آمار ریاضی ۱
فصل چهارم: برآوردگر نااریب با کمترین واریانس بطور یکنواخت
بخش اول

استاد: دکتر اکبر اصغرزاده

تعریف

فرض کنید $f(x;\theta) \overset{iid}{\sim} X_1, \ldots, X_n$، $\theta \in \Theta$. برآوردگر $T^* = t(X_1, \ldots, X_n)$ را یک برآوردگر نااریب با کمترین واریانس بطور یکنواخت ($UMVUE^*$) برای $\alpha(\theta)$ گویند هرگاه:

۱. T^* برای $\alpha(\theta)$ نااریب باشد یعنی $E(T^*) = \alpha(\theta)$.

۲. برای هر برآوردگر نااریب دیگر T داشته باشیم:

$$Var_\theta(T^*) \leq Var_\theta(T), \quad \forall \theta \in \Theta.$$

تذکر. به عبارتی $UMVUE$ بهترین برآوردگر نااریب است.

*** UMVUE: Uniformly Minimum Variance Unbiased Estimator**

قضیه رائو-بلاکول

فرض کنید $X_1, \ldots, X_n \overset{iid}{\sim} f(x;\theta)$، $\theta \in \Theta$. همچنین فرض کنید که S یک آماره بسنده برای θ و T برآوردگر نااریب برای $\alpha(\theta)$ باشد. تعریف کنید $T^* = E(T|S)$.

۱) T^* یک آماره است و تابعی از آماره بسنده S می‌باشد.

۲) T^* یک برآوردگر نااریب برای $\alpha(\theta)$ است یعنی $E(T^*) = \alpha(\theta)$.

۳) $Var_\theta(T^*) \le Var_\theta(T), \quad \forall\, \theta \in \Theta$

۴) $Var_\theta(T^*) = Var_\theta(T) \;\Leftrightarrow\; Pr(T^* = T) = 1$

تذکر. در واقع این قضیه نشان می‌دهد که با شرطی کردن هر برآوردگر نااریب روی یک آماره بسنده می‌توان برآوردگر نااریب بهتر (با واریانس کمتر) پیدا کرد.

اثبات ۱. داریم $T^* = E(T|S)$، لذا T^* تابعی از S است و چون S برای θ بسنده است، بنابراین توزیع شرطی نمونه تصادفی $X_1, \ldots, X_n$ به شرط S و نیز توزیع شرطی هر آماره‌ی دیگر مانند T به شرط S وابسته به θ نیست. بنابراین $E(T|S)$ وابسته به θ نیست. لذا T^* یک آماره است (T^* یک آماره بسنده است).

اثبات ۲. از فرمول امید ریاضی مکرر یعنی $E(X) = E[E(X|Y)]$ داریم:

$$E(T^*) = E[E(T|S)] = E(T) = \alpha(\theta).$$

اثبات ۳. از فرمول بسط واریانس یعنی $Var(X) = Var[E(X|Y)] + E[Var(X|Y)]$ داریم:

$$Var(T) = Var[E(T|S)] + E[Var(T|S)].$$

لذا

$$Var(T) = Var(T^*) + \underbrace{E[Var(T|S)]}_{\geq 0} \geq Var(T^*)$$

$$\Rightarrow \; Var_\theta(T^*) \leq Var_\theta(T), \;\; \forall\, \theta \in \Theta.$$

اثبات ۴. از فرمول واریانس شرطی یعنی

$$Var(X|Y) = E\Big[\Big(X - E(X|Y)\Big)^{۲}|y\Big] = E(X^{۲}|y) - E^{۲}(X|y),$$

نتیجه می‌شود:

$$Var_\theta(T^*) = Var_\theta(T) \;\Leftrightarrow\; E[Var(T|S)] = ۰$$
$$\Leftrightarrow\; E\Big[E\Big(T - E(T|S)\Big)^{۲}|S\Big] = ۰$$
$$\Leftrightarrow\; E\Big(T - E(T|S)\Big)^{۲} = ۰ \Leftrightarrow\; E(T - T^*)^{۲} = ۰,$$
$$\Leftrightarrow\; Var(T - T^*) \overset{*}{=} ۰$$
$$\Leftrightarrow\; P\Big(T - T^* = E(T - T^*)\Big) \overset{**}{=} ۱$$
$$\Leftrightarrow\; P(T - T^* = ۰) = ۱$$
$$\Leftrightarrow\; P(T = T^*) = ۱.$$

$$* \, Var(T-T^*) = E(T-T^*)^2 - \underbrace{E^2(T-T^*)}_{0}$$

$$\Big(E(T-T^*) = E(T) - E(T^*) = \alpha(\theta) - \alpha(\theta) = 0\Big)$$

$$** \, Var(X) = 0 \;\Leftrightarrow\; P(X=\mu_x) = 1$$

مثال. اگر $X_1, \dots, X_n \overset{iid}{\sim} b(1,p), \quad 0 < p < 1$. می‌دانیم $S = \sum_{i=1}^{n} X_i$ یک آماره بسنده برای p و $T = X_1$ یک برآوردگر نااریب برای p است $\Big(E(T) = E(X_1) = p\Big)$.

از روی آماره بسنده S و برآوردگر نااریب T، یک برآوردگر نااریب دیگر با واریانس کمتر برای p بیابید.

