۲۶ فرض کنید در دو جمعیت میانگین مصرف روزانه پروتئین به ترتیب ۱۲۵ و ۱۰۰ گرم باشد. اگر مقادیر مصرف روزانه پروتئین در دو جمعیت دارای توزیع نرمال با انحراف معیار ۱۵ گرم باشد، احتمال اینکه نمونه‌های تصادفی و مستقل ۲۵ نفری از هر جمعیت، دارای تفاوت بین میانگین‌های نمونه کمتر از ۱۲ گرم باشد را بیابید.

۲۷ میانگین نمره هوش دانشجویان سال اول، در یک کالج بخصوص ۵۴۰ با انحراف معیار ۵۰ است. احتمال اینکه دو گروه از دانشجویان انتخابی به طور تصادفی به ترتیب شامل ۳۲ و ۵۰ دانشجو در حد متوسط نمره‌هایشان (الف) بیش از ۲۰ نمره، (ب) مقداری بین ۵ و ۱۰ نمره، فرق داشته باشند را بیابید.

۲۸ میانگین و انحراف معیار نمرات درس آمار دانشجویان یک دانشگاه به ترتیب ۷۲ و ۸ می‌باشد. دو نمونه تصادفی مستقل ۳۶ و ۴۰ تایی انتخاب می‌کنیم، احتمال اینکه تفاضل میانگین نمرات این دو نمونه بین ۲ و ۵ باشد را بیابید.

۲۹ یک نوع گلوله فلزی ساخت کارخانه‌ای به طور متوسط دارای وزن ۰/۵۰ اونس با انحراف معیار ۰/۰۲ اونس است. مطلوب است محاسبه احتمال اینکه دو بسته ۱۰۰۰تایی از این نوع گلوله‌ها بیش از ۲ اونس با یکدیگر تفاوت وزنی داشته باشند.

۳۰ مقادیر زیر را محاسبه کنید

$$F_{۰/۹۵}(۷,۱۵)\ ,F_{۰/۹۵}(۱۵,۷), F_{۰/۹۹}(۲۴,۱۹)\ , F_{۰/۹۹}(۱۹,۲۴)\ , F_{۰/۹۰}(۲۸,۱۲)$$

۳۱ از دو جمعیت نرمال با واریانسهای ۲۰ و ۳۰ به ترتیب نمونه‌های تصادفی ۸ و ۱۰ تایی انتخاب کرده‌ایم. احتمال اینکه واریانس نمونه اول بیش از ۲ برابر واریانس نمونه دوم باشد را بیابید.

۳۲ اگر نمونه‌های تصادفی مستقل با اندازه‌های $n_۱=n_۲=۸$ از جمعیتهای نرمال با واریانسهای یکسان به دست آمده باشند، احتمال اینکه یکی از این دو واریانس نمونه حداقل ۷ برابر بزرگتر از دیگری باشد را بیابید.

۳۳ اگر $S_۱^۲$ و $S_۲^۲$ به ترتیب واریانسهای دو نمونه تصادفی مستقل از دو جمعیت نرمال باشند و بدانیم که واریانس جمعیت دوم سه برابر واریانس جمعیت اول است و نمونه‌هایی به اندازه $n_۱=۸$ و $n_۲=۱۲$ انتخاب شده باشد، مطلوب است محاسبه

$$P(S_۱<\sqrt{۱/۶۳}\,S_۲)$$

فصل هفتم

نظریه برآوردیابی

۱.۷ استنباط آماری

همانگونه که در فصل ششم اشاره شد هدف از یک بررسی آماری، جمع آوری و تنظیم اطلاعات از قسمتی از جمعیت و تجزیه و تحلیل و نتیجه‌گیری از روی آنها در مورد کل جمعیت می‌باشد. این تجزیه و تحلیل و نتیجه‌گیری را استنباط آماری می‌گویند.

تعریف ۱.۷ استنباط آماری روشی است که بوسیله آن براساس نتایج حاصله از نمونه انتخابی از جمعیت در مورد کل جمعیت یا پارامترهای ناشناخته جمعیت نتیجه گیریهایی انجام می‌گیرد.

استنباط آماری دارای دو شاخه مهم زیر است

۱ـ برآورد پارامترهای مجهول جمعیت.

۲ـ آزمون فرضهای آماری در مورد پارامترهای مجهول جمعیت.

برآورد پارامتر مجهول جمعیت خود نیز به دو روش انجام می‌گیرد الف- برآورد نقطه‌ای، ب- برآورد فاصله‌ای. برای روشن شدن موضوع به مثال زیر توجه کنید.

مثال ۱.۱.۷ الف- یک بازرس می‌خواهد متوسط میزان پروتئین موجود در یک غذای کنسرو شده را تعیین کند. برای این منظور او یک نمونه n تایی از این غذای کنسرو شده را جمع آوری می‌کند و با اندازه‌گیری میزان پروتئین موجود در آنها و محاسبه متوسط این مقادیر می‌خواهد متوسط میزان پروتئین در کل کنسروها را تعیین کند. این عمل را برآوردیابی گویند.

ب- دو شرکت A و B یک نوع غذای کنسرو شده را تولید می‌کنند. شرکت A مدعی است که میزان پروتئین موجود در کنسروهای این شرکت از شرکت B بیشتر است. این ادعای شرکت A را یک فرض آماری گویند. با جمع آوری دو نمونه از این دو شرکت و مقایسه آنها ادعای شرکت A را رد یا قبول می‌کنیم که این عمل را آزمون فرض آماری گویند.

در این فصل به مبحث برآورد پارامترهای مجهول جمعیت خواهیم پرداخت و در فصل بعد مبحث آزمون فرضهای آماری را در نظر می‌گیریم.

۲.۷ برآورد پارامتر مجهول جمعیت

فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی n تایی با مقادیر مشاهده شده x_1، x_2 ... و x_n از جمعیت X با توزیع احتمال $f_\theta(x)$ باشند که این توزیع احتمال به پارامتر مجهول θ بستگی دارد. هدف از برآوردیابی یافتن کمیتی از روی مقادیر مشاهده شده x_1، x_2 ... و x_n بعنوان تقریبی از پارامتر مجهول θ می‌باشد. این عمل به دو روش انجام می‌گیرد.

