض کنید g_nها توابعی در کلاس $\mathcal{G}$ بوده و به g صعود کنند یعنی $g_n \uparrow g$. آنگاه چون g_n ها نامنفی و صعودی هستند، بنا به قضیه همگرایی یکنوا داریم

$$\int_E g d\mu = \int_E \lim_n g_n d\mu = \lim_n \int_E g_n d\mu \leq \nu(E) \Rightarrow g \in \mathcal{G}$$

لذا کلاس $\mathcal{G}$ تحت حد توابع غیرنزولی نیز بسته است. حال اگر برای هر g در $\mathcal{G}$، $\alpha = \sup \int g d\mu$ $(\alpha \leq \nu(\Omega))$ قرار دهیم و $g_n \in \mathcal{G}$ را طوری انتخاب کنیم که $\int g_n d\mu > \alpha - \frac{1}{n}$. انگاه اگر $f_n = \max(g_1 \dots g_n)$ و $f = \lim f_n$، f در $\mathcal{G}$ قرار دارد و $\int f d\mu = \lim_n \int f_n d\mu \geq \lim_n \int g_n d\mu = \alpha$. بنابراین f عضوی از $\mathcal{G}$ می‌باشد که در آن $\int f d\mu$ ماکسیمم است. (چرا f در $\mathcal{G}$ قرار دارد؟ چون $\mathcal{G}$ نسبت به عمل ماکسیمم گیری متناهی و حد توابع غیرنزولی و مثبت بسته است).
تعریف کنید $\nu_{ac} = \nu_{ac}(E) = \int_E f d\mu$ و $\nu_s = \nu_s(E) = \nu(E) - \nu_{ac}(E)$ آنگاه

$$\nu(E) = \nu_{ac}(E) + \nu_s(E) = \int_E f d\mu + \nu_s(E) \qquad (۴\text{-}۱۲)$$

از آنجایی که:

$$\begin{cases} \nu \; finite \Rightarrow \nu(E) < \infty \\ \mathrm{f} \in \mathcal{g} \Rightarrow \nu_{ac}(E) = \int_E f d\mu \leq \nu(E) < \infty \end{cases} \Rightarrow \nu_s(E) < \infty$$

پس ν_s و ν_{ac} اندازه‌های متناهی و البته نامنفی‌اند. از آنجایی که ν نسبت به μ مطلقاً پیوسته است، می توان $\nu_{ac} \ll \mu$ را نتیجه گرفت. زیرا:

$$\mu(A) = 0 \Rightarrow 0 \leq \nu_{ac}(\mathrm{A}) = \int_E f d\mu \leq \nu(A) = 0 \;\Rightarrow \nu_{ac}(\mathrm{A}) = 0$$

حال فرض کنید که ν_s نسبت به μ منفرد نباشد، در نتیجه از لم (۱۲-۲)، یک مجموعه A و یک ε مثبت وجود دارد به‌طوریکه $\mu(A) > 0$ و برای تمام $E \subset A$ها $\varepsilon\mu(E) \leq \nu_s(E)$ است.
پس برای هر E

$$\begin{aligned} \int_E (f + \varepsilon I_A) d\mu &= \int_E f d\mu + \varepsilon\mu(E \cap A) \leq \int_E f d\mu + \nu_s(E \cap A) \\ &= \int_{E\cap A} f d\mu + \nu_s(E \cap A) + \int_{E-A} f d\mu \\ &= \nu(E \cap A) + \int_{E-A} f d\mu \leq \nu(E \cap A) + \nu(E - A) \\ &= \nu(E) \end{aligned}$$

به عبارت دیگر $\left(f + \varepsilon I_A\right) = g^*$ در $\mathcal{g}$ قرار دارد و

$$\int (f + \varepsilon I_A)\, d\mu = \int f d\mu + \varepsilon \int I_A d\mu = \alpha + \varepsilon\mu(A) > \alpha$$

که با ماکسیمم بودن f در تناقض است. بنابراین، μ و ν_s دو به دو منفرد می‌باشند. بنابراین sای وجود دارد که $\nu_s(s) = \mu(s^c) = 0$. ولی از آن‌جا که $\nu << \mu$ پس $\nu_s(s^c) \leq \nu(s^c) = 0$ و هم‌چنین $\nu_s(\Omega) = 0$. در نتیجه در رابطه (۱۲-۴)، ν_s حذف می شود و داریم:

$$\nu(E) = \int_E f d\mu$$

خود رابطه (۱۲-۴) را تجزیه لبگ گویند که در واقع ν را به دو قسمت مطلقاً پیوسته و قسمت منفرد (ویژه) تقسیم می کند.

۱۲–۵ مسائل

۱) فرض کنید μ، υ، υ_n اندازه های متناهی روی (Ω, F) باشند وبه ازای هر A ، $\upsilon(A)=\sum_n \upsilon_n(A)$. قرار دهید $\upsilon(A)=\int_A f\,d\mu+\upsilon'(A)$ و $\upsilon_n(A)=\int_A f_n d\mu+\upsilon_n'(A)$ تجزیه های (۱۲-۴) باشند. در اینجا υ' و υ_n' نسبت به μ منفرد هستند. نشان دهید که $f=\sum_n f_n$ به جز روی یک مجموعه μ -اندازه صفر، و به ازای هر A ، $\upsilon'(A)=\sum_n \upsilon_n'(A)$ می‌باشد. همچنین نشان دهید $\upsilon \ll \mu$ اگر و فقط اگر به ازای هر n، $\upsilon_n \ll \mu$.

۲) اگر $A^+ \cup A^-$ تجزیه هان φ باشد ممکن است تجزیه دیگری مانند $A_1^+ \cup A_1^-$ وجود داشته باشد. چنین مثالی بسازید و نشان دهید که $\varphi(A^+ \vartriangle A_1^+) = \varphi(A^- \vartriangle A_1^-) = 0$.

۳) μ، υ و ρ را اندازه های σ متناهی روی (Ω, F) در نظر بگیرید. فرض کنید که مشتقات رادون نیکودایم متناهی و نامنفی باشند.

الف) نشان دهید که $\mu \ll \upsilon$ و $\rho \ll \mu$ ایجاب می کند که $\rho \ll \upsilon$ و

$$\frac{d\upsilon}{d\rho}=\frac{d\upsilon}{d\mu}\frac{d\mu}{d\rho}$$

ب) نشان دهید که $\upsilon \equiv \mu$ ایجاب می کند که

$$\frac{d\upsilon}{d\mu}=I_{[d\mu/d\upsilon>0]}\left(\frac{d\mu}{d\upsilon}\right)^{-1}$$

ج) فرض کنید که $\rho \ll \mu$ و $\rho \ll \upsilon$، و A مجموعه ای باشد جایی‌که $d\upsilon/d\rho > 0 = d\mu/d\rho$. نشان دهید که $\mu \ll \upsilon$ اگر و فقط اگر $\rho(A) = 0$ در حالتی که

$$\frac{d\upsilon}{d\mu}=I_{[d\mu/d\rho>0]}\frac{d\upsilon/d\rho}{d\mu/d\rho}$$

فصل سیزدهم

احتمال شرطی

۱۳–۱ احتمال شرطی در حالت گسسته

ابتدا احتمال شرطی یک مجموعه A نسبت به مجموعه B را بررسی می کنیم، که به صورت زیر تعریف می شود:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}\ , \qquad P(B) \neq 0$$

در حالتی که $P(B) = 0$ باشد، این رابطه قابل تعریف نمی باشد. اکنون احتمال شرطی را بر این اساس در نظر بگیرید که مشاهده گر یک سری اطلاعات جزیی در اختیار داشته باشد. یک فضای احتمال $(\Omega, \mathcal{F}, P)$ عملکرد یک مکانیسم را که به وسیله شانس P توصیف می کند، به این صورت که این مکانیسم ω ای تولید می کند که براساس P توزیع می شود. $P(A)$ برای مشاهده گر احتمال این است که نقطه ω تولید شده در A بیفتد. فرض کنید که ω در B باشد و مشاهده گر فقط همین اطلاع را داشته باشد. از دیدگاه مشاهده گر، با داشتن این اطلاعات جزیی درباره ω، احتمال اینکه ω در A هم باشد برابر با $P(A|B)$ است که به $P(A)$ ترجیح داده می شود. از سوی دیگر اگر $\omega \in B^c$ باشد، در این صورت احتمال مربوطه برابر با $P(A|B^c)$ می شود و خواهیم داشت:

$$f(\omega) = \begin{cases} P(A|B), & \omega \in B \\ P(A|B^c), & \omega \in B^c \end{cases}$$

مشاهده گر می داند که ω به B یا B^c تعلق دارد؛ بنابراین احتمال اینکه $\omega \in A$ باشد برابر با $f(\omega)$ است.

اگرچه مشاهده گر در حالت کلی نمی داند ω چیست، اما می تواند مقدار $f(\omega)$ را با توجه به تعلق به B یا B^c محاسبه کند. عکس این حالت، یعنی با دانستن مقدار $f(\omega)$ هنگامی می تواند بفهمد که ω در B یا در B^c است، که $P(A|B) = P(A|B^c)$ باشد، یعنی A و B مستقل باشند که در این حالت احتمال شرطی برابر با همان احتمال غیر شرطی $P(A)$ خواهد بود.

مجموعه های B و B^c فضای نمونه Ω را افراز می کنند، حال می خواهیم این ایده را برای یک افراز کلی از Ω داشته باشیم.

فرض کنید... $B_1, B_2,$ یک افراز متناهی یا قابل شمارش از فضای نمونه Ω و $\mathcal{G}$ سیگما میدان تولید شده توسط B_i ها (شامل تمام اجتماعات B_i ها) باشد.

برای هر $A\epsilon\mathcal{F}$، تابع $f(\omega)$ را به صورت زیر در نظر بگیرید:

$$f(\omega) = \frac{P(A \cap B_i)}{P(B_i)}\ , \qquad \omega\ \epsilon B_i, \qquad i = 1,2, \dots$$

اگر مشاهده گر بداند که کدام B_i از افراز Ω شامل ω است در این صورت با داشتن این اطلاع احتمال پیشامد $\{\omega\ \epsilon\ A\}$ برابر با $f(\omega)$ است.

تابع یا متغیر تصادفی f احتمال شرطی A به شرط $\mathcal{G}$ نامیده می شود و با نماد $P[A \parallel \mathcal{G}]$ یا $P[A \parallel \mathcal{G}]_\omega$ نشان داده می شود. $P[A \parallel \mathcal{G}]$ تابعی است که مقدار آن روی B_i برابر با همان احتمال شرطی معمول $P(A|B_i)$ می باشد. برای کامل شدن تعریف باید مشخص کنیم که اگر $P(B_i) = 0$ باشد، چه اتفاقی می افتد؟ در این حالت می توان $P[A \parallel \mathcal{G}]$ را برابر با هر مقدار ثابت دلخواهی روی B_i در نظر گرفت، که این مقدار اگرچه دلخواه است، اما باید برای همه B_i ها یکسان تعریف شود.

مثال ۱۳-۱: فرایند پواسن را در نظر بگیرید.

فرض کنید $0 \le s \le t$، $A = [N_s = 0]$ و $B_i = [N_t = i], i = 0,1, \dots$، از آن جایی که در فرایند پواسن نموها مستقل هستند و دارای توزیع پواسن می باشند داریم

$$\begin{aligned} P(A|B_i) &= P(N_s = 0|N_t = i) = \frac{P(N_s = 0, N_t = i)}{P(N_t = i)} \\ &= \frac{P(N_s = 0)P(N_t - N_s = i)}{P(N_t = i)} \\ &= \frac{(e^{-\lambda s}\ (\lambda s)^0/0!)(e^{-\lambda(t-s)}\ \left(\lambda(t-s)\right)^i/i!)}{\left(e^{-\lambda t}\,\frac{(\lambda t)^i}{i!}\right)} = (1 - \frac{s}{t})^i \end{aligned}$$

در نتیجه:

$$P[N_s = 0 \parallel \mathcal{G}]_\omega = (1 - \tfrac{s}{t})^i, \quad \omega\ \epsilon B_i, i = 0,1, \dots \qquad (۱۳-۱)$$

چون روی B_i، $N_t(\omega) = i$ است عبارت فوق را می توان به صورت زیر نوشت:

$$P[A \parallel \mathcal{G}] = \begin{cases} (1-\frac{s}{t})^{N_t(\omega)}, & \omega \in B_i \\ 0 \quad , & \omega \notin B_i \end{cases}$$

در این مثال لازم نیست مقدار ω را بدانیم بلکه کافیست $N_t(\omega)$ را بدانیم. در این جا مشاهده‌گر تعداد حوادثی که در بازه زمانی $[0,t]$ رخ داده است را می داند، ولی جایگاه این حوادث را در بازه $[0,t]$ که همان ω است نمی داند. عبارت (۱۴-۱) بیانگر این احتمال است که هیچ یک از پیشامدها تا قبل از زمان s رخ ندهد.

مثال ۱۳-۲: فرض کنید $X_0, X_1, \dots$ یک زنجیر مارکف با فضای حالت S باشد. پیشامدهای

$$[X_0 = i_0, X_1 = i_1, \dots, X_n = i_n], \quad i_0, \dots, i_n \in S$$

یک افراز متناهی یا قابل شمارش از فضای نمونه Ω تشکیل می دهند. فرض کنید $\mathcal{G}_n$ سیگما میدان تولید شده توسط این افراز باشد. در این صورت

$$\begin{aligned} P[X_{n+1} = j \parallel \mathcal{G}_n]_\omega &= P[X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n] \\ &= P[X_{n+1} = j | X_n = i_n] = p_{i,j} \quad \forall \omega \epsilon [X_0 = i_0, X_1 = i_1, \dots, X_n = i_n] \end{aligned}$$

از طرفی ${\mathcal{G}_n}^0$ را سیگما میدان تولید شده به وسیله پیشامدهای $[X_n = i], \quad i \in S$ که فضای نمونه Ω را افراز می کنند، در نظر بگیرید. سیگما میدان تولید شده کوچکتر از سیگما میدان $\mathcal{G}_n$ است. از طرفی

$$P[X_{n+1} = j \parallel {\mathcal{G}_n}^0]_\omega = p_{i,j} \quad \forall \omega \epsilon [X_n = i]$$

در نتیجه با توجه به خاصیت مارکفی داریم

$$P[X_{n+1} = j \parallel \mathcal{G}_n]_\omega = P[X_{n+1} = j \parallel {\mathcal{G}_n}^0]_\omega$$

در هر دو مثال اخیر لازم نیست خود ω را بدانیم، تنها کافی است با توجه به تعلق به B_i ها، احتمال آن را محاسبه کنیم.

۱۳-۲ احتمال شرطی در حالت کلی

احتمال شرطی را ابتدا روی افراز Ω تعریف کرده، سپس به سیگما میدان تولید شده توسط این افراز تعمیم دادیم؛ اما از آن‌جایی که یک سیگما میدان لزومی ندارد حتما توسط یک افراز تولید شود، تابع احتمال شرطی را روی سیگما میدان در حالت کلی تعریف می کنیم. فضای اندازه (Ω, F, P) را در نظر گرفته و $\mathcal{G}$ را سیگما میدانی در $\mathcal{F}$ در نظر بگیرید. اکنون هدف تعریف $P[A \parallel \mathcal{G}]$ است. بدین منظور A را در $\mathcal{F}$ ثابت در نظر گرفته و اندازه متناهی v را روی سیگما میدان $\mathcal{G}$ به صورت زیر تعریف می کنیم:

$$\forall A\epsilon\mathcal{F}:\ \upsilon(G) = P(A\cap G),\qquad G\epsilon\mathcal{G}$$

در این صورت $P(G) = 0$ ایجاب می کند $P(A\cap G) = 0$ و در نتیجه $\nu(G) = 0$ می باشد.

$$P(G) = 0 \Rightarrow \nu(G) = 0 \Rightarrow\ \nu \ll P$$

از آنجایی که ν و P هر دو اندازه متناهی و $\nu \ll P$، می توان قضیه رادون- نیکودایم را برای اندازه های ν و P روی فضای اندازه پذیر $(\Omega,\mathcal{G})$ به کار برد. در نتیجه بنابر قضیه رادون نیکودایم، یک تابع یا متغیر تصادفی f (اندازه پذیر $\mathcal{G}$ و انتگرال پذیر نسبت به P) وجود دارد به طوری که

$$P(A\cap G) = \upsilon(G) = \int_G f dp\,,\qquad \forall G\epsilon\mathcal{G}$$

بنابراین در حقیقت احتمال شرطی مشتق رادون نیکودایم دو اندازه است.

تابع f یا $P[A \parallel \mathcal{G}]$ متغیری تصادفی با دو ویژگی زیر می باشد:

الف) $P[A \parallel \mathcal{G}]$ انتگرال پذیر و $\mathcal{G}$ اندازه پذیر است.

ب) $P[A \parallel \mathcal{G}]$ در رابطه زیر صدق می کند:

$$\int_G P[A \parallel \mathcal{G}]\, dp = P(A\cap G),\qquad G\epsilon\mathcal{G}$$

شرط (الف) ایجاب می کند که مقدار $P[A \parallel \mathcal{G}]$ فقط به مجموعه های درون سیگما میدان $\mathcal{G}$ بستگی داشته باشد. پس در حالت کلی هم، به دانستن مقدار ω نیاز نداریم و تنها کافی است که بدانیم G شامل ω هست یا نه.

در حالت کلی متغیرهای تصادفی $P[A \parallel \mathcal{G}]$ زیادی وجود دارند که با احتمال ۱ باهم برابر بوده و هرکدام را یک صورت از احتمال شرطی می گویند.

شرط (ب) تعبیر قماربازی دارد. فرض کنید که یک قمارباز پس از مشخص شدن اینکه نتیجه، عضوی از $\mathcal{G}$ است، تصمیم بگیرد روی پیشامد A شرط بندی کند (اگر $A \notin \mathcal{G}$ ، مشاهده گر نمی داند که A رخ داده است یانه). شخص باید مبلغ $P[A \parallel \mathcal{G}]$ واحد (به عنوان نرخ ورودی) بپردازد. اگر A رخ دهد ۱ واحد برنده می شود، در غیر این صورت هیچ مقداری دریافت نمی کند. حال اگر مشاهده گر تصمیم به شرط بندی بگیرد و مبلغ ورودی را بپردازد، در صورتی که A رخ دهد $1 - P[A \parallel \mathcal{G}]$ سود، و در صورتی که A رخ ندهد $P[A \parallel \mathcal{G}]$ واحد ضرر خواهد کرد. در نتیجه سود او برابر است با:

$$(1 - P[A \parallel \mathcal{G}])I_A - P[A \parallel \mathcal{G}]I_{A^c} = I_A - P[A \parallel \mathcal{G}]$$

و اگر از شرط بندی صرف نظر کند سود برابر صفر خواهد بود.

فرض کنید این فرد تصمیم گیری خود را بر این مبنا قرار دهد که اگر G رخ دهد شرط بندی را قبول و در غیر این صورت نپذیرد. به بیان دیگر می تواند تصمیم را پس از انجام آزمایش و مشاهده برآمد ممکن بگیرد. سود مورد انتظار با این استراتژی (امید ریاضی سود) برابر است با

$$\int_G (I_A - P[A \parallel \mathcal{G}])dp = \int_G I_A \, dp - \int_G P[A \parallel \mathcal{G}]dp = P(A \cap G) - \int_G P[A \parallel \mathcal{G}]dp = 0$$

صفر شدن این رابطه نشان دهنده عادلانه بودن این استراتژی است. به این معنی که مشاهده گر به ازای هر $G\epsilon\mathcal{G}$ (که برای تصمیم گیری انتخاب می شود) متوسط سودش صفر باشد.

مثال۱۳-۳ فرض کنید $A\epsilon\mathcal{G}$ و $\mathcal{G} = \mathcal{F}$ باشد، تابع نشانگر $I_A = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in A^c \end{cases}$ را نیز در نظر بگیرید. می خواهیم نشان دهیم $P[A||\mathcal{G}] = \mathrm{I_A}$ ، صورتی از احتمال شرطی است. از آن جایی که $\mathrm{I_A}$ انتگرال پذیر و $\mathcal{G}$ اندازه پذیر می باشد، شرط (الف) برقرار است. از طرفی $\int_G \mathrm{I_A}\, dp = P(A \cap G)$ ، در نتیجه با توجه به شرط (ب) $P[A||\mathcal{G}] = \mathrm{I_A}$ می باشد یعنی $\mathrm{I_A}$ یک صورت از احتمال شرطی $P[A||\mathcal{G}]$ است.

مثال ۱۳-۴: فرض کنید $\mathcal{G}$ کوچکترین سیگما میدان باشد یعنی $\mathcal{G} = \{\phi, \Omega\}$. در این صورت هر تابع $\mathcal{G}$ اندازه پذیر، یک تابع ثابت می باشد. می خواهیم نشان دهیم $P[A||\mathcal{G}] = \mathrm{P(A)}$، یک صورت از احتمال شرطی است. از آنجایی که $\mathrm{P(A)}$ انتگرال پذیر و$\mathcal{G}$ اندازه پذیر است، شرط (الف) برقرار است. از طرف دیگر،

$$\int_G P(A)dp = P(A \cap G) = \begin{cases} \int_\phi P(A)dp = 0 = P(A \cap \phi) \\ \int_\Omega P(A)dp = P(A) = P(A \cap \Omega) \end{cases}$$

بنابراین تابع $\mathrm{P(A)}$ صورتی از احتمال شرطی $P[A||\mathcal{G}]$ می باشد. در حقیقت مشاهده گر از انجام آزمایش

$\mathcal{G} = \{\Phi, \Omega\}$ اطلاعی درباره A به دست نمی آورد.

با توجه به دو مثال فوق در می یابیم که هر چه سیگما میدان کوچکتر شود شرط (الف) سخت تر برقرار می شود و توابع کمتری در آن صدق می کنند. در حقیقت وقتی سیگما میدان به کوچکترین اندازه ممکن $\{\Phi, \Omega\}$ برسد تنها توابع ثابت در آن صدق می کنند که به خاطر شرط (ب) تابع $P[A||\mathcal{G}] = \mathrm{P(A)}$ را بر می گزینیم. توجه کنید که تابع ثابت $\mathrm{P(A)}$ دارای کمترین اطلاع در مورد $P[A||\mathcal{G}]$ می باشد. از طرفی با کوچک شدن سیگما میدان شرط (ب) راحت تر برقرار می شود. هرچه سیگما میدان بزرگتر شود برقراری شرط (ب) سخت تر است و زمانی که بزرگترین سیگما میدان را داریم تنها تابع نشانگر $I_A = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in A^c \end{cases}$ در شرط (ب) صادق است.

مثال ۱۳-۵: فرض کنید Ω سطحی دوبعدی در R^2 باشد ($\Omega = \{(x,y): x\epsilon R, y\epsilon R\}$) و $\mathcal{F}$ گردایه $\mathfrak{R}^2$ از مجموعه های بورل دوبعدی باشد. سیگما میدان $\mathcal{G}$ را شامل نوارهای عمودی مجموعه های حاصلضربی

$E \times R^1 = [(x,y): x \in E]$ در نظر بگیرید که در آن E یک مجموعه بورل خطی باشد. اگر مشاهده گر برای هر عضو $\mathcal{G}$ بداند که (x,y) عضو آن هست یا نه، در حقیقت مقدار x را می‌داند. آزمایش $\mathcal{G}$ مربوط به تعیین مولفه اول مختصات نمونه ای (تعیین x) است. فرض کنید P اندازه احتمال روی $\mathfrak{R}^2$ با چگالی $f(x,y)$ نسبت به اندازه لبگ دو بعدی باشد.

$$P(A) = \iint_A f(x,y)dxdy$$

A را نوار افقی $R^1 \times F = [(x,y): y \in F]$، که در آن F یک مجموعه بورل خطی است، در نظر بگیرید ($A = \{R^1 \times F, F \in \mathfrak{R}\}$). در اینجا نشان می دهیم که احتمال شرطی $P[A||\mathcal{G}]$ می تواند به فرم صریح زیر محاسبه شود.

$$\varphi(x,y) = \frac{\int_F f(x,t)dt}{\int_{R^1} f(x,t)dt}$$

و $\varphi(x,y)$ در نقاطی که مخرج به سمت صفر میل می کند، برابر با صفر است. این نقاط تشکیل یک مجموعه با احتمال صفر را می دهند. می خواهیم نشان دهیم که تابع $\varphi(x,y)$ صورتی از احتمال شرطی $P[A||\mathcal{G}]$ است، یعنی در شرط (الف) و (ب) صدق می کند.

از آنجایی که تابع $\varphi(x,y)$ تنها تابعی برحسب x است و $\mathcal{G}$ شامل عبارات $E \times R^1$ است، $\varphi(x,y)$ نسبت به $\mathcal{G}$ اندازه پذیر است.

به منظور بررسی شرط ب باید نشان دهیم:

$$\int_{E\times R^1} \varphi(x,y)dp(x,y) = P(A \cap (E \times R^1)) \qquad (۲-۱۴)$$

چون $A = \{R^1 \times F, F \in \mathfrak{R}\}$ است، پس سمت راست رابطه برابر $P[E \times F]$ می شود.

$$\begin{aligned} P\big(A \cap (E \times R^1)\big) &= P\big((R^1 \times F) \cap (E \times R^1)\big) = P\big((R^1 \cap E) \times (F \cap R^1)\big) \\ &= P[E \times F] \end{aligned}$$

و بنابر قضیه فوبینی داریم:

$$\int_{E\times R^1}\varphi(x,y)dp(x,y)=\int_E\int_{R^1}\varphi(x,y)f(x,y)dydx=\int_E\varphi(x,y)\int_{R^1}f(x,y)dydx$$
$$=\int_E\frac{\int_F f(x,t)dt}{\int_{R^1}f(x,t)dt}\int_{R^1}f(x,y)dydx=\int_E\int_F f(x,t)dtdx$$
$$=\iint\limits_{E\times F}f(x,y)dxdy=P[E\times F]$$

بنابراین،

$$\int_{E\times R^1}\varphi(x,y)dp(x,y)=P(A\cap(E\times R^1))=P[E\times F]$$

و $\varphi(x,y)$ صورتی از احتمال شرطی $P[A||\mathcal{G}]$ است که شرطی کردن احتمالات به شرط پیشامدهایی با احتمال صفر ($\{x\}\times R^1$) را پوشش می دهد.

مثال ۱۳-۶: مجموعه A طبق تعریف از سیگما میدان $\mathcal{G}$ مستقل است اگر از هر عضو آن مستقل باشد، یعنی

$$P[A\cap G]=\mathrm{P(A)P(G)}\ \ \forall G\in\mathcal{G}$$

این تعریف معادل آن است که $P[A||\mathcal{G}]=\mathrm{P(A)}$. زیرا،

$$P[A\cap G]=\int\limits_{\mathrm{G}}P[A||\mathcal{G}]dP=\int\limits_{\mathrm{G}}P[A]dP=\mathrm{P(A)}\int\limits_{\mathrm{G}}dP=\mathrm{P(A)P(G)}$$

فرض کنید $X:(\Omega,\mathcal{F})\to(R,\Re)$. در این صورت سیگما میدان $\sigma(X)$ که توسط متغیر تصادفی X تولید می شود، شامل مجموعه هایی به صورت زیر است:

$$\sigma(X)=\{\omega:X(\omega)\epsilon H\},H\epsilon\Re^i.$$

که کوچکترین سیگما میدانی در Ω است که X نسبت به آن متغیر تصادفی می باشد. احتمال شرطی A به شرط متغیر تصادفی X را می توان به صورت $P(A\parallel\sigma(X))$ یا به طور معادل با نماد $P(A\parallel X)$ نشان داد. این تعریف را می توان به بردار متغیرهای تصادفی و یا به طور معادل به هر مجموعه متناهی از متغیرهای تصادفی تعمیم داد.

برای هر مجموعه متناهی از متغیرهای تصادفی $[X_t,t\in T]$ ، سیگما میدان $\sigma[X_t,t\in T]$ که توسط X_t ها تولید می شود، کوچکترین سیگما میدانی است که X_t ها نسبت به آن اندازه پذیرند و احتمال شرطی

$P[A\parallel\sigma(X_t),t\epsilon T]$ یا $P[A\parallel X_t,t\epsilon T]$، احتمال شرطی A نسبت به مجموعه ای از متغیرهای تصادفی است.

مثلا احتمال شرطی در مثال زنجیره مارکوف به صورت زیر است:

$$P[X_{n+1} = j | X_0, \dots, X_n] = P[X_{n+1} = j | X_n]$$

یعنی احتمال شرطی $\{X_{n+1} = j\}$ برای کسی که X_n را می داند برابر کسی است که $X_0, \dots, X_n$ را می داند.

مثال۱۳-۷: فرض کنید X و Y بردارهای تصادفی مستقل j و k بعدی باشند، و μ توزیع X روی R^{j} باشد. به ازاء هر $H \in \Re^j$ و $J \in \Re^{j+k}$ داریم:

$$P[X \in H, (X,Y) \in J] = \int_H P[(x,Y) \in J] \mu(dx)$$

به ازاءهر $x \in R^j$، قرار می دهیم:

$$f(x) = P[(x,Y) \in J] = P[\omega' : (x, Y(\omega') \in J].$$

می خواهیم نشان دهیم $f(X(\omega))$ صورتی از احتمال شرطی$P[(X,Y) \in J \parallel X]_\omega$ است. $f(X(\omega))$ نسبت به $\sigma(X(\omega))$ اندازه پذیر است و از آن جایی که μ توزیع X است یک تغییر متغیر، نتیجه می‌دهد

$$\int_{[X \in H]} f\big(X(\omega)\big) P(d\omega) = \int_H f(x) \mu(dx) = P([(X,Y) \in J] \cap [X \in H])$$

از آن جایی که $[X \in H]$ عضو $\sigma(X)$ است، بنابراین:

$$f(X(\omega)) = P[(X,Y) \in J \parallel X]_\omega$$

مثال ۱۳-۸: فرض کنید X و Y متغیرهای تصادفی مستقل و Y دارای تابع توزیع F باشد. در نظر بگیرید:

$$J = [(u,v) : \max\{u,v\} \le m]$$

در این صورت$f(x)$ معرفی شده در مثال قبل به صورت زیر می باشد:

$$f(x) = P[(x,Y) \in J] = \begin{cases} 0, & m < x \\ F(m), & m \ge x \end{cases}$$

حال اگر $M = \max\{X,Y\}$، داریم

$$\begin{aligned} f\big(X(\omega)\big) &= P[(X,Y) \in J \parallel X]_\omega = P[\max\{X,Y\} \le m \parallel X]_\omega = P[M \le m \parallel X]_\omega \\ &= I_{[X \le m]} F(m), \quad w.p.1 \end{aligned}$$

حال قضیه ای را مطرح می کنیم که برای بررسی احتمالات شرطی مفید است.

قضیه ۱۳-۱: فرض کنید $\mathcal{P}$ یک π سیستم باشد که سیگما میدان $\mathcal{G}$ را تولید می کند، و فضای نمونه Ω به صورت اجتماع متناهی یا شمارا از مجموعه هایی در $\mathcal{P}$ باشد. در این صورت تابع انتگرال پذیر fصورتی از احتمال شرطی $P[A||\mathcal{G}]$ است. اگر fاندازه پذیر $\mathcal{G}$ باشد و

$$\int_G f dp = P(A \cap G), \quad \forall G \epsilon \mathcal{P}$$

اثبات: دو اندازه متناهی$\upsilon_1(G) = \int_G f dp$ و $\upsilon_2(G) = P(A \cap G)$ را روی سیگما میدان $\mathcal{G}$ در نظر بگیرید. از آن جایی که به ازاء هر $G \epsilon \mathcal{P}$ داریم $\int_G f \mathrm{dP} = \mathrm{P}(\mathrm{A} \cap \mathrm{G})$، یا به طور معادل $\upsilon_1(G) = \upsilon_2(G)$ است، پس شرط دوم برقرار است. شرط اول نیز جزء شرایط قضیه است. در نتیجه اثبات کامل می شود.

مثال ۱۳-۹: مجموعه $[X_t : t \geq 0]$ از متغیرهای تصادفی یک فرایند مارکف در زمان پیوسته است اگر برای $k \geq 1$، $0 \leq t_1 \leq \cdots \leq t_k \leq u$ و $H \in \mathfrak{R}^1$ عبارت زیر با احتمال یک برقرار باشد:

$$P\left[X_u \epsilon H \middle| X_{t_1}, \dots, X_{t_k}\right] = P\left[X_u \epsilon H \middle| X_{t_k}\right] \qquad (۳\text{-}۱۴)$$

رابطه بالا مشابه تعریف فوق برای حالت زمان-گسسته است. فرض کنید ، با در نظر گرفتن $t \leq u$ سمت راست عبارت (۱۴-۳) به عنوان یک صورت از احتمال شرطی سمت چپ، داریم

$$\int_G P[X_u \epsilon H \parallel X_t] dp = P[(X_u \epsilon H) \cap G]$$

برای $0 \leq t_1 \leq \cdots \leq t_k = t \leq u$ و $G \in \sigma(X_{t_1}, \dots, X_{t_k})$. حال t ، u و H را ثابت، و $t_1, \dots, t_k$ و k را متغییر در نظر می گیریم. گردایه $\mathcal{P} = \cup\, \sigma(X_{t_1}, \dots, X_{t_k})$ را درنظر بگیرید. اگر $A \in \sigma(X_{t_1}, \dots, X_{t_k})$ و $B \in \sigma(X_{s_1}, \dots, X_{s_i})$، آنگاه $A \in \sigma(X_{r_1}, \dots, X_{r_j})$، که در آن r_α ها جایی است که s_γ ها و t_β ها با هم ادغام می شوند. در نتیجه $\mathcal{P}$ یک π -سیستم است. از آن جایی که $\mathcal{P}$ سیگما میدان$\sigma[X_s : s \leq t]$ را تولید می کند و $P[X_u \epsilon H \parallel X_t]$ نسبت به این سیگما میدان اندازه پذیر است، با توجه به رابطه (۱۴-۳) و قضیه ۱۴-۱ نتیجه می شود که $P[X_u \epsilon H \parallel X_t]$ صورتی از احتمال شرطی $P[X_u \epsilon H \parallel X_s, s \leq t]$ است. یعنی با احتمال یک داریم

$$P[X_u \epsilon H \parallel X_t] = P[X_u \epsilon H \parallel X_s, s \leq t], \qquad t \leq u$$

این مثال بیان می کند که برای محاسبه احتمالات شرطی آینده در فرایند مارکوف زمان-پیوسته،$\sigma(X_t)$ زمان حال معادل با$\sigma[X_s : s \leq t]$ زمان حال و گذشته است.

