فصل ۲

برآورد نقطه ای

فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی n تایی با مقادیر مشاهده شده $x_1, x_2, \ldots, x_n$ از جامعه آماری $X_1, X_2, \ldots, X_N$ با توزیع احتمال $f_\theta(x)$ است که این توزیع احتمال به پارامتر مجهول θ وابسته است. به عبارت دیگر

$X_1, X_2, \ldots, X_N \sim f_\theta(x)$: جامعه آماری

$X_1, X_2, \ldots, X_n \sim f_\theta(x)$: نمونه تصادفی

θ: پارامتر نامعلوم

هدف از برآوردیابی، یافتن کمیتی از مقادیر $x_1, x_2, \ldots, x_n$ به عنوان تقریبی (یا برآوردی) از پارامتر مجهول θ است.

در روش **برآورد نقطه ای**، براساس مقادیر مشاهده شده نمونه تصادفی $x_1, x_2, \ldots, x_n$، تابعی را به عنوان تقریبی از پارامتر مجهول جامعه ارائه می دهیم.

$X_1, X_2, \ldots, X_N \sim f_\theta(x)$: جامعه آماری

$X_1, X_2, \ldots, X_n \sim f_\theta(x)$: نمونه تصادفی

$T = g(X_1, X_2, \ldots, X_n)$: برآوردگر (تابعی از نمونه تصادفی بدون عنصر مجهول)

$t = g(x_1, x_2, \ldots, x_n)$: برآورد (مقدار مشاهده شده برآوردگر)

مثال ۲-۱. آزمایش بررسی قد افراد یک اداره را درنظر بگیرید.

$X_۱, X_۲, \ldots, X_N \sim N(\mu, ۱۰۰)$: جامعه آماری

$X_۱, X_۲, X_۳, X_۴ \sim N(\mu, ۱۰۰)$: نمونه تصادفی ۴تایی

$T = g(X_۱, X_۲, X_۳, X_۴) = \overline{X} = \frac{۱}{۴}\sum_{i=۱}^{۴} X_i$: برآوردگر

اگر مقادیر مشاهده شده نمونه به صورت $x_۱ = ۱۷۲$، $x_۲ = ۱۶۸$، $x_۳ = ۱۶۵$ و $x_۴ = ۱۷۵$ باشد، آن گاه مقدار مشاهده شده برآوردگر $\overline{X}$ برابر است با

$$\overline{x} = \frac{۱}{۴}\sum_{i=۱}^{۴} x_i = ۱۷۰$$

بنابراین با استفاده از نمونه فوق، میانگین قد افراد را ۱۷۰ تخمین می زنیم.

مثال ۲-۲. آزمایش انتخاب یک عدد تصادفی در بازه (a, b) را در نظر بگیرید.

$X_۱, X_۲, \ldots, X_N \sim U(a, b)$: جامعه آماری

$X_۱, X_۲, X_۳, X_۴, X_۵ \sim U(a, b)$: نمونه تصادفی ۵تایی

$\boldsymbol{\theta} = (a, b)$

$T = g(X_۱, X_۲, X_۳, X_۴, X_۵) = (\min_{i\in\{۱,۲,۳,۴,۵\}} X_i, \max_{i\in\{۱,۲,۳,۴,۵\}} X_i)$: برآوردگر

اگر مقادیر مشاهده شده نمونه به صورت $x_۱ = ۱/۷$، $x_۲ = ۱/۴$، $x_۳ = ۱/۹$، $x_۴ = ۱/۳$ و $x_۵ = ۱/۲$ باشد، آن گاه مقدار مشاهده شده برآوردگر T برابر است با

$$t = (۱/۲, ۱/۹)$$

بنابراین با استفاده از نمونه فوق، مقدار a با ۱/۲ و مقدار b با ۱/۹ تخمین زده می شود.

تمرین ۲-۳. در مثال ۲-۲ برآوردگرهای دیگری برای θ بیان کرده و مقدار آن ها را برای نمونه موجود بدست آورید.

۱.۲ روش های برآوردیابی

روش های مختلفی برای برآورد نقطه ای یک پارامتر مجهول وجود دارد. در اینجا دو تا از معروف ترین این روش ها را شرح می دهیم.

۱.۱.۲ روش گشتاوری

این روش یکی از قدیمی ترین روش های برآوردیابی است که در سال ۱۸۹۴ توسط کارل پیرسن، آماردان انگلیسی استفاده شد. روش برآورد گشتاوری که به روش MM معروف است مبتنی بر بکارگیری گشتاورهای توزیع و گشتاورهای نمونه ای می باشد. در حالت کلی، فرض کنید

$X_1, X_2, \ldots, X_N \sim f_{\boldsymbol{\theta}}(x)$: جامعه آماری

$X_1, X_2, \ldots, X_n \sim f_{\boldsymbol{\theta}}(x)$: نمونه تصادفی

$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k) \in \Theta \subseteq \mathbf{R}^k$: پارامتر نامعلوم

بطوری که k گشتاور اول این توزیع وجود دارند. طبق تعریف داریم

$$\mu_r = \mathrm{E}_{\boldsymbol{\theta}}(X^r) \qquad r = 1, 2, \ldots, k$$

که تابعی از $\boldsymbol{\theta}$ است. اگر r–امین گشتاور نمونه ای بصورت زیر تعریف شود

$$M_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r = \overline{X^r} \qquad r = 1, 2, \ldots, k$$

آن گاه برآورد MM پارامترهای مجهول $\theta_1, \theta_2, \ldots, \theta_k$ از تشکیل و حل k معادله به صورت

$$\mu_r = M_r \qquad r = 1, 2, \ldots, k \tag{۱-۲}$$

بدست می آیند.

توجه داشته باشید که

- برآورد گشتاوری پارامتر θ لزوما یکتا نیست و معمولا به صورت $\widetilde{\theta}$ نشان داده می شود.
- برآورد MM برای $\theta = \phi(\mu_1, \mu_2, \ldots, \mu_r)$ به صورت $\widetilde{\theta} = \phi(M_1, M_2, \ldots, M_r)$ است.
- یک روش برای برآورد تابعی از پارامترهای مجهول به صورت $\gamma(\theta_1, \theta_2, \ldots, \theta_k)$ استفاده از تابع $\gamma(\widetilde{\theta}_1, \widetilde{\theta}_2, \ldots, \widetilde{\theta}_k)$ است.
- طبق قانون اعداد بزرگ، می دانیم $M_r \xrightarrow[\mathrm{P}]{} \mu_r$. درواقع، M_rها سازگار هستند.