الف- برآورد نقطه‌ای در این روش از روی مقدار مشاهده شده x_1، x_2 ... و x_n یک نمونه تصادفی، تنها مقدار مشاهده شده یک آماره را بعنوان تقریبی از پارامتر مجهول جمعیت ارایه می‌دهیم. اگر بر اساس نمونه تصادفی X_1، X_2، ... و X_n آماره مورد نظر ما برای تخمین پارامتر مجهول $T=T(X_1,...,X_n)$ باشد آنگاه متغیر تصادفی $T=T(X_1,...,X_n)$ را یک برآوردگر[1] پارامتر θ و عدد $t=T(x_1,...,x_n)$ را یک برآورد[2] پارامتر θ گویند. برای مثال برای به دست آوردن متوسط قد افراد یک اداره اگر یک نمونه تصادفی چهار تایی X_1، X_2، X_3 و X_4 با مقادیر مشاهده شده $x_1=172$، $x_2=168$، $x_3=165$ و $x_4=175$ را جمع آوری کنیم آنگاه آماره $\overline{X}=\frac{1}{4}\sum_{i=1}^{4}X_i$ یک برآوردگر میانگین قد افراد اداره یعنی μ است و مقدار مشاهده شده آن یعنی $\overline{x}=\frac{1}{4}\sum_{i=1}^{4}x_i=170$ یک برآورد نقطه‌ای μ است.

چون یک برآوردگر تابعی از نمونه تصادفی است بنابراین برای یک پارامتر مجهول می‌توان برآوردگرهای زیادی را معرفی کرد. حال این سوال مطرح می‌شود که در بین

ویژگیهای برآوردگر خوب

۱- نااریبی: $\hat{\theta}$ را یک برآورد کننده نااریب برای θ گوئیم هرگاه $E[\hat{\theta}]=\theta$

۲- کارا بودن: اگر دو برآوردگر نااریب $\hat{\theta}_1$ و $\hat{\theta}_2$ برای θ داشته باشیم و $Var(\hat{\theta}_1)<Var(\hat{\theta}_2)$ باشد آنگاه گوئیم $\hat{\theta}_1$ کاراتر از $\hat{\theta}_2$ است.

1- Estimator 2- Estimate

$E[\bar{x}]=E[\frac{x_1+x_2+\dots+x_n}{n}]=\dots=\mu$

$\mu_{\bar{x}}=\mu$

برآوردگرهای یک پارامتر کدامیک بهترین است. در پاسخ به این سوال به تعاریف زیر توجه کنید.

تعریف ۲.۷ برآوردگر $T=T(X_1,...,X_n)$ را یک برآوردگر نااریب برای θ گویند هر گاه

$$\boxed{E(T)=\theta} \qquad (1.7)$$

به عبارت دیگر اگر متوسط مقدار برآوردگر T به ازای نمونه‌های مختلف برابر پارامتر θ باشد، آنگاه T را برای θ نااریب گویند.

مثال ۱.۲.۷ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی از جمعیتی با میانگین μ و واریانس σ^2 باشند.

الف- تحت چه شرطی برآوردگر $T=\sum_{i=1}^{n}a_iX_i$ که در آن a_iها مقادیر ثابتی هستند، یک براوردگر نااریب برای μ است.

ب- نشان دهید که واریانس نمونه $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ یک برآوردگر نااریب برای واریانس جمعیت σ^2 است.

حل الف- می‌دانیم که $E(X_i)=\mu, i=1,2,...,n$ و بایستی $E(T)=\mu$ باشد. بنابراین طبق قوانین امید ریاضی داریم که

$$\mu=E(T)=E(\sum_{i=1}^{n}a_iX_i)=\sum_{i=1}^{n}a_iE(X_i)=\sum_{i=1}^{n}a_i\mu=\mu(\sum_{i=1}^{n}a_i)$$

بنابراین با شرط $\sum_{i=1}^{n}a_i=1$، برآوردگر T برای μ نااریب است. اگر قرار دهیم $a_i=\frac{1}{n}, i=1,...,n$ آنگاه $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ برای μ نااریب است.

ب- به تمرین ۲ مراجعه کنید.

تعریف ۳.۷ اگر T_1 و T_2 دو برآوردگر نااریب θ باشند و $Var(T_1)<Var(T_2)$ آنگاه برآوردگر T_1 را کارآتر از برآوردگر T_2 گویند. بنابراین در بین برآوردگرهای نااریب θ آن برآوردگری که دارای کمترین واریانس باشد را کارآترین برآوردگر گویند.

مثال ۲.۲.۷ در مثال ۱.۲.۷ (الف) نشان دهید که در بین تمام برآوردگرهای $T=\sum_{i=1}^{n}a_iX_i$ با شرط $\sum_{i=1}^{n}a_i=1$، برآوردگر $\overline{X}$ دارای کمترین واریانس است.

$$\sigma^2 \sum_{i=1} \left((a_i-\frac{1}{n})^2+2(a_i-\frac{1}{n})\frac{1}{n}+\frac{1}{n^2}\right)=\sigma^2\left(\sum_{i=1}^{n}(a_i-\frac{1}{n})^2+\sum_{i=1}^{n}2(a_i-\frac{1}{n})\frac{1}{n}+\sum_{i=1}^{n}\frac{1}{n^2}\right)$$

حل طبق خواص واریانس برای متغیرهای تصادفی مستقل داریم که

$$Var(T)=Var(\sum_{i=1}^{n} a_iX_i)=\sum_{i=1}^{n} a_i^2 Var(X_i)=\sigma^2\sum_{i=1}^{n} a_i^2$$

$$=\sigma^2\sum_{i=1}^{n}(a_i-\frac{1}{n}+\frac{1}{n})^2=\sigma^2\left[\sum_{i=1}^{n}(a_i-\frac{1}{n})^2+\frac{2}{n}\underbrace{\sum_{i=1}^{n}(a_i-\frac{1}{n})}_{\text{بنابر شرط داده شده برابر صفر است}}+\frac{1}{n}\right]$$

$$=\sigma^2\left[\sum_{i=1}^{n}(a_i-\frac{1}{n})^2+\frac{1}{n}\right] \qquad 2a_i-\frac{1}{n}=0 \Rightarrow a_i=\frac{1}{n}$$

مقدار درون کروشه موقعی کمترین مقدار خود را به دست می‌آورد که $\sum_{i=1}^{n}(a_i-\frac{1}{n})^2=0$ و یا $a_i=\frac{1}{n}, i=1,\dots,n$، یعنی $\overline{X}$ در بین برآوردگرهای فوق دارای کمترین واریانس است.

