آمار ریاضی ۱
فصل دوم: بسندگی

استاد: دکتر اکبر اصغرزاده

دانشگاه مازندران

تعریف آماره

تابعی از نمونه تصادفی که در ساختار آن پارامتر مجهول وجود ندارد.

آماره بسنده

آماره‌ای که حاوی تمام اطلاعاتی است که نمونه‌ی تصادفی آن را در بر دارد.

فرض کنید که $X_1, \ldots, X_n$ یک نمونه‌ی تصادفی از جامعه‌ای با تابع چگالی $f(x, \theta)$ باشد که $\theta \in \Theta$ پارامتر مجهول است. آماره $S = g(X_1, \ldots, X_n)$ با یافته $s = g(x_1, \ldots, x_n)$ (مقداری که در عمل مشاهده می‌شود) را یک آماره بسنده یا کافی برای θ گویند هرگاه این آماره شامل همه اطلاعاتی باشد که نمونه‌ی تصادفی درباره θ می‌دهد.

تعریف آماره بسنده: $(Sufficient\ Statistic)$

S با یافته s را یک آماره بسنده برای θ گویند هرگاه توزیع شرطی نمونه تصادفی $X_۱, \ldots, X_n$ به شرط آماره $S = s$ برای تمام مقادیر s وابسته به θ نباشد.

– تعریف فوق بدین معنی است که با داشتن آماره بسنده S، نمونه تصادفی $X_۱, \ldots, X_n$ اطلاعات بیشتری درباره‌ی پارامتر θ در اختیار ما قرار نمی‌دهد. به عبارتی خود این آماره بسنده S حاوی همه‌ی اطلاعاتی است که نمونه‌ی تصادفی درباره θ در بر دارد.

مثال ۱. اگر $X_۱, \ldots, X_n \overset{iid}{\sim} b(۱, p)$ باشد، نشان دهید که آماره $S = \sum_{i=۱}^{n} X_i$ یک آماره‌ی بسنده برای p است.

حل: تابع جرم احتمال X:

$$P(X = x) = p^x\,(۱ - p)^{۱-x}, \qquad x = ۰, ۱$$

$$
\begin{aligned}
&P(X_۱ = x_۱, \ldots, X_n = x_n | S = s) \\
&= \frac{P(X_۱ = x_۱, \ldots, X_n = x_n,\ S = s)}{P(S = s)} \\
&= \frac{P(X_۱ = x_۱, \ldots, X_n = x_n,\ X_۱ + \cdots + X_n = x_۱ + \cdots + x_n)}{P(S = s)}
\end{aligned}
$$

تحت شرط $x_۱ + \cdots + x_n = s$ داریم

$$P(X_۱ = x_۱, \ldots, X_n = x_n,\ X_۱ + \cdots + X_n = x_۱ + \cdots + x_n) = P(X_۱ = x_۱, \ldots, X_n = x_n),$$

همچنین $S = \sum_{i=۱}^{n} X_i \sim b(n, p)$، و

$$P(S = s) = \binom{n}{s} p^s (۱ - p)^{n-s}, \quad s = ۰, ۱, \ldots, n,$$

لذا

$$P(X_۱ = x_۱, \ldots, X_n = x_n | S = s) = \begin{cases} \frac{P(X_۱ = x_۱, \ldots, X_n = x_n)}{P(S = s)} & x_۱ + \cdots + x_n = s \\ ۰ & x_۱ + \cdots + x_n \neq s \end{cases}$$

$$
= \begin{cases} \frac{\prod_{i=1}^{n} P(X_i = x_i)}{P(S=s)} & x_1 + \cdots + x_n = s \\ 0 & x_1 + \cdots + x_n \neq s \end{cases}
$$

$$
= \begin{cases} \frac{\prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}}{\binom{n}{s} p^s (1-p)^{n-s}} & x_1 + \cdots + x_n = s \\ 0 & x_1 + \cdots + x_n \neq s \end{cases}
$$

$$
= \begin{cases} \frac{p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}}{\binom{n}{s} p^s (1-p)^{n-s}} & \sum_{i=1}^{n} x_i = s \\ 0 & \sum_{i=1}^{n} x_i \neq s \end{cases}
$$

$$
= \begin{cases} \frac{1}{\binom{n}{s}} & \sum_{i=1}^{n} x_i = s \\ 0 & \sum_{i=1}^{n} x_i \neq s \end{cases}
$$

که به p وابسته نیست. لذا S یک آماره بسنده برای p است.

