

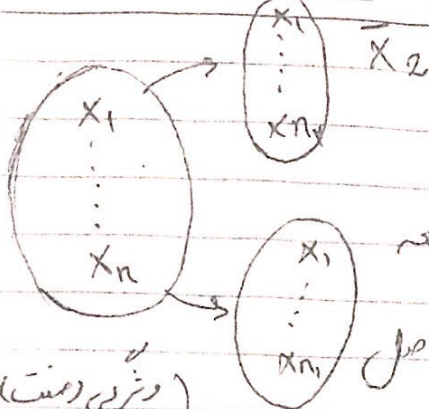
جزوه روشنی آماری

Subject:
Year:

Month:

Date:

نقشه اول



$$\bar{X}_1 = \frac{1}{n} \sum_{i=1}^n x_i \quad (n: \text{حجم نمونه})$$

هدف اصلی درس: نمونه گیری و پس تحلیل جامعه

و ارتباط پیرامون جامعه براساس نتایج حاصل

(x : ویژگی (مت)

(N : حجم جامعه)

از نمونه

جامع آماری: مجموعه ای از تمام عناصر است که بر سوره تعریف شده ای تعلق دارند عبارت دیگر مجموعه ای از افراد یا چیزهایی که می خواهیم یک یا چند ویژگی را بر آنها مطالعه کنیم.

نمونه: قسمتی از جامعه آماری است که طبق ضوابط انتخاب میشود و مطالعه ای آن

به جای مطالعه کل جامعه مقدور است؛ نمونه گیری باعث کاهش هزینه و افزایش

سرعت میشود همچنین با ابراز توان و کیفیت کار، عدم صدمه وارد کردن به واحدهای

جامعه

پارامتر: ویژگی عددی جامعه مربوط به صفت مورد بررسی مثل میانگین (م) جامعه:

که μ واریانس جامعه σ^2 است نسبت در جامعه که در صورتی که جامعه آماری بزرگ باشد مقدار دقیق آن را می توان محاسبه کرد. آماره \bar{x} تابعی از نمونه تصادفی می باشد.

پارامتر: بعضی مواقع علاقمند به اطلاع از برخی ویژگی های عددی یک جامعه می باشیم

Sadra

* پارامتر مربوط به جامعه و آماره مربوط به نمونه است *

Subject:

Year:

Month:

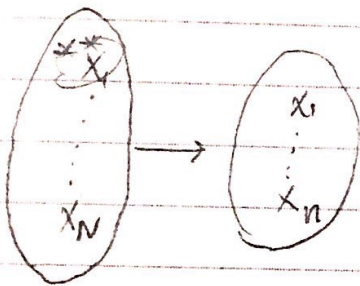
* * * * * چگونه می‌توانیم چون مقیاس‌های مختلف داریم

نشان می‌دهیم. این برآورد یکی از آماره‌های نمونه است. انفرادی سوال این

است که کدام آماره به عنوان برآورد کننده انتخاب شود. بنابراین برای

نمیتوانیم ویژگی‌های یک برآورد کننده خوب را بیان کنیم

کدام را به عنوان $\hat{\theta}$ انتخاب کنیم.



$$\begin{cases} T_2(x_1, \dots, x_n) = \bar{x} \\ T_1(x_1, \dots, x_n) = \min(x_1, \dots, x_n) \\ T_3(x_1, \dots, x_n) = \text{median}(x_1, \dots, x_n) \end{cases}$$

یکی از آماره \rightarrow برآورده $\hat{\theta} \rightarrow \theta$ پارامتر جامعه

۱. برآورد کننده نااریب باشد یعنی مقدار واقعی θ هر چه باشد داشته باشیم.

* برای پارامترها امید ریاضی نداریم. $E(\hat{\theta}) = \theta$ \rightarrow عدالت

نمونه تصادفی: نمونه x_1, \dots, x_n را از جامعه آماری با صفت x ^{می‌گیریم} نمونه

تصادفی می‌گیریم به طوری x_1, \dots, x_n دو به دو از هم مستقلند و هم توزیع با

x هستند. $x_1 \equiv x$ ^{توزیع} $\rightarrow F(\mu, \sigma^2)$ که توزیع $x \sim$ $x_n \equiv x$

$$x_1 \sim F(\mu, \sigma^2)$$

$$\vdots$$

$$x_n \sim F(\mu, \sigma^2)$$

Sadra

نکته: از بین آماره‌های داده‌شده، آنکه بهترین تقریب از بقیه است به عنوان برآورد کننده پارامتر جامعه انتخاب می‌کنیم که بهترین تعریف دارد: برآوردی بهتر است که خواص زیر را داشته باشد.

① نااریبی: فرض کنید پارامتر مورد بررسی جامعه θ باشد برانقصورت $\hat{\theta}$ را یک

برآورد کننده نااریب می‌گوییم هرگاه:

$$E(\hat{\theta}) = \theta$$

مثال: فرض کنید x_1, \dots, x_n یک نمونه تصادفی از جامعه‌ای با میانگین

μ و واریانس σ^2 باشد، تحت چه شرایطی برآورد کننده (برآوردگر)

$(a_i \text{ ثابت})$ $T = \sum_{i=1}^n a_i x_i$ یک برآورد کننده نااریب برای μ

است پس باید $E(T) = \mu$

$$E(x) = \sum x \cdot f(x)$$

$$E(T) = E\left(\sum a_i x_i\right) = E(a_1 x_1 + \dots + a_n x_n)$$

$$= E(a_1 x_1) + E(a_2 x_2) + \dots + E(a_n x_n)$$

$$= a_1 E(x_1) + a_2 E(x_2) + \dots + a_n E(x_n)$$

$$= a_1 \mu + a_2 \mu + \dots + a_n \mu = \mu (a_1 + \dots + a_n)$$

* شرط بی‌تاریکی $\sum a_i = 1$

a_i هالیه باشد

Sadra

$$E(x^2) = \text{Var}(x) + E^2(x)$$

$$V(x) = E(x)^2 - E^2 x$$

$$V(\bar{x}) = \frac{\sigma^2}{n} \rightarrow \text{داریانس میانگین}$$

Subject:

Year:

Month:

Date:

مثال: فرض کنید x_1, \dots, x_n توزیع تصادفی از جامعه با صفت x و میانگین

μ و واریانس σ^2 باشد نشان دهید:

$$E(\bar{x}) = \mu$$

$$E(S^2) = \sigma^2$$

نویسار استوار

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$E(\bar{x}) = \mu$$

$$E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} E\left(\sum x_i\right) = \frac{1}{n} E(x_1, \dots, x_n)$$

$$= \frac{1}{n} E(x_1) + E(x_2) + \dots + E(x_n) = \frac{1}{n} \times n(\mu) = \mu$$

$$E(S^2) = \sigma^2$$

$$E\left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n (x_i - \bar{x})^2\right)$$

د.~

$$E(S^2) = E\left(\frac{1}{n-1} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right) = \frac{1}{n-1} \sum_{i=1}^n [E(x_i^2) + E(\bar{x}^2) - 2E(x_i\bar{x})] \quad \leftarrow \text{خواص امید ریاضی}$$

$(x_i\bar{x}) \quad (*)$

از طرفی می دانیم $v(x_i) = E(x_i)^2 - E^2(x_i)$

① $E(x_i)^2 = \sigma^2 + E^2(x_i) = \sigma^2 + \mu^2$ ② $E(\bar{x})^2 = v(\bar{x}) = \frac{\sigma^2}{n} + \mu^2$

از طرفی $\sum_{i=1}^n 2E(x_i\bar{x})$ را می توان به صورت زیر نوشت:

③ $2E\left(\sum_{i=1}^n x_i\bar{x}\right) = 2nE(\bar{x})^2 = 2n\frac{\sigma^2}{n} + 2n\mu^2$

با جایگذاری ①، ②، ③ در (*)

$$E(S^2) = \frac{1}{n-1} \left[\sum_{i=1}^n (\sigma^2 + \mu^2) + \sum_{i=1}^n \left(\frac{\sigma^2}{n} + \mu^2\right) - 2n\frac{\sigma^2}{n} - 2n\mu^2 \right]$$

Sadra

$$= \frac{1}{n-1} \left[\underbrace{\sigma^2 + n\sigma^2}_{\text{ناتمام است}} \right] = \sigma^2 \quad 4.$$

X_1, \dots, X_n نمره تصادفی که دو به دو مستقلند و کوواریانسها صفر

Subject

Year:

Month:

Date:

محاسبه واریانس میانگین

$$V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} V(X_1 + \dots + X_n) = \frac{1}{n^2} [V(X_1) + \dots + V(X_n)]$$

$$= \frac{1}{n^2} [\sigma^2 + \dots + \sigma^2] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\begin{cases} V(X+Y) = V(X) + V(Y) + 2\text{Cov}(X, Y) \\ V(aX) = a^2 V(X) \\ V(a) = 0 \end{cases}$$

خواص
واریانس

مثال

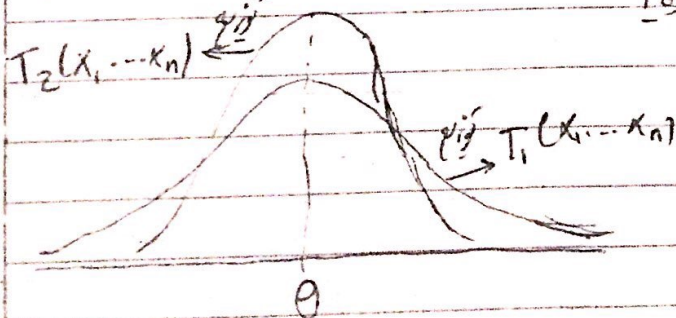
$$X \sim F(\mu, \sigma^2) \quad X_1, \dots, X_n \quad \begin{matrix} \rightarrow E(\bar{X}) = \mu \\ \rightarrow V(\bar{X}) = \frac{\sigma^2}{n} \end{matrix}$$

۲) اکنون سوال این است که بین دو آماره که هر دو ناربی هستند برای پارامتر مورد

تقریب جامعه (مثلاً θ) کدام را به عنوان برآورد کننده انتخاب کنیم. فرض کنید X متغیر تصادفی

از توزیع با پارامتر θ و واریانس σ^2 باشد و X_1, \dots, X_n نمونه‌های تصادفی

از X باشد و $T_1(X_1, \dots, X_n)$ و $T_2(X_1, \dots, X_n)$ هر دو برای θ ناربی باشند و توزیع



به صورت شکل زیر داشته باشند

$$E(T_1(X_1, \dots, X_n)) = \theta$$

$$E(T_2(X_1, \dots, X_n)) = \theta$$

با توجه به شکل ترجیح می‌دهیم برآورد کننده ناربی را انتخاب کنیم و برآوردی کمتری

Sadra

* mean square error

« میانگین مربع خطا »

Subject:

Year:

Month:

Date:

به عبارت دیگر واریانس کوچکتری داشته باشد برای این مثال چون T_2

واریانس کوچکتری دارد پس T_2 را بر T_1 ترجیح می‌دهیم.

نکته: در صورتیکه بین برآوردهای ارب می‌خواهیم برآورد خوب انتخاب کنیم

معمولاً برآوردی را انتخاب می‌کنیم که MSE^* کوچکتری داشته باشد به صورت

تقریب بشود:

$$MSE = \text{var}(\hat{\theta}) + \text{bias}^2(\hat{\theta})$$

به معنای

$$E(\hat{\theta}) \neq \theta \quad \text{یا} \quad E(\hat{\theta}) = \theta + \text{bias}$$

اربی

نکته: ثابت کنید:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = \text{var}(\hat{\theta}) + \text{bias}^2(\hat{\theta})$$

و بتوان برآورد میانگین بگیریم

« ادامه مثال قبل »

$$E(S^2) = E\left(\frac{1}{n-1} \sum_{i=1}^n (x_i^2 - \bar{x}^2 - 2x_i\bar{x})\right) = \frac{1}{n-1} \sum_{i=1}^n [E(x_i^2) + E(\bar{x}^2) - 2E(x_i\bar{x})]$$

← خواص امید ریاضی

$$v(x_i) = E(x_i^2) - E^2(x_i)$$

از طرف می‌دانیم

$$\textcircled{1} E(x_i^2) = \sigma^2 + E^2(x_i) = \sigma^2 + \mu^2 \quad \textcircled{2} E(\bar{x}^2) = v(\bar{x}) = \frac{\sigma^2}{n} + \mu^2$$

$$\sum_{i=1}^n 2E(x_i\bar{x})$$

از طرفی

$$\textcircled{3} 2E\left(\sum_{i=1}^n x_i\bar{x}\right) = 2nE(\bar{x})^2 = 2n\frac{\sigma^2}{n} + 2n\mu^2$$

اجایگذاری $\textcircled{1}$, $\textcircled{2}$, $\textcircled{3}$ در $*$

$$E(S^2) = \frac{1}{n-1} \left[\sum_{i=1}^n (\sigma^2 + \mu^2) + \sum_{i=1}^n \left(\frac{\sigma^2}{n} + \mu^2\right) - 2n\frac{\sigma^2}{n} - 2n\mu^2 \right]$$

Sadra

$$= \frac{1}{n-1} \left[\sigma^2 + n\sigma^2 \right] = \sigma^2$$

ذکر: برآورده های به نازیب باشد دارای کمترین واریانس و البته در فضای پارامتر تعریف شده باشد که برآورده خوب یا برآورده نازیب با واریانس کمینه نامیده میشود. خصوصیات دیگری برای برآورده شده تعریف میشود که به شرح زیر است:

(۳) سازگاری: سازگاری یعنی اگر حجم نمونه را افزایش دهیم $\hat{\theta}$ به θ میل کند

احتمال اختلاف آنها به اندازه ϵ هم صفر باشد یعنی

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \epsilon) = 0 \quad \forall \epsilon > 0$$

کمترین: از نامساوی چیسف ثابت کنید عبارت بالا معادل است با $\lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = 0$

کاری یک بر آورده: یک معیار برای مقایسه دو برآورده کاری است. فرض

کنید $u(x)$ و $v(x)$ دو برآورده پارامتر θ باشند در انصوت کاری این دو برآورده

به صورت زیر تعریف می شود:

$$e = \frac{\text{Var}(u)}{\text{Var}(v)}$$

$e < 1$ u کارآزاده است

$e > 1$ $u \sim v$

مثال) برای یک نمونه تصادفی دو تایی x_1, x_2 از توزیع $N(\mu, \sigma^2)$ داریم:

$$u(x) = \frac{x_1 + x_2}{2}, \quad T(x) = 2x_1 - x_2$$

۱) آیا $u(x)$ و $T(x)$ برای μ برآورد کننده هستند؟

$$= \frac{1}{2} [E(x_1) + E(x_2)]$$

۲) کدام برآورده بهتر است؟

$$= \frac{1}{2} [2\mu] = \mu$$

۳) کاری این دو برآورده را مقایسه کنید.

$$E(T(x)) = E(2x_1 - x_2) = 2E(x_1) - E(x_2) = 2\mu - \mu = \mu$$

$$\textcircled{2} \quad \text{Var}(u(x)) = \text{Var}\left(\frac{x_1 + x_2}{2}\right) = \frac{1}{4} \text{Var}(x_1 + x_2) = \frac{1}{4} [\text{Var}(x_1) + \text{Var}(x_2) + 2\text{Cov}(x_1, x_2)] = \frac{1}{4} [1 + 1 + 0] = \frac{1}{2}$$

$$\text{Var}(T(x)) = \text{Var}(2x_1 - x_2) = 4\text{Var}(x_1) + \text{Var}(x_2) - 4\text{Cov}(x_1, x_2) = 4 + 1 - 0 = 5$$

Sadra

$u(x)$ نسبت به $T(x)$ واریانس کوچکتری دارد پس بهتر است.