ب- برآورد فاصله‌ای در برآورد نقطه‌ای موقعی برآوردگر $T=T(X_1,\dots,X_n)$ برای پارامتر θ یک برآوردگر خوب است که مقدار مشاهده شده آن یعنی برآورد $t=T(x_1,\dots,x_n)$ به پارامتر θ نزدیک باشد. اما چون با تغییر نمونه مقدار برآورد تغییر می‌یابد پس برآورد نقطه‌ای دارای خطای زیاد می‌باشد. بنابراین به جای برآورد نقطه‌ای می‌توان از برآورد فاصله که دارای خطای کمتری است استفاده کرد.

در روش برآورد فاصله‌ای، یک فاصله (L,U) از اعداد حقیقی را به عنوان تقریبی از پارامتر مجهول θ ارایه می‌دهیم که این فاصله با یک احتمال زیاد پارامتر θ را در برداشته باشد. این فاصله را فاصله اطمینان گویند و اگر احتمال قرار گرفتن پارامتر θ در این فاصله $1-\alpha$ باشد، آن را یک فاصله اطمینان $(1-\alpha)100\%$ گویند. L را حد پائین فاصله و U را حد بالای فاصله و $1-\alpha$ را ضریب اطمینان فاصله گویند. چگونگی به دست آوردن فاصله اطمینان در مثال زیر تشریح شده است.

مثال ۳.۲.۷ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی از جمعیت نرمال با میانگین μ و واریانس σ^2 باشند، که در آن σ^2 مقداری معلوم (مثلاً $\sigma^2=4$) ولی μ پارامتر مجهول است. یک برآورد نقطه‌ای و یک فاصله اطمینان $(1-\alpha)100\%$ برای μ پیدا کنید.

حل پارامتر μ میانگین کل جمعیت است بنابراین طبق مثالهای ۱.۲.۷ و ۲.۲.۷ معقول به نظر می‌رسد که آن را با میانگین نمونه $\bar{x}$ برآورد کنیم بنابراین برآوردگر نقطه‌ای μ عبارت است از

$$\hat{\mu}=\overline{X}$$

برای به دست آوردن یک فاصله اطمینان $(1-\alpha)100\%$ برای μ بایستی فاصله (L,U) را به گونه‌ای تعیین کنیم که

$$\boxed{P(L<\mu<U)=1-\alpha} \qquad (۲.۷)$$

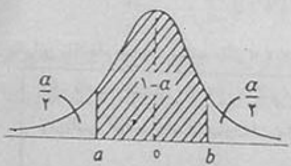
برای این منظور ابتدا تابعی را تعیین می‌کنیم که بستگی به نمونه تصادفی و پارامترهای معلوم و مجهول جمعیت داشته باشد به طوری که توزیع این تابع به پارامترهای مجهول بستگی نداشته باشد. این تابع را **تابع محور** می‌نامند. با استفاده از مطالب بخش ۲.۶ تابع محور مناسب برای این مثال عبارت است از

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

حال برای به دست آوردن فاصله اطمینان برای μ اعداد $a<b$ را به گونه‌ای انتخاب می‌کنیم که

$$P(a<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<b)=1-\alpha$$

با توجه به شکل ۱.۷ نقطه b نقطه‌ای روی محور افقی نمودار تابع چگالی نرمال استاندارد است که مساحت زیر منحنی قبل از این نقطه $1-\frac{\alpha}{2}$ می‌باشد. این نقطه را با $z_{1-\frac{\alpha}{2}}$ نمایش می‌دهیم که از جدول (III) قابل محاسبه است. نقطه a قرینه نقطه b می‌باشد. با قرار دادن این مقادیر a و b در رابطه فوق داریم که



شکل ۱.۷

$$P(-z_{1-\frac{\alpha}{2}}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_{1-\frac{\alpha}{2}})=1-\alpha$$

با حل نامساوی درون پرانتز نسبت به پارامتر μ داریم که

$$P(\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})=1-\alpha$$

با مقایسه این رابطه با رابطه (۲.۷) نتیجه می‌شود که یک فاصله اطمینان $(1-\alpha)100\%$ برای میانگین μ عبارت است از

$$\mu\in(\bar{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$

در بخشهای بعد با استفاده از مطالب و مثالهای این بخش پارامترهای مختلف جمعیت را برآورد خواهیم کرد.

۳.۷ برآورد میانگین جمعیت

فرض کنید که از یک جمعیت X با میانگین μ و واریانس σ^2 یک نمونه تصادفی به اندازه n انتخاب کرده باشیم و بخواهیم میانگین جمعیت یعنی μ را برآورد کنیم. طبق مطالب بخش قبل بهترین برآوردگر نقطه‌ای μ عبارت است از $\hat{\mu}=\bar{X}$ برای به دست آوردن یک فاصله اطمینان برای μ دو حالت زیر را در نظر می‌گیریم.