مثال ۲. اگر $X_1, X_2, X_3 \overset{iid}{\sim} b(1,p)$، نشان دهید که آماره $U = X_1 + X_2 + 2X_3$ برای p بسنده نیست.

$$\begin{aligned}
&P(X_1 = 1, X_2 = 1, X_3 = 0 | U = 2) \\
&= \frac{P(X_1 = 1, X_2 = 1, X_3 = 0,\ U = 2)}{P(U = 2)} \\
&= \frac{P(X_1 = 1, X_2 = 1, X_3 = 0)}{P(X_1 = 1, X_2 = 1, X_3 = 0) + P(X_1 = 0, X_2 = 0, X_3 = 1)} \\
&= \frac{p.p.(1-p)}{p.p.(1-p) + (1-p)(1-p)p} = p
\end{aligned}$$

لذا توزیع شرطی نمونه به شرط $U = 2$ به پارامتر p بستگی دارد و لذا آماره U یک آماره بسنده برای p نیست.

$$U = 2 \iff (X_1 = 1, X_2 = 1, X_3 = 0) \text{ یا } (X_1 = 0, X_2 = 0, X_3 = 1)$$

مثال ۳. اگر $X_1, \dots, X_n \overset{iid}{\sim} p(\lambda)$، نشان دهید آماره‌ی $S = \sum_{i=1}^n X_i$ یک آماره‌ی بسنده برای λ است.

حل: داریم $S = \sum_{i=1}^n X_i \sim p(n\lambda)$،

$$P(X_1 = x_1, \dots, X_n = x_n | S = s) = \frac{P(X_1 = x_1, \dots, X_n = x_n, S = s)}{P(S = s)}$$

$$= \begin{cases} \frac{P(X_1=x_1,\dots,X_n=x_n)}{P(S=s)}, & \sum_{i=1}^n x_i = s \\ 0, & o.w \end{cases}$$

$$= \begin{cases} \frac{\prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}}{\frac{(n\lambda)^s e^{-n\lambda}}{s!}}, & \sum_{i=1}^n x_i = s \\ 0, & o.w \end{cases}$$

$$= \begin{cases} \frac{\frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!}}{\frac{(n\lambda)^s e^{-n\lambda}}{s!}}, & \sum_{i=1}^n x_i = s \\ 0, & o.w \end{cases} = \begin{cases} \frac{s!}{n^s \prod_{i=1}^n x_i!}, & \sum_{i=1}^n x_i = s \\ 0, & o.w \end{cases}$$

که به پارامتر λ وابسته نیست و لذا آماره S یک آماره بسنده برای λ است.

تذکر

در حالت کلی برای اینکه آماره‌ی S یک آماره‌ی بسنده برای پارامتر θ باشد، بایستی نسبت چگالی توأم نمونه تصادفی به چگالی احتمال آماره S یعنی نسبت زیر وابسته به پارامتر θ نباشد.

$$\frac{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{f_S(s)}$$

قضیه فاکتورگیری (قضیه دسته بندی فیشر–نیمن)

اگر $X_1,\ldots,X_n \overset{iid}{\sim} f(x,\theta)$ و $\theta\in\Theta$، در اینصورت آماره S با یافته s را یک آماره‌ی بسنده برای θ گویند، اگر و فقط اگر داشته باشیم:

$$f(\underline{x},\theta)=g(s,\theta).\,h(\underline{x})$$

که

- $f(\underline{x},\theta)$ چگالی توأم نمونه
- $g(s,\theta)$ تابعی از یافته آماره بسنده و θ
- $h(\underline{x})$ تابعی از $\underline{x}$ است.

مثال ۴. اگر $X_1, \ldots, X_n \overset{iid}{\sim} f(x,\theta)$ که $0 < x < 1$, $f(x,\theta) = \theta\, x^{\theta-1}$. آماره‌ی بسنده برای θ را پیدا کنید.

حل. تابع چگالی توأم نمونه $\underline{x} = (x_1, \ldots, x_n)$:

$$\begin{aligned}
f(\underline{x},\theta) &= \prod_{i=1}^{n} f(x_i,\theta) \\
&= \prod_{i=1}^{n} \theta\, x_i^{\theta-1} \\
&= \theta^n \Big(\prod_{i=1}^{n} x_i\Big)^{\theta-1} \\
&= \theta^n \Big(\prod_{i=1}^{n} x_i\Big)^{\theta} \Big(\prod_{i=1}^{n} x_i\Big)^{-1} \\
&= g\Big(\prod_{i=1}^{n} x_i, \theta\Big)\, h(\underline{x})
\end{aligned}$$

لذا از قضیه فاکتورگیری، آماره بسنده برای θ می‌شود $S(\underline{X}) = \prod_{i=1}^{n} X_i$.

