برآورد فاصله‌ای یک و دونمونه‌ای

فردوس گرجی

برآورد نقطه‌ای

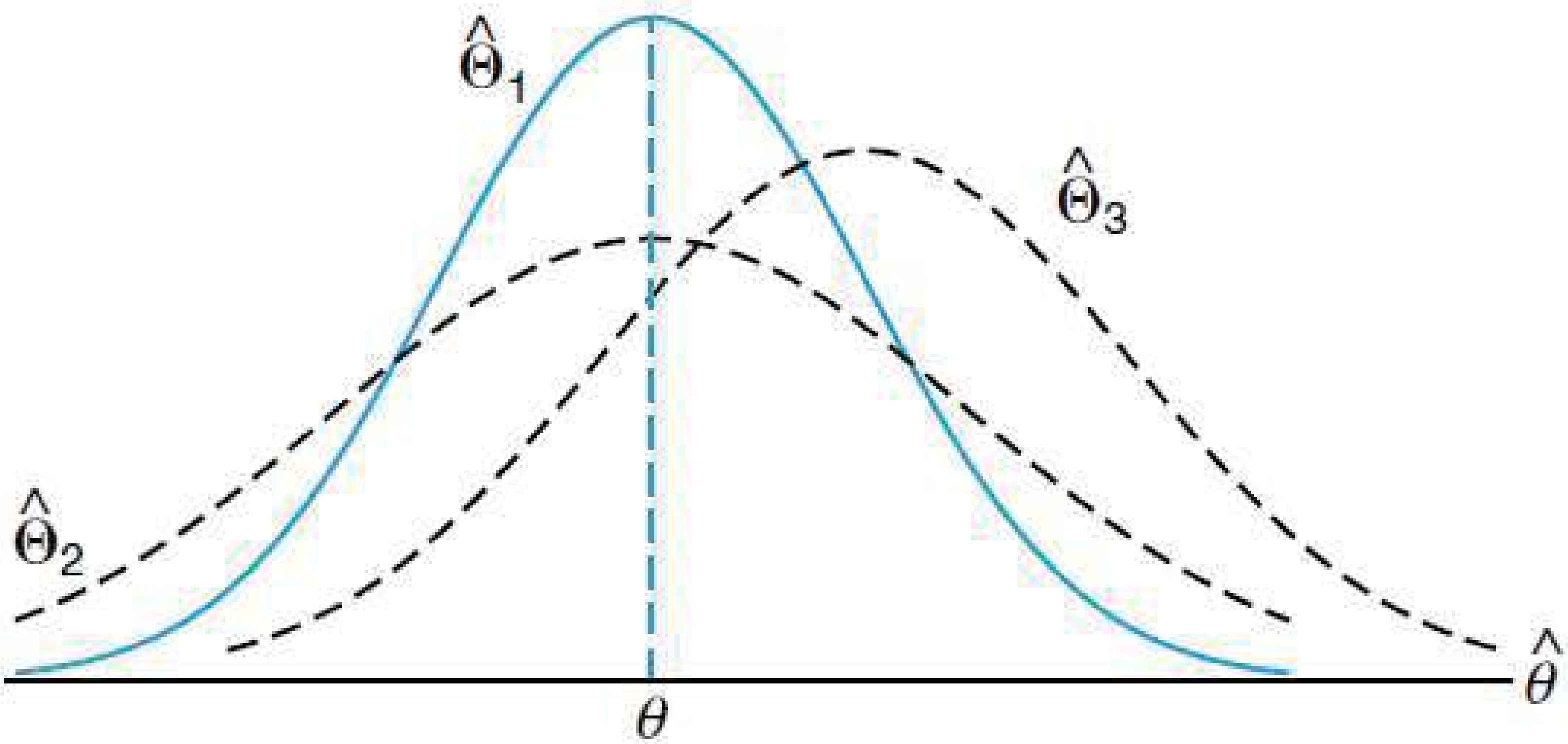
فرض کنید به دنبال یافتن پارامتر مجهول θ از جامعه هستیم. با تعریف آماره T از نمونه تصادفی به عنوان برآوردگر، مقدار به دست آمده، یعنی t را به عنوان برآورد θ یا $\hat{\theta}$ در نظر می‌گیریم.
مثلا مقدار آماره $\bar{X}$، یعنی $\bar{x}$ را به عنوان برآورد میانگین جامعه، یعنی $\hat{\mu}$ در نظر می‌گیریم. یا مقدار آماره $\frac{X}{n}$، یعنی $\frac{x}{n}$ را به عنوان برآورد پارامتر نسبت، $\hat{p}$ در آزمایش دوجمله‌ای در نظر می‌گیریم.

برآوردگر نااریب

اگر مقدار مورد انتظار برآوردگری برابر با پارامتر جامعه باشد، $\mu_T = E(T) = \theta$، آن را برآوردگر نااریب گوییم. مثلا برآوردگر $S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i - \bar{X})^۲$ برای واریانس جامعه ($\sigma^۲$) نااریب و برآوردگر $S_b^۲ = \frac{۱}{n}\sum_{i=۱}^{n}(X_i - \bar{X})^۲$ اریب است.

برآوردگر کارا

اگر $T_۱ = \hat{\theta}_۱$ و $T_۲ = \hat{\theta}_۲$ دو برآوردگر نااریب برای پارامتر θ باشند، معمولا برآوردگری مناسب‌تر است که واریانس کمتری داشته باشد. زیرا با تغییر نمونه تصادفی میزان تغییر مقدار آن احتمالا کمتر است.
اگر $\sigma^۲_{\hat{\theta}_۲} < \sigma^۲_{\hat{\theta}_۱}$، اصطلاحا می‌گوییم $\hat{\theta}_۱$ برآوردگری کاراتراز $\hat{\theta}_۲$ برای θ است.
تعریف: از بین تمام برآوردگرهای نااریب ممکن برای پارامتر θ، برآوردگری که کمترین واریانس را دارد، کاراترین برآوردگر برای θ نامیده می‌شود.
مثلا در توزیع نرمال، میانه نمونه $\tilde{X}$ و میانگین نمونه $\bar{X}$، هر دو برآورگری نااریب برای پارامتر میانگین جامعه، μ هستند. اما واریانس $\bar{X}$ کمتر و لذا برآوردگری کاراتر است.

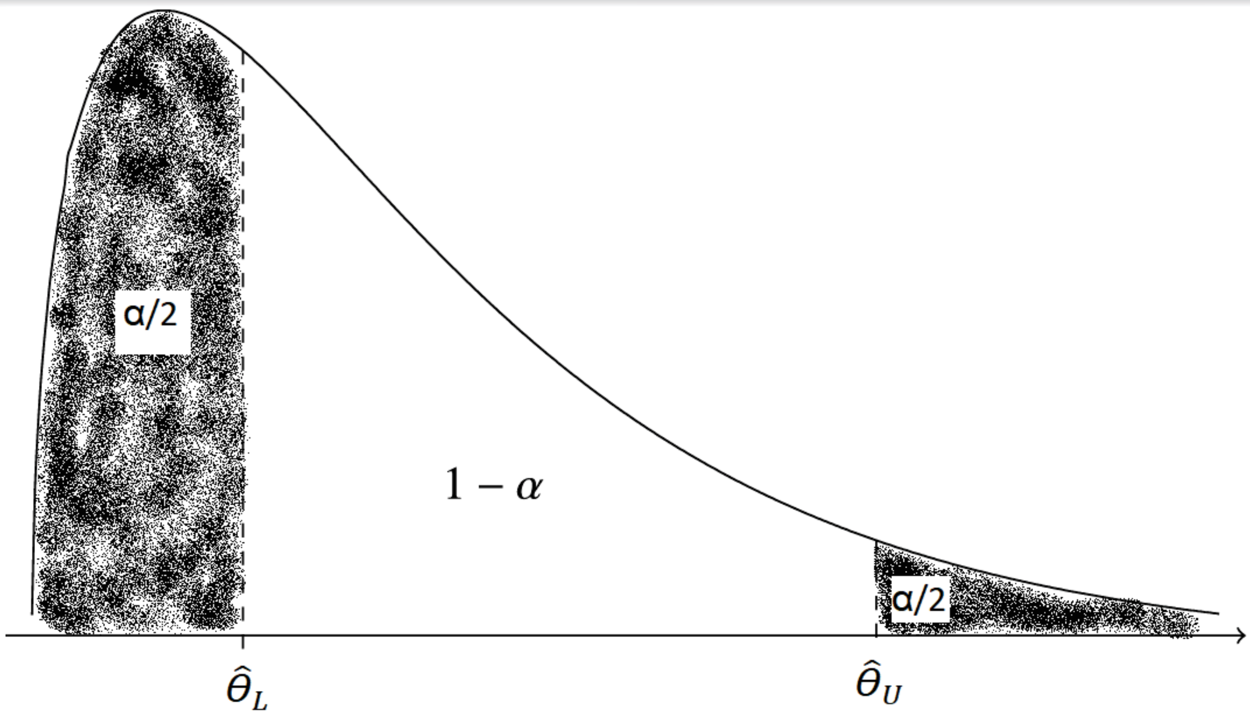

$\hat{\Theta}_1$
$\hat{\Theta}_3$
$\hat{\Theta}_2$
θ
$\hat{\theta}$

برآورد فاصله‌ای

حتی در کاراترین برآوردگر ها هم نمی‌توان انتظار داشت مقدار برآورد نقطه‌ای به دست آمده با برابر با مقدار واقعی پارامتر جامعه باشد. برای همین، در خیلی از مواقع از برآورد فاصله‌ای استفاده می‌کنیم. به این ترتیب که با استفاده از برآوردگر نقطه‌ای، بازه‌ای را معرفی می‌کنیم که با احتمال زیاد، مقدار واقعی پارامتر را در بر بگیرد.