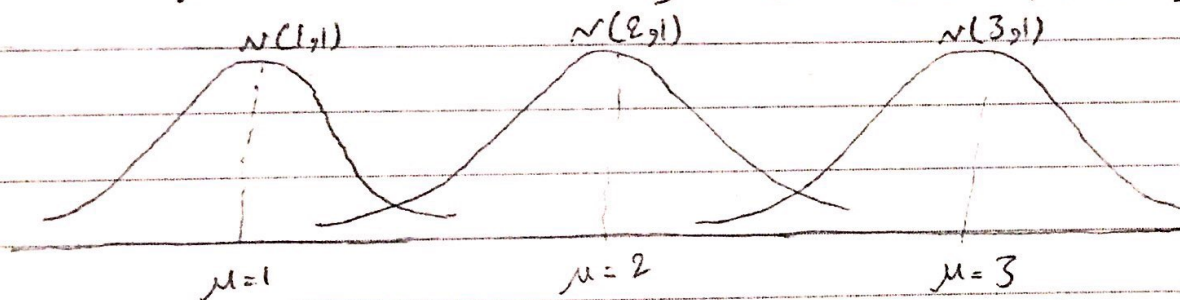
③ $e = \frac{\text{var}(u)}{\text{var}(T)} = \frac{\frac{1}{2}}{5} = \frac{1}{10} < 1$. u ، در برابر کارآزایی است.
 "بخش دوم"

- توزیع نرمال فرض کنید X یک غیر تصادفی پیوسته با تابع چگالی زیر باشد:

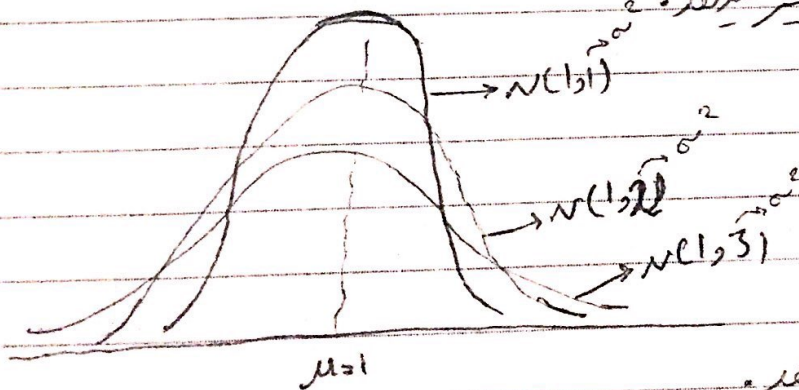
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < +\infty$$

و بماند $N(\mu, \sigma^2)$ نشان میدهیم به μ میانگین یا پارامتر مکان و به σ^2

واریانس یا پارامتر شکل (شیاس) میگویند. $E(X) = \mu$ و $\text{var}(X) = \sigma^2$



* پس μ مکان توزیع را تغییر میدهد.

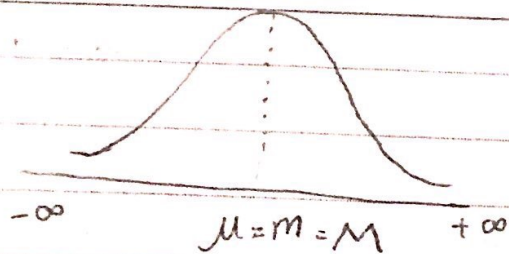


* پس σ شکل توزیع را تغییر میدهد.

- با توجه به خواص خوبی که توزیع نرمال دارد معمولاً این توزیع مورد توجه است؛ از جمله

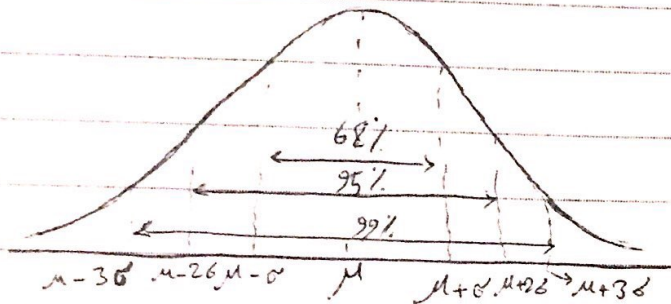
خواص این توزیع میتوان به موارد زیر اشاره کرد:

① دامنه تعریف X ، کل اعداد حقیقی است $(-\infty, +\infty)$



② توزیع متعارف است و شکل زیر دارد:

③ بازه بندی توزیع مشخص است یعنی اگر $X \sim N(\mu, \sigma^2)$ داریم:



④ توزیع نرمال تحت ترکیبات خطی بسته است.
دو ویژگی:

if $X \sim N(\mu, \sigma^2)$; $a, b \Rightarrow aX + b \sim N$

$$\begin{cases} E(aX + b) = aE(X) + b = a\mu + b \\ V(aX + b) = a^2 V(X) = a^2 \sigma^2 \end{cases}$$

if $X_1 \sim N(\mu_1, \sigma_1^2)$

$X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1, X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
مثلاً X_2

⑤ تحت شرط بودن بسته است:

$(X_1, X_2) \sim N_2$ و $X_1 | X_2 \sim N_1$ $X_2 | X_1 \sim N_1$

⑥ تحت حاشیه سازی هم بسته است:

$(X_1, X_2) \sim N_2 \Rightarrow X_1 \sim N_1$
 $X_2 \sim N_1$

Sadra

$$E(S_n^2) \neq \sigma^2$$

مثال نشان دهید

$$E(S_n^2) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

نکته: بعد از ساده کردن S^2 ، می توان از فرمول زیر

هم استفاده کرد:

$$S^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - n \bar{x}^2 \right]^*$$

فرمول معروف است (حقیقا با اثبات)

$$S^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 + \sum_{i=1}^n \bar{x}^2 - 2 \sum_{i=1}^n x_i \bar{x} \right]$$

$\underbrace{\sum_{i=1}^n \bar{x}^2}_{n \bar{x}^2} \quad \underbrace{- 2 \sum_{i=1}^n x_i \bar{x}}_{- 2 n \bar{x}^2}$

«مهم ترین» مثال دیگر اگر $(\mu, \sigma^2) \sim X$ باشد و x_1, \dots, x_n نمونه

صادق از X باشند آنوقت $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ و $X \sim N(\mu, \sigma^2)$

$$\begin{cases} x_1 \sim N(\mu, \sigma^2) \\ \vdots \\ x_n \sim N(\mu, \sigma^2) \end{cases}$$

طبق خاصیت کزنرمال تحت ترکیبات خطی

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + \dots + x_n)$$

پس بر این ترکیب نرمال خطی است.

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$\xrightarrow{E(\bar{X})} \mu(\bar{X})$

$$X \sim N(\mu, \sigma^2) \xrightarrow[\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i]{\text{نمونه تصادفی } x_1, \dots, x_n} \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

تعی حد درزی: اگر X متغیر تصادفی از هر توزیع با میانگین μ و واریانس σ^2 باشد

و X_1, \dots, X_n نمونه تصادفی "به حد کافی بزرگ" از X باشد آنگاه \bar{X} دارای

توزیع نرمال است. $N(\mu, \frac{\sigma^2}{n})$ $\bar{X} \approx$ $\xrightarrow[n \rightarrow \infty]{X_1, \dots, X_n}$ $F(\mu, \sigma^2)$ هر توزیع

معمولاً $n \geq 30$

توزیع نرمال استاندارد: هر توزیع نرمالی را میتوان به توزیع نرمال استاندارد

تبدیل کرد، فرض کنید X دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشد،

آنگاه متغیر تصادفی جدید Z که به صورت $Z = \frac{X - \mu}{\sigma}$

تعریف میشود دارای توزیع $N(0, 1)$ است که به توزیع نرمال استاندارد معروف است

و مقادیر $P(Z \leq z)$ به ازای z های مختلف با استفاده از نرم افزار محاسبه شده

است و در انتهای این کتاب آماری موجود است.

مثال: یک شرکت تولیدی لاستیک اتومبیل، لاستیکهای تولید میکند به طول عمر این لاستیکها

دارای توزیع نرمال با میانگین 24 ماه و انحراف معیار 2 ماه است. احتمال اینکه در یک نمونه

تصادفی 25 تای از لاستیکها میانگین طول عمر کمتر از 25 ماه باشد را بیابید.

$$X \sim N(24, 4) \quad \mu = 24, \sigma = 2, n = 25$$

$$\bar{X} \sim N(24, \frac{4}{25})$$

$$P(\bar{X} < 25) = P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{25 - 24}{\frac{2}{5}}\right) = P(Z < 2.5)$$

Sadra

7

= 0.9938

مثال فرض کنید مقدار سالهای تحصیل در بین افراد بالغ در شهری دارای میانگین

۱۱ سال و انحراف معیار سه سال باشد، احتمال اینکه در یک نمونه تصادفی

صد نفری از افراد متوسط مقدار سالهای تحصیل بین ۱۱ تا ۱۲ سال باشد را بیابید.

چون جامعه نرمال نیست اما n به اندازه کافی بزرگ است پس از قضیه مرکزی استفاده میکنیم.

$$P(11 < \bar{X} < 12)$$

$$X \sim F(11, 1, 9) \xrightarrow{n \geq 30} \bar{X} \sim N(11, 1, \frac{9}{100})$$

$$E(x) = \mu = 11,1 \quad V(\bar{x}) = \frac{\sigma^2}{n} = \frac{9}{100}$$

$$P\left(\frac{11 - 11,1}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{12 - 11,1}{\frac{\sigma}{\sqrt{n}}}\right) = P(-0,33 < Z < 3) \quad P(Z \leq z)$$

$$= P(Z < 3) - P(Z < -0,33)$$

$$= \Phi(3) - \Phi(-0,33)$$

$$= 0,9987 - 0,3707 = 0,628$$

توزیع های دو: فرض کنید جامعه ای دارای توزیع نرمال است و X متغیر تصادفی مورد

$$X \sim N(\mu, \sigma^2) \quad \text{بررسی:}$$

و X_1, \dots, X_n نمونه ای تصادفی از این جامعه، طبق تعریف نمونه تصادفی:

$$X_1, \dots, X_n \sim N(\mu, \sigma^2) \rightarrow X_i \sim N(\mu, \sigma^2) \quad i=1, \dots, n$$

در صورت استفاده از برآورد در میان و یا عدد محدودیت بود که باید از درجه آزادی به همان مقدار کم کنیم *

Subject: _____
Year: _____ Month: _____ Date: _____

از طرفی (اگر) $Z_i = \frac{x_i - \mu}{\sigma}$ انون تابع جدید زیر داری

توزیع گای دو است: $Z_i^2 = \frac{(x_i - \mu)^2}{\sigma^2} \sim \chi_1^2$

$\sum_{i=1}^n Z_i^2 \sim \chi_n^2$
 (مجموع گای دو) (خودم)
 توزیع

نکته: اگر در جامعه آن بر مجهول باشد و در رابطه: $\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$ به جای μ از \bar{x} استفاده می شود از درجه آزادی، یک کمتر شود

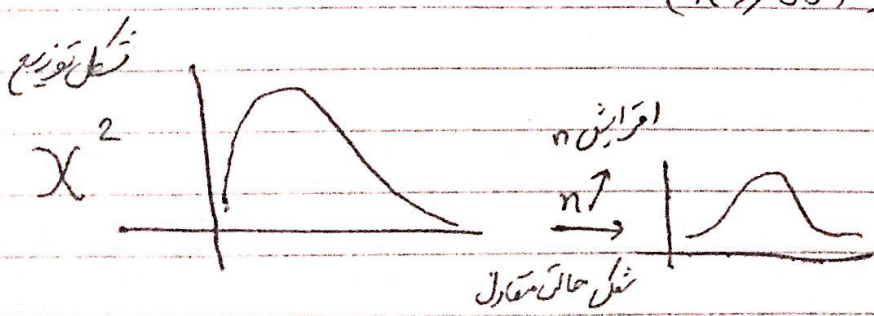
میانگین جامعه $\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma^2} \sim \chi_{n-1}^2$
 میانگین نمونه

نکته: همچنین ثابت می شود که:

if $y \sim \chi_n^2 \Rightarrow E(y) = n$

$V(y) = 2n$

نکته: توزیع گای دو به درجه آزادی حساس است و با افزایش درجه آزادی توزیع گای دو به نرمال گسترده می شود ($n \geq 30$).



Sadra

Subject:

Year: 17

Month:

Date:

$$X \sim N(\mu, \sigma^2)$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \xrightarrow{\frac{\sigma^2}{\sigma^2}} \frac{\sigma^2}{n-1} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2$$

توزیع S: $\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2$

$$\Rightarrow \frac{n-1 S^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$X_i \sim N(\mu, \sigma^2)$$

$$\Rightarrow \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$$

طبقاً * در مقادیر

انفونی داریم:

$$E\left(\frac{(n-1)S^2}{\sigma^2}\right) = n-1$$

$$\frac{(n-1)}{\sigma^2} E(S^2) = n-1 \Rightarrow E(S^2) = \sigma^2$$

بهین ترتیب S^2 برای σ^2 ناریت است.

بهین ترتیب:

$$V\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

$$\frac{(n-1)^2}{\sigma^4} V(S^2) = 2(n-1) \Rightarrow V(S^2) = \frac{2\sigma^4}{(n-1)}$$

$$\lim_{n \rightarrow \infty} V(S^2) = 0 \Rightarrow S^2 \text{ برای } \sigma^2 \text{ ناریت است.}$$

تبرین) برای S_n^2 تمام مراحل بالا را طی کن. (توزیع $X \sim N(\mu, \sigma^2)$ ، X_1, \dots, X_n)

و نشان بدهید S_n^2 برای σ^2 ناریت است و کارایی S_n^2 را با S^2 مقایسه کنید.

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Sadra

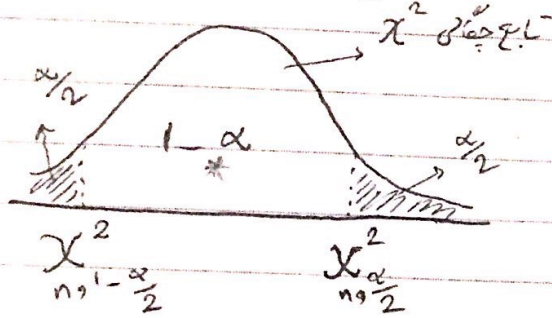
Subject:

Year:

Month:

Date:

نکته: توزیع های دو به صورت چوله به راست است بنابراین:



بازای مقایسه مختلف α ، مقایسه $\chi^2_{\alpha/2}$

و $\chi^2_{1-\alpha/2}$ را می توانید از جدول آستادی کت ب

استخراج کنید.

نکته: دقت کنید اگر در مسئله ای میانگین معلوم باشد، نگاه رفرمولهای S^2 و S_n^2 به

جای \bar{X} از μ استفاده میشود نگاه:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 \quad , \quad S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

$$\Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi_n^2 \quad , \quad \Rightarrow \frac{nS_n^2}{\sigma^2} \sim \chi_n^2$$

تمرین) به عنوان تمرین در حالتی که μ معلوم است، $E(S^2)$ و $E(S_n^2)$ را بیابید و

واریانس آنها را مقایسه کنید.

توزیع تصادفی

توزیع t استودنت: فرض کنید $X \sim N(\mu, \sigma^2)$ و $X_1, \dots, X_n \sim N(\mu, \sigma^2)$

آنگاه میدانیم: $Z_i = Z_i = \frac{X_i - \mu}{\sigma} \sim N(0, 1)$

$$\sum_{i=1}^n Z_i^2 \sim \chi_n^2$$

Subject:

Year:

Month:

Date:

$$\frac{N(0,1)}{\sqrt{\chi^2_{df}/df}} = t_{df}$$

در این صورت:

$$T = \frac{Z}{\sqrt{\frac{U}{n}}} \sim t_n$$

توزیع نرمال

توزیع t

* مقدار دلی بختراز

$$\sqrt{\chi^2_{df}/df} = \sqrt{\chi^2_n/n}$$

degree of freedom

نرمال

درجه آزادی

اگر درجه آزادی بالا برود شکل نرمال شبیه شد و بالا می رود

درجه آزادی

$$n \nearrow \quad t \approx N$$

مثال) از جامعه‌ای با توزیع نرمال نمونه‌ای به حجم n انتخاب می‌شود، نشان

دهی $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ داریم توزیع t است و با n-1 درجه آزادی است.

