نکات کلیدی فصل هشتم

● به طور کلی آماره‌ای که برای برآورد کردن پارامتر جامعه به کار می‌بریم، «برآوردکننده» یا «برآوردگر نقطه‌ای آن پارامتر» می‌نامیم. ولی از آنجا که برآورد نقطه‌ای دارای خطای زیاد می‌باشد و تعداد آن با تغییر حجم نمونه تغییر خواهد کرد و به ندرت می‌توان انتظار داشت که برآوردهای نقطه‌ای برابر با پارامترهایی باشند که هدف برآورد آنهاست؛ سعی می‌شود تا به جای برآورد نقطه‌ای، یک برآورد فاصله‌ای تعیین شود.

● به دست آوردن فاصله اطمینان در جوامع نرمال، نمونه‌های بزرگ و نمونه‌های کوچک با واریانس‌های معلوم و نامعلوم متفاوت است.

۱– فاصله اطمینان μ برای نمونه‌های بزرگ

حد پایین و حد بالای فاصله‌های اطمینان با ضریب اطمینان

$$L = \bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad , \quad U = \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

ب) اگر انحراف معیار جامعه نامعلوم باشد:

$$L = \bar{X} - t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n}} \quad , \quad U = \bar{X} + t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n}}$$

● لازم به ذکر است که فاصله اطمینان فوق را حتی وقتی جامعه دقیقاً نرمال نباشد می‌توان به کار برد؛ ولی در این حالت ضریب اطمینان فاصله دقیقاً $(1-\alpha)$ نیست بلکه تقریباً $(1-\alpha)$ است و این ویژگی مهم توزیع t است.

۲– برآورد واریانس جامعه

فرض کنید نمونه‌ای تصادفی به اندازه n و واریانس S^2، از جامعه‌ای نرمال با واریانس σ^2 استخراج کرده‌ایم در این صورت حدود فاصله اطمینانی با ضریب $(1-\alpha)$، برای σ^2 از رابطه‌های زیر به دست می‌آیند:

$$(L, U) = \left(\frac{(n-1)S^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}}, \frac{(n-1)S^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}} \right)$$

که $\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}$ و $\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}$ به ترتیب نقطه $\frac{\alpha}{2}$ درصد و نقطه $1-\frac{\alpha}{2}$ درصد توزیع χ^2 با $(n-1)$ درجه آزادی هستند که در جدول ضمیمه موجود است.

H_0	آماره آزمون	H_1	ناحیه بحرانی	ناحیه قبول
$\mu \geq \mu_0$		$\mu < \mu_0$	$Z < -z_{\alpha}$	$Z \geq -z_{\alpha}$
$\mu \leq \mu_0$	$Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$\mu > \mu_0$	$Z > z_{\alpha}$	$Z \leq z_{\alpha}$
$\mu = \mu_0$	σ معلوم	$\mu \neq \mu_0$	$\lvert Z \rvert > z_{\frac{\alpha}{2}}$	$\lvert Z \rvert \leq z_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$		$\mu < \mu_0$	$T < -t_{(\alpha, n-1)}$	$T \geq t_{(\alpha, n-1)}$
$\mu \leq \mu_0$	$T = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$\mu > \mu_0$	$T > t_{(\alpha, n-1)}$	$T \leq t_{(\alpha, n-1)}$
$\mu = \mu_0$	σ نامعلوم	$\mu \neq \mu_0$	$\lvert T \rvert > t_{\left(\frac{\alpha}{2}, n-1\right)}$	$\lvert T \rvert \leq t_{\left(\frac{\alpha}{2}, n-1\right)}$

$(۱-\alpha)$ برای μ به شرح زیر است:

الف) اگر انحراف معیار جامعه معلوم باشد:

$$L=\overline{X}-z_{\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}\quad,\quad U=\overline{X}+z_{\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}$$

ب) اگر انحراف معیار جامعه معلوم نباشد:

$$L=\overline{X}-z_{\frac{\alpha}{۲}}\frac{S}{\sqrt{n}}\quad,\quad U=\overline{X}+z_{\frac{\alpha}{۲}}\frac{S}{\sqrt{n}}$$

۲- فاصله اطمینان μ برای نمونه‌های کوچک

حدود اطمینان برای μ با ضریب اطمینان $(۱-\alpha)$ وقتی $n<30$ عبارت است از:

الف) اگر انحراف معیار جامعه معلوم باشد:

نکات کلیدی فصل نهم

۱- برآورد نسبت جامعه

الف) برآوردکننده نقطه‌ای نسبت جامعه

$\overline{P}=\frac{X}{n}$، بهترین برآوردکننده P نسبت به سایر برآوردکننده‌هاست و میانگین و واریانس $\overline{P}$ از رابطه‌های زیر به دست می‌آیند.

$$E(\overline{P})=p$$

$$\sigma^{۲}_{p}=\frac{P(۱-P)}{n}$$

ب) برآورد فاصله‌ای p برای نمونه‌های بزرگ:

اگر جامعه مورد نظر نامتناهی، یا متناهی بزرگ باشد، حد بالا و پایین فاصله اطمینان P عبارت‌اند از:

$$(L,U)=\overline{P}\pm z_{\frac{\alpha}{۲}}\sqrt{\frac{\overline{P}(۱-\overline{P})}{n}}$$

نکات کلیدی فصل دهم

در این فصل آزمون‌های آماری را مورد بررسی قرار دادیم:

الف) آزمون‌های دوطرفه
$$\begin{cases}H_{۰}:\mu=\mu_{۰}\\H_{۱}:\mu\neq\mu_{۰}\end{cases}$$

ب) آزمون‌های یک‌طرفه راست
$$\begin{cases}H_{۰}:\mu\leq\mu_{۰}\\H_{۱}:\mu>\mu_{۰}\end{cases}$$

ج) آزمون‌های یک‌طرفه چپ
$$\begin{cases}H_{۰}:\mu\geq\mu_{۰}\\H_{۱}:\mu<\mu_{۰}\end{cases}$$

جدول فوق، آماره‌های مناسب برای هر آزمون و نواحی بحرانی مربوطه برای پارامترهای توزیع نرمال را ارائه می‌دهد:

$H_{۰}$	آماره آزمون	$H_{۱}$	ناحیه بحرانی	ناحیه قبول
$\mu_{۱}\geq\mu_{۲}$	$Z=\frac{(\overline{X}_{۱}-\overline{X}_{۲})}{\sqrt{\frac{\sigma_{۱}^{۲}}{n_{۱}}+\frac{\sigma_{۲}^{۲}}{n_{۲}}}}$ $\sigma_{۱}$ و $\sigma_{۲}$ معلوم	$\mu_{۱}<\mu_{۲}$	$Z>z_{\alpha}$	$Z\leq z_{\alpha}$
$\mu_{۱}\leq\mu_{۲}$		$\mu_{۱}>\mu_{۲}$	$Z<Z_{\alpha}$	$Z\geq-z_{\alpha}$
$\mu_{۱}=\mu_{۲}$		$\mu_{۱}\neq\mu_{۲}$	$\lvert Z\rvert>Z_{\frac{\alpha}{۲}}$	$\lvert Z\rvert\leq z_{\frac{\alpha}{۲}}$
$\mu_{۱}\geq\mu_{۲}$	$T=\frac{(\overline{X}_{۱}-\overline{X}_{۲})}{S_{P}\left(\sqrt{\frac{۱}{n_{۱}}+\frac{۱}{n_{۲}}}\right)}$ $\sigma_{۱}=\sigma_{۲}$ اما نامعلوم $S_{P}^{۲}=\frac{(n_{۱}-۱)S_{۱}^{۲}+(n_{۲}-۱)S_{۲}^{۲}}{n_{۱}+n_{۲}-۲}$	$\mu_{۱}<\mu_{۲}$	$T>t_{(\alpha,n_{۱}+n_{۲}-۲)}$	$T\leq t_{(\alpha,n_{۱}+n_{۲}-۲)}$
$\mu_{۱}\leq\mu_{۲}$		$\mu_{۱}>\mu_{۲}$	$T<-t_{(\alpha,n_{۱}+n_{۲}-۲)}$	$T\geq-t_{(\alpha,n_{۱}+n_{۲}-۲)}$
$\mu_{۱}=\mu_{۲}$		$\mu_{۱}\neq\mu_{۲}$	$\lvert T\rvert>t_{(\frac{\alpha}{۲},n_{۱}+n_{۲}-۱)}$	$\lvert T\rvert\leq t_{(\frac{\alpha}{۲},n_{۱}+n_{۲}-۲)}$

نکات کلیدی فصل یازدهم

در این فصل آزمون‌های مربوط به نسبت جامعه p را وقتی اندازه نمونه و جامعه بزرگ باشد بررسی کردیم و سپس آزمون‌هایی را درباره واریانس جامعه شرح دادیم. به طور خلاصه خواهیم داشت:

$H_۰$	آماره آزمون	$H_۱$	ناحیه بحرانی	ناحیه قبول
$p \geq p_۰$	$Z=\dfrac{\bar{p}-p_۰}{\sigma_{\bar{p}}}$	$p < p_۰$	$Z < -z_\alpha$	$Z \geq -z_\alpha$
$p \leq p_۰$		$p > p_۰$	$Z > z_\alpha$	$Z \leq z_\alpha$
$p = p_۰$		$p \neq p_۰$	$\lvert Z\rvert > z_{\frac{\alpha}{۲}}$	$\lvert Z\rvert \leq z_{\frac{\alpha}{۲}}$
$\sigma^۲ \geq \sigma_۰^۲$	$\chi^۲=\dfrac{(n-۱)S^۲}{\sigma_۰^۲}$	$\sigma^۲ < \sigma_۰^۲$	$X^۲ < \chi^۲_{(۱-\alpha, n-۱)}$	
$\sigma^۲ \leq \sigma_۰^۲$		$\sigma^۲ > \sigma_۰^۲$	$X^۲ > \chi^۲_{(\alpha, n-۱)}$	
$\sigma^۲ = \sigma_۰^۲$		$\sigma^۲ \neq \sigma_۰^۲$	$X^۲ < \chi^۲_{(۱-\frac{\alpha}{۲}, n-۱)}$ $X^۲ > \chi^۲_{(\frac{\alpha}{۲}, n-۱)}$	

