# آمار ریاضی ۱
# فصل دوم: آماره بسنده مینیمال

استاد: دکتر اکبر اصغرزاده

دانشگاه مازندران

### آماره بسنده مینیمال

آماره بسنده مینیمال آماره بسنده‌ای است که بیشترین فشردگی را روی فضای مشاهدات نمونه‌ای $\chi$ اعمال می‌کند (فضای مشاهدات نمونه‌ای را به کمترین کلاس ممکن افراز می‌کند).

### تعریف آماره بسنده مینیمال

آماره $S(X)$ را یک آماره بسنده مینیمال برای $\theta$ گویند هرگاه:

الف) $S(X)$ برای $\theta$ آماره بسنده باشد.

ب) تابعی از هر آماره بسنده دیگر باشد.

* نکته: هر تابع یک به یک از آماره بسنده مینیمال، یک آماره بسنده مینیمال است.

### قضیه ۱

اگر $X_1, \ldots, X_n \overset{iid}{\sim} f(x, \theta)$، $\theta \in \Theta$. آماره $S(\underline{X})$ برای $\theta$ بسنده مینیمال است اگر و فقط اگر به ازای هر $\underline{x}$ و $\underline{y}$ متعلق به $\chi$ که $\chi$ فضای مشاهدات نمونه‌ای است، داشته باشیم:

$$\frac{f(\underline{x}, \theta)}{f(\underline{y}, \theta)} \text{ مستقل از } \theta \text{ باشد} \Leftrightarrow S(\underline{x}) = S(\underline{y})$$

مثال ۱. اگر $X_۱,\dots,X_n \overset{iid}{\sim} b(۱,p),\ ۰<p<۱$، نشان دهید که آماره $S(\underline{X})=\sum_{i=۱}^{n}X_i$ یک آماره بسنده مینیمال برای $p$ است.

$$f(\underline{x},p)=\prod_{i=۱}^{n}f(x_i,p)=\prod_{i=۱}^{n}p^{x_i}\,(۱-p)^{۱-x_i}=p^{\sum_{i=۱}^{n}x_i}\,(۱-p)^{n-\sum_{i=۱}^{n}x_i}$$

$$\forall\ \underline{x},\ \underline{y}\in\chi:\ \frac{f(\underline{x},p)}{f(\underline{y},p)}=\frac{p^{\sum_{i=۱}^{n}x_i}\,(۱-p)^{n-\sum_{i=۱}^{n}x_i}}{p^{\sum_{i=۱}^{n}y_i}\,(۱-p)^{n-\sum_{i=۱}^{n}y_i}}$$
$$=p^{\sum_{i=۱}^{n}x_i-\sum_{i=۱}^{n}y_i}\,(۱-p)^{\sum_{i=۱}^{n}y_i-\sum_{i=۱}^{n}x_i}$$

کسر فوق وابسته به $p$ نیست اگر و فقط اگر داشته باشیم $\sum_{i=۱}^{n}x_i=\sum_{i=۱}^{n}y_i$.
لذا آماره $S(\underline{X})=\sum_{i=۱}^{n}X_i$ برای $p$ بسنده مینیمال است.

مثال ۲. اگر $X_1, \ldots, X_n \overset{iid}{\sim} f(x,\theta)$ که $f(x,\theta) = \theta\, x^{\theta-1},\ 0 < x < 1,\ \theta > 0$. آماره بسنده مینیمال را برای $\theta$ پیدا کنید.

$$f(\underline{x},\theta) = \prod_{i=1}^{n} \theta\, x_i^{\theta-1} = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1},$$

$$\forall\ \underline{x},\ \underline{y} \in \chi:\ \frac{f(\underline{x},\theta)}{f(\underline{y},\theta)} = \frac{\theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}}{\theta^n \left(\prod_{i=1}^{n} y_i\right)^{\theta-1}} = \left(\frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} y_i}\right)^{\theta-1}.$$

کسر فوق وابسته به $\theta$ نیست اگر و فقط اگر $\prod_{i=1}^{n} y_i = \prod_{i=1}^{n} x_i$. لذا آماره بسنده مینیمال برای $\theta$ می‌شود $S(\underline{X}) = \prod_{i=1}^{n} X_i$.

* نکته: آماره $T = \sum_{i=1}^{n} \ln X_i$ نیز یک آماره بسنده مینیمال برای $\theta$ است. چون $T$ تابعی یک به یک از $S$ است و داریم:

$$T = \ln S = \ln(\prod_{i=1}^{n} X_i) = \sum_{i=1}^{n} \ln X_i.$$

مثال ۳. اگر $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$، آماره بسنده مینیمال $(\mu, \sigma^2)$ را پیدا کنید.