توجه داشته باشید که هدف از سازگاری برآوردگر T_n بررسی رفتار حدی دنباله $\{T_n\}_{n=1}^{\infty}$ از برآوردگرها است. دنباله برآوردگرهای $\{T_n\}$ برای $\gamma(\theta)$ سازگار است اگر وقتی n سمت بی نهایت میل کند داشته باشیم

$$T_n \xrightarrow[\mathrm{P}]{} \gamma(\theta)$$

و این معادل با این است که برای هر $\epsilon > 0$ داریم

$$\lim_{n\to\infty} \mathrm{P}(|T_n - \gamma(\theta)| > \epsilon) = 0$$

.

- می توان نشان داد که اگر رابطه بین $\theta_1, \theta_2, \dots, \theta_r$ و $\mu_1, \mu_2, \dots, \mu_r$ یک به یک باشد و برای $r = 1, 2, \dots, k$، $\theta_r = \phi_r(\mu_1, \mu_2, \dots, \mu_k)$ توابعی پیوسته از $\mu_1, \mu_2, \dots, \mu_k$ باشند آن گاه

$$\widetilde{\theta}_r = \phi_r(M_1, M_2, \dots, M_k) \qquad r = 1, 2, \dots, k$$

جواب معادلات گشتاوری در (۲-۱) بوده و به ازای $r = 1, 2, \dots, k$، برآوردگر $\widetilde{\theta}_r$ برای θ_r سازگار است.

- این روش در حالتی که گشتاورهای μ_r وجود نداشته باشند، کارایی ندارد.

با استفاده از برآوردگر گشتاورهای مراتب مختلف می توان پارامترهای مجهول را برآورد کرد.

مثال ۲-۴. فرض کنید $X \sim N(0, \theta)$. با استفاده از یک نمونه تصادفی n تایی برآورد گشتاوری θ را بیابید.

حل: داریم $\mu = \mathrm{E}(X) = 0$ و $\theta = \mathrm{E}(X^2) = \mu_2$. پس با استفاده از روش گشتاوری داریم

$$\widetilde{\theta} = \widetilde{\mu}_2 = M_2 = \overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

مثال ۲-۵. فرض کنید $X \sim N(\mu, \sigma^2)$. با استفاده از یک نمونه تصادفی n تایی برآورد گشتاوری $\theta = (\mu.\sigma^2)$ را بیابید.

حل: داریم $\mu = \mathrm{E}(X)$ و $\sigma^2 = \mathrm{E}(X^2) - \mu^2 = \mu_2 - \mu^2$. پس با استفاده از روش گشتاوری بدست می آوریم

$$\mu_1 = \mathrm{E}(X) = \mu \quad \Longrightarrow \widetilde{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
$$\mu_2 = \mathrm{E}(X^2) = \mu^2 + \sigma^2 \quad \Longrightarrow \widetilde{\sigma^2} = \widetilde{\mu_2} - \widetilde{\mu}^2 = \overline{X^2} - \overline{X}^2.$$

مثال ۲-۶. فرض کنید X دارای توزیعی با تابع چگالی زیر باشد

$$f_\theta(x) = \theta x^{\theta - 1} \qquad 0 < x < 1,\ \theta > 0.$$

با استفاده از یک نمونه تصادفی n تایی برآورد گشتاوری $\frac{\theta}{\theta+1}$ و θ را بیابید.

حل: داریم

$$\mu = \mathrm{E}(X) = \int_0^1 \theta x^{\theta} \mathrm{dx} = \frac{\theta}{\theta+1}.$$

از طرفی برآورد MM برای μ به صورت $M_1 = \overline{X}$ است و بنابراین

$$\overline{X} = \widetilde{\frac{\theta}{\theta+1}} \quad \Longrightarrow \quad \widetilde{\theta} = \frac{\overline{X}}{1-\overline{X}}$$

تمرین ۲-۷. اگر $X \sim E(\mu)$ با استفاده از یک نمونه n تایی از X برآورد گشتاوری پارامتر μ را بدست آورید.

تمرین ۲-۸. اگر $X \sim P(\lambda)$ با استفاده از یک نمونه n تایی از X برآورد گشتاوری پارامتر λ را بدست آورید.

۲.۱.۲ روش درستنمایی ماکزیمم

یکی از قدیمی ترین و پراهمیت ترین روش های برآوردیابی نقطه ای روش درستنمایی ماکزیمم است که به اختصار روش ML نامیده می شود. این روش که متداول ترین روش در بین استفاده کنندگان علمی از آمار است اولین بار توسط گوس در سال ۱۸۲۱ بکارگرفته شد و پس از آن در سال ۱۹۲۵ به صورت گسترده تری توسط فیشر در انتقاد از روش گشتاوری مورد استفاده قرارگرفت. البته روش ML بهترین روش برآوردیابی نیست.

روش ML مبتنی بر استفاده از تابع درستنمایی است.
فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از متغیرتصادفی X با چگالی $f_{\boldsymbol{\theta}}(x)$ باشد. همچنین $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k) \in \Theta \subseteq \mathbf{R}^k$. اگر $x_1, x_2, \ldots, x_n$ یافته های این نمونه تصادفی باشد، با توجه به تعریف نمونه تصادفی چگالی هر X_i، $i = 1, 2, \ldots, n$، به صورت $f_{\boldsymbol{\theta}}(x_i)$ است و چگالی توام X_iها برابر است با

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_{\boldsymbol{\theta}}(x_i)$$

که به آن تابع درستنمایی برای یافته های نمونه تصادفی می گویند. تابع درستنمایی را با $L(\boldsymbol{\theta}, \mathrm{x})$ نیز نشان می دهند.