قضیه ۱۴-۲ (ویژگی های احتمال شرطی):

الف) به ازاء هر A، $0 \leq P[A \parallel \mathcal{G}] \leq 1$.

ب) اگر $A_1, A_2, \ldots$ یک دنباله متناهی یا قابل شمارش از مجموعه های ناسازگار باشد، آنگاه:

$$P\left[\bigcup_n A_n \parallel \mathcal{G}\right] = \sum_n P[\,A_n \parallel \mathcal{G}]$$

ج) اگر $A \subset B$، آنگاه:

$$P[A \parallel \mathcal{G}] \leq P[B \parallel \mathcal{G}]$$

و

$$P[B - A \parallel \mathcal{G}] = P[B \parallel \mathcal{G}] - P[A \parallel \mathcal{G}].$$

د)

$$P\left[\bigcup_{i=1}^{n} A_i \parallel \mathcal{G}\right] = \sum_{i=1}^{n} P[\,A_i \parallel \mathcal{G}] - \sum_{i<j} P[A_i \cap A_j \parallel \mathcal{G}] + \cdots$$

ه) اگر $A_n \uparrow A$ ، آنگاه $P[A_n \parallel \mathcal{G}] \uparrow P[A \parallel \mathcal{G}]$،

و) اگر $A_n \downarrow A$ ، آنگاه $P[A_n \parallel \mathcal{G}] \downarrow P[A \parallel \mathcal{G}]$،

ز) اگر $P(A) = 1$، آنگاه $P[A \parallel \mathcal{G}] = 1$ و اگر $P(A) = 0$ ، آنگاه $P[A \parallel \mathcal{G}] = 0$ می باشد.

اثبات:

الف) برای هر صورت احتمال شرطی داریم

$$\int_G P[A||\mathcal{G}]dP = P(A \cap G) \geq 0, \forall G \epsilon \mathcal{G}$$

از آن جایی که $P[A||\mathcal{G}]$ اندازه پذیر $\mathcal{G}$ است، در نتیجه از رابطه بالا نتیجه می شود که $P[A||\mathcal{G}]$ ، به جز روی یک مجموعه که احتمالش صفر است، باید نامنفی باشد. از طرف دیگرداریم

$$\int_G P[A||\mathcal{G}]\,dp = P(A \cap G) \leq P(G) = \int_G 1\,dp \qquad , \forall G \in \mathcal{G}$$

چون P سیگما متناهی است نتیجه می‌شود که $P[A||\mathcal{G}] \leq 1$.

ب) باید نشان دهیم که $\sum_n P[\,A_n \parallel \mathcal{G}]$ صورتی از احتمال شرطی $P[\bigcup_n A_n \parallel \mathcal{G}]$ می باشد. از آن جایی که $\sum_n P[\,A_n \parallel \mathcal{G}]$ اندازه پذیر $\mathcal{G}$ است، کافی است که در شرط (ب) صدق کند. از آنجایی که A_n ها مجزا و $P[A_n||\mathcal{G}]$ نامنفی است، برای هر $G \in \mathcal{G}$ داریم:

$$\int_G \sum_n P[A_n \parallel \mathcal{G}]\, dp = \sum_n \int_G P[A_n \parallel \mathcal{G}]\, dp$$
$$= \sum_n P(A_n \cap G)$$
$$= P\left(\left(\bigcup_n A_n\right) \cap G\right)$$

در نتیجه $P[\bigcup_n A_n \parallel \mathcal{G}] = \sum_n P[A_n \parallel \mathcal{G}]$.

ج) از آن جایی که $A \subset B$ ، داریم $B = A \cup (B - A)$، در نتیجه:

$$P[B \parallel \mathcal{G}] = P\big[\big(A \cup (B - A)\big) \parallel \mathcal{G}\big] = P[A \parallel \mathcal{G}] + P[B - A \parallel \mathcal{G}],$$

بنابراین $P[B - A \parallel \mathcal{G}] = P[B \parallel \mathcal{G}] - P[A \parallel \mathcal{G}]$.

از طرفی $P[B - A \parallel \mathcal{G}] = P[B \parallel \mathcal{G}] - P[A \parallel \mathcal{G}] \geq 0$، پس $P[B \parallel \mathcal{G}] \geq P[A \parallel \mathcal{G}]$.

د) مانند اثبات قسمت الف داریم:

$$\int_G \Big(\sum_{i=1}^{n} P[A_i \parallel \mathcal{G}] - \sum_{i<j} P[A_i \cap A_j \parallel \mathcal{G}] + \cdots\Big) dp$$
$$= \int_G \sum_{i=1}^{n} P[A_i \parallel \mathcal{G}]\, dp - \int_G \sum_{i<j} P[A_i \cap A_j \parallel \mathcal{G}]\, dp + \cdots$$
$$= \sum_i \int_{\mathrm{G}} P[A_i \parallel \mathcal{G}] dp - \sum_{i<j} \int_{\mathrm{G}} P\big[A_i \cap A_j \parallel \mathcal{G}\big] dp + \cdots$$
$$= \sum_{i=1}^{n} P[A_i \cap G] - \sum_{i<j} P\big[(A_i \cap A_j) \cap G\big] + \cdots = P\left[\left(\bigcup_{i=1}^{n} A_i\right) \cap G\right]$$

بنابراین عبارت $\sum_{i=1}^{n} P[A_i \parallel \mathcal{G}] - \sum_{i<j} P[A_i \cap A_j \parallel \mathcal{G}] + \cdots$ صورتی از احتمال شرطی $P[\bigcup_{i=1}^{n} A_i \parallel \mathcal{G}]$ است.

ه) باید نشان دهیم $\lim_{n\to\infty} \mathrm{P}[\mathrm{A_n} \parallel \mathcal{G}] = P[\mathrm{A} \parallel \mathcal{G}]$. با توجه به $\mathcal{G}$ اندازه پذیر بودن $\lim_{n\to\infty} \mathrm{P}[\mathrm{A_n} \parallel \mathcal{G}]$ و

$$\int_G \lim_{n\to\infty} \mathrm{P}[\mathrm{A_n} \parallel \mathcal{G}]\ \mathrm{dp} = \lim_{n\to\infty} \int_{\mathrm{G}} \mathrm{P}[\mathrm{A_n} \parallel \mathcal{G}]\ dp = \lim_{n\to\infty} \mathrm{P}(\mathrm{A_n} \cap \mathrm{G}) = \mathrm{P}(\mathrm{A} \cap \mathrm{G})$$

شرط (الف) و (ب) برقرار است و $\lim_{n\to\infty} \mathrm{P}[\mathrm{A_n} \parallel \mathcal{G}]$ صورتی از احتمال شرطی $\mathrm{P}[\mathrm{A} \parallel \mathcal{G}]$ است.

و) اثبات این قسمت مشابه قسمت (ه) می باشد.

ز) باید نشان دهیم که ۱ صورتی از احتمال شرطی $P[\mathrm{A} \parallel \mathcal{G}]$ است. می دانیم اگر $P(A) = 1$ باشد، آنگاه A مستقل از هر پیشامدی مانند G است که در رابطه $P(A \cap G) = P(A)P(G) = P(G)$ صدق کند. بنابراین $\int_G 1\,\mathrm{dp} = \mathrm{P(G)} = \mathrm{P(A \cap G)}$ ، و ۱ صورتی از احتمال شرطی $P[\mathrm{A} \parallel \mathcal{G}]$ است.

حال اگر $P(A) = 0$ باشد، از آنجایی که $\mathrm{A \cap G}$ زیرمجموعه A است داریم $P(A \cap G) \leq P(A) = 0$ ، در نتیجه $P(A \cap G) = 0$. با توجه به اینکه $\int_G 0\,\mathrm{dp} = 0 = \mathrm{P(A \cap G)}$ ، صفر صورتی از احتمال شرطی $P[\mathrm{A} \parallel \mathcal{G}]$ است.

۱۳–۳ توزیع های احتمال شرطی

فرض کنید X متغیر تصادفی روی فضای اندازه $(\Omega, \mathcal{F}, P)$ باشد و $\mathcal{G}$ یک سیگما میدان در $\mathcal{F}$ باشد.

قضیه ۱۳–۳: برای هر مجموعه H در بورل $\mathcal{R}^1$ و $\epsilon\omega\Omega$، تابع $\mu(H, \omega)$ با دو ویژگی زیر وجود دارد:

الف) به ازاء هر $\epsilon\omega\Omega$ ، $\mu(., \omega)$ یک اندازه احتمال روی $\mathcal{R}^1$ است.

ب) به ازاء هر $H\epsilon\mathcal{R}^1$، $\mu(H, .)$ ورژنی از احتمال شرطی $P[X \in H \parallel \mathcal{G}]$ است.

اندازه احتمال $\mu(., \omega)$ توزیع شرطی X به شرط $\mathcal{G}$ است.

اگر $\mathcal{G} = \sigma(Z)$ باشد، $\mu(., \omega)$ توزیع شرطی X به شرط متغیر تصادفی Z است.

مثال ۱۳–۱۰: فرض کنید X و Y دو متغیر تصادفی با توزیع توأم υ در R^2، و چگالی $f(x, y)$ نسبت به اندازه لبگ باشند:

$$P[(X, Y) \in A] = \upsilon(A) = \iint_A f(x, y)dxdy$$

قرار دهید $g(x, y) = f(x, y) / \int_{R^1} f(x, t)dt$، و $\mu(H, x) = \int_H g(x, y)dy$ دارای چگالی $g(., x)$ باشد. اگر $\int_{R^1} \mathrm{f(x, t)dt} = 0$ باشد، $\mu(., x)$ را برابر یک مقدار دلخواه روی خط در نظر می گیریم. بنابراین $\mu(H, X(\omega))$ درواقع توزیع شرطی Y به شرط X خواهد بود. در حقیقت رابطه (۱۳-۲)، معادل است با:

$$\int_{E\times R^1} \mu(F,x)d\upsilon(x,y) = \upsilon(E\times F)$$

که با تغییر متغیر نتیجه می شود:

$$\int_{[X\in E]} \mu\big(F, X(\omega)\big)\, P(d\omega) = P[X\in E, Y\in F]$$

بنابراین $\mu\big(F, X(\omega)\big)$ صورتی از $P[Y\in F\|X]_\omega$ است.

۱۳–۴ مسائل

۱) در فرایند پواسن نشان دهید برای $0 < s < t$ توزیع شرطی N_s به شرط N_t توزیع دوجمله ای با پارامترهای N_t و s/t است.

$$P[N_s = k\|N_t] = \begin{cases} \binom{N_t}{k}(s/t)^k(1-s/t)^{N_t-k}, k\le N_t \\ 0, k > N_t \end{cases}.$$

۲) فرم زیر از قضیه بیز را اثبات کنید

$$P(G|A) = \frac{\int_G P[A\|\mathrm{g}]dp}{\int_\Omega P[A\|\mathrm{g}]dp}, \forall G\in \mathrm{g}.$$

۳) فرض کنید X و Y متغیرهای تصادفی مستقل با تابع توزیع F (که مثبت و پیوسته است) باشند. احتمال شرطی $[X\le x]$ به شرط متغیر تصادفی $M = \max\{X,Y\}$ را به دست آورید.

۴) در تمرین (۱۳-۲) نشان دهید که توزیع شرطی X به شرط M به صورت زیر می باشد؛ که در آن μ توزیع مرتبط با F (مثبت و پیوسته) است.

$$\frac{1}{2}I_{[M\in H]}(\omega) + \frac{1}{2}\frac{\mu(H\cap(-\infty,M(\omega)])}{\mu\big((-\infty,M(\omega)]\big)}$$

فصل چهاردهم

امید شرطی

۱۴–۱ تعریف

فرض کنید X یک متغیر تصادفی انتگرالپذیر روی $(\Omega,\mathcal{F},P)$ و $\mathcal{G}$ سیگما میدانی در $\mathcal{F}$ باشد. در اینصورت یک متغیر تصادفی $E[X\|\mathcal{G}]$ وجود دارد که مقدار امید شرطی X به شرط $\mathcal{G}$ نامیده می‌شود و دارای دو خاصیت زیر می‌باشد:

i. $E[X\|\mathcal{G}]$ نسبت به $\mathcal{G}$ اندازه‌پذیر و انتگرالپذیر است

ii. $\int_G E[X\|\mathcal{G}]\, dP = \int_G X\, dP \qquad G\epsilon\mathcal{G}$ (۱-۱۴)

لازم به ذکر است که نمی‌توان نتیجه گرفت $E[X\|\mathcal{G}] = X$ زیرا X نسبت به $\mathcal{F}$ اندازه‌پذیر است ولی $E[X\|\mathcal{G}]$ نسبت به $\mathcal{G}$ اندازه پذیر می‌باشد.

برای اثبات وجود چنین متغیر تصادفی دو حالت پیش می‌آید که هر کدام را جداگانه بررسی می‌کنیم:

۱- $X \geq 0$: یک اندازه مانند ν را روی $\mathcal{G}$ بصورت $\nu(G) = \int_G X\, dP$ تعریف می‌کنیم. چون X انتگرالپذیر است، با کاربرد قضیه رادون نیکودایم تابع f وجود دارد که نسبت به $\mathcal{G}$ اندازه‌پذیر است و

$\nu(G) = \int_G f\, dP$، مشخصا این f خواص i و ii را دارا می‌باشد.

۲- اگر X لزوما نامنفی نباشد خواهیم داشت:

$$\int\limits_G (E[X^+\|\mathcal{G}] - E[X^-\|\mathcal{G}])\, dP = \int\limits_G E[X^+\|\mathcal{G}]\, dP - \int\limits_G E[X^-\|\mathcal{G}]\, dP$$

$$= \int_G X^+\, dP - \int_G X^-\, dP = \int_G (X^+ - X^-)\, dP = \int_G X\, dP$$

پس در این حالت نیز تابع $E[X^+\|\mathcal{G}] - E[X^-\|\mathcal{G}]$ وجود دارد که خاصیت i و ii را دارا می‌باشد.

به طور کلی متغیرهای تصادفی زیادی هستند که مانند $E[X\|\mathcal{G}]$ عمل می‌کنند. هر یک از آنها یک صورت از مقدار امید شرطی نامیده می‌شوند و هر دو صورت با احتمال یک با یکدیگر برابرند.

در حالت خاصی که $\mathcal{G}$ کوچکترین سیگما میدان ممکن ($\mathcal{G} = \{\emptyset, \Omega\}$) باشد، خواهیم داشت:

$E[X\|\{\emptyset,\Omega\}] = E[X]$. حال نشان می‌دهیم $E[X]$ صورتی از امید شرطی است.

چون $\mathcal{G}$ تنها شامل دو عضو است فقط برای این دو عضو بررسی می‌کنیم:

$$\int_G E(X)\,dP = \begin{cases} \int_\emptyset E(X)\,dP = 0 = \int_\emptyset X\,dP \\ \int_\Omega E(X)\,dP = E(X)\int_\Omega dP = E(X) = \int_\Omega X\,dP \end{cases}$$

در حالت خاصی که $\mathcal{G}$ بسیار بزرگ و برابر با $\mathcal{F}$ انتخاب شود، می‌توان نتیجه گرفت $E[X\|\mathcal{F}] = X$ و در نتیجه می‌توان گفت X صورتی از امید شرطی است (X انتگرال پذیر و نسبت به $\mathcal{F}$ اندازه‌پذیر است، همچنین همواره $\int_G E(X)\,dP = \int_G X\,dP$)

زمانیکه $\mathcal{G}$ بزرگتر می‌شود و به سمت $\mathcal{F}$ می‌رود شرط i ضعیف‌تر می‌شود و برقراری آن ساده‌تر می‌شود، چون هر چه تعداد اعضای $\mathcal{G}$ بیشتر باشد توابع بیشتری نسبت به $\mathcal{G}$ اندازه‌پذیر خواهند بود. (زیرا باید تصویر معکوس آن تابع در $\mathcal{G}$ باشد و هر چه $\mathcal{G}$ بزرگتر باشد احتمال اینکه نقش معکوس آن تابع یکی از اعضای $\mathcal{G}$ باشد بیشتر است) در همین حین با بزرگتر شدن $\mathcal{G}$ شرط ii قوی‌تر شده و برقراری آن سخت‌تر می‌شود زیرا شرط ii باید برای همه $G\epsilon\mathcal{G}$ ها برقرار باشد و هر چه تعداد اعضای $\mathcal{G}$ بیشتر شوند احتمال خرابکاری تعدادی از G ها در رابطه ii بیشتر می‌شود که ممکن است رابطه ii را برهم زده و برقرار نسازند.

توجه: مقدار $E[X\|\mathcal{G}]_\omega$ در ω ، بصورت مقدار امید شرطی X تعبیر می‌شود برای کسی که $\forall G\epsilon\mathcal{G}$ می‌داند که آیا G شامل نقطه ω هست یا نه.

شرط ii را می‌توان چنین بیان کرد که اگر به مشاهده‌گر، با داشتن اطلاعات جزئی موجود در $\mathcal{G}$ پیشنهاد شرط‌بندی با این شرایط داده شود که مقدار ورودی $E[X\|\mathcal{G}]$ را بپردازد و در صورت برد مقدار X را پس بگیرد و چنانچه او استراتژی شرط‌بندی را بر مبنای وقوع G پایه‌ریزی کند، آنگاه عبارت $\int_G (X - E[X\|\mathcal{G}])\,dP = 0$ اذعان می‌دارد که بازی عادلانه است.

مثال ۱۴-۱ : فرض کنید $B_1, B_2, \ldots$ یک افراز متناهی و شمارشپذیر از Ω باشد که سیگما میدان $\mathcal{G}$ را تولید می‌کند. با فرض اینکه f= $E[X\|\mathcal{G}]$ باشد، اگر f بخواهد نسبت به $\mathcal{G}$ اندازه‌پذیر باشد باید f^{-1} عضوی از $\mathcal{G}$ باشد. از طرفی اگر f ثابت باشد f^{-1} عضو $\mathcal{G}$ است و f نسبت به $\mathcal{G}$ اندازه‌پذیر است. ولی اگر f ثابت نباشد f^{-1} عضو $\mathcal{G}$ نیست و f نسبت به $\mathcal{G}$ اندازه‌پذیر نخواهد بود.

اگر $P(B_i) = 0$ شود، مقدار $E[X\|\mathcal{G}]$ روی B_i ثابت ولی دلخواه است.

$E[X\|\mathcal{G}] = a_i$ را نامگذاری کنید:$(G = B_i)$

$$\int_{B_i} E[X\|\mathcal{G}]\,dP = \int_{B_i} X\,dP \quad \Rightarrow$$

$$\int_{B_i} \alpha_i\,dP = \int_{B_i} X\,dP \quad \Rightarrow \quad \int I_{B_i}\alpha_i\,dP = \int_{B_i} X\,dP$$

$$\Rightarrow \quad \alpha_i P(B_i) = \int_{B_i} X\, dP$$

$$\Rightarrow \quad \alpha_i = \frac{1}{P(B_i)} \int_{B_i} X\, dP \qquad \omega \epsilon B_i \quad , P(B_i) > 0 \qquad (۲-۱۴)$$

مثال ۱۴-۲: برای مجموعه A و تابع I_A مفاهیم $E[I_A \| \mathcal{G}]$ و $P[A \| \mathcal{G}]$ با احتمال ۱یکسانند.

اثبات: با در نظر گرفتن $X = I_A$ (با کاربرد خواص صورت احتمال شرطی) داریم

$$\int_G P[A\|\mathcal{G}]\, dP = P(A \cap G) = \int_G I_A\, dP = \int_G X\, dP = \int_G E[I_A\|\mathcal{G}]\, dP \qquad (۳-۱۴)$$

بطور کلی می‌توان گفت برای تابع ساده $X = \sum_i a_i I_{A_i}$ نیز با احتمال ۱ خواهیم داشت:

$$E[X\|\mathcal{G}] = \sum_i a_i P(A_i\|\mathcal{G})$$

در حقیقت باید نشان دهیم $\sum_i a_i P(A_i\|\mathcal{G})$ صورتی از امید شرطی است :

$$\int_G \sum_i a_i P(A_i\|\mathcal{G})\, dP = \sum_i a_i \int_G P(A_i\|\mathcal{G})\, dP = \sum_i a_i \int_G I_{A_i}\, dP =$$

$$= \int_G \sum_i a_i I_{A_i}\, dP = \int_G X\, dP$$

توجه: اگر $[X_t, t\epsilon T]$ مجموعه‌ای از متغیرهای تصادفی باشد، $E[X\|X_t, t\epsilon T]$ به صورت $E[X\|\mathcal{G}]$ با در نظر گرفتن $\sigma[X_t, t\epsilon T]$ به عنوان سیگما میدان در نقش $\mathcal{G}$ تعریف می‌شود.

۱۴–۲ ویژگی های امید شرطی

لم ۱۴–۱: اگر P یک π سیستم باشد که سیگما میدان $\mathcal{F}$ را تولید کند و Ω اجتماع شمارشپذیری از مجموعه‌های داخل P باشد، آنگاه برای هر f و g انتگرالپذیر که $\int_A f\, dP = \int_A g\, dP$ برای هر $A\epsilon P$ برقرار باشد، داریم

$$f = g \qquad (\mathrm{a.e})$$

قضیه زیر بیان می‌کند که شرایط i و ii روی π سیستم‌ها نیز برقرار است.

قضیه ۱۴–۱: با فرض اینکه P یک π سیستم باشد که سیگما میدان $\mathcal{G}$ را تولید کند و Ω اجتماع شمارشپذیری از مجموعه‌های داخل $\mathcal{G}$ باشد، اگر تابع انتگرالپذیر f نسبت به $\mathcal{G}$ اندازه‌پذیر باشد و داشته باشیم:

$$\int_G E[X\|\mathcal{G}]\, dP = \int_G f\, dP = \int_G X\, dP \qquad \forall G\epsilon P$$

آنگاه تابع f صورتی از امید شرطی است.

اثبات: برای اثبات با توجه به برقراری شرایط لم (۱۵-۱)، کافیست از لم قبل از قضیه برای توابع f و $E[X\|\mathcal{G}]$ استفاده کنیم.

تمامی تساوی و نامساویهای قضیه بعد با احتمال ۱ برقرار هستند.

قضیه ۱۴-۲: با فرض اینکه X و Y و X_n انتگرالپذیر باشند خواهیم داشت :

i) $X = a \quad \Rightarrow \quad \int_G X\,dP = \int_G a\,dP$

ii) $E[aX + bY\|\mathcal{G}] = aE[X\|\mathcal{G}] + bE[Y\|\mathcal{G}]$

iii) $X \le Y \quad \Rightarrow \quad E[X\|\mathcal{G}] \le E[Y\|\mathcal{G}]$

iv) $|E[X\|\mathcal{G}]| \le E[\,|X|\,\|\mathcal{G}]$

v) $lim X_n = X \quad , \quad |X_n| \le Y \quad \Rightarrow \quad lim E[X_n\|\mathcal{G}] = E[X\|\mathcal{G}]$

اثبات:

i) این قسمت با توجه به برابری $X = a$ از خواص انتگرال مشخصا برقرار است.

ii) برای آنکه نشان دهیم $aE[X\|\mathcal{G}] + bE[Y\|\mathcal{G}]$ برابر با $E[aX + bY\|\mathcal{G}]$ است، کافیست نشان دهیم که ورژنی از امید شرطی است که برای اینکار باید برقراری ۲ شرط را بررسی کنیم. با توجه به انتگرالپذیری و اندازه‌پذیری $E[X\|\mathcal{G}]$ و $E[Y\|\mathcal{G}]$ مشخصا شرط اول برقرار است و کافیست نشان دهیم:

$$\int_G aE[X\|\mathcal{G}] + bE[Y\|\mathcal{G}]\,dP = \int_G [aX + bY]\,dP$$

$$\int_G aE[X\|\mathcal{G}] + bE[Y\|\mathcal{G}]\,dP = a\int_G E[X\|\mathcal{G}]\,dP + b\int_G E[Y\|\mathcal{G}]\,dP =$$
$$= a\int_G X\,dP + b\int_G Y\,dP = \int_G aX\,dP + \int_G bY\,dP =$$
$$= \int_G [aX + bY]\,dP$$

iii) چون $X \le Y$ پس $0 \le Y - X$ و می‌توان نوشت:

$$\int_G E[Y\|\mathcal{G}] - E[X\|\mathcal{G}]\,dP = \int_G E[Y\|\mathcal{G}]\,dP - \int_G E[X\|\mathcal{G}]\,dP =$$
$$= \int_G Y\,dP - \int_G X\,dP \ge 0 \qquad \Rightarrow$$
$$\Rightarrow \quad E[Y\|\mathcal{G}] - E[X\|\mathcal{G}] \ge 0 \quad \Rightarrow \quad E[Y\|\mathcal{G}] \ge E[X\|\mathcal{G}]$$

از قسمت iii می‌توان نتیجه گرفت که اگر $X = Y$ با احتمال ۱ برقرار باشد، آنگاه داریم

$$E[Y\|\mathcal{g}] = E[X\|\mathcal{g}]$$

iv) چون همواره داریم $-|X| \leq X \leq |X|$ می‌توان نوشت:

$$-E[|X|\|\mathcal{g}] \leq E[X\|\mathcal{g}] \leq E[|X|\|\mathcal{g}]$$

با در نظر گرفتن $E[|X|\|\mathcal{g}]$=b با کاربرد تعریف قدرمطلق می توان نوشت :

$$-b \leq X \leq b \quad \Leftrightarrow \quad |X| \leq b$$

$$-E[|X|\|\mathcal{g}] \leq E[X\|\mathcal{g}] \leq E[|X|\|\mathcal{g}] \quad \Leftrightarrow \quad |E[X\|\mathcal{g}]| \leq E[|X|\|\mathcal{g}]$$

v) با در نظر گرفتن $\mathrm{Z_n} = \underset{\mathrm{k \geq n}}{Sup}\,|\mathrm{X}_k - \mathrm{X}|$ با احتمال ۱ داریم $Z_n \downarrow 0$ و می‌توان نوشت:

$$|E[X_n\|\mathcal{g}] - E[X\|\mathcal{g}]| = |E[X_n - X\|\mathcal{g}]| \leq E[|X_n - X|\|\mathcal{g}] \leq E[Z_n\|\mathcal{g}]$$

پس کافیست نشان دهیم $E[Z_n\|\mathcal{g}] \downarrow 0$ ، چون دنباله Z_n ها ناصعودی است. از طرفی $E[Z_n\|\mathcal{g}]$ کراندار است و Z_n ها مثبت و نزولی اند و چون هر دنباله کراندار و یکنوا حد دارد پس $E[Z_n\|\mathcal{g}]$ دارای حد نامنفی است که آن حد را Z نامگذاری می‌کنیم و کافیست نشان دهیم که با احتمال ۱ داریم : $Z = 0$

چون Z بزرگترین در دنباله Z هاست پس می‌توان گفت :

$$\mathrm{Z} = \lim E[Z_n\|\mathcal{g}] \leq E[Z_n\|\mathcal{g}] \leq Z_n \quad \Rightarrow \quad \int_{\mathrm{G}} Z\,dP \leq \int_{\mathrm{G}} Z_n\,dP \quad \Rightarrow$$

$$\Rightarrow \quad E(Z) \leq E(Z_n)$$

با توجه به برقراری شرایط قضیه همگرایی مغلوب و کاربرد آن داریم: $E(Z_n) \to 0$ و از آنجا که $E(Z) \leq 0$ است پس با توجه به نامنفی بودن Z با احتمال، ۱، $\mathrm{Z} = 0$ برقرار است و از آنجا خواهیم داشت:

$$|E[X_n\|\mathcal{g}] - E[X\|\mathcal{g}]| \leq 0 \quad \Rightarrow \quad limE[X_n\|\mathcal{g}] = E[X\|\mathcal{g}]$$

لم ۱۴-۲: برای هر تابع نامنفی X که نسبت به $\mathcal{g}$ اندازه‌پذیر باشد، دنباله ای از تابع ساده نامنفی $\{\mathrm{X}_n\}$ که نسبت به $\mathcal{g}$ اندازه‌پذیر باشند، وجود دارد که $|X_n| \leq |X|$ و $\lim X_n = X$.

قضیه ۱۴-۳: اگر X نسبت به $\mathcal{g}$ اندازه‌پذیر و Y و XY انتگرالپذیر باشند آنگاه با احتمال ۱ خواهیم داشت:

$$E[XY\|\mathcal{g}] = XE[Y\|\mathcal{g}] \qquad (۴-۱۴)$$

اثبات: ابتدا برای حالت $X = I_{G_0}$ نشان می‌دهیم که $XE[Y\|\mathcal{g}]$ صورتی از امید شرطی است.

مشخصا $I_{G_0}E[Y\|\mathcal{g}]$ انتگرالپذیر و نسبت به $\mathcal{g}$ اندازه‌پذیر است پس شرط i برقرار است پس برای آنکه صورتی از امید شرطی باشد، کافیست برقراری شرط ii بررسی شود یعنی باید نشان دهیم:

$$\int_{G} I_{G_0} E[Y\|\mathcal{g}]\, dP = \int_{G} I_{G_0} Y\, dP$$

$$\int_{G} I_{G_0} E[Y\|\mathcal{g}]\, dP = \int_{G\cap G_0} E[Y\|\mathcal{g}]\, dP = \int_{G\cap G_0} Y\, dP = \int_{G} I_{G_0} Y\, dP$$

پس رابطه $E[XY\|\mathcal{g}] = XE[Y\|\mathcal{g}]$ اگر X تابع نشانگر باشد برقرار است. با کاربرد بخش ii قضیه (۱۴-۲)، نشان می‌دهیم که اگر X تابع ساده $\sum_i a_i I_{G_i}$ که نسبت به $\mathcal{g}$ اندازه‌پذیر باشد نیز قضیه (۱۴-۳)، برقرار است.

اگر $X = \sum_i a_i I_{G_i}$ آنگاه داریم:

$$E[XY\|\mathcal{g}] = E\left[\sum_i a_i I_{G_i} Y\|\mathcal{g}\right] = \sum_i a_i\, E\left[I_{G_i} Y\|\mathcal{g}\right] =$$

$$= \sum_i a_i I_{G_i}\, E[Y\|\mathcal{g}] = XE[Y\|\mathcal{g}]$$

با استفاده از بخش(۷) قضیه (۱۴-۲) و کاربرد لم (۱۴-۲) و همچنین انتگرالپذیری $|XY|$ می‌توان نوشت:

$$|X_n Y| \le |XY| \;\;\Rightarrow\;\; E[XY\|\mathcal{g}] = \lim E[X_n Y\|\mathcal{g}] \qquad (۱۴\text{-}۵)$$

چون بر طبق لم (۱۴-۲)، X_n تابع ساده است پس می‌توان نوشت:

$$E[X_n Y\|\mathcal{g}] = X_n E[Y\|\mathcal{g}] \qquad (۱۴\text{-}۶)$$

همچنین مجددا با کاربرد لم (۱۴-۲)، خواهیم داشت:

$$\lim X_n E[Y\|\mathcal{g}] = XE[Y\|\mathcal{g}] \qquad (۱۴\text{-}۷)$$

حال با توجه به روابط (۱۴-۵)، (۱۴-۶)، (۱۴-۷)، خواهیم داشت:

$$E[XY\|\mathcal{g}] = \lim E[X_n Y\|\mathcal{g}] = \lim X_n E[Y\|\mathcal{g}] = XE[Y\|\mathcal{g}]$$

که اثبات را کامل می‌کند.

تذکر: امید شرطی گرفتن می‌تواند به عنوان میانگین‌گیری یا هموارسازی در نظر گرفته شود و این حقیقت ذهن ما را به این سمت سوق می‌دهد که میانگین گرفتن از X نسبت به $\mathcal{g}_2$ و سپس میانگین گرفتن از نتیجه حاصل نسبت به یک سیگما میدان کوچکتر مانند $\mathcal{g}_1$ ممکن است نتیجه یکسانی را نسبت به حالتی که از همان ابتدا نسبت به سیگما میدان $\mathcal{g}_1$ میانگین‌گیری شود به ما بدهد، جهت بررسی این موضوع قضیه (۱۴-۴)، در زیر ارائه می‌شود.

قضیه ۱۴-۴ : اگر X انتگرالپذیر باشد و $\mathcal{g}_1$ و $\mathcal{g}_2$ دو سیگما میدان باشند که $\mathcal{g}_1 \subset \mathcal{g}_2$ آنگاه با احتمال ۱ خواهیم داشت:

$$E[E[X\|\mathcal{g}_2]\|\mathcal{g}_1] = E[X\|\mathcal{g}_1] \qquad (۱۴\text{-}۸)$$

اثبات: چون $E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1]$ نسبت به $\mathcal{G}_1$ اندازه‌پذیر است، برای آنکه نشان دهیم صورتی از امید شرطی نسبت به $\mathcal{G}_1$ است، کافیست نشان دهیم:

$$\int_{\mathrm{G}} E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1]\, dP = \int_{\mathrm{G}} X\, dP \qquad \forall G\epsilon\mathcal{G}_1$$

چون $E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1]$ نسبت به $\mathcal{G}_1$ اندازه‌پذیر است پس با توجه به برقراری شرط ii در ابتدای فصل، برای آن می‌توان نوشت:

$$\int_{\mathrm{G}} E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1]\, dP = \int_{\mathrm{G}} E[X\|\mathcal{G}_2]\, dP \qquad \forall G\epsilon\mathcal{G}_1 \qquad (۹\text{-}۱۴)$$

چون $E[X\|\mathcal{G}_2]$ نسبت به $\mathcal{G}_2$ اندازه‌پذیر است پس با توجه به برقراری شرط ii برای آن می‌توان نوشت:

$$\int_{\mathrm{G}} E[X\|\mathcal{G}_2]\, dP = \int_{\mathrm{G}} X\, dP \qquad \forall G\epsilon\mathcal{G}_2 \qquad (۱۰\text{-}۱۴)$$

که با توجه به اینکه رابطه (14-10)، برای هر $G\epsilon\mathcal{G}_2$ برقرار است، و $\mathcal{G}_1 \subset \mathcal{G}_2$ است، پس رابطه

(14-10)، برای هر $G\epsilon\mathcal{G}_1$ نیز برقرار خواهد بود که از ترکیب (14-9) و (14-10) خواهیم داشت:

$$\int_{\mathrm{G}} E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1]\, dP = \int_{\mathrm{G}} E[X\|\mathcal{G}_2]\, dP = \int_{\mathrm{G}} X\, dP \qquad \forall G\epsilon\mathcal{G}_1$$

نکته ۱۴–۱ : در حالت خاص، اگر $\mathcal{G}_2$ بسیار بزرگ باشد یعنی $\mathcal{G}_2 = \mathcal{F}$ باشد، با احتمال ۱ خواهیم داشت

$$E[X\|\mathcal{G}_2] = E[X\|\mathcal{F}] = X$$

$$\Rightarrow \qquad E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1] = E[E[X\|\mathcal{F}]\|\mathcal{G}_1] = E[X\|\mathcal{G}_1] \qquad (۱۱\text{-}۱۴)$$

به این معناست که در حالت خاص $\mathcal{G}_2 = \mathcal{F}$ نیز، قضیه (۱۴-۴)، برقرار است.