الف- واریانس جمعیت معلوم است در این حالت اگر جمعیت نرمال باشد آنگاه یک تابع محور مناسب عبارت است از

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1) \qquad (۳.۷)$$

و بنابراین با انجام عملیات مشابه مثال ۳.۲.۷ داریم که

> یک فاصله اطمینان $100(1-\alpha)$٪ برای میانگین جمعیت نرمال μ موقعی که واریانس σ^2 معلوم است عبارت است از
>
> $$\mu\in(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$
>
> که در آن $\bar{x}$ میانگین نمونه تصادفی و $z_{1-\frac{\alpha}{2}}$ مقدار متغیر نرمال استاندارد است که سطح زیر منحنی در سمت چپ آن $1-\frac{\alpha}{2}$ باشد.

توجه کنید که اگر جمعیت نرمال نباشد ولی $n\geq 30$ باشد آنگاه رابطه (۳.۷) هنوز به طور تقریبی برقرار است و در نتیجه فاصله اطمینان فوق یک فاصله اطمینان $100(1-\alpha)$٪ برای μ است.

ب- واریانس جمعیت نامعلوم است در این حالت اگر جمعیت نرمال باشد آنگاه طبق مطالب بخش ۴.۶ یک تابع محور مناسب برای ساختن فاصله اطمینان برای μ عبارت است از

$$T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t_{(n-1)} \qquad (۴.۷)$$

حال با استفاده از این تابع محور و انجام عملیات مشابه با مثال ۳.۲.۷ و جایگزینی s به جای σ و $t_{1-\frac{\alpha}{2}}(n-1)$ به جای $z_{1-\frac{\alpha}{2}}$ داریم که

> یک فاصله اطمینان $100(1-\alpha)$٪ برای میانگین جمعیت نرمال μ موقعی که واریانس σ^2 نامعلوم است عبارت است از
>
> $$\mu\in(\bar{x}-t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}\ ,\ \bar{x}+t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}})$$
>
> که در آن $\bar{x}$ و s به ترتیب میانگین و انحراف استاندارد نمونه تصادفی n تایی و $t_{1-\frac{\alpha}{2}}(n-1)$ مقدار متغیر $t_{(n-1)}$ است که سطح زیر منحنی در سمت چپ آن $1-\frac{\alpha}{2}$ باشد.

توجه کنید که برای محاسبه s^2 می‌توان از فرمول زیر استفاده کرد.

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}(\sum_{i=1}^{n}x_i)^2\right] \qquad (۵.۷)$$

همچنین توجه کنید که اگر $n>30$ باشد آنگاه $t_{1-\frac{\alpha}{2}}(n-1)=z_{1-\frac{\alpha}{2}}$ و بنابراین در حالتی که σ نامعلوم و $n>30$ باشد می‌توان از فاصله اطمینان حالت (الف) با قرار دادن s به جای σ استفاده کرد.

مثال ۱.۳.۷ یک تولیدکننده لامپهای روشنایی، لامپهایی را تولید می‌کند که انحراف معیار طول عمر آنها ۴۰ ساعت است. اگر یک نمونه تصادفی ۳۶ تایی دارای حد متوسط عمر ۸۷۰ ساعت باشند، یک برآورد نقطه‌ای و یک فاصله اطمینان ۹۵ درصدی برای حد متوسط عمر تمام لامپهای تولیدی این کارخانه را به دست آورید.

حل در این مثال $\sigma=40$، $n=36$، $\bar{x}=870$ و $1-\alpha=0/95$ می‌باشد بنابراین برآورد نقطه‌ای عبارت است از $\hat{\mu}=\bar{x}=870$. برای به دست آوردن فاصله اطمینان داریم که

$$1-\alpha=0/95 \Rightarrow 1-\frac{\alpha}{2}=0/975$$

و با استفاده از جدول (III) $z_{1-\frac{\alpha}{2}}=z_{0/975}=1/96$، بنابراین

$$\mu \in \left(\bar{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$$=\left(۸۷۰ - ۱/۹۶\frac{۴۰}{\sqrt{۳۶}}\ ,\ ۸۷۰ + ۱/۹۶\frac{۴۰}{\sqrt{۳۶}}\right)=(۸۵۶/۹۳\ ,\ ۸۸۳/۰۷)$$

یعنی ۹۵ درصد اطمینان داریم که حد متوسط عمر تمام لامپهای تولیدی این کارخانه در فاصله فوق قرار دارد.

مثال ۲.۳.۷ اگر طول قد کارمندان یک اداره دارای توزیع نرمال باشد، یک فاصله اطمینان ۹۵ درصدی برای میانگین طول قد کارمندان این اداره پیدا کنید در حالیکه یک نمونه ۵ تایی از بین کارمندان انتخاب شده باشد و مقادیر ۱۶۰، ۱۷۰، ۱۶۵، ۱۷۵ و ۱۸۰ به دست آمده باشد.

حل در این مثال σ نامعلوم است و $n=۵$، $\sum_{i=1}^{۵} x_i = ۸۵۰$، $\sum_{i=1}^{۵} x_i^2 = ۱۴۴۷۵۰$ و $1-\alpha = ۰/۹۵$ می‌باشد، بنابراین

$$\bar{x} = \frac{۱}{۵}\sum_{i=1}^{۵} x_i = \frac{۸۵۰}{۵} = ۱۷۰$$

$$s^2 = \frac{۱}{۵-۱}\left[\sum_{i=1}^{۵} x_i^2 - \frac{۱}{۵}\left(\sum_{i=1}^{۵} x_i\right)^2\right] = \frac{۱}{۴}\left[۱۴۴۷۵۰ - \frac{(۸۵۰)^2}{۵}\right] = ۶۲/۵$$

و با استفاده از جدول (V) داریم که $t_{1-\frac{\alpha}{2}}(n-1) = t_{۰/۹۷۵}(۴) = ۲/۷۸$ در نتیجه

$$\mu \in \left(\bar{x} - t_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\ ,\ \bar{x} + t_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right)$$

$$=\left(۱۷۰ - ۲/۷۸\sqrt{\frac{۶۲/۵}{۵}}\ ,\ ۱۷۰ + ۲/۷۸\sqrt{\frac{۶۲/۵}{۵}}\right)=(۱۶۰/۱۷\ ,\ ۱۷۹/۸۳)$$

یعنی ۹۵ درصد اطمینان داریم که میانگین طول قد کارمندان این اداره در فاصله فوق قرار دارد.