①

$$X \sim N(\mu, \sigma^2) \quad X_1, \dots, X_n \Rightarrow Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

قبلاً بدویم $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

②

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

①, ② $\Rightarrow T = \frac{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2} / (n-1)}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

Sadra

درجه آزادی \rightarrow درج ب
مردم

Subject: معادلات و توابع

Year: Month: Date:

مثال: نمرات یک کلاس از دانشجویان طراری توزیع نرمال با میانگین 15 است. اگر از این کلاس یک نفر 20 نمره انتخاب کنیم و مشاهده کنیم که انحراف استاندارد نمرات (انحراف بسیار و غیر وارثی) 4.28 است. احتمال اینکه میانگین نمرات

$$\mu = 15$$

$$n = 20$$

$$S = 4.28$$

این افراد از 17 بیشتر باشد را بیابید.

$$P(\bar{X} > 17) = P\left(\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} > \frac{17 - 15}{\frac{4.28}{\sqrt{20}}}\right)$$

$$= P(T > 2.109) = 1 - P(T < 2.109) \quad t_{n-1}$$

$$= 0.025$$

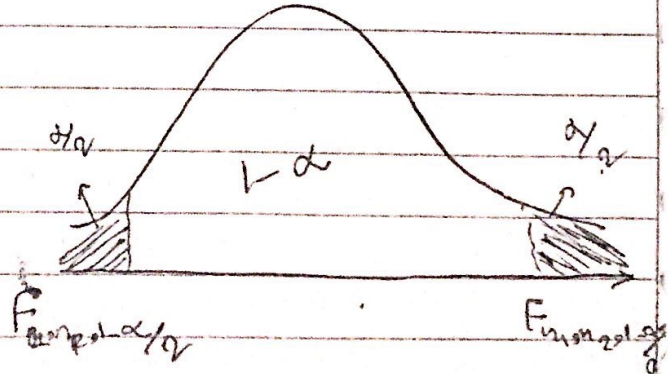
از جدول

توزیع فیشر:

if $u_1 \sim \chi^2_{n_1}$ و $u_2 \sim \chi^2_{n_2}$

$$\Rightarrow F = \frac{u_1/n_1}{u_2/n_2} \sim F_{n_1, n_2} \rightarrow \text{درجه آزادی خروجی} \quad \text{درجه آزادی ورودی}$$

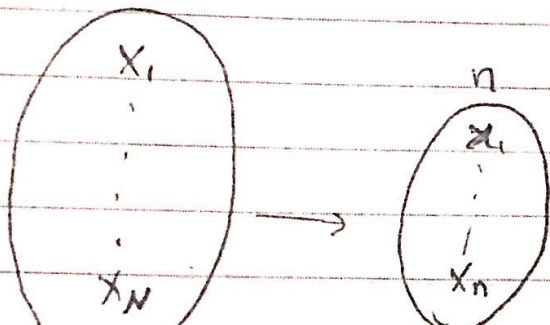
$$F_{n_1, n_2, \alpha} = \frac{1}{F_{n_2, n_1, 1-\alpha}}$$



Sadra

بر آورد نقطه ای: ها نظریه قبل از اشاره شد یکی از اهداف اصلی این درس بر آورد

نقطه ای به بار می آوریم " میانگین جامعه " و " واریانس جامعه " و " نیت "



میانگین نمونه

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - n\bar{X}^2 \right]$$



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

$$E(\bar{X}) = \mu, \quad V(\bar{X}) = \frac{\sigma^2}{n}$$

همچنین قبلاً دیدیم

$$\lim_{n \rightarrow \infty} V(\bar{X}) = 0$$

و ثابت می شود که $V(X)$ یعنی بر آورد کننده های نا اید ب کمترین واریانس است

پس \bar{X} را به عنوان بر آورد نقطه ای خوب می گیریم.

نکته: توزیع \bar{X} بستگی به توزیع نمونه دارد مثلاً اگر جامعه نرمال باشد

دیدیم چون \bar{X} تابعی خطی از نمونه تصادفی است و توزیع نرمال تحت ترکیبات خطی

بسته است \bar{X} هم دارای توزیع نرمال است.

$$X \sim N(\mu, \sigma^2) \xrightarrow{\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i} \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

(P) نت همان افعال موفقیت است

Subject:

Year:

Month: _____

Date: _____

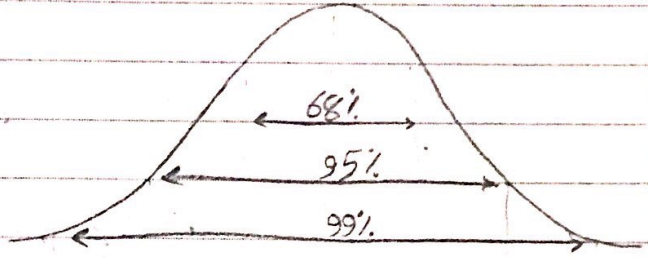
نکته: اگر $X \sim F$ (تابع هر توزیعی داشته باشد) به طوریکه نمودار استخراج

شده به اندازه کافی بزرگ باشد نگاه طبق قضیه حد مرکزی :

$$X \sim F(\mu, \sigma^2) \xrightarrow{x_1, \dots, x_n} \bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

محکم دلائل سے مزین متنوع و منفرد موضوعات پر مشتمل مفت آن لائن مکتبہ

$n \nearrow$
 $n \geq 30$


$$\mu - \frac{3\sigma}{\sqrt{n}} \quad \mu - \frac{2\sigma}{\sqrt{n}} \quad \mu - \frac{\sigma}{\sqrt{n}} \quad \mu \quad \mu + \frac{\sigma}{\sqrt{n}} \quad \mu + \frac{2\sigma}{\sqrt{n}} \quad \mu + \frac{3\sigma}{\sqrt{n}}$$

مرکز نقطه ای ثبت جامعه P : گاهی در محل هدف برآورد ثبت جامعه است،

برای مثال فرض کنید ویرگی شاغل بودن یا نبودن افراد جامعه مورد نظر است که مطالعه‌ی

کل جامعه امکان پذیر نیست و من خواهم از روی یک نمونه تصادفی P را برگزینم و در آن
تعدادشان n (تعداد افراد دارای صفت خاص)

$N \rightarrow 1001110101 \rightarrow P: X$

بسم الله الرحمن الرحيم
الحمد لله رب العالمين

$1010010 \rightarrow X \sim \text{bin}(n, p)$
 $\underbrace{\hspace{1.5cm}}_n$ $\underbrace{\hspace{1.5cm}}_{p \text{ independent}}$

X تعداد افراد که در نمونه شامل هستند

از طرفی میدانیم در توزیع دو جمله ای $X \sim \text{bin}(n, p)$

①

$$E(x) = np$$

(2)

$$v(x) = n p \frac{q}{r}$$

$$q = 1 - p$$

$\frac{1}{2} = \frac{x}{\text{Sadra}^2}$

چون تعداد افراد هستند که در نمونه صفت خاص را دارند پس
شماره اول

نسبت در نمونه است (نسبت شایع نمونه) اکنون میتوان نشان داد که نسبت نمونه

$$\hat{p} = p$$

برآورد نقطه ای خوبی برای نسبت در جامعه است.

\hat{p} نسبت جامعه p نسبت نمونه

$$E(\hat{p}) = E(p) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) \stackrel{(1)}{=} \frac{1}{n} \cdot np = p \quad (3)$$

از (1) و (2)

یعنی $\hat{p} = p$ برآورد ناریب برای p است.

$$V(\hat{p}) = V(p) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n^2} \cdot npq = \frac{pq}{n} \quad (4)$$

$$V(\hat{p}) \xrightarrow{n \rightarrow \infty} 0$$

و ثابت میشود \hat{p} دارای منطبق واریانس بین برآوردهای ناریب p است

پس $\hat{p} = p$ به عنوان "برآورد خوب" p شناخته میشود.

نکته: اگر n به اندازه کافی بزرگ باشد میتوان نشان داد \hat{p} دارای توزیع

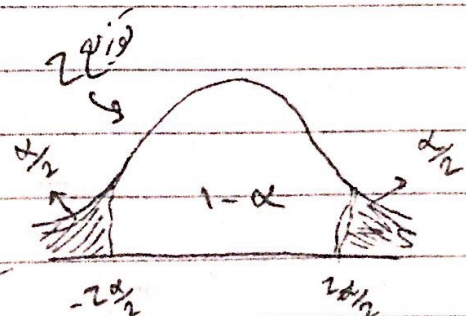
تقریبی نرمال در باشد:

$$\hat{p} \stackrel{(3, 4)}{\approx} N\left(p, \frac{pq}{n}\right)$$

و

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$\sim N(0, 1)$



Sadra

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

با قابلیت Z :

$$P(-Z_{\frac{\alpha}{2}} < \frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}} < Z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$\Rightarrow P(|\hat{P} - P| < Z_{\frac{\alpha}{2}} \sqrt{\frac{Pq}{n}}) = 1 - \alpha$$

نکته: چون $\frac{Pq}{n}$ مجهول است معمولاً از برآورد آن استفاده می‌شود یعنی

حاشیه خطا برابر است با $Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}\hat{q}}{n}}$ و معمولاً با "S.E." نشان می‌دهند.
برای خطا

مثال: شرکتی به از طریق نظرسنجی فروش در یک ماهانه پیشنهادهای ویژه‌ای را

از طریق پست برای افراد فرستاد یک پیشنهاد که آزمایش برای نمونه تصادفی شامل 250 نفر

انتخاب شدند و به وسیله پست برای افراد نمونه ارسال می‌شود بر مبنای این نمونه

70 نفر تصمیم به خرید می‌گیرند یک برآورد نقطه‌ای برای نسبت اعضای که انتظار می‌رود از

طریق پست خرید کنند بدست آورید و حاشیه خطای 95٪ را برای این برآورد بدست

تقدیر

برآورد $X = 70$ تعداد افراد که خرید کردند

$$\hat{p} = \frac{70}{250} = 0.28$$

چون $0.025 \leq \alpha = 0.05 \leq 1 - \alpha = 0.975$

$$Z_{0.025} = 1.96 \approx 2$$

$$S.E.(\hat{P}) = Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}\hat{q}}{n}} = 2 \sqrt{\frac{0.28(1-0.28)}{250}} = 0.056$$

Sadra

مثال: فرض کنید داده‌های زیر از جامعه نرمال استخراج شده است. میانگین جامعه را برآورد کنید و حاشیه خطا برآورد را بدست آورید. ($\sigma = 1$)

۱ ۲ ۴ ۳ ۵

$$\text{حاشیه خطا} = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{1}{\sqrt{5}}$$

$$\hat{\mu} = \bar{X} = \frac{1+2+4+3+5}{5} = 3$$

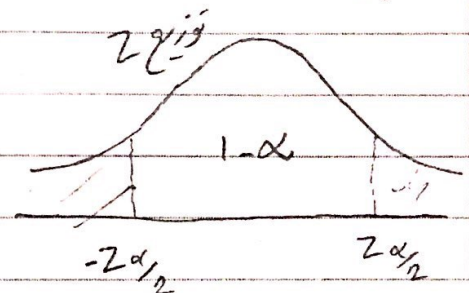
$$1-\alpha = 0.95 \rightarrow \frac{\alpha}{2} = 0.025$$

$$\alpha = 0.05$$

$$Z_{0.025} = 1.96$$

میدانیم اگر جامعه نرمال باشد (μ, σ^2) $\bar{X} \sim N$ پس:

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



$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1 - \alpha$$

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

$$\Rightarrow P(|\bar{X} - \mu| < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

برآورد نقطه‌ای واریانس جامعه (σ^2): یکی دیگر از پارامترهای مهم برای استنباط

پیرامون جامعه σ^2 است که قبلاً دیدیم:

$$E(S^2) = \sigma^2$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

تقریب
نمودار

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

اگر جامعه نرمال با n

(وقتی از Z استفاده می‌کنیم یعنی از توزیع نرمال استفاده می‌کنیم)

Subject:

Year:

Month:

Date:

ثابت می‌شود S^2 برآورد خوب σ^2 به همین برآوردکننده‌های ناریب

$$\hat{\sigma}^2 = \sigma^2$$

است.

* دقت کمتری ناریب تحت جذر گرفتن حفظ نمی‌شود

$$E(S^2) = \sigma^2 \Rightarrow E(S) = \sigma \quad \text{و} \quad E(S) \neq \sigma$$

مقدار ناریب به n بستگی دارد بطوریکه

ناریب با افزایش n به سمت صفر می‌رود.

$$E(S) \approx \sigma$$

$n \uparrow$

تعیین حجم نمونه: برای تعیین حجم نمونه معمولاً بستگی به مطالعه مقدار خطا (d) مشخص

می‌شود و براساس d، حجم نمونه مشخص می‌شود:

$$\textcircled{1} \quad P(|\bar{X} - \mu| < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

خطا

$$\textcircled{2} \quad |\bar{X} - \mu| \leq d \quad \text{از طرفی}$$

$$\textcircled{1}, \textcircled{2} \Rightarrow d = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow d^2 = Z_{\alpha/2}^2 \frac{\sigma^2}{n} \Rightarrow n = Z_{\alpha/2}^2 \frac{\sigma^2}{d^2}$$

نکته: اگر σ^2 مجهول باشد می‌توانیم یک نمونه اولیه استخراج کنیم

$$\hat{\sigma}^2 = S^2 \quad \hat{n} = Z_{\alpha/2}^2 \frac{\hat{\sigma}^2}{d^2}$$

کنیم پس

$$\textcircled{2} \quad \text{اگر مسئله "بیت جامعه" باشد داریم} \quad P(|\hat{p} - p| < Z_{\alpha/2} \sqrt{\frac{pq}{n}}) = 1 - \alpha$$

Sadria

$$\Rightarrow \hat{p} - p \leq d \Rightarrow d = Z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

$$\Rightarrow d^2 = Z_{\alpha/2}^2 \frac{pq}{n} \Rightarrow n = Z_{\alpha/2}^2 \frac{pq}{d^2}$$

مثلاً p و q مجهولند و در نتیجه pq مجهولات و باید از برای آن استفاده

کرد، در بدترین حالت میتوان $p = q = \frac{1}{2}$

$$n = Z_{\alpha/2}^2 \cdot \frac{1}{4d^2}$$

یا نمونه اولیه استخراج میکنیم و p را حساب میکنیم.

$$\hat{n} = Z_{\alpha/2}^2 \frac{\hat{p}\hat{q}}{d^2}$$

مثال) یک زیست شناس در خواهر میانگین نمرات موجود در حجم آب دریاچان را

محاسبه کند در محاسبات سال گذشته معلوم شده است که انحراف معیار نمرات موجود

در دریاچه 4 است، چه تعداد از نمونه این زیست شناس برای برآورد میانگین باید

بگیرد تا 90٪ مطمئن باشد که خطای برآورد از 0.1 بیشتر نخواهد بود.

$$n = ? , 1 - \alpha = 0.9 , d = 0.18 , \sigma = 4$$

$$1 - \alpha = 0.9 \Rightarrow \alpha = 0.1 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow Z_{0.05} = 1.64$$

Sadra

$$n = (1,64)^2 \times 16 = 67,75 \Rightarrow n \approx 68$$

مثال) یک بررسی به منظور برآورد نسبت کسانی که دید خاص ندارند در یک برنامه بهداشت

عمومی انجام میشود چند نفر را باید معاینه کرد؟ اگر وزارت بهداشت بخواهد با 98٪ اطمینان

خطای برآورد کمتر از 0,05 باشد و متغیر:

$$d = 0,05$$

$$1 - \alpha = 0,98$$

الف) هیچ اطلاعی از p نداریم.

$$\rightarrow \alpha = 0,02$$

ب) p براساس نمونه اولیه 3/4 تخمین زده شده است.

$$\rightarrow \frac{\alpha}{2} = 0,01$$

$$\rightarrow Z_{0,01} = 2,33$$

$$n = ? \Rightarrow n = (2,33)^2 \times \frac{1}{4(0,05)} = 543$$

$$n = (2,33)^2 \times \frac{(0,3)(0,7)}{(0,05)^2} = 456$$

نکته: از حالت های بدست آوردن فاصله اطمینان بداند (برآورد فاصله ای (فاصله اطمینان، بازه اطمینان)؛ چون با احتمال 1، برآورد

نقطه ای برابر پارامتر نیست معمولاً علاقه مند به ارائه یک فاصله برای پارامتر می باشیم که با

اطمینان $(1-\alpha)$ 100٪ (مثلاً 95٪، $\alpha = 0,05$ ، $1-\alpha = 0,95$ ، اطمینان 95٪) مطمئن

باشیم که فاصله مذکور پارامتر را در بر دارد. برای مثال اگر پارامتر مورد نظر 0 باشد، معمولاً

یک فاصله اطمینان براساس برآورد نقطه ای محاسبه می شود و به صورت زیر بیان می شود.

$$\text{Sadra} \quad \text{فاصله اطمینان} \rightarrow (L, U) : P(L < \theta < U) = 1 - \alpha$$

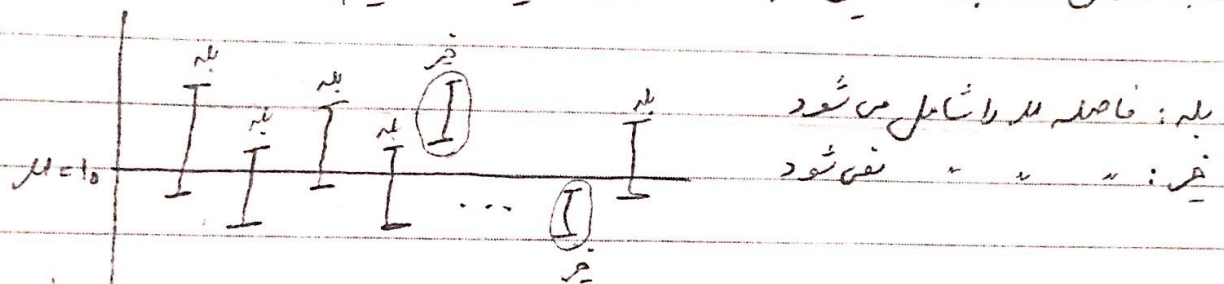
14 که ضریب اطمینان

تفسیر فاصله اطمینان، مقدر این است که اگر مثلاً فاصله اطمینان 95٪ برای پارامتر

میانگین جامعه محاسبه شده است، انتظار داریم اگر 100 بار نمونه گیری انجام دهیم و 100 فاصله

اطمینان برای هر بدست آوریم، 95 بار فاصله در شامل شوند (فرض کنید در واقعاً

جامعه ای، باشد از این جامعه 100 بار نمونه گیری کرده ایم)



از 100 فاصله اطمینان 95 فاصله، در شامل می شوند

برآورد فاصله ای برای پارامتر میانگین جامعه: در این بخش در حالت های مختلف برآورد

فاصله ای را برای پارامتر هر که مجهول است بدست می آوریم.