$$\hat{\theta} \quad \rightarrow \quad \hat{\theta}_L, \hat{\theta}_U \quad \rightarrow \quad \hat{\theta}_L < \theta < \hat{\theta}_U$$

$$P(\hat{\theta}_L < \theta < \hat{\theta}_U) = ۱ - \alpha$$



برآورد فاصله‌ای برای میانگین جامعه

برآوردگر نقطه‌ای

می‌دانیم که اگر جامعه‌ای با میانگین μ و واریانس معلوم $\sigma^۲$ نرمال باشد و یا نرمال نباشد ولی تعداد اعضای نمونه، n، به اندازه کافی بزرگ باشد ($n \geq ۳۰$)، آماره $\hat{\mu} = \bar{X}$ که یک برآوردگر نااریب برای پارامتر میانگین جامعه (μ) است دارای توزیع نرمال با میانگین μ و واریانس $\frac{\sigma^۲}{n}$ می‌باشد. پس

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(۰, ۱)$$

برای برآورد فاصله‌ای μ می‌توان نوشت:

$$P(Z_L < Z < Z_U) = ۱ - \alpha$$

$$P(-z_{\frac{\alpha}{۲}} < Z < z_{\frac{\alpha}{۲}}) = ۱ - \alpha$$

$$\rightarrow P(-z_{\frac{\alpha}{۲}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\frac{\alpha}{۲}}) = ۱ - \alpha$$

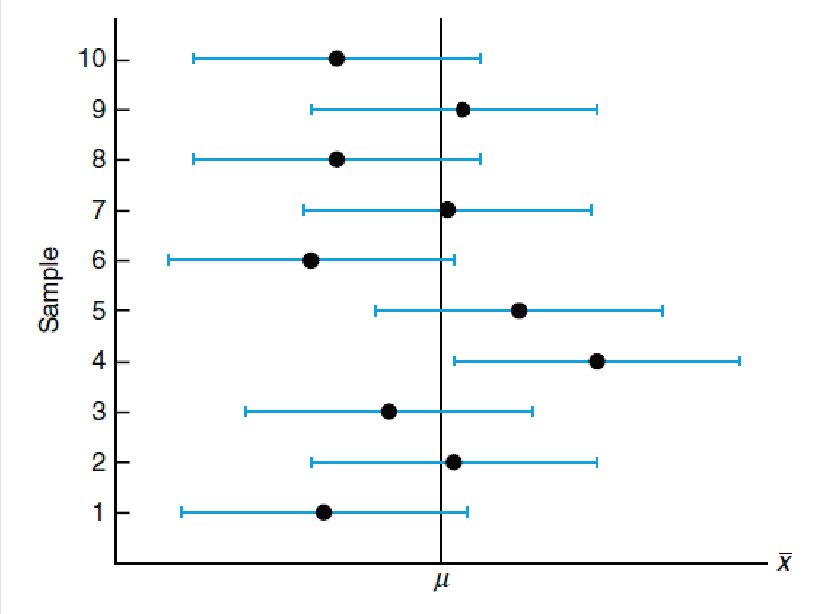
$$P(\underbrace{\bar{X} - z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}}_{\hat{\mu}_L} < \mu < \underbrace{\bar{X} + z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}}_{\hat{\mu}_U}) = ۱ - \alpha;$$

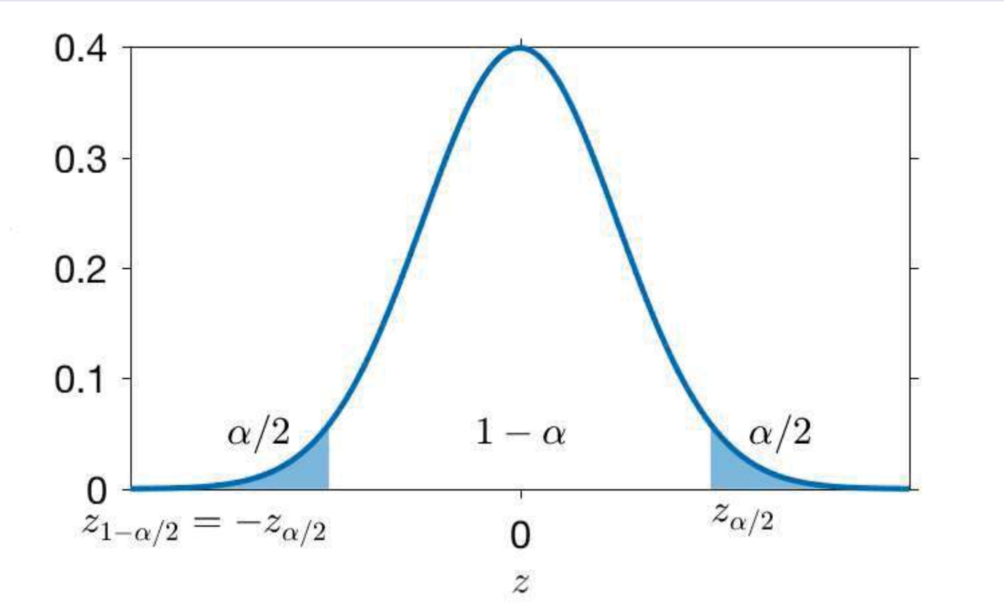
فاصله اطمینان μ وقتی $\sigma^۲$ معلوم است

اگر $\bar{X}$ میانگین یک نمونه تصادفی از اندازه n از جامعه‌ای با واریانس معلوم $\sigma^۲$ باشد، یک فاصله اطمینان $(۱-\alpha)$۱۰۰٪ی برای μ عبارت است از

$$\bar{X} - z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$$

که در آن $z_{\frac{\alpha}{۲}}$ مقدار zی است که مساحت سمت راست آن $\frac{\alpha}{۲}$ است.





مثال ۱

مقدار متوسط به دست آمده از اندازه‌گیری ماده ای در ۳۶ نقطه از یک رودخانه، ۲/۶ میلی‌گرم در هر میلی‌لیتر است. با فرض اینکه انحراف معیار مقدار این ماده در رودخانه (انحراف معیار جامعه) ۰/۳ است، فاصله اطمینان ۹۵ درصدی و ۹۹ درصدی را برای میانگین این ماده در رودخانه بیابید.