نکته ۱۴–۲ : در حالت خاص، اگر $\mathcal{G}_1 = \{\emptyset, \Omega\}$ و $\mathcal{G}_2 = \mathcal{G}$ باشد، قضیه (۱۴-۴)، حالت خاصی از صورت امید شرطی برای $G = \Omega$ می‌شود.

$$E[E[X\|\mathcal{G}_2]\|\mathcal{G}_1] = E[E[X\|\mathcal{G}]\|\{\emptyset, \Omega\}] = E[E[X\|\mathcal{G}]] = E[X]$$
$$= E[X\|\{\emptyset, \Omega\}] = E[X\|\mathcal{G}_1]$$

که سومین تساوی به دلیل عبارت $E[E[X\|\mathcal{G}]] = \int_{\Omega} E[X\|\mathcal{G}]\, dP = \int_{\Omega} X\, dP = E[X]$ برقرار است.

نکته ۱۴–۳: اگر $\mathcal{G}_1 \subset \mathcal{G}_2$ باشد، آنگاه $E[X\|\mathcal{G}_1]$ که نسبت به $\mathcal{G}_1$ اندازه‌پذیر است، نسبت به $\mathcal{G}_2$ نیز اندازه‌پذیر می باشد، بنابراین گرفتن امید شرطی نسبت به $\mathcal{G}_2$ تغییری در $E[X\|\mathcal{G}_1]$ ایجاد نمی‌کند. یعنی داریم

$$E[E[X\|\mathcal{G}_1]\|\mathcal{G}_2] = E[X\|\mathcal{G}_1] \qquad (۱۳\text{-}۱۴)$$

نتیجه: بنابراین با دقت در نکته (۱۴-۳) و قضیه (۱۴-۴)، به آسانی مشخص می‌شودکه زمانیکه $\mathcal{G}_1 \subset \mathcal{G}_2$ است چه امید شرطی را ابتدا نسبت به $\mathcal{G}_1$ و سپس نسبت به $\mathcal{G}_2$ بگیریم و چه امید شرطی را ابتدا نسبت به $\mathcal{G}_2$ و سپس نسبت به $\mathcal{G}_1$ بگیریم پاسخ یکسان و برابر با $E[X\|\mathcal{G}_1]$ خواهد بود.

۱۴–۳ نامساوی جنسن برای مقادیر امید شرطی

اگر φ یک تابع محدب و X و $\varphi(\mathrm{X})$ هر دو انتگرالپذیر باشند، آنگاه با احتمال ۱خواهیم داشت:

$$\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}]) \leq \mathrm{E}[\varphi(\mathrm{X})\|\mathcal{G}] \qquad (۱۴\text{-}۱۲)$$

اثبات: با در نظر گرفتن یک خط پوشش $(x_0, \varphi(x_0))$ برای هر x_0 به صورت زیر داریم

$$\varphi(x_0) + \mathrm{A}(x_0)(\mathrm{x} - x_0) \leq \varphi(\mathrm{x})$$

شیب $\mathrm{A}(x_0)$ می‌تواند به عنوان مشتق راست $\varphi(\mathrm{X})$ در نظر گرفته شود که در x_0 غیر صعودی است.

$$\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}]) + \mathrm{A}(\mathrm{E}[\mathrm{X}\|\mathcal{G}])(\mathrm{X} - \mathrm{E}[\mathrm{X}\|\mathcal{G}]) \leq \varphi(\mathrm{X})$$

نخست با فرض اینکه $\mathrm{E}[\mathrm{X}\|\mathcal{G}]$ کراندار باشد، هر ۳ جمله سمت چپ نامساوی فوق انتگرال پذیرند و ضمنا اگر φ یک تابع محدب در R^1 باشد، آنگاه φ و A در مجموعه‌های کراندار،کراندار هستند. حال با گرفتن امید شرطی نسبت به $\mathcal{G}$ از طرفین خواهیم داشت:

$$\mathrm{E}[\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}])\|\mathcal{G}] + \mathrm{E}[\mathrm{A}(\mathrm{E}[\mathrm{X}\|\mathcal{G}])(\mathrm{X} - \mathrm{E}[\mathrm{X}\|\mathcal{G}])\|\mathcal{G}] \leq \mathrm{E}[\varphi(\mathrm{X})\|\mathcal{G}]$$

چون $\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}])$ و $\mathrm{A}(\mathrm{E}[\mathrm{X}\|\mathcal{G}])$ نسبت به $\mathcal{G}$ اندازه‌پذیر هستند بنابراین گرفتن امید شرطی نسبت به $\mathcal{G}$ تاثیری بر آنها ندارد، و با خارج کردن آنها از امید شرطی داریم

$$\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}]) + \mathrm{A}(\mathrm{E}[\mathrm{X}\|\mathcal{G}])\left(\int_{\mathrm{G}} X\, dP \ - \int_{\mathrm{G}} \mathrm{E}[\mathrm{X}\|\mathcal{G}]\, dP\ \right)] \leq \mathrm{E}[\varphi(\mathrm{X})\|\mathcal{G}]$$

که با توجه به عبارت $\int_{\mathrm{G}} X\, dP = \int_{\mathrm{G}} \mathrm{E}[\mathrm{X}\|\mathcal{G}]\, dP$، جمله میانی فوق برابر با صفر خواهد بود و در نتیجه داریم

$$\varphi(\mathrm{E}[\mathrm{X}\|\mathcal{G}]) \leq \mathrm{E}[\varphi(\mathrm{X})\|\mathcal{G}]$$

چون در حالت قبل فرض کردیم $\mathrm{E}[\mathrm{X}\|\mathcal{G}]$ کراندار باشد، برای حالت کلی با در نظر گرفتن

$G_n = [|\mathrm{E}[\mathrm{X}\|\mathcal{G}]| \leq n]$ داریم

$$E[I_{G_n}X\|\mathcal{g}] = I_{G_n}E[X\|\mathcal{g}]$$

$$I_{G_n}E[X\|\mathcal{g}] = \begin{cases} E[X\|\mathcal{g}] & \omega \in G_n \\ 0 & \omega \notin G_n \end{cases} \quad \leftrightarrow \quad E[X\|\mathcal{g}] \le n \quad \rightarrow$$ کراندار است

که نشان می‌دهد رابطه جنسن برای $I_{G_n}X$ برقرار است و می‌توان نوشت:

$$\varphi(E[I_{G_n}X\|\mathcal{g}]) \le E[\varphi(I_{G_n}X)\|\mathcal{g}] \qquad \Rightarrow \qquad \varphi(I_{G_n}E[X\|\mathcal{g}]) \le E[\varphi(I_{G_n}X)\|\mathcal{g}] \quad (۱)$$

$$\lim_{n\to\infty} I_{G_n}E[X\|\mathcal{g}] \;\rightarrow\; E[X\|\mathcal{g}] \qquad \Rightarrow$$

$$\lim_{n\to\infty} \varphi(I_{G_n}E[X\|\mathcal{g}]) \;\rightarrow\; \varphi(E[X\|\mathcal{g}]) \qquad (*)$$

$$E[\varphi(I_{G_n}X)\|\mathcal{g}] = E[I_{G_n}\varphi(X) + I_{G_n^c}\varphi(0)\|\mathcal{g}] = I_{G_n}E[\varphi(X)\|\mathcal{g}] + I_{G_n^c}\varphi(0) \quad (**)$$

$$\Rightarrow \quad \lim_{n\to\infty} E[\varphi(I_{G_n}X)\|\mathcal{g}] \quad \rightarrow \quad E[\varphi(X)\|\mathcal{g}] \qquad (***)$$

که برابری (***) به این دلیل است که $\varphi(0)$ مقدار ثابتی می‌باشد و با حدگیری از طرفین (**) بدست آمده است.

چون (*) حد جمله سمت چپ و (***) حد سمت راست نامساوی (۱) می‌باشند، پس از (*) و (***) خواهیم داشت

$$\varphi(E[X\|\mathcal{g}]) \le E[\varphi(X)\|\mathcal{g}]$$

نکته ۱۴-۴ : چون $|X|$ تابعی محدب است، اگر $\varphi(X) = |X|$ باشد، آنگاه نامساوی جنسن بخش(iv) قضیه (۱۴-۲)، را نتیجه می‌دهد، یعنی:

$$|E[X\|\mathcal{g}]| \le E[\,|X|\,\|\mathcal{g}]$$

۱۴-۴ توزیع های شرطی و امیدهای شرطی

قضیه ۱۴-۵ : فرض کنید $\mu(.,\omega)$ توزیع شرطی متغیر تصادفی X نسبت به $\mathcal{g}$ باشد، اگر $\varphi: R \to R$ تابع بورل باشد که برای آن $\varphi(X)$ انتگرالپذیر باشد، آنگاه $\int_{R^1} \varphi(X)\,\mu(dx,\omega)$صورتی از امید شرطی $E[\varphi(X)\|\mathcal{g}]_\omega$ است.

اثبات: در مرحله اول فرض می‌کنیم $\varphi = I_H\,,\, H\epsilon R^1$ باشد، با احتمال ۱ خواهیم داشت:

$$\int_R I_H\,\mu(dx,\omega) = \int_{R\cap H}\mu(dx,\omega) = \int_H \mu(dx,\omega)$$

$$= P(X\epsilon H\|\mathcal{g}) = P(H\|\mathcal{g}) = E[I_H\|\mathcal{g}]$$

در مرحله دوم فرض می‌کنیم تابع ساده به صورت $\varphi(x) = \sum_i a_i I_{A_i}(x)$ باشد، با احتمال ۱ خواهیم داشت:

$$\int_{\mathrm{R}} \sum_i a_i I_{A_i}(x)\, \mu(dx,\omega) = \sum_i a_i \int_{\mathrm{R}} I_{A_i}(x) \mu(dx,\omega) = \sum_i a_i \mathrm{E}\left[I_{A_i}(x) \| \mathcal{g}\right]$$

$$= \mathrm{E}\left[\sum_i a_i\, I_{A_i}(x) \| \mathcal{g}\right]$$

در مرحله سوم فرض می‌کنیم $\varphi \geq 0$ باشد، با کاربرد لم (۱۴-۲)، تابع ساده $\varphi_{\mathrm{n}}(\mathrm{x})$ نامنفی وجود دارد که به $\varphi(\mathrm{x})$ همگراست و داریم

$$\int_{\mathrm{R}} \varphi_{\mathrm{n}}(\mathrm{x})\, \mu(dx,\omega) = E[\varphi_{\mathrm{n}}(\mathrm{X}) \| \mathcal{g}]$$

با توجه به برابری رابطه فوق، حد طرفین نیز با هم برابر است، حال از طرفین حد می‌گیریم.

برای حد طرف چپ داریم

$$\varphi_{\mathrm{n}}(\mathrm{x}) \to \varphi(\mathrm{x}) \qquad \Rightarrow \qquad \int_{\mathrm{R}} \varphi_{\mathrm{n}}(\mathrm{x})\, \mu(dx,\omega) \to \int_{\mathrm{R}} \varphi(\mathrm{x})\, \mu(dx,\omega) \qquad (۱۴\text{-}۱۴)$$

با توجه به انتگرالپذیری $\varphi(\mathrm{x})$ و رابطه $\varphi_{\mathrm{n}}(\mathrm{x}) \leq \varphi(\mathrm{x})$ با کاربرد قسمت (۷) قضیه (۱۴-۲)، همچنین برای حد طرف راست نیز خواهیم داشت:

$$\int_{\mathrm{R}} \varphi(\mathrm{x})\, \mu(dx,\omega) = E[\varphi(\mathrm{x}) \| \mathcal{g}] \qquad (۱۵\text{-}۱۴)$$

که از برابری (۱۴-۱۴) و (۱۴-۱۵) اثبات برای حالت $\varphi \geq 0$ کامل است.

در مرحله چهارم برای حالت کلی φ اثبات می‌کنیم:

$$\varphi(\mathrm{x}) = \varphi^{+}(\mathrm{x}) - \varphi^{-}(\mathrm{x}) \qquad \Rightarrow$$

$$\int_{\mathrm{R}} \varphi(\mathrm{x})\, \mu(dx,\omega) = \int_{\mathrm{R}} \varphi^{+}(\mathrm{x}) - \varphi^{-}(\mathrm{x})\, \mu(dx,\omega)$$

$$= \int_{\mathrm{R}} \varphi^{+}(\mathrm{x})\, \mu(dx,\omega) - \int_{\mathrm{R}} \varphi^{-}(\mathrm{x})\, \mu(dx,\omega)$$

$$= E[\varphi^{+}(\mathrm{x}) \| \mathcal{g}] - E[\varphi^{-}(\mathrm{x}) \| \mathcal{g}] = E[\varphi^{+}(\mathrm{x}) - \varphi^{-}(\mathrm{x}) \| \mathcal{g}]$$

$$= E[\varphi(\mathrm{x}) \| \mathcal{g}]$$

که اثبات برای حالت کلی φ کامل است.

از اثبات بالا این نتیجه بدست می‌آید که $\int_{\mathrm{R}} \varphi(\mathrm{x})\, \mu(dx,\omega)$ نسبت به $\mathcal{g}$ اندازه‌پذیر است و با احتمال ۱ متناهی است. (چون با احتمال ۱ صورتی از $E[\varphi(\mathrm{x}) \| \mathcal{g}]$ است و $\varphi(\mathrm{x})$ انتگرال پذیر است پس امید شرطی $\varphi(\mathrm{x})$ متناهی است).

اگر X انتگرال پذیر باشد، آنگاه با استفاده از قضیه بالا برای $\varphi(\mathrm{x}) = X$ با احتمال ۱ می‌توان نوشت:

$$E[\mathrm{x} \| \mathcal{g}]_{\omega} = \int_{-\infty}^{+\infty} \mathrm{x}\, \mu(dx,\omega)$$

و برای حالت $X = \varphi(\mathrm{x})$ با احتمال ۱ می‌توان نوشت:

$$E[\varphi(\mathrm{x})\|\mathcal{g}]_\omega = \int_{-\infty}^{+\infty} \varphi(\mathrm{x})\mu(dx,\omega) \qquad (۱۶\text{-}۱۴)$$

با توجه به نابرابری جنسن برای حالت امید غیر شرطی، $\int_{-\infty}^{+\infty} \varphi(\mathrm{x})\mu(dx,\omega)$ کمترین مقدار خود را اگر φ محدب باشد در $\varphi[\int_{-\infty}^{+\infty} \mathrm{x}\mu(dx,\omega)]$ خواهد داشت یعنی به عبارت دیگر $E[\varphi(\mathrm{x})\|\mathcal{g}]_\omega$ در $\varphi(E[\mathrm{x}\|\mathcal{g}]_\omega)$ مینیمم می‌شود که اثباتی مجدد بر نامساوی جنسن برای امید شرطی به شمار می‌رود.

۱۴–۵ مسائل

۱- نشان دهید اگر X و Y مستقل باشند آنگاه برابری $E[Y\|X] = E[Y]$ برقرار است و یا به عبارت دیگر اگر X و Y مستقل باشند آنگاه برابری $\mathrm{E[XY] = E[X]E[Y]}$ برقرار است. با ذکر مثال در G که شامل ۳ نقطه باشد نشان دهید عکس روابط بالا برقرار نیست.

۲- در رابطه (۱۴-۸) قضیه ۱۴-۴ نشان دهید امید شرطی سمت راست رابطه، صورتی از امید شرطی سمت چپ است.

۳- برای X و Y کراندار نشان دهید : $E\big[Y\,E[X\|\mathcal{g}]\big] = E\big[X\,E[Y\|\mathcal{g}]\big]$

۴- نشان دهید اگر $\mathcal{g}_1 \subset \mathcal{g}_2$ و $E[X^2] < \infty$ باشد :

$$E[(X - E[X\|\mathcal{g}_2])^2] \le E[(X - E[X\|\mathcal{g}_1])^2]$$

(با بزرگتر شدن سیگما میدان ، پراکندگی X در اطراف میانگین شرطی‌اش کمتر می‌شود)

۵- با تعریف $Var[X\|\mathcal{g}] = E[(X - E[X\|\mathcal{g}])^2\|\mathcal{g}]$ اثبات کنید که رابطه زیر برقرار است:

$$Var[X] = E\big[Var[X\|\mathcal{g}]\big] + Var\big[E[X\|\mathcal{g}]\big]$$

فصل پانزدهم

مارتینگل

۱۵–۱ مارتینگل ها

تعریف ۱۵-۱ : فرض کنید $X_1, X_2, \cdots$ دنباله ای از متغیرهای تصادفی روی فضای احتمال $(\Omega, \mathcal{F}, P)$ باشند و فرض کنید $\mathcal{F}_1, \mathcal{F}_2, \cdots$ دنباله ای از زیر-سیگما میدان‌های $\mathcal{F}$ باشد. دنباله $\{(X_n, \mathcal{F}_n): n = 1,2, \cdots\}$ یک مارتینگل است اگر چهار شرط زیر برقرار باشند:

i) $\mathcal{F}_\mathrm{n} \subset \mathcal{F}_{n+1}$ ،

ii) X_n نسبت به $\mathcal{F}_n$ اندازه پذیر باشد،

iii) X_n انتگرال پذیر باشد؛ یعنی: $E[|X_n|] < \infty$ ،

iv) با احتمال ۱ داشته باشیم: $E[X_{n+1}||\mathcal{F}_n] = X_n$.

شروط بالا بیان می کنند که محتوای اطلاعاتی ما از فرایند، لحظه به لحظه و با وقوع هر یک از مقادیر X_{n+1} نسبت به مرحله n ام فرایند افزایش می یابد، X_n هماهنگ با $\mathcal{F}_n$ است و همچنین با داشتن تمامی اطلاعات تا مرحله n ام فرایند، تنها پیش بینی موجود از آینده فرایند در مرحله $n+1$ ام، همان X_n خواهد بود. به عبارت دیگر در هر مرحله از فرایند، مقدار مورد انتظار از مرحله بعدی، همان مقدار فرایند در مرحله کنونی است و مقدار متغیر در این مرحله از فرایند، حاوی بیشترین میزان اطلاعات از فضای اطلاعاتی فرایند است. به عنوان تعریفی دیگر، دنباله $X_1, X_2, \cdots$ یک مارتینگل وابسته به سیگما میدان‌های $\mathcal{F}_1, \mathcal{F}_2, \cdots$ خواهد بود.

نکته: اگر X_n یک مارتینگل نسبت به $\mathcal{F}_n$ باشد، آنگاه X_n نسبت به $\sigma\{X_1, X_2, \cdots, X_n\}$ نیز یک مارتینگل است.

اثبات: چهار شرط لازمه برای مارتینگل بودن را بررسی می کنیم. کوچکترین σ- میدان

$\mathcal{g}_n = \sigma\{X_1, X_2, \cdots, X_n\}$ ساخته شده از متغیرهای تصادفی نیز یک فضای پیشامد برای تعریف مارتینگل ها ایجاد می کند. واضح است که $\mathcal{g}_n \subseteq \mathcal{g}_{n+1}$ (شرط i برقرار است). بدیهی است $X_n \in \sigma\{X_1, X_2, \cdots, X_n\}$ ، یعنی X_n نسبت به $\mathcal{g}_n$ اندازه پذیر است (شرط ii برقرار است). شرط iii که در فرضیات نکته جای دارد. حال اگر

ویژگی (iv) مارتینگل ها برقرار باشد، آنگاه $X_1, X_2, \cdots, X_n$ نسبت به $\mathcal{g}_n$ ها مارتینگل هستند که در زیر به اثبات آن می پردازیم:

$$E[X_{n+1}||\mathcal{g}_n] = E[E[X_{n+1}||\mathcal{F}_n]||\mathcal{g}_n] = E[X_n||\mathcal{g}_n] = X_n$$

برابری آخر نیز با توجه به $E[X_{n+1}||\mathcal{F}_n] = X_n$ و اندازه پذیری X_n نسبت به $\mathcal{g}_n$ حاصل شده است. در نتیجه با تعریف سیگما میدان $\mathcal{g}_n = \sigma\{X_1, X_2, \cdots, X_n\}$ ، تمامی شروط مارتینگل برقرار است.

■

برای سیگما میدان خاص $\mathcal{g}_n$ ویژگی چهارم مارتینگل به صورت زیر خواهد بود:

$$E[X_{n+1}||X_1, X_2, \cdots, X_n] = X_n$$

از آنجا که X_n نسبت به $\mathcal{F}_n$ اندازه پذیر است، اگر و تنها اگر $\sigma\{X_1, X_2, \cdots, X_n\} \subset \mathcal{F}_n$ باشد، بنابراین $\sigma\{X_1, X_2, \cdots, X_n\}$ کوچکترین سیگما میدان است که X_n ها نسبت به آن مارتینگل هستند.

مثال ۱۵-۱ : فرض کنید $\mathrm{Z}_1, Z_2, \cdots$ دنباله ای از متغیرهای تصادفی مستقل و هم توزیع روی فضای احتمال $(\Omega, \mathcal{F}, P)$ با میانگین متناهی $E[\mathrm{Z}_1] = \mu < \infty$ باشند. بگیرید $X_n = \sum_{i=1}^{n} Z_i$ در این صورت برای هر $n \geq 1$، $\sigma\{X_1, X_2, \cdots, X_n\} \subseteq \mathcal{F}_n$. زیرا با فرض اندازه پذیر بورل بودن تابعی همچون f (لزوماً این تابع پیوسته نیست) و با توجه به تعریف ۹-۳ داریم:

$$\sigma\{f(Z)\} = \{\omega \in \Omega \,|f(Z(\omega))\epsilon\beta, \forall\, \beta \in \mathcal{B}(\mathcal{R}^1)\}$$

$$= \{Z^{-1}(f^{-1}(\beta)) \;|\, \beta \in \mathcal{B}(\mathcal{R}^1)\}$$

$$\subseteq \{Z^{-1}(\beta) \;|\, \beta \in \mathcal{B}(\mathcal{R}^1)\} = \sigma\{Z\}$$ چون f اندازه پذیر بورل است

و بدیهی است که $E|X_n| < \infty$. همچنین داریم:

$$E[X_{n+1}||\mathcal{F}_n] = E[\mathrm{Z}_1 + Z_2 + \cdots + \mathrm{Z}_n + Z_{n+1}||\mathcal{F}_n]$$

$$= X_n + E[Z_{n+1}||\mathcal{F}_n]$$

$$= X_n + E[Z_{n+1}]$$

$$= X_n + \mu$$

حال اگر $\mu = 0$ باشد آنگاه X_n یک مارتینگل است. (در حالت کلی تر می توان بیان کرد که اگر $\mu > 0$ باشد آنگاه X_n یک زیرمارتینگل و اگر $\mu < 0$ باشد آنگاه X_n یک زبرمارتینگل خواهد بود).

■

بنابراین طبق مثال فوق می توان نتیجه گرفت که دنباله ای از متغیرهای تصادفی مستقل با میانگین صفر، یک مارتینگل است.

باید توجه داشت که شرط اساسی مارتینگل بودن در ویژگی (iv) تعریف ۱ قرار دارد. البته شرط (iii) نیز مورد نیاز است که وجود $E[X_{n+1}||\mathcal{F}_n]$ را تضمین نماید. با این حال شرط (iv) بیان می کند که X_n یک ورژن از $E[X_{n+1}||\mathcal{F}_n]$ است. از آنجایی که X_n نسبت به $\mathcal{F}_n$ اندازه پذیر است، این شرط معادل با رابطه زیر است.

$$\int_A X_{n+1}\,\mathrm{dP} = \int_A X_n\,\mathrm{dP} \qquad\qquad A \in \mathcal{F}_n$$

اثبات: نشان می دهیم که،

$$E[X_{n+1}||\mathcal{F}_n] = X_n \quad \Leftrightarrow \quad \int_A X_{n+1}\,\mathrm{dP} = \int_A X_n\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_n$$

اثبات شرط کافی: بنابه فرض مسئله، X_n ورژنی از امید شرطی $E[X_{n+1}||\mathcal{F}_n]$ است. در نتیجه بنا به تعریف امید شرطی در فصل پیش،

$$\int_A E[X_{n+1}||\mathcal{F}_n]\,\mathrm{dP} = \int_A X_n\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_n$$

اثبات شرط لازم: اگر رابطه $\int_A X_{n+1}\,\mathrm{dP} = \int_A X_n\,\mathrm{dP}$ برای هر $A \in \mathcal{F}_n$ برقرار باشد، آنگاه با توجه به برقراری رابطه انتگرالی و اندازه پذیری X_n و نسبت به $\mathcal{F}_n$ ، X_n ورژنی از امید شرطی $E[X_{n+1}||\mathcal{F}_n]$ خواهد بود.

همچنین از آنجایی که $\mathcal{F}_n$ ها تودرتو و صعودی هستند در نتیجه برای $A \in \mathcal{F}_n$ و $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ داریم:

$$\int_A X_n\,\mathrm{dP} = \int_A X_{n+1}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_n$$

$$\int_A X_{n+1}\,\mathrm{dP} = \int_A X_{n+2}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_{n+1}$$

$$\vdots$$

$$\int_A X_{n+k-1}\,\mathrm{dP} = \int_A X_{n+k}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_{n+k-1}$$

$$\Rightarrow \quad \int_A X_n\,\mathrm{dP} = \int_A X_{n+1}\,\mathrm{dP} = \int_A X_{n+2}\,\mathrm{dP} = \cdots = \int_A X_{n+k}\,\mathrm{dP}$$

بنابراین X_n ای که نسبت به $\mathcal{F}_n$ مارتینگل است، یک ورژن از امید شرطی $E[X_{n+k}||\mathcal{F}_n]$ می باشد و در نتیجه:

$$\forall k \geq 1: \qquad E[X_{n+k}||\mathcal{F}_n] = X_n \qquad \mathrm{w.p.}\,1$$

باید توجه داشت که اگر $A = \Omega$ باشد آنگاه بر اساس رابطه فوق داریم:

$$\int_\Omega X_n\,\mathrm{dP} = \int_\Omega X_{n+1}\,\mathrm{dP} = \int_\Omega X_{n+2}\,\mathrm{dP} = \cdots = \int_\Omega X_{n+k}\,\mathrm{dP}$$

$$\Rightarrow \quad E[X_1] = E[X_2] = \cdots$$

بنابراین *امید ریاضی در مارتینگل ها ثابت است*. یعنی

$$E[\mathrm{X}_1] = E[\mathrm{X}_2] = \cdots = E[\mathrm{X}_n]$$

شرایط تعریف مارتینگل ها برای تفاضلات $\Delta_n = \mathrm{X}_n - \mathrm{X}_{n-1}$ نیز با تعریف $\Delta_1 = \mathrm{X}_1$ قابل انجام است و با استفاده از خاصیت خطی بودن این تبدیلات داریم:

$$E[\Delta_{n+1}||\mathcal{F}_n] = 0$$

اثبات: با توجه به اندازه پذیر بودن X_n نسبت به $\mathcal{F}_n$ و همچنین مارتینگل بودن X_n ها نسبت به $\mathcal{F}_n$، برابری زیر برقرار است.

$$E[\Delta_{n+1}||\mathcal{F}_n] = E[\mathrm{X}_{n+1} - \mathrm{X}_n||\mathcal{F}_n] = E[\mathrm{X}_{n+1}||\mathcal{F}_n] - E[\mathrm{X}_n||\mathcal{F}_n] = \mathrm{X}_n - \mathrm{X}_n = 0$$

باید توجه داشت که از انجایی که $X_k = \Delta_1 + \Delta_2 + \cdots + \Delta_k$ و $\Delta_k = \mathrm{X}_k - \mathrm{X}_{k-1}$، مجموعه های $X_1, X_2, \cdots, X_n$ و $\Delta_1, \Delta_2, \cdots, \Delta_n$ میدان های یکسانی تولید می کنند. یعنی:

$$\sigma\{X_1, X_2, \cdots, X_n\} = \sigma\{\Delta_1, \Delta_2, \cdots, \Delta_n\}$$

مثال ۱۵-۲: فرض کنید $\Delta_1, \Delta_2, \cdots$ متغیرهای تصادفی مستقل و انتگرال پذیر باشند به طوری که برای $n \geq 2$ داشته باشیم، $E[\Delta_n] = 0$. اگر $\mathcal{F}_n = \sigma\{\Delta_1, \Delta_2, \cdots, \Delta_n\}$ باشد، در این صورت با توجه به استقلال این متغیرها داریم:

$$E[\Delta_{n+1}||\mathcal{F}_n] = E[\Delta_n] = 0$$

و اگر Δ یک متغیر تصادفی دیگر باشد که از Δ_n ها مستقل است، و اگر $\mathcal{F}_n$ با $\sigma\{\Delta, \Delta_1, \Delta_2, \cdots, \Delta_n\}$ جایگزین شود، در این صورت $X_n = \Delta_1 + \Delta_2 + \cdots + \Delta_n$ ها هنوز نسبت به $\mathcal{F}_n$ جدید مارتینگل هستند و این که اجازه دهیم $\mathcal{F}_n$ یک σ- میدان بزرگتر از $\sigma\{X_1, X_2, \cdots, X_n\}$ باشد طبیعی و قانع کننده است. به عبارت دیگر، همیشه این طور نیست که $\sigma\{X_1, X_2, \cdots, X_n\}$ در شرایط مارتینگل بودن صدق کند بلکه سیگما میدان های بزرگتر نیز می توانند در این شرایط صادق باشند.

مثال ۱۵-۳: فرض کنید $(\Omega, \mathcal{F}, P)$ یک فضای احتمال و ν یک اندازه متناهی روی $\mathcal{F}$ باشد و $\mathcal{F}_1, \mathcal{F}_2, \cdots$ یک دنباله غیر نزولی از سیگما میدان ها روی فضای $\mathcal{F}$ باشد. اگر برای هر $A \in \mathcal{F}_n$، اندازه P یک اندازه غالب ν باشد (یعنی: برای هر $A \in \mathcal{F}_n$، اگر $\mathrm{P}(A) = 0$، آنگاه $\nu(A) = 0$)، در این صورت چگالی یا مشتق رادون نیکودیم

مثل X_n برای v نسبت به P روی $\mathcal{F}_n$ وجود دارد به طوری که، نسبت به $\mathcal{F}_n$ اندازه پذیر و نسبت به P انتگرال پذیر می باشد و در رابطه زیر صدق می کند.

$$\int_A X_n \, \mathrm{dP} = \mathrm{v}(A) \quad , \quad A \in \mathcal{F}_n$$

حال اگر $A \in \mathcal{F}_n$ باشد و $\mathcal{F}_1, \mathcal{F}_2, \cdots$ یک دنباله غیر نزولی از سیگما میدان ها روی فضای $\mathcal{F}$ باشند، آنگاه $A \in \mathcal{F}_{n+1}$ و بنابراین داریم:

$$\int_A X_{n+1} \, \mathrm{dP} = \mathrm{v}(A) \quad , \quad A \in \mathcal{F}_{n+1}$$

و با توجه به یکتایی مشتق رادون نیکودیم داریم:

$$\int_A X_{n+1} \, \mathrm{dP} = \int_A X_n \, \mathrm{dP} \quad , \quad A \in \mathcal{F}_n$$

بنابراین X_n ها نسبت به $\mathcal{F}_n$ مارتینگل هستند.

مثال ۱۵-۴: فرض کنید اندازه v در مثال ۱۲-۳، یک اندازه احتمال مانند Q روی $\mathcal{F}$ باشد و $\mathrm{Y}_1, \mathrm{Y}_2, \cdots$ متغیرهای تصادفی روی فضای اندازه پذیر $(\Omega, \mathcal{F})$ و $\mathcal{F}_n = \sigma\{\mathrm{Y}_1, \mathrm{Y}_2, \cdots, \mathrm{Y}_n\}$ باشند. فرض کنید که تحت اندازه P توزیع بردار تصادفی $(\mathrm{Y}_1, \mathrm{Y}_2, \cdots, \mathrm{Y}_n)$ دارای چگالی $p_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)$ نسبت به اندازه لبگ n بعدی باشد و بردار تصادفی $(\mathrm{Y}_1, \mathrm{Y}_2, \cdots, \mathrm{Y}_n)$ دارای چگالی $q_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)$ تحت Q باشد. اگر p_n همه جا مثبت(مخالف صفر) باشد، آنگاه مشتق رادون نیکودیم Q نسبت به P روی $\mathcal{F}_n$ به صورت زیر تعریف می شود.

$$Q(A) = \int_A X_n \, \mathrm{dP} \quad ; \quad X_n = \frac{q_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)}{p_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)}$$

به طوری که در این رابطه، $A \equiv [\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n] \in H$ و H استوانه های n بعدی در فضای اعداد حقیقی را نشان می دهد.