توجه کنید به کمک بعضی از ماشینهای محاسبه می‌توان به راحتی مقادیر $\bar{x}$ و s را به طور مستقیم محاسبه نمود.

خطای برآورد میانگین چون اغلب مقدار برآورد نقطه‌ای $\bar{x}$ دقیقاً مساوی μ نیست بنابراین برآورد نقطه‌ای دارای خطا است. با استفاده از حدود فاصله اطمینان می‌توان میزان این خطا یعنی $|\bar{x} - \mu|$ را در دو حالت زیر تعیین کرد.

الف - اگر واریانس σ^2 معلوم باشد آنگاه

$$-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow |\bar{x} - \mu| < z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

ب - اگر واریانس σ^2 نامعلوم باشد آنگاه

$$-t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} < \bar{x} - \mu < t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} \Rightarrow |\bar{x} - \mu| < t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}$$

بنابراین اگر $\bar{x}$ را به عنوان برآورد نقطه‌ای μ به کار ببریم آنگاه $۱۰۰(1-\alpha)$٪ مطمئن هستیم که

الف - خطای برآورد کمتر از $z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ است اگر σ معلوم باشد.

ب - خطای برآورد کمتر از $t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}$ است اگر σ معلوم نباشد.

برای مثال در مثال ۱.۳.۷ چون $z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} = ۱/۹۶\frac{۴۰}{\sqrt{۳۶}} = ۱۳/۰۷$

بنابراین ۹۵ درصد مطمئن هستیم که خطای برآورد میانگین از ۱۳/۰۷ کمتر است.

تعیین اندازه نمونه در یک بررسی آماری یکی از مهمترین مراحل، تعیین اندازه نمونه n قبل از عمل نمونه گیری می‌باشد. اگر یک حداکثر مقدار خطای e برای برآورد میانگین μ برای نمونه گیر قابل تحمل باشد آنگاه بوسیله خطای برآورد می‌توان اندازه نمونه n را در دو حالت زیر تعیین کرد.

الف - اگر واریانس σ^2 معلوم باشد آنگاه

$$z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le e \Rightarrow n \ge \left(z_{1-\frac{\alpha}{2}}\sigma/e\right)^2$$

ب - اگر واریانس σ^2 نامعلوم باشد آنگاه

$$t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} \le e \Rightarrow n \ge \left(t_{1-\frac{\alpha}{2}}(n-1)s/e\right)^2$$

(که در این حالت ابتدا به وسیله یک نمونه مقدماتی مقداری برای s به دست می آوریم و سپس اندازه نمونه واقعی n را از فرمول فوق محاسبه می‌کنیم. در هر دو حالت n را برابر کوچکترین عدد صحیح که در نامساویهای فوق صدق کند انتخاب می‌کنیم. بنابراین)

اگر $\bar{x}$ را به عنوان برآورد نقطه‌ای μ به کار ببریم آنگاه $(۱-\alpha)۱۰۰\%$ مطمئن هستیم که خطای برآورد از مقدار مشخص e کمتر است موقعی که اندازه نمونه از رابطه زیر محاسبه گردد

الف - اگر واریانس $\sigma^۲$ معلوم باشد $n \ge (z_{۱-\frac{\alpha}{۲}}\sigma/e)^۲$

ب - اگر واریانس $\sigma^۲$ نامعلوم باشد $n \ge (t_{۱-\frac{\alpha}{۲}}(n-۱)s/e)^۲$

مثال ۳.۳.۷ در مثال ۱.۳.۷ اگر بخواهیم با اطمینان ۹۵ درصد خطای برآورد حد متوسط طول عمر لامپها از ۱۰ ساعت کمتر باشد، اندازه نمونه را تعیین کنید.

حل $n \ge (z_{۱-\frac{\alpha}{۲}}\sigma/e)^۲ = (\frac{۱/۹۶\times ۴۰}{۱۰})^۲ = ۶۱/۴۷$

بنابراین اندازه نمونه بایستی n=۶۲ باشد.

مثال ۴.۳.۷ در مثال ۲.۳.۷ اگر بخواهیم با اطمینان ۹۵ درصد خطای برآورد میانگین طول قد کارمندان اداره از ۵ سانتیمتر کمتر باشد، اندازه نمونه را تعیین کنید.

حل اگر داده‌های مثال ۲.۳.۷ را به عنوان یک نمونه مقدماتی در نظر بگیریم آنگاه $s^۲=۶۲/۵$ و در نتیجه

e=۵

$n \ge (t_{۱-\frac{\alpha}{۲}}(n-۱)s/e)^۲ = (\frac{۲/۷۸}{۵})^۲ \times ۶۲/۵ = ۱۹/۳۲۱$

بنابراین اندازه نمونه بایستی n=۲۰ باشد.

۴.۷ برآورد واریانس جمعیت

فرض کنید که از یک جمعیت نرمال با میانگین μ و واریانس $\sigma^۲$ یک نمونه تصادفی به اندازه n انتخاب کرده باشیم و بخواهیم واریانس جمعیت یعنی $\sigma^۲$ را برآورد کنیم. بهترین برآوردگر نقطه‌ای $\sigma^۲$ عبارت است از واریانس نمونه $S^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})^۲$ یعنی $\hat{\sigma}^۲=S^۲$