مثال ۵. اگر $X_1, \ldots, X_n \overset{iid}{\sim} p(\lambda)$، نشان دهید آماره‌ی $S = \sum_{i=1}^{n} X_i$ یک آماره‌ی بسنده برای λ است.

حل. از تعریف قبلا اثبات شد. اینجا از قضیه فاکتورگیری استفاده می‌کنیم.

$$\begin{aligned} f(\underline{x}, \lambda) &= \prod_{i=1}^{n} f(x_i, \lambda) \\ &= \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \\ &= \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \\ &= \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda} \times \frac{1}{\prod_{i=1}^{n} x_i!} = g(\sum_{i=1}^{n} x_i, \lambda)\, h(\underline{x}) \end{aligned}$$

لذا از قضیه فاکتورگیری، آماره بسنده λ می‌شود $S(\underline{X}) = \sum_{i=1}^{n} X_i$.

مثال ۶. اگر $X_1, \ldots, X_n \overset{iid}{\sim} b(1, p)$، آماره‌ی بسنده p را بیابید.

حل.

$$f(x,p) = p^x\,(1-p)^{1-x}, \quad x = 0, 1, \quad 0 < p < 1$$

$$\begin{aligned} f(\underline{x}, p) &= \prod_{i=1}^{n} f(x_i, p) \\ &= \prod_{i=1}^{n} p^{x_i}\,(1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^{n} x_i}\,(1-p)^{n-\sum_{i=1}^{n} x_i} \\ &= g(\sum_{i=1}^{n} x_i, p)\, h(\underline{x}) \end{aligned}$$

لذا آماره‌ی بسنده برای p می‌شود $S(\underline{X}) = \sum_{i=1}^{n} X_i$.

مثال ۷. اگر $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$، آماره‌ی بسنده را برای پارامتر (پارامترها) در هریک از حالات زیر به‌دست آورید.

الف) σ^2 معلوم باشد (آماره‌ی بسنده برای μ).
ب) μ معلوم باشد (آماره‌ی بسنده برای σ^2).
ج) هردو مجهول باشند (آماره‌ی بسنده برای هردو).

حل.

$$f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R}, \ \mu \in \mathbb{R}, \ \sigma > 0$$

$$\begin{aligned} f(\underline{x}, \mu, \sigma^2) &= \prod_{i=1}^{n} f(x_i, \mu, \sigma^2) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2}{2\sigma^2}} \end{aligned}$$

الف) اگر $\sigma^۲$ معلوم باشد مثلا $\sigma^۲ = \sigma_۰^۲$

$$f(\underline{x}, \mu) = \left(\frac{۱}{\sqrt{۲\pi}\sigma_۰}\right)^n e^{\frac{n\mu^۲ - ۲\mu \sum_{i=۱}^n x_i}{۲\sigma_۰^۲}} . e^{-\frac{\sum_{i=۱}^n x_i^۲}{۲\sigma_۰^۲}}$$
$$= g(\sum_{i=۱}^n x_i, \mu) \times h(\underline{x})$$

لذا در این حالت آماره‌ی بسنده برای μ می‌شود $S(\underline{X}) = \sum_{i=۱}^n X_i$.

ب) اگر μ معلوم باشد، مثلا $\mu = \mu_۰$، آنگاه

$$f(\underline{x}, \sigma^۲) = \left(\frac{۱}{\sqrt{۲\pi}\sigma}\right)^n e^{-\frac{۱}{۲\sigma^۲} \sum_{i=۱}^n (x_i - \mu_۰)^۲}$$
$$= g(\sum_{i=۱}^n (x_i - \mu_۰)^۲, \sigma^۲) \times ۱$$

لذا در این حالت آماره بسنده برای $\sigma^۲$ می‌شود $S(\underline{X}) = \sum_{i=۱}^n (X_i - \mu_۰)^۲$. در حالت خاص اگر $\mu = ۰$ آنگاه $S(\underline{X}) = \sum_{i=۱}^n X_i^۲$ آماره بسنده برای $\sigma^۲$ می‌باشد.