واریانس جامعه معلوم

حالت اول: (فاصله اطمینان برای هر وقتی جامعه نرمال است)، n بزرگ یا کوچک

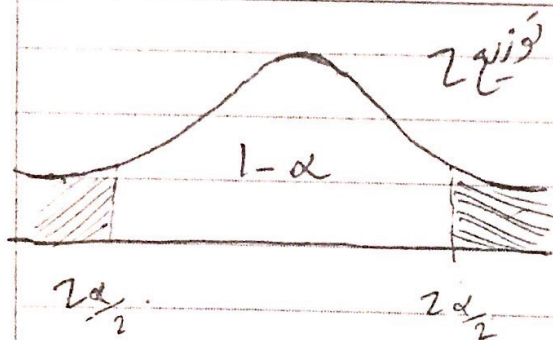
فرض کنید X_1, \dots, X_n نمونه های تصادفی از جامعه نرمال با میانگین μ و واریانس σ^2 باشد. قبلاً دیدیم \bar{X} برآورد خوب برای μ است و وقتی جامعه نرمال است \bar{X} به ترتیب

خطی از X_1, \dots, X_n است نیز نرمال است: X_1, \dots, X_n نمونه تصادفی

$$X \sim N(\mu, \sigma^2) \xrightarrow{X_1, \dots, X_n} \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

$N(0, 1)$

Sadra



$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1 - \alpha$$

باجایگزینی Z :

$$P\left(-Z_{\alpha/2} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\alpha/2}\right) = 1 - \alpha$$

با توجه داشتن μ در وسط و انتقال بقیه جملات برطرفین نامی داریم:

$$P\left(\underbrace{\bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}_{\text{مرکز یا من فاصله اطمینان}} < \mu < \underbrace{\bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}_{\text{مرکز یا بالا فاصله اطمینان}}\right) = 1 - \alpha$$

$$\boxed{U - L = \text{طول فاصله اطمینان}}$$

پس یک فاصله اطمینان $\% (1 - \alpha) 100$ برای μ

$$\boxed{\mu = \left(\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)} \quad \text{①} \quad \text{وقتی جامعه نرمال است و σ^2 معلوم:}$$

مثال ۱: بر مبنای یک نمونه تصادفی شامل ۲۵ مشاهده از جامعه نرمال با میانگین μ معلوم و σ^2 معلوم و جامعه نرمال

و انحراف معیار ۸ میانگین جامعه ۴۲٫۷ بدست آمده است یک فاصله اطمینان

۹۵٪ برای میانگین جامعه بسازید.

$$n = 25, \bar{X} = 42,7, \sigma = 8, 1 - \alpha = 0,95 \Rightarrow \alpha = 0,05 \Rightarrow \frac{\alpha}{2} = 0,025$$

$$Z_{0,025} = 1,96 \quad \mu: \left(42,7 \pm 1,96 \frac{8}{\sqrt{25}} \right) = (39,6, 45,8)$$

نکته: اگر واریانس جامعه معلوم نباشد فقط در صورتیکه بزرگ باشد چون میتوان

ازاریبی t نسبت به z چشم پوشی کرد در رابطه فاصله اطمینان ① میتوان نری

$$\text{به جای } z \text{ استفاده کرد: } \mu = \left(\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \quad ②$$

جامعه نرمال و n مجهول

حالت دوم: اگر جامعه نرمال نباشد اما بزرگ باشد ثابت می شود در حالتیکه

واریانس معلوم است از رابطه ① و اگر واریانس مجهول است میتوان از رابطه ②

استفاده کرد طبق قضیه حد مرکزی

فرض کنید X_1, \dots, X_n نمونه ای از جامعه ای بایه توزیع باشد چون فرض

کرده ایم n بزرگ است طبق قضیه حد مرکزی $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$\text{پس } Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ داریم:}$$

$$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1 - \alpha$$

با جایگذاری Z :

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

باشد داشتن μ و σ و انتقال به سمت راست و چپ داریم Sadra

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

پس یک فاصله اطمینان $1 - \alpha$ برای μ و σ جابجه نرمال نیست،

$$\mu: \left(\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) \quad \text{نیز σ^2 و n بزرگ و معلوم:}$$

σ^2 معلوم و n بزرگ

و σ^2 مجهول و n بزرگ

$$\mu: \left(\bar{X} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}} \right) \quad \text{نیز σ^2 و n مجهول}$$

حالت سوم: (n کوچک باشد و توزیع جابجه نرمال باشد) فرض کنید X_1, \dots, X_n

نمونه‌های تصادفی از $X \sim N(\mu, \sigma^2)$ باشد اگر σ^2 معلوم باشد:

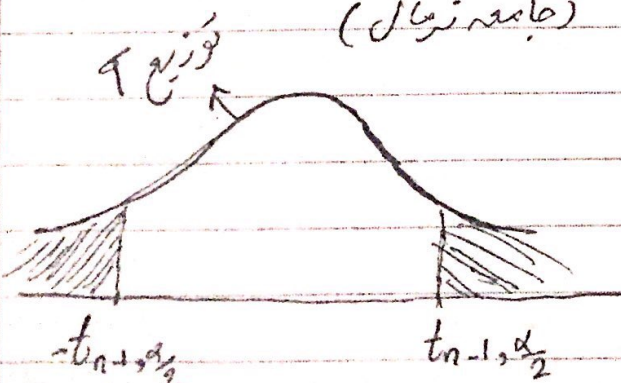
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \rightarrow Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

و دقیقاً مشابه مطالب حالت ① و ② تکرار می‌شود و داریم:

$$\mu: \left(\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) \quad \text{جابجه نرمال، σ^2 معلوم، n کوچک}$$

اما اگر σ^2 معلوم نباشد:

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1} \quad \text{قبلاً دیدیم: (جابجه نرمال)}$$



Sadra

* n بزرگ و از آمار بی چشم پوشی کرده و در هر دو حالت از آن استفاده

شود *

Subject:

Year:

Month:

Date:

$$P(-t_{n-1, \frac{\alpha}{2}} < T < t_{n-1, \frac{\alpha}{2}}) = 1 - \alpha$$

$$P(-t_{n-1, \frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} < t_{n-1, \frac{\alpha}{2}}) = 1 - \alpha$$

باینده داشتن الموسط و انتقال سایر جملات به طرفین نامساوی داریم:

$$P(\bar{X} - t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}}) = 1 - \alpha$$

پس یک فاصله اطمینان $1 - \alpha$ (۱۰۰٪) برای الموسط جامعه نریال است = n

نویس و نه مجهول:

$$\mu = \left(\bar{X} \pm t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right)$$

نه مجهول و
نویس، جامعه نریال

خلاصه مطالب

$$\mu = \left(\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right) \text{ جامعه نریال باشد } \sigma^2 \text{ معلوم باشد}$$

$$\mu = \left(\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \text{ نریال } n$$

$$\mu = \left(\bar{X} \pm t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \text{ مجهول باشد } n$$

$$\mu = \left(\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right) \text{ معلوم } \sigma^2 \text{ نریال } n$$

$$\mu = \left(\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \text{ مجهول } n$$

جامعه نریال نباشد

Sadra

کوچک $n < 30$ ، بزرگ $n > 30$

Subject:

Year:

Month:

Date:

مثال اداره بهداشت یک نفر ساحل در نظر دارد به مورد میزان باکتری در واحد حجم آب ساحل یک دریای تحقیق کند. برای این منظور ۱۰ نمونه از واحد حجم آب را آزمایش و میزان باکتری موجود در آب را به شرح زیر گزارش میکند

175 190 215 198 184 207 310 193 196 180

یک برگه در نقطه ای برای میانگین جامعه و یک فاصله اطمینان ۹۵٪ برای میانگین میزان

باکتری بدست آورید ، فرض کنید جامعه نرمال است .

جامعه نرمال

$n=10$

کوچک

شماره جدول

$$\mu: \left(\bar{X} \pm t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right)$$

$$\hat{\mu} = \bar{X} = \frac{175 + \dots + 180}{10} = 194,8 \quad , \quad S^2 = \frac{1}{9} [(175 - 194,8)^2 + \dots +$$

$$(180 - 194,8)^2] = (13,14)^2$$

$$t_{9, 0,025} = 2,262 \quad \leftarrow 1 - \alpha = 0,95 \Rightarrow \alpha = 0,05 \Rightarrow \frac{\alpha}{2} = 0,025$$

$$\mu: \left(194,8 \pm 2,262 \frac{13,14}{\sqrt{10}} \right)$$

Sadra

برآورد فاصله‌ای برای واریانس جامعه (سه): در محل معمولاً علاوه بر پارامترهای

پارامترهای نیز مورد نظر است. قبلاً دیدیم برآورد نقطه‌ای خوب برای σ^2 وقتی n

بجهول است عبارت است از تغییرات نمونه یعنی:

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

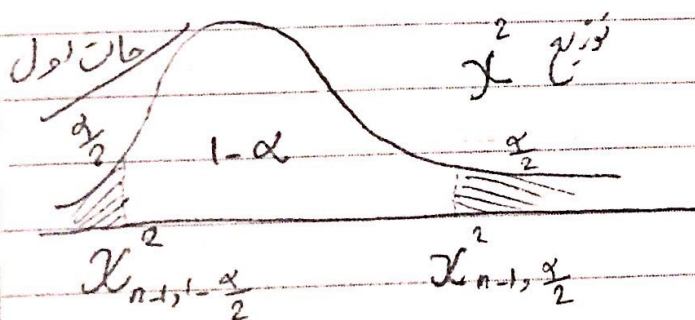
و وقتی μ معلوم است:

$$\hat{\sigma}^2 = S_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

همچنین دیدیم: (تحت فرض نرمال بودن $n > 30$) \rightarrow حالت اول

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2 \xrightarrow{\downarrow} \frac{nS_n^2}{\sigma^2} \sim \chi_n^2$$

حالت دوم $\leftarrow \mu$ معلوم



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

یا بگذاریم χ^2 :

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

Sadra

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

پس یک فاصله اطمینان $(1-\alpha)$ برای σ^2 وقتی μ مجهول است:

$$\sigma^2: \left(\frac{(n-1)S^2}{\chi^2_{n-1, \frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1, 1-\frac{\alpha}{2}}} \right) \rightarrow \text{مگر در فاصله اطمینان وقتی μ مجهول است.}$$

سه اول کپی به بعد بزرگ

نکته: فاصله اطمینان $(1-\alpha)$ برای انحراف معیار عبارت است

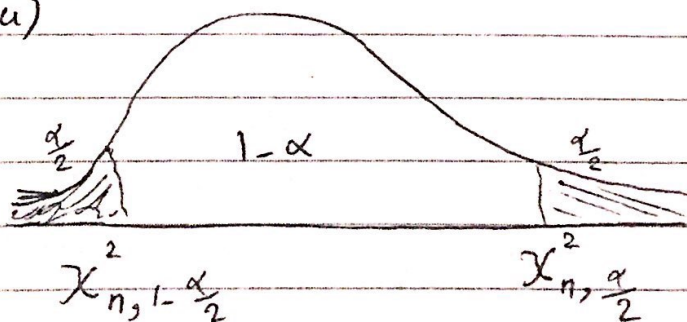
از جذر رابطه بالا:

$$\sigma: \left(\sqrt{\frac{(n-1)S^2}{\chi^2_{n-1, \frac{\alpha}{2}}}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{n-1, 1-\frac{\alpha}{2}}}} \right)$$

حالت دوم

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

$$\chi^2 = \frac{nS_n^2}{\sigma^2} \sim \chi_n^2$$



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

اما بگذاریم χ^2

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

۱۸

$$P\left(\underbrace{\frac{nS_n^2}{\chi_{n, \frac{\alpha}{2}}^2}}_L < \sigma^2 < \underbrace{\frac{nS_n^2}{\chi_{n, 1-\frac{\alpha}{2}}^2}}_U \right) = 1 - \alpha$$

پس یک فاصله اطمینان $1 - \alpha$ برای σ^2 وقتی σ نامعلوم است:

$$\sigma^2 = \left(\frac{nS_n^2}{\chi_{n, \frac{\alpha}{2}}^2}, \frac{nS_n^2}{\chi_{n, 1-\frac{\alpha}{2}}^2} \right) \rightarrow \text{بر آورد فاصله ای برای } \sigma^2 \text{ وقتی } \sigma \text{ نامعلوم}$$

نکته: فاصله اطمینان $1 - \alpha$ برای انحراف معیار وقتی σ نامعلوم است از جذر رابطه بالا:

$$\sigma: \left(\sqrt{\frac{nS_n^2}{\chi_{n, \frac{\alpha}{2}}^2}}, \sqrt{\frac{nS_n^2}{\chi_{n, 1-\frac{\alpha}{2}}^2}} \right)$$

مثال: ده نمونه زیر از جامعه ای در دسترس می باشد فاصله اطمینان 95٪ برای میانگین

جامعه واریانس جامعه بدست آورید: 230 225 229 227 228

(افترض مثال بدون) 232 225 226 228 226

- واریانس جامعه مجهول، حجم نمونه $n < 30$ ، جامعه نرمال:

$$\mu: \left(\bar{X} \pm t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \quad \begin{cases} 1 - \alpha = 95 \\ \alpha = 0.05 \rightarrow t_{9, 0.025} = 2.26 \\ \frac{\alpha}{2} = 0.025 \end{cases} \text{ جدول}$$

$$\bar{X} = \frac{230 + \dots + 226}{10} = 227.6$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{9} \left[(230 - 227.6)^2 + \dots + (226 - 227.6)^2 \right]$$

Sadra

$$= 5,26 \rightarrow S = \sqrt{5,26} = 2,27$$

$$\mu: (227,6 \pm 2,262 \frac{2,27}{\sqrt{10}}) = (226, 229,2)$$

برای بدست آوردن فاصله اطمینان برای σ^2 :

$$\sigma^2: \left(\frac{(n-1)S^2}{\chi^2_{n-1, \frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1, 1-\frac{\alpha}{2}}} \right) = \left(\frac{9 \times 5,26}{19,22}, \frac{9 \times 5,26}{2,7} \right)$$

محاسبه مقدار راز جدول بدست آمده

$$= (1,56, 4,14)$$

مثال: مقدار آهن موجود در یک کیلوگرم سنگ آهن متغیر تصادفی با توزیع نرمال است. اگر

برای یک نمونه 26 تایی داشته باشیم $S^2 = 93,24$ که فاصله اطمینان با ضریب اطمینان

$$n=26, S^2=93,24, \begin{cases} 1-\alpha=0,9 \\ \alpha=0,1 \\ \frac{\alpha}{2}=0,05 \end{cases} \quad 90\% \text{ برای } \sigma^2 \text{ بدست آورید.}$$

توزیع نرمال، σ^2 مجهول

$$\chi^2_{25, 0,05} = 37,65$$

$$\chi^2_{25, 0,95} = 14,6$$

$$\sigma^2: \left(\frac{25 \times 93,24}{37,65}, \frac{25 \times 93,24}{14,6} \right) = (61,91, 159,65)$$

فاصله اطمینان برای σ^2 جامعه: قبلاً دیدیم برآورد خوب برای σ^2 جامعه نسبت نمونه است:

$$\hat{p} = p = \frac{x}{n} \rightarrow \text{تعداد افرادی که در نمونه صفت خاص را دارند}$$

Sadra

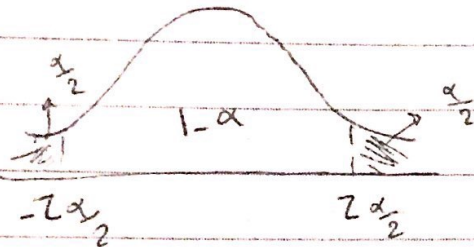
نمونه

از طرف دیگر اگر n بزرگ باشد:

$$\hat{p} \approx N(p, \frac{pq}{n})$$

$$E(\hat{p}) = p \quad \text{var}(\hat{p}) = \frac{pq}{n} \Rightarrow Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$$

توزیع



$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1 - \alpha$$

با جایگزینی Z :

$$P(-Z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} < Z_{\alpha/2}) = 1 - \alpha$$

$$\Rightarrow P\left(\underbrace{\hat{p} - Z_{\alpha/2} \sqrt{\frac{pq}{n}}}_L < p < \underbrace{\hat{p} + Z_{\alpha/2} \sqrt{\frac{pq}{n}}}_U\right) = 1 - \alpha$$

پس فاصله اطمینان $1 - \alpha$ برای p :

$$\left[p : \left(\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{pq}{n}} \right) \right]$$

چون $\text{var}(\hat{p}) = \frac{pq}{n}$ معمولاً مجهول است از برای برآورد آن در فرمول بالا استفاده

می شود.

$$\widehat{\text{var}(\hat{p})} = \frac{\hat{p}\hat{q}}{n}, \quad p : \left(\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right)$$

Sadra

مثال) از نیروی کار در شهر بزرگی نمونه تصادفی مرکب از دو هزار نفر انتخاب شده اند، افراد نمونه

مورد ~~مورد~~ مصاحبه قرار گرفتند و معلوم شده است که 165 نفر بیکارند: $n=2000$
 $x=165$

الف) نرخ بیکاری را بر بنای این داده ها برآورد کنید.

$$\hat{p} = \frac{165}{2000} = 0.0825$$

ب) فاصله اطمینان 95٪ برای نرخ بیکاری محاسبه کنید.

$$P: (\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}) = (0.0825 \pm 1.96 \sqrt{\frac{0.0825(1-0.0825)}{2000}})$$

$$(1-\alpha) = 0.95 \rightarrow \alpha = 0.05 \rightarrow \frac{\alpha}{2} = 0.025 \rightarrow Z_{0.025} = 1.96$$

مثال) برای برآورد کردن نسبت صاحبان اتومبیل که ماشین خود را به مبلغ بیش از یک مقدار معین بیمه کرده اند، نمونه‌ای تصادفی مرکب از 400 صاحب اتومبیل برگزیده شده است، اگر حق بیمه اتومبیل 56 نفر از این مقدار معین باشد یک فاصله اطمینان 95٪ برای نسبت این افراد بسازید.