اثبات: بنا به مثال قبل داریم، $\int_A X_n \, \mathrm{dP} = Q(A)$ و از طرفی بنا به تعریف تابع چگالی در قضیه رادون نیکودیم داریم:

$$\int_{[\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n] \in H} p_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n) \, \mathrm{d}\mathbf{y} = P(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)$$

$$\int_{[\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n] \in H} q_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n) \, \mathrm{d}\mathbf{y} = Q(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n)$$

حال با استفاده از رابطه $\int_A X_n \, \mathrm{dP} = \int_{[\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n] \in H} X_n p_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n) \, \mathrm{d}\mathbf{y}$ داریم:

$$\int_A X_n \, \mathrm{dP} = Q(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n) = \int_{[\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n] \in H} q_n(\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_n) \, \mathrm{d}\mathbf{y}$$

$$\Longrightarrow \int_{[y_1,y_2,\cdots,y_n]\in H} X_n p_n(y_1,y_2,\cdots,y_n)\,d\mathbf{y} = \int_{[y_1,y_2,\cdots,y_n]\in H} q_n(y_1,y_2,\cdots,y_n)\,d\mathbf{y}$$

و با توجه به برابری انتگرالی این دو عبارت داریم:

$$X_n p_n(y_1,y_2,\cdots,y_n) = q_n(y_1,y_2,\cdots,y_n) \qquad w.p.1$$

$$\Longrightarrow\ X_n = \frac{q_n(y_1,y_2,\cdots,y_n)}{p_n(y_1,y_2,\cdots,y_n)} \qquad w.p.1$$

در این حالت، X_n همان نسبت درستنمایی است که توان q_n را نسبت به p_n می سنجد. حال اگر $(Y_1,Y_2,\cdots,Y_n)$ متغیرهای تصادفی مستقل باشند، آنگاه:

$$p(y_1,y_2,\cdots,y_n) = p(y_1)p(y_2)\cdots p(y_n) \qquad \text{و}$$

$$q(y_1,y_2,\cdots,y_n) = q(y_1)q(y_2)\cdots q(y_n)$$

و اگر $P(y_n \in H) \neq Q(y_n \in H)$ باشد، آنگاه با تعریف $Z_n = I_{[y_n\in H]}$ و با استناد به قانون قوی اعداد بزرگ داریم

$$\bar{Z}_n \xrightarrow{P} E[Z_n] = P(y_n \in H) \qquad w.P.1$$

$$\bar{Z}_n \xrightarrow{Q} E[Z_n] = Q(y_n \in H) \qquad w.Q.1$$

$$\Longrightarrow P(y_n \in H) = 1 \quad , \quad Q(y_n \in H) = 1$$

و این اتفاق بر روی مجموعه های مجزا در $\mathcal{F}$ اتفاق می افتد. چون P و Q طبق فرض برابر نیستند پس تنها میدان تحقیق آنها باید متفاوت باشد که بتوانند مقادیر اندازه احتمال یکسانی را برای آرگومان های یکسان داشته باشند. در نتیجه اگرچه P بر Q مغلوب است اما این دو اندازه توأما منفرد[2] هستند. بنابراین P روی $\mathcal{F}_n$ منجر به صفر شدن Q می شود اما روی $\mathcal{F}$ تکیه گاه آنها از هم جدا می باشد.

مثال ۱۵-۵: فرض کنید که Z یک متغیر تصادفی انتگرال پذیر روی $(\Omega,\mathcal{F},P)$ باشد و $\mathcal{F}_n$ ها،دنباله ای از

σ- میدان‌های غیر نزولی $\mathcal{F}$ باشند. اگر،

$$X_n = E[Z||\mathcal{F}_n]$$

آنگاه، سه شرط اول مارتینگل برقرار است و تنها باید شرط (iv) را بررسی نمود. با استفاده از تعریف فوق داریم

$$E[X_{n+1}||\mathcal{F}_n] = E[E[z||\mathcal{F}_{n+1}]||\mathcal{F}_n] = E[z||\mathcal{F}_n] = X_n$$

---
[2] Mutually Singular

در نتیجه X_n یک مارتینگل است. رابطه فوق بیان می کند که، هر چه اطلاعات موجود از فرایند افزایش یابد(با تعریف Z این امر امکان پذیر است)، بازهم انتظارمان از مقدار فرایند در زمان های آینده، برابر با آن چیزی است که در آخرین مرحله از فرایند ثبت شده است.

۱۵–۲ زیرمارتینگل ها

تعریف ۱۵-۲ : فرض کنید $X_1, X_2, \cdots$ دنباله ای از متغیرهای تصادفی روی فضای احتمال $(\Omega, \mathcal{F}, P)$ باشند و فرض کنید $\mathcal{F}_1, \mathcal{F}_2, \cdots$ زیر سیگما میدان های $\mathcal{F}$ باشد. دنباله $\{(X_n, \mathcal{F}_n) : n = 1,2, \cdots\}$ یک زیرمارتینگل است اگر چهار شرط زیر برقرار باشند:

i) $\mathcal{F}_\text{n} \subset \mathcal{F}_{n+1}$ ،

ii) X_n نسبت به $\mathcal{F}_n$ اندازه پذیر باشد،

iii) X_n انتگرال پذیر باشد؛ یعنی: $E[|X_n|] < \infty$ ،

iv) با احتمال ۱ داشته باشیم: $E[X_{n+1}||\mathcal{F}_n] \geq X_n$.

مشاهده می شود که شرط های اول تا سوم با تعریف مارتینگل یکسان اند اما تفاوت در شرط چهارم وجود دارد و این شرط بیان می کند که با داشتن تمامی اطلاعات تا مرحله n ام فرایند، پیش بینی موجود از آینده فرایند در مرحله $n+1$ ام، بیشتر از چیزی است که در مرحله n ام اتفاق افتاده است. مانند مارتینگل ها در اینجا نیز به عنوان تعریفی دیگر، دنباله $X_1, X_2, \cdots$ یک مارتینگل وابسته به سیگما میدان‌های $\mathcal{F}_1, \mathcal{F}_2, \cdots$ خواهد بود.

همچنین مانند مارتینگل ها می توان کوچکترین σ- میدان $\mathcal{g}_n = \sigma\{X_1, X_2, \cdots, X_n\}$ ساخته شده از متغیرهای تصادفی را نیز برای تعریف مارتینگل ها ایجاد نمود. در اینجا با داشتن ویژگی های $\mathcal{g}_n \subseteq \mathcal{g}_{n+1}$ و اندازه پذیری X_n نسبت به $\mathcal{g}_n$ داریم

$$E[X_{n+1}||\mathcal{g}_n] = E[E[X_{n+1}||\mathcal{F}_n]||\mathcal{g}_n] \geq E[X_n||\mathcal{g}_n] = X_n$$

بنابراین با عمل انتگرال گیری از دو طرف رابطه زیر خواهیم داشت:

$$E[X_{n+1}||\mathcal{g}_n] \geq X_n$$

$$\Longrightarrow \int_A E[X_{n+1}||\mathcal{g}_n]\,\mathrm{dP} \geq \int_A X_n\,\mathrm{dP}$$

$$\Longrightarrow \int_A X_{n+1}\,\mathrm{dP} \geq \int_A X_n\,\mathrm{dP}$$

در نتیجه با تعریف σ- میدان $\mathcal{g}_n = \sigma\{X_1, X_2, \cdots, X_n\}$ نیز تمامی شروط زیرمارتینگل برقرار است.

برای σ- میدان خاص $\mathcal{g}_n$ ویژگی چهارم زیرمارتینگل به صورت زیر خواهد بود:

$$E[X_{n+1} || X_1, X_2, \cdots, X_n] \geq X_n$$

این شرط معادل با رابطه زیر است:

$$\int_A X_{n+1}\,\mathrm{dP} \geq \int_A X_n\,\mathrm{dP} \qquad\qquad A \in \mathcal{F}_n$$

همچنین از آنجایی که $\mathcal{F}_n$ ها تودرتو و صعودی هستند در نتیجه برای $A \in \mathcal{F}_n$ و $\mathcal{F}_\mathrm{n} \subset \mathcal{F}_{n+1}$ داریم:

$$\int_A X_n\,\mathrm{dP} \leq \int_A X_{n+1}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_n$$

$$\int_A X_{n+1}\,\mathrm{dP} \leq \int_A X_{n+2}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_{n+1}$$

$$\vdots$$

$$\int_A X_{n+k-1}\,\mathrm{dP} \leq \int_A X_{n+k}\,\mathrm{dP} \quad , \quad A \in \mathcal{F}_{n+k-1}$$

$$\Rightarrow \int_A X_n\,\mathrm{dP} \leq \int_A X_{n+1}\,\mathrm{dP} \leq \int_A X_{n+2}\,\mathrm{dP} \leq \cdots \leq \int_A X_{n+k}\,\mathrm{dP}$$

باید توجه داشت که اگر $A = \Omega$ باشد آنگاه بر اساس رابطه فوق داریم:

$$\int_\Omega X_n\,\mathrm{dP} \leq \int_\Omega X_{n+1}\,\mathrm{dP} \leq \int_\Omega X_{n+2}\,\mathrm{dP} \leq \cdots \leq \int_\Omega X_{n+k}\,\mathrm{dP}$$

$$\Rightarrow E[X_1] \leq E[X_2] \leq \cdots$$

بنابراین *امید ریاضی در زیر مارتینگل ها نانزولی است.*

مثال ۱۵-۶: فرض کنید $\Delta_1, \Delta_2, \cdots$ متغیرهای تصادفی مستقل و انتگرال پذیر باشند به طوری که برای $n \geq 2$ داشته باشیم، $E[\Delta_n] \geq 0$. اگر $\mathcal{F}_n = \sigma\{\Delta_1, \Delta_2, \cdots, \Delta_n\}$ برقرار باشد، در این صورت مجموع های جزئی $\Delta_1 + \Delta_2 + \cdots + \Delta_n$ تشکیل یک زیرمارتینگل می دهند.

$S_1 = \Delta_1$

$S_2 = \Delta_1 + \Delta_2$

$\vdots$

$S_n = \Delta_1 + \Delta_2 + \cdots + \Delta_n$

کافی است نشان دهیم این مجموع های جزئی در ویژگی (iv) زیرمارتینگل ها صدق می کنند. با توجه به استقلال Δ_n ها و همچنین، اندازه پذیری S_n نسبت به $\mathcal{F}_n$ داریم:

$$E[S_{n+1}||\mathcal{F}_n] = E[S_n + \Delta_{n+1}||\mathcal{F}_n] = E[S_n||\mathcal{F}_n] + E[\Delta_{n+1}||\mathcal{F}_n] = S_n + E[\Delta_{n+1}] \geq S_n$$

مثال ۱۵-۷: فرض کنید X_n ها نسبت به $\mathcal{F}_n$ ها مارتینگل باشند. پس $|X_n|$ نسبت به $\mathcal{F}_n$ اندازه پذیر و همچنین انتگرال پذیر است. با استفاده از نامساوی کوشی شوارتز داریم:

$$E[|X_{n+1}|||\mathcal{F}_n] \geq |E[X_{n+1}||\mathcal{F}_n]| = |X_n|$$

بنابراین $|X_n|$ نسبت به $\mathcal{F}_n$ زیرمارتینگل است. در حالت کلی ، توابعی همچون $|x|^p$ (با $p \geq 1$ و $E|x|^p < \infty$)

تشکیل زیر مارتینگل نسبت به $\mathcal{F}_n$ خواهند داد.

تذکر: اگر $X_1, X_2, \cdots, X_n$ تولید کننده فضای پیشامدهای $\mathcal{F}_n$ باشند، آنگاه $|X_1|, |X_2|, \cdots, |X_n|$ سیگما میدانی کوچکتر از $\mathcal{F}_n$ تولید خواهند کرد.

تذکر: به طور کلی $\sigma(f[X]) \subseteq \sigma(X)$ ، (سیگما میدان القا شده توسط هرتابعی از X کوچکتر از سیگما میدان القا شده توسط خود X است).

۱۵–۳ توابعی از مارتینگل ها

قضیه ۱۵-۱:

$i)$ اگر $X_1, X_2, \cdots$ ها نسبت به $\mathcal{F}_1, \mathcal{F}_2, \cdots$ مارتینگل و φ یک تابع محدب باشد و $\varphi(X_n)$ انتگرال پذیر باشد، آنگاه $\varphi(X_1), \varphi(X_2), \cdots$ نسبت به $\mathcal{F}_1, \mathcal{F}_2, \cdots$ زیرمارتینگل هستند.

$ii)$ اگر $X_1, X_2, \cdots$ ها نسبت به $\mathcal{F}_1, \mathcal{F}_2, \cdots$ زیرمارتینگل و φ یک تابع محدب غیرنزولی باشد و $\varphi(X_n)$ انتگرال پذیر باشد، آنگاه $\varphi(X_1), \varphi(X_2), \cdots$ نسبت به $\mathcal{F}_1, \mathcal{F}_2, \cdots$ زیرمارتینگل هستند.

اثبات:

ابتدا بند (i) قضیه را اثبات می کنیم. بنا به فرض بند (i) $E[X_{n+1}||\mathcal{F}_n] = X_n$ و φ یک تابع محدب است. بنابراین:

$$\varphi(X_n) = \varphi(E[X_{n+1}||\mathcal{F}_n]) \leq E[\varphi(X_{n+1})||\mathcal{F}_n]$$

$$\Rightarrow \varphi(X_n) \le E[\varphi(X_{n+1})||\mathcal{F}_n]$$

بنا به فرض بند (ii) $E[X_{n+1}||\mathcal{F}_n] \ge X_n$ و φ یک تابع محدب غیر نزولی است. بنابراین با اعمال تابع φ بر روی نابرابری مذکور و سپس بنا به محدب بودن تابع φ داریم:

$$\varphi(X_n) \le \varphi(E[X_{n+1}||\mathcal{F}_n]) \le E[\varphi(X_{n+1})||\mathcal{F}_n]$$

$$\Rightarrow \varphi(X_n) \le E[\varphi(X_{n+1})||\mathcal{F}_n]$$

مثال ۱۵-۸ : فرض کنید دنباله X_n ها نسبت به $\mathcal{F}_n$ ها مارتینگل باشند. آنگاه دنباله e^{X_n} ها نسبت به $\mathcal{F}_n$ یک زیر مارتینگل هستند.

اثبات: چهار شرط لازمه برای زیر مارتینگل بودن را بررسی می کنیم. طبق تعریف مارتینگل و فرض مثال، شرط *i* برقرار است. طبق فرض مثال، X_n نسبت به $\mathcal{F}_n$ اندازه پذیر است و می دانیم $\sigma\{X_1, X_2, \cdots, X_n\} \subset \mathcal{F}_n$ (کوچکترین سیگما میدان القا شده توسط X_n ها) . از طرفی دیگر، σ-میدان القا شده توسط تابعی همچون $f(x) = e^x$ یعنی $\mathcal{g}_n = \sigma\{e^{X_1}, e^{X_2}, \cdots, e^{X_n}\}$ زیر مجموعه ای از σ-میدان القا شده توسط خود X یعنی $\sigma\{X_1, X_2, \cdots, X_n\}$ است. بنابراین $\mathcal{g}_n \subset \sigma\{X_1, X_2, \cdots, X_n\} \subset \mathcal{F}_n$ حال می توان اندازه پذیر بودن دنباله e^{X_n} ها نسبت به $\mathcal{F}_n$ را نتیجه گرفت یعنی شرط *ii* برقرار است. شرط *iii* بدیهی است. همچنین با فرض $Y_n = e^{X_n}$ می توان نوشت

$$E[Y_{n+1}||\mathcal{F}_n] = E[e^{X_{n+1}}||\mathcal{F}_n] \ge e^{E[X_{n+1}||\mathcal{F}_n]} \ge e^{X_n} = Y_n$$

بنابراین Y_n ها نسبت به $\mathcal{F}_n$ یک زیر مارتینگل هستند.

■

۱۵-۴ زمان های توقف

فرض کنید τ یک متغیر تصادفی با مقادیر طبیعی باشد که $\tau: \Omega \to \Re$. آنگاه τ یک زمان توقف است اگر

$[\tau = k] \in \mathcal{F}_k$ به عبارت دیگر $[\omega: \tau(\omega) = k] \in \mathcal{F}_k$.

اگر τ زمان توقف باشد آنگاه $[\tau = k]$ بدین معناست که فرایند در نقطه X_k توقف دارد. برای اینکه τ یک متغیر تصادفی خوش تعریف باشد، باید نسبت به $\mathcal{F}_k$ که اطلاعات X_k را دربر دارد، اندازه پذیر باشد که تعریف فوق را نتیجه می دهد. در واقع τ روی اندیس فرایند X_t تعریف شده است. اما با توجه به اینکه $\mathcal{F}_k$ ها صعودی اند، پس

$[\tau = k] \in \mathcal{F}_k$ معادل با $[\tau \le k] \in \mathcal{F}_k$ می باشد. حال سیگما میدان مربوط به این متغیر تصادفی را با $\mathcal{F}_\tau$ تعریف می کنیم که در آن $[\tau \le k]$ با $[\tau = k]$ قابل جایگذاری است.

$$\mathcal{F}_\tau = \{A \in \mathcal{F}: A \cap [\tau \le \mathrm{k}] \in \mathcal{F}_k\,, \quad 1 \le k < \infty\}$$

در واقع این σ-میدان، اطلاعات موجود در فرایند در زمان های توقف کوچکتر یا مساوی k که شامل $\mathrm{X}_1, X_2, \cdots, X_k$ می شود را بر روی پیشامد A محدود می کند. یعنی این سیگما میدان ، مجموعه پیشامدهایی است که A رخ بدهد و فرایند در X_1 متوقف شود، یا A رخ بدهد و فرایند در X_2 متوقف شود و ... و A رخ بدهد و فرایند در X_k متوقف شود و در مجموع، پیشامد وقوع A است به طوری که A ، تا مرحله k ام فرایند رخ دهد و زمانی که A رخ می دهد، فرایند متوقف گردد و این k برای تمامی مقادیر بین ۱ تا بی نهایت تعریف می شود تا سیگما میدان $\mathcal{F}_\tau$ تمامی حالات وقوع A را حتی تا مرحله بی نهایت از فرایند نیز شامل شود.

طبق تعریف $[\tau = \mathrm{j}] \in \mathcal{F}_j$ و در نتیجه با توجه به $\{A \cap [\tau = \mathrm{j}]\} \subseteq [\tau = \mathrm{j}]$ داریم

$$\{A \cap [\tau = \mathrm{j}]\} \in \mathcal{F}_j$$

که نتیجه می دهد $[\tau = \mathrm{j}]$ نسبت به $\mathcal{F}_\tau$ اندازه پذیر است.

تعریف: اگر τ یک زمان توقف برای دنباله X_n باشد و A یک پیشامد باشد. آنگاه مجموعه A را مقدم بر τ گوییم هرگاه وقتی $\tau = \mathrm{n}$ باشد، با مشاهده متغیر های تصادفی $\mathrm{X}_1, X_2, \cdots, X_n$ بتوانیم بگوییم A اتفاق افتاده یا خیر.

اکنون مجموعه همه پیشامد های مقدم بر τ را با $\mathcal{F}_\tau$ نشان می دهیم. و در قسمت های الف ، ب و ج ذیل نشان می دهیم که $\mathcal{F}_\tau$ به فرم $\mathcal{F}_\tau = \{A \in \mathcal{F}: A \cap [\tau = \mathrm{k}] \in \mathcal{F}_k\,, 1 \le k < \infty\}$ یک سیـگما میدان است. البته $[\tau \le \mathrm{k}]$ با $[\tau = \mathrm{k}]$ قابل جایگذاری است.

الف: بدیهی است که $\emptyset \in \mathcal{F}_\tau$

ب: اگر $A \in \mathcal{F}_\tau$ آنگاه $\{\acute{A} \cap [\tau = \mathrm{n}]\} = \{\tau = \mathrm{n}\} - A \cap \{\tau = \mathrm{n}\} \in \mathcal{F}_n$ بنابراین $\acute{A} \in \mathcal{F}_\tau$

ج: اگر $A_i \in \mathcal{F}_\tau$ برای... $i = 1,2,$ آنگاه $\{(\cup_i A_i) \cap [\tau = \mathrm{n}]\} = \cup_i \{A_i \cap [\tau = \mathrm{n}]\} \in \mathcal{F}_n$ بنابراین $\cup_i A_i \in \mathcal{F}_\tau$

■

در واقع $\mathcal{F}_\tau$ ، سیگما میدان ساخته شده برای پیشامدهایی است که تا زمان های توقف مختلف رخ می دهند. به بیان دیگر، سیگما میدان این فضا است که تا زمان توقف X_1 ($\tau = 1$) چه پیشامدهایی رخ می دهند، ... و تا زمان توقف X_k ($\tau = \mathrm{k}$) چه پیشامدهایی رخ می دهند. در واقع اطلاعات $\mathcal{F}_\tau$ ، شامل مقادیر $\tau(\omega), X_1(\omega), \cdots, X_{\tau(\omega)}(\omega)$ می باشد، که نشان می دهد $X_{\tau(\omega)}(\omega)$ یک متغیر تصادفی است که فضای Ω را افزار می کند. چون

$$I_A(\omega) = I_A(\omega') \quad \forall A \in \mathcal{F}_\tau \qquad \Leftrightarrow \qquad X_i(\omega) = X_i(\omega') \;\;:\;\; i \le \tau(\omega) = \tau(\omega')$$

به عبارت دیگر، برای هر X_i با $i \leq \tau(\omega)$ ، ω و ω' در یک افراز قرار می گیرند اگر و تنها اگر A به طور همزمان با ω و ω' رخ دهد و چون ω و ω' هم زمان رخ می دهند(در یک افراز قرار دارند) لذا $\tau(\omega) = \tau(\omega')$.

حال نشان می دهیم که X_{τ_1}, X_{τ_2} مارتینگل هستند. کافی است نشان دهیم که تحت فرض مارتینگل بودن X_n ها، $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$. بنابراین برای $\tau_1 \leq \tau_2$ ، $\forall A \in \mathcal{F}_{\tau_1} \quad : \quad A \cap [\tau_1 \leq k] \in \mathcal{F}_k$ ، از طرفی

$$A \cap [\tau_1 \leq k] \cap [\tau_2 \leq k] = A \cap [\tau_2 \leq k] \in \mathcal{F}_{\tau_2} \quad \Longrightarrow \quad A \in \mathcal{F}_{\tau_2}$$

$$\Rightarrow \ \mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$$

نکته: فرض کنید $X_1, X_2, \cdots, X_n$ نسبت $\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$ مارتینگل باشند و τ زمان توقف باشد، آنگاه $X_{\tau\wedge n}$ ($\tau\wedge n = Min(\tau, n)$) نسبت به $\mathcal{F}_\tau$ مارتینگل است.

اثبات: به ساده گی می توان نوشت

$$\begin{aligned} {X_{\tau\wedge n}}^{-1}(\beta) &= \{\omega : X_{\tau\wedge n}(\omega) \in \beta\} \\ &= \{X_\tau \in \beta \cap \tau_2 \leq \mathrm{n}\} \cup \{X_\tau \in \beta \cap \tau_2 \leq \mathrm{n}\} \\ &= \left(\bigcup_{i=1}^{n} \{\omega : X_\mathrm{i} \in \beta, \tau = \mathrm{i}\}\right) \cup \{\omega : X_\mathrm{n} \in \beta, \tau > \mathrm{n}\} \in \mathcal{F}_\mathrm{n} \end{aligned}$$

بنابراین $X_{\tau\wedge n}$ اندازه پذیر است.

از طرفی دیگر $X_{\tau\wedge n}$ انتگرال پذیر است زیرا

$$|X_{\tau\wedge n}| \leq \sum_{i=1}^{\mathrm{n}} |\mathrm{X_i}| \Rightarrow E|X_{\tau\wedge n}| < \infty$$

همچنین داریم

$$X_{\tau\wedge n} = X_\tau I_{\tau<n} + X_\mathrm{n} I_{\tau\geq n} = \sum_{i=1}^{\mathrm{n-1}} X_\mathrm{i} I_{\tau=i} + X_\mathrm{n} I_{\tau\geq n}$$

بنابراین می توان نوشت

$$\begin{aligned} E[X_{\tau\wedge n} || \mathcal{F}_{n-1}] &= \sum_{i=1}^{\mathrm{n-1}} X_\mathrm{i} I_{\tau=i} + I_{\tau\geq n} E[X_n || \mathcal{F}_{n-1}] \\ &= \sum_{i=1}^{\mathrm{n-2}} X_\mathrm{i} I_{\tau=i} + X_{\mathrm{n-1}} I_{\tau=\mathrm{n-1}} + I_{\tau>\mathrm{n-1}} X_{\mathrm{n-1}} \end{aligned}$$

$$= X_{\tau} I_{\tau<n-1} + X_{n-1} I_{\tau\geq n-1}$$
$$= X_{\tau\wedge(n-1)}$$

قضیه ۱۵-۲: اگر $X_1, X_2, \cdots, X_n$ نسبت $\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$ زیرمارتینگل باشند و $1 \leq \tau_1 \leq \tau_2 \leq n$ زمان های توقف باشند، آنگاه X_{τ_1}, X_{τ_2} نسبت به $\mathcal{F}_{\tau_1}$ و $\mathcal{F}_{\tau_2}$ زیرمارتینگل است.

اثبات: در ابتدا با $A = X_{\tau}^{-1}(\beta)$ داریم

$$A \cap [\tau_1 \leq n] = \{X_\tau \in \beta \cap \tau_2 \leq \mathrm{n}\} = \bigcup_{k=0}^{n} \{X_\mathrm{k} \in \beta \cap \tau_2 = \mathrm{k}\} \in \mathcal{F}_\mathrm{n}$$

بنابراین X_{τ_1} نسبت به $\mathcal{F}_{\tau_1}$ اندازه پذیر است. به همین ترتیب X_{τ_2} نسبت به $\mathcal{F}_{\tau_2}$ نیز اندازه پذیر است. از طرفی $\left|X_{\tau_i}\right|$ توسط $\sum_{k=1}^{\mathrm{n}} |\mathrm{X_k}|$ مغلوب می شود پس $\left|X_{\tau_i}\right|$ انتگرال پذیر می باشد چون شرط انتگرال پذیری X_n را داشته ایم. بنابراین مجموع متناهی آنها نیز اندازه پذیر می باشدپس باید نشان دهیم که، $E[X_{\tau_2} || \mathcal{F}_{\tau_1}] \geq X_{\tau_1}$ و یا برای هر $A \in \mathcal{F}_{\tau_1}$ ، $\int_A X_{\tau_2} - X_{\tau_1} dp \geq 0$.

$$A \in \mathcal{F}_{\tau_1} \implies A \cap [\tau_1 < k \leq \tau_2] = A \cap [\tau_1 \leq \mathrm{k}-1] \cap \; [\tau_2 \leq \mathrm{k}-1]^c \in \mathcal{F}_{k-1}$$

حال اگر $\Delta_k = \mathrm{X}_k - \mathrm{X}_{k-1}$ تعریف شود، آنگاه:

$$\int_A X_{\tau_2} - X_{\tau_1} dp = \int_A \sum_{k=1}^{n} I_{[\tau_1<k\leq\tau_2]} \Delta_k \, dp$$
$$= \sum_{k=1}^{n} \int_A I_{[\tau_1<k\leq\tau_2]} \Delta_k dp$$

و بنابراین با توجه به زیرمارتینگل بودن Δ_k ها نسبت به $\mathcal{F}_k$ نتیجه حاصل می شود.

$$\int_A X_{\tau_2} - X_{\tau_1} dp \;= \sum_{k=1}^{n} \int_{A\cap[\tau_1<k\leq\tau_2]} \Delta_k dp \geq 0 \quad ; \;\; A \cap [\tau_1 < k \leq \tau_2] \in \mathcal{F}_{k-1}$$

زیرا،

$$A \cap [\tau_1 < k \leq \tau_2] = \; A \cap [\tau_1 \leq k-1] \cap \; [\tau_2 \leq k-1]^c \in \mathcal{F}_{k-1}$$

و از طرفی طبق فرض مان می دانیم که $E[X_k - X_{k-1} || \mathcal{F}_k] \geq 0$. بنابراین می توان نوشت که

$$E[X_{\tau_2} || \mathcal{F}_{\tau_1}] \geq X_{\tau_1}$$

۱۵–۵ نامساوی‌ها

در این بخش دو نامساوی اساسی مورد مطالعه قرار می گیرد.

قضیه ۱۵-۳: اگر $X_1, X_2, \cdots, X_n$ زیرمارتینگل باشند، آنگاه برای $\alpha \geq 0$ داریم:

$$P[\max_{i\leq n} X_i \geq \alpha] \leq \frac{1}{\alpha} E[|X_n|]$$

اثبات: برای $\alpha \geq 0$ دلخواه، فرض کنید $A_1 = [X_1 \geq \alpha]$، $A_k = [\max_{i\leq k} X_i \geq \alpha]$ که نتیجه می دهد:

$$A = \bigcup_{k=1}^{n} A_k = [\max_{i\leq n} X_i \geq \alpha]$$

از آنجا که X_n زیرمارتینگل است، پس $A_k \in \mathcal{F}_k = \sigma\{X_1, X_2, \cdots, X_k\}$ و بنابراین با توجه به اندازه پذیری A_k نسبت به $\mathcal{F}_k$ داریم:

$$\int_A X_n \, dP = \sum_{k=1}^{n} \int_{A_k} X_n dp = \sum_{k=1}^{n} \int_{A_k} E[X_n || \mathcal{F}_k] dp \geq \sum_{k=1}^{n} \int_{A_k} X_k dp$$

حال می دانیم که X_n زیرمارتینگل است و در نتیجه:

$$\int_{A_k} X_k \, dP \geq \int_{A_k} X_{k-1} \, dPdP \geq \cdots \geq \int_{A_k} X_1 \, dP \quad ; \; A_k \in \mathcal{F}_k, A_k \in \mathcal{F}_j : \; j < k$$

$$\Longrightarrow \int_{A_k} X_k \, dP \geq \int_{A_k} \max(X_1, X_2, \cdots, X_k) \, dP \geq \int_{A_k} \alpha \, dP = \alpha P(A)$$

$$\Longrightarrow \alpha P(A) \leq \int_A X_n \, dP = \int_{[\max_{i\leq n} X_i \geq \alpha]} X_n \, dP \leq \int_{[\max_{i\leq n} X_i \geq 0]} X_n \, dP$$

$$\leq \int_{[X_n \geq 0]} X_n \, dP = E[X_n^+]$$

$$\leq E[X_n^+] + E[X_n^-] = E[|X_n|]$$

بنابراین قضیه اثبات شده است.

مثال ۱۵-۹: اگر $S_1, S_2, \cdots$ مجموع های جزئی متغیرهای تصادفی مستقل با میانگین صفر باشند، در این صورت بنا به مثال۱۵-۱، یک مارتینگل تشکیل می دهند. اگر واریانس ها را متناهی در نظر بگیریم، آنگاه $S_1^2, S_2^2, \cdots$ یک زیرمارتینگل را تشکیل می دهند و رابطه حاکم بر قضیه ۱۴-۳ برای این فرایند تصادفی به صورت زیر خواهد بود:

$$P\left[\max_{i\leq n} S_i^2 \geq \alpha\right] \leq \frac{1}{\alpha^2} E[S_n^2]$$

یا به عبارت دیگر، با توجه به صفر بودن میانگین ها و برابری واریانس و امید توان دوم متغیرها:

$$P\left[\max_{i\le n}|S_i| \ge \alpha\right] \le \frac{1}{\alpha^2}Var[S_n]$$

که این رابطه همان نامساوی کلموگروف است.

دومین نامساوی مهم و اساسی این فصل ابتدا نیاز به داشتن تصوری از فراتقاطع[۳] ها دارد. فرض کنید بازه $[\alpha,\beta]$ یک فاصله باشد که $\alpha<\beta$ و $X_1,X_2,\cdots,X_n$ متغیرهای تصادفی باشند. تعداد فراتقاطع های $[\alpha,\beta]$ که به وسیله $X_1(\omega),X_2(\omega),\cdots,X_n(\omega)$ ایجاد می شوند، برابر با تعداد دفعاتی است که دنباله از زیر α به بالای β می رود. به طور خاص تعداد فراتقاطع ها برابر با تعداد جفت های u و v است که اولا مقدار صحیح داشته باشند و ثانیا در روابط زیر صدق کنند:

برای $1\le u<v\le n$، داشته باشیم:
$$\begin{cases} X_i \le \alpha & i=u \\ \alpha < X_i < \beta & u<i<v \\ X_i \ge \beta & i=v \end{cases}$$

قضیه ۱۵-۴: اگر $X_1,X_2,\cdots,X_n$ زیرمارتینگل باشند، آنگاه تعداد فراتقاطع ها در بازه $[\alpha,\beta]$ در رابطه زیر صدق می کند:

$$E[U] \le \frac{E[|X_n|]+\alpha}{\beta-\alpha}$$

اثبات: بنا به قضیه۲ برای $\tau_2 \ge \tau_1$ داریم: $E[X_{\tau_2}] \ge E[X_{\tau_1}]$. حال τ_i ها را در فرایند فراتقاطع ها به صورت زیر تعریف می کنیم.

τ_1 = اولین j که $X_j \le \alpha$ باشد و در غیر این صورت برابر با n است.

τ_k (برای k فرد)= کوچکترین j که $\tau_{k-1} < j \le n$ و $X_j \le \alpha$ باشد و در غیر این صورت برابر با n است.

τ_k (برای k زوج)= کوچکترین j که $\tau_{k-1} < j \le n$ و $X_j \ge \beta$ باشد و در غیر این صورت برابر با n است.

و در نتیجه، U برابر با بزرگترین i است که در رابطه $X_{\tau_{2i-1}} \le \alpha < \beta \le X_{\tau_{2i}}$ صدق کند.

به منظور اثبات قضیه، فاصله α تا β را به اندازه α به پایین منتقل می کنیم و مقادیر متغیر تصادفی $X_k-\alpha$ را که به مقدار منفی می رسند برابر با صفر در نظر می گیریم. به عبارت دیگر، متغیر تصادفی Y_k را به صورت زیر تعریف می کنیم و این متغیر تصادفی مقادیر نامنفی را در بازه $[0,\theta]$ اختیار می کند به طوری که $\theta=\beta-\alpha$.

$$Y_k = \max\{X_k-\alpha,0\} = (X_k-\alpha)^+$$

---
[۳] Upcrossing

بنابراین U برابر با تعداد فراتقاطع ها در بازه $[0,\theta]$ برای زیرمارتینگل های $Y_1, Y_2, \cdots, Y_n$ می باشد (دقت کنید که تابع $(X_k-\alpha)^+$ یک تابع محدب صعودی از زیرمارتینگل هاست، بنابراین Y_nها نیز زیر مارتینگل اند) به طوری که زمان های توقف برای این فرایند به صورت زیر تعریف می شوند:

$$\tau_0 = 1$$

τ_k (برای k فرد): کوچکترین j که $\tau_{k-1} < j \le n$ و $Y_j = 0$ باشد و در غیر این صورت برابر با n است.

τ_k (برای k زوج): کوچکترین j که $\tau_{k-1} < j \le n$ و $Y_j \ge \theta$ باشد و در غیر این صورت برابر با n است.