ج) اگر μ و $\sigma^{۲}$ هردو مجهول باشند، آنگاه

$$f(\underline{x},\mu,\sigma^{۲})=\left(\frac{۱}{\sqrt{۲\pi\sigma}}\right)^{n}e^{-\frac{\sum_{i=۱}^{n}x_i^{۲}-۲\mu\sum_{i=۱}^{n}x_i+n\mu^{۲}}{۲\sigma^{۲}}}$$
$$=g(\sum_{i=۱}^{n}x_i,\ \sum_{i=۱}^{n}x_i^{۲},\ \mu,\ \sigma^{۲})\times ۱$$

لذا در این حالت آماره بسنده توأم برای $(\mu,\sigma^{۲})$ می‌شود $S(\underline{X})=(\sum_{i=۱}^{n}X_i,\ \sum_{i=۱}^{n}X_i^{۲})$.

برای توزیع نرمال $N(\mu,\sigma^۲)$ نشان دادیم که آماره $\left(\sum_{i=۱}^{n} X_i,\ \sum_{i=۱}^{n} X_i^۲\right)$ آماره‌ی بسنده توأم برای $(\mu,\sigma^۲)$ است. می‌توان آماره‌ی بسنده توأم دیگری نیز معرفی کرد.

$$f(\underline{x},\mu,\sigma^۲)=(۲\pi\sigma^۲)^{-\frac{n}{۲}}\,e^{-\frac{۱}{۲\sigma^۲}\sum_{i=۱}^{n}(x_i-\mu)^۲}$$

می‌دانیم که

$$\sum_{i=۱}^{n}(x_i-\mu)^۲=\sum_{i=۱}^{n}(x_i-\bar{x})^۲+n(\bar{x}-\mu)^۲$$

یا

$$\sum_{i=۱}^{n}(x_i-\mu)^۲=n(\bar{x}-\mu)^۲+(n-۱)s^۲$$

$$s^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\bar{x})^۲$$

لذا چگالی توأم را می‌توان به‌صورت زیر نوشت:

$$\begin{aligned}f(\underline{x},\mu,\sigma^۲)&=(۲\pi\sigma^۲)^{-\frac{n}{۲}}\,e^{\frac{-n(\bar{x}-\mu)^۲}{۲\sigma^۲}-\frac{(n-۱)s^۲}{۲\sigma^۲}}\\&=g(\bar{x},s^۲,\mu,\sigma^۲)\,h(\underline{x})\end{aligned}$$

لذا آماره‌ی $(\bar{X}, S^۲)$ نیز یک آماره بسنده (توأم) برای $(\mu, \sigma^۲)$ است.

آماره‌ی بسنده دیگر $(\bar{X}, \bar{X^۲})$ می‌باشد، که $\bar{X^۲} = \frac{۱}{n}\sum_{i=۱}^{n} X_i^۲$.

مثال ۸. اگر $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$ که $\theta > 0$. آماره‌ی بسنده را برای θ پیدا کنید.
حل.

$$f(x, \theta) = \frac{1}{\theta}, \quad 0 < x < \theta$$

چگالی توأم نمونه تصادفی:

$$\begin{aligned} f(\underline{x}, \theta) &= \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \frac{1}{\theta}, \quad 0 < x_i < \theta, \quad i = 1, \ldots, n \\ &= \frac{1}{\theta^n}, \quad 0 < x_{(1)} < x_{(2)} < \cdots < x_{(n)} < \theta \\ &= \frac{1}{\theta^n}, \quad x_{(n)} < \theta \end{aligned}$$

لذا

$$f(\underline{x}, \theta) = \frac{1}{\theta^n} I_{(x_{(n)}, \infty)}(\theta) = g(x_{(n)}, \theta) \times h(\underline{x}).$$

لذا از قضیه فاکتورگیری آماره‌ی بسنده برای θ می‌شود $S = X_{(n)} = \max(X_1, \ldots, X_n)$.

مثال ۹. اگر $X_1, \ldots, X_n \overset{iid}{\sim} f(x,\mu,\sigma)$ که

$$f(x,\mu,\sigma) = \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}}, \quad x > \mu,\ \mu \in \mathbb{R},\ \sigma > 0.$$

آماره‌ی بسنده توأم را برای (μ,σ) به‌دست آورید.
حل. این توزیع، توزیع نمایی دوپارامتری است که اگر $\mu = 0$ باشد به توزیع نمایی تبدیل می‌شود.

$$f(x,\sigma) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}, \quad x > 0, \quad \sigma > 0$$

تابع چگالی توأم نمونه:

$$f(\underline{x},\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma} e^{-\frac{x_i-\mu}{\sigma}} \qquad x_i > \mu,\ i = 1,\ldots,n$$
$$= \frac{1}{\sigma^n} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma}} \qquad \mu < x_{(1)}$$

که $x_{(1)} = \min(x_1,\ldots,x_n)$.