دقت کنید که هر چه اندازه $L(\boldsymbol{\theta})$ بیشتر باشد، شانس مشاهده $x_1, x_2, \ldots, x_n$ بیشتر است. بنابراین، در یک روش برآوردیابی براساس تابع درستنمایی،

برآورد $\boldsymbol{\theta}$ را در Θ طوری تعیین می کنیم که $L(\theta)$ بیشترین باشد.

توجه داشته باشید که

- تابع درستنمایی $L(\theta)$ لزوما مشتق پذیر نیست.
- تابع درستنمایی $L(\theta)$ یک تابع چگالی احتمال نیست.

تعریف ۲-۹. فرض کنید $T(\mathbf{X})$ برآوردگری برای θ باشد به طوری که برای هر $\theta \in \Theta$، $\mathrm{P}(T(\mathbf{X}) \in \Theta) = ۱$. اگر برای هر $\theta \in \Theta$، داشته باشیم

$$L(T(\mathbf{X})) \geq L(\theta)$$

آن گاه $T(\mathbf{X})$ یک برآورد درستنمایی ماکزیمم برای θ نامیده می شود که آن را با $\widehat{\theta}$ نمایش می دهیم.

براساس تعریف بالا، لزومی ندارد برآورد $\widehat{\theta}$ یکتا باشد و داریم

$$L(\widehat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

یا بطور معادل

$$l(\widehat{\theta}) = \sup_{\theta \in \Theta} l(\theta) \qquad \text{where} \quad l(\theta) = \ln L(\theta).$$

مثال ۲-۱۰. فرض کنید طول عمر یک نوع لامپ متغیر تصادفی با چگالی زیر باشد.

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

برآوردگری برای θ بدست آورید.

حل: فرض کنیم $x_۱, x_۲, \ldots, x_n$ یافته های یک نمونه تصادفی از X باشد. طبق روش گشتاوری داریم

$$\mu_۱ = \mathrm{E}(X) = \int_۰^\infty \theta x e^{-\theta x} \mathrm{dx} = -xe^{-\theta x}|_۰^\infty + \int_۰^\infty e^{-\theta x} \mathrm{dx} = \frac{-۱}{\theta} e^{-\theta x}|_۰^\infty = \frac{۱}{\theta}$$

$$M_۱ = \overline{X} \Longrightarrow \widetilde{\theta} = \frac{۱}{\overline{X}}$$

و براساس روش درستنمایی ماکزیمم بدست می آوریم

$$\begin{aligned}
&L(\theta)=\prod_{i=۱}^{n} f_\theta(x_i)=\prod_{i=۱}^{n}\theta e^{-\theta x_i}=\theta^n e^{-\theta\sum_{i=۱}^{n}x_i}\quad \textit{if}\quad \forall\, i\quad x_i>۰\\
&\Longrightarrow\ l(\theta)=\ln L(\theta)=n\ln\theta-\theta\sum_{i=۱}^{n}x_i\\
&\Longrightarrow\ l'(\theta)=\frac{n}{\theta}-\sum_{i=۱}^{n}x_i=۰\ \Longrightarrow \theta=\frac{n}{\sum_{i=۱}^{n}x_i}=\frac{۱}{\overline{X}}\\
&\Longrightarrow\ l''(\theta)=\frac{-n}{\theta^{۲}}<۰\ \Longrightarrow\ \widehat{\theta}=\frac{۱}{\overline{X}}.
\end{aligned}$$

در مثال بالا دیدیم که برآورد گشتاوری و درستنمایی ماکزیمم برای بعضی از پارامترهای توزیع می توانند برابر باشند.

مثال ۲-۱۱. فرض کنید X یک متغیرتصادفی با تابع چگالی زیر باشد

$$f_p(x)=p^x(۱-p)^{۱-x}\qquad x=۰,۱.$$

با استفاده از یک نمونه n تایی برآورد درستنمایی ماکزیمم پارامتر $\theta=p$ را برای $\Theta_۱=(۰,۱)$ و $\Theta_۲=\{\frac{۱}{۲},\frac{۱}{۳},\frac{۱}{۴}\}$ بدست آورید.

حل:

$$L(p)=\prod_{i=۱}^{n}p^{x_i}(۱-p)^{۱-x_i}=p^{\sum_{i=۱}^{n}x_i}(۱-p)^{n-\sum_{i=۱}^{n}x_i}$$

در حالت $\Theta_۱=(۰,۱)$ داریم

$$\begin{aligned}
&l(p)=\sum_{i=۱}^{n}x_i\ln p+(n-\sum_{i=۱}^{n}x_i)\ln(۱-p)\\
&\Longrightarrow\ l'(p)=\frac{\sum_{i=۱}^{n}x_i}{p}-\frac{n-\sum_{i=۱}^{n}x_i}{۱-p}=۰\\
&\Longrightarrow\ \frac{\sum_{i=۱}^{n}x_i(۱-p)-p(n-\sum_{i=۱}^{n}x_i)}{p(۱-p)}=۰\Longrightarrow\ p=\frac{\sum_{i=۱}^{n}x_i}{n}=\overline{x}\\
&l''(p)=\frac{-\sum_{i=۱}^{n}x_i}{p^{۲}}-\frac{n-\sum_{i=۱}^{n}x_i}{(۱-p)^{۲}}\ \Longrightarrow\ l''(\overline{x})<۰\ \Longrightarrow \widehat{p}=\overline{x}.
\end{aligned}$$

در حالت $\Theta_۲=\{\frac{۱}{۲},\frac{۱}{۳},\frac{۱}{۴}\}$ نمی توان از مشتق گیری استفاده کرد و با فرض اینکه $n=۵$ و یافته نمونه

به صورت $۱,۱,۰,۱,۰$ باشد داریم

$$L(\frac{۱}{۲}) = (\frac{۱}{۲})^{۳}(\frac{۱}{۲})^{۲} = (\frac{۱}{۲})^{۵} = ۰/۰۳۱۲۵$$
$$L(\frac{۱}{۳}) = (\frac{۱}{۳})^{۳}(\frac{۲}{۳})^{۲} = ۰/۰۱۶۴۶$$
$$L(\frac{۱}{۴}) = (\frac{۱}{۴})^{۳}(\frac{۳}{۴})^{۲} = ۰/۰۰۸۷۸۹$$
$$\Longrightarrow \max\{L(\frac{۱}{۲}), L(\frac{۱}{۳}), L(\frac{۱}{۴})\} = L(\frac{۱}{۲}) \Longrightarrow \widehat{p} = \frac{۱}{۲}$$

توجه کنید که برآورد درستنمایی ماکزیمم دارای خاصیت پایایی است بطوری که برآورد تابعی از پارامترهای مجهول، $\gamma(\theta_۱, \theta_۲, \dots, \theta_k)$، به صورت $\gamma(\widehat{\theta}_۱, \widehat{\theta}_۲, \dots, \widehat{\theta}_k)$ است.