اگر $F=\dfrac{MSR}{MSE} > F_{(k-۱, N-k, \alpha)}$ آنگاه فرض برابری میانگین را در سطح معنی‌داری α رد می‌کنیم:

$$SSR=\sum_{i=۱}^{k} n(\bar{x}_i-\bar{x})^۲$$

$$SST=\sum_{I=۱}^{k}\sum_{j=۱}^{n}(x_{ij}-\bar{x})^۲$$

$$SSE=\sum_{i=۱}^{k}\sum_{j=۱}^{n}(x_{ij}-\bar{x}_i)^۲$$

که $\bar{X}$ میانگین کل داده‌ها، $\bar{X}_i$ میانگین نمونه i ام، k تعداد گروه‌ها، N تعداد کل داده‌ها و X_{ij} j امین مقدار شده در نمونه i ام است.

داشتیم:

$$\bar{x}=\frac{۱}{n}\sum_{i=۱}^{n}x_i \qquad \bar{y}=\frac{۱}{n}\sum_{i=۱}^{n}y_i$$

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

$$S_{yy}=\sum_{i=۱}^{n}(y_i-\bar{y})^۲=\sum_{i=۱}^{n}y_i{}^۲-n\bar{y}^۲$$

$$S_{xy}=\sum_{i=۱}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=۱}^{n}x_iy_i-n\bar{x}\,\bar{y}$$

$$\hat{\beta}=\frac{S_{xy}}{S_x}, \quad \hat{\alpha}=\bar{y}-\hat{\beta}\,\bar{x}$$

فرمول‌های دیگری نیز برای محاسبه مجموع توان‌های دوم وجود دارند که عبارت‌اند از:

$$SSR=\sum_{i=۱}^{k}\frac{T_i^۲}{n}-\frac{T^۲}{N}$$

$$SST=\sum_{i=۱}^{k}\sum_{j=۱}^{n}x^۲{}_{ij}-\frac{T^۲}{N}$$

$$SSE=SST-SSR$$

نکات کلیدی فصل دوازدهم

در این فصل میانگین‌های k جامعه نرمال را که با $\mu_1, \mu_2, \dots, \mu_k$ نشان می‌دهیم با هم مقایسه کردیم و بررسی کردیم که آیا اختلافی بین میانگین‌ها وجود دارد یا نه.

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

H_1: حداقل یکی از $\mu_1, \mu_2, \dots, \mu_k$ ها با بقیه اختلاف دارد.

برای این منظور جدول واریانس را تشکیل می‌دهیم:

منبع تغییر	درجه آزادی	مجموع مربعات	میانگین مربعات	F
بین گروه‌ها	$k-1$	SSR	$MSR = \frac{SSR}{k-1}$	MSR
درون گروه‌ها	$N-k$	SSE	$MSE = \frac{SSE}{N-k}$	MSE
جمع	$N-1$	SST		

نکات کلیدی فصل سیزدهم

- گاهی با مسأله پیش‌بینی مقدار یک متغیر مانند y، با استفاده از مقدار متغیر دیگری مانند x مواجه هستیم. برای این منظور باید ابتدا یک رابطه بین مقادیر x و y به دست آورده و با استفاده از آن، مقدار y را پیش‌بینی کنیم یا تخمین بزنیم. ساده‌ترین رابطه‌ای که می‌توان بین مقادیر دو متغیر در نظر گرفت رابطه خطی است که با استفاده از یک نمونه تصادفی که مجموعه زوج‌های مرتبی از مقادیر x و y است شرح دادیم. مدلی که برای رابطه x و y در نظر گرفتیم یک رابطه خطی به شکل زیر است:

$$y_i = \alpha + \beta x_i + e_i \qquad i = 1, 2, \dots, n$$

و باید پارامترهای β_0 و β_1 را برآورد کردیم که برای محاسبه این دو مقدار مجهول نیاز به دست آوردن فرمول‌های زیر

و خط رگرسیون کمترین مربعات عبارت است از:

$$\hat{y} = \hat{\alpha} + \hat{\beta}x$$

و در نهایت ضریب همبستگی x و y را از رابطه زیر به دست آوردیم.

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$ یا

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\left[n\sum x_i^2 - \left(\sum x_i\right)^2\right]\left[n\sum y_i^2 - \left(\sum y_i\right)^2\right]}}$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{j=1}^{n}(y_i - \bar{y})^2}}$$