لذا

$$
\begin{aligned}
f(\underline{x}, \mu, \sigma) &= \sigma^{-n} e^{-\frac{n(\bar{x}-\mu)}{\sigma}}\, I_{(-\infty, x_{(1)})(\mu)} \\
&= g(\bar{x}, x_{(1)}, \mu, \sigma) \times 1
\end{aligned}
$$

لذا آماره‌ی بسنده توأم برای (μ, σ) می‌شود:

$$S = (\bar{X}, X_{(1)}) \quad \text{یا} \quad S = \Big(\sum_{i=1}^{n} X_i,\ X_{(1)}\Big)$$

آماره‌ی بسنده در خانواده نمایی

فرض کنید $X_1, \dots, X_n$ نمونه‌ای تصادفی متعلق به خانواده نمایی با تابع چگالی احتمال زیر باشد:

$$f(x,\theta) = a(\theta)\, b(x)\, e^{c(\theta)\, d(x)},$$

آنگاه چگالی توأم نمونه به‌صورت زیر خواهد بود:

$$\begin{aligned} f(\underline{x},\theta) &= \prod_{i=1}^{n} f(x_i,\theta) = \prod_{i=1}^{n} a(\theta)\, b(x_i)\, e^{c(\theta)\, d(x_i)} \\ &= [a(\theta)]^n\, e^{c(\theta)\sum_{i=1}^{n} d(x_i)} \prod_{i=1}^{n} b(x_i) \\ &= g(\sum_{i=1}^{n} d(x_i),\theta) \times h(\underline{x}) \end{aligned}$$

که در آن $h(\underline{x}) = \prod_{i=1}^{n} b(x_i)$. لذا در خانواده نمایی تک پارامتری، آماره بسنده برای پارامتر θ می‌شود $S(\underline{X}) = \sum_{i=1}^{n} d(X_i)$.

آماره‌ی بسنده (توأم) در خانواده نمایی k پارامتری

فرض کنید $X_1, \dots, X_n$ نمونه‌ای تصادفی متعلق به خانواده نمایی k پارامتری با تابع چگالی احتمال زیر باشد:

$$f(x, \underline{\theta}) = a(\underline{\theta})\, b(x)\, e^{\sum_{i=1}^{k} c_i(\underline{\theta})\, d_i(x)},$$

چگالی توأم نمونه:

$$\begin{aligned} f(\underline{x}, \underline{\theta}) &= \prod_{i=1}^{n} f(x_i, \underline{\theta}) = \prod_{i=1}^{n} a(\underline{\theta})\, b(x_i)\, e^{\sum_{j=1}^{k} c_j(\underline{\theta})\, d_j(x_i)} \\ &= [a(\underline{\theta})]^n\, e^{c_1(\theta)\sum_{i=1}^{n} d_1(x_i) + \cdots + c_k(\theta)\sum_{i=1}^{n} d_k(x_i)} \prod_{i=1}^{n} b(x_i) \\ &= g\Big(\sum_{i=1}^{n} d_1(x_i), \dots, \sum_{i=1}^{n} d_k(x_i), \theta\Big) \times h(\underline{x}) \end{aligned}$$

لذا در خانواده نمایی k پارامتری آماره‌ی بسنده می‌شود

$$S(\underline{X}) = \Big(\sum_{i=1}^{n} d_1(X_i), \dots, \sum_{i=1}^{n} d_k(X_i)\Big).$$

مثال ۱۰. هرکدام از توزیع‌های زیر به خانواده توزیع‌های نمایی تک پارامتری تعلق دارد.

۱. $X \sim b(n,p)$ (توزیع دو جمله‌ای) با $d(x) = x$.
لذا آماره‌ی بسنده برای p می‌شود $S(\underline{X}) = \sum_{i=1}^{n} X_i$.

۲. $X \sim N(\mu, 1)$ (توزیع نرمال) با $d(x) = x$.
لذا آماره‌ی بسنده برای μ می‌شود $S(\underline{X}) = \sum_{i=1}^{n} X_i$.

۳. $X \sim N(0, \sigma^2)$ با $d(x) = x^2$.
لذا آماره‌ی بسنده برای σ^2 می‌شود $S(\underline{X}) = \sum_{i=1}^{n} X_i^2$.