برای به دست آوردن یک فاصله اطمینان برای $\sigma^۲$ طبق مطالب بخش ۳.۶ تابع محور مناسب عبارت است از

$$Y=\frac{(n-۱)S^۲}{\sigma^۲}\sim \chi^۲_{(n-۱)} \qquad (۶.۷)$$

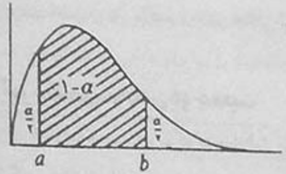
حال با استفاده از این تابع محور، اعداد a و b را چنان تعیین می‌کنیم که

$$P(a<\frac{(n-۱)S^۲}{\sigma^۲}<b)=۱-\alpha$$

با توجه به شکل ۲.۷ مقادیر a و b عبارتند از

$$a=\chi^۲_{\frac{\alpha}{۲}}(n-۱) \quad , b=\chi^۲_{۱-\frac{\alpha}{۲}}(n-۱)$$

با قرار دادن این مقادیر در رابطه فوق و انجام عملیات مشابه مثال ۳.۲.۷ داریم که



شکل ۲.۷

یک فاصله اطمینان $(۱-\alpha)۱۰۰\%$ برای واریانس جمعیت نرمال $\sigma^۲$ عبارت است از

$$\sigma^۲\in\left(\frac{(n-۱)s^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}(n-۱)} , \frac{(n-۱)s^۲}{\chi^۲_{\frac{\alpha}{۲}}(n-۱)}\right)$$

که در آن $s^۲$ واریانس یک نمونه تصادفی n تایی است.

مثال ۱.۴.۷ طول یک لوله ساختمانی دارای توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ است. یک نمونه تصادفی ۲۵ تایی از لوله‌ها جمع آوری شده است و مقادیر $\sum_{i=۱}^{۲۵}x_i=۲۵۲/۷$ و $\sum_{i=۱}^{۲۵}x_i^۲=۲۵۷۹/۷$ حاصل شده است. یک برآورد نقطه‌ای و یک فاصله اطمینان ۹۰ درصدی برای واریانس واقعی لوله‌ها به دست آورید.

حل با توجه به مقادیر داده شده و رابطه (۵.۷) داریم که

$$s^۲=\frac{۱}{۲۵-۱}\left[۲۵۷۹/۷-\frac{(۲۵۲/۷)^۲}{۲۵}\right]=۱/۰۶$$

بنابراین برآورد نقطه‌ای $\sigma^۲$ عبارت است از $\hat{\sigma}^۲=s^۲=۱/۰۶$

همچنین با استفاده از جدول (IV) داریم که

$$۱-\alpha=۰/۹۰ \Rightarrow \begin{cases}\frac{\alpha}{۲}=۰/۰۵\\ ۱-\frac{\alpha}{۲}=۰/۹۵\end{cases} \Rightarrow \begin{cases}\chi^۲_{\frac{\alpha}{۲}}(n-۱)=\chi^۲_{۰/۰۵}(۲۴)=۱۳/۸\\ \chi^۲_{۱-\frac{\alpha}{۲}}(n-۱)=\chi^۲_{۰/۹۵}(۲۴)=۳۶/۴\end{cases}$$

بنابراین

$$\sigma^2\in\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)=\left(\frac{24\times 1.06}{36.4},\frac{24\times 1.06}{13.8}\right)=(0.699,\ 1.843)$$

یعنی ۹۰ درصد اطمینان داریم که واریانس واقعی لوله‌ها در فاصله فوق قرار دارد.

۵.۷ برآورد تفاضل میانگین دو جمعیت

فرض کنید دو جمعیت داریم که جمعیت اول دارای میانگین μ_1 و واریانس σ_1^2 و جمعیت دوم دارای میانگین μ_2 و واریانس σ_2^2 باشد. یک نمونه تصادفی n_1 تایی از جمعیت اول انتخاب کرده و میانگین نمونه آن را $\bar{X}_1$ و واریانس نمونه آن را S_1^2 می‌نامیم و یک نمونه تصادفی n_2 تایی از جمعیت دوم انتخاب کرده و میانگین نمونه آن را $\bar{X}_2$ و واریانس نمونه آن را S_2^2 می‌نامیم. همچنین فرض کنید که این دو نمونه تصادفی از یکدیگر مستقل باشند. می‌خواهیم اختلاف میانگین دو جمعیت یعنی $\mu_1-\mu_2$ را برآورد کنیم. بهترین برآوردگر نقطه‌ای $\mu_1-\mu_2$ عبارت از اختلاف میانگینهای دو نمونه می‌باشد یعنی

$$\widehat{\mu_1-\mu_2}=\bar{X}_1-\bar{X}_2$$

برای به دست آوردن فاصله اطمینان برای $\mu_1-\mu_2$ دو حالت زیر را در نظر می‌گیریم.

الف- واریانس دو جمعیت معلوم باشند در این حالت σ_1^2 و σ_2^2 مقادیر معلومی هستند. اگر دو جمعیت نرمال باشند، با توجه به مطالب بخش ۵.۶ و رابطه (۵.۶) تابع محور مناسب عبارت است از

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1) \tag{7.7}$$

حال با استفاده از این تابع محور و انجام عملیات مشابه مثال ۳.۲.۷ به دست می‌آوریم که

یک فاصله اطمینان $100(1-\alpha)\%$ برای تفاضل میانگین دو جمعیت نرمال که واریانسهای آنها معلوم است عبارت است از

$$\mu_1-\mu_2\in\left(\bar{x}_1-\bar{x}_2-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ \bar{x}_1-\bar{x}_2+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

که در آن $\bar{x}_1$ و $\bar{x}_2$ به ترتیب میانگینهای نمونه‌های تصادفی n_1 تایی و n_2 تایی از دو جمعیت می‌باشند.

توجه کنید که اگر دو جمعیت نرمال نباشند اما $n_1\ge 30$ و $n_2\ge 30$ باشند آنگاه طبق مطالب بخش ۵.۶، رابطه (۷.۷) هنوز به طور تقریبی برقرار است و بنابراین فاصله اطمینان فوق نیز در این حالت یک فاصله اطمینان $100(1-\alpha)\%$ برای $\mu_1-\mu_2$ است.