راه حل:

$$\text{فاصله اطمینان ۹۵ درصد} \rightarrow \alpha = ۰/۰۵ \rightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

$$\rightarrow z_{\frac{\alpha}{۲}} = z_{۰/۰۲۵} = ۱/۹۶, z_{-\frac{\alpha}{۲}} = z_{-۰/۰۲۵} = -۱/۹۶$$

$$\text{برآورد نقطه‌ای}\ \mu = \bar{x} = ۲/۶$$

$$\rightarrow ۲/۶ - ۱/۹۶(\frac{۰/۳}{\sqrt{۳۶}}) < \mu < ۲/۶ + ۱/۹۶(\frac{۰/۳}{\sqrt{۳۶}}) \quad \rightarrow \quad ۲/۵ < \mu < ۲/۷$$

$$\text{فاصله اطمینان ۹۹ درصد} \rightarrow \alpha = ۰/۰۱ \rightarrow \frac{\alpha}{۲} = ۰/۰۰۵$$

$$\rightarrow z_{\frac{\alpha}{۲}} = z_{۰/۰۰۵} = ۲/۵۷۵, z_{-\frac{\alpha}{۲}} = z_{-۰/۰۰۵} = -۲/۵۷۵$$

$$\text{برآورد نقطه‌ای}\ \mu = \bar{x} = ۲/۶$$

$$\rightarrow ۲/۶ - ۲/۵۷۵(\frac{۰/۳}{\sqrt{۳۶}}) < \mu < ۲/۶ + ۲/۵۷۵(\frac{۰/۳}{\sqrt{۳۶}}) \quad \rightarrow ۲/۴۷ < \mu < ۲/۷۳$$

مثال ۲

مقادیر اندازه‌گیری شده انرژی برخورد در ده نمونه برش فولاد A238 در دمای ۶۰ درجه سانتی گراد به صورت زیر است:

$$۶۴/۱, ۶۴/۷, ۶۴/۵, ۶۴/۶, ۶۴/۵, ۶۴/۳, ۶۴/۶, ۶۴/۸, ۶۴/۲, ۶۴/۳$$

فرض کنید انرژی برخورد چنین نمونه‌هایی دارای توزیع نرمال با انحراف معیار ۱ ژول است. یک فاصله اطمینان ۹۵ درصدی برای میانگین توزیع انرژی برخورد بیابید.

راه حل: با استفاده از انرژی‌های اندازه‌گیری شده داریم: $\bar{X} = ۶۴/۴۶$

$$\text{فاصله اطمینان ۹۵ درصد} \rightarrow \alpha = ۰/۰۵ \rightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

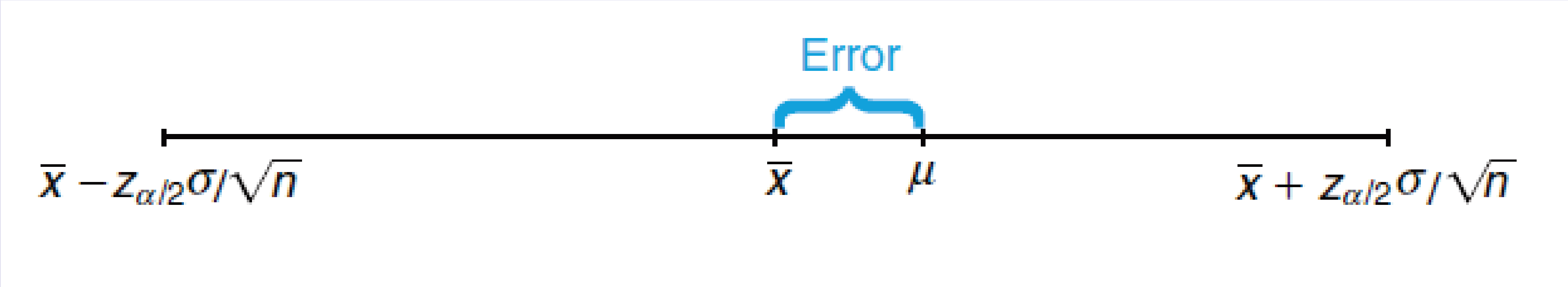
$$\rightarrow z_{\frac{\alpha}{۲}} = z_{۰/۰۲۵} = ۱/۹۶, z_{-\frac{\alpha}{۲}} = z_{-۰/۰۲۵} = -۱/۹۶$$

$$\rightarrow ۶۴/۴۶ - ۱/۹۶(\frac{۱}{\sqrt{۱۰}}) < \mu < ۶۴/۴۶ + ۱/۹۶(\frac{۱}{\sqrt{۱۰}})$$

$$\rightarrow \quad ۶۳/۸۴ < \mu < ۶۵/۰۸$$

قضیه

اگر $\bar{x}$ به عنوان برآورد از μ به کار برده شود، آنگاه می‌توانیم $(۱ - \alpha)۱۰۰\%$ مطمئن باشیم که خطا از $z_{\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}$ بیشتر نیست.



قضیه

اگر $\bar{x}$ به عنوان برآورد از μ به کار برده شود و واریانس جامعه معلوم باشد، اگر اندازه نمونه

$$n = \left(\frac{z_{\frac{\alpha}{۲}}\sigma}{e}\right)^{۲}$$

باشد، آنگاه می‌توانیم $(۱ - \alpha)۱۰۰\%$ مطمئن باشیم که خطا از مقدار مشخص از پیش تعریف شده e بیشتر نمی‌شود.

مثال ۳

در مثال ۱، اندازه نمونه‌ها چقدر باید باشد اگر بخواهیم ۹۵ درصد مطمئن باشیم که خطای برآورد کمتر از ۰/۰۵ است؟

راه حل:

$$\sigma = ۰/۳$$

$$n \left(\frac{(۱/۹۶)(۰/۳)}{۰/۰۵} \right)^{۲} = ۱۳۸/۳ \simeq ۱۳۹$$

توجه کنید که در مثال قبل، با تعداد نمونه ۳۶، با احتمال ۹۵ درصد مطمئنیم که خطا از ۰/۰۹۸ کمتر است.

برآورد فاصله‌ای برای میانگین جامعه، μ

حال اگر جامعه‌ای با توزیع نرمال و میانگین μ و واریانس نا معلوم باشد، آماره $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ دارای توزیع t با $n-۱$ درجه آزادی می‌باشد. پس

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim T(n-۱)$$

برای برآورد فاصله‌ای μ می‌توان نوشت:

$$P(T_L < T < T_U) = ۱ - \alpha$$
$$P(-t_{\frac{\alpha}{۲}} < T < t_{\frac{\alpha}{۲}}) = ۱ - \alpha$$
$$\rightarrow P(-t_{\frac{\alpha}{۲}} < \frac{\bar{X} - \mu}{S/n} < t_{\frac{\alpha}{۲}}) = ۱ - \alpha$$
$$P(\underbrace{\bar{X} - t_{\frac{\alpha}{۲}}\frac{S}{\sqrt{n}}}_{\hat{\mu}_L} < \mu < \underbrace{\bar{X} + t_{\frac{\alpha}{۲}}\frac{S}{\sqrt{n}}}_{\hat{\mu}_U}) = ۱ - \alpha;$$

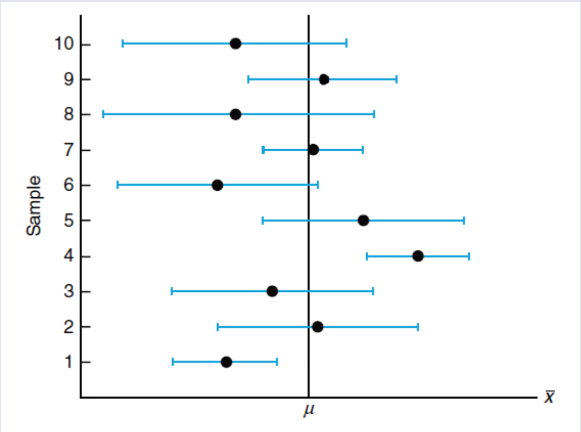
برآورد فاصله‌ای برای میانگین جامعه، μ

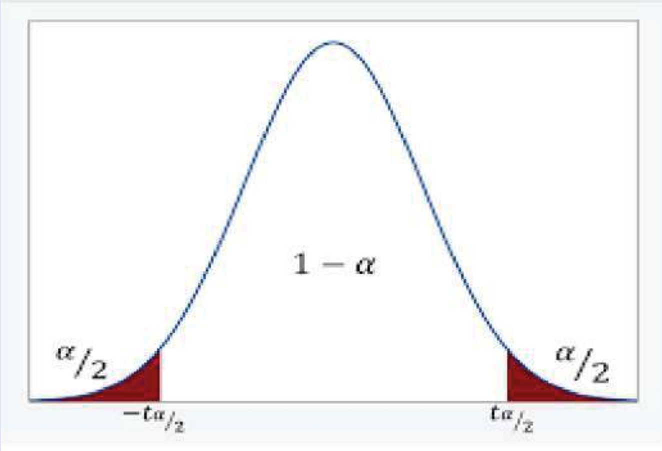
فاصله اطمینان μ وقتی $\sigma^{۲}$ نامعلوم است

اگر $\bar{X}$ و S به ترتیب میانگین و انحراف معیار یک نمونه تصادفی از اندازه n از جامعه‌ای نرمال با واریانس نامعلوم باشد، یک فاصله اطمینان $(۱-\alpha)$٪۱۰۰ی برای μ عبارت است از

$$\bar{X} - t_{\frac{\alpha}{۲}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{۲}} \frac{S}{\sqrt{n}}$$

که در آن $t_{\frac{\alpha}{۲}}$ مقدار tی است که مساحت سمت راست آن $\frac{\alpha}{۲}$ است.