تمرین ۲-۱۲. در مثال ۲-۱۱ اگر $\Theta = (\frac{۱}{۴}, \frac{۱}{۲})$، برآورد درستنمایی ماکزیمم پارامتر p را بدست آورید.

تمرین ۲-۱۳. اگر $X \sim Bin(m, P_N)$ براساس یک نمونه n تایی برآورد $\boldsymbol{ML}$ و $\boldsymbol{MM}$ پارامتر P_N را با فرض معلوم بودن مقدار m بدست آورید.

تمرین ۲-۱۴. اگر $X \sim Be(P_N)$ براساس یک نمونه n تایی برآورد $\boldsymbol{ML}$ و $\boldsymbol{MM}$ پارامتر P_N را بدست آورید.

تمرین ۲-۱۵. فرض کنید $X \sim U(۰, \theta)$. براساس یک نمونه n تایی برآورد $\boldsymbol{ML}$ و $\boldsymbol{MM}$ پارامتر θ را بدست آورید. برآورد $\boldsymbol{ML}$ $\frac{\theta}{۱+\theta}$ را نیز محاسبه کنید.

تمرین ۲-۱۶. فرض کنید $X \sim U(a, b)$. براساس یک نمونه n تایی برآورد $\boldsymbol{ML}$ و $\boldsymbol{MM}$ پارامترهای a و b را بدست آورید. اگر نمونه ای به صورت ۱/۱۵ ,۱/۲۳ ,۱/۰ ,۱/۱ ,۱/۱۲ ثبت شده باشد، مقدار پارامترها را برآورد کنید.

تمرین ۲-۱۷. فرض کنید $X \sim N(۰, \theta)$. براساس یک نمونه n تایی برآورد $\boldsymbol{ML}$ و $\boldsymbol{MM}$ پارامتر θ را بدست آورید.

۲.۲ برخی ویژگی برآوردهای نقطه ای

با توجه به اینکه برآوردگر تابعی از نمونه تصادفی است، پس برای یک پارامتر مجهول می توان برآوردگرهای زیادی را معرفی نمود. حال این سوال مطرح می شود که در بین برآوردگرهای یک پارامتر، کدام یک بهترین است.

معیارهای مختلفی برای سنجش بهترین برآوردگر وجود دارد که در اینجا به ذکر دو معیار می پردازیم.

۱.۲.۲ نااریبی

برآوردگر $T = g(X_۱, X_۲, \dots, X_n)$ را یک برآوردگر نااریب برای پارامتر θ گویند هرگاه

$$\mathrm{E}(T) = \theta$$

به عبارت دیگر، اگر متوسط مقدار برآوردگر T به ازای نمونه های مختلف، برابر با پارامتر θ باشد آن گاه T را برای θ نااریب گویند. برای روشن شدن این مفهوم به مثال زیر توجه کنید.

مثال ۲-۱۸. فرض کنید که جامعه آماری با چهار عضو و مقادیر زیر داشته باشیم:

$$X_۱, X_۲, X_۳, X_۴ \qquad N = ۴$$
$$x_۱ = ۱\ ,\ x_۲ = ۲\ ,\ x_۳ = ۳\ ,\ x_۴ = ۴ \qquad \Longrightarrow\ \mu = ۲/۵$$

اگر بخواهیم نمونه ای با سه عضو از این جامعه انتخاب کنیم، تعداد نمونه های سه عضوی ممکن برابر با ۴ نمونه با برآوردهای زیر است:

$$X_۱, X_۲, X_۳ \quad \Longrightarrow \bar{x}_۱ = ۲$$
$$X_۱, X_۲, X_۴ \quad \Longrightarrow \bar{x}_۲ = ۲/۳۳$$
$$X_۱, X_۳, X_۴ \quad \Longrightarrow \bar{x}_۳ = ۲/۶۶$$
$$X_۲, X_۳, X_۴ \quad \Longrightarrow \bar{x}_۴ = ۳$$

واضح است که مقادیر $\overline{x}$ بدست آمده کوچکتر یا بزرگتر از مقدار μ است. بدین معنا که برآورد بدست آمده توسط هر نمونه تصادفی می تواند بزرگتر، کوچکتر و یا مساوی با پارامتر جامعه باشد و لزومی ندارد که همیشه برابر با μ باشد زیرا برآورد بدست آمده تقریب، تخمین و یا برآوردی از پارامتر جامعه است. این در حالی است که داریم:

$$\frac{\overline{x}_۱ + \overline{x}_۲ + \overline{x}_۳ + \overline{x}_۴}{۴} = \frac{۱۰}{۴} = ۲/۵ = \mu$$

یعنی به طور متوسط میانگین نمونه ای با میانگین جامعه برابر است.

- **توزیع دوجمله ای:** فرض کنید

$$X_۱, X_۲, \dots, X_N \sim Bin(m, P_N)$$
$$X_۱, X_۲, \dots, X_n \sim Bin(m, P_N)$$

به طوری که m معلوم و $\theta = P_N$ نامعلوم باشد. می دانیم که P_N، نسبت موفقیت و یا احتمال موفقیت است پس می توان آن را به صورت زیر نمایش داد:

$$P_N = \frac{T}{N}$$

به طوری که T، تعداد موفقیت ها در جامعه بوده و N حجم جامعه است. برآوردگر مطلوب برای پارامتر P_N (نسبت موفقیت ها در جامعه)، نسبت موفقیتها در نمونه است که آن را با P_n نمایش می دهیم و داریم $P_n = \frac{t}{n}$ که در آن t تعداد موفقیت ها در نمونه و n حجم نمونه است.