مثال ۱۱. توزیع نرمال $N(\mu, \sigma^2)$ به توزیع نمایی دوپارامتری تعلق دارد با

$$d_1(x) = x, \qquad d_2(x) = x^2$$

لذا آماره‌ی بسنده توأم برای (μ, σ^2) در توزیع نرمال می‌شود

$$S = \Big(\sum_{i=1}^{n} d_1(X_i),\ \sum_{i=1}^{n} d_2(X_i)\Big) = \Big(\sum_{i=1}^{n} X_i,\ \sum_{i=1}^{n} X_i^2\Big).$$

رابطه بین MLE و آماره‌ی بسنده

اگر S یک آماره‌ی بسنده برای پارامتر θ، $\hat{\theta}$ برآورد ML (MLE) θ باشد که از حل معادله درست‌نمایی بدست می‌آید، آن‌گاه $\hat{\theta}$ تابعی از S خواهد بود.

اثبات. S یک آماره‌ی بسنده برای θ است، لذا از قضیه فاکتورگیری داریم:

$$f(\underline{x}, \theta) = g(s, \theta)\, h(\underline{x})$$

یا

$$L(\theta) = g(s, \theta)\, h(\underline{x}) \quad \text{(تابع درست‌نمایی نمونه)}$$

لگاریتم تابع درست‌نمایی نمونه:

$$\begin{aligned} l(\theta) = \ln L(\theta) &= \ln[g(s, \theta)\, h(\underline{x})] \\ &= \ln[g(s, \theta)] + \ln[h(\underline{x})] \end{aligned}$$

$$l'(\theta) = \frac{d}{d\theta} l(\theta) = \frac{d}{d\theta} \ln[g(s, \theta)]$$

می‌دانیم که MLE پارامتر θ، $\hat{\theta}$ از حل معادله

$$\frac{d}{d\theta} l(\theta) = ۰ \qquad \text{(معادله درستنمایی)}$$

به‌دست می‌آید. به عبارتی $\hat{\theta}$ از حل معادله

$$\frac{d}{d\theta} \ln g(s, \theta) = ۰$$

حاصل می‌شود و لذا $\hat{\theta}$ تابعی از S (آماره‌ی بسنده) خواهد بود.

مثال ۱۲. فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} Exp(\theta)$ با تابع چگالی احتمال زیر باشد.

$$f(x,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0,\ \theta > 0.$$

می‌دانیم خانواده فوق متعلق به خانواده نمایی است و $S(\underline{X}) = \sum_{i=1}^{n} X_i$ آماره‌ی بسنده برای θ می‌باشد.
برای به‌دست آوردن برآورد ML داریم:

$$L(\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \theta^{-n}\, e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

$$l(\theta) = \ln L(\theta) = -n \ln \theta - \frac{\sum_{i=1}^{n} x_i}{\theta}$$

$$l'(\theta) = \frac{-n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} = 0$$

$$\Rightarrow \quad \frac{n}{\theta} = \frac{\sum_{i=1}^{n} x_i}{\theta^2} \Rightarrow \theta = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\Rightarrow \quad \hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n}$$

لذا $\hat{\theta}$ تابعی از آماره‌ی بسنده است.

رابطه میان آماره‌ی بسنده و آماره‌ی دلخواه

اگر S یک آماره‌ی بسنده برای θ و $T = T(X_۱, \ldots, X_n)$ یک آماره‌ی دلخواه باشد، در اینصورت توزیع شرطی T به شرط $S = s$ به پارامتر θ وابسته نیست.

اثبات:

$$P(T \leq t|S = s) = P[T(X_۱, \ldots, X_n) \leq t|S = s]$$
$$= \int_A \cdots \int f_{X_۱,\ldots,X_n|S}(x_۱, \ldots, x_n|s)dx_۱ \ldots dx_n$$

که

$$A = \{(x_۱, \ldots, x_n) : T(x_۱, \ldots, x_n) \leq t\}.$$

چون S بسنده است لذا چگالی شرطی $f_{X_۱,\ldots,X_n|S}(x_۱, \ldots, x_n|s)$ به θ وابسته نیست و لذا عبارت فوق وابسته به θ نخواهد بود.

بسندگی تابعی از آماره بسنده

۱. اگر S یک آماره‌ی بسنده برای θ و S خود تابعی از آماره دیگر T باشد مثلا $S=\phi(T)$، آنگاه T نیز یک آماره‌ی بسنده برای θ خواهد بود.