چگالی توأم نمونه:

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

$$
\forall\, \underline{x},\ \underline{y} \in \chi :\ \frac{f(\underline{x}, \mu, \sigma^2)}{f(\underline{y}, \mu, \sigma^2)} = \frac{f(\underline{x}, \mu, \sigma^2)}{f(\underline{y}, \mu, \sigma^2)} = \frac{(2\pi\sigma^2)^{-\frac{n}{2}}\, e^{\frac{-n(\bar{x}-\mu)^2}{2\sigma^2} - \frac{(n-1)s_x^2}{2\sigma^2}}}{(2\pi\sigma^2)^{-\frac{n}{2}}\, e^{\frac{-n(\bar{x}-\mu)^2}{2\sigma^2} - \frac{(n-1)s_y^2}{2\sigma^2}}},
$$

که مستقل از پارامترهای $\mu$ و $\sigma^2$ است اگر و تنها اگر داشته باشیم:

$$
\bar{x} = \bar{y}, \quad s_x^2 = s_y^2,
$$

یا $(\bar{x}, s_x^2) = (\bar{y}, s_y^2)$. لذا آماره بسنده مینیمال می‌شود $(\bar{x},\ s^2)$.

همچنین داریم

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

$$\forall\, \underline{x},\ \underline{y} \in \chi:\ \frac{f(\underline{x}, \mu, \sigma^۲)}{f(\underline{y}, \mu, \sigma^۲)} = e^{\frac{۱}{۲\sigma^۲}(\sum_{i=۱}^{n} x_i^۲ - \sum_{i=۱}^{n} y_i^۲)}\, e^{-\frac{\mu}{\sigma^۲}(\sum_{i=۱}^{n} x_i - \sum_{i=۱}^{n} y_i)},$$

که مستقل از $\mu$ و $\sigma^۲$ است اگر داشته باشیم:

$$\sum_{i=۱}^{n} x_i = \sum_{i=۱}^{n} y_i,\quad \sum_{i=۱}^{n} x_i^۲ = \sum_{i=۱}^{n} y_i^۲,$$

یا $(\sum_{i=۱}^{n} x_i,\ \sum_{i=۱}^{n} x_i^۲) = (\sum_{i=۱}^{n} y_i,\ \sum_{i=۱}^{n} y_i^۲)$. پس آماره بسنده مینیمال دیگر می‌شود:

$$(\sum_{i=۱}^{n} X_i,\ \sum_{i=۱}^{n} X_i^۲) \quad \text{یا} \quad (\bar{X},\ \bar{X^۲}).$$

مثال ۴. اگر $X_۱, \ldots, X_n \overset{iid}{\sim} f(x, \alpha, \beta)$ باشد که

$$f(x, \alpha, \beta) = \alpha\, \beta^{\alpha}\, x^{-\alpha-۱}, \quad x \geq \beta,\ \alpha,\ \beta > ۰.$$

آماره بسنده مینیمال برای $(\alpha,\ \beta)$ را به‌دست آورید.

حل. چگالی توأم نمونه تصادفی

$$f(\underline{x}, \alpha, \beta) = \prod_{i=۱}^{n} f(x_i, \alpha, \beta) = \prod_{i=۱}^{n} \alpha\, \beta^{\alpha}\, x_i^{-\alpha-۱}$$

$$= \alpha^{n}\, \beta^{n\alpha} \left( \prod_{i=۱}^{n} x_i \right)^{-\alpha-۱}, \quad x_i \geq \beta, \quad i = ۱, \ldots, n.$$

لذا

$$
\begin{aligned}
f(\underline{x}, \alpha, \beta) &= \alpha^n \beta^{n\alpha} \Big(\prod_{i=۱}^{n} x_i\Big)^{-\alpha-۱}, \ \ x_{(۱)} \geq \beta \\
&= \alpha^n \beta^{n\alpha} \Big(\prod_{i=۱}^{n} x_i\Big)^{-\alpha-۱} I_{(۰,\ x_{(۱)})}(\beta) \\
&= \alpha^n \beta^{n\alpha} \Big(\prod_{i=۱}^{n} x_i\Big)^{-\alpha} I_{(۰,\ x_{(۱)})}(\beta) \Big(\prod_{i=۱}^{n} x_i\Big)^{-۱} \\
&= g(\alpha, \beta, \prod_{i=۱}^{n} x_i,\ x_{(۱)})\, h(\underline{x}).
\end{aligned}
$$

از قضیه فاکتورگیری آماره بسنده برای $(\alpha, \beta)$ می‌شود:

$$
S = (\prod_{i=۱}^{n} X_i,\ X_{(۱)}).
$$

آماره بسنده مینیمال:

$$\forall\ \underline{x},\ \underline{y} \in \chi:\ \frac{f(\underline{x},\alpha,\beta)}{f(\underline{y},\alpha,\beta)} = \frac{\alpha^n\,\beta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-\alpha-1} I_{(0,\ x_{(1)})}(\beta)}{\alpha^n\,\beta^{n\alpha}\left(\prod_{i=1}^n y_i\right)^{-\alpha-1} I_{(0,\ y_{(1)})}(\beta)}$$

$$= \left(\frac{\prod_{i=1}^n x_i}{\prod_{i=1}^n y_i}\right)^{-\alpha-1} \frac{I_{(0,\ x_{(1)})}(\beta)}{I_{(0,\ y_{(1)})}(\beta)},$$