مثال ۵

کشاورزی به تصادف ده هندوانه از مزرعه‌اش را وزن می‌کند و اندازه‌های زیر را بر حسب پوند به دست می‌آورد:

$$۷/۷۲, ۹/۵۸, ۱۲/۳۸, ۷/۷۷, ۱۱/۲۷, ۸/۸۰, ۱۱/۱۰, ۷/۸۰, ۱۰/۱۷, ۶/۰۰$$

با فرض این که وزن هندوانه‌ها دارای توزیع نرمال است، یک فاصله اطمینان ۹۵ درصدی برای میانگین وزن هندوانه‌های مزرعه بیابید.

راه حل:

با استفاده از وزن‌های اندازه‌گیری شده داریم: $\bar{X} = ۹/۲۵۹, S^{۲} = ۳/۹۶۱۵$

ضمنا $\alpha = ۰/۰۵$ و با توجه به حجم نمونه، $n = ۱۰$ و این که واریانس جامعه نامعلوم است، از توزیع T با ۹ درجه آزادی استفاده می‌کنیم؛ $t_{\frac{\alpha}{۲}} = t_{۰/۰۲۵} \simeq ۲/۲۶۲$

$$\left(\bar{X} - t_{\frac{\alpha}{۲},n-۱}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{۲},n-۱}\frac{S}{\sqrt{n}}\right)$$

$$\rightarrow \left(۹/۲۵۹ - ۲/۲۶۲ \times \frac{\sqrt{۳/۹۶۱۵}}{\sqrt{۱۰}} < \mu < ۹/۲۵۹ + ۲/۲۶۲ \times \frac{\sqrt{۳/۹۶۱۵}}{\sqrt{۱۰}}\right)$$

$$\rightarrow (۷/۸۳۵۳ < \mu < ۱۰/۶۸۲۷)$$

مثال ۶

با اندازه‌گیری قطر یک نمونه تصادفی ۲۰ تایی از میلگردهای تولیدی یک کارخانه ، مقادیر $\bar{X} = ۱۵/۶$ و $S^{۲} = ۸/۴$ به دست آمده‌اند. مطلوب است فاصله اطمینان ۹۹ درصدی برای میانگین قطر میلگردهای تولید شده در این کارخانه.

راه حل: داریم $\alpha = ۰/۰۱$ و با توجه به حجم نمونه، $n = ۲۰$ و این که واریانس جامعه نامعلوم است، از توزیع T با ۱۹ درجه آزادی استفاده می‌کنیم؛ $t_{\frac{\alpha}{۲}} = t_{۰/۰۰۵} \simeq ۲/۸۶۱$

$$\left(\bar{X} - t_{\frac{\alpha}{۲},n-۱}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{۲},n-۱}\frac{S}{\sqrt{n}}\right)$$

$$\rightarrow \left(۱۵/۶ - ۲/۸۶۱ \times \frac{\sqrt{۸/۴}}{\sqrt{۲۰}} < \mu < ۱۵/۶ + ۲/۸۶۱ \times \frac{\sqrt{۸/۴}}{\sqrt{۲۰}}\right)$$

$$\rightarrow (۱۳/۷۴۵۹ < \mu < ۱۷/۴۵۴۱)$$

خطای استاندارد برآوردگر نقطه‌ای و طول فاصله اطمینان

طول فاصله اطمینان به دست آمده با خطای استاندارد برآوردگر نقطه‌ای (یعنی انحراف معیار آن)، رابطه دارد.

در حالت $\sigma^۲$ معلوم:

$$\bar{x} \pm z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{۲}} s.e.(\bar{x})$$

در حالت $\sigma^۲$ نامعلوم:

$$\bar{x} \pm t_{\frac{\alpha}{۲}} \frac{s}{\sqrt{n}} = \bar{x} \pm t_{\frac{\alpha}{۲}} s.e.(\bar{x})$$

برآورد فاصله‌ای برای تفاضل میانگین دو جامعه، $\mu_۱ - \mu_۲$

فرض کنید دو جامعه داشته باشیم.
جامعه‌ی اول دارای میانگین $\mu_۱$ و واریانس $\sigma_۱^۲$ باشد.
جامعه‌ی دوم دارای میانگین $\mu_۲$ و واریانس $\sigma_۲^۲$ باشد.
یک نمونه‌ی تصادفی n تایی $X_۱, \ldots, X_n$ از جامعه‌ی اول انتخاب کرده و میانگین این نمونه را با $\bar{X}$ و واریانس آن را با $S_۱^۲$ نمایش می‌دهیم.
یک نمونه‌ی تصادفی m تایی $Y_۱, \ldots, Y_m$ از جامعه‌ی دوم انتخاب کرده و میانگین این نمونه را با $\bar{Y}$ و واریانس آن را با $S_۲^۲$ نشان می‌دهیم.
فرض کنید نمونه‌گیری از دو جامعه مستقل از یکدیگر باشد.
می‌خواهیم فاصله اطمینانی برای $\mu_۱ - \mu_۲$ پیدا کنیم.

فاصله اطمینان اختلاف میانگین‌ها $\mu_1 - \mu_2$

حالت اول: واریانس دو جامعه σ_1^2 و σ_2^2 معلوم باشد

می‌دانیم در صورت نرمال بودن جامعه‌ها و یا بزرگ بودن اندازه نمونه ها داریم:

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right) \to Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0, 1)$$

پس

$$P(-z_{\frac{\alpha}{2}} < Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} < z_{\frac{\alpha}{2}}) = 1 - \alpha$$

فاصله اطمینان $\mu_1 - \mu_2$ وقتی σ_1^2 و σ_2^2 معلوم است

اگر $\bar{X}_1$ و $\bar{X}_2$ میانگین‌های نمونه‌های تصادفی مستقل از اندازه های n و m از جامعه‌هایی با واریانس‌های معلوم σ_1^2 و σ_2^2 باشند، فاصله اطمینان $100(1-\alpha)$٪ی برای $\mu_1 - \mu_2$ عبارت است از

$$(\bar{X} - \bar{Y}) - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

که در آن $z_{\frac{\alpha}{2}}$ مقدار zی است که مساحت سمت راست آن $\frac{\alpha}{2}$ است.

مثال ۷

برای مقایسه دو موتور خودرو A و B، در شرایط یکسان، مسافت پیموده شده آن‌ها با مقدار بنزین مشخصی بر حسب مایل بر گالن اندازه‌گیری شده است. ۵۰ آزمایش با موتور A و ۷۵ آزمایش با موتور B انجام شده است. مقدار متوسط مسافت طی شده با این موتورها به ترتیب ۳۶ و ۴۲ مایل بر گالن برای A و B به دست آمده است. یک فاصله اطمینان ۹۶ درصدی برای $\mu_B - \mu_A$ بیابید. فرض کنید انحراف معیار جامعه برای مسافت پیموده شده دو موتور A و B به ترتیب ۶ و ۸ مایل بر گالن است.

راه حل:

$$\bar{x}_B - \bar{x}_A = ۴۲ - ۳۶ = ۶, \quad \alpha = ۰/۰۴ \to z_{\frac{\alpha}{۲}} = z_{۰/۰۲} = ۲/۰۵$$

$$۶ - ۲/۰۵\sqrt{\frac{۶۴}{۷۵} + \frac{۳۶}{۵۰}} < \mu_B - \mu_A < ۶ + ۲/۰۵\sqrt{\frac{۶۴}{۷۵} + \frac{۳۶}{۵۰}}$$

$$\to ۳/۴۳ < \mu_B - \mu_A < ۸/۵۷$$

حالت دوم: واریانس دو جامعه نامعلوم، اما با یکدیگر مساوی باشد، $\sigma_1^2 = \sigma_2^2$

می‌دانیم در صورت نرمال بودن جامعه‌ها داریم:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim T(n+m-2)$$

پس

$$P(-t_{\frac{\alpha}{2}} < T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} < t_{\frac{\alpha}{2}}) = 1 - \alpha$$

فاصله اطمینان $\mu_1 - \mu_2$ وقتی σ_1^2 و σ_2^2 نامعلوم هستند ولی $\sigma_1^2 = \sigma_2^2$

اگر $\bar{X}_1$ و $\bar{X}_2$ میانگین‌های نمونه‌های تصادفی مستقل از اندازه های n و m از جامعه‌های تقریبا نرمال با واریانس‌های نامعلوم اما مساوی باشند، فاصله اطمینان $(1-\alpha)100\%$ی برای $\mu_1 - \mu_2$ عبارت است از

$$(\bar{X} - \bar{Y}) - t_{\frac{\alpha}{2}} S_p\sqrt{\frac{1}{n} + \frac{1}{m}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + t_{\frac{\alpha}{2}} S_p\sqrt{\frac{1}{n} + \frac{1}{m}}$$

که در آن $t_{\frac{\alpha}{2}}$ مقدار ی است که مساحت سمت راست آن $\frac{\alpha}{2}$ است.