توجه کنید که نسبت موفقیت ها در نمونه، برای پارامتر نسبت موفقیت ها در جامعه ناااریب است. به عبارت دیگر

$$\mathrm{E}(P_n) = P_N$$

• **توزیع نرمال:** فرض کنید $X_۱, X_۲, \ldots, X_N \sim N(\mu, \sigma^۲)$. پس این جامعه دارای دو پارامتر است که آن ها را با $\theta_۱ = \mu$ و $\theta_۲ = \sigma^۲$ نمایش می دهیم. درواقع داریم:

$$\theta_۱ = \mu = \frac{۱}{N}\sum_{i=۱}^{N} X_i \quad \text{or} \quad \mu = \mathrm{E}(X)$$

$$\theta_۲ = \sigma^۲ = \frac{۱}{N}\sum_{i=۱}^{N} (X_i - \mu)^۲ \quad \text{or} \quad \sigma^۲ = \mathrm{Var}(X)$$

⋆ میانگین نمونه یعنی $\overline{X}$، برآوردگر مطلوب برای پارامتر $\theta_۱ = \mu$ است به طوری که

$$\overline{X} = \frac{۱}{n}\sum_{i=۱}^{n} X_i$$

داریم

$$\mathrm{E}(\overline{X}) = \mathrm{E}(\frac{۱}{n}\sum_{i=۱}^{n} X_i) = \frac{۱}{n}\sum_{i=۱}^{n} \mathrm{E}(X_i) = \frac{۱}{n}\sum_{i=۱}^{n} \mu = \mu.$$

پس برآوردگر $\overline{X}$ یک برآوردگر نااریب برای μ است.

⋆ وقتی μ نامعلوم باشد واریانس نمونه $S^۲_{n-۱}$، برآوردگر مطلوب برای پارامتر $\theta_۲ = \sigma^۲$ است به طوری که

$$S^۲_{n-۱} = \frac{۱}{n-۱}\sum_{i=۱}^{n} (X_i - \overline{X})^۲$$

برآوردگر $S^۲_{n-۱}$ یک برآوردگر نااریب برای $\sigma^۲$ است و داریم

$$\frac{(n-۱)S^۲_{n-۱}}{\sigma^۲} \sim \chi^۲(n-۱) \quad \Longrightarrow \mathrm{E}(\frac{(n-۱)S^۲_{n-۱}}{\sigma^۲}) = n-۱ \quad \Longrightarrow \mathrm{E}(S^۲_{n-۱}) = \sigma^۲$$

شایان ذکر است که واریانس نمونه ای می تواند به صورت $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ نیز تعریف شود اما

$$\mathrm{E}(S_n^2) = \frac{n-1}{n}\mathrm{E}(S_{n-1}^2) = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

و S_n^2 برای پارامتر واریانس جامعه نااریب نیست. از اینرو S_n^2 برآوردگر مطلوبی برای پارامتر واریانس جامعه نرمال نخواهد بود.

تمرین ۲-۱۹. در توزیع نرمال اگر μ معلوم باشد، برآورد نقطه ای برای σ^2 بدست آورید.

تمرین ۲-۲۰. در توزیع نرمال آیا S_{n-1} یک برآورد نااریب برای σ است؟ پاسخ خود را ثابت کنید.

تمرین ۲-۲۱. در توزیع نرمال آیا $\frac{1}{S_{n-1}^2}$ یک برآورد نااریب برای $\frac{1}{\sigma^2}$ است؟ پاسخ خود را ثابت کنید.

۲.۲.۲ کمترین واریانس

اگر T_1 و T_2 دو برآوردگر نااریب برای پارامتر θ باشند به طوری که

$$\mathrm{Var}(T_1) < \mathrm{Var}(T_2),$$

آن گاه برآوردگر T_1 از برآوردگر T_2 کاراتر است. بنابراین در بین برآوردگرهای نااریب برای θ، برآوردگری که دارای کمترین واریانس است را کاراترین برآوردگر می نامند.

مثال ۲-۲۲. فرض کنید

$$X_1, X_2, \ldots, X_N \sim N(\mu, \sigma^2)$$
$$X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$$

به طوری که σ^2 معلوم و $\theta = \mu$ نامعلوم باشد. اگر $T_1 = \frac{X_1 + X_n}{2}$ و $T_2 = \overline{X}$، کدام یک برآوردگر کاراتری برای μ است؟

حل: طبق خواص امید ریاضی و واریانس هر تابع از متغیرهای تصادفی داریم:

$$\mathrm{E}(T_1) = \mathrm{E}(\frac{X_1 + X_n}{2}) = \frac{\mathrm{E}(X_1) + \mathrm{E}(X_n)}{2} = \frac{2\mu}{2} = \mu$$
$$\mathrm{E}(T_2) = \mathrm{E}(\overline{X}) = \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{n\mu}{n} = \mu$$

پس هر دوی این برآوردگرها، برای پارامتر $\theta = \mu$ نااریب هستند.

$$\mathrm{Var}(T_۱) = \mathrm{Var}(\frac{X_۱ + X_n}{۲}) = \frac{\mathrm{Var}(X_۱) + \mathrm{Var}(X_n)}{۴} = \frac{۲\sigma^۲}{۴} = \frac{\sigma^۲}{۲}$$
$$\mathrm{Var}(T_۲) = \mathrm{Var}(\overline{X}) = \mathrm{Var}(\frac{۱}{n}\sum_{i=۱}^{n} X_i) = \frac{۱}{n^۲}\sum_{i=۱}^{n}\mathrm{Var}(X_i) = \frac{n\sigma^۲}{n^۲} = \frac{\sigma^۲}{n}$$

واضح است که

$$\mathrm{Var}(T_۱) = \frac{\sigma^۲}{۲} > \mathrm{Var}(T_۲) = \frac{\sigma^۲}{n}$$

و بنابراین برآوردگر $T_۲ = \overline{X}$ کاراتر از $T_۱ = \dfrac{X_۱ + X_n}{۲}$ است.