اثبات. S برای θ بسنده است لذا

$$\begin{aligned} f(\underline{x},\theta) &= g(s,\theta)\,h(\underline{x}) \\ & g(\phi(t),\theta)\,h(\underline{x}) \\ &= g^*(t,\theta)\,h(\underline{x}) \end{aligned}$$

$$\Rightarrow \quad f(\underline{x},\theta) = g^*(t,\theta)\,h(\underline{x})$$

لذا از عکس قضیه فاکتورگیری آماره‌ی T برای θ بسنده است.

بسندگی تابعی از آماره بسنده

۲. اگر S برای θ بسنده و $U=\psi(S)$ یک تابع یک به یک از S باشد، آنگاه U نیز برای θ بسنده است.

اثبات.

$$U=\psi(S) \Leftrightarrow S=\psi^{-1}(U)$$

* بنا بر نکته قبل، چون S بسنده و تابعی از U است، لذا U نیز برای θ بسنده خواهد بود.

مثال ۱۳. اگر $X_۱, X_۲, X_۳ \overset{iid}{\sim} b(۱, p)$ نشان دهید که آماره‌ی $W = X_۱ + ۵X_۲ + X_۳$ برای p بسنده است.

حل. $S = X_۱ + X_۲ + X_۳$ آماره‌ی بسنده دلخواه برای p است. لذا کافی است نشان دهیم که S تابعی از W است.
فضای نمونه:

$$\chi = \{(۰, ۰, ۰)\ (۰, ۱, ۰)\ (۱, ۰, ۰)\ (۰, ۰, ۱)\ (۰, ۱, ۱)\ (۱, ۰, ۱)\ (۱, ۱, ۰)\ (۱, ۱, ۱)\}$$

(x_1, x_2, x_3)	W	S
(0,0,0)	0	0
(0,0,1)	1	1
(0,1,0)	5	1
(1,0,0)	1	1
(1,1,0)	6	2
(1,0,1)	2	2
(0,1,1)	6	2
(1,1,1)	7	3

از جدول مشاهده می‌شود که آماره‌ی بسنده S تابعی از آماره‌ی W است، مثلا $S = \phi(W)$ (چون برای هر W فقط یک S داریم، لذا W نیز یک آماره‌ی بسنده است).

توجه کنید که W تابعی از S نیست.

سؤال. در مثال قبل آماره‌های $S = X_۱ + X_۲ + X_۳$ و $W = X_۱ + ۵X_۲ + X_۳$ برای p بسنده‌اند. کدام آماره بهتر است؟

$$\text{تکیه‌گاه}\, S = \{۰, ۱, ۲, ۳\}$$
$$\text{تکیه‌گاه}\, W = \{۰, ۱, ۲, ۵, ۶, ۷\}$$

آماره‌ی S فضای مشاهدات نمونه‌ای χ را به ۴ کلاس زیر افراز می‌کند.

$$C_۰ = \{(۰, ۰, ۰)\} \qquad C_۱ = \{(۰, ۰, ۱)\,(۰, ۱, ۰)\,(۰, ۰, ۱)\}$$
$$C_۳ = \{(۱, ۱, ۱)\} \qquad C_۲ = \{(۱, ۱, ۰)\,(۰, ۱, ۱)\,(۱, ۰, ۱)\}$$

آماره‌ی W فضای مشاهدات نمونه‌ای χ را به ۶ کلاس زیر افراز می‌کند.

$$D_۰ = \{(۰, ۰, ۰)\} \qquad D_۱ = \{(۰, ۰, ۱)\,(۱, ۰, ۰)\}$$
$$D_۲ = \{(۱, ۰, ۱)\} \qquad D_۵ = \{(۰, ۱, ۰)\}$$
$$D_۶ = \{(۱, ۱, ۰)\,(۰, ۱, ۱)\} \qquad D_۷ = \{(۱, ۱, ۱)\}$$

بین دو آماره‌ی بسنده S و W آماره‌ای بهتر است که میزان فشردگی آن بیشتر باشد. به عبارتی آماره‌ای بهتر است که فضای مشاهدات نمونه‌ای را به کلاس‌های کمتری افراز کند. در مثال قبل S تابعی از W است و میزان فشردگی داده‌های نمونه‌ای برای S بیشتر است.

بعدا نشان خواهیم داد که در مثال فوق آماره S یک آماره‌ی بسنده مینیمال است. یعنی S آماره بسنده‌ای است که بیشترین فشردگی ممکن را روی داده‌ها اعمال می‌کند.