مثال ۸

دو نمونه تصادفی مستقل به اندازه‌های $n = ۱۰$ و $m = ۸$ از نوع خاصی از قطعه تراشکاری شده که توسط دو دستگاه A و B به دست آمده‌اند، داریم. میانگین و انحراف معیار طول قطعات نمونه حاصل از دستگاه A $\bar{x}_A = ۳/۱, S_۱ = ۰/۵$ و میانگین و انحراف معیار طول قطعات نمونه حاصل از دستگاه B $\bar{x}_B = ۲/۷, S_۲ = ۰/۷$ است. با فرض نرمال بودن توزیع طول قطعات و برابری واریانس آن‌ها، یک فاصله اطمینان ۹۵ درصدی برای $\mu_A - \mu_B$ به دست آورید.

راه حل:

$$\bar{x}_A - \bar{x}_B = ۳/۱ - ۲/۷ = ۰/۴, \quad \alpha = ۰/۰۵ \to t_{\alpha/۲} = ۲/۱۲$$

$$S_p = ۰/۵۹۵۸$$

$$۰/۴ - ۲/۱۲ S_p\sqrt{\frac{۱}{۱۰} + \frac{۱}{۸}} < \mu_A - \mu_B < ۰/۴ + ۲/۱۲ S_p\sqrt{\frac{۱}{۱۰} + \frac{۱}{۸}}$$

$$\to ۰/۲ < \mu_A - \mu_B < ۱$$

نکته

در فاصله اطمینان تفاضل میانگین‌ها با فرض واریانس‌های مجهول و مساوی، دو جامعه باید نرمال باشند. اما کمی عدول از فرض برابری واریانس‌ها و یا نرمال بودن جامعه‌ها، درجه اطمینان فاصله به دست آمده به طور جدی تغییر نمی کند. خصوصا اگر دو جامعه نرمال باشند ولی واریانس‌های مجهول و نابرابر داشته باشند، با شرط تساوی اندازه نمونه‌ها، هنوز هم نتایج به دست آمده معقول هستند.

نکته

در مورد فاصله اطمینان برای تفاضل میانگین دو جامعه، $\mu_۱ - \mu_۲$، اگر حدود فاصله به دست آمده هردو مثبت (یا هر دو منفی) باشند، با کمی خطا می توان ادعا کرد که میانگین جامعه اول از میانگین جامعه دوم بیشتر (یا کمتر) است.

فاصله اطمینان نسبت، p، در آزمایش دوجمله‌ای

کاربردها

تخمین میزان درصد افرادی که به یک شخص در انتخابات رای می‌دهند.
تخمین درصد قطعات معیوب تولید شده در یک کارخانه
تخمین احتمال بهبودی پس از دریافت نوعی دارو

برآورد نقطه‌ای

$$\hat{P} = \frac{X}{n}, \quad X = \text{تعداد موفقیت‌ها در } n \text{ آزمایش}$$

نکته

فرض می‌کنیم مقدار واقعی نسبت جامعه، خیلی نزدیک صفر یا یک نباشد، خصوصا وقتی n کوچک است.
به طور کلی برای اطمینان، لازم است $n\hat{p}$ و $n\hat{q}$ هردو بزرگتر یا مساوی با ۵ باشند.

فرض کنیم

$$Y_۱, Y_۲, \ldots, Y_n \sim Bernoulli, \quad Y_i = ۰ \text{ یا } ۱$$

داریم:

$$\hat{P} = \frac{X}{n} = \frac{\sum_{i=۱}^{n} Y_i}{n} = \bar{Y} \quad (\text{میانگین نمونه‌ای } Y_i \text{ها})$$

$$n \text{ بزرگ} \quad \rightarrow \quad \hat{P} \sim N(\mu_{\hat{P}}, \sigma^۲_{\hat{P}}) \qquad \text{طبق قضیه حد مرکزی}$$

$$\mu_{\hat{P}} = E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{np}{n} = p$$

$$\sigma^۲_{\hat{P}} = \sigma^۲_{\frac{X}{n}} = \frac{۱}{n^۲}\sigma^۲_X = \frac{npq}{n^۲} = \frac{pq}{n}$$

بنابراین

$$Z = \frac{\hat{P} - p}{\sqrt{pq/n}} \sim N(۰, ۱)$$

$$P(-z_{\frac{\alpha}{۲}} < Z < z_{\frac{\alpha}{۲}}) = P(-z_{\frac{\alpha}{۲}} < \frac{\hat{P} - p}{\sqrt{pq/n}} < z_{\frac{\alpha}{۲}}) = ۱ - \alpha$$

$$P\left(\hat{P} - z_{\frac{\alpha}{۲}}\sqrt{\frac{pq}{n}} < p < \hat{P} + z_{\frac{\alpha}{۲}}\sqrt{\frac{pq}{n}}\right) = ۱ - \alpha$$
$$P\left(\hat{P} - z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{P} + z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = ۱ - \alpha$$

فاصله اطمینان برای p در نمونه بزرگ

اگر $\hat{p}$ نسبت موفقیت‌ها در نمونه‌ای تصادفی از اندازه n باشد و $\hat{q} = ۱ - \hat{p}$، فاصله اطمینان تقریبی $(۱ - \alpha)$۱۰۰٪ی برای پارامتر دوجمله‌ای p عبارت است از

$$\hat{P} - z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{P} + z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

با شرط $n\hat{p}, n\hat{q} \geq ۵$

نکته

وقتی توزیع فوق هندسی را با توزیع دوجمله‌ای تقریب می‌زنیم، باز هم می‌توانیم از فاصله اطمینان فوق برای پارامتر دوجمله‌ای استفاده کنیم.

مثال ۱۲

نمونه‌ای تصادفی متشکل از ۴۰۰ قطعه تولید شده در یک کارخانه را انتخاب و تست کرده و فهمیدیم که ۳۲ قطعه معیوب است. یک فاصله اطمینان ۹۵ درصدی برای نسبت واقعی قطعات معیوب بیابید.

راه حل:

$$\hat{p} = \frac{x}{n} = \frac{۳۲}{۴۰۰} = ۰/۰۸; \hat{q} = ۰/۹۲; \alpha = ۰/۰۵; z_{\frac{\alpha}{۲}} = ۱/۹۶$$

$$۰/۰۸ - ۱/۹۶\sqrt{\frac{۰/۰۸ \times ۰/۹۲}{۴۰۰}} < p < ۰/۰۸ + ۱/۹۶\sqrt{\frac{۰/۰۸ \times ۰/۹۲}{۴۰۰}}$$

$$۰/۰۵۳ < p < ۰/۱۰۷$$

مثال ۱۳

برای کشف میزان اثربخشی یک داروی جدید، از بین بیماران داوطلب، ۵۰۰ نفر به تصادف انتخاب شده و دارو را مصرف کردند. پس از اتمام دوره درمان، مشاهده شد که ۴۲۱ نفر بهبود یافتند. یک فاصله اطمینان ۹۵ درصدی برای احتمال اثربخشی دارو (نسبت واقعی افرادی که بهبود پیدا می‌کنند به کل بیماران) بیابید.

راه حل:

$$\hat{p} = \frac{۴۲۱}{۵۰۰} = ۰/۸۴۲; \hat{q} = ۰/۱۵۸; \alpha = ۰/۰۵; z_{\frac{\alpha}{۲}} = ۱/۹۶; n = ۵۰۰$$

$$۰/۸۴۲ - ۱/۹۶\sqrt{\frac{۰/۸۴۲ \times ۰/۱۵۸}{۵۰۰}} < p < ۰/۸۴۲ + ۱/۹۶\sqrt{\frac{۰/۸۴۲ \times ۰/۱۵۸}{۵۰۰}}$$

$$۰/۸۱۰ < p < ۰/۸۷۴$$

مثال ۱۴

در یک نمونه تصادفی از ۵۰۰ خانواده ساکن یک شهر، معلوم شده است که ۳۴۰ خانواده از بینندگان ثابت یک برنامه تلویزیونی در شبکه استانی خود هستند. یک فاصله اطمینان ۹۹ درصدی برای نسبت واقعی تمام خانواده]ایی که از بینندگان ثابت برنامه مذکور هستند بیابید (تقریب فوق هندسی با دوجمله‌ای).