مثال ۲-۲۳. نتایج مربوط به مقاومت نمونه های بتنی تهیه شده در یک کارگاه ساختمانی بر حسب کیلوگرم بر سانتیمتر مربع به صورت زیر ارائه شده است،

۲۰۱, ۲۰۵, ۱۹۹, ۱۹۰, ۲۱۸, ۲۴۲, ۱۸۳, ۲۲۱, ۲۳۵, ۲۱۴, ۲۲۲,
۲۱۱, ۲۴۱, ۲۰۶, ۲۰۹, ۱۹۷, ۲۱۳, ۲۲۴, ۲۳۱, ۲۱۷, ۲۰۲, ۱۹۵.

اگر بدانیم که مقاومت بتن از توزیع نرمال پیروی می کند، برآورد نقطه ای امیدریاضی و واریانس آن ها را بدست آورید.

حل.

$$\widehat{\mu} = \overline{x} = \frac{۱}{۲۲}\sum_{i=۱}^{۲} ۲x_i = ۲۱۲/۵$$
$$\widehat{\sigma^۲} = S^۲_{n-۱} = \frac{۱}{۲۱}\sum_{i=۱}^{۲۲}(x_i - \overline{x})^۲ = \frac{۱}{۲۱}(\sum_{i=۱}^{۲۲} x_i^۲ - ۲۲\overline{x}^۲) = ۲۷۴/۵$$

تمرین ۲-۲۴. فرض کنید $U_۱$ و $U_۲$ دو برآوردگر نااریب و مستقل برای θ باشند که واریانس آن ها به ترتیب برابر با ۲ و ۳ است. اگر $T = aU_۱ + bU_۲$، ضرایب a و b را طوری تعیین کنید که T برآورد نااریب با کمترین واریانس در کلاس برآوردگر های $aU_۱ + bU_۲$ برای θ باشد.

تمرین ۲-۲۵. فرض کنید $U_۱$ و $U_۲$ دو برآوردگر نااریب و مستقل برای θ باشند. بر حسب $U_۱$ و $U_۲$ یک برآوردگر نااریب برای $\theta^۲$ و $\theta(۱ - \theta)$ معرفی کنید.

۳.۲ برآورد نقطه ای پارامتر دو جامعه

همانطور که قبلاً هم بیان شد، در استنباط آماری هدف شناخت جامعه آماری براساس یک یا چند مشخصه (متغیر) است. یکی از راه های شناخت یک جامعه آماری، مقایسه آن با یک یا چند جامعه آماری دیگر است. از اینرو، در این قسمت به مقایسه دو جامعه آماری می پردازیم.

۱.۳.۲ مقایسه دو جامعه آماری مستقل نرمال

دو جامعه مستقل را در نظر بگیرید که براساس یک مشخصه بررسی و باهم مقایسه می گردند. فرض کنید توزیع آماری مشخصه مورد نظر در دو جامعه نرمال باشد.

- جامعه اول: $X_1, X_2, \dots, X_N \sim N(\mu_1, \sigma_1^2)$
  حجم جامعه: N
  مشخصه (متغیر) مورد بررسی: X
  میانگین جامعه: μ_1
  واریانس جامعه: σ_1^2
- جامعه دوم: $Y_1, Y_2, \dots, Y_M \sim N(\mu_2, \sigma_2^2)$
  حجم جامعه: M
  مشخصه (متغیر) مورد بررسی: Y
  میانگین جامعه: μ_2
  واریانس جامعه: σ_2^2

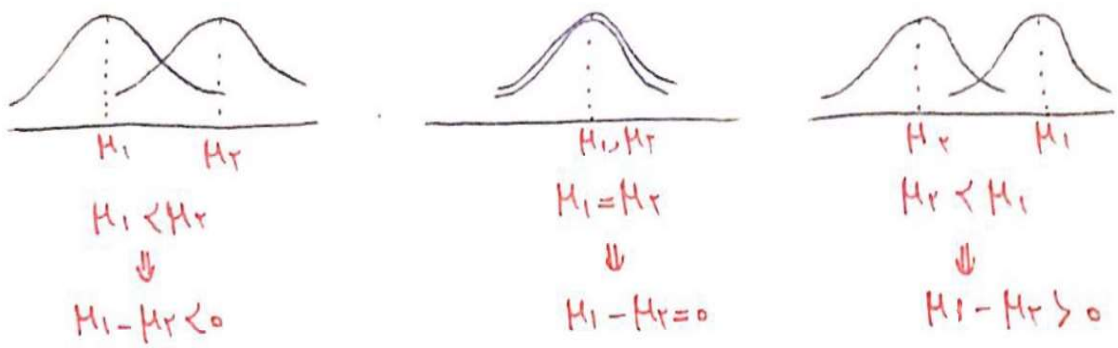
برای مثال، جامعه اول را دانشجویان دختر و جامعه دوم را دانشجویان پسر در نظر گرفته و فرض کنید مشخصه مورد بررسی، نمرات پایانترم آن ها باشد. هدف از مقایسه این دو جامعه، شناخت بهتر آن ها است.

برای مقایسه دو جامعه نرمال باید به بررسی پارامترهای دو جامعه پرداخته شود:

$$X_1, X_2, \dots, X_N \sim N(\mu_1, \sigma_1^2)$$
$$Y_1, Y_2, \dots, Y_M \sim N(\mu_2, \sigma_2^2)$$

به عبارت دیگر باید به مقایسه میانگین های دو جامعه (یعنی μ_1 و μ_2) و یا واریانس های دو جامعه (یعنی σ_1^2 و σ_2^2) پرداخت.