نشان می دهیم که τ_k ها با تعریف فوق، زمان های توقف هستند. کافی است نشان دهیم که $[\tau_k = j] \in \mathcal{F}_j$ بدین منظور ابتدا فرض کنید که k عددی فرد باشد. در این صورت:

$$[\tau_k = j] = \cup_{i=1}^{j-1}[\tau_k = j, \tau_{k-1} = i] = \cup_{i=1}^{j-1}[\,\tau_{k-1} = i, Y_{i+1} > 0, \cdots, Y_{j-1} > 0, Y_j = 0]$$

بنابراین با استناد به اندازه پذیری هر یک از اجزای عبارت سمت راست معادله بالا نسبت به $\mathcal{F}_i$، $\mathcal{F}_{i+1}$ و ... $\mathcal{F}_j$ در نتیجه، $[\tau_k = j]$ نسبت به $\mathcal{F}_j$ اندازه پذیر است. یعنی: $[\tau_k = j] \in \mathcal{F}_j$.

از طرفی برای $[\tau_k = n]$ نیز داریم

$$[\tau_k = n] = [\tau_k \le n-1]^c \in (\mathcal{F}_{n-1} \subset \mathcal{F}_n)$$

در نتیجه τ_k برای k فرد، یک زمان توقف است و برای k زوج نیز به طور مشابه اثبات می شود.

از طرفی چون τ_k اکیدا صعودی است تا به نقطه k=n برسد داریم:

$$Y_n = Y_{\tau_n} \ge Y_{\tau_n} - Y_0 = Y_{\tau_n} - Y_{\tau_{n-1}} + Y_{\tau_{n-1}} - Y_{\tau_{n-2}} + Y_{\tau_{n-2}} - \cdots - Y_{\tau_0}$$
$$= \sum_{k=2i}(Y_{\tau_k} - Y_{\tau_{k-1}}) + \sum_{k=2i-1}(Y_{\tau_k} - Y_{\tau_{k-1}}) \quad ; \;\; 1 \le k \le n$$

حال با گرفتن امید ریاضی از طرفین نامعادله و استناد به قضیه۱۵-۲ که نتیجه ی آن در این بخش، نامنفی بودن امید ریاضی آخرین عبارت سمت راست نامعادله می باشد، داریم:

$$E[Y_n] \ge E[\sum_{k=2i}(Y_{\tau_k} - Y_{\tau_{k-1}})]$$

می دانیم تمامی جملات امید ریاضی سمت راست نامعادله، مثبت هستند و از طرفی $Y_{\tau_k} - Y_{\tau_{k-1}} \ge \theta$ در نتیجه:

$$E[\sum_{k=2i}(Y_{\tau_k} - Y_{\tau_{k-1}})] \ge E[\sum_{k=2i}\theta] = \theta E[\sum_{k=2i}\ ]$$

واضح است که $\sum_{k=2i}$ برابر با تعداد فراتقاطع ها در طول بازه مورد مطالعه است. چرا که تعداد اندیس های زوج شده در فرایند را شمارش می کند و می دانیم تعداد فراتقاطع ها برابر با تعداد اندیس های زوج برای فرایند Y_k است. در نتیجه:

$$E[Y_n] \geq E[\sum_{k=2i}(Y_{\tau_k} - Y_{\tau_{k-1}})] \geq \theta \sum_{k=2i} \ = \theta E[U]$$

$$\Longrightarrow E[U] \leq \frac{E[Y_n]}{\theta}$$

حال برای $E[Y_n]$ نیز داریم:

$$E[Y_n] = \int Y_n dp = \int (X_n - \alpha) I_{[X_n > \alpha]} dp = \int_{[X_n > \alpha]} (X_n - \alpha) dp = \int_{[X_n - \alpha > 0]} (X_n - \alpha) dp$$

$$= E[|X_n - \alpha|^+] \leq E[|X_n - \alpha|^+] + E[|X_n - \alpha|^-]$$

$$= E[|X_n - \alpha|] \leq E[|X_n| + |\alpha|] = E[|X_n|] + |\alpha|$$

در نتیجه با توجه به نامنفی بودن α قضیه اثبات می شود و داریم:

$$E[U] \leq \frac{E[|X_n|] + |\alpha|}{\beta - \alpha} = \frac{E[|X_n|] + \alpha}{\beta - \alpha}$$

۱۵–۶ قضایای همگرایی

قضیه همگرایی مارتینگل ها منسوب به دوب[۴] شکل های متفاوتی دارد که ساده ترین شکل آن به صورت زیر است.

قضیه ۱۵-۵: فرض کنید $X_1, X_2, \cdots, X_n$ زیرمارتینگل باشند. اگر $k = SupE[|X_n|] < \infty$ ، آنگاه X_n با احتمال ۱ به X میل می کند به طوری که X یک متغیر تصادفی است که $E[|X|] \leq k$.

اثبات: برای α و β ثابت، فرض کنید U_n تعداد فراتقاطع های فرایند $X_1, X_2, \cdots, X_n$ در طول بازه $[\alpha, \beta]$ باشد. بر اساس قضیه۴ و قرار دادن $E[|X_n|] \leq SupE[|X_n|] = k$ داریم:

$$E[U_n] \leq \frac{E[|X_n|] + |\alpha|}{\beta - \alpha} \leq \frac{k + \alpha}{\beta - \alpha}$$

---
[۴] Doob

از آنجا که U_n غیر نزولی است و $E[U_n]$ کراندار است، از قضیه همگرایی یکنوا نتیجه می گیریم که $SupU_n$ انتگرال پذیر است و تقریبا همه جا مقادیر متناهی اختیار می کند. حال فرض کنید X^* و X_* به ترتیب $limsup$ و $liminf$ دنباله متغیرهای تصادفی $X_1, X_2, \cdots$ باشند که ممکن است مقادیر بی نهایت را نیز بگیرند.

اگر $X_* < \alpha < \beta < X^*$ باشد، آنگاه U_n باید به سمت بی نهایت برود. زیرا همیشه X^* و X_* به ترتیب در بالا و پایین β و α قرار می گیرند و بی نهایت بار نیز این امر تکرار می شود بنابراین همیشه یک مقدار U به مقادیر قبلی آن افزوده می شود. اما از آنجا که U_n با احتمال ۱ کراندار است، پس:

$$P[X_* < \alpha < \beta < X^*] = 0.$$

حال با فرض مجموعه A به فرم $A = \{\omega: X_n(\omega)$ به یک حد متناهی و یا نا متناهی همگرا نیست$\}$ نشان می دهیم که احتمال این مجموعه برابر با صفر است.

$$A = \{\omega: -\infty \le X_* < X^* \le \infty\}$$
$$= \bigcup_{\substack{\alpha<\beta \\ \alpha,\beta\in Q}} \{\omega: -\infty \le X_* < \alpha < \beta < X^* \le \infty\}$$
$$= \bigcup_{\substack{\alpha<\beta \\ \alpha,\beta\in Q}} \{U_n = \infty\}$$

بنابراین داریم

$$P(A) \le \sum_{\substack{\alpha<\beta \\ \alpha,\beta\in Q}} P(U_n = \infty) = 0$$

پس نمی توان رابطه $X_* < X^*$ را پذیرفت و در نتیجه X^* و X_* با احتمال ۱ برابر هستند و بنابراین دنباله $X_1, X_2, \cdots, X_n$ به مقدار یک متغیر تصادفی مانند X میل می کند که این مقدار می تواند نامتناهی نیز باشد. بر اساس لم فاتو و فرض مسئله داریم که: $E[|X|] \le liminf E[|X_n|] \le k$ بنابراین $E[|X|] \le k$.

■

حال اگر X_n ها تشکیل یک زیرمارتینگل بدهند آنگاه داریم: $E[|X_1|] \le E[|X_2|] \le \cdots$ و بنابراین $E[|X_n|]$ نزولی نخواهد بود، بلکه صعود می کند و حد هم دارد و حد آن برابر با سوپریمم آن است به طوری که، $limE[|X_n|] = k$.

تذکر: فرض متناهی بودن k در این قضیه، فرضی حیاتی است چرا که اگر X_n ها تشکیل یک مارتینگل نامنفی بدهند، آنگاه با توجه به $E[|X_n|] = E[X_n] = E[X_1] < \infty$ ، k لزوما متناهی خواهد بود.

پیش از ارائه دومین قضیه همگرایی در این بخش باید لم زیر را که در اثبات قضیه مورد استفاده قرار می گیرد، مورد توجه قرار داد.

لم ۱۵-۱ : اگر Z انتگرال پذیر و $\mathcal{F}_n$ دنباله ای از سیگما میدان دلخواه در $\mathcal{F}$ باشند، آنگاه متغیر تصادفی $E[Z||\mathcal{F}_n]$ به طور یکنواخت انتگرال پذیر است.

تذکر: تابع f_n به طور یکنواخت انتگرال پذیر است اگر $\lim\limits_{\alpha} Sup_n \int_{[|f_n|>\alpha]} |f_n|\, d\,\mu \;=\; 0$

اثبات: از آنجا که $|E[Z||\mathcal{F}_n]| \le E[|Z|||\mathcal{F}_n]$. بنابراین اگر $E[|Z|||\mathcal{F}_n]$ به طور یکنواخت انتگرال‌پذیر باشد در نتیجه $E[Z||\mathcal{F}_n]$ نیز به طور یکنواخت انتگرال پذیر است. به عبارت دیگر اگر تابعی مانند g_n وجود داشته باشد که $|g_n| \le f_n$ آنگاه، $[|g_n| > \alpha] \subseteq [f_n > \alpha]$ و بنابراین:

$$\lim_{\alpha} Sup_n \int_{[|g_n|>\alpha]} |g_n|\, d\,\mu \;\le \lim_{\alpha} Sup_n \int_{[|g_n|>\alpha]} f_n\, d\,\mu \;\le \lim_{\alpha} Sup_n \int_{[|f_n|>\alpha]} f_n\, d\,\mu$$

و اگر عبارت سمت راست به صفر میل کند، آنگاه عبارت سمت چپ نیز به صفر میل می کند؛ یعنی g_n انتگرال پذیر یکنواخت است. پس کافی است نشان دهیم $\underset{\alpha}{lim}\, Sup_n \int_{[E[|Z|||\mathcal{F}_n]>\alpha]} E[|Z|||\mathcal{F}_n]\, d\,\mu$ به صفر میل می کند تا ثابت شود که $\lim\limits_{\alpha} Sup_n \int_{[|E[Z||\mathcal{F}_n]|>\alpha]} |E[Z||\mathcal{F}_n]|\, d\,\mu$ به صفر میل می کند و در نتیجه $E[Z||\mathcal{F}_n]$ انتگرال‌پذیر یکنواخت است.

ابتدا تعریف می کنیم، $A_{\alpha n} = [E[|Z|||\mathcal{F}_n] > \alpha]$ و چون $E[|Z|||\mathcal{F}_n]$ نسبت به $\mathcal{F}_n$ اندازه پذیر است، پس $A_{\alpha n}$ نیز نسبت به $\mathcal{F}_n$ اندازه پذیر خواهد بود. با استناد به اندازه پذیری $A_{\alpha n}$ نسبت به $\mathcal{F}_n$ و تعریف امید ریاضی شرطی در فصل های پیشین داریم:

$$\int_{A_{\alpha n}} E[|Z|||\mathcal{F}_n]\, dp = \int_{A_{\alpha n}} |Z|\, dp$$

پس کافی است نشان دهیم $\lim\limits_{\alpha} Sup_n \int_{A_{\alpha n}} |Z|\, dp = 0$.

بنا به تعریف $limsup$ و همچنین تعریف همگرایی دنباله ها، معادل با شرط انتگرالی فوق الذکر، باید نشان دهیم که:

$$\forall \varepsilon > 0 \quad \exists \alpha \;:\; \forall n \quad \int_{A_{\alpha n}} |Z|\, dp < \varepsilon$$

می دانیم که $\int_{A_{\alpha n}} |Z|\, dp$ تابعی از $A_{\alpha n}$ یا به عبارت ساده تر، A است و همچنین $|Z|$ تابعی مثبت است لذا بنا به قضیه رادون نیکودیم، $|Z|$ می تواند چگالی اندازه ای مانند μ باشد که توسط اندازه P مغلوب است. یعنی:

$$\mu(A) = \int_A |Z|\, dp \qquad \Longrightarrow \qquad \mu \ll P$$

با استناد به روش $\varepsilon - \delta$ در قضیه پیوستگی به طور مطلق و بکارگیری آن برای دو اندازه فوق داریم:

$$\forall\varepsilon>0 \quad \exists\delta \ : \ P(A)<\delta \quad \Longrightarrow \quad \mu(A)=\int_A |Z|\,dp<\varepsilon$$

اما برای α به قدر کافی بزرگ، بنابه نامساوی مارکوف داریم:

$$P[E[|Z|||\mathcal{F}_n]\geq\alpha]\leq\frac{1}{\alpha}E[E[|Z|||\mathcal{F}_n]=\frac{1}{\alpha}E[|Z|]<\delta$$

و این مقدار به n بستگی ندارد؛ لذا برای هر n برقرار است. در نتیجه برای $\alpha\to\infty$:

$$P[A_{\alpha n}]\leq\frac{1}{\alpha}E[|Z|]<\delta$$

$$\Longrightarrow \ P[A_{\alpha n}]<\delta$$

$$\Longrightarrow \ \mu(A)=\int_A |Z|\,dp<\varepsilon$$

$$\Longrightarrow \ \forall\varepsilon>0 \quad \exists\alpha \ : \ \forall n \quad \int_{A_{\alpha n}}|Z|\,dp<\varepsilon$$

$$\Longrightarrow \ \lim_{\alpha} Sup_n \int_{A_{\alpha n}}|Z|\,dp<\varepsilon$$

و بنابراین، لم اثبات شده است.

قضیه ۱۵-۶: اگر $\mathcal{F}_n\uparrow\mathcal{F}_\infty$ و Z انتگرال پذیر باشد، آنگاه با احتمال برابر با ۱ داریم

$$E[Z||\mathcal{F}_n]\to E[Z||\mathcal{F}_\infty]$$

تذکر: اگر $\mathcal{F}_n$ دنباله ای صعودی از σ- میدان ها باشد، آنگاه $\bigcup_{n=1}^{\infty}\mathcal{F}_n$ را با $\mathcal{F}_\infty$ نشان می دهیم و می‌توان نوشت $\mathcal{F}_n\uparrow\mathcal{F}_\infty$. البته لزومی ندارد که $\mathcal{F}_\infty$ دقیقا همان $\bigcup_{n=1}^{\infty}\mathcal{F}_n$ باشد ولی لازم است که $\bigcup_{n=1}^{\infty}\mathcal{F}_n$ مولد $\mathcal{F}_\infty$ باشد.

اثبات: $X_n=E[Z||\mathcal{F}_n]$ یک مارتینگل نسبت به $\mathcal{F}_n$ می باشد و تحت شرط انتگرال پذیری Z، بر اساس لم ۱۵-۱، $E[Z||\mathcal{F}_n]$ انتگرال پذیر یکنواخت می باشد و بنابراین X_n نیز انتگرال پذیر یکنواخت است.

همچنین داریم

$$|X_n|=|E[Z||\mathcal{F}_n]|\leq E[|Z|||\mathcal{F}_n]$$

$$\Longrightarrow E[|X_n|]\leq E\big[E\big[|Z|\big||\mathcal{F}_n\big]\big]=E[|Z|]$$

تحت شرط انتگرال پذیری Z، $E[|Z|]<\infty$ و در نتیجه $E[|X_n|]<\infty$ و بنابراین $SupE[|X_n|]<\infty$. با توجه به اینکه X_n یک مارتینگل نسبت به $\mathcal{F}_n$ می باشد، با استناد به قضیه۱۵-۵، X_n با احتمال ۱ به X میل می کند به طوری که X یک متغیر تصادفی انتگرال پذیر است. پس ثابت شد که $X_n\to X$ و X انتگرال پذیر است.

حال باید نشان دهیم که $X = E[Z||\mathcal{F}_\infty]$ و برای این منظور باید نشان دهیم که X ورژنی از $E[Z||\mathcal{F}_\infty]$ است. با توجه به انتگرال پذیری یکنواخت X_n و $X_n \to X$ و همچنین استناد به قضیه همگرایی یکنوا، $\int X_n \to \int X$. بنابراین :

$$\lim_n \int_A X_n \, dp = \int_A X \, dp \quad : A \epsilon \mathcal{F}_k \quad , \quad k \le n$$

بنابراین با استفاده از ویژگی مارتینگل بودن X_n ها داریم:

$$\int_A X_n \, dp = \int_A E[Z||\mathcal{F}_n] \, dp = \int_A Z \, dp \quad , \; A \epsilon \mathcal{F}_n \;\; \Longrightarrow \;\; A \epsilon \bigcup_{k=1}^{n} \mathcal{F}_k$$

بنابراین برای هر $A \epsilon \bigcup_{k=1}^{n} \mathcal{F}_k$ رابطه زیر برقرار است:

$$\lim_n \int_A X_n \, dp = \int_A X \, dp = \lim_n \int_A E[Z||\mathcal{F}_n] \, dp = \int_A Z \, dp \quad , \quad A \epsilon \bigcup_{k=1}^{\infty} \mathcal{F}_k$$

می توان نشان داد که $\bigcup_{k=1}^{\infty} \mathcal{F}_k$ ها تشکیل یک π -سیستم می دهند. پس رابطه بالا که به ازای هر A متعلق به π –سیستم برقرار است، به ازای هر A متعلق به σ- میدان تولید شده توسط این π –سیستم نیز برقرار است. حال چون $X_n \to X$ و X_n نسبت به $\mathcal{F}_n$ اندازه پذیر است و در نتیجه نسبت به $\mathcal{F}_{n+1}$ ، $\mathcal{F}_{n+2}$ ، ... و $\mathcal{F}_\infty$ نیز اندازه پذیر است، پس حد آن هم نسبت به $\mathcal{F}_\infty$ اندازه پذیر می شود. در نتیجه، $A \epsilon \mathcal{F}_\infty$ و اندازه پذیری X نسبت به $\mathcal{F}_\infty$ نتیجه می دهد که X ورژنی از $E[Z||\mathcal{F}_\infty]$ است و قضیه اثبات می شود.

۱۵–۷ مارتینگل وارون

تعریف ۱۵-۳: دنباله متغیرهای تصادفی $X_{-1}, X_{-2}, \cdots$ نسبت به σ- میدان های $\mathcal{F}_{-1}, \mathcal{F}_{-2}, \cdots$ یک مارتینگل وارون نامیده می شود اگر در شروط زیر صدق کند.

i) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ ، $n < -1$

ii) X_n برای $n \le -1$ نسبت به $\mathcal{F}_n$ اندازه پذیر باشد،

iii) X_n برای $n \le -1$ انتگرال پذیر باشد؛ یعنی: $E[|X_n|] < \infty$ ،

iv) با احتمال ۱ برای $n < -1$ داشته باشیم: $E[X_{n+1}||\mathcal{F}_n] = X_n$.

قضیه ۱۵-۷: برای مارتینگل وارون، حد $\lim_{n\to\infty} X_{-n} = X$ وجود دارد و انتگرال پذیر است و برای هر n، $E[X_{-n}] = E[X]$.

اثبات: اثبات این قضیه همانند اثبات قضیه ۱۲-۵ می باشد. فرض کنید X^* و X_* به ترتیب $limsup$ و $liminf$ دنباله متغیرهای تصادفی $X_{-1}, X_{-2}, \cdots$ باشند که ممکن است مقادیر بی نهایت را نیز بگیرند.

برای دنباله متناهی $X_{-n}, \cdots, X_{-2}, X_{-1}$ اگر $X_* < \alpha < \beta < X^*$ باشد، آنگاه U_n در اینجا نیز باید به سمت بی نهایت برود. اما از آنجا که U_n با احتمال ۱ کراندار است و $E[U_n] \leq \frac{E[|X_{-1}|]+|\alpha|}{\beta-\alpha}$، پس:

$$P[X_* < \alpha < \beta < X^*] = 0.$$

حال مجددا داریم

$$[X_* < X^*] = \cup_{\alpha,\beta\in Q}[X_* < \alpha < \beta < X^*]$$

$$\Longrightarrow P[X_* < X^*] = P\left[\cup_{\alpha,\beta\in Q}[X_* < \alpha < \beta < X^*]\right] \leq \sum_{\alpha,\beta\in Q} P[X_* < \alpha < \beta < X^*] = 0$$

پس نمی توان رابطه $X_* < X^*$ را پذیرفت و در نتیجه $P[X_* = X^*] = 1$ و بنابراین دنباله $X_{-n}, \cdots, X_{-2}, X_{-1}$ به مقدار یک متغیر تصادفی مانند X میل می کند و بنابراین حد این دنباله وجود دارد و چون تعداد فراتقاطع ها متناهی است، پس X_* نمی تواند از مقدار α کوچکتر و X^* از مقدار β بزرگتر باشد. از آنجا که در مارتینگل ها داشتیم، $X_n = E[X_{n+k}||\mathcal{F}_n]$، در اینجا نیز برای مارتینگل های وارون داریم: $X_{-n} = E[X_{-n+k}||\mathcal{F}_{-n}]$ و با قرار دادن $k=n-1$ داریم: $X_{-n} = E[X_{-1}||\mathcal{F}_{-n}]$

حال تحت انتگرال پذیری $X_{-1}, X_{-2}, \cdots$ و بر اساس لم۱ $X_{-n} = E[X_{-1}||\mathcal{F}_{-n}]$ به طور یکنواخت انتگرال پذیر است و در نتیجه، $X_{-n} \to X$ ب که X انتگرال پذیر است و $\int X_{-n} \to \int X$. بنابراین : $E[X_{-n}] \to E[X]$. اما از طرفی داریم : $E[X_{-1}] = E[X_{-2}] = \cdots = E[X_{-n}]$ و چون تمامی امیدهای ریاضی باهم برابر هستند پس در نتیجه، حدشان هم با خودشان برابر است یعنی

$$E[X_{-n}] = E[X].$$

تذکر: اگر $\mathcal{F}_n$ دنباله ای نزولی از σ- میدان ها باشد که $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots$، آنگاه $\cap_{n=1}^{\infty} \mathcal{F}_n = \mathcal{F}_0$ یک سیگما میدان است و می توان نوشت $\mathcal{F}_n \downarrow \mathcal{F}_0$.

قضیه ۱۵-۸: اگر $\mathcal{F}_n \downarrow \mathcal{F}_0$ و Z یک متغیر تصادفی انتگرال پذیر باشد، آنگاه با احتمال ۱:

$$E[Z||\mathcal{F}_n] \to E[Z||\mathcal{F}_0]$$

اثبات: اگر $X_{-n} = E[Z||\mathcal{F}_n]$ آنگاه $\cdots, X_{-2}, X_{-1}$ نسبت به $\cdots, \mathcal{F}_2, \mathcal{F}_1$ یک مارتینگل است. با استفاده از قضیه ۱۳-۷، $E[Z||\mathcal{F}_n]$ *برای* n بزرگ به سمت متغیر تصادفی انتگرال پذیر X میل می کند. کافی است نشان دهیم $X = E[Z||\mathcal{F}_0]$. می دانیم اگر Z یک متغیر تصادفی انتگرال پذیر و $\cdots, \mathcal{F}_2, \mathcal{F}_1$ ها دلخواه باشند، بر اساس لم۱۵-۱، $E[Z||\mathcal{F}_n]$ به طور یکنواخت انتگرال پذیر است و چون X برابر با حد $E[Z||\mathcal{F}_n]$ برای $n \geq k$ و k ثابت می باشد، در نتیجه $E[Z||\mathcal{F}_n]$ نسبت به $\mathcal{F}_n$ اندازه پذیر است. بنابراین نسبت به $\mathcal{F}_k$ هم اندازه پذیر و X هم نسبت به $\mathcal{F}_k$ اندازه پذیر است. از طرفی چون X نسبت به $\mathcal{F}_k$ اندازه پذیر است و k دلخواه است، پس می توانیم k را بزرگ اختیار کنیم. وقتی k زیاد می شود، n نیز افزایش می یابد و در واقع، اجبار n به سمت بی نهایت میل می کند و عمل حدگیری صورت می پذیرد. هرچه k بزرگتر شود، X نسبت به $\mathcal{F}_k$ های کوچکتری اندازه پذیر می‌شود و در نتیجه وقتی k به اندازه کافی بزرگ شود، X نسبت به کوچکترین $\mathcal{F}_k$ که همان $\mathcal{F}_0$ است اندازه پذیر می شود.

با استفاده از انتگرال پذیری یکنواخت $E[Z||\mathcal{F}_n]$ و قضیه همگرایی یکنوا، $E[Z||\mathcal{F}_n] \to X$ و در نتیجه $\int E[Z||\mathcal{F}_n] \to \int X$.

بنابراین اگر $A \in \mathcal{F}_0$ باشد، آنگاه:

$$\begin{aligned}\int_A X\,dp &= \lim_n \int_A E[Z||\mathcal{F}_n]\,dp\\ &= \lim_n \int_A E[E[Z||\mathcal{F}_n]||\mathcal{F}_0]\,dp = \lim_n \int_A E[Z||\mathcal{F}_0]\,dp\\ &= \int_A E[Z||\mathcal{F}_0]\,dp\end{aligned}$$

و برابری آخر با توجه به استقلال جمله $\lim_n \int_A E[Z||\mathcal{F}_0]\,dp$ از مقدار n صورت گرفته است. در نتیجه:

$$\int_A X\,dp = \int_A E[Z||\mathcal{F}_0]\,dp \quad , \quad A \in \mathcal{F}_0$$

بنابراین X یک ورژن از امید شرطی $E[Z||\mathcal{F}_0]$ می باشد.

توجه: قضایای ۱۵-۶ و ۱۵-۸ در یک امتدادند اما اختلاف اساسی بین قضایای ۱۵-۵ و ۱۵-۷ وجود دارد. در قضیه ۱۵-۷، مارتینگل های وارون یک متغیر نهایی یعنی X_{-1} دارند و بنابراین در اثبات همگرایی، نیازی به فرض کراندار بودن $E[X_n]$ نمی باشد. به عبارت دیگر، اثبات قضیه ۱۵-۷ که X در آن انتگرال پذیر است برای زیرمارتینگل ها مناسب نیست.

در کاربرد قضیه ۱۵-۸، معمولا $\mathcal{F}_n$ را $\sigma\{Y_n, Y_{n+1}, \cdots\}$ می گیرند که در آن $Y_1, Y_2, \cdots$ متغیرهای تصادفی هستند. در قضیه ۱۵-۶، $\mathcal{F}_n$ اغلب $\sigma\{Y_1, \cdots, Y_n\}$ است. در این حالت $\mathcal{F}_\infty = \sigma\{Y_1, Y_2, \cdots\}$ در حقیقت بزرگتر از $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ است.

کاربردها: مشتقات

قضیه ۱۵-۹: فرض کنید $(\Omega, \mathcal{F}, P)$ یک فضای احتمال و ν یک اندازه متناهی روی $\mathcal{F}$ باشد و $\mathcal{F}_n \uparrow \mathcal{F}_\infty \subset \mathcal{F}$. فرض کنید P اندازه غالب ν روی $\mathcal{F}_n$ و X_n مشتق رادون نیکودیم متناظر با آن باشد. آنگاه با احتمال ۱، $X_n \longrightarrow X$ به طوری که X یک متغیر تصادفی انتگرال پذیر است.

i) اگر P بر ν روی $\mathcal{F}_\infty$ غالب باشد، آنگاه، X مشتق رادون نیکودیم متناظر با آن است.

ii) اگر P و ν اندازه های منفرد روی $\mathcal{F}_\infty$ باشند، آنگاه با احتمال برابر با ۱، $X = 0$.

اثبات: قبلا نشان دادیم که X_n به عنوان چگالی رادون نیکودیم روی فضای $\mathcal{F}_n$ یک مارتینگل تشکیل می‌دهد. در نتیجه باید نسبت به $\mathcal{F}_n$ اندازه پذیر باشد و همچنین داریم: $\int_A \mathrm{X_n}\, \mathrm{dp} = \nu(\mathrm{A})$ به طوری که $A \in \mathcal{F}_n$. از آنجا که $\mathrm{X_n}$ یک چگالی است پس مقادیر نامنفی اختیار می کند و بنابراین:

$$E[|X_n|] = E[X_n] = \nu(\Omega) < \infty \quad , \Omega \in \mathcal{F}_n$$

و نامساوی آخر بنا بر متناهی بودن اندازه ν برقرار است. بنابراین $E[|\mathrm{X_n}|] = \nu(\Omega)$ و $SupE[|X_n|] < \infty$ و در نتیجه $\mathrm{X_n}$ به سمت حد خود یعنی X میل می کند به طوری که X نسبت به $\mathcal{F}_\infty$ انتگرال پذیر است، چون حد آن در $\mathcal{F}_\infty$ تعریف شده و امید ریاضی قدرمطلق آن متناهی است. حال در ادامه اثبات برای بخش اول قضیه، می دانیم که P بر ν روی $\mathcal{F}_\infty$ غالب است. فرض کنید Z مشتق رادون نیکودایم ν روی این فضا باشد که برای $A \in \mathcal{F}_\infty$ داریم: $\int_A Z\, dp = \nu(A)$. اگر A را از $\mathcal{F}_\infty$ طوری انتخاب کنیم که به $\mathcal{F}_n$ نیز تعلق داشته باشد، آنگاه برای این پیشامد A داریم: $\int_A Z\, dp = \int_A X_n\, dp$ و بنابراین $\mathrm{X_n}$ ورژنی از امید شرطی $E[Z||\mathcal{F}_n]$ می باشد. یعنی:

$\mathrm{X_n} = E[Z||\mathcal{F}_n]$. با توجه به اینکه Z چگالی رادون نیکودایم است پس در نتیجه با استناد به قضیه ۱۴-۶، انتگرال پذیر است و بنابه فرض مسئله هم می دانیم که $\mathcal{F}_n \uparrow \mathcal{F}_\infty$. در نتیجه:

$$E[Z||\mathcal{F}_n] \longrightarrow E[Z||\mathcal{F}_\infty]$$

که بنابر اندازه پذیری Z نسبت به $\mathcal{F}_\infty$ و برابری $\mathrm{X_n} = E[Z||\mathcal{F}_n]$ داریم:

$$\mathrm{X_n} \longrightarrow Z$$

که استفاده از یکتایی حد $\mathrm{X_n} \longrightarrow X$ نشان می دهد X با احتمال ۱ همان Z، مشتق رادون نیکودایم است.

حال برای اثبات بخش دوم قضیه، می دانیم که P و ν اندازه های منفرد روی $\mathcal{F}_\infty$ هستند، بنابراین یک مجموعه مانند S از $\mathcal{F}_\infty$ وجود دارد که $\nu(S) = 0$ و $\mathrm{P}(S) = 1$. بنابر لم فاتو داریم::

$$\int_A X\, dp = \int_A lim X_n\, dp = \int_A liminf X_n\, dp \leq liminf \int_A X_n\, dp$$

و از طرفی اگر $A \in \mathcal{F}_k$ باشد آنگاه برای $n = k, k+1, \dots$ داریم: $\int_A X_n \, dp = \nu(A)$ و در نتیجه $liminf \int_A X_n \, dp = liminf \nu(A) = \nu(A)$. بنابراین :

$$\int_A X \, dp \le \nu(A) \ , \ A \in \cup_{k=1}^{\infty} \mathcal{F}_k$$

اکنون با استناد به قضیه کلاس یکنوا، این نامساوی برای هر $A \in \mathcal{F}_\infty = \sigma(\cup_{k=1}^{\infty} \mathcal{F}_k)$ برقرار است. بنابراین:

$$\int X \, dp = \int_S X \, dp \le \nu(S) = 0 \ \ , \ \ S \in \mathcal{F}_\infty$$

$$\Longrightarrow \int X \, dp = 0$$

$$\Longrightarrow X = 0 \ \ a.e.$$

قضیه کلاس یکنوا: اگر $\mathcal{F}_0$ یک میدان باشد و M یک کلاس یکنوا باشد(کلاس M از زیرمجموعه های Ω یکنواست اگر تحت اشتراک و اجتماع یکنوا بسته باشد)، آنگاه اگر $\mathcal{F}_0$ دارای خاصیت کلاس M باشد، $\sigma(\mathcal{F}_0)$ هم دارای همان خاصیت است.

در قضیه ۱۵-۹، کلاس ما شامل $\mathcal{F}_k$ هایی است که یک کلاس یکنوا هستند چون اجتماع و اشتراک $\mathcal{F}_k$ ها نیز به همین کلاس تعلق دارند و تمامی آنها دارای ویژگی $\int_A X \, dp \le \nu(A)$ می باشند.

۱۵–۸ نسبت درستنمایی

فرض کنید که $\nu = Q$ یک اندازه احتمال باشد و $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$ باشد که در آن $Y_1, \dots, Y_n$ متغیرهای تصادفی باشند. مشتق رادون نیکودیم یا نسبت درستنمایی X_n به شکل $X_n = \frac{q_n(Y_1,\dots,Y_n)}{p_n(Y_1,\dots,Y_n)}$ است که در آن q_n و p_n روی $\mathfrak{R}^n$ تعریف شده اند.با استفاده از قضیه ۱۴-۹، X_n به X همگراست و X نسبت به $\mathcal{F}_\infty = \sigma(Y_1, Y_2, \dots)$ انتگرال پذیر و اندازه پذیر است. اگر Y_n ها تحت P و Q مستقل باشند و اگر چگالی ها متفاوت باشد، در این صورت P و Q روی $\sigma(Y_1, Y_2, \dots)$ منفرد می باشند و در این حالت X=0 است و X_n به صفر میل می کند. توجیه آماری ما در این حالت بدین صورت است که هرچه X_n کوچکتر شود، P به عنوان توضیح و تفسیر مشاهدات $Y_1, \dots, Y_n$ نسبت به Q ترجیح داده می شود و زمانی که X_n با احتمال ۱ به سمت صفر می رود، در واقع P اندازه احتمال مربوط به Y_n هاست که آنها را کنترل و نقش Q حذف می شود.

حال ممکن است چنین به نظر بیاید که فرایند در زمان دلخواه، مثلا در مرحله n ام متوقف شود. در این صورت اگر τ زمان توقف آن باشد، دیگر X_τ یک نسبت درستنمایی نباشد و X_τ دیگر نتواند مشتق رادون نیکودایم با خواص X_n باشد. نشان می دهیم که چنین تصوری برقرار نیست و X_τ همچنان مشتق رادون نیکودایم باقی می ماند. پس فرض کنید که P اندازه مغلوب بر Q روی هر یک از $\mathcal{F}_n$ ها باشد ولی نسبت درستنمایی X_n شکل خاصی نداشته باشد.