راه حل:

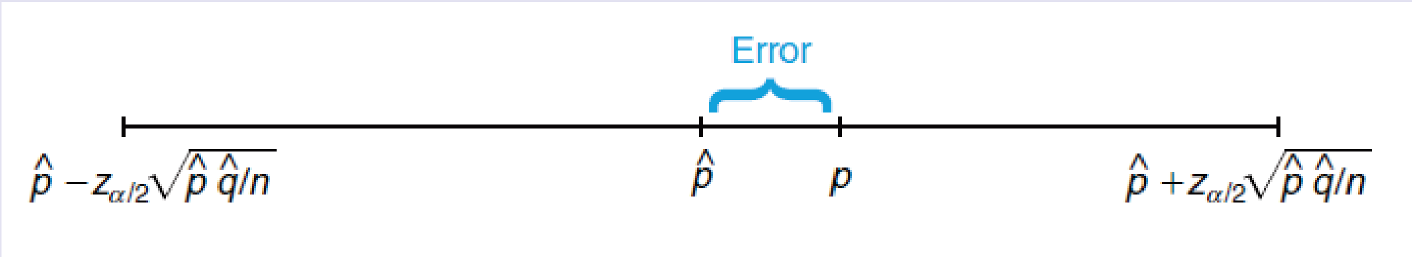
$$\hat{p} = \frac{۳۴۰}{۵۰۰} = ۰/۶۸; \hat{q} = ۰/۳۲; \alpha = ۰/۰۱; z_{\frac{\alpha}{۲}} = ۲/۵۷۵; n = ۵۰۰$$

$$۰/۶۸ - ۲/۵۷۵\sqrt{\frac{۰/۶۸ \times ۰/۳۲}{۵۰۰}} < p < ۰/۶۸ + ۲/۵۷۵\sqrt{\frac{۰/۶۸ \times ۰/۳۲}{۵۰۰}}$$

$$۰/۶۲۶ < p < ۰/۷۳۴$$

قضیه

اگر $\hat{p}$ به عنوان برآورد p به کار برده شود، آنگاه می‌توانیم $(۱-\alpha)۱۰۰٪$ مطمئن باشیم که خطا از $z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$ بیشتر نیست.



مثال ۱۵

در مثال ۱۲، ۹۵ درصد اطمینان داریم که خطا از $۰/۰۲۷ = ۱/۹۶\sqrt{\frac{۰/۰۸\times۰/۹۲}{۴۰۰}}$ بیشتر نمی شود.

مثال ۱۶

در مثال ۱۳، ۹۵ درصد اطمینان داریم که خطا از $۰/۰۳۲ = ۱/۹۶\sqrt{\frac{۰/۸۴۲\times۰/۱۵۸}{۵۰۰}}$ بیشتر نمی شود.

مثال ۱۷

در مثال ۱۴، ۹۹ درصد اطمینان داریم که خطا از $۰/۰۵۴ = ۲/۵۷۵\sqrt{\frac{۰/۶۸\times۰/۳۲}{۵۰۰}}$ بیشتر نمی شود.

قضیه

اگر $\hat{p}$ به عنوان برآورد p به کار برده شود، می‌توانیم $(\alpha - ۱)$٪۱۰۰ مطمئن باشیم وقتی اندازه نمونه تقریبا $n = \frac{z^{۲}_{\frac{\alpha}{۲}}\hat{p}\hat{q}}{e^{۲}}$ است، خطا از مقدار مشخص از پیش تعریف شده e بیشتر نمی‌شود.

نکته

در قضیه فوق نیاز است برای تعیین اندازه نمونه، ابتدا برآورد خامی از p داشته باشیم. در غیر این صورت می‌توانیم نمونه‌ای مقدماتی با اندازه $n \geq ۳۰$ برای برآورد اولیه p داشته باشیم و سپس مقدار n را برای خطای داده شده به دست آوریم و یا از قضیه بعد استفاده کنیم.

مثال ۱۸

در مثال ۱۳، تعداد نمونه‌ها چقدر باشد تا ۹۵ درصد مطمئن باشیم که خطا از ۰/۰۲ کمتر است؟

راه حل:

$$n = \frac{(۱/۹۶)^{۲}(۰/۸۴۲)(۰/۱۵۸)}{۰/۰۲^{۲}} = ۱۲۷۷/۶۸ \simeq ۱۲۷۸$$

مثال ۱۹

در مثال ۱۴، تعداد نمونه‌ها چقدر باشد تا ۹۹ درصد مطمئن باشیم که خطا از ۰/۰۳ کمتر است؟

راه حل:

$$n = \frac{(۲/۵۷۵)^{۲}(۰/۶۸)(۰/۳۲)}{۰/۰۳^{۲}} = ۱۶۰۳/۱۳۸ \simeq ۱۶۰۴$$

نکته

گاهی به دست آوردن برآورد اولیه برای p و یا انتخاب اولیه حدسی p جهت تعیین اندازه نمونه دشوار و یا حتی غیر عملی است. در چنین مواقعی، یک کران بالا برای n در نظر گرفته می‌شود که خطای حاصل، از مقدار مورد نظر با درجه اطمینان لازم بیشتر نشود.

برای تعداد لازم نمونه داریم: $n = \frac{z^2_{\frac{\alpha}{2}}\hat{p}\hat{q}}{e^2}$. این مقدار به ازای $p = \frac{1}{2}$ بیشترین مقدار ممکن را دارد. زیرا با مشتق گیری از آن نسبت به $\hat{p}$ داریم:

$$\frac{d}{d\hat{p}}\frac{z^2_{\frac{\alpha}{2}}\hat{p}(1-\hat{p})}{e^2} = \frac{z^2_{\frac{\alpha}{2}}}{e^2}(-2p+1) = 0 \quad \rightarrow \quad p = \frac{1}{2}$$

$$\frac{d^2}{d\hat{p}^2}\frac{z^2_{\frac{\alpha}{2}}\hat{p}(1-\hat{p})}{e^2} = \frac{z^2_{\frac{\alpha}{2}}}{e^2}(-2) \leq 0 \quad \rightarrow$$ نقطه اکسترمم، ماکسیمم نسبی است.

$$\max n = \frac{z^2_{\frac{\alpha}{2}}}{4e^2}.$$

قضیه

اگر $\hat{p}$ به عنوان برآورد p به کار برده شود، وقتی اندازه نمونه $n = \frac{z^2_{\frac{\alpha}{2}}}{4e^2}$ است، می‌توانیم $100(1-\alpha)\%$ مطمئن باشیم که خطا از مقدار مشخص از پیش تعریف شده e بیشتر نمی‌شود.

مثال ۲۰

در مثال ۱۳، اندازه نمونه چقدر باشد تا حداقل ۹۵ درصد مطمئن باشیم که خطای برآورد ما از ۰/۰۲ کمتر است؟

راه حل: اگر هیچ نمونه مقدماتی برای برآورد p و یا هیچ حدس اولیه‌ای از آن نداشته باشیم، اندازه لازم برای نمونه به صورت زیر محاسبه می‌شود:

$$n = \frac{z^{۲}_{\frac{\alpha}{۲}}}{۴e^{۲}} = \frac{۱/۹۶^{۲}}{۴(۰/۰۲)^{۲}} = ۲۴۰۱$$

مثال ۲۱

در مثال ۱۴، اندازه نمونه چقدر باشد تا حداقل ۹۹ درصد مطمئن باشیم که خطای برآورد ما از ۰/۰۳ کمتر است؟

راه حل: در صورت نداشتن نمونه مقدماتی برای برآورد p و یا حدس اولیه‌ای برای آن می‌نویسیم:

$$n = \frac{z^{۲}_{\frac{\alpha}{۲}}}{۴e^{۲}} = \frac{۲/۵۷۵^{۲}}{۴(۰/۰۳)^{۲}} = ۱۸۴۱/۸۴ \simeq ۱۸۴۲$$

مثال ۲۲

در یک نمونه تصادفی از سالمندان یک شهر، مشاهده شد که ۱۲۱۹ نفر دارای مشکل فشار خون، و ۲۳۱۳ نفر دارای فشار خون طبیعی هستند. یک فاصله اطمینان ۹۵ درصدی برای درصد واقعی سالمندانی که در این شهر مشکل فشار خون دارند بیابید. حداکثر میزان خطا با ضریب اطمینان ۹۵ درصد چقدر است؟ حجم نمونه چقدر باید باشد تا با احتمال ۹۵ درصد، خطا از ۰/۰۱ کمتر باشد.