ب- واریانس دو جمعیت نامعلوم اما مساوی باشند در این حالت $\sigma_1^2=\sigma_2^2=\sigma^2$ می‌باشد و بایستی ابتدا σ^2 یعنی واریانس مشترک دو جمعیت را با واریانس مشترک دو نمونه به صورت زیر برآورد کنیم.

$$\hat{\sigma}^2=S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \tag{8.7}$$

اگر دو جمعیت نرمال باشند آنگاه طبق مطالب بخش ۵.۶ و رابطه (۷.۶) تابع محور مناسب عبارت است از

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{(n_1+n_2-2)}$$

حال با استفاده از این تابع محور و انجام عملیات مشابه مثال ۳.۲.۷ به دست می‌آوریم که

یک فاصله اطمینان $100(1-\alpha)\%$ برای تفاضل میانگین دو جمعیت نرمال که واریانسهای آنها نامعلوم اما مساوی هستند عبارت است از

$$\mu_1-\mu_2\in$$

$$\left(\bar{x}_1-\bar{x}_2-t_{1-\frac{\alpha}{2}}(n_1+n_2-2)s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}_1-\bar{x}_2+t_{1-\frac{\alpha}{2}}(n_1+n_2-2)s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

که در آن s_p^2 واریانس مشترک دو نمونه تصادفی n_1 و n_2 تایی از دو جمعیت می‌باشد.

توجه کنید که اگر $n_1\ge 30$ و $n_2\ge 30$ و جمعیتها غیر نرمال با واریانسهای نامعلوم باشند آنگاه می‌توان از فاصله اطمینان قسمت (الف) با قرار دادن s_1 و s_2 به جای σ_1 و σ_2 استفاده کرد.

مثال ۱.۵.۷ دو شرکت A و B لامپهای روشنایی تولید می‌کنند که به ترتیب دارای انحراف معیارهای ۲۷ و ۳۱ ساعت است. یک نمونه تصادفی ۴۰ تایی از لامپهای تولیدی هر شرکت انتخاب می‌کنیم و به ترتیب متوسط طول عمر لامپها برای این دو نمونه را ۶۴۹ و ۶۳۵ به دست می‌آوریم. با

$1-\alpha = 0.96 \to \alpha = 0.04 \to \frac{\alpha}{2} = 0.02$

$Z_{1-\frac{\alpha}{2}} = Z_{1-0.02} = Z_{0.98} = 2.055$

ساختن یک فاصله اطمینان ۹۶ درصدی برای اختلاف متوسط طول عمر لامپهای دو شرکت، چه نتیجه‌ای به دست می‌آورید.

حل در این مثال $\sigma_1=27$ ، $\sigma_2=31$ ، $n_1=n_2=40$ ، $\bar{x}_1=649$ و $\bar{x}_2=635$ و $z_{1-\frac{\alpha}{2}}=z_{0.98}=2.054$. بنابراین طبق حالت (الف) داریم که

$$\mu_1-\mu_2\in$$

$$\left(649-635-2.054\sqrt{\frac{(27)^2}{40}+\frac{(31)^2}{40}},\ 649-635+2.054\sqrt{\frac{(27)^2}{40}+\frac{(31)^2}{40}}\right)$$

$$=(14-13.351,\ 14+13.351)=(0.649,\ 27.351)$$

چون تمامی فاصله مثبت است پس ۹۶ درصد اطمینان داریم که متوسط طول عمر لامپهای شرکت A از شرکت B بیشتر است.

مثال ۲.۵.۷ دو آزمایشگاه ۱ و ۲ به طور مستقل برای اندازه‌گیری چربی موجود در شیرهای پاستوریزه اقدام می‌نمایند. هر یک تعدادی نمونه انتخاب کرده و نتایج در جدول زیر ثبت شده است. با فرض نرمال بودن دو جمعیت و مساوی بودن واریانسها، یک فاصله اطمینان ۹۵ درصدی برای اختلاف میانگین چربی موجود در شیرهای دو کارخانه به دست آورید. چه نتیجه‌ای می‌گیرید؟

آزمایشگاه ۱	۳	۵	۷	۳	۸	۶	۸	۹	۴	۷
آزمایشگاه ۲	۹	۸	۸	۴	۷	۶	۸	۶		

حل از جدول فوق مقادیر زیر را به دست می‌آوریم

$$n_1=10,\ \sum x_{1i}=60,\ \sum x_{1i}^2=402,\ \bar{x}_1=6,\ s_1^2=4.67$$

$$n_2=8,\ \sum x_{2i}=56,\ \sum x_{2i}^2=410,\ \bar{x}_2=7,\ s_2^2=2.57$$

بنابراین

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}=\frac{9(4.67)+7(2.57)}{16}=3.75$$

در نتیجه $s_p=1.937$ و همچنین

$$t_{1-\frac{\alpha}{2}}(n_1+n_2-2)=t_{0.975}(16)=2.12$$

بنابراین با استفاده از حالت (ب) داریم که

$$\mu_1-\mu_2\in\left(6-7-2.12(1.937)\sqrt{\frac{1}{10}+\frac{1}{8}},\ 6-7+2.12(1.937)\sqrt{\frac{1}{10}+\frac{1}{8}}\right)$$

$$=(-1-1.948,\ -1+1.948)=(-2.948,\ 0.948)$$

چون فاصله شامل مقادیر منفی و مثبت است پس ۹۵ درصد اطمینان داریم که میانگین چربی موجود در شیرهای دو کارخانه با یکدیگر مساوی است.