فرض کنید τ یک متغیر تصادفی با مقادیر صحیح مثبت باشد و نشان دهنده زمان توقف فرایند باشد. پس می توان فرض کرد که $[\tau = n] \in \mathcal{F}_n$ که خاصیت زمان توقف بودن τ است و شرطی است که مانع از بیش برآوردی در مراحل بعدی فرایند نسبت به حال می شود. پس تصادفی بودن آن را تضمین می کند و ما در حقیقت با توجه به اطلاعاتی که در مرحله n ام فرایند داریم زمان توقف را تصادفا انتخاب می کنیم. فرض کنید $\mathcal{F}_\tau$ کلاس مجموعه هایی مانند G باشد که در آن $\mathcal{F}_\tau = \{G | G \cap [\tau = n] \in \mathcal{F}_n \quad , \mathrm{n} \geq 1\}$.

در واقع $\mathcal{F}_\tau$ نشان دهنده اطلاعاتی است که فرایند در زمان توقف دارد. دانستن $(\mathrm{Y}_1(\omega), \mathrm{Y}_2(\omega), \dots, \mathrm{Y}_{\tau(\omega)}(\omega))$ معادل دانستن این است که هر G که در میدان $\mathcal{F}_\tau$ قرار دارد، آیا ω را در خود دارد یا خیر. در حقیقت $\mathcal{F}_\tau$ بوسیله مجموعه های:

$$[\tau = n, (\mathrm{Y}_1, \dots, \mathrm{Y}_\mathrm{n}) \in \mathrm{H}] \;\; , \;\; \mathrm{n} \geq 1 \;\; , \mathrm{H} \in \mathfrak{R}^n$$

تولید می شود. حال با این فرض نشان می دهیم که X_τ یک نسبت درستنمایی(مشتق رادون نیکودایم) برای Q نسبت به P روی $\mathcal{F}_\tau$ می باشد. اگر $\mathrm{H} \in \mathfrak{R}^1$ آنگاه $\left[X_\tau \in H\right] = \cup_{n=1}^{\infty}([\tau = n] \cap [X_n \in H])$. با کمی توجه در می یابیم که چون X_n نسبت به $\mathcal{F}_n$ اندازه پذیر است پس $[X_n \in H] \in \mathcal{F}_n$ و از طرفی $\mathcal{F}_n = \sigma(\mathrm{Y}_1, \dots, \mathrm{Y}_\mathrm{n})$ پس در نتیجه X_n به شکل $(\mathrm{Y}_1, \dots, \mathrm{Y}_\mathrm{n})$ می باشد و گفتیم که مولد $\mathcal{F}_\tau$ همان $[\tau = n, (\mathrm{Y}_1, \dots, \mathrm{Y}_\mathrm{n}) \in \mathrm{H}]$ ها هستند پس اجتماع مولدها هم در $\mathcal{F}_\tau$ خواهد بود. پس: $\left[X_\tau \in H\right] \in \mathcal{F}_\tau$. بنابراین X_τ نسبت به $\mathcal{F}_\tau$ اندازه پذیر است. علاوه بر این با استفاده از این مطلب که X_n مشتق رادون نیکودایم Q است، داریم:

$$\int_G X_\tau \, dp = \sum_{n=1}^{\infty} \int_{G \cap [\tau = n]} X_n \, dp = \sum_{n=1}^{\infty} Q(G \cap [\tau = n]) = Q(G)$$

بنابراین:

$$\int_G X_\tau \, dp = Q(G) \;\; , \quad G \in \mathcal{F}_\tau$$

$$\Longrightarrow \; X_\tau = \frac{dQ}{dP}$$

و در نتیجه، X_τ مشتق رادون نیکودیم Q روی $\mathcal{F}_\tau$ می باشد.

به عنوان یکی از کاربردهایی قضایای همگرایی در مارتینگل ها می توان به قضیه دوفینیتی اشاره کرد.

قضیه دوفینیتی۱۵-۱۰: اگر متغیرهای تصادفی $X_1, \mathrm{X}_2, \dots$ تبادل پذیر باشند و مقادیر صفر و یک را اختیار کنند، آنگاه یک متغیر تصادفی θ وجود دارد که:

$$P[X_1 = u_1, \dots, X_n = u_n || \theta] = \theta^{\mathrm{s}} (1 - \theta)^{\mathrm{n-s}}$$

به طوری که $s = u_1 + u_2 + \cdots + u_n$ و با استفاده از انتگرال گیری بر روی تابع چگالی یکنواخت θ از طرفین معادله، داریم:

$$P[X_1 = u_1, \dots, X_n = u_n] = \mathrm{E}[\theta^{s}(1-\theta)^{n-s}]$$

اثبات: فرض کنید $S_{\mathrm{m}} = X_1 + X_2 + \cdots + X_m$. آنگاه برای $t \leq m$ داریم:

$$P[S_{\mathrm{m}} = t] = \sum_{u_1+u_2+\cdots+u_m=t} P[X_1 = u_1, \dots, X_m = u_m]$$

که با توجه به تبادل پذیری متغیرهای تصادفی تعداد $\binom{\mathrm{m}}{t}$ جمله در سمت راست معادله قرار دارد و در نتیجه:

$$P[X_1 = u_1, \dots, X_m = u_m | S_{\mathrm{m}} = t] = \binom{\mathrm{m}}{t}^{-1}$$

همین نتایج برای n ثابت و $s \leq n \leq m$ و $u_1 + u_2 + \cdots + u_n = s \leq t \leq m$ برقرار است به طوری که برای باقی جملات $u_{n+1} + \cdots + u_m$ مجموع برابر با t-s خواهد بود و همچنین:

$$P[X_1 = u_1, \dots, X_n = u_n | S_{\mathrm{m}} = t] = \frac{\binom{\mathrm{m-n}}{t-s}}{\binom{\mathrm{m}}{t}}$$

$$= \frac{(t)_s (m-t)_{n-s}}{(m)_n} = f_{n,s,m}(\frac{t}{m})$$

به طوری که،

$$f_{n,s,m}(x) = \prod_{i=0}^{s-1}(x - \frac{i}{m}) \prod_{i=0}^{n-s-1}(1 - x - \frac{i}{m}) / \prod_{i=0}^{n-1}(1 - \frac{i}{m})$$

اگر به قسمت شرط احتمال فوق، شرط های بیشتری بر پایه

$$S_{\mathrm{m+1}} = t_1 ، S_{\mathrm{m+2}} = t_2 \text{ و } \dots S_{\mathrm{m+j}} = t_j$$

را نیز اضافه کنیم، بازهم معادله برقرار است و خواهیم داشت:

$$P\big[X_1 = u_1, \dots, X_n = u_n \big| S_{\mathrm{m}}, \dots, S_{\mathrm{m+j}}\big] = f_{n,s,m}(\frac{S_{\mathrm{m}}}{m})$$

حال فرض کنید، $\mathcal{g}_m = \sigma(S_{\mathrm{m}}, S_{\mathrm{m+1}}, \dots)$ و $\mathcal{g} = \bigcap_m \mathcal{g}_m$ باشند. حال n و $u_1, \dots, u_n$، را ثابت فرض کنید به طوری که $u_1 + u_2 + \cdots + u_n = s$. اگر $j \to \infty$ آنگاه با استفاده از قضیه ۱۳-۶:

$$P[X_1 = u_1, \dots, X_n = u_n | \mathcal{g}_m] = f_{n,s,m}(\frac{S_{\mathrm{m}}}{m})$$

و اگر $m \to \infty$ آنگاه با استفاده از قضیه ۱۲-۸:

$$P[X_1 = u_1, \dots, X_n = u_n | \mathcal{g}]_\omega = \lim_{\mathrm{m}} f_{n,s,m}(\frac{S_{\mathrm{m}}(\omega)}{m})$$

و این رابطه برای ω خارج از مجموعه ای با اندازه صفر برقرار است. حال فرض کنید دو نقطه حدی متفاوت برای $\frac{S_{\mathrm{m}}(\omega)}{m}$ وجود دارد، آنگاه با توجه به اینکه فاصله هر $\frac{S_{\mathrm{m}}(\omega)}{m}$ از مقدار بعدی آن کوچکتر از $\frac{2}{m}$ است در نتیجه نقاط حدی آن باید یک بازه ناتباهیده را پر کنند. اما $\lim_k x_{nk} = x$ ایجاب می کند که:

$$\lim_k f_{n,s,m_k}(x_{m_k}) = x^s(1-x)^{n-s}$$

که در نتیجه باید $x^s(1-x)^{n-s}$ بر روی بازه مزبور ثابت باشد و این امکان پذیر نیست. پس در نتیجه $\frac{S_{\mathrm{m}}(\omega)}{m}$ باید به یک حدی مانند $\theta(\omega)$ روی این بازه میل کند و این نشان می دهد که با احتمال ۱:

$$P[X_1 = u_1, \dots, X_n = u_n | \mathcal{g}] = \lim_m f_{n,s,m}(\frac{S_m(\omega)}{m}) = \theta^s(1-\theta)^{n-s}$$

و با گرفتن امید ریاضی از طرفین بر روی $\sigma(\theta)$ اثبات کامل است. برای اثبات این قضیه در حالت کلی تر صفحه۲۲۸ از کتاب Rick Durrett(2013) را ببینید .

۱۵–۹ قضیه حد مرکزی

فرض کنید $\mathrm{X}_1, X_2, \cdots$ نسبت به $\mathcal{F}_1, \mathcal{F}_2, \cdots$ مارتینگل باشد و متغیر تصادفی کراندار Y_{n} را به صورت

$Y_{\mathrm{n}} = X_{\mathrm{n}} - X_{\mathrm{n-1}}$ تعریف کنید. بنابراین، $\mathrm{E}[\mathrm{Y_n}||\mathcal{F}_{\mathrm{n-1}}] = 0$ و اگر $\sigma_{\mathrm{n}}^2 = \mathrm{E}[\mathrm{Y}_{\mathrm{n}}^2||\mathcal{F}_{\mathrm{n-1}}]$ باشد و زمان توقف به صورت $\nu_t = \min[\mathrm{n}\colon \sum_{\mathrm{k=1}}^{\mathrm{n}} \sigma_{\mathrm{k}}^2 \ge \mathrm{t}]$ تعریف شود، آنگاه برای t به قدر کافی بزرگ، $X_{\nu_t}/\sqrt{t}$ دارای توزیع تقریبی نرمال می باشد.

به بیان دیگر، می توان این رابطه را به صورت قضیه ۱۵-۱۲ مطرح کرد. اما به منظور اثبات قضیه ۱۵-۱۲ باید از قضیه ۱۵-۱۱ استفاده نمود که صورت این قضیه بدون اثبات ارائه شده است.

قضیه ۱۵-۱۱: فرض کنید برای هر n ثابت، $\mathrm{X}_{n1}, X_{n2}, \cdots$ نسبت به $\mathcal{F}_{n1}, \mathcal{F}_{n2}, \cdots$ مارتینگل باشد و داشته باشیم $Y_{\mathrm{nk}} = X_{\mathrm{nk}} - X_{\mathrm{n,k-1}}$ به طوری که $\mathrm{E}[\mathrm{Y_{nk}}||\mathcal{F}_{\mathrm{n,k-1}}] = 0$ و $\sigma_{nk}^2 = \mathrm{E}[\mathrm{Y}_{\mathrm{nk}}^2||\mathcal{F}_{\mathrm{n,k-1}}]$.

اگر $\sum_n \sigma_{nk}^2 \xrightarrow{P} \sigma^2$ به طوری که σ یک مقدار ثابت و مثبت باشد و همچنین اگر برای هر ε، $\sum_{k=1}^{n} E[Y_{nk}^2 I_{[|Y_{nk}|\geq\varepsilon]}] \to 0$ آنگاه: $\sum_{k=1}^{\infty} Y_{nk} \Longrightarrow \sigma N$.

قضیه ۱۵-۱۲: فرض کنید $Y_n = X_n - X_{n-1}$ متغیرهای تصادفی به طور یکنواخت کراندار باشند و در رابطه $E[Y_n||\mathcal{F}_{n-1}] = 0$ صدق کنند و همچنین $\sum_n \sigma_n^2 = \infty$ باشد، آنگاه $X_{\nu_t}/\sqrt{t} \Longrightarrow N$.

اثبات: فرض کنید که $Y_{nk} = I_{[\nu_n\geq k]} Y_k/\sqrt{n}$ و $\mathcal{F}_{nk} = \mathcal{F}_k$ باشد. از

$[\nu_n \geq k] = [\sum_{j=1}^{k-1} \sigma_j^2 < n] \in \mathcal{F}_{k-2}$ داریم $E[Y_{nk}||\mathcal{F}_{n,k-1}] = 0$ و

$\sigma_{nk}^2 = E[Y_{nk}^2|\mathcal{F}_{n,k-1}] = I_{[\nu_n\geq k]}\, \sigma_k^2/n$. اگر k کران $|Y_k|$ باشد، آنگاه :

$$1 \leq \sum_{k=1}^{\infty} \sigma_{nk}^2 = \frac{\sum_{k=1}^{\nu_n} \sigma_k^2}{n} \leq 1 + K^2/n$$

بنابراین با قرار دادن $\sigma = 1$ همگرایی $\sum_n \sigma_{nk}^2 \xrightarrow{P} \sigma^2$ حاصل می شود. برای n به اندازه کافی بزرگ به طوری که در رابطه $k/\sqrt{n} < \varepsilon$ صدق کند، مجموع $\sum_{k=1}^{n} E[Y_{nk}^2 I_{[|Y_{nk}|\geq\varepsilon]}]$ به سمت صفر میل می‌کند. بنابراین شرایط قضیه ۱۴-۱۰ صادق است و در نتیجه :

$$\sum_{k=1}^{\nu_n} Y_k/\sqrt{n} = \sum_{k=1}^{\infty} Y_{nk} \Longrightarrow N$$

۱۵–۱۰مسائل

۱) هر زیر مارتینگل X_n می تواند به فرم یکتای $X_n = M_n + A_n$ تجزیه شود طوری که M_n یک مارتینگل و A_n یک دنباله افزایشی پیش بینی پذیر با $A_0 = 0$ است. (راهنمایی: تجزیه دابس)

۲) فرض کنید متغیرهای تصادفی $X_1, \dots, X_n$ مستقل با میانگین صفر و $Var(X_1) = \sigma^2 < \infty$ و $M_n = S_n^2 - n\sigma^2$ باشد، به طوری که $S_n = \sum_{i=1}^{n} X_i$. نشان دهید که M_n مارتینگل است.

۳) فرض کنید متغیرهای تصادفی نامنفی $X_1, \dots, X_n$ مستقل و هم توزیع با میانگین ۱ باشند به طوری که $X_0 = 1$. اگر $M_n = \prod_{i=1}^{n} X_i$ باشد نشان دهید که M_n مارتینگل است.

۴) فرض کنید $(\Omega, \mathcal{F}, P)$ فضای احتمال باشد. فیلتر $\mathbb{F}$ را به صورت $\mathbb{F} = \{\mathcal{F}_n : n = 0,1,\dots\}$ و متغیر تصادفی Y را به صورت $\mathrm{Y}: \Omega \to \mathbb{R}$ تعریف می کنیم. فرایند تصادفی X_n را به صورت $X_n = E[Y|\mathcal{F}_n]$ در نظر بگیرید. نشان دهید که X_n مارتینگل است.

۵) فرض کنید $(\Omega, \mathcal{F}, P)$ فضای احتمال باشد. فیلتر $\mathbb{F}$ را به صورت $\mathbb{F} = \{\mathcal{F}_n : n = 0,1,\dots\}$ تعریف می کنیم. اگر S یک فرایند قدم زدن تصادفی روی Z و T اولین زمانی باشد که $S_{\mathrm{n}} = 4$ است، نشان دهید T یک زمان توقف است.

۶) نشان دهید متغیر تصادفی $T: \Omega \to \mathbb{N} \cup \{+\infty\}$ یک زمان توقف است اگر و تنها اگر

$$\{T = n\} \epsilon \mathcal{F}_n \text{ برای هر } n = 0,1,.$$

فصل شانزدهم

فرایندهای تصادفی: قضایای وجودی کلموگروف

۱۶–۱ مقدمه

یک فرآیند تصادفی (Stochastic Process) مجموعه $[X_t: t \in T]$ از متغیرهای تصادفی است که روی فضای احتمال $(\Omega, \mathcal{F}, \mathcal{P})$ تعریف می‌شوند. مثلاً مارتینگل‌ها نمونه‌ای از فرآیندهای تصادفی هستند که در آنها $T = \{1,2,3, \dots\}$ می‌باشد. فرآیند پواسن نیز نمونه دیگری از فرآیندهای تصادفی است: $[N_t: t \geq 0]$ که در آن $T = [0, \infty)$. برای این فرآیندها نقاط T نشان دهنده زمان هستند. در اغلب حالت‌ها، T مجموعه اعداد صحیح است و زمان، گسسته می‌باشد یا اینکه T یک بازه از R است و زمان، پیوسته می‌باشد. در این فصل بطور کلی T را دلخواه در نظر می‌گیریم.

۱۶–۲ توزیع‌های متناهی‌البعد

یک فرآیند معمولا برحسب توزیع‌هایی که در فضای اقلیدسی القاء می‌کند، توصیف می شود. برای هر k تایی $(t_1, \dots, t_k)$ از عناصر مجزای T، بردار تصادفی $(X_{t_1}, \dots, X_{t_k})$ روی R^k توزیعی مانند $\mu_{t_1 \dots t_k}$ دارد:

$$\mu_{t_1 \dots t_k}(H) = P\left[\left(X_{t_1}, \dots, X_{t_k}\right) \in H\right], \qquad H \in \mathcal{R}^k \qquad (16.1)$$

این اندازه‌های احتمال $\mu_{t_1 \dots t_k}$ توزیع‌های متناهی‌البعد (finite-dimensional distributions) از فرآیند تصادفی $[X_t: t \in T]$ می‌باشند.

مجموعه توزیع‌های متناهی‌البعد بطور کامل خصوصیات فرآیند را مشخص نمی‌کنند. با این وجود اولین مرحله در تئوری عمومی فرآیند تصادفی، ساختن فرآیندهایی با مجموعه‌ای از توزیع‌های متناهی‌البعد است.

حال رابطه (16.1) را در نظر بگیرید. این رابطه دو خاصیت سازگار از سیستم $\mu_{t_1 \dots t_k}$ را نتیجه می‌دهد:

اول: فرض کنید که H در (16.1) به شکل $H = H_1 \times \dots \times H_k$ ،$(H_i \in \mathcal{R}^1)$، باشد و جایگشت Π از $(1,2, \dots, k)$ را نیز در نظر بگیرید. از آنجایی که $\left[\left(X_{t_1}, \dots, X_{t_k}\right) \in (H_1 \times \dots \times H_k)\right]$ و $\left[\left(X_{t_{\Pi_1}}, \dots, X_{t_{\Pi_k}}\right) \in \left(H_{\Pi_1} \times \dots \times H_{\Pi_k}\right)\right]$ پیشامدهای یکسانی هستند خواهیم داشت:

$$P\left[\left(X_{t_1},\dots,X_{t_k}\right)\in\left(H_1\times\dots\times H_k\right)\right]=P\left[\left(X_{t_{\Pi_1}},\dots,X_{t_{\Pi_k}}\right)\in\left(H_{\Pi_1}\times\dots\times H_{\Pi_k}\right)\right]$$

$$\Rightarrow \mu_{t_1\dots t_k}\left(H_1\times\dots\times H_k\right)=\mu_{t_{\Pi_1}\dots t_{\Pi_k}}\left(H_{\Pi_1}\times\dots\times H_{\Pi_k}\right) \qquad (16.2)$$

(رابطه بالا براساس این خاصیت نوشته شده که $(X,Y)\in A\times B \iff (Y,X)\in B\times A$ و همانطور که از (16.2) مشاهده می‌شود، عبارات سمت چپ و راست، ورودی‌های یکسان ندارند. هدف ما در ادامه کار این خواهد بود که ورودی چپ و راست معادله را یکسان کنیم)

برای مثال، اگر $\mu_{s,t}=\nu\times\nu'$ باشد، لزوماً $\mu_{t,s}=\nu'\times\nu$ می‌شود.

دوم: شرط دوم سازگاری بصورت زیر است:

$$\mu_{t_1\dots t_{k-1}}\left(H_1\times\dots\times H_{k-1}\right)=\mu_{t_1\dots t_{k-1}t_k}\left(H_1\times\dots\times H_{k-1}\times R^1\right) \qquad (16.3)$$

(در اینجا نیز ورودی‌ها متفاوتند و مجدداً می‌خواهیم کاری کنیم که ورودی‌ها یکسان شوند)

این رابطه واضح است زیرا $\left(X_{t_1},\dots,X_{t_{k-1}}\right)$ در $H_1\times\dots\times H_{k-1}$ قرار دارد اگر و فقط اگر $\left(X_{t_1},\dots,X_{t_{k-1}},X_{t_k}\right)$ در $H_1\times\dots\times H_{k-1}\times R^1$ باشد. حال می‌خواهیم این دو شرط را به روش دیگری بیان کنیم. تعریف کنید:

$$\varphi_\Pi:\ R^k\longrightarrow R^k$$

به طوریکه:

$$\varphi_\Pi(x_1,\dots,x_k)=\left(x_{\Pi_1^{-1}},\dots,x_{\Pi_k^{-1}}\right)$$

φ_Π ، جایگشت Π را روی مؤلفه‌ها اعمال می‌کند، برای مثال اگر Π، x_3 را به اولین جایگاه بفرستد در این صورت $\Pi^{-1}1=3$ می‌شود $\left(x_1,x_2,\boxed{x_3},x_4\right)\xrightarrow{\Pi}\left(x_{\boxed{3}},\dots\right)$.

از آنجایی که $\varphi^{-1}{}_\Pi\left(H_1\times\dots\times H_k\right)=H_{\Pi_1}\times\dots\times H_{\Pi_k}$ ، از (16.2) نتیجه می‌شود که:

$$\mu_{t_{\Pi_1}\dots t_{\Pi_k}}\varphi^{-1}{}_\Pi(H)=\mu_{t_1\dots t_k}(H) \qquad for\ rectangles\ H$$

اما در این صورت:

$$\mu_{t_1\dots t_k}=\mu_{t_{\Pi_1}\dots t_{\Pi_k}}{\varphi_\Pi}^{-1} \qquad (16.4)$$

و در نتیجه ورودی ها یکسان شد.

بطور مشابه، اگر $\varphi\ :\ R^k \longrightarrow R^{k-1}$، تصویر زیر را ایجاد کند

$$\varphi(x_1, \dots, x_k) = (x_1, \dots, x_{k-1})$$

که آخرین عضو قطع شده است، در این صورت از (16.3) نتیجه می‌شود که:

$$\mu_{t_1 \dots t_{k-1}} = \mu_{t_1 \dots t_k} \varphi^{-1} \tag{16.5}$$

(پس ما تا اینجا دو عملگر معرفی کردیم. یکی φ_Π که جایگشت ایجاد می‌کند و دیگری φ که عمل قطع کردن عضو یا اعضای آخر را انجام می‌دهد. می‌خواهیم از اینجا به بعد دو نماد قبل را (جایگشت و قطع کردن) به یک نماد و یک عملگر تبدیل کنیم.)

روابط (16.4) و (16.5) بسط مشترکی دارند. فرض کنید که $(u_1, \dots, u_m)$ یک m تایی از عناصر مجزای T هستند و هر عضو از $(t_1, \dots, t_k)$ نیز یک عضو از $(u_1, \dots, u_m)$ است. همچنین فرض کنید $k \leq m$ باشد، در این صورت $(t_1, \dots, t_k)$ برش اولیه بعضی از جایگشت‌های $(u_1, \dots, u_m)$ است. یعنی یک جایگشت Π از $(1,2, \dots, m)$ وجود دارد به طوریکه

$$(u_{\Pi^{-1}1}, \dots, u_{\Pi^{-1}k}, \dots, u_{\Pi^{-1}m}) = (t_1, \dots, t_k, t_{k+1}, \dots, t_m)$$

جاییکه $t_{k+1}, \dots, t_m$ عناصری از $(u_1, \dots, u_m)$ هستند که در $(t_1, \dots, t_k)$ ظاهر نمی‌شوند و در واقع حذف می‌شوند. حال تعریف کنید: $\psi: R^m \longrightarrow R^k$ که در آن:

$$\psi(x_1, \dots, x_m) = (x_{\Pi^{-1}1}, \dots, x_{\Pi^{-1}k}) \tag{16.6}$$

که ψ، Π را روی مؤلفه ها اعمال می‌کند و سپس با یک برش، k تای اول آن‌ها را منعکس می‌کند. به عنوان مثال:

$$\psi(H_1 \times R \times H_2) \xrightarrow{\text{مرحله اول}} \begin{cases} H_1 \times H_2 \times R \\ H_2 \times H_1 \times R \\ R \times H_1 \times H_2 \\ \vdots \end{cases} \xrightarrow{\text{مرحله دوم}} \begin{cases} H_1 \times H_2 \\ H_2 \times H_1 \end{cases}$$

حال از آنجایی که $\psi\left(x_{u_1}, \dots, x_{u_m}\right) = \left(x_{t_1}, \dots, x_{t_k}\right)$، بنابراین،

$$\mu_{t_1 \dots t_k} = \mu_{u_1 \dots u_m} \psi^{-1} \tag{16.7}$$

این رابطه، روابط (16.4) و (16.5) را به عنوان حالت خاص در بر دارد. همچنین از آنجا که ψ شامل یک عمل جایگشت و به دنبال آن یک دنباله از تقطیع‌ها به شکل $(x_1, \dots, x_l) \to (x_1, \dots, x_{l-1})$ است پس خود یک نتیجه از (16.4) و (16.5) نیز می‌باشد.

مثال:

$$\mu_{1,2}(A \times B) = \mu_{1,2,3,4}\psi^{-1}(A \times B)$$
$$\Longrightarrow \mu_{1,2}(A \times B) = \mu_{1,2,3,4}(B \times R \times A \times R)$$

فضای حاصل‌ضربی

ساختار استاندارد فرآیندهای عمومی شامل فضاهای حاصل‌ضربی است. ابتدا فضای حاصل‌ضربی را معرفی می‌کنیم. فرض کنید T یک مجموعه دلخواه از اندیس‌ها باشد و فرض کنید R^T مجموعه همه توابع حقیقی روی T باشد (یعنی همه نگاشت‌ها از T به R). هدف این است که نشان دهیم دو مجموعه روبرو معادلند: $R^T = \{x | \, x : T \longrightarrow R\}$

- اگر $T = \{1,2, \dots, k\}$ باشد، یک تابع حقیقی روی T می‌تواند بوسیله k تایی $(x_1, \dots, x_k)$ از اعداد حقیقی شناخته یا معرفی شود و بنابراین R^T می‌تواند بوسیله فضای اقلیدسی k بعدی R^k معرفی شود.
- اگر $T = \{1,2, \dots\}$ باشد، یک تابع حقیقی روی T، دنباله $\{x_1, x_2 \, \dots\}$ از اعداد حقیقی است.
- اگر T یک بازه باشد، R^T شامل همه توابع حقیقی است که اگر چه غیر معمول و غیر قابل نوشتن است ولی همگی روی این بازه تعریف می‌شوند.

به عنوان مثال، $T = \{1,2\}$ را در نظر می‌گیریم. نشان می‌دهیم که:

$$R^{\{1,2\}} = \{x | \, x : \{1,2\} \longrightarrow R\} = R^2$$

که xها همه توابعی هستند که می‌توانند روی دو نقطه ۱ و ۲ اعمال شوند. مثلاً فرض کنید:

$$\begin{cases} 1 \overset{x}{\rightarrow} x(1) = 15 \\ 2 \overset{x}{\rightarrow} x(2) = 16 \end{cases} \Rightarrow \{1,2\} \overset{x}{\rightarrow} (15,16)$$

که می‌بینیم برد x حداکثر می‌تواند دو نقطه از R باشد.

بنابراین چون $R^{\{1,2\}}$ شامل مجموعه همه توابع حقیقی است، پس توابع بیشماری هستند که روی $\{1,2\}$ اعمال می‌شوند و به ما زوج مرتب می‌دهند. پس همه این زوج مرتب‌ها به ما R^2 را خواهند داد، زیرا تناظر یک به یک بین دو مجموعه برقرار است.

به هر حال مجموعه T هر چه که باشد، اعضای R^T را با x نشان خواهیم داد $(x \in R^T)$. مقادیر x در t را با $x(t)$ یا x_t نمایش می‌دهیم؛ بسته به اینکه آیا x را به عنوان یک تابع از t با دامنه T ببینیم یا به عنوان یک بردار با مؤلفه‌هایی که با اعضای t از T، اندیس‌گذاری شده‌اند.

حالت دنباله‌ای	$\longrightarrow \quad (x_1, x_2) \in R^2$
معمول نیست	$\longrightarrow \quad \big(x(1), x(2)\big) \in R^2$

همانطور که R^k می‌تواند به عنوان ضرب کارتزین k تا از Rها (در حقیقت یک کپی kتایی R) در نظر گرفته شود، R^T نیز می‌تواند به عنوان یک فضای حاصل‌ضرب در نظر گرفته شود (یک کپی kتایی از R).

حال برای هر t، نگاشت $Z_t : R^T \longrightarrow R^1$ را بصورت زیر تعریف کنید:

$$Z_t(x) = x(t) = x_t \tag{16.8}$$

به عنوان مثال:

$$Z_1 : R^2 \longrightarrow R$$

$$Z_1(x_1.x_2) = x_1$$

$$Z_2(x_1.x_2) = x_2$$

Z_tها را توابع مختصاتی (Coordinate functions) یا تصویرها (Projections) می‌نامیم. Z_tها در واقع متغیرهای تصادفی خواهند بود که روی فضای احتمالی که روی R^T ساخته می‌شود تعریف می‌شوند و به آنها متغیرهای مختصاتی (Coordinate Variables) گفته می‌شود. در ادامه بطور مکرر مقدار $Z_t(x)$ با $Z(t,x)$ نمایش داده خواهد شد. اگر x ثابت فرض شود، $Z(.,x)$ یک تابع حقیقی روی T خواهد بود و در واقع چیزی جز $x(.)$ نخواهد بود، یعنی خود x. اگر t ثابت فرض شود، $Z(t,.)$ یک تابع حقیقی روی R^T خواهد بود و معادل می‌شود با تابع Z_t که بوسیله رابطه (16.8) معرفی شد.

$$If\ x\ is\ fixed \quad \Rightarrow \quad Z(.,x) : T \longrightarrow R$$

$$t \longrightarrow Z(t,x) = x(t)$$

$$If\ t\ is\ fixed \quad \Rightarrow \quad Z(t,.) : R^T \longrightarrow R$$

$$Z_t(x) = x(t)$$

حال می‌خواهیم ایده سیگما میدان مجموعه‌های برل k بعدی را به فضای R^T تعمیم دهیم. فرض کنید $\mathcal{R}^T$، سیگما میدانی باشد که بوسیله همه توابع مختصاتی Z_t، $t \in T$، تولید می‌شود.

$$(t: fix) \quad \rightarrow \quad \mathcal{R}^T = \sigma[Z_t : t \in T]$$

این سیگما میدان بوسیله مجموعه‌هایی به فرم زیر تولید می‌شود:

$$[x \in R^T : Z_t(x) \in H] = [x \in R^T :\ x_t \in H] \quad for \quad t \in T, H \in \mathcal{R}^1$$

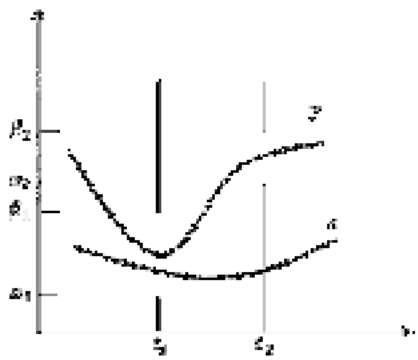
$$H = (3,5) \Rightarrow \ \{x \in \mathcal{R}^2 :\ Z_1(x) = x_1 \ \in (3,5)\} = (3,5) \times \mathcal{R}$$

اگر $T = \{1,2, \dots , k\}$ باشد آنگاه $\mathcal{R}^T$ بر $\mathcal{R}^k$ منطبق است. فرض کنید کلاس $\mathcal{R}_0^T$ شامل همه مجموعه‌های به فرم زیر باشد:

$$A = \left[x \in R^T :\ \left(Z_{t_1}(x), \dots , Z_{t_k}(x)\right) \in H\right] \qquad (16.9)$$

$$= \left[x \in R^T :\ \left(x_{t_1}, \dots , x_{t_k}\right) \in H\right]$$

به طوریکه k یک عدد صحیح است، $(t_1, \dots , t_k)$ یک kتایی از نقاط مجزای T است و $H \in \mathcal{R}^k$. مجموعه‌هایی به این فرم که اعضای $\mathcal{R}_0^T$ هستند، مجموعه‌های متناهی البعد یا سیلندرها (cylinders) نامیده می‌شوند (شکل ۱۶-۱). البته $\mathcal{R}_0^T$ مولد $\mathcal{R}^T$ است. خود $\mathcal{R}_0^T$ یک سیگما میدان نیست و منطبق بر $\mathcal{R}^T$ نمی‌باشد (مگر آنکه T متناهی باشد)، اما بحث زیر نشان می‌دهد که یک میدان است.



شکل ۱۶-۱: اگر T یک بازه باشد، سیلندر $[x \in R^T :\ \alpha_1 < x(t_1) < \beta_1, \alpha_2 < x(t_2) < \beta_2\ ]$ شامل توابعی است که از دو دریچه نشان داده شده، عبور کند. y در سیلندر قرار می‌گیرد اما Z درون سیلندر قرار نمی‌گیرد (البته نیازی نیست این توابع الزاماً پیوسته باشند).