حل. از قضیه رائو-بلاکول برآوردگر نااریب با واریانس کمتر برای p می‌شود:

$$T^* = E(T|S) = E(X_1|S).$$

داریم

$$\begin{aligned}
E(X_1|S=s) &= \sum_{x_1=0,1} x_1\, f_{X_1|S}(x_1|s) = 0.f_{X_1|S}(0|s) + 1.f_{X_1|S}(1|s)\\
&= f_{X_1|S}(1|s) = P(X_1=1|S=s)\\
&= \frac{P(X_1=1, S=s)}{P(S=s)}\\
&= \frac{P(X_1=1, \sum_{i=1}^n X_i = s)}{P(\sum_{i=1}^n X_i = s)}\\
&= \frac{P(X_1=1, \sum_{i=2}^n X_i = s-1)}{P(\sum_{i=1}^n X_i = s)}\\
&= \frac{P(X_1=1)\, P(\sum_{i=2}^n X_i = s-1)}{P(\sum_{i=1}^n X_i = s)}.
\end{aligned}$$

چون $\sum_{i=1}^{n} X_i \sim b(n,p)$ و $\sum_{i=2}^{n} X_i \sim b(n-1,p)$ بنابراین نتیجه می‌شود:

$$E(X_1|S=s) = \frac{p\binom{n-1}{s-1}p^{s-1}(1-p)^{n-s}}{\binom{n}{s}p^s\,(1-p)^{n-s}} = \frac{\binom{n-1}{s-1}}{\binom{n}{s}} = \frac{\frac{(n-1)!}{(s-1)!(n-s)!}}{\frac{n!}{s!(n-s)!}} = \frac{s}{n}.$$

$$\Rightarrow\ E(X_1|S=s) = \frac{s}{n}\ \Rightarrow\ T^* = E(T|S) = \frac{S}{n}.$$

لذا برآوردگر نااریب با واریانس کمتر برای p می‌شود $T^* = \frac{S}{n}$.

در ادامه نشان می‌دهیم که $Var(T^*)$ واقعا از $Var(T)$ کمتر است یعنی

$$Var(T^*) \leq Var(T), \quad \forall T.$$

داریم

$$Var(T^*) = Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n}$$
$$Var(T) = Var(X_1) = p(1-p).$$

چون $\frac{p(1-p)}{n} \leq p(1-p), \quad \forall\, n \geq 1$، لذا $Var(T^*) \leq Var(T)$.

مثال. اگر $\circ < p < ۱$ $X_۱, \ldots, X_n \stackrel{iid}{\sim} b(۱, p),$.

الف) نشان دهید که $T = X_۱ - X_۱ X_۲$ یک برآوردگر نااریب برای $\alpha(p) = p(۱ - p)$ است.

ب) از روی برآوردگر نااریب T و آماره بسنده $S = \sum_{i=۱}^{n} X_i$، یک برآوردگر نااریب دیگر برای $\alpha(p)$ با واریانس کمتر به‌دست آورید.

حل. الف)

$$\begin{aligned} E(T) = E(X_۱ - X_۱ X_۲) &= E(X_۱) - E(X_۱ X_۲) \\ &= E(X_۱) - E(X_۱)E(X_۲) \\ &= p - p.p \\ &= p(۱ - p) \\ &= \alpha(p). \end{aligned}$$

یعنی T برای $\alpha(p)$ نااریب است.

ب) برآوردگر ناریب با واریانس کمتر برای $\alpha(p)$ می‌شود $T^* = E(T|S)$.

برد T عبارت است از $S_T = \{۰, ۱\}$. لذا $E(T|S)$ به‌صورت زیر محاسبه می‌شود.

$$
\begin{aligned}
E(T|S=s) &= \sum_{t=۰}^{۱} t\, f_{T|S}(t|s) = f_{T|S}(۱|s) = P(T=۱|S=s)\\
&= P(X_۱=۱,\, X_۲=۰|S=s)\\
&= \frac{P(X_۱=۱\, X_۲=۰, \sum_{i=۱}^{n} X_i = s)}{P(\sum_{i=۱}^{n} X_i = s)}\\
&= \frac{P(X_۱=۱\, X_۲=۰, \sum_{i=۳}^{n} X_i = s-۱)}{P(\sum_{i=۱}^{n} X_i = s)}\\
&= \frac{P(X_۱=۱)\, P(X_۲=۰)\, P(\sum_{i=۳}^{n} X_i = s-۱)}{P(\sum_{i=۱}^{n} X_i = s)} \qquad \Big(\sum_{i=۳}^{n} X_i \sim b(n-۲,p)\Big)\\
&= \frac{p(۱-p)\binom{n-۲}{s-۱} p^{s-۱}(۱-p)^{n-s-۱}}{\binom{s}{n} p^s\,(۱-p)^{n-s}}\\
&= \frac{\frac{(n-۲)!}{(s-۱)!\,(n-s-۱)!}}{\frac{n!}{s!(n-s)!}} = \frac{s(n-s)}{n(n-۱)}.
\end{aligned}
$$

لذا برآوردگر نااریب با واریانس کمتر برای p می‌شود $T^* = \frac{S(n-S)}{n(n-۱)}$.

تذکر. برآوردگر نااریب با واریانس کمتری که از قضیه رائو-بلاکول به‌دست می‌آید لزوماً یکتا نیست مگر اینکه آماره بسنده S آماره کامل هم باشد.