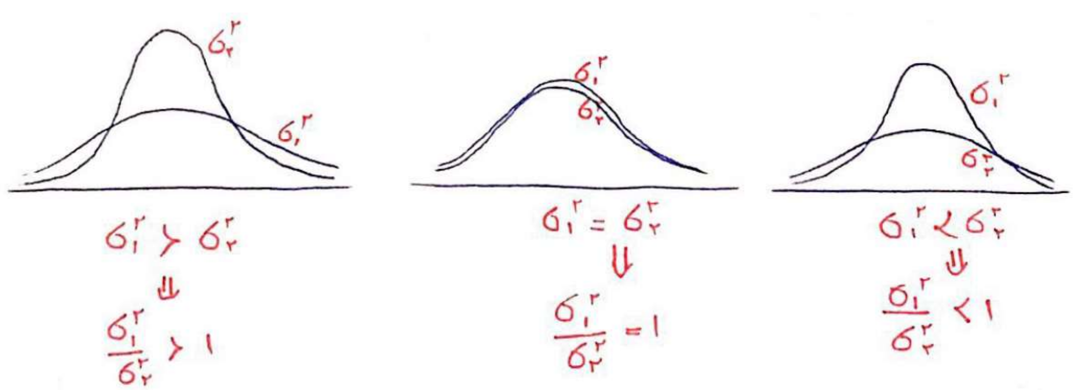
براساس روابط بین میانگین های دو جامعه، ممکن است یکی از سه حالت زیر را داشته باشیم:



شکل ۲-۱: مقایسه دو جامعه نرمال براساس میانگین آن ها

پس می توان پارامتری را بر روی دو جامعه به صورت $\theta = \mu_1 - \mu_2$ تعریف کرد و با برآورد آن به صورت نقطه ای و یا فاصله ای، به مقایسه دو جامعه نرمال مستقل پرداخت.

برای مقایسه دو جامعه نرمال مستقل براساس واریانس های دو جامعه یعنی σ_1^2 و σ_2^2، یکی از سه حالت زیر ممکن است رخ دهد:



شکل ۲-۲: مقایسه دو جامعه نرمال براساس واریانس آن ها

پس جهت مقایسه دو جامعه نرمال مستقل براساس واریانس آن ها می توان پارامتر $\theta = \frac{\sigma_1^2}{\sigma_2^2}$ را تعریف کرد و آن را به صورت نقطه ای و فاصله ای برآورد کرد.

برآورد نقطه ای پارامتر $\theta = \mu_1 - \mu_2$

دو جامعه نرمال مستقل به صورت زیر داریم:

$$X_1, X_2, \ldots, X_N \sim N(\mu_1, \sigma_1^2)$$
$$Y_1, Y_2, \ldots, Y_M \sim N(\mu_2, \sigma_2^2)$$

برای برآوردیابی ابتدا از هر جامعه یک نمونه تصادفی به صورت زیر می گیریم

$$X_1, X_2, \ldots, X_n \sim N(\mu_1, \sigma_1^2)$$
$$Y_1, Y_2, \ldots, Y_m \sim N(\mu_2, \sigma_2^2)$$

می دانیم برآورد نقطه ای مطلوب برای μ_1، $\overline{X}$ و برآورد نقطه ای مطلوب برای μ_2، $\overline{Y}$ است. با استفاده از این واقعیت می توان برآورد نقطه ای برای پارامتر $\theta = \mu_1 - \mu_2$ را $\overline{X} - \overline{Y}$ درنظر گرفت. جهت بررسی مطلوبیت این برآوردگر داریم

$$\mathrm{E}(\overline{X} - \overline{Y}) = \mathrm{E}(\overline{X}) - \mathrm{E}(\overline{Y}) = \mu_1 - \mu_2$$

پس برآوردگر $\overline{X} - \overline{Y}$ برای پارامتر $\theta = \mu_1 - \mu_2$ ناااریب است. همچنین داریم

$$\mathrm{Var}(\overline{X} - \overline{Y}) = \mathrm{Var}(\overline{X}) + \mathrm{Var}(\overline{Y}) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}$$

و $\overline{X} - \overline{Y}$ یک برآوردگر مطلوب برای $\theta = \mu_1 - \mu_2$ است.

تمرین ۲-۲۶. با توجه به مطالبی که تا کنون گفته شد، ثابت کنید که چرا $\overline{X} - \overline{Y}$ برآورد نقطه ای برای پارامتر $\theta = \mu_1 - \mu_2$ است.

برآورد نقطه ای $\theta = \frac{\sigma_1^2}{\sigma_2^2}$

دو جامعه نرمال مستقل به صورت زیر داریم:

$$X_1, X_2, \ldots, X_N \sim N(\mu_1, \sigma_1^2)$$
$$Y_1, Y_2, \ldots, Y_M \sim N(\mu_2, \sigma_2^2)$$

برای برآوردیابی ابتدا از هر یک از جوامع بالا، یک نمونه تصادفی به صورت زیر می گیریم

$$X_۱, X_۲, \ldots, X_n \sim N(\mu_۱, \sigma_۱^۲)$$
$$Y_۱, Y_۲, \ldots, Y_m \sim N(\mu_۲, \sigma_۲^۲)$$

می دانیم برآوردگر نقطه ای مطلوب برای $\sigma_۱^۲$، $S_۱^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i - \overline{X})^۲$ و برآوردگر نقطه ای مطلوب برای $\sigma_۲^۲$، $S_۲^۲ = \frac{۱}{m-۱}\sum_{i=۱}^{m}(Y_i - \overline{Y})^۲$ است. از اینرو می توان برآوردگر $\frac{S_۱^۲}{S_۲^۲}$ را برای $\frac{\sigma_۱^۲}{\sigma_۲^۲}$ درنظر گرفت.

با محاسبات پیچیده ای، می توان نشان داد که $\frac{S_۱^۲}{S_۲^۲}$ یک برآوردگر نقطه ای مطلوب برای پارامتر $\frac{\sigma_۱^۲}{\sigma_۲^۲}$ است.

مثال ۲–۲۷. برای یک درس، یک گروه صبح و یک گروه بعداز ظهر تشکیل می شود. در آخر ترم از کلاس صبح ۸ نفر و از کلاس بعداز ظهر ۹ نفر را به تصادف انتخاب کرده و از آن ها امتحان گرفته شده است. نتایج به صورت زیر بدست آمده است:

$$Class\ ۱: \quad ۱۲,\ ۷,\ ۱۵,\ ۱۲,\ ۱۰,\ ۸,\ ۷,\ ۹$$
$$Class\ ۲: \quad ۱۰,\ ۱۱,\ ۵,\ ۱۶,\ ۱۸,\ ۲,\ ۹,\ ۷,\ ۳$$

با فرض اینکه نمرات دو کلاس مستقل از یکدیگر بوده و هر دو کلاس از توزیع نرمال تبعیت کنند، یک براورد نقطه ای برای تفاضل میانگین ها و نسبت واریانس های آن ها بدست آورید.