ابتدا توجه کنید که اگر به عنوان مثال، $T = 2$ باشد، آنگاه:

$$A = \{[x \in R^2, x_1 \in H\ \ ]\ or\ [x \in R^2, x_2 \in H^0\ ]\}\,; \quad H, H^0 \in \mathcal{R}$$

حال نشان می‌دهیم که مجموعه بالا مولد سیگما میدان $\mathcal{R}^2$ است:

$$\sigma\, \overbrace{\{[x \in R^2, x_1 \in H]\ \ or\ [x \in R^2, x_2 \in H^0]\,;\ \ H, H^0 \in \mathcal{R}\}}^{u} = \mathcal{R}^2$$

عضوی از هر دو سمت می‌گیریم و نشان می‌دهیم به طرف دوم هم متعلق است:

$$[(x_1, x_2)\,,\ \ x_1 \in H] = H \times R\ \in\ \mathcal{R}^2 \ \ \Rightarrow generator\ of\ \sigma[\dots] \Rightarrow u \in \mathcal{R}^2$$

$$[a,b]\times[c,d]=\{(x_1,x_2)|x_1\in[a,b]\}\cap\{(x_1,x_2)|x_2\in[c,d]\}\ \in u$$
$$\Rightarrow \mathcal{R}^2\subset u\ \Rightarrow \mathcal{R}^2\subset\sigma(u)$$

متمم مجموعه A می‌شود:

$$R^T-A=\left[x\in R^T:\left(x_{t_1},\dots,x_{t_k}\right)\in R^k-H\right]$$

بنابراین $\mathcal{R}_0^T$ تحت عمل متمم‌گیری بسته است. حال فرض کنید که A بوسیله رابطه (16.9) داده شده باشد و B نیز بصورت زیر تعریف شود:

$$B=\left[x\in R^T:\left(x_{s_1},\dots,x_{s_k}\right)\in I\right] \tag{16.10}$$

به طوریکه $I\in\mathcal{R}^k$. به عنوان مثال:

$$A=[(x_1,\dots,x_{100})|(x_1,x_{56})\in[2,8]\times[11,21]]$$
$$B=\left[(x_1,\dots,x_{100})|(x_3,x_4,x_{56})\in[8,9]\times[23,46]\times[180,944]\right]$$

فرض کنید $(u_1,\dots,u_m)$ یک mتایی شامل همه t_αها و s_βها باشد. حال $(t_1,\dots,t_k)$ باید برش اولیه یک جایگشت از $(u_1,\dots,u_m)$ باشد و اگر ψ مانند (16.6) تعریف شده باشد و $H'=\psi^{-1}H$ باشد، آنگاه $H'\in\mathcal{R}^m$ و A بصورت زیر خواهد بود:

$$A=\left[x\in R^T:\left(x_{u_1},\dots,x_{u_m}\right)\in H'\right] \tag{16.11}$$
$$A=[(x_1,\dots,x_{100})|(x_1,x_3,x_4,x_{56})\in[2,8]\times R\times R\times[11,21]]$$

به طور مشابه B می‌تواند به فرم زیر نوشته شود:

$$B=\left[x\in R^T:\left(x_{u_1},\dots,x_{u_m}\right)\in I'\right],\qquad I'\in\mathcal{R}^m \tag{16.12}$$

و در این صورت:

$$A\cup B=\left[x\in R^T:\left(x_{u_1},\dots,x_{u_m}\right)\in H'\cup I'\right] \tag{16.13}$$

از آنجایی که $H'\cup I'$ متعلق به $\mathcal{R}^m$ است، $A\cup B$ یک سیلندر است. این بحث نشان می‌دهد که $\mathcal{R}_0^T$ یک میدان است بطوریکه $\mathcal{R}^T=\sigma(\mathcal{R}_0^T)$.

ساختن اندازه احتمال

ما $\mathcal{R}^T$ را طوری ساختیم که Z_tها نسبت به آن اندازه‌پذیر باشند. پس Z_tها توابعی اندازه پذیر روی فضای اندازه‌پذیر $(R^T,\mathcal{R}^T)$ هستند. اگر P یک اندازه احتمال روی $\mathcal{R}^T$ باشد، در این صورت، $[Z_t: t\in T]$ یک فرآیند تصادفی روی $(R^T,\mathcal{R}^T,P)$ است. فرآیندی که متغیرهای آن مختصات‌ها هستند.

نکته۱۶-۱- می‌توان فرآیند تصادفی را تعمیمی از متغیر تصادفی در نظر گرفت.

نکته۱۶-۲- $\mathcal{R}^{\mathrm{T}}$ بوسیله $\mathcal{R}_0^{\mathrm{T}}$ تولید می‌شود، البته ابتدا بوسیله مجموعه شامل $\mathrm{Z_t}$ها تولید می‌شود.

قضایای وجودی کلموگروف:

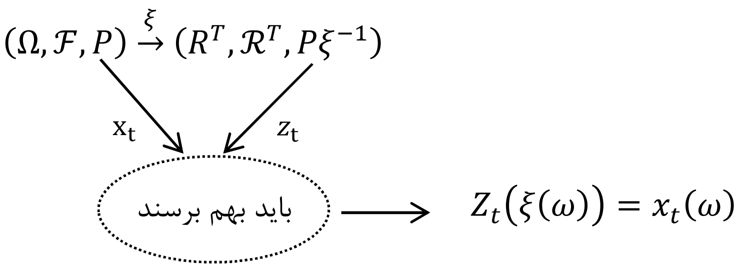
این قضیه به دو صورت بیان می‌شود:

قضیه ۱۶-۱: اگر $\mu_{t_1\dots t_k}$ یک سیستم از توزیع هایی باشد که در شرایط (16.2) و (16.3) صدق کنند، یک اندازه احتمال P در $\mathcal{R}^T$ وجود دارد، بطوریکه فرایند متغیرهای مختصاتی $[Z_t: t\in T]$ در $(R^T,\mathcal{R}^T,P)$ دارای مقدار $\mu_{t_1\dots t_k}$ به عنوان توزیع متناهی البعد است.

قضیه ۱۶-۲: اگر $\mu_{t_1\dots t_k}$ یک سیستم از توزیع هایی باشد که در شرایط (16.2) و (16.3) صدق کنند، یک فرآیند تصادفی $[x_t: t\in T]$ دارای مقدار $\mu_{t_1\dots t_k}$ به عنوان توزیع متناهی البعد در فضای احتمال $(\Omega,\mathcal{F},P)$ وجود دارد.

هر دو قضیه معادلند و ما اینجا این موضوع را اثبات می‌کنیم:

به روشنی می‌توان دید که قضیه اول، قضیه دوم را نتیجه می‌دهد. در واقع قضیه دوم حالت خاصی از قضیه اول است. در اینجا نشان می‌دهیم که قضیه دوم هم، قضیه اول را نتیجه می‌دهد.



برای اینکه X_t و Z_t بتوانند به هم برسند باید یک نقشه بین دو فضای بالا وجود داشته باشد. فرض کنید که $[x_t: t\in T]$ روی $(\Omega,\mathcal{F},P)$ توزیع‌های متناهی البعد $\mu_{t_1\dots t_k}$ را دارا باشد و یک نگاشت $\xi\,:\,\Omega\longrightarrow\mathcal{R}^T$ را بصورت زیر تعریف کنید:

$$Z_t(\xi(\omega))=x_t(\omega),\qquad t\in T \tag{16.14}$$

برای هر ω ، $\xi(\omega)$ یکی از اعضای R^T است.

واضح است که:

$$\xi^{-1}\left[x \in R^T: \left(Z_{t_1}(x), \dots, Z_{t_k}(x)\right) \in H\right] \longrightarrow$$
$$= \left[\omega \in \Omega: \left(Z_{t_1}(\xi(\omega)), \dots, Z_{t_k}(\xi(\omega))\right) \in H\right]$$
$$= \left[\omega \in \Omega: \left(X_{t_1}(\xi(\omega)), \dots, X_{t_k}(\xi(\omega))\right) \in H\right] \qquad (16.15)$$

A و $\mathcal{R}^T$ هستند و چون $\xi(\omega)$ در $\mathcal{R}^T$ است پس باید ξ^{-1} روی مولد $\mathcal{R}^T$ اثر کند تا اگر در $\mathcal{F}$ بیفتد به معنی اندازه‌پذیر بودن ξ نسبت به $\mathcal{F}$ باشد.

از آنجاییکه X_tها متغیرهای تصادفی‌اند که نسبت به $\mathcal{F}$ اندازه‌پذیرند، این مجموعه در $\mathcal{F}$ می‌افتد اگر $H \in \mathcal{R}^k$ باشد. بنابراین $\xi^{-1}A \in \mathcal{F}$ می‌شود که در آن $A \in \mathcal{R}_0^T$ است و بنابراین، ξ نسبت به $\mathcal{F}/\mathcal{R}^T$ اندازه‌پذیر می‌شود.

با استفاده از رابطه (16.15) و این حقیقت که $[x_t: t \in T]$ دارای توزیع‌های متناهی البعد $\mu_{t_1 \dots t_k}$ می‌باشد، $P\xi^{-1}$ در رابطه زیر صدق می‌کند: (از طرفین رابطه (16.15) احتمال می‌گیریم)

$$P\xi^{-1}\left[x \in R^T: \left(Z_{t_1}(x), \dots, Z_{t_k}(x)\right) \in H\right]$$
$$= P\left[\omega \in \Omega: \left(X_{t_1}(\xi(\omega)), \dots, X_{t_k}(\xi(\omega))\right) \in H\right] = \mu_{t_1 \dots t_k}(H) \qquad (16.16)$$

در نتیجه فرآیند $[Z_t: t \in T]$ نیز روی $(R^T, \mathcal{R}^T, P\xi^{-1})$ دارای توزیع‌های متناهی البعد $\mu_{t_1 \dots t_k}$ می‌باشد.

بنابراین اثبات یکی از صورت‌های قضیه وجودی کولموگروف معادل اثبات دیگری است.

مثال ۱۶-۱: فرض کنید که T متناهی است، به عنوان مثال $T = \{1,2,\dots,k\}$، بنابراین $(R^k, \mathcal{R}^k) \equiv (R^T, \mathcal{R}^T)$ و با فرض اینکه $P = \mu_{1,2,\dots,k}$ ، چنین P ای در شرایط قضیه ۱۶-۱ صدق می‌کند.

مثال ۱۶-۲: فرض کنید که $T = \{1,2,\dots\}$ و

$$\mu_{t_1 \dots t_k} = \mu_{t_1} \times \dots \times \mu_{t_k} \qquad (16.17)$$

که در آن μ_1، μ_2، ... توزیع‌هایی روی R هستند. به راحتی می‌توان دید که شرایط سازگاری (16.2) و (16.3) در اینجا برقرار هستند؛

اثبات (16.2):

$$\mu_{t_1 \dots t_k} = \mu_{t_1} \times \dots \times \mu_{t_k}$$
$$\mu_{t_{\Pi_1} \dots t_{\Pi_k}} = \underbrace{\mu_{t_{\Pi_1}} \times \dots \times \mu_{t_{\Pi_k}}}_{\text{یک جایگشت از } \mu_{t_1} \times \dots \times \mu_{t_k}}$$

عبارت‌های سمت راست با هم برابرند بنابراین عبارت‌های سمت چپ نیز برابر می‌شوند

اثبات (16.3):

$$\mu_{t_1 \dots t_{k-1}}(H_1 \times \dots \times H_{k-1}) = \mu_{t_1}(H_1) \times \dots \times \mu_{t_{k-1}}(H_{k-1})$$

$$\mu_{t_1 \dots t_{k-1} t_k}(H_1 \times \dots \times H_{k-1} \times R) = \mu_{t_1}(H_1) \times \dots \times \mu_{t_{k-1}}(H_{k-1}) \times \mu_{t_k}(R)$$

عبارت‌های سمت راست برابرند بنابراین عبارت‌های سمت چپ نیز برابر می‌شوند.

یادآوری (قضیه): اگر $\{\mu_n\}$ یک دنباله متناهی یا نامتناهی از اندازه احتمال روی $\mathcal{R}$ باشد، آنگاه روی یک فضای احتمال $(\Omega, \mathcal{F}, P)$، یک دنباله از متغیرهای تصادفی وجود دارد به قسمی که $\{X_n\}$ دارای توزیع μ_n باشد.

بنابراین با استفاده از قضیه فوق روی یک فضای احتمال $(\Omega, \mathcal{F}, P)$، یک دنباله مستقل $X_1, X_2, \dots$ از متغیرهای تصادفی وجود دارد که توزیع‌های آنها به ترتیب می‌شود $\mu_1, \mu_2, \dots$ که در این صورت رابطه (16.17)، توزیع $\left(X_{t_1}, \dots, X_{t_k}\right)$ می‌شود.

نتیجه: اگر X و Y مستقل باشند، خودبه‌خود توزیع توأم آنها حاصلضرب توزیع‌ها می‌شود.

مثال ۱۶- ۳: اگر T یک زیر مجموعه از R باشد، به عنوان مثال $[3,12]$، (یعنی رابطه ترتیب روی T تعریف شده)، در اینصورت با در نظر گرفتن این فاصله، ما باید بتوانیم μ را برای هر حالتی از این فاصله حساب کنیم. به عنوان مثال $\mu_{3.5,4,1.5,6}$. با استفاده از نتیجه این مثال می توان $\mu_{1.5,3.5,4,6}$ را محاسبه کرده و μ مورد نظر نامرتب را همان μ مرتب تعریف کرد.

$$\mu_{3.5,4,1.5,6} = \mu_{1.5,3.5,4,6}$$

اولین اثبات از قضایای کلموگروف:

قضیه ۱۶-۱ را در نظر بگیرید. اگر A یک سیلندر باشد که با رابطه (16.9) تعریف شد، حال تعریف زیر را ارائه می‌دهیم:

$$P(A) = \mu_{t_1 \dots t_k}(H) \tag{16.18}$$

این تعریف خیلی عادی است و طبیعی است که چنین تعریفی بکنیم ولی این تعریف باید خوش‌تعریف باشد یعنی با تغییر اندیس‌های $t_1, \dots, t_k$ و H که در رابطه (16.9) حضور دارند، این احتمال تغییر نکند و وابسته به نوع نمایش نباشد، به عنوان مثال:

$$\left\{x \middle| \left(x_{t_1}, x_{t_5}\right) \in (3,8) \times (12,18)\right\} \equiv \left\{x \middle| \left(x_{t_1}, x_{t_3}, x_{t_5}\right) \in (3,8) \times R \times (12,18)\right\}$$

در واقع اینجا این سوال برایمان پیش می‌آید که آیا این تعریف سازگار است یا نه؟ زیرا A به عنوان یک سیلندر نمایش‌های دیگری هم دارد. فرض کنید که مثلا A بر سیلندر B که با تعریف (16.10) معرفی شد منطبق باشد می‌خواهیم ببینیم آیا می‌توان نتیجه گرفت که $P(A) = P(B)$؟ چون ما تا اینجا هنوز جنس P را نمی‌شناسیم و نمی‌توانیم مستقیم بگوییم که چون $A = B$ است پس $P(A) = P(B)$، زیرا هنوز مشخص نیست که P احتمال باشد.

همانطور که قبلا مشاهده شد اگر $(u_1, \dots, u_m)$ شامل همه t_α و s_β ها باشد، می‌توان آن را با (16.11) نیز نمایش داد که در آن $H' = \psi^{-1}H$ می‌باشد و ψ نیز در رابطه (16.6) تعریف شده است. از آنجا که شرایط سازگاری (16.2) و (16.3) بصورت کلی‌تر در رابطه (16.7) بیان شدند پس:

$$P(A) = \mu_{t_1 \dots t_k}(H) = \mu_{u_1 \dots u_m}(H')$$

بطور مشابه رابطه (16.10) نیز فرم (16.12) را خواهد داشت و

$$P(B) = \mu_{s_1 \dots s_j}(I) = \mu_{u_1 \dots u_m}(I')$$

از آنجا که u_γها مجزا هستند برای هر سری از اعداد حقیقی $\mathfrak{z}_1, \dots, \mathfrak{z}_m$ نقاطی مانند x از R^T وجود دارند که:

$$\left(x_{u_1}, \dots, x_{u_m}\right) = (\mathfrak{z}_1, \dots, \mathfrak{z}_m)$$

از این رابطه نتیجه می‌شود که اگر سیلندرهای (16.11) و (16.12) منطبق باشند، در اینصورت $H' = I'$ می‌شود. بنابراین از $A = B$ می‌توان نتیجه گرفت که:

$$P(A) = \mu_{u_1 \dots u_m}(H') = \mu_{u_1 \dots u_m}(I') = P(B)$$

و تعریف (16.18) به راستی سازگار است.

خلاصه:

$$A = \left[x \middle| x \in R^T : \left(x_{u_1}, \dots, x_{u_m}\right) \in \psi^{-1}H = H'\right]$$

$$B = \left[x \middle| x \in R^T : \left(x_{u_1}, \dots, x_{u_m}\right) \in \psi^{-1}I = I'\right]$$

$$since\ A = B \implies H' = I' \implies \mu_{u_1,\dots,u_m}(H') = \mu_{u_1,\dots,u_m}(I') \implies P(A) = P(B)$$

(از اینجا به بعد می‌خواهد خواص تابع احتمال‌زا را در P چک کند چون ما هنوز نمی‌دانیم که P احتمال است.) حال سیلندرهای مجزای A و B را در نظر بگیرید. بطور معمول مجموعه اندیس‌ها را معادل هم و مشابه هم در نظر

می‌گیریم. حال فرض کنید که A با فرمول (16.11) و B با فرمول (16.12) داده شده باشند، بطوریکه (16.13) برقرار باشد.

اگر $H' \cap I' \neq \emptyset$ باشد، در اینصورت $A \cap B$ هم باید غیر تهی باشد. بنابراین چون A و B مجزا در نظر گرفته شده‌اند، $H' \cap I' = \emptyset$ می‌شود و

$$P(A \cup B) = \mu_{u_1 \dots u_m}(H' \cup I')$$
$$= \mu_{u_1 \dots u_m}(H') + \mu_{u_1 \dots u_m}(I') = P(A) + P(B)$$

بنابراین P روی $\mathcal{R}_0^T$ جمع‌پذیر متناهی است. روشن است که $P(R^T) = 1$.

فرض کنید که بتوان نشان داد که P خاصیت جمع‌پذیری شمارش‌پذیر داشته باشد، پس P اندازه احتمال می‌شود، (روی $\mathcal{R}_0^T$)، بنابراین P توسیعی به یک اندازه احتمال روی $\mathcal{R}^T$ خواهد داشت.

ضمناً P تعریف زیر را روی $\mathcal{R}_0^T$ داشت و بنابراین فرآیند مختصاتی $[Z_t: t \in T]$ توزیع‌های متناهی البعد مورد نیاز را خواهد داشت. در اینجا چون ثابت شده P تابع احتمال است پس می‌توان حرف توزیع را به میان آورد.

$$\Rightarrow P\left[x \in R^T: \left(Z_{t_1}(x), \dots, Z_{t_k}(x)\right) \in H\right] = \mu_{t_1 \dots t_k}(H)$$

پس حالا کافی است که نشان دهیم P روی $\mathcal{R}_0^T$ جمع‌پذیر شمارش پذیر است. برای اثبات این موضوع از این نکته استفاده می‌کنیم که اگر P یک اندازه احتمال با خاصیت جمع‌پذیری متناهی باشد که روی میدان $\mathcal{F}$ تعریف می‌شود و اگر $A_n \downarrow \emptyset$ برای $A_n \in \mathcal{F}$، نتیجه بدهد $P(A_n) \downarrow 0$، در این صورت P خاصیت جمع‌پذیری شمارش‌پذیر دارد. فرض کنید که $A_1 \supset A_2 \supset \cdots$ و برای هر n، $P(A_n) \geq \varepsilon > 0$ باشد. هدف این است که نشان دهیم $\bigcap_n A_n \neq \emptyset$ است. (برهان خلف: فرض کرده‌ایم که P خاصیت جمع‌پذیری شمارش‌پذیر ندارد ولی در عین حال $A_n \downarrow \emptyset$ نتیجه می‌دهد $P(A_n) = 0$. نشان می‌دهیم که چنین فرض غلط است و $P(A_n) \neq 0$ است)

از آنجاییکه $A_n \in \mathcal{R}_0^T$ و از آنجا که مجموعه اندیس‌هایی که در تعیین یک سیلندر نقش دارند همیشه می‌توانند جایگشت و یا گسترش داشته باشند،

توضیح:

$A_2 \subset A_1$ در نتیجه:

$$A_1 = \{(x_7, x_{13}, x_{14}) \in H_1 \times H_2 \times H_3\}$$
$$\supset A_2 = \{(x_7, x_{13}, x_{14}, x_{18}) \in H_1 \times H_2 \times H_3 \times H_4\}$$

بنابراین با افزایش اندیس A تعداد اندیس‌های xها هم می‌توانند افزایش پیدا کنند. می‌دانیم که A_2 زیر مجموعه A_1 است.

بنابراین دنباله‌ای از $t_1, t_2, \ldots$ که نقاطی در T هستند وجود دارد که:

$$A_n = \left[x \in R^T : \left(x_{t_1}, \ldots, x_{t_n}\right) \in H_n\right] \;\; ; \; H_n \in \mathcal{R}^n$$

با این تعریف A_1 تنها درگیر t_1 می‌شود. در حالیکه ممکن است در اصل A_1 درگیر ۵ تا اندیس باشد، ما می‌توانیم به جای بقیه اندیس‌ها R بگذاریم و مثلا به M برسیم: $A_1 \supset M$ و M همان مجموعه‌ای خواهد بود که یک اندیس دارد و ما می‌توانیم M را همان A_1 تعریف شده در بالا بگیریم.

البته $P(A_n) = \mu_{t_1 \ldots t_n}(H_n)$. حال با استفاده از خاصیت regularity درون H_n یک مجموعه فشرده مثل K_n وجود دارد بطوریکه:

$$\mu_{t_1 \ldots t_n}(H_n - K_n) < {}^{\varepsilon}\!/_{2^{n+1}}$$

اگر $B_n = \left[x \in R^T : \left(x_{t_1}, \ldots, x_{t_n}\right) \in K_n\right]$ باشد، در این صورت از آنجا که:

$$A_n - B_n = \left\{x \middle| \left(x_{t_1}, \ldots, x_{t_n}\right) \in H_n - K_n\right\}$$

$$\Rightarrow P(A_n - B_n) = \mu_{t_1 \ldots t_n}(H_n - K_n)$$

$$\Rightarrow P(A_n - B_n) < {}^{\varepsilon}\!/_{2^{n+1}}$$

حال فرض کنید $C_n = \bigcap_{k=1}^{n} B_k$، بنابراین: $C_n \subset B_n \subset A_n$ و در نتیجه:

$$P(A_n - C_n) = P(A_n - \textstyle\bigcap_{k=1}^{n} B_k) = P(A_n \bigcap (\bigcup_{k=1}^{n} B_k^c))$$

$$= P(\textstyle\bigcup_{k=1}^{n} (A_n \bigcap B_k^c))$$

$$\leq \textstyle\sum_{k=1}^{n} P(A_n - B_k)$$

$$\leq \textstyle\sum_{k=1}^{\infty} P(A_n - B_k)$$

$$< \textstyle\sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} \sum_{k=1}^{\infty} \frac{1}{2^k} = \frac{\varepsilon}{2}$$

$$\Rightarrow P(A_n - C_n) < \frac{\varepsilon}{2}$$

از طرفی:

$$P(A_n - C_n) = P(A_n) - P(C_n) \quad since \;\; C_n \subset A_n$$

و می‌دانیم که $P(A_n) \geq \varepsilon$ و $P(A_n - C_n) < \frac{\varepsilon}{2}$ است. بنابراین $P(C_n) > \frac{\varepsilon}{2} > 0$ و این نتیجه می‌دهد که $C_n \neq \emptyset$ است.

حال یک نقطه مثل $x^{(n)} \in C_n$ در R^T انتخاب می‌کنیم. اگر $n \geq k$ باشد، آنگاه:

$x^{(n)} \in C_n \subset C_k \subset B_k$ و بنابراین بر اساس تعریف B_k:

$$\left(x_{t_1}^{(n)}, \dots, x_{t_k}^{(n)}\right) \in K_k$$

از آنجاییکه K_k فشرده است پس کراندار است، بنابراین زیردنباله $\left(x_{t_k}^{(1)}, x_{t_k}^{(2)}, \dots\right)$ به ازای هر k کراندار است. حال اینجا کار ما یک اشکال دارد و آن این است که برای هر k یک زیر دنباله جداگانه خواهیم داشت، بنابراین باید با استفاده از روشی خاص، زیر دنباله‌ای را بیرون بکشیم که به ازای هر k همگرا باشد. این روش روش قطری است.

با استفاده از این روش، یک زیر دنباله صعودی مثل $n_1, n_2, \dots$ از اعداد صحیح انتخاب می‌کنیم، به‌طوریکه $\lim x_{t_k}^{(n_i)}$ برای هر k وجود داشته باشد. حال در R^T، نقطه‌ای مانند x وجود دارد که t_kامین مؤلفه آن جواب این حد است:

$$\lim x_{t_k}^{(n_i)} = x_{t_k}$$

در این صورت $\forall k$، $\left(x_{t_1}, \dots, x_{t_k}\right)$ حد $\left(x_{t_1}^{(n_i)}, \dots, x_{t_k}^{(n_i)}\right)$ است وقتی $i \longrightarrow \infty$ و بنابراین در K_k می‌افتد. چون K_k فشرده است و حد هر دنباله‌ای در یک مجموعه فشرده در خود مجموعه قرار دارد و این به آن معنی است که طبق تعریف B_k، $x = \left(x_{t_1}, \dots, x_{t_k}\right)$ در B_k قرار دارد و بنابراین چون $\forall k$، $B_k \subset A_k$ است پس در A_k قرار دارد. بنابراین $x \in \bigcap_{k=1}^{\infty} A_k$ (چون $\forall k$ برقرار است پس در اشتراک می‌افتد) که نشان می‌دهد $\bigcap_k A_k \neq \emptyset$ است پس $A_n \neq \emptyset$ که نتیجه می‌دهد $P(A_n) \neq 0$ که این در تناقض با فرض است. پس فرض خلف از ابتدا اشتباه بوده و در نتیجه P خاصیت جمع‌پذیری شمارش‌پذیر دارد.

نابسندگی $\mathcal{R}^T$

قضیه ۱۶-۳: با فرض اینکه $[X_t : t \in T]$ متعلق به خانواده ای از توابع حقیقی بر Ω باشند

i) اگر $A \in \sigma[X_t : t \in T]$ و $\omega \in A$ و اگر $X_t(\omega) = X_t(\omega')$ برای تمام $t \in T$، آنگاه $\omega' \in A$.

ii) اگر $A \in \sigma[X_t : t \in T]$، برای برخی زیر مجموعه‌های شمارش پذیر S از T داریم $A \in \sigma[X_t : t \in S]$.

اثبات:

تعریف می کنیم $Z_t(\xi(\omega)) = X_t(\omega)$ برای $\xi: \Omega \longrightarrow R^T$. حال برای $\mathcal{F} = \sigma[X_t: t \in T]$ می‌دانیم ξ بر $\mathcal{F}/_{\mathcal{R}^T}$ اندازه پذیر است زمانیکه $\mathcal{F}$ شامل گردایه $[\xi^{-1}M: M \in \mathcal{R}^T] \subset \mathcal{F}$ $\longrightarrow$ $[\xi^{-1}M: M \in \mathcal{R}^T]$ باشد.

(I) مفهوم اندازه‌پذیر بودن ξ نسبت به $\mathcal{F}/_{\mathcal{R}^T}$ یعنی:

$$\forall M \in R^T;\ \xi^{-1}M \in \mathcal{F}$$

$$\Rightarrow \mathcal{F} \text{ شامل گردایه } [\xi^{-1}M: M \in \mathcal{R}^T]$$

گردایه k بعدی یک سیگما میدان است و با استفاده از (16.15) شامل مجموعه‌های زیر است:

$$\left[\omega \in \Omega:\ \left(X_{t_1}(\omega), \dots, X_{t_k}(\omega)\right) \in H\right],\ H \in \mathcal{R}^k$$

بنابراین شامل سیگما میدان $\mathcal{F}$ که $\mathcal{F} \subset [\xi^{-1}M: M \in \mathcal{R}^T]$ را تولید می کنند است. (II) چون کلاس $[\xi^{-1}M: M \in \mathcal{R}^T]$ شامل مجموعه‌های بالاست پس شامل سیگما میدانی که تولید می‌کنند نیز می‌باشد.

با استفاده از (I) و (II) داریم:

$$\mathcal{F} = \sigma[X_t: t \in T] = [\xi^{-1}M: M \in \mathcal{R}^T]$$

اثبات (i): با استفاده از فرض داریم: $\omega \in A = \xi^{-1}M$

عضوی از $[\xi^{-1}\mathrm{M}: \mathrm{M} \in \mathcal{R}^{\mathrm{T}}]$ | عضوی از $\sigma[\mathrm{X_t}: \mathrm{t} \in \mathrm{T}]$

$$\Rightarrow\ X_t(\omega) = X_t(\omega')\ \ \forall t \in T\ \Rightarrow\ Z_t(\xi(\omega)) = Z_t(\xi(\omega'))\ \forall t \in T$$

$$\Rightarrow\ \xi(\omega) = \xi(\omega') \Rightarrow\ \omega' \in A = \xi^{-1}M$$

اثبات (ii) :

برای $S \subset T$ داریم $\mathcal{F}_S = \sigma[X_t: t \in T]$، حال می خواهیم نشان دهیم $\mathcal{F} = \mathcal{F}_T$: $\bigcup_S \mathcal{F}_S = \mathcal{F}_T$ بر $\mathcal{G} =$ $\bigcup_S \mathcal{F}_S$ منطبق است. نتایج را برای اجتماع زیرمجموعه‌های شمارش پذیر S از T بسط می دهیم.

مشاهده می شود $\bigcup_S \mathcal{F}_S \subset \mathcal{F}$. حال نشان می دهیم $\mathcal{F} \subset \mathcal{G} = \bigcup_S \mathcal{F}_S$

برای این منظور ابتدا نشان می‌دهیم $\mathcal{G}$ یک سیگما میدان است:

1. $\Omega \in \mathcal{F}_s$ since $\mathcal{F}_s$ is σ–field $\Rightarrow$ $\Omega \in \mathcal{G}$

2. If $A \in \mathcal{G} \Rightarrow A \in \cup_s \mathcal{F}_s \Rightarrow A \in \mathcal{F}_s$ for some countable $S \Rightarrow A^c \in \mathcal{F}_s$ for some $S \Rightarrow A^c \in \cup_s \mathcal{F}_s = \mathcal{G} \Rightarrow A^c \in \mathcal{G}$
3. If $A_1, A_2, \dots \in \mathcal{G} \Rightarrow A_n \in \mathcal{F}_{s_n}$ for some countable $S_n \Rightarrow \cup_n A_n \in \mathcal{F}_s$ for $S = \cup_n S_n \Rightarrow \cup_n A_n \in \mathcal{G}$

بنابراین $\mathcal{G}$ یک سیگما میدان است و چون شامل مجموعه‌های $[X_t \in H]$ است، شامل سیگما میدان $\mathcal{F}$ است که تولید می کنند $\mathcal{F} \subset \cup_s \mathcal{F}_s = \mathcal{G}$ پس $\mathcal{F} = \cup_s \mathcal{F}_s$ و اگر $A \in \mathcal{F} = \sigma[X_t: t \in T]$ برای برخی زیرمجموعه-های شمارش پذیر نتیجه می دهد $A \in \sigma[X_t: t \in S]$.

نکته ۱۶ -۳: از این قضیه نتیجه می‌شود که خیلی از توابعی که در R^T هست در $\mathcal{R}^k$ نیست. مثلا فرض کنید که $T = [0, \infty)$ باشد. زیر مجموعه C از R^T را در نظر بگیرید که برابر با توابع پیوسته باشد:

C: تابعی پیوسته از $R^{T=[0,\infty)}$

ولی نشان می‌دهیم که C در $\mathcal{R}^T$ نیست و این بی‌کفایتی $\mathcal{R}^T$ را می‌رساند. زیرا توابع پیوسته توابع مهمی هستند ولی در $\mathcal{R}^T$ وارد نمی‌شوند. برای اثبات از برهان خلف استفاده می‌کنیم. فرض کنید که C در $\mathcal{R}^T$ باشد. با استفاده از قسمت (ii) قضیه (فرض کنید $\Omega = R^T$ و قرار دهید $[Z_t: t \in T]$ در نقش $[X_t: t \in T]$) بنابراین C در $\sigma[Z_t: t \in S]$ قرار می‌گیرد، برای بعضی S های شمارش‌پذیر که $S \subset [0, \infty)$. حال با استفاده از قسمت (i) اگر $x \in C$ باشد و $\forall t \in S\ , Z_t(x) = Z_t(y)$، در این صورت $y \in C$. از آنجا که فرض کرده بودیم $C \in \mathcal{R}^T$ است نتیجه می‌شود که یک مجموعه شمارش‌پذیر مانند S وجود دارد بطوریکه:

If $x \in C$ and $x(t) = y(t)\ \ \forall t \in S \ \Rightarrow\ y \in C$

یعنی yهم پیوسته است، ولی این درست نیست، زیرا مجموعه شمارش‌پذیر S هر چه که باشد برای هر تابع پیوسته x واضح است که تابع‌هایی مانند y وجود دارند که نقاط ناپیوستگی دارند ولی روی یک مجموعه شمارش‌پذیر با برابرند. بنابراین یک نتیجه غلط بدست آورده‌ایم و C نمی‌توانسته از اول در $\mathcal{R}^T$ باشد.

مثال ۱۶- ۴:

$$\left.\begin{aligned} & x(t) = \frac{1}{t^2+1} \\ & y(t) = \begin{cases} \dfrac{1}{t^2+1} & t \in S = Z \\ 0 & o.w. \end{cases} \end{aligned}\right\} \Rightarrow x(t) = y(t)\ \forall t \in S$$

بحث بالا نشان می‌دهد که یک مجموعه مثل A در R^T نمی‌تواند در $\mathcal{R}^T$ باشد مگر اینکه یک زیر مجموعه شمارش‌پذیر S از T وجود داشته باشد با این خاصیت که اگر

$$x \in A \ , x(t) = y(t) \ \forall t \in S \ \Rightarrow \ y \in A$$

بنابراین A نمی‌تواند در $\mathcal{R}^T$ باشد اگر ما قادر باشیم (حداقل) تنها یکی از نقاط $x \in A$ را از A خارج کنیم در واقع تعریفش را تغییر دهیم مثل x و y در بالا و C چنین مجموعه‌ای است که می‌توان در تعریف توابع پیوسته دستکاری کرد.