راه حل:

$$\hat{p} = \frac{x}{n} = \frac{۱۲۱۹}{۳۵۳۲} = ۰/۳۴۵; \quad \hat{q} = ۱ - ۰/۳۴۵ = ۰/۶۵۵$$

$$\alpha = ۰/۰۵, \quad z_{\frac{\alpha}{۲}} = z_{۰/۰۲۵} = ۱/۹۶;$$

$$۰/۳۴۵ - ۱/۹۶\sqrt{\frac{۰/۳۴۵ \times ۰/۶۵۵}{۳۵۳۲}} < p < ۰/۳۴۵ + ۱/۹۶\sqrt{\frac{۰/۳۴۵ \times ۰/۶۵۵}{۳۵۳۲}}$$

$$۰/۳۲۹ < p < ۰/۳۶۱, \quad CI : ۰/۳۴۵ \pm ۰/۰۱۶$$

با احتمال ۹۵ درصد مطمئنیم که خطا از ۰/۰۱۶ بیشتر نیست.

$$n_۱ = \frac{۱/۹۶^۲ \times ۰/۳۴۵ \times ۰/۶۵۵}{۰/۰۱^۲} = ۸۶۸۱/۰۵۵۶ \simeq ۸۶۸۲$$ با استفاده از نمونه قبلی

$$n_۲ = \frac{۱/۹۶^۲}{۴ \times ۰/۰۱^۲} = ۹۶۰۴$$ بدون فرض اولیه

فاصله اطمینان برای تفاضل بین دو نسبت

فرض کنید دو جامعه داشته باشیم.
جامعه‌ی اول دارای پارامتر نسبت p_1 باشد.
جامعه‌ی دوم دارای پارامتر نسبت p_2 باشد.
یک نمونه‌ی تصادفی n_1 تایی از جامعه‌ی اول انتخاب کرده و با شمارش تعداد موفقیت‌ها در آن، یعنی x_1 برآورد نقطه‌ای $\hat{p} = \frac{x_1}{n_1}$ را می‌سازیم.
یک نمونه‌ی تصادفی n_2 تایی از جامعه‌ی دوم انتخاب کرده و با شمارش تعداد موفقیت‌ها در آن، یعنی x_2 برآورد نقطه‌ای $\hat{p} = \frac{x_2}{n_2}$ را می‌سازیم.
فرض کنید نمونه‌گیری از دو جامعه مستقل از یکدیگر باشد.
می‌خواهیم فاصله اطمینانی برای $p_1 - p_2$ پیدا کنیم.
فرض می‌کنیم که n_1p_1، n_1q_1، n_2p_2 و n_2q_2 همگی بزرگتر یا مساوی ۵ باشند.

کاربرد

بررسی اختلاف درصد افراد مبتلا به بیماری‌های ریوی در افراد سیگاری و غیرسیگاری
مقایسه درصد قبولی کنکور در دو مدرسه
درصد کالاهای معیوب در یک خط تولید قبل و بعد از تغییرات در دستگاه‌ها

می دانیم اگر n_1 و n_2 بزرگ باشند، طبق قضیه حد مرکزی، توزیع $\hat{p}_1$ و $\hat{p}_2$ تقریبا نرمال به ترتیب با میانگین‌های p_1 و p_2 و واریانس‌های $\frac{p_1 q_1}{n_1}$ و $\frac{p_2 q_2}{n_2}$ است. با فرض مستقل بودن دو نمونه، داریم:

$$\hat{p}_1 - \hat{p}_2 \sim N(\mu_{\hat{p}_1 - \hat{p}_2}, \sigma^2_{\hat{p}_1 - \hat{p}_2}), \quad \mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2, \quad \sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

بنابراین می‌توان نوشت:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0, 1)$$

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = P\left(-z_{\frac{\alpha}{2}} < \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left((\hat{p}_1 - \hat{p}_2) - z_{\frac{\alpha}{2}}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\frac{\alpha}{2}}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}\right) = 1 - \alpha$$

$$P\left((\hat{p}_1 - \hat{p}_2) - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}\right) = 1 - \alpha$$

فاصله اطمینان برای $p_۱ - p_۲$ در نمونه بزرگ

اگر $\hat{p}_۱$ و $\hat{p}_۲$ به ترتیب نسبت موفقیت‌ها در نمونه‌های تصادفی از اندازه‌های $n_۱$ و $n_۲$ باشند و $\hat{q}_۱ = ۱ - \hat{p}_۱$ و $\hat{q}_۲ = ۱ - \hat{q}_۲$، فاصله اطمینان تقریبی $(۱ - \alpha)۱۰۰$ برای تفاضل دو پارامتر نسبت، $p_۱ - p_۲$، عبارت است از:

$$(\hat{p}_۱ - \hat{p}_۲) - z_{\frac{\alpha}{۲}} \sqrt{\frac{\hat{p}_۱ \hat{q}_۱}{n_۱} + \frac{\hat{p}_۲ \hat{q}_۲}{n_۲}} < p_۱ - p_۲ < (\hat{p}_۱ - \hat{p}_۲) + z_{\frac{\alpha}{۲}} \sqrt{\frac{\hat{p}_۱ \hat{q}_۱}{n_۱} + \frac{\hat{p}_۲ \hat{q}_۲}{n_۲}}$$

مثال ۲۳

در آزمایش، یک نمونه تصادفی ۴۰۰ تایی از نوعی محصول را به مدت یک ساعت در دمای ۱۵۰ درجه سانتی قرار داده و مشاهده می‌شود که $\hat{p}_۱ = ۴۰\%$ آن‌ها کارایی خود را از دست می‌دهند. پس از ایجاد یک تغییر در مکانیزم تولید، یک نمونه تصادفی ۳۰۰ تایی از محصولات جدید در شرایط مشابه قرار گرفته و مشاهده می‌شود که $\hat{p}_۲ = ۳۰\%$ آن‌ها کارایی خود را از دست می‌دهند. یک فاصله اطمینان ۹۰ درصدی برای نسبت واقعی $p_۱ - p_۲$ بیابید.

راه حل:

$$(\hat{p}_۱ - \hat{p}_۲) = ۰/۴ - ۰/۳ = ۰/۱$$

$$\alpha = ۰/۱, \quad z_{\frac{\alpha}{۲}} = z_{۰/۰۵} = ۱/۶۴۵$$

$$(\hat{p}_۱ - \hat{p}_۲) - z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}_۱\hat{q}_۱}{n_۱} + \frac{\hat{p}_۲\hat{q}_۲}{n_۲}} < p_۱ - p_۲ < (\hat{p}_۱ - \hat{p}_۲) + z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}_۱\hat{q}_۱}{n_۱} + \frac{\hat{p}_۲\hat{q}_۲}{n_۲}}$$

$$۰/۱ - ۱/۶۴۵\sqrt{\frac{۰/۴(۱ - ۰/۴)}{۴۰۰} + \frac{۰/۳(۱ - ۰/۳)}{۳۰۰}} < p_۱ - p_۲$$

$$< ۰/۱ + ۱/۶۴۵\sqrt{\frac{۰/۴(۱ - ۰/۴)}{۴۰۰} + \frac{۰/۳(۱ - ۰/۳)}{۳۰۰}}$$

$$۰/۰۴ < p_۱ - p_۲ < ۰/۱۶, \quad ۰/۱ \pm ۰/۰۶$$

مثال ۲۴

فرض کنید پس از پخش مستمر یک انیمیشن از شبکه کودک، تحقیقی صورت گرفته که در آن میزان علاقه کودکان را نسبت به شخصیت اصلی داستان بررسی می‌کند. به این ترتیب دو نمونه تصادفی مستقل ۵۰ نفری از بین کودکان دختر و پسر انتخاب شده و تعداد افرادی که شخصیت مورد نظر، شخصیت محبوب آن‌هاست به ترتیب $x_۱ = ۱۱$ و $x_۲ = ۲۳$ نفر به دست آمده است. یک فاصله اطمینان ۹۵ درصدی برای مقدار واقعی $p_۲ - p_۱$ (نسبت در پسران منهای نسبت در دختران) بیابید.