حل: فرض می کنیم X نمرات کلاس صبح و Y نمرات کلاس بعدازظهر باشد. داریم $X \sim N(\mu_۱, \sigma_۱^۲)$ و $Y \sim N(\mu_۲, \sigma_۲^۲)$. طبق اطلاعات مساله داریم

$$n = ۸, \quad \overline{x} = \frac{۱}{۸}\sum_{i=۱}^{۸} x_i = \frac{۸۰}{۸} = ۱۰, \quad S_۱^۲ = \frac{۱}{۷}\sum_{i=۱}^{۸}(x_i - \overline{x})^۲ = ۸$$
$$m = ۹, \quad \overline{y} = \frac{۱}{۹}\sum_{i=۱}^{۹} y_i = \frac{۸۱}{۹} = ۹, \quad S_۲^۲ = \frac{۱}{۸}\sum_{i=۱}^{۹}(y_i - \overline{y})^۲ = ۳۰$$

بنابراین برای $\theta_۱ = \mu_۱ - \mu_۲$ و $\theta_۲ = \frac{\sigma_۱^۲}{\sigma_۲^۲}$ بدست می آوریم

$$\widehat{\theta}_۱ = \overline{x} - \overline{y} = ۱۰ - ۹ = ۱ \qquad \widehat{\theta}_۲ = \frac{S_۱^۲}{S_۲^۲} = \frac{۸}{۳۰}.$$

دقت کنید که مقادیر بدست آمده فقط تقریبی برای $\theta_۱$ و $\theta_۲$ هستند که با استفاده از نمونه مشاهده شده بدست آمده اند.

تمرین ۲-۲۸. با توجه به مطالبی که تا کنون گفته شد، ثابت کنید که چرا $\frac{S_1^2}{S_2^2}$ برآورد نقطه ای برای پارامتر $\theta_2 = \frac{\sigma_1^2}{\sigma_2^2}$ است.

۲.۳.۲ مقایسه دو جامعه آماری مستقل برنولی

دو جامعه مستقل درنظر بگیرید که براساس یک مشخصه بررسی و با هم مقایسه می گردند. فرض کنید توزیع آماری مشخصه موردنظر در دو جامعه برنولی است. یعنی

$$X_1, X_2, \ldots, X_N \sim Be(P_1)$$
$$Y_1, Y_2, \ldots, Y_M \sim Be(P_2)$$

برای مثال، جامعه اول را دانشجویان دختر و جامعه دوم را دانشجویان پسر در نظر گرفته و فرض کنید مشخصه مورد بررسی، رضایت یا عدم رضایت ایشان از تدریس مجازی باشد.

برای مقایسه دو جامعه برنولی باید به بررسی پارامترهای دو جامعه (یعنی مقایسه نسبت موفقیت در دو جامعه، P_1 و P_2) پرداخت. پارامتر $\theta = P_1 - P_2$ را بدین منظور تعریف می کنیم.

برآورد نقطه ای $\theta = P_1 - P_2$

دو جامعه برنولی مستقل به صورت زیر داریم:

$$X_1, X_2, \ldots, X_N \sim Be(P_1)$$
$$Y_1, Y_2, \ldots, Y_M \sim Be(P_2)$$

برای برآوردیابی ابتدا از هریک از جوامع بالا یک نمونه تصادفی به صورت زیر می گیریم:

$$X_1, X_2, \ldots, X_n \sim Be(P_1)$$
$$Y_1, Y_2, \ldots, Y_m \sim Be(P_2)$$

می دانیم که برآوردگر مطلوب نقطه ای برای P_1، $p_1 = \frac{t_1}{n}$ و برآوردگر نقطه ای مطلوب برای P_2، $p_2 = \frac{t_2}{m}$ بوده به طوری که t_1 تعداد موفقیت در نمونه از جامعه اول و t_2 تعداد موفقیت در نمونه جامعه دوم است. از اینرو می توان برآوردگر $p_1 - p_2$ را برای پارامتر $P_1 - P_2$ در نظر گرفت. داریم:

$$\mathrm{E}(p_1 - p_2) = \mathrm{E}(p_1) - \mathrm{E}(p_2) = P_1 - P_2$$

پس برآوردگر $p_۱ - p_۲$ یک برآوردگر نااریب برای $P_۱ - P_۲$ است.

$$\mathrm{Var}(p_۱ - p_۲) = \mathrm{Var}(p_۱) + \mathrm{Var}(p_۲) = \frac{P_۱(۱ - P_۱)}{n} + \frac{P_۲(۱ - P_۲)}{m}$$

فصل ۳

برآورد فاصله ای

همانطورکه در فصل های قبل نیز گفته شد، در مساله برآوردیابی به دنبال برآورد پارامتر جامعه با استفاده از یک نمونه تصادفی هستیم.

در روش برآورد نقطه ای، پارامتر جامعه توسط یک برآوردگر، برآورد شده و مقدار آن با استفاده از مقدار برآوردگر به ازای نمونه مشاهده شده، تخمین زده می شود. اما در برآورد فاصله ای به جای تقریب پارامتر با یک عدد، از یک محدوده یا فاصله استفاده می کنیم که با احتمال بسیار بالا پارامتر جامعه را دربرمی گیرد.

تعریف ۳-۱. فاصله (L, U) به طوری که $L < U$ و L یا U متغیرهای تصادفی هستند را یک فاصله تصادفی می نامند.

برای مثال، فرض کنید $X \sim N(۰, ۴)$. در این صورت $(۲X - ۱, ۲X)$ یک فاصله تصادفی است. همچنین اگر یک نمونه n تایی از متغیر تصافی X را به صورت $X_۱, X_۲, \ldots, X_n$ درنظر بگیریم، $(\overline{X}, \infty)$ نیز یک فاصله تصادفی است.

تعریف ۳-۲. فرض کنید X یک متغیر تصادفی از توزیعی با پارامتر θ بوده و (L, U) یک فاصله تصادفی باشد. در این صورت اگر

$$\mathrm{P}(L < \theta < U) = ۱ - \alpha$$