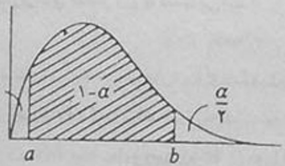
۶.۷ برآورد نسبت واریانس دو جمعیت

دو جمعیت بخش قبل و نمونه‌های تصادفی به دست آمده از آنها را در نظر بگیرید. در این بخش می‌خواهیم نسبت واریانس دو جمعیت یعنی $\frac{\sigma_2^2}{\sigma_1^2}$ را برآورد کنیم. بهترین برآوردگر نقطه‌ای $\frac{\sigma_2^2}{\sigma_1^2}$ عبارت از نسبت واریانسهای دو نمونه می‌باشد یعنی

$$\widehat{\frac{\sigma_2^2}{\sigma_1^2}}=\frac{S_2^2}{S_1^2}$$

برای به دست آوردن یک فاصله اطمینان برای $\frac{\sigma_2^2}{\sigma_1^2}$ اگر دو جمعیت نرمال باشند آنگاه طبق مطالب بخش ۶.۶ تابع محور مناسب عبارت است از

$$F=\frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1) \qquad (9.7)$$



شکل ۳.۷

حال با استفاده از این تابع محور و توجه به شکل ۳.۷، با انجام عملیات مشابه مثال ۳.۲.۷ یک فاصله اطمینان $100(1-\alpha)\%$ برای $\frac{\sigma_2^2}{\sigma_1^2}$ به صورت زیر به دست می‌آید.

$$\frac{\sigma_2^2}{\sigma_1^2}\in\left(\frac{s_2^2}{s_1^2}F_{\frac{\alpha}{2}}(n_1-1,n_2-1),\ \frac{s_2^2}{s_1^2}F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right)$$

با توجه به اینکه $F_{\frac{\alpha}{2}}(n_1-1,n_2-1)=\frac{1}{F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)}$ بنابراین به دست می‌آوریم که

یک فاصله اطمینان $(1-\alpha)100$٪ برای نسبت واریانس دو جمعیت نرمال عبارت است از

$$\frac{\sigma_2^2}{\sigma_1^2}\in\left(\frac{s_2^2}{s_1^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)},\ \frac{s_2^2}{s_1^2}F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right)$$

که در آن s_1^2 و s_2^2 به ترتیب واریانسهای نمونه‌های n_1 و n_2 تایی از دو جمعیت می‌باشند.

مثال ۱.۶.۷ از یک درس یک کلاس صبح و یک کلاس بعد از ظهر تشکیل شده است. در آخر ترم از کلاس صبح ۸ نفر و از کلاس بعدازظهر ۹ نفر به تصادف انتخاب کرده و از آنها امتحان بعمل آورده‌ایم و نتایج در جدول زیر یادداشت گردیده‌اند.

نمرات کلاس صبح	۹	۷	۸	۱۰	۱۲	۱۵	۷	۱۲	
نمرات کلاس بعدازظهر	۳	۷	۹	۲	۱۸	۱۶	۵	۱۱	۱۰

فرض کنید نمرات دو کلاس از یکدیگر مستقل بوده و از توزیع نرمال پیروی کنند. یک فاصله اطمینان ۹۰ درصدی برای نسبت واریانسها و نسبت انحراف معیارهای نمرات دو کلاس به دست آورید. چه نتیجه‌ای می‌گیرید؟

حل از جدول فوق مقادیر زیر به دست می‌آیند

$$n_1=8\ ,\ \sum x_{1i}=80\ ,\ \sum x_{1i}^2=856\ ,\ s_1^2=8$$
$$n_2=9\ ,\ \sum x_{2i}=81\ ,\ \sum x_{2i}^2=969\ ,\ s_2^2=30$$

همچنین از جدول (VI) داریم که

$$1-\alpha=0/90 \Rightarrow \begin{cases} F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)=F_{0/95}(7,8)=3/50 \\ F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)=F_{0/95}(8,7)=3/73 \end{cases}$$

بنابراین فاصله اطمینان ۹۰ درصدی برای $\frac{\sigma_2^2}{\sigma_1^2}$ عبارت است از

$$\frac{\sigma_2^2}{\sigma_1^2}\in\left(\frac{30}{8}\times\frac{1}{3/73},\ \frac{30}{8}\times 3/50\right)=(1/005,\ 13/125)$$

حال اگر از مقادیر این فاصله جذر بگیریم فاصله اطمینان ۹۰ درصدی برای نسبت انحراف معیارها

یعنی $\frac{\sigma_2}{\sigma_1}$ به صورت زیر به دست می‌آید

$$\frac{\sigma_2}{\sigma_1}\in(1/003,\ 3/623)$$

چون این فاصله تماماً از یک بزرگتر است پس ۹۰ درصد اطمینان داریم که $\sigma_2^2>\sigma_1^2$ می‌باشد.

۷.۷ تمرینات

۱ فرض کنید X_1 و X_2 دو برآوردگر نااریب و مستقل پارامتر θ با واریانسهای ۲ و ۳ باشند. اگر $T=a_1X_1+a_2X_2$ آنگاه ضرایب a_1 و a_2 را به گونه‌ای پیدا کنید که T یک برآوردگر نااریب با کمترین واریانس برای θ باشد. با مقایسه واریانس برآوردگرهای X_1 و X_2 و T نتیجه بگیرید که کدامیک از این ۳ برآوردگر نااریب، بهتر می‌باشد.

۲ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی n تایی از جمعیت X با میانگین μ و واریانس σ^2 باشند. اگر $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ و $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ به ترتیب میانگین و واریانس این نمونه باشند، مطلوب است

الف - نشان دهید که $E(\overline{X})=\mu$ و $Var(\overline{X})=\frac{\sigma^2}{n}$

ب - نشان دهید که $S^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\right]$

ج - با استفاده از قسمت (ب) نشان دهید که $E(S^2)=\sigma^2$

۳ نمونه‌ای به اندازه ۲۵ لامپ روشنایی از یک دسته بزرگ از لامپهای ۴۰واتی گرفته شده است و میانگین عمر لامپهای نمونه ۱۴۱۰ ساعت است. با فرض اینکه عمر لامپها دارای توزیع نرمال با انحراف معیار ۲۰۰ ساعت باشد، یک فاصله اطمینان ۹۵٪ برای میانگین عمر لامپهای این دسته به دست آورید.

۴ الف - از یک جمعیت نرمال با واریانس ۴ یک نمونه تصادفی به اندازه ۲۵ انتخاب کرده‌ایم و میانگین این نمونه ۲۰ شده است. یک فاصله اطمینان ۹۰ درصدی برای میانگین این جمعیت پیدا کنید.

ب - اگر بخواهیم طول فاصله اطمینان را به نصف کاهش دهیم، اندازه نمونه را چه مقدار باید انتخاب کنیم؟