ب) اگر نخواهیم 95٪ مطمئن باشیم که $d=0.008$ (خطا برآورد) چه حجم نمونه‌ای باید اختیار کنیم؟

$$\hat{p} = \frac{56}{400} = 0.14$$

$$P: \hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.14 \pm 1.96 \sqrt{\frac{0.14(1-0.14)}{400}})$$

$$= (0.106, 0.174)$$

پس 95٪ اطمینان داریم نسبت صاحبان اتومبیل که اتومبیل خود را بیش از مبلغ معین بیمه می‌کنند

$$n = \frac{\hat{p} \hat{q} Z_{\frac{\alpha}{2}}^2}{d^2} = \frac{(0.14)(1-0.14)(1.96)^2}{(0.008)^2} \approx 7228$$

تعریف آزمون فرض: هر وقت بخواهیم یک ادعا یا فرضیه‌ای را از روی داده‌های نمونه

بررسی کنیم مسئله آزمون فرض نامیده می‌شود. برای آزمون فرض بر هر مسئله اولین گام تشکیل

فرضیات است که به طور قراردادی به این صورت است که ادعا در فرض مقابل قرار می‌گیرد و

آنچه به تجربه بر مبنای مسئله وجود دارد در فرض صفر به صورت زیر می‌باشد: $H_0: \mu = 1200$ (فرض صفر)
 طول عمر لامپ‌های تولید شده کارخانه
 $H_1: \mu > 1200$ (فرض مقابل)
 مهندس فرآیند تولید پیشنهاد می‌دهد که طول عمر لامپ‌ها بزرگتر شود
 $H_1: \mu > 1200$ (فرض مقابل H_0)

پس از تشکیل فرضیه برای آزمون کردن نمونه‌ای تصادفی استخراج می‌شود، سطح آزمون

تعیین می‌کنیم. چون فرضیه‌ای که در H_0 قرار می‌دهیم معمولاً به تجربه برقرار بوده و اکنون

بر اساس نمونه H_0 رد شود به معنای این خطا، خطای بی است و همین معنای خطای

نوع اول نامیده می‌شود به عنوان سطح آزمون در نظر گرفته می‌شود و قبل از نتیجه گیری کنترل

می‌شود که با α نشان می‌دهند و معمولاً عدد کوچکی مثل $\frac{1}{100}$ ، $\frac{1}{10}$ ، $\frac{1}{5}$ ، $\frac{1}{2}$ ، ...

فرض می‌کنند پس احتمال رد H_0 (به شرطی که H_0 درست باشد) سطح آزمون می‌گویند:

$$\alpha = P(R | H_0) \quad \text{سطح آزمون}$$

reject

Sadra

Subject:

Year:

Month:

Date:

و در مرحله آخر فرض درست بودن H_0 و بر اساس نمونه استخراج شده و با توجه به مسئله آماره

آزمون "شکلی" می‌شود و تصمیم گیری انجام می‌شود.

خطای نوع دوم معمولاً با β نشان می‌دهیم و به صورت زیر تعریف می‌شود:

$$\beta = P(RH_0 | H_1 \text{ درست باشد})$$

$$1 - \beta = \text{توان آزمون}$$

Reject

Accept

RH_0 H_0 رد می‌شود	AH_0 H_0 پذیرفته می‌شود	تصمیم / وضعیت
خطای نوع اول	تصمیم صحیح است.	آر. H_0 درست باشد
$\alpha = \text{احتمال}$	$1 - \alpha = \text{احتمال}$	
تصمیم صحیح است.	خطای نوع دوم:	آر. H_0 غلط باشد
توان آزمون $1 - \beta = \text{احتمال}$	$\beta = \text{احتمال}$	

مقدار احتمال (p-value): یک روش کلیه یک برای آزمون کردن فرضیه‌ها محاسبه

مقدار احتمال است که به صورت زیر تعریف می‌شود:

کوچکترین مقداری که برای α می‌توان در نظر گرفت $p\text{-value} = \Rightarrow$ مقدار احتمال

تا H_0 رد شود، مثلاً فرض کنید T آماره آزمون تحت H_0 و مقدار مشاهده شده آن باشد $Sadad$

Subject:

Year:

Month:

Date:

$$\left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta > \theta_0 \end{array} \right.$$

$$P\text{-Value} = P(T_0 > t)$$

$$\left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta < \theta_0 \end{array} \right.$$

$$P\text{-Value} = P(T_0 < t)$$

برای تصمیم گیری:

$$\text{if } P\text{-Value} \leq \alpha \Rightarrow R H_0$$

$$\text{اگر فرضیه به صورت } \left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{array} \right. \text{ باشد } P\text{-Value عبارت است از:}$$

$$P\text{-Value} = 2 \min \left\{ P(T_0 > t), P(T_0 < t) \right\}$$

یا

$$P(|T_0| > t)$$

نکته: مصادیق H_0 قرار میگیرد.

آزمون فرض برای میانگین (μ) جامعه: برای میانگین جامعه به فرضیه زیر ممکن است مورد

$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{array} \right. \equiv \left\{ \begin{array}{l} H_0: \mu \leq \mu_0 \\ H_1: \mu > \mu_0 \end{array} \right.$$

تقریباً باشد.

فرضیه‌های یکطرفه

$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{array} \right. \equiv \left\{ \begin{array}{l} H_0: \mu \geq \mu_0 \\ H_1: \mu < \mu_0 \end{array} \right.$$

Sadra

* معادل $z_{\alpha/2}$ if $12.1 > z_{\alpha/2}$

Subject: _____
Year: _____ Month: _____ Date: _____

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases} \rightarrow \text{نفریه دوطرفه}$$

نمونه آزمون را به دو نظر منسیریم نمودن استخراج می شود x_1, \dots, x_n و آزمون

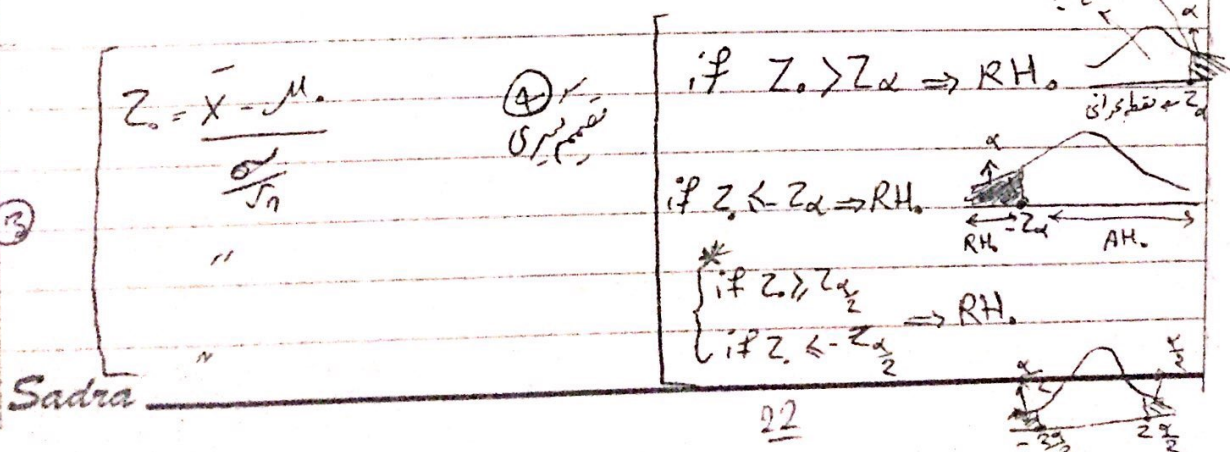
برای حالت های مختلف آماره های آزمون بدست آورده می شود

حالت اول: اگر جامعه نرمال باشد و پارامترش جامعه معلوم باشد، بدینیم:
ایا n بزرگ باشد و پارامترش معلوم

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \leftarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \leftarrow x_1, \dots, x_n \sim N(\mu, \sigma^2)$$

$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ آماره آزمون تحت H_0 می نویسیم پس از جواب آماره

① حالت اول نفریه دوطرفه	$H_0: \mu = \mu_0$	α	در مورد نفریه تقسیم می کنیم
	$H_1: \mu > \mu_0$		
	$H_1: \mu < \mu_0$		
② تقریباً یقیناً طرح آزمون	$H_0: \mu = \mu_0$	α	
	$H_1: \mu < \mu_0$		
	$H_1: \mu > \mu_0$		
③ تقریباً دفاعیه آماره	$H_0: \mu = \mu_0$	α	
	$H_1: \mu \neq \mu_0$		



Sadra

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Subject:

Year:

Month:

Date:

حالت دوم: n بزرگ باشد (حاصد نژاد یا غیر نژاد) و واریانس مجهول باشد در انصورت

قبلاً دیدیم: $Z = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim N(0, 1)$ $\xrightarrow[\text{آزمون}]{\text{ایجاد}}$ $Z_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$

①

②

③

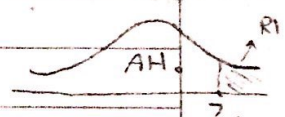
④

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$$

α

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$\text{if } Z_0 > Z_{\alpha} \Rightarrow RH_0$$

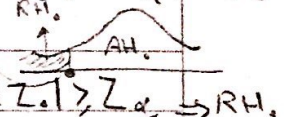


$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$$

α

"

$$\text{if } Z_0 < -Z_{\alpha} \Rightarrow RH_0$$



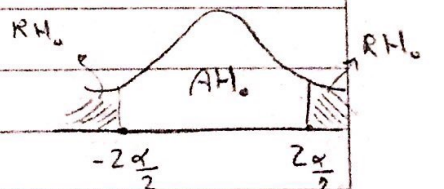
$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

α

"

$$\text{if } |Z_0| > Z_{\frac{\alpha}{2}} \Rightarrow RH_0$$

$$\begin{cases} Z_0 > Z_{\frac{\alpha}{2}} \\ Z_0 < -Z_{\frac{\alpha}{2}} \end{cases}$$



حالت سوم: وقتی حاصد نژاد است و n کوچک و واریانس مجهول:

براین نیز فرض کنید نمونه تصادفی X_1, \dots, X_n را اختیار است.

①

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

قبلاً دیدیم: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ میانگین نمونه

تغییرات نمونه: $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Subject:

Year:

Month:

Date:

$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{array} \right. , \left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{array} \right. , \left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{array} \right. \quad \text{انگشت فرضیه از}$$

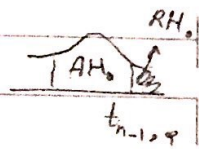
موردتقریبی. پس آماره آزمون با توجه به رابطه ①

$$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

انگشت برای فرضیه اول چون \bar{X} باید به اندازه کافی

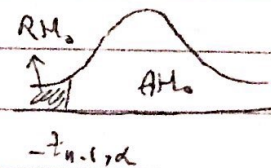
از μ_0 بزرگتر باشد تا H_0 را رد کنیم پس قاعده

$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{array} \right. \quad \alpha \quad T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad \text{if } T_0 > t_{n-1, \alpha} \Rightarrow R H_0$$



برای فرضیه دوم: برای این فرضیه باید \bar{X} به اندازه کافی از μ_0 کوچکتر باشد که H_0 را

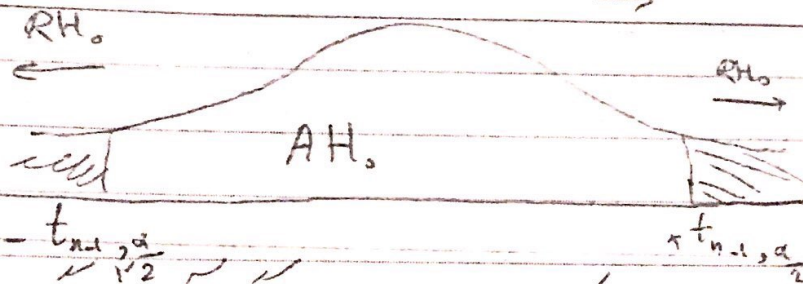
$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{array} \right. \quad \alpha \quad T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad \text{if } T_0 < -t_{n-1, \alpha} \Rightarrow R H_0$$



فرضیه سوم: برای این فرضیه باید \bar{X} به اندازه کافی از μ_0 کوچکتر یا بزرگتر باشد

$$\left\{ \begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{array} \right. \quad \alpha \quad T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad \text{if } |T_0| > t_{n-1, \frac{\alpha}{2}} \Rightarrow R H_0$$

Sadra



مثال) تعداد زیاری از بیماران قبلاً یک بیماری بخصوص را برآورده و گزارش کرده اند که مدت

درمان بیماری به روش استاندارد را در میانگین 15 روز و انحراف معیار 3 روز است. ادعا

شده که روش جدید میتواند مدت درمان را کوتاه تر کند. برای روش جدید درمان

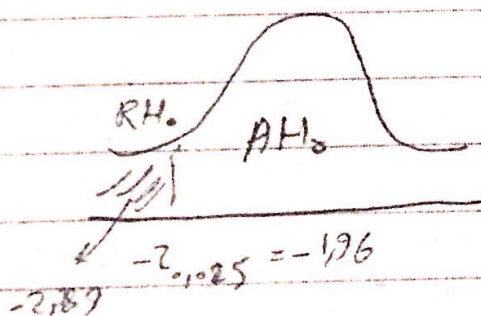
نمونه تصادفی به حجم 70 انتخاب شده است. پس از انجام روش جدید درمان بر روی

این 70 نفر میانگین مدت درمان 14 روز شده است. آیا در سطح معنی داری (سطح آزمون)

$$\mu = 15, \sigma = 3, n = 70, \bar{X} = 14 \quad \text{روش جدید بهتر است؟} \quad \alpha = 0.025$$

چشم نمونه برابر انحراف معیار معلوم

$$\begin{cases} H_0: \mu = 15 \\ H_1: \mu < 15 \end{cases} \quad \rightarrow \quad Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{14 - 15}{\frac{3}{\sqrt{70}}} = -2.789$$



$$-2.789 < -1.96 \Rightarrow RH.$$

مثال) نمونه تصادفی از پرونده های فوایدان شرکتی نشان میدهد که فشار شات برای فعلی

معنی از ماشین چاپ به ترتیب روزهای کاری شده است. اگر تعداد روزهای بایستی از توزیع نرمال

میوه‌های در سطح آزمون ۱۰/۰ می‌تواند ارعارد که میانگین زمان بایگانی از ۱۰/۵

بسته است $n=8$ 10 12 19 14 15 18 11 13

$$\begin{cases} H_0: \mu = 10.5 \\ H_1: \mu > 10.5 \end{cases}$$

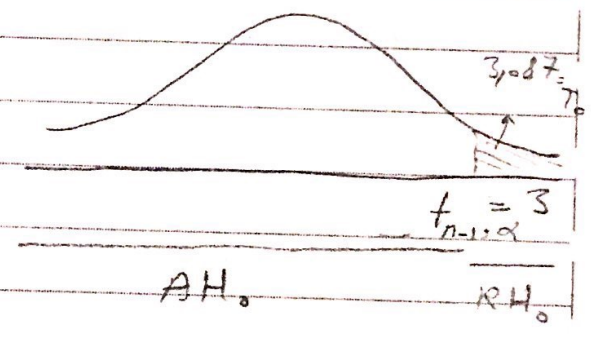
جامعه نرمال، واریانس جامعه مجهول

$$\bar{X} = \frac{10 + \dots + 13}{8} = 14, S^2 = \frac{1}{7} ((10-14)^2 + \dots + (13-14)^2) =$$

10.286

$$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{14 - 10.5}{\frac{\sqrt{10.286}}{\sqrt{8}}} = 3.087$$

$$\alpha = 0.01 \quad t_{n-1, \alpha} = t_{7, 0.01} = 3$$



$$if T_0 > t_{n-1, \alpha} \Rightarrow RH_0$$

$3.087 > 3 \Rightarrow RH_0$ یعنی میانگین نرمال بایگانی قطعات بیشتر از ۱۰/۵ روز است.