۱. فرض کنید $X_1, \ldots, X_n$ یک نمونه تصادفی با تابع چگالی احتمال $f(x, \theta)$ باشد.

الف) نشان دهید نمونه تصادفی $\underline{X} = (X_1, \ldots, X_n)$ یک آماره‌ی بسنده برای θ می‌باشد.
ب) نشان دهید آماره‌های ترتیبی $(X_{(1)}, \ldots, X_{(n)})$ یک آماره‌ی بسنده برای θ می‌باشد.
ج) نشان دهید نمونه تصادفی $(X_1, \ldots, X_{n-1})$ نمی‌تواند آماره‌ی بسنده θ باشد.

۲. در هرکدام از توزیع‌های زیر آماره‌ی بسنده پارامتر θ را به‌دست آورید.
الف) $X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta,\ \theta)$ (توزیع بتا با $\alpha = \beta = \theta$).
ب) $X_1, \ldots, X_n \overset{iid}{\sim} \Gamma(\theta,\ \theta)$ (توزیع گاما با $\alpha = \beta = \theta$).
ج) $X_1, \ldots, X_n \overset{iid}{\sim} U(\theta,\ \theta + 1)$ (توزیع یکنواخت پیوسته).
د) $X_1, \ldots, X_n \overset{iid}{\sim} f(x, \theta)$ که

$$f(x,\theta) = \frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}, \qquad x > 0,\ \ \theta > 0.$$

۳. اگر $X_1, X_2 \overset{iid}{\sim} p(\theta)$ باشد. نشان دهید آماره $T = X_1 + 2X_2$ نمی‌تواند آماره‌ی بسنده θ باشد.

۴. فرض کنید $X_1, \ldots, X_n \overset{iid}{\sim} f(x,,\theta)$ که

$$f(x,\theta) = \begin{cases} \frac{2x}{\theta}, & 0 < x < \theta \\ 0, & o.w. \end{cases}$$

الف) آیا خانواده فوق به خانواده نمایی تعلق دارد؟ بررسی کنید.
ب) آماره‌ی بسنده پارامتر θ را بیابید.

۵. اگر متغیر تصادفی X دارای تابع احتمال زیر باشد:

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad \theta > 0,$$

الف) آیا خانواده فوق به خانواده نمایی تعلق دارد؟ بررسی کنید.
ب) آماره‌ی بسنده پارامتر θ را بیابید.

۶. فرض کنید $X_1, \ldots, X_n \overset{iid}{\sim} f(x,\theta)$ که

$$f(x,\theta) = \theta\, x^{\theta-1}, \quad 0 < x < 1, \;\; \theta > 0,$$

الف) آماره‌ی بسنده θ را به‌دست آورید.
ب) آیا آماره‌ی $T(\underline{X}) = \sum_{i=1}^{n} \ln X_i$ می‌تواند آماره بسنده θ باشد؟ چرا؟

۷. فرض کنید $X_۱, \ldots, X_n \overset{iid}{\sim} N(\sigma,\ \sigma)$ که $\sigma > ۰$.

الف) آماره‌ی بسنده را برای σ به‌دست آورید.

ب) نشان دهید $(\bar{X}, S^۲)$ نیز یک آماره‌ی بسنده برای σ است.

۸. فرض کنید $X \sim N(۰, \sigma^۲)$ باشد. آیا آماره $T(X) = |X|$ یک آماره‌ی بسنده برای $\sigma^۲$ است؟ چرا؟

۹. اگر $X_۱, \ldots, X_n \overset{iid}{\sim} f(x,\theta)$ که

$$f(x,\theta) = \theta\,(۱-\theta)^{x-۱}, \quad x = ۱, ۲, \ldots, \quad ۰ < \theta < ۱,$$

الف) آیا خانواده فوق به خانواده نمایی تک پارامتری تعلق دارد؟ بررسی کنید.

ب) آماره‌ی بسنده پارامتر θ را به‌دست آورید.

۱۰. فرض کنید $X_۱, \ldots, X_n$ یک نمونه تصادفی با تابع چگالی زیر باشد.

$$f_{X_i}(x|\theta) = \begin{cases} e^{i\theta - x} & x \geq i\theta \\ ۰ & o.w \end{cases}$$

نشان دهید آماره‌ی $T = \min_i (X_i/i)$ یک آماره‌ی بسنده برای θ می‌باشد.