که مستقل از $\alpha$ و $\beta$ است اگر و فقط اگر :

$$\prod_{i=1}^n x_i = \prod_{i=1}^n y_i,\quad x_{(1)} = y_{(1)},$$

یا

$$(\prod_{i=1}^n x_i,\ x_{(1)}) = (\prod_{i=1}^n y_i,\ y_{(1)}).$$

لذا آماره بسنده مینیمال برای $(\alpha,\beta)$ می‌شود:

$$S = (\prod_{i=1}^n X_i,\ X_{(1)}).$$

## آماره بسنده مینیمال در خانواده نمایی

$$X_1, \ldots, X_n \overset{iid}{\sim} f(x,\theta), \quad \theta \in \Theta \subseteq \mathbb{R}^1$$

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

$$\forall\ \underline{x},\ \underline{y};\ \frac{f(\underline{x},\theta)}{f(\underline{y},\theta)} = \frac{[a(\theta)]^n \left(\prod_{i=1}^n b(x_i)\right) e^{c(\theta) \sum_{i=1}^n d(x_i)}}{[a(\theta)]^n \left(\prod_{i=1}^n b(y_i)\right) e^{c(\theta) \sum_{i=1}^n d(y_i)}}$$

$$= \frac{\prod_{i=1}^n b(x_i)}{\prod_{i=1}^n b(y_i)}\, e^{c(\theta)\left[\sum_{i=1}^n d(x_i) - \sum_{i=1}^n d(y_i)\right]},$$

که مستقل از پارامتر $\theta$ است اگر و فقط اگر $\sum_{i=1}^n d(y_i) = \sum_{i=1}^n d(x_i)$.
لذا در خانواده نمایی تک پارامتری، آماره

$$S = \sum_{i=1}^n d(X_i),$$

یک آماره بسنده مینیمال برای $\theta$ است.

تذکر. می‌توان نشان داد که در خانواده نمایی $k$ پارامتری آماره بسنده مینیمال برای $\underline{\theta} = (\theta_1, \ldots, \theta_n)$ می‌شود:

$$(\sum_{i=1}^{n} d_1(x_i), \ldots, \sum_{i=1}^{n} d_k(x_i)).$$

تمرین. از خاصیت فوق استفاده کرده و آماره بسنده مینیمال را برای پارامتر یا پارامترهای مجهول در هریک از توزیع‌های زیر بیابید.

الف) $p(\lambda)$ توزیع پواسن
ب) $Ge(p)$ توزیع هندسی
ج) $N(\mu, \sigma^2)$ توزیع نرمال
د) $\Gamma(\alpha,\ \beta)$ توزیع گاما
ه) $Beta(\alpha,\ \beta)$ توزیع بتا

۱. در هریک از توزیع‌های زیر، بر اساس نمونه تصادفی $X_1, \ldots, X_n$ آماره بسنده مینیمال $\theta$ را پیدا کنید.

$$a)\ f(x|\theta) = \frac{2x}{\theta^2},\ \ 0 < x < \theta,\ \ \theta > 0,$$

$$b)\ f(x|\theta) = \frac{\theta}{(1+x)^{1+\theta}},\ \ 0 < x < \infty,\ \ \theta > 0,$$

$$c)\ f(x|\theta) = e^{-(x-\theta)}\ \exp(e^{-(x-\theta)}),\ \ -\infty < x < \infty,\ \ -\infty < \theta < \infty,$$

$$d)\ f(x|\theta) = \frac{1}{\pi[1+(x-\theta)^2]},\ \ -\infty < x < \infty,\ \ -\infty < \theta < \infty,$$

$$e)\ f(x|\theta) = \frac{1}{2}e^{-|x-\theta|},\ \ -\infty < x < \infty,\ \ -\infty < \theta < \infty.$$



۲. فرض کنید $X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta,\ \theta)$ که $\theta > 0$.

آماره بسنده مینیمال $\theta$ را بیابید.

۳. اگر $X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta,\ 1)$ با تابع چگالی احتمال زیر باشد.

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

الف) آماره بسنده مینیمال $\theta$ را بیابید.
ب) آیا $\sum_{i=1}^{n} \ln X_i$ آماره بسنده مینیمال $\theta$ است؟ چرا؟

۴. فرض کنید $X_1, \ldots, X_n \overset{iid}{\sim} U(\theta,\ \theta+1)$. آماره بسنده مینیمال $\theta$ را بیابید.