مثال) نمونه شامل ۲۰ رسید فروش کن ب یک فروشگاهت ب درل میانگین ۱۲۱ دلار

و انحراف معیار ۱۰/۲ دلار است. از این داده‌ها برای آزمون این فرض که میانگین فروش

کن ب کمتر از ۱۲۵ دلار است استفاده کنید. سطح آزمون را ۰/۰۵ در نظر بگیرید. فرض

$n=20 < 30$, $\bar{x} = 121$, $S = 10.2$ کنید جامعه نرمال است.

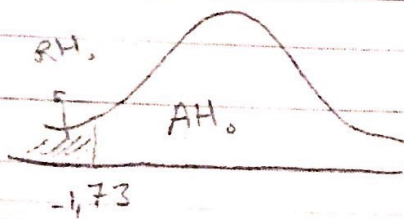
Sadra $T_0 = \frac{121 - 125}{\frac{10.2}{\sqrt{20}}} = -1.75$

Subject:

Year:

Month:

Date:



$$-t_{n-1, 0.05} = -t_{19, 0.05} = -1.73$$

$$-1.75 < -1.73 \Rightarrow RH.$$

مثال: از یک نمونه تصادفی شامل 36 مشاهده نتایج $\bar{X} = 80.4$ و $S = 16.2$ برآورد

آزمون فرضیه های زیر را با استفاده از روش آزمون (تست) و از روش P-value

$$H_0: \mu = 74$$

$$H_1: \mu > 74$$

$$\alpha = 0.05$$

$$H_0: \mu = 76$$

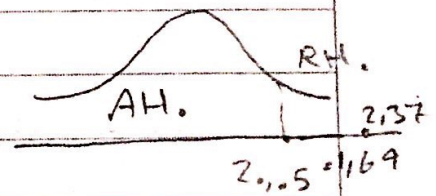
$$H_1: \mu \neq 76$$

الان

$$n = 36 > 30, \bar{X} = 80.4, S = 16.2$$

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{80.4 - 74}{\frac{16.2}{\sqrt{36}}} = 2.37$$

$$|Z_0| = 2.37 > Z_{\alpha} = Z_{0.05} = 1.64 \Rightarrow RH.$$



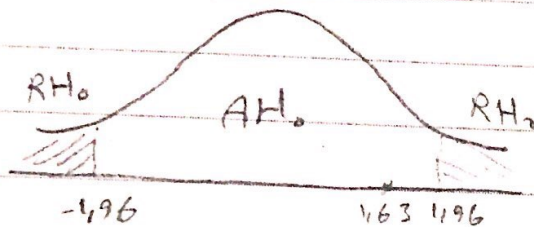
$$p(Z \leq z) = \Phi(z), p\left(\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \geq 2.37\right) = 1 - \Phi(2.37) = 0.0089$$

$$\text{if } p\text{-value} < \alpha \Rightarrow RH.$$

$$0.0089 < 0.05 \Rightarrow RH.$$

$$Z_0 = \frac{80,4 - 76}{\frac{16,2}{\sqrt{36}}} = 1,63$$

$$\begin{aligned} & \text{if } Z_0 > Z_{\frac{\alpha}{2}} \\ & \text{or} \\ & Z < -Z_{\frac{\alpha}{2}} \Rightarrow R H_0 \end{aligned}$$



$$1,63 > 1,96 \Rightarrow A H_0$$

$$\rightarrow 2 \min \{P(Z_0 > z_0), P(Z_0 < -z_0)\}$$

$$P\text{-value} = P(|Z_0| > 1,63) = P(Z_0 > 1,63) + P(Z_0 < -1,63)$$

$$= (1 - \Phi(1,63)) + (\Phi(-1,63)) = 0,1032 < 0,05 \Rightarrow A H_0$$

آن‌ها فرض برای نیت در جامعه: فرض نیت هدف مطالعه نیت در جامعه است و در خصوص

نیت جامعه ارجایی شده است که مطالعه کل جامعه امکان پذیر نیست و از روی نمونه خواهم ارجا

لا بررسی کنیم به روش زیر ممکن است مورد نظر باشد:

$$\begin{cases} H_0: p = p_0 \\ H_1: p > p_0 \end{cases}$$

$$\begin{cases} H_0: p = p_0 \\ H_1: p < p_0 \end{cases}$$

$$\begin{cases} H_0: p = p_0 \\ H_1: p \neq p_0 \end{cases}$$

برای این مقوله نمونه را استخراج می‌شود و مقایسه می‌شود:

$$\hat{p} = p \stackrel{\text{نمونه}}{\sim} N(p, \frac{p q_0}{n})$$

به حق: H_0

$$\hat{p} \sim N(p_0, \frac{p_0 q_0}{n})$$

Sadra

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \sim N(0,1) \rightarrow \text{آزمون پارامتر}$$

پس داریم:

فرضیه آزمون	سطح آزمون	آماره	تقریب
$H_0: p = p_0$	α	$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$	if $Z_0 > Z_{\alpha} \Rightarrow RH_0$
$H_1: p > p_0$	α	"	if $Z_0 < -Z_{\alpha} \Rightarrow RH_0$
$H_0: p = p_0$	α	"	if $Z_0 > Z_{\alpha/2}$ or $Z_0 < -Z_{\alpha/2} \Rightarrow RH_0$
$H_1: p \neq p_0$	α	"	"

مثال: بررسی یک سرشماری در 5 سال پیش. 20% خانواده‌های یک شهر زیر خط فقر

زندگی می‌کردند. برای معلوم شدن این موضوع که آیا این درصد تغییری کرده است یا خیر نمونه‌ای

تصادفی به حجم 400 خانواده مطالعه شد و معلوم شد 70 نفر زیر خط فقرند. آیا این یافته‌ها

دلالت بر این دارد که درصد خانواده‌های زیر خط فقر با 5 سال پیش فرق کرده است؟ این

$p = 0.2$ فرض H_0 داشته باشیم

آزمون دایره‌ای و از روش مستقیم و از روش مقدار احتمال انجام دهید

$$\begin{cases} H_0: p = 0.2 \\ H_1: p \neq 0.2 \end{cases} \quad n = 400 \quad \alpha = 0.05$$

$X = 70 \rightarrow \hat{p} = \frac{70}{400} = 0.175$

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.175 - 0.2}{\sqrt{\frac{0.2 \times 0.8}{400}}} = -1.25$$

Sadra $\sqrt{\frac{p_0 q_0}{n}}$

Subject:

Year:

Month:

Date:

تدریس برای آزمون دو طرفه

$$\alpha = 0.025 \rightarrow Z_{0.025} = 1.96$$

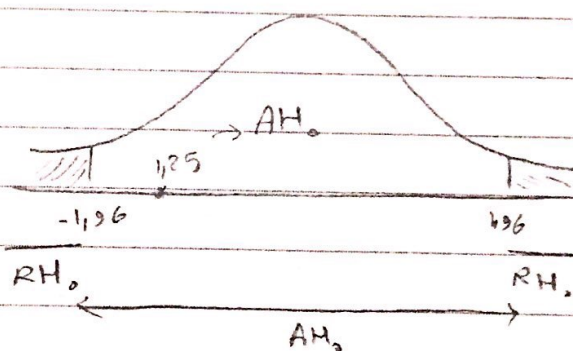
Conclusion \rightarrow if $|Z_0| > Z_{\alpha/2} = 1.96$

$$\Rightarrow 1.25 > 1.96 \Rightarrow A H_0$$

$$p\text{-value} = P(|Z_0| > 1.25) = P(Z_0 > 1.25) + P(Z_0 < -1.25)$$

$$= 1 - P(Z_0 < 1.25) + P(Z_0 < -1.25) = 1 - \Phi(1.25) + \Phi(-1.25) = 0.21$$

$$p\text{-value} < 0.05 \Rightarrow A H_0$$

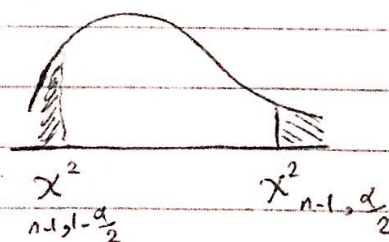


آزمون برای واریانس جامعه: در این روش برای واریانس جامعه و وقت

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 \quad \hat{\sigma}^2 = S^2 \quad \text{این عبارت است از:}$$

از طرفی در این روش جامعه نرمال است و n بزرگ:

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



برای آزمون فرض: فرضهای مختلف داریم:

Sadra

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \rightarrow \text{مرسوم}$$

Subject:

Year:

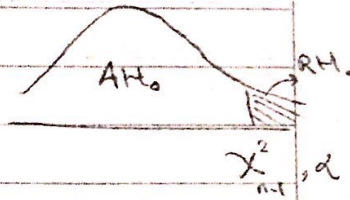
Month:

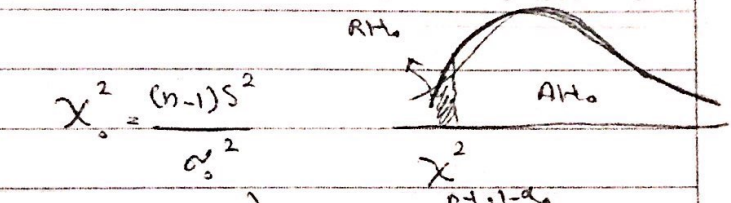
Date:

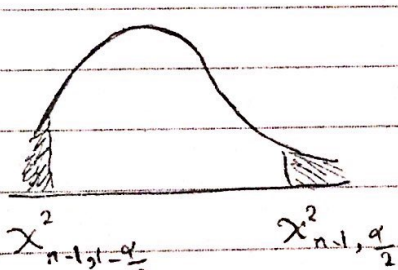
$$S_n^2 = \frac{1}{n} \sum (x_i - \mu)^2 \rightarrow \text{مرسوم}$$

نتیجه گیری: σ^2 (نظریه) و σ_0^2 (طرح از مومن)

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \end{cases} \propto \chi^2_0 = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{if } \chi^2_0 > \chi^2_{n-1, \alpha} \Rightarrow \text{RH.}$$

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 < \sigma_0^2 \end{cases} \propto \chi^2_0 = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{if } \chi^2_0 < \chi^2_{n-1, 1-\alpha}$$


$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases} \propto \chi^2_0 = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{if } \chi^2_0 < \chi^2_{n-1, 1-\alpha/2} \text{ or } \chi^2_0 > \chi^2_{n-1, \alpha/2}$$


$$\text{if } \chi^2_0 > \chi^2_{n-1, \frac{\alpha}{2}} \text{ or } \chi^2_0 < \chi^2_{n-1, \frac{\alpha}{2}}$$


نکته: اگر مرسوم باشد می توانیم برآورد خوب برای σ^2 و S_n^2 است یعنی:

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2, \quad \frac{nS_n^2}{\sigma^2} \sim \chi^2_n$$

در روابط بالا باید از معادسی آزمون استفاده کنید و از توزیع χ^2 دو باید به کار ببرید.

$$\chi^2_0 = \frac{nS_n^2}{\sigma_0^2} \quad \text{است چایی } \chi^2_{n-1} \text{ چایی (n-1) } n \text{ بزرگم چون مرسوم}$$

Sadra

مثال ۱: فرض کنید برای نمونه تصادفی به حجم $n = 25$ از توزیع نرمال، فاصله اطمینان

برای میانگین جامعه به صورت $(106,8, 115,2)$ باشد. آیا این نتیجه استقاده شده است؟

الف) مقدار انحراف بسیار قوی را بدست آورید.

هت یا نه؟

ب) در سطح آزمون $\alpha = 0,05$ فرض انحراف بسیار جاد به برابر 3 ~~در سطح 3~~

الف $n = 25$

$\mu = (106,8, 115,2)$

$$\begin{cases} \mu_0 = 106,8 \\ \mu_1 = 115,2 \end{cases}$$

$$\begin{cases} H_0: \mu = 106,8 \\ H_1: \mu = 115,2 \end{cases}$$

جامعه نرمال

حجم نمونه کوچک

در بایوس مجهول

آزمون یک طرفه

$$\mu = \left(\bar{X} \pm t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right) \sim \begin{cases} \bar{X} - t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} = 106,8 \\ \bar{X} + t_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}} = 115,2 \end{cases}$$

$$t_{24, 0,025} = 2,064 \sim \begin{cases} \bar{X} - 2,064 \frac{S}{5} = 106,8 \\ \bar{X} + 2,064 \frac{S}{5} = 115,2 \end{cases}$$

$$2\bar{X} = 222 \Rightarrow \bar{X} = 111, \quad \hat{\mu} = 111 \Rightarrow 111 + 2,064 \frac{S}{5} = 115,2$$

$$\Rightarrow \frac{2,064 S}{5} = 4,2 \Rightarrow 2,064 S = 21 \Rightarrow S = \frac{21}{2,064} = 10,17$$

$$\chi^2 = \frac{(n-1) S^2}{\sigma_0^2} \xrightarrow{\text{انحراف جاد}} \sqrt{\chi^2} = \frac{\sqrt{(n-1)} S}{\sigma_0} = \frac{\sqrt{24} \times 10,17}{3} = 16,61$$



$$3,62 \leftarrow \sqrt{12,4} = \chi_{24, 0,025}^2 \quad \chi_{24, 0,025}^2$$

Subject:

Year:

Month:

Date:

صحت با این رسم

استباط پیرامون دو جامعه

در این قسمت روشهای استباط آماری برای مقایسه دو جامعه بر مبنای نمونه های

مستقل را مورد بحث قرار میدهم و مت گنید شما باید دو جامعه داشته باشیم ماهی

مستقل دو بیمار (روشن درمان) مثل دو نوع لاستیک یا نوع دارو مقایسه میشوند

مثلاً فرض کنید بیمار برای مطالعه ای اثر نوع دارو در اختیار یک محقق قرار میدهند

هدف مقایسه اثر دارو شماره یک با اثر دارو شماره 2 بر مردان بیادان است

ابتداءً طور تصادفی بیمار به دو گروه تقسیم میشوند و به گروه 1 دارو شماره 1 و به گروه 2

دارو شماره 2 داده میشود مشاهدات حاصل از دارو شماره یک هیچ ارتباطی با مشاهدات

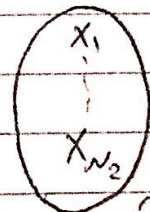
حاصل از دارو شماره 2 ندارد و تقسیم بیمارها به دو گروه کاملاً تصادفی انجام شده است

پس در این بحث خواهیم دو بیمار یا دو جامعه را بر مبنای نمونه های مستقل مقایسه کنیم

برای استباط پیرامون دو جامعه با فرض اینکه به هر دلیلی

سرشاری امکان پذیر نیست ابتدا دو نمونه از هر جامعه استخراج

میکشیم استخراج دو نمونه کاملاً مستقل است



جمع n_2 جامعه 2

جمع n_1 جامعه 1



نمونه 2

جمع n_1

جمع n_2

x_1, \dots, x_{n_1} از جامعه اول

مجموعه دو از هم مستقل

از هم مستقل

y_1, \dots, y_{n_1} از جامعه دوم

مجموعه دو از هم مستقل

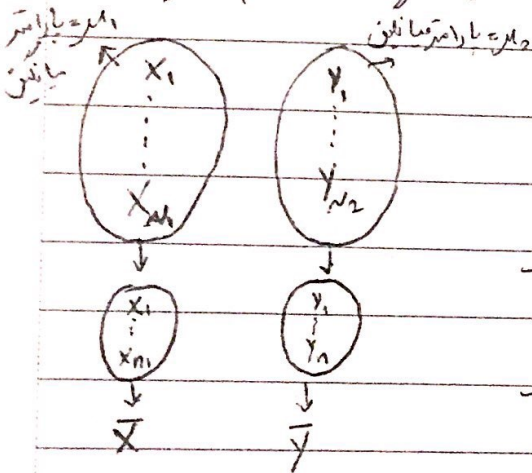
Sadra

1. مشابه استنباط پیرامون یک جامعه می‌خواهیم از روی نمونه‌های تصادفی در جامعه را با هم

2. مقایسه کنیم و پارامترهای مهم میانگین و واریانس نسبت به دو رانندگی را بررسی قرار می‌دهیم.

3. مقایسه میانگین دو جامعه $(\mu_1 - \mu_2)$ یا $(\mu_2 - \mu_1)$ (تفاوت در رانندگی): ابتدا نمونه

4. مستقل از دو جامعه به ترتیب با میانگین μ_1 و μ_2 استخراج می‌شود و این میانگین نمونه



5. برای درستی برای میانگین‌های جامعه است.

$$\hat{\mu}_1 = \bar{X} \rightarrow E(\bar{X}) = \mu_1$$

$$\hat{\mu}_2 = \bar{Y} \rightarrow E(\bar{Y}) = \mu_2$$

6. از خواص امید ریاضی:

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

7. پس $\bar{Y} - \bar{X}$ یک برآورد کننده ناآریب برای $\mu_2 - \mu_1$ است.