راه حل:

$$(\hat{p}_۲ - \hat{p}_۱) = \frac{x_۲}{n_۲} - \frac{x_۱}{n_۱} = \frac{۲۳}{۵۰} - \frac{۲۳}{۵۰} = ۰/۴۶ - ۰/۲۲ = ۰/۲۴$$

$$\alpha = ۰/۰۵, \quad z_{\frac{\alpha}{۲}} = z_{۰/۰۲۵} = ۱/۹۶$$

$$(\hat{p}_۲ - \hat{p}_۱) - z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}_۱\hat{q}_۱}{n_۱} + \frac{\hat{p}_۲\hat{q}_۲}{n_۲}} < p_۲ - p_۱ < (\hat{p}_۲ - \hat{p}_۱) + z_{\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}_۱\hat{q}_۱}{n_۱} + \frac{\hat{p}_۲\hat{q}_۲}{n_۲}}$$

$$۰/۲۴ - ۱/۹۶\sqrt{\frac{۰/۴۶(۱ - ۰/۴۶)}{۵۰} + \frac{۰/۲۲(۱ - ۰/۲۲)}{۵۰}} < p_۱ - p_۲$$

$$< ۰/۲۴ + ۱/۹۶\sqrt{\frac{۰/۴۶(۱ - ۰/۴۶)}{۵۰} + \frac{۰/۲۲(۱ - ۰/۲۲)}{۵۰}}$$

$$۰/۰۶ < p_۱ - p_۲ < ۰/۴۲, \quad ۰/۲۴ \pm ۰/۱۸$$

فاصله اطمینان برای واریانس جامعه، σ^2

برآوردگر نقطه‌ای

می‌دانیم اگر جامعه‌ای دارای توزیع نرمال باشد و یک نمونه تصادفی n تایی $X_1, X_2, \ldots, X_n$ از آن انتخاب کنیم، واریانس نمونه که به صورت $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})$ تعریف می‌شود، یک برآوردگر ناااریب برای واریانس جامعه، σ^2، بوده و دارای توزیع خی-۲ با $n-1$ درجه آزادی است:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

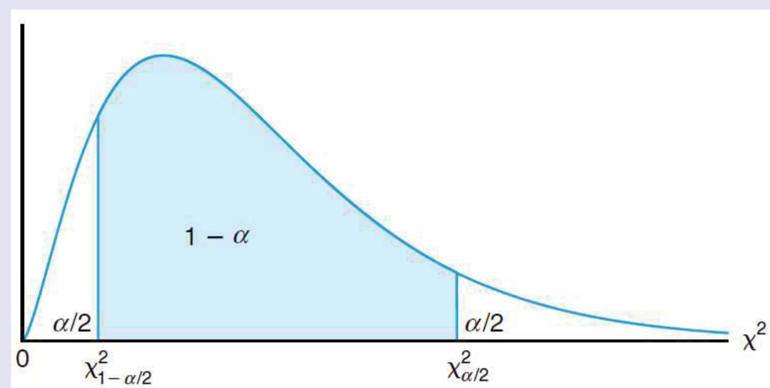
بنابراین می‌توان نوشت:

$$P(\chi^2_{1-\frac{\alpha}{2}} < \chi^2 < \chi^2_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(\chi^2_{1-\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P\left(\frac{\chi^2_{1-\frac{\alpha}{2}}}{(n-1)S^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{\frac{\alpha}{2}}}{(n-1)S^2}\right) = 1-\alpha$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}\right) = 1-\alpha$$

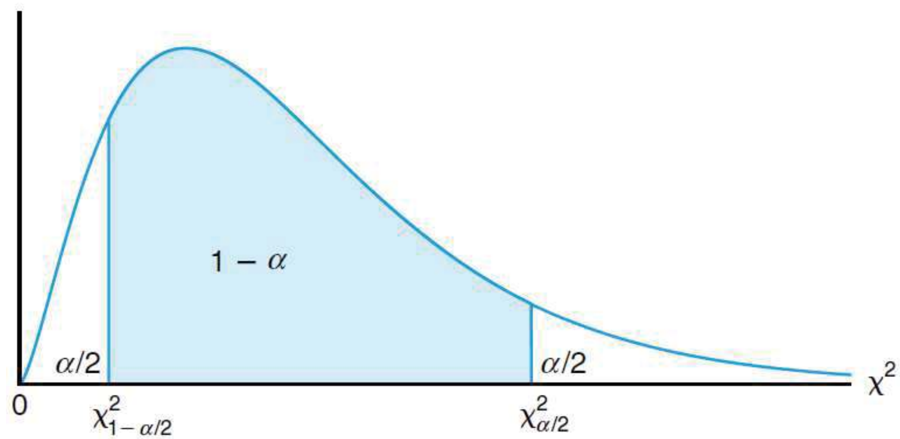


فاصله اطمینان برای واریانس جامعه، $\sigma^۲$

اگر $S^۲$ واریانس نمونه‌ای تصادفی با اندازه n از جامعه‌ای نرمال باشد، فاصله اطمینان $(۱-\alpha)۱۰۰٪$ برای $\sigma^۲$ عبارت است از:

$$\frac{(n-۱)S^۲}{\chi^۲_{\frac{\alpha}{۲}}} < \sigma^۲ < \frac{(n-۱)S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}}$$

که در آن $\chi^۲_{\frac{\alpha}{۲}}$ و $\chi^۲_{۱-\frac{\alpha}{۲}}$ مقادیر $\chi^۲$ با $\nu=(n-۱)$ درجه آزادی هستند که مساحت زیر منحنی خی-۲ در سمت راست آن‌ها، به ترتیب $\frac{\alpha}{۲}$ و $۱-\frac{\alpha}{۲}$ است.
با جذر گرفتن از کران‌های فاصله اطمینان $\sigma^۲$ یک واصله اطمینان $(۱-\alpha)۱۰۰٪$ برای انحراف معیار جامعه، σ، به دست می‌آید.



مثال ۲۵

در یک نمونه تصادفی ده تایی از جرم ذرات معلق در یک محلول بر حسب میلی‌گرم، اعداد زیر به دست آمده است:

$$۹۷, ۷۵, ۱۲۴, ۱۰۶, ۱۲۰, ۱۳۱, ۹۴, ۹۷, ۹۶, ۱۰۲$$

اگر بدانیم توزیع جرم ذرات معلق نرمال است، یک فاصله اطمینان ۹۵ درصدی برای واریانس و انحراف معیار جرم ذرات معلق در محلول بیابید.

راه حل:

$$\bar{x} = ۱۰۴/۲, \quad s^۲ = ۲۷۷, \quad \alpha = ۰/۰۵, \quad \chi^۲_{\frac{\alpha}{۲}} = ۱۹/۰۲۳, \quad \chi^۲_{۱-\frac{\alpha}{۲}} = ۲/۷$$

$$\frac{(n-۱)S^۲}{\chi^۲_{\frac{\alpha}{۲}}} < \sigma^۲ < \frac{(n-۱)S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}}$$

$$\frac{۹ \times ۲۷۷}{۱۹/۰۲۳} < \sigma^۲ < \frac{۹ \times ۲۷۷}{۲/۷}$$

$$۱۳۳ < \sigma^۲ < ۹۲۳$$

$$۱۱/۵۳ < \sigma < ۳۰/۳۸$$

مثال ۲۶

یک نمونه تصادفی ۲۰ تاییاز بلبرینگ‌های فولادی با قطری اسمی ۲ میلیمتر انتخاب کرده و قطر آن‌ها را اندازه‌گیری کرده‌ایم. اعداد به دست آمده (بر حسب میلیمتر) به شرح زیر است:

$$۲/۰۲۱/۹۴۲/۰۹۱/۹۵۱/۹۸۲/۰۰۲/۰۳۲/۰۴۲/۰۸۲/۰۷,$$

$$۱/۹۹۱/۹۶۱/۹۹۱/۹۵۱/۹۹۱/۹۹۲/۰۳۲/۰۵۲/۰۱۲/۰۳$$

با فرض نرمال بودن داده‌ها، یک فاصله اطمینان ۹۵ درصدی برای $\sigma^۲$ و σ بیابید.

راه حل:

$$\bar{x} = ۲/۰۰۹۵, \quad s^۲ = ۰/۰۴۳۵, \quad \alpha = ۰/۰۵, \quad \chi^۲_{\frac{\alpha}{۲}} = ۳۲/۸۵, \quad \chi^۲_{۱-\frac{\alpha}{۳}} = ۸/۹۱$$

$$\frac{(n-۱)S^۲}{\chi^۲_{\frac{\alpha}{۲}}} < \sigma^۲ < \frac{(n-۱)S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}}$$

$$\frac{۱۹ \times ۰/۰۴۳۵}{۳۲/۸۵} < \sigma^۲ < \frac{۱۹ \times ۰/۰۴۳۵}{۸/۹۱}$$

$$۰/۰۰۱۱ < \sigma^۲ < ۰/۰۰۴۰$$

$$۰/۰۳۳ < \sigma < ۰/۰۶۳$$