۱۶–۳ مسائل

۱) عملگر ψ_i را بر ۵ تایی زیر اعمال کرده و تمام نتایج ۳ تایی مطلوب را بدست آورید.

$$\left(x_1, x_2, x_3, x_4, x_5\right)$$

۲) تساوی قضیه‌های وجودی کلموگروف را اثبات کنید.

۳) با فرض این که داشته باشیم $\mu_{t_1,t_2,\dots,t_k} = \mu_{t_1} + \mu_{t_2} + \dots + \mu_{t_k}$، آیا این اندازه در شرایط قضیه‌های کلموگروف صدق می‌کند؟

۴) با ذکر دو مثال بی‌کفایتی $\mathcal{R}^T$ نسبت به R^T را نشان دهید.

فصل هفدهم

حرکت براونی

۱۷–۱ مقدمه

تعریف۱۷-۱: فرایند تصادفی $\{B_t: \mathrm{t} \geq 0\}$ یک حرکت براونی نامیده می شود اگر :

۱) $B(0) = 0$

۲) فرایند $B(t)$ یک تابع پیوسته از t باشد.

۳) نموهای فرایند در فاصله های زمانی متعامد، متغیرهای تصادفی مستقل از یکدیگر باشند. یعنی برای هر $t_1 < t_2 < t_3 < t_4$ ، متغیرهای تصادفی $B(t_4) - B(t_3)$ و $B(t_3) - B(t_2)$ و $B(t_2) - B(t_1)$ از هم مستقل باشند.

۴) نمو فرایند در بازه $[s, t]$ دارای توزیع نرمال با میانگین صفر و واریانس $t - s$ باشد. به عبارت دیگر:

$$B(t) - B(s) = B(t - s) \sim N(0, t - s) \ \ , t > s$$

بنابراین $P(B(t) - B(s) \leq a) = \int_{-\infty}^{a} \frac{e^{\frac{-x^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dx$.

نکته ۱۷-۱:

۱) $B(t) \sim N(0, t)$

۲) اگر t مقدار ثابتی باشد، آنگاه $B(t, \omega)$ یک متغیر تصادفی است.

۳) اگر ω ثابت باشد، آنگاه $B(t, \omega)$ یک تابع از t است و مسیر نمونه ای نامیده می شود.

۴) از آنجا که $B(t) - B(s) \sim N(0, t - s)$ در نتیجه،

$$\Delta B(t) = B(t_2) - B(t_1) \sim N(0, t_2 - t_2) = N(0, \Delta t)$$
$$\Rightarrow \Delta B(t) \sim N(0, \Delta t)$$

۵) می توان با استفاده از بند (۴) در تعریف مشتق این فرایند از رابطه زیر استفاده نمود.

$$\Delta B(t) = B(t_2) - B(t_1) = B(\Delta t) + B(t) - B(t)$$

$$\Rightarrow \Delta B(t) = B(\Delta t + t) - B(t)$$

۲–۱۷ کاربردها

حرکت براونی در معادلات پویای قیمت سهام

فرض کنید $S(t)$ قیمت سهام در زمان t باشد. در اینصورت $S(\Delta t + t)$ قیمت سهام در زمان $\Delta t + t$ و $\Delta S(t)$ مقدار نمو قیمت سهام در این بازه است. آنگاه:

$$\Delta S(t) = S(\Delta t + t) - S(t)$$

$$\frac{\boldsymbol{S(\Delta t+t)-S(t)}}{\boldsymbol{S(t)}} = \frac{\boldsymbol{\Delta S(t)}}{\boldsymbol{S(t)}} = \boldsymbol{M\Delta t + \sigma \Delta B(t)} \qquad (۱-۱۷)$$

حال اگر $t\epsilon[0,\infty)$ باشد، آنگاه

$$\frac{\mathrm{d}S(t)}{S(t)} = M\Delta t + \sigma \Delta B(t) \qquad (۲-۱۷)$$

معادلات (۱۷-۱) و (۱۷-۲) معادلات دیفرانسیل تصادفی نامیده می شوند که در هردوی این معادلات، $M\Delta t$ جمله غیرتصادفی معادله و مقدار انحراف نامیده می شود در حالی که $\sigma\Delta B(t)$ جمله تصادفی معادله بوده و انتشار نام دارد.

۳–۱۷ گام تصادفی متقارن

فرایند گام تصادفی متقارن یک حرکت براونی می باشد. به منظور درک بیشتر این مسئله، فرض کنید بازه زمانی $[0,\mathrm{T}]$ به n بخش برابر با طول های $\Delta t = \frac{T}{n}$ افراز گردد. به عبارت دیگر، بازه $[0,\mathrm{T}]$ شامل زیربازه های $\{0\}$، $[0,\Delta t]$، $[0,2\Delta t]$ ، ... و $[0,(n-1)\Delta t]$ می باشد. حال فرض کنید،

$$\mathrm{X_k} \sim \begin{cases} 0.5 & \mathrm{x} = \sqrt{\Delta \mathrm{t}}, -\sqrt{\Delta \mathrm{t}} \\ 0 & \text{جاهای دیگر} \end{cases}$$

همچنین

$$E(\mathrm{X_k}) = 0.5\sqrt{\Delta \mathrm{t}} - 0.5\sqrt{\Delta \mathrm{t}} = 0$$

$$Var(\mathrm{X_k}) = E\left(\mathrm{X_k}^2\right) = 0.5\Delta \mathrm{t} + 0.5\Delta \mathrm{t} = \Delta \mathrm{t}$$

در این صورت، موقعیت ذره در لحظه T برابر با $\mathrm{S_n}$ است به طوری که $\mathrm{S_n} = \sum_{k=1}^{n} \mathrm{X_k}$ و همچنین $\mathrm{X_k}$ ها نیز از یکدیگر مستقل هستند. بنابراین

$$E(\mathrm{S_n}) = E\left(\sum_{k=1}^{n} \mathrm{X_k}\right) = 0$$

$$Var(\mathrm{S_n}) = Var\left(\sum_{k=1}^{n} \mathrm{X_k}\right) = \sum_{k=1}^{n} Var(\mathrm{X_k}) = n\Delta \mathrm{t} = \mathrm{T}$$

توجه داشته باشید که $S(0) = 0$ و $\mathrm{S_n} = \sum_{k=1}^{n} \mathrm{X_k}$ پیوسته است. همچنین $S(t) - S(s)$ و $S(m) - S(v)$ برای $s < t < v < m$ مستقل هستند. در نتیجه به منظور اثبات براونی بودن این فرایند، باید شرط (۴) تعریف (۱-۱۷)، برقرار باشد. برای این منظور از تابع مولدگشتاور $\mathrm{S_n}$ استفاده می‌کنیم.

$$M_{\mathrm{S_n}}(\mathrm{m}) = \mathrm{E}[\mathrm{e}^{\mathrm{mS_n}}] = \mathrm{E}\left[\mathrm{e}^{\mathrm{m}\sum_{k=1}^{n}\mathrm{X_k}}\right] = \prod_{\mathrm{k=1}}^{\mathrm{n}} \mathrm{E}[\mathrm{e}^{\mathrm{mX_k}}] = \left[\mathrm{E}[\mathrm{e}^{\mathrm{mX_1}}]\right]^{\mathrm{n}}$$

$$\mathrm{E}[\mathrm{e}^{\mathrm{mX_1}}] = 0.5\mathrm{e}^{\mathrm{m}\sqrt{\Delta \mathrm{t}}} + 0.5\mathrm{e}^{-\mathrm{m}\sqrt{\Delta \mathrm{t}}}$$

با استفاده از بسط تیلور دو جمله عبارت فوق، داریم

$$\mathrm{e}^{\mathrm{m}\sqrt{\Delta \mathrm{t}}} \cong 1 + \mathrm{m}\sqrt{\Delta \mathrm{t}} + \frac{\mathrm{m}^2\Delta \mathrm{t}}{2}$$

$$\mathrm{e}^{-\mathrm{m}\sqrt{\Delta \mathrm{t}}} \cong 1 - \mathrm{m}\sqrt{\Delta \mathrm{t}} + \frac{\mathrm{m}^2\Delta \mathrm{t}}{2}$$

در نتیجه داریم

$$\mathrm{E}[\mathrm{e}^{\mathrm{mX_1}}] = \frac{1}{2} + \frac{\mathrm{m}\sqrt{\Delta \mathrm{t}}}{2} + \frac{\mathrm{m}^2\Delta \mathrm{t}}{4} + \frac{1}{2} - \frac{\mathrm{m}\sqrt{\Delta \mathrm{t}}}{2} + \frac{\mathrm{m}^2\Delta \mathrm{t}}{4} = 1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}$$

$$M_{\mathrm{S_n}}(\mathrm{m}) = \left[1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}\right]^{\mathrm{n}}$$

$$\ln M_{\mathrm{S_n}}(\mathrm{m}) = \mathrm{n}\ln\left[1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}\right] = \frac{\ln\left[1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}\right]}{\frac{1}{\mathrm{n}}}$$

$$\lim_{\mathrm{n}\to\infty} \ln M_{\mathrm{S_n}}(\mathrm{m}) = \lim_{\mathrm{n}\to\infty} \frac{\ln\left[1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}\right]}{\frac{1}{\mathrm{n}}}$$

$$= \lim_{\mathrm{n}\to\infty} \frac{\frac{-\mathrm{m}^2\mathrm{T}}{2\mathrm{n}^2}}{-\left[1 + \frac{\mathrm{m}^2\mathrm{T}}{2\mathrm{n}}\right]\frac{1}{\mathrm{n}^2}}$$

$$= \lim_{n\to\infty} \frac{m^2T}{2\left[1+\frac{m^2T}{2n}\right]} = \frac{m^2T}{2}$$

$$\Rightarrow \lim_{n\to\infty} M_{S_n}(m) = e^{\frac{m^2T}{2}} \qquad (۳-۱۷)$$

و در نتیجه، $S_n \sim N(0,T)$ و بنابراین ، S_n برای $n \to \infty$ یک فرایند براونی است.

۱۷–۴ کواریانس حرکت براونی

فرض کنید $s < t$ باشد. در این صورت

$$Cov\big(B(t), B(s)\big) = E[B(t)B(s)] = \min\{s,t\} = s \qquad (۴-۱۷)$$

قضیه ۱۷-۱: فرض کنید B و B^* دو حرکت براونی مستقل باشند و $-1 < \rho < 1$، آنگاه فرایند

$Z(t) = \rho B(t) + \sqrt{1-\rho^2}\, B^*(t)$ یک حرکت براونی است.

اثبات: شرط های (۱) و (۲) از تعریف (۱۷-۱) برقرار هستند. به منظور اثبات مسئله، شرط (۳) و (۴) را بررسی می کنیم.

اگر $s < t < v < m$ باشد، آنگاه

$$Z(t) - Z(s) = [\rho B(t) + \sqrt{1-\rho^2}\, B^*(t)] - [\rho B(s) + \sqrt{1-\rho^2}\, B^*(s)]$$

$$= \rho[B(t) - B(s)] + \sqrt{1-\rho^2}[B^*(t) - B^*(s)]$$

$$Z(m) - Z(v) = \rho[B(m) - B(v)] + \sqrt{1-\rho^2}[B^*(m) - B^*(v)]$$

حال از آنجا که $B^*(t) - B^*(s)$ از $B^*(m) - B^*(v)$ و $B(t) - B(s)$ از $B(m) - B(v)$ مستقل هستند و فرایندهای $B^*(t)$ و $B(t)$ نیز از یکدیگر مستقل هستند. در نتیجه نموهای $Z(t) - Z(s)$ و $Z(m) - Z(v)$ مستقل از یکدیگر می باشند.

حال به منظور اثبات برقراری شرط (۴) داریم

$$E[Z(t)] = E\left[\rho B(t) + \sqrt{1-\rho^2}\, B^*(t)\right] = \rho E[B(t)] + \sqrt{1-\rho^2} E[B^*(t)] = 0$$

$$Var[Z(t)] = Var\left[\rho B(t) + \sqrt{1-\rho^2}\, B^*(t)\right]$$

$$= \rho^2 Var[B(t)] + (1-\rho^2) Var[B^*(t)]$$

$$= \rho^2 t + (1-\rho^2)t = t$$

$$\Rightarrow Z(t) \sim N(0,t)$$

تذکر: ضریب همبستگی فرایندهای $Z(t)$ و $B(t)$ برابر با ρ است.

$$Corr\big(Z(t),B(t)\big) = \frac{Cov\big(Z(t),B(t)\big)}{\sqrt{Var\big(Z(t)\big)Var\big(B(t)\big)}}$$

$$= \frac{E[Z(t)B(t)] - E[Z(t)]E[B(t)]}{\sqrt{t.t}}$$

$$= \frac{E[Z(t)B(t)]}{t} = \frac{E\left[\left(\rho B(t) + \sqrt{1-\rho^2}\, B^*(t)\right) B(t)\right]}{t}$$

$$= \frac{\rho E[B(t)^2] + \sqrt{1-\rho^2} E[B(t)]E[B^*(t)]}{t}$$

$$= \frac{\rho Var[B(t)]}{t} = \frac{\rho t}{t} = \rho$$

۱۷–۵ مشتق ناپذیری حرکت براونی

بازه زمانی به طول $\Delta t = \frac{1}{n}$ را با شروع از نقطه t به صورت $[t, t+\Delta t]$ در نظر بگیرید.فرایند تصادفی $X_n = \frac{\Delta B(t)}{\Delta t}$ میزان تغییرات حرکت براونی $B(t)$ در بازه مورد نظر است. اگر X_n متناهی باشد، آنگاه $B(t)$ مشتق پذیر و اگر نامتناهی باشد آنگاه $B(t)$ مشتق ناپذیر است.

اثبات:

$$X_n = \frac{\Delta B(t)}{\Delta t} = \frac{B(t+\Delta t) - B(t)}{\Delta t} = \frac{B\left(t+\frac{1}{n}\right) - B(t)}{\frac{1}{n}} = n\left[B\left(t+\frac{1}{n}\right) - B(t)\right]$$

و از انجا که ترکیب خطی از متغیر نرمال است در نتیجه X_n دارای توزیع نرمال می باشد.

$$E[X_n] = nE\left[B\left(t+\frac{1}{n}\right) - B(t)\right] = 0$$

$$Var(X_n) = n^2 Var\left[B\left(t+\frac{1}{n}\right) - B(t)\right] = n^2\left[Var\left[B\left(t+\frac{1}{n}\right)\right] + Var[B(t)]\right]$$

$$= n^2\left[t + \frac{1}{n} + t - 2t\right] = n$$

$$\Rightarrow X_{\mathrm{n}} \sim \mathrm{N}(0,\mathrm{n}) \quad , \quad Z = \frac{X_{\mathrm{n}}}{\sqrt{\mathrm{n}}} \sim \mathrm{N}(0,1) \quad \Rightarrow \quad X_{\mathrm{n}} = \sqrt{\mathrm{n}}\mathrm{Z} \sim \mathrm{N}(0,\mathrm{n})$$

اگر k یک عدد مثبت باشد، آنگاه $P[|X_{\mathrm{n}}| > k] \to 1$ زیرا

$$P[|X_{\mathrm{n}}| > k] = P\left[\left|\sqrt{\mathrm{n}}\mathrm{Z}\right| > k\right] = P\left[|\mathrm{Z}| > \frac{k}{\sqrt{\mathrm{n}}}\right]$$

$$\lim_{n\to\infty} P[|X_{\mathrm{n}}| > k] = P[|\mathrm{Z}| > 0] = 1$$

یعنی X_{n} متناهی نبوده، بنابراین فرایند $\mathrm{B(t)}$ مشتق پذیر نمی باشد.

۱۷-۶ تغییرات حرکت براونی

فرض کنید بازه زمانی $[0,\mathrm{T}]$ به n بخش برابر با طول های $\Delta t = \frac{T}{n}$ افراز گردد. تغییرات مرتبه اول حرکت براونی یک متغیر تصادفی است به صورت زیر تعریف می شود.

$$\sum_{k=0}^{\mathrm{n}-1}|\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k})| \tag{۱۷-۵}$$

و همچنین داریم

$$E\left[\sum_{k=0}^{\mathrm{n}-1}[\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k})]\right] = \sum_{k=0}^{\mathrm{n}-1}\mathrm{E}|\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k})| = 0$$

$$\begin{aligned} Var\left[\sum_{k=0}^{\mathrm{n}-1}[\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k})]\right] &= \sum_{k=0}^{\mathrm{n}-1}\mathrm{Var}|\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k})| \\ &= \sum_{k=0}^{\mathrm{n}-1}[\mathrm{Var}(\mathrm{B}(\mathrm{t_{k+1}})) + \mathrm{Var}(\mathrm{B}(\mathrm{t_k})) - 2\mathrm{Cov}(\mathrm{B}(\mathrm{t_{k+1}}),\mathrm{B}(\mathrm{t_k}))] \\ &= \sum_{k=0}^{\mathrm{n}-1}[\mathrm{t_{k+1}} - \mathrm{t_k}] = \sum_{k=0}^{\mathrm{n}-1}\Delta t = \mathrm{n}\Delta t = T \end{aligned}$$

همچنین تغییرات مرتبه دوم فرایند نیز از طریق رابطه زیر محاسبه می شود.

$$\sum_{k=0}^{\mathrm{n}-1}(\mathrm{B}(\mathrm{t_{k+1}}) - \mathrm{B}(\mathrm{t_k}))^2 \tag{۱۷-۶}$$

همچنین داریم

$$E\left[\sum_{k=0}^{n-1}\left(B(t_{k+1})-B(t_k)\right)^2\right]=\sum_{k=0}^{n-1}E\left(B(t_{k+1})-B(t_k)\right)^2$$

با توجه به این نکته که اگر $X\sim N(0,\sigma^2)$ داشته باشد، آنگاه :

$$E[|X|^p]=\begin{cases}0 & p \text{ فرد باشد} \\ 1\times 3\times 5\times \ldots \times (p-1)\sigma^2 & p \text{ زوج باشد}\end{cases}$$

داریم

$$\sum_{k=0}^{n-1}E\left(B(t_{k+1})-B(t_k)\right)^2=\sum_{k=0}^{n-1}E\left(\Delta B(t_k)\right)^2=\sum_{k=0}^{n-1}\Delta t=n\Delta t=T$$

همچنین،

$$\mathrm{Var}\left[\sum_{k=0}^{n-1}\left(B(t_{k+1})-B(t_k)\right)^2\right]=\sum_{k=0}^{n-1}\mathrm{Var}\left(B(t_{k+1})-B(t_k)\right)^2$$

$$=\sum_{k=0}^{n-1}\left\{E\left(B(t_{k+1})-B(t_k)\right)^4-E\left(B(t_{k+1})-B(t_k)\right)^2\right\}$$

$$=\sum_{k=0}^{n-1}\{3(\Delta t)^2-(\Delta t)^2\}=2n(\Delta t)^2=2T\Delta t$$

و اگر $\Delta t\to 0$ آنگاه $\mathrm{Var}\left[\sum_{k=0}^{n-1}\left(B(t_{k+1})-B(t_k)\right)^2\right]\to 0$.

۷-۱۷ حرکت براونی انعکاس یافته

زمانی که بجای مقدار فرایند، میزان تغییرات ان در هر جهتی مورد نظر باشد، از شکل انعکاس یافته فرایند براونی استفاده می کنیم. فرض کنید $Z_t=|B(t)|$ باشد. آنگاه مقدار Z_t همان تصویر فرایند $B(t)$ بر روی محور افق و برای مقادیر منفی $B(t)$ است. از آنجا که $B(t)\sim N(0,t)$ در نتیجه

$$G_z=P(Z_t\le z)=P(|B(t)|\le z)=P(-z\le B(t)\le z)$$

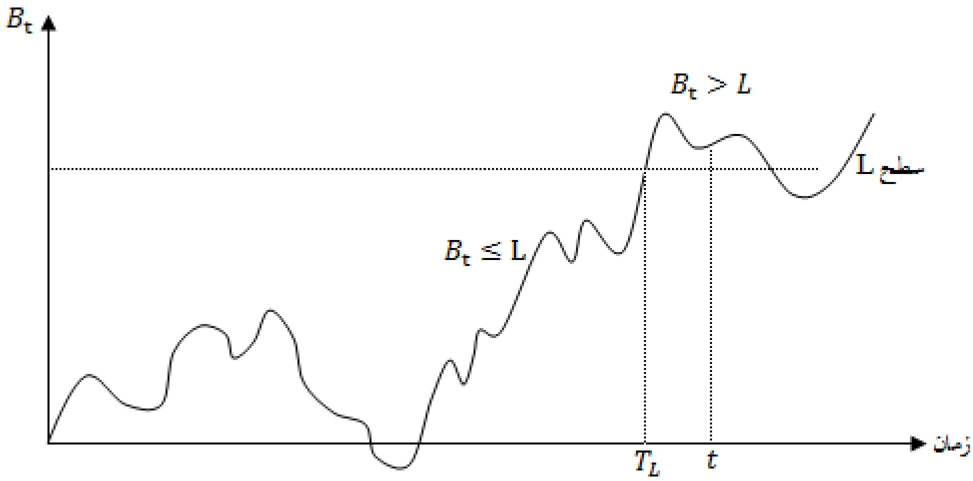
$$=P\left(\frac{-z}{\sqrt{t}}\le\frac{B(t)}{\sqrt{t}}\le\frac{z}{\sqrt{t}}\right)=\phi\left(\frac{z}{\sqrt{t}}\right)-\phi\left(\frac{-z}{\sqrt{t}}\right)=2\phi\left(\frac{z}{\sqrt{t}}\right)-1$$

$$\frac{dG_z}{dz}=g(z)=\frac{z}{\sqrt{t}}\varphi\left(\frac{z}{\sqrt{t}}\right)$$

$$\Rightarrow \mathrm{g(z)} = \frac{\mathrm{z}}{\sqrt{t}}\frac{1}{\sqrt{2\pi}}\, e^{\frac{-z^2}{2t}} \qquad (۷-۱۷)$$

۱۷–۸ اولین عبور فرایند از یک حد

در این بخش، توزیع احتمال عبور فرایند از یک نقطه مرزی محاسبه می شود. برای این منظور فرض کنیدL معرف نقطه مرزی، $B(t)$ معرف فرایند براونی، T_{L} معرف زمانی که فرایند به نقطه L می رسد و t نیز معرف زمانی بعد از زمان T_{L} باشد.



$$P(B_t > L) = P(B_t > L|T_L \leq t)P(T_L \leq t) + P(B_t > L|T_L > t)P(T_L > t)$$

پیشامد $B_t > L$ بدین معناست که فرایند در زمان t در بالای حد مرزی L قرار دارد و می دانیم که طبق فرض $T_L \leq t$ است. از این رو $P(B_t > L|T_L > t)P(T_L > t) = 0$. در نتیجه

$$P(B_t > L) = P(B_t > L|T_L \leq t)P(T_L \leq t)$$

حال، برای فرایند در زمان t دو حالت اتفاق می افتد و آن به این صورت است که فرایند در زمان t برای $T_L \leq t$ یا بالای حد مرزی L قرار دارد و یا پایین حد مرزی. در نتیجه

$$P(B_t > L) = \frac{1}{2}P(T_L \leq t)$$

$$\Rightarrow P(T_L \leq t) = 2P(B_t > L) = 2[1 - P(B_t \leq L)] = 2[1 - P\left(\frac{B_t}{\sqrt{t}} > \frac{L}{\sqrt{t}}\right)]$$

$$H(t) = P(T_L \leq t) = 2[1 - \phi(\frac{L}{\sqrt{t}})]$$

$$h(t) = \frac{\partial H(t)}{\partial t} = \frac{L}{\sqrt{t}} \varphi(\frac{L}{\sqrt{t}}) \qquad (۸-۱۷)$$

۱۷-۹ ماکسیمم فرایند براونی

در این بخش توزیع ماکسیمم یک فرایند براونی مورد بررسی قرار می گیرد. فرض کنید M_{t} ماکسیمم فرایند براونی در بازه زمانی $[0, \mathrm{T}]$ باشد. برای یافتن توزیع آن از توزیع احتمال فرایند برای عبور از نقطه مرزی که در بخش قبل اشاره شد، استفاده می کنیم.

$$\begin{aligned} P(M_t > m) &= P(T_m \leq t) \\ &= 2P(B_t \geq m) \\ &= 2[1 - P(B_t \leq m)] \\ &= 2\left[1 - P\left(\frac{B_t}{\sqrt{t}} > \frac{m}{\sqrt{t}}\right)\right] \\ &= 2[1 - \phi\left(\frac{m}{\sqrt{t}}\right)] \end{aligned}$$

$$\begin{aligned} \Rightarrow R(m) &= P(M_t \leq m) \\ &= 1 - 2\left[1 - \phi\left(\frac{m}{\sqrt{t}}\right)\right] = 2\phi\left(\frac{m}{\sqrt{t}}\right) - 1 \end{aligned}$$

و

$$r(m) = \frac{\partial R(m)}{\partial m} = \frac{2}{\sqrt{t}} \phi\left(\frac{m}{\sqrt{t}}\right) \qquad (۹-۱۷)$$

قضیه ۱۷-۲: نموهای $X_{\mathrm{t}} - \mathrm{X_s}$ فرایند تصادفی X_{t} از مجموعه اطلاعات($\mathcal{F}_s$) تا لحظه s مستقل است و بنابراین مجموعه نموهای فرایند بر روی بازه های زمانی نامتقاطع مستقل هستند.

اثبات: فرض کنید $t_1 < \mathrm{t}_2 < \mathrm{t}_3$. باید نشان دهیم که $X_{\mathrm{t}_3} - \mathrm{X}_{t_2}$ از نمو $X_{\mathrm{t}_2} - \mathrm{X}_{t_1}$ مستقل است.

$$E\left[e^{\theta_1(X_{\mathrm{t}_2} - \mathrm{X}_{t_1}) + \theta_2(X_{\mathrm{t}_3} - \mathrm{X}_{t_2})}\right] = E\left[E\left(e^{\theta_1(X_{\mathrm{t}_2} - \mathrm{X}_{t_1}) + \theta_2(X_{\mathrm{t}_3} - \mathrm{X}_{t_2})} \middle| \mathcal{F}_{\mathrm{t}_2}\right)\right]$$

$$= E\left[e^{\theta_1(X_{\mathrm{t}_2} - \mathrm{X}_{t_1})} E\left(e^{\theta_2(X_{\mathrm{t}_3} - \mathrm{X}_{t_2})} \middle| \mathcal{F}_{\mathrm{t}_2}\right)\right]$$

$$= E\left[e^{\theta_1(X_{\mathrm{t}_2} - \mathrm{X}_{t_1})} E\left(e^{\theta_2(X_{\mathrm{t}_3} - \mathrm{X}_{t_2})}\right)\right] = E\left[e^{\theta_1(X_{\mathrm{t}_2} - \mathrm{X}_{t_1})}\right] E\left[e^{\theta_2(X_{\mathrm{t}_3} - \mathrm{X}_{t_2})}\right]$$

و برابری آخر نیز با استفاده از اندازه پذیری $e^{\theta_1(X_{t_2}-X_{t_1})}$ نسبت به $\mathcal{F}_{t_2}$ و استقلال $e^{\theta_2(X_{t_3}-X_{t_2})}$ از $\mathcal{F}_{t_2}$ حاصل شده است. بنابراین نموهای $X_{t_2}-X_{t_1}$ و $X_{t_3}-X_{t_2}$ مستقل هستند.

قضیه ۱۷-۳(ویژگی مارکوفی): فرایند حرکت براونی $\{B_t\ , t\geq 0\}$ تحت فیلتر $\mathcal{F}_t$ یک فرایند مارکوف می باشد.

پیش از اثبات این قضیه باید به نکته زیر توجه داشت:

نکته ۱۷-۲: فرض کنید $f: s\to R$ یک تابع کراندار و اندازه پذیر باشد.آنگاه X_t نسبت به $\mathcal{F}_s$ برای $s<t$ ویژگی مارکوفی دارد اگر

$$E[f(X_t)|\mathcal{F}_s]=E[f(X_t)|\sigma(X_s)]$$

و به این معناست که $E[f(X_t)|\sigma(X_s)]$ به X_s بستگی دارد.

اثبات: فرض کنید f تابع کراندار و اندازه پذیر باشد و $0\leq s\leq t\leq T$.

$$\begin{aligned}E[f(B_t)|\mathcal{F}_s]&=E[f(B_t-B_s+B_s)|\mathcal{F}_s]\\&=E[f(Z+B_s)|\mathcal{F}_s]\end{aligned}$$

به طوری که $Z=B_t-B_s\sim N(0,t-s)$ و Z از $\mathcal{F}_s$ مستقل است و همچنین B_s نسبت به $\mathcal{F}_s$ اندازه‌پذیر است. بنابراین

$$\begin{aligned}E[f(B_t)|\mathcal{F}_s]&=E[f(Z+B_s)|\mathcal{F}_s]\\&=E[f(Z+B_s)]\\&=\frac{1}{\sqrt{2\pi(t-s)}}\int_{-\infty}^{\infty}f(Z+B_s)e^{\frac{-z^2}{2(t-s)}}dz=g(B_s)\end{aligned}$$

و در نتیجه $E[f(B_t)|\mathcal{F}_s]$ تابعی از B_s است.

قضیه ۱۷-۳(ویژگی مارتینگلی): اگر فرایند $\{B_t\ , t\geq 0\}$ یک حرکت براونی باشد، آنگاه مارتینگل است و فرایند $\{e^{B_t-\frac{1}{2}t}\ , t\geq 0\}$ مارتینگل نمایی است.

اثبات: فرض کنید $s<t$ و $\mathcal{F}_s$ فیلتر باشد. آنگاه

$$\begin{aligned}E[B_t|\mathcal{F}_s]&=E[B_t-B_s+B_s|\mathcal{F}_s]\\&=E[B_t-B_s|\mathcal{F}_s]+E[B_s|\mathcal{F}_s]\end{aligned}$$

با توجه به استقلال B_t-B_s از $\mathcal{F}_s$ و اندازه پذیری B_s نسبت به $\mathcal{F}_s$ داریم:

$$E[B_t-B_s|\mathcal{F}_s]+E[B_s|\mathcal{F}_s]=E[B_t-B_s]+B_s=B_s$$

در نتیجه B_t مارتینگل است.

همچنین،

$$E\left[e^{B_t-\frac{1}{2}t}\middle|\mathcal{F}_\mathrm{s}\right]=E\left[e^{B_t-B_s+B_s-\frac{1}{2}t-\frac{1}{2}s+\frac{1}{2}s}\middle|\mathcal{F}_\mathrm{s}\right]=E\left[e^{(B_t-B_s)-\frac{1}{2}(t-s)+(B_s-\frac{1}{2}s)}\middle|\mathcal{F}_\mathrm{s}\right]$$

با توجه به استقلال $e^{(B_t-B_s)-\frac{1}{2}(t-s)}$ از $\mathcal{F}_\mathrm{s}$ و اندازه پذیری $e^{(B_s-\frac{1}{2}s)}$ نسبت به $\mathcal{F}_\mathrm{s}$ داریم:

$$E\left[e^{(B_t-B_s)-\frac{1}{2}(t-s)+(B_s-\frac{1}{2}s)}\middle|\mathcal{F}_\mathrm{s}\right]=e^{(B_s-\frac{1}{2}s)}e^{-\frac{1}{2}(t-s)}E[e^{(B_t-B_s)}]$$

$$=e^{(B_s-\frac{1}{2}s)}e^{-\frac{1}{2}(t-s)}e^{\frac{1}{2}(t-s)}=e^{(B_s-\frac{1}{2}s)}$$

بنابراین $e^{B_t-\frac{1}{2}t}$ مارتینگل نمایی است.

۱۷–۱۰ مسائل

۱) فرض کنید $\mathrm{B_t}$ یک فرایند براونی باشد. نشان دهید $B_\mathrm{t}{}^2-\mathrm{t}$ یک مارتینگل است.

۲) اگر μ_t مقدار ماکسیمم حرکت براونی در فاصله $[0,\mathrm{t}]$ باشد چگالی μ_t را پیدا کنید

۳) اگر $\mathrm{B_{t_1}}$ و $\mathrm{B_{t_2}}$ متغیرهای تصادفی وابسته‌ای باشند که :

$$X_1=\mathrm{B_{t_1}}\sim N(0,t_1)\,,X_2=\mathrm{B_{t_2}}\sim N(0,t_2)$$

نشان دهید :

$$\mathrm{P}[\mathrm{B_{t_1}}\le a_1\,,\mathrm{B_{t_2}}\le a_2]=\frac{1}{2\pi\sqrt{t_1(t_2-t_1)}}\int_{-\infty}^{a_1}\int_{-\infty}^{a_2}\mathrm{e}^{-\frac{\mathrm{x}_1^2}{2\mathrm{t}_1}-\frac{(\mathrm{x}_2-\mathrm{x}_1)}{2(t_2-\mathrm{t}_1)}}\mathrm{dx_1dx_2}$$

۴) نشان دهید که برای هر $\mathrm{x}<0$ احتمال اینکه $\{B_t^x\}$ حداقل یک صفر در فاصله زمانی $(0,\mathrm{t})$ داشته باشد از رابطه زیر حساب می‌شود :

$$\frac{|\mathrm{x}|}{2\pi}\int_0^t u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\,du$$

(راهنمایی : $f_{T(x)}(t)=\frac{|\mathrm{x}|}{2\pi}t^{-\frac{3}{2}}e^{-\frac{x^2}{2t}}$)

منابع

عین الله *پاشا* , (۱۳۸۹) , احتمال ۱ , انتشارات **دانشگاه پیام نور**

رامداس بات , ترجمه محمد **حسین علامت ساز , ابوالقاسم بزرگ نیا** (۱۳۷۵) ,نظریه احتمال مدررن

Alan F. Karr *(1993). Probability (Springer Texts in Statistics)*

Boris V. Gnedenko *(1998). Theory of Probability*

Erhan Cinlar *(2011). Probability and Stochastics*

Krishna B. Athreya, Soumendra N. Lahiri *(2006). Measure theory and probability theory*

Kai Lai chung*, (2001). A course in probability theory*

Patrick Billingsley *(1995)*. Probability and Measure, Third Edition (Wiley Series in Probability and Statistics)

Robert B. Ash, Catherine A. Doléans-Dade *(2000). Probability & Measure Theory, Second Edition*

Ramdas Bhat .B *(1981). Modern Probability Theory*

Sidney Resnick *(1999).A Probability Path*