8. پس یک برآورد کننده برای $\mu_2 - \mu_1$ عبارت است از $\bar{Y} - \bar{X}$.

$$\hat{\mu}_2 - \mu_1 = \bar{Y} - \bar{X} \rightarrow \text{"برآورد نقطه‌ای"}$$

9. برآورد فاصله‌ای و فاصله اطمینان برای $\mu_2 - \mu_1$:

10. برای بدست آوردن برآورد فاصله‌ای حالت‌های مختلف را بررسی می‌کنیم:

11. حالت اول: نمونه‌ها به اندازه کافی بزرگ باشند و پارامترهای جامعه نرمال باشد و

$$V(ax+by) = a^2 V(x) + b^2 V(y) - 2ab \text{cov}(x,y)$$

Year: Month: Date:



Sa Su Mo Tu We Th Fr

در این حالت معلوم می‌باشد که نرمال باشد واریانس معلوم.

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

چون از هم مستقل اند

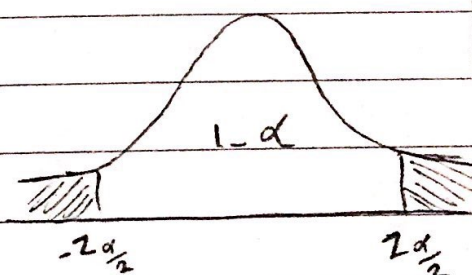
$$V(\bar{Y} - \bar{X}) = V(\bar{Y}) + V(\bar{X}) = \frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}$$

بنابراین در این حالت چون نمونه به اندازه کافی بزرگ هستند (یا جامعه نرمال):

$$\begin{aligned} \bar{X} &\sim N(\mu_1, \frac{\sigma_1^2}{n_1}) \\ \bar{Y} &\sim N(\mu_2, \frac{\sigma_2^2}{n_2}) \end{aligned} \quad \xrightarrow{\text{طبق خواص نرمال}} \quad \bar{Y} - \bar{X} \sim N\left(\mu_2 - \mu_1, \frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}\right)$$

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

فاصله اطمینان با ضریب اطمینان $1 - \alpha$ مورد نظر



$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

با جایگزینی Z :

$$P\left(-z_{\alpha/2} < \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}} < z_{\alpha/2}\right) = 1 - \alpha$$

با توجه داشتن $\mu_2 - \mu_1$ در صورت انتقال سایر جمله‌ها به طرفین داریم:

(2)

Year: Month: Date:



Sa Su Mo Tu We Th Fr

$$P((\bar{Y} - \bar{X}) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}} < \mu_2 - \mu_1 < (\bar{Y} - \bar{X}) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}})$$

با این فاصله می توان $\mu_2 - \mu_1$ را با $(1-\alpha) \%$ اطمینان و n_1 و n_2 بزرگ

باشند و این ها معلوم نیست زیرا است:

$$\mu_2 - \mu_1 = \left((\bar{Y} - \bar{X}) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}} \right)$$

نکته: فقط در حالتی که n_1 و n_2 هر دو بزرگ هستند و این ها محمول بر آن صورت می توان

در رابطه بالا به جای σ_1^2 و σ_2^2 از S_1^2 و S_2^2 استفاده کرد:

$$\mu_2 - \mu_1 = \left((\bar{Y} - \bar{X}) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{S_2^2}{n_2} + \frac{S_1^2}{n_1}} \right)$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \quad \text{و} \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

مثال: برای مقایسه سن ازدواج زنان در دو ایل A و B از هر ایل نمونه ای شامل 100

زن استخراج شده است و سن ازدواج آنها ثبت کرده ایم و خلاصه آماره ها به شرح زیر بدست

A B $\mu_B - \mu_A =$

حجم نمونه $n_1 = 100$ $n_2 = 100$ آمده است. $\bar{y} - \bar{x} = 20.7 - 18.5 = 2.2$ بر آورد نقطه ای

میانگین $\bar{x} = 18.5$ $\bar{y} = 20.7$ سن ازدواج در ایل A کمتر از ایل B است.

واریانس $S_A = 6.2$ $S_B = 5.8$

بر آورد فاصله ای: $\alpha = 0.05$ $1 - \alpha = 0.95$ و $Z_{0.975} = 1.96$



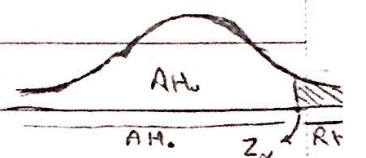
$$\mu_B - \mu_A = (\bar{y} - \bar{x}) \pm 1.96 \sqrt{\frac{S_B^2}{n_2} + \frac{S_A^2}{n_1}}$$

$$= (12.7 - 18.5) \pm 1.96 \sqrt{\frac{(5.8)^2}{100} + \frac{(6.2)^2}{100}}$$

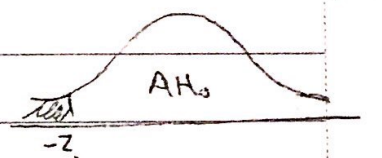
چون کران پایین و کران بالا $\rightarrow (3.88, 0.52)$
فاصله اطمینان مثبت است پس با اطمینان 95٪ سن ازدواج زنان B از ایل A بیشتر است

اگر زمون فرض در این حالت: (n_1, n_2) برابرند و واریانس ها مجهول (به فرضیه مورد نظر است) بیشتر فرضیه
تست آماری (فرضیه تقسیم) آماره آزمون سطح آزمون شکل فرضیه

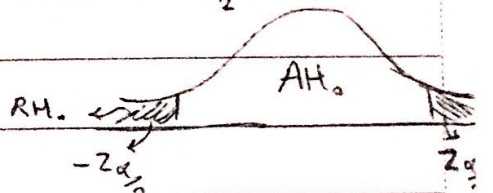
$$\begin{cases} H_0: \mu_2 - \mu_1 = S_0 \\ H_1: \mu_2 - \mu_1 > S_0 \end{cases} \quad \alpha \quad Z_0 = \frac{(\bar{y} - \bar{x}) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}} \quad \text{if } Z_0 > Z_{\alpha} \Rightarrow RH.$$



$$\begin{cases} H_0: \mu_2 - \mu_1 = S_0 \\ H_1: \mu_2 - \mu_1 < S_0 \end{cases} \quad \alpha \quad // \quad \text{if } Z_0 < -Z_{\alpha} \Rightarrow RH.$$



$$\begin{cases} H_0: \mu_2 - \mu_1 = S_0 \\ H_1: \mu_2 - \mu_1 \neq S_0 \end{cases} \quad \alpha \quad // \quad \text{if } |Z_0| > Z_{\alpha/2} \Rightarrow RH. *$$



$$\text{if } Z_0 > Z_{\alpha/2} \text{ or } Z_0 < -Z_{\alpha/2} \Rightarrow RH.$$

مثال: آیا مثال قبل خواهد قوی برای قبول این ادعا که متوسط سن ازدواج در وایل متفاوت

$$\begin{cases} H_0: \mu_B - \mu_A = 0 \\ H_1: \mu_B - \mu_A \neq 0 \end{cases}$$

است وجود دارد؟ سطح آزمون را 2٪ در نظر بگیرید.

3

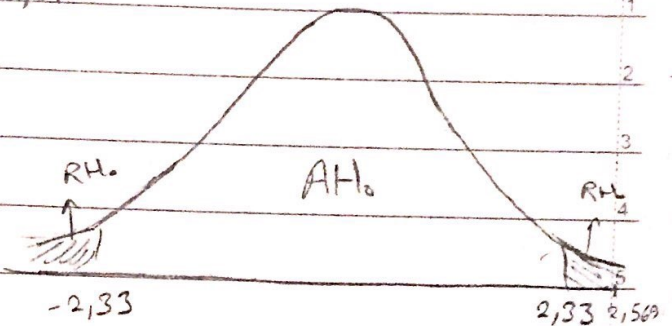
Year: Month: Date:



Sa Su Mo Tu We Th Fr

$$\alpha = 0.02 \rightarrow \alpha/2 = 0.01$$

$$Z_0 = \frac{(20.7 - 18.5) - (0)}{\sqrt{\frac{(5.8)^2}{100} + \frac{(6.3)^2}{100}}} = 2.569$$



$$Z_{0.01} = 2.33$$

$$2.569 > 2.33 \Rightarrow RH_1$$

تفسیر: چون H_0 رد می شود به قدری که از دواج دوایلی متفاوت است.

حالت دوم: اگر n_1 یا n_2 کوچک باشند و جوامع نرمال و پارامترهای آنها مجهول، در این حالت

دو زیر حالت در نظر می گیریم ① زیر حالت اول: دو جامعه از نظر پراکندگی یکسان هستند

② زیر حالت دوم: دو جامعه از نظر پراکندگی یکسان نیستند.

برای زیر حالت اول داریم:

$$X_1, \dots, X_n \rightarrow \bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}\right)$$

نمونه های تصادفی از جامعه نرمال اول

$$Y_1, \dots, Y_n \rightarrow \bar{Y} \sim N\left(\mu_2, \frac{\sigma_1^2}{n_2} + \frac{\sigma_2^2}{n_2}\right)$$

نمونه های تصادفی از جامعه نرمال دوم

$$\frac{(n_1 - 1) S_1^2}{\sigma^2} \sim \chi_{n_1 - 1}^2 \quad S_1^2 = \frac{1}{n_1 - 1} \sum (x_i - \bar{x})^2$$

$$\frac{(n_2 - 1) S_2^2}{\sigma^2} \sim \chi_{n_2 - 1}^2 \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2$$

TAHERIAN

جمع دو توزیع گامی دوگانه (نقشه آزمون) با هم جمع می‌شود

Year: Month: Date:



Sa Su Mo Tu We Th Fr

چون $\sigma_1^2 = \sigma_2^2 = \sigma^2$ پس یک برآورد ترکیبی از دو نمونه به صورت زیر برای آن می‌توان

برای آورد $E(S_1^2) = \sigma_1^2 = \sigma^2$

$$E(S_2^2) = \sigma_2^2 = \sigma^2 \quad \left\{ \rightarrow S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \right.$$

$$E(S_p^2) = \frac{(n_1-1)\sigma^2 + (n_2-1)\sigma^2}{n_1+n_2-2} = \frac{\sigma^2(n_1+n_2-2)}{n_1+n_2-2} = \sigma^2$$

از طرفی طبق خواص توزیع نرمال:

$$\bar{y} - \bar{x} \sim N(\mu_2 - \mu_1, \sigma^2 (\frac{1}{n_1} + \frac{1}{n_2}))$$

$$E(\bar{y} - \bar{x}) = \mu_2 - \mu_1$$

$$V(\bar{y} - \bar{x}) = V(\bar{y}) + V(\bar{x}) - 2\text{Cov}(\bar{y}, \bar{x}) = \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1} = \sigma^2 (\frac{1}{n_1} + \frac{1}{n_2})$$

با استاندارد کردن:

$$Z = \frac{(\bar{y} - \bar{x}) - (\mu_2 - \mu_1)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0, 1) \quad (1)$$

همچنین چون:

$$\frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi_{n_2-1}^2, \quad \frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi_{n_1-1}^2$$

بنابراین:

$$\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{\sigma^2} \sim \chi_{(n_2-1)+(n_1-1)}^2 \quad (2)$$

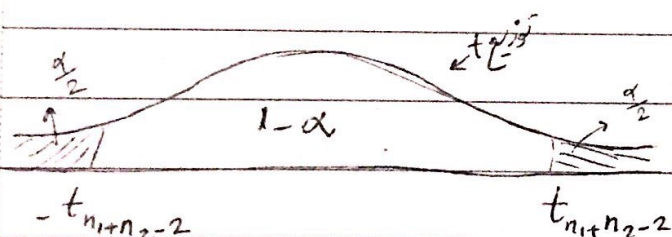
آنون یک توزیع نرمال استاندارد (1) و یک توزیع گامی دو (2) داریم و می‌دانیم:

$$(\bar{y} - \bar{x}) - (\mu_2 - \mu_1)$$

$$T = \frac{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}} = t_{n_1+n_2-2}$$

$$\Rightarrow T = \frac{(\bar{y} - \bar{x}) - (\mu_2 - \mu_1)}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2} \quad \star$$

برای بدست آوردن فاصله اطمینان $100(1-\alpha)\%$ برای $\mu_2 - \mu_1$ داریم:



$$P(-t_{n_1+n_2-2, \alpha/2} < T < t_{n_1+n_2-2, \alpha/2}) = 1 - \alpha$$

$$\Rightarrow P(-t_{n_1+n_2-2, \alpha/2} < \frac{(\bar{y} - \bar{x}) - (\mu_2 - \mu_1)}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{n_1+n_2-2, \alpha/2}) = 1 - \alpha$$

$$\Rightarrow P((\bar{y} - \bar{x}) - t_{n_1+n_2-2, \alpha/2} S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_2 - \mu_1 < (\bar{y} - \bar{x}) + t_{n_1+n_2-2, \alpha/2} S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}) = 1 - \alpha$$

بن فاصله اطمینان $100(1-\alpha)\%$ برای حالتی که: جوامع زیر نرمالند، σ_1^2 و σ_2^2 مجهول و یکی برابر

$$\mu_2 - \mu_1 = (\bar{y} - \bar{x}) + t_{n_1+n_2-2, \frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

برای آزمون نقیه های زیر هم از آمار سی (*) استفاده می کنیم: نقیه ها

فرضیه صفر فرضیه بیهوده ازجه آمار آزمون سطح آزمون

$$H_0: \mu_2 - \mu_1 = \delta_0$$

$$H_1: \mu_2 - \mu_1 > \delta_0$$

$$T_0 = \frac{(\bar{y} - \bar{x}) - \delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\text{if } T_0 > t_{n_1+n_2-2, \alpha}$$

$\Rightarrow RH_0$

$$H_0: \mu_2 - \mu_1 = \delta_0$$

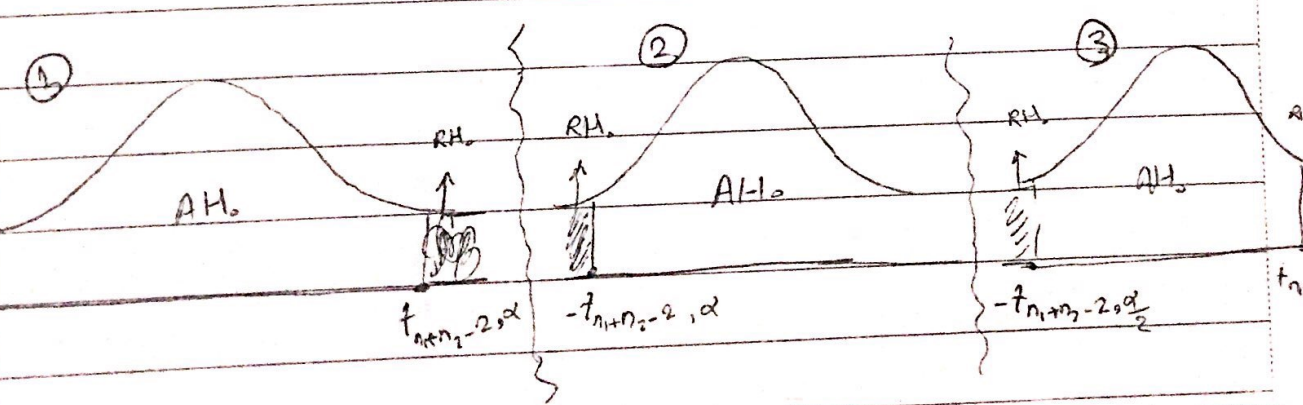
$$H_1: \mu_2 - \mu_1 < \delta_0$$

$$\text{if } T_0 < -t_{n_1+n_2-2, \alpha} \Rightarrow RH_0$$

$$H_0: \mu_2 - \mu_1 = \delta_0$$

$$H_1: \mu_2 - \mu_1 \neq \delta_0$$

$$\text{if } |T_0| < t_{n_1+n_2-2, \frac{\alpha}{2}} \Rightarrow R$$



برای زیر حالت دوم داریم:

جوامع نرمال بودند که کوچک و درایانش ها مجهول و نا برابر:

5

Year: Month: Date:



Sa Su Mo Tu We Th Fr

$$\sigma_1^2 \neq \sigma_2^2$$

در این زیر حالت از آماره‌ی زیر استفاده می‌شود:

$$T^* = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \approx t_{\min(n_1-1, n_2-1), \frac{\alpha}{2}}^*$$

یا $d.f.$ را به صورت زیر بدست می‌آوریم در متن کتاب:

$$d.f. = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}$$

برای آزمون کرین نرفیه‌های بالا از آماره آزمون:

$$T_0^* = \frac{(\bar{Y} - \bar{X}) - S_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (**)$$

استفاده می‌کنیم و حاصل اطمینان را در این حالت:

$$\mu_2 - \mu_1 : (\bar{Y} - \bar{X}) + t_{d.f., \frac{\alpha}{2}}^* \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

برای آزمون نرفیه‌های زیر هم از آماره‌ی T_0^* استفاده می‌کنیم:



فرضیه ها

از آمار آماری

تأثیر و مقایسه

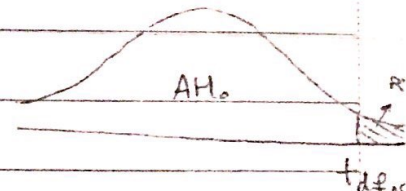
$$H_0: \mu_2 - \mu_1 = S_0$$

$$H_1: \mu_2 - \mu_1 > S_0$$

$$H_2: \mu_2 - \mu_1 < S_0$$

$$T_0 = \frac{(\bar{y} - \bar{x}) - S_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$\text{if } T_0 > t_{df, \alpha}^* \Rightarrow RH_0$$



$$H_0: \mu_2 - \mu_1 = S_0$$

$$H_1: \mu_2 - \mu_1 > S_0$$

$$H_2: \mu_2 - \mu_1 < S_0$$

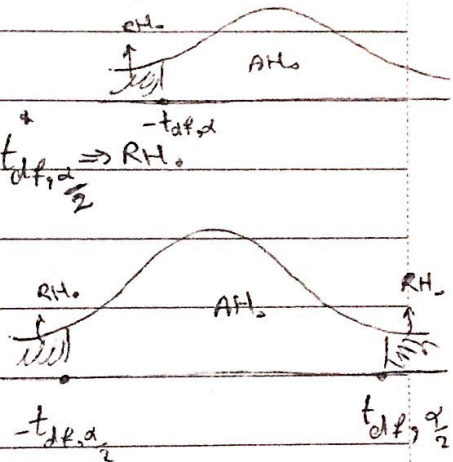
$$\text{if } T_0 < -t_{df, \alpha}^* \Rightarrow RH_0$$

$$H_0: \mu_2 - \mu_1 = S_0$$

$$H_1: \mu_2 - \mu_1 \neq S_0$$

$$H_2: \mu_2 - \mu_1 < S_0$$

$$\text{if } |T_0| > t_{df, \alpha/2}^* \Rightarrow RH_0$$



مقایسه واریانس های دو جامعه: هدف مقایسه σ_1^2 و σ_2^2 یعنی واریانس های دو جامعه

است. قبلاً دیدیم، (فرض نکریم) بودن دو جامعه n_1 و n_2 بزرگ:

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} \sim \chi_{n_1 - 1}^2$$

$$\Rightarrow F = \frac{\left(\frac{(n_1 - 1)S_1^2}{\sigma_1^2} \right) / (n_1 - 1)}{\left(\frac{(n_2 - 1)S_2^2}{\sigma_2^2} \right) / (n_2 - 1)}$$

$$\frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi_{n_2 - 1}^2$$

6

(توزیع F دو هردوشت و مقدره طاق و مسا ندارد)

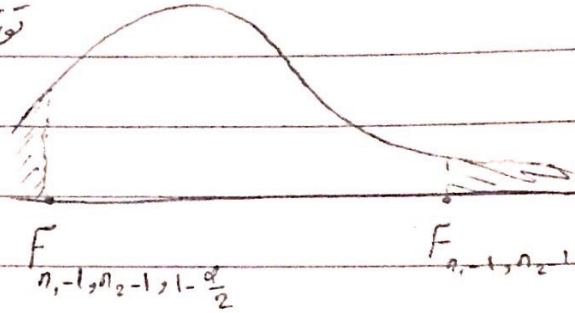
Year: Month: Date:



Sa Su Mo Tu We Th Fr

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1-1, n_2-1}$$

توزیع



$$P\left(F_{(n_1-1, n_2-1), 1-\frac{\alpha}{2}} < \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} < F_{(n_1-1, n_2-1), \frac{\alpha}{2}}\right) = 1 - \alpha$$

بنابراین فاصله اطمینان $1 - \alpha$ برای نسبت وابستگی ها:

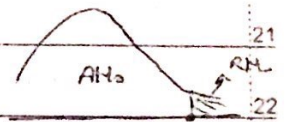
$$\frac{\sigma_2^2}{\sigma_1^2} : \left(\frac{S_2^2}{S_1^2} F_{(n_1-1, n_2-1), 1-\frac{\alpha}{2}}, \frac{S_2^2}{S_1^2} F_{(n_1-1, n_2-1), \frac{\alpha}{2}} \right)$$

$$F_{(n_1-1, n_2-1), 1-\frac{\alpha}{2}} = \frac{1}{F_{(n_2-1, n_1-1), \frac{\alpha}{2}}}$$

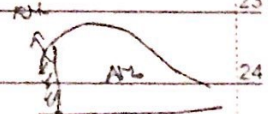
برای آزمون رین نزقیه های مربوط به وابستگی دو جامعه داریم:

از $+$ و $-$ در جدول

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases} \alpha \quad F_0 = \frac{S_1^2}{S_2^2} \quad \text{if } F_0 > F_{(n_1-1, n_2-1), \alpha} \Rightarrow RH$$



$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 < \sigma_1^2 \end{cases} \alpha \quad // \quad \text{if } F_0 < F_{(n_1-1, n_2-1), 1-\alpha} \Rightarrow RH$$



$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases} \alpha \quad // \quad \text{if } F_0 < F_{(n_1-1, n_2-1), 1-\frac{\alpha}{2}} \text{ or } F_0 > F_{(n_1-1, n_2-1), \frac{\alpha}{2}} \Rightarrow RH$$



TAHERIAN



مقایسه نسبت های دو جامعه: فرض کنید هدف مقایسه نسبت صفت حاصل در دو جامعه است.

مثلاً مقایسه نرخ بیماری در شهرستان دماوند، با نسبت افراد دارای ماشین ظرفشویی

در منطقه یک دماوند تهران.

با فرض بزرگ بودن نمونه ها (n_1 و n_2) پس رویت به اندازه کافی بزرگ از دو جامعه است.

پس می بینیم. ابتدا دیدیم نسبت بزرگ برآورده خوبی برای نسبت جامعه است یعنی:
 نسبت نمونه برای جامعه؟

$$\hat{p}_1 = \frac{X}{n_1}, \quad \hat{p}_2 = \frac{Y}{n_2}$$

نمونه ها مستقل \rightarrow

$$E(\hat{p}_1) = p_1, \quad E(\hat{p}_2) = p_2$$

طبق خواص امید ریاضی:

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$

پس برآورده نقطه ای برای $p_1 - p_2$ است.

و چون نمونه های اندازه کافی بزرگ هستند پس می دانیم:

$$\hat{p}_1 \sim N\left(p_1, \frac{p_1 q_1}{n_1}\right) \quad \text{پس} \quad \sqrt{(\hat{p}_1 - \hat{p}_2)^2} = \sqrt{(\hat{p}_1)^2 + (\hat{p}_2)^2}$$

$$\hat{p}_2 \sim N\left(p_2, \frac{p_2 q_2}{n_2}\right) \quad = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)$$

و از خواص توزیع نرمال:

7

Year: Month: Date:



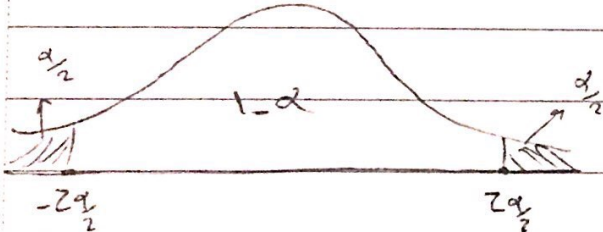
Sa Su Mo Tu We Th Fr

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

1. با استفاده از کسین:

$$\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

نکته: چون $\frac{p_1 q_1}{n_1}$ و $\frac{p_2 q_2}{n_2}$ مجهولند در کل از برآوردهای آنها یعنی $\frac{\hat{p}_1 \hat{q}_1}{n_1}$ و $\frac{\hat{p}_2 \hat{q}_2}{n_2}$ استفاده



2. معادله از رابطه 1

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

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

$$= 1 - \alpha$$

یعنی 'فاصله اطمینان' $(1 - \alpha)$ برای $p_1 - p_2$

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

آگر خردش نبود از $\frac{\hat{p}_1 \hat{q}_1}{n_1}$ و $\frac{\hat{p}_2 \hat{q}_2}{n_2}$ استفاده می‌کنیم یعنی

برای آزمون کسین فرضیه $H_0: p_1 - p_2 = 0$ (آیا دو جامعه دارای نسبت برابر هستند)

بیشتر

$$H_0: p_1 - p_2 = 0 \Rightarrow p_1 = p_2 = p$$



تعیین زوایا: در این مقایسات نمونه‌ها دایره هستند بر این مثال فرض کنید هدف مطالعه
(مغش)

تأثیر دو محیط متفاوت بر روی ظرفیت یادگیری کودکان قبل از ورود است. در این گونه

مسائل بهتر است از دو قطرهای یکسان استفاده کرد که از نظر سن و عوامل ترتیبی یکسان

هستند یا فرض کنید هدف مقایسه نمرات باینز و باراندت جوانان است که از یک گروه

دانشجویی را هر دو نمونه استفاده می‌شود. پس دو نمونه از دو جامعه مستقل نداریم به عبارتی دو نمونه

واسته هستند و یک جامعه یا گروه داریم که دو بار مورد آزمایش قرار گرفته‌اند.

دانشجو گروه‌های باینز گروه‌های باراندت

1	X_1	Y_1	$\rightarrow (X_1, Y_1)$	این تحلیل این دو نوع داده‌ها
...	
n	X_n	Y_n	$\rightarrow (X_n, Y_n)$	

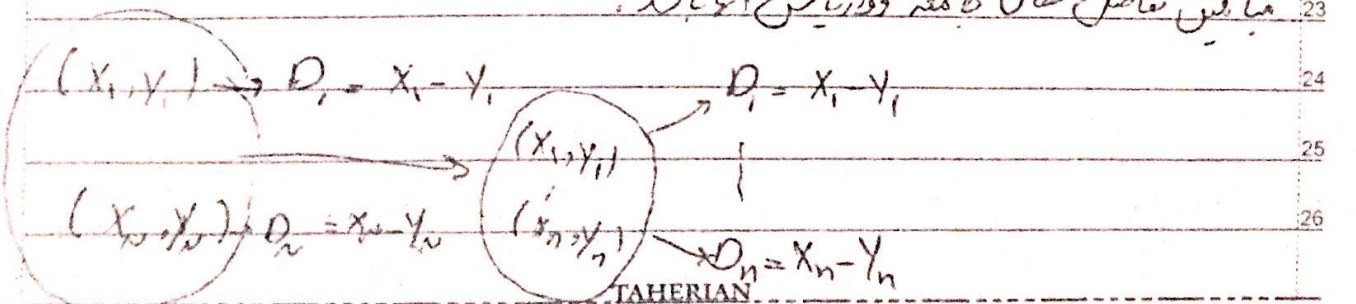
ابتدا تفاضل‌ها محاسبه می‌شوند.

X_1	Y_1	$D_1 = X_1 - Y_1$	$X_1 \dots X_n$
...
X_n	Y_n	$D_n = X_n - Y_n$	$Y_1 \dots Y_n$

$D_i = X_i - Y_i$

بطوریکه D_1, D_2, \dots, D_n مستقل هستند. فرض کنید δ و δ^2 به ترتیب

میانگین تفاضل‌های جامعه و واریانس آنها باشد.





$$\bar{D} = \frac{1}{N} \sum_{i=1}^N D_i$$

$$S_D^2 = \frac{1}{N} \sum_{i=1}^N (D_i - \bar{D})^2$$

واریانس فاضل ها جامعه

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i$$

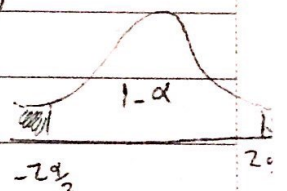
$E(\bar{D}) = \delta$ و میانگین فاضل نمونه

$$S_D^2 = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2$$

$E(S_D^2) = \sigma_D^2$ و واریانس فاضل نمونه

با فرض نرمال بودن جامعه یا n به اندازه کافی بزرگ:

$$\bar{D} \sim N\left(\delta, \frac{\sigma_D^2}{n}\right) \rightarrow Z = \frac{\bar{D} - \delta}{\frac{\sigma_D}{\sqrt{n}}}$$



و مثلاً برای جامعه فاضل احاطه ۹۵٪ برای میانگین فاضل جامعه:

$$\delta : \bar{D} \pm Z_{\alpha/2} \frac{\sigma_D}{\sqrt{n}} \xrightarrow[\text{مجموعه بزرگ}]{\text{مجموعه بزرگ یا n بزرگ و \sigma_D معلوم}} \bar{D} \pm Z_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

و مثلاً یک جامعه اگر جامعه نرمال، n کوچک و σ_D مجهول باشد از رابطه زیر برای محاسبه فاضل:

$$\delta : \bar{D} \pm t_{n-1, \alpha/2} \frac{S_D}{\sqrt{n}}$$

اطمینان استفاده می کنیم:

برای آزمون فرض فرضیه های زیر:

①
$$\begin{cases} H_0: \delta = \delta_0 \\ H_1: \delta > \delta_0 \end{cases}$$

②
$$\begin{cases} H_0: \delta = \delta_0 \\ H_1: \delta < \delta_0 \end{cases}$$

③
$$\begin{cases} H_0: \delta = \delta_0 \\ H_1: \delta \neq \delta_0 \end{cases}$$

! به صورتیکه n بزرگ باشد جامعه نرمال و σ_D معلوم:

$$Z_0 = \frac{\bar{D} - \delta_0}{\frac{\sigma_D}{\sqrt{n}}}$$

9

Year: Month: Date:



Sa Su Mo Tu We Th Fr

۲۱ در صورتیکه μ بزرگ باشد و μ_0 مجهول:

$$Z_0 = \frac{\bar{D} - \delta_0}{\frac{S_D}{\sqrt{n}}} \rightarrow \text{آماره آزمون}$$

۲۲ اگر μ کوچک و μ_0 مجهول و σ معلوم باشد:

$$T_0 = \frac{\bar{D} - \delta_0}{\frac{S_D}{\sqrt{n}}} \rightarrow \text{آماره آزمون}$$

۱۱ برای حالت ۱ و ۲ فرضیه های I، II، III:

$$I \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } Z_0 > Z_{\alpha} \Rightarrow RH_0$$

$$II \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } Z_0 < -Z_{\alpha} \Rightarrow RH_0$$

$$III \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } |Z_0| > Z_{\frac{\alpha}{2}} \Rightarrow RH_0$$

۱۹ برای حالت سوم:

$$I \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } T_0 > t_{n-1, \alpha} \Rightarrow RH_0$$

$$II \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } T_0 < -t_{n-1, \alpha} \Rightarrow RH_0$$

$$III \text{ فرضیه: } \text{قاعده تقسیم} \quad \text{if } |T_0| > t_{n-1, \frac{\alpha}{2}} \Rightarrow RH_0$$

1 مثال: در تحقیق به منظور مقایسه سیگارهای با کنترل، تعداد افراد تحت سیگاری و انفرادی که

2

3

3	تحت سیگار	انفرادی سیگار	
57	36	21	سیگار طبعی
55	44	11	سیگار کنترل
6			صفت خاص

4

5

سیگار را ترک نکردند به صورت زیر است

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

یک فاصله اطمینان 95٪ برای تفاضل بین نسبت

افراد که سیگار را ترک کرده اند در دو گروه دیگر

سیگار کنترل: گروه دوم

سیگار طبعی: گروه اول

$$\hat{p} = \frac{21}{57}$$

$$\hat{p} = \frac{11}{55}$$

$$p_1 - p_2 = (\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \quad p_1 - p_2 > 0 \Rightarrow p_1 > p_2$$

$$\Rightarrow p_1 - p_2 : \left(\frac{21}{57} - \frac{11}{55} \right) \pm 1.96 \sqrt{\frac{\frac{21}{57} \times \frac{36}{57}}{57} + \frac{\frac{11}{55} \times \frac{44}{55}}{55}} = \left(\frac{1.33}{57} \pm \frac{0.14}{55} \right)$$

16 با توجه به فاصله اطمینان به دست آمده چون صفر در بازه اطمینان نیست نتیجه می گیریم موقعیت سیگار

18 طبی از سیگار غیر طبعی بیشتر است. (چون بازه اطمینان مثبت است)

20 مثال: آماره های خلاصه زیر از دو نمونه تصادفی مستقل از دو جامعه شبه - شده است. با توجه

21 بر این اطلاعات فرقی لازم نیست همچنین یک فاصله اطمینان 95٪ برای پاریس به دست آورید

22 (فرض نال)

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases} \quad \begin{matrix} n_1 = 12 & \bar{x} = 249 & s_1 = 19 \\ n_2 = 15 & \bar{y} = 233 & s_2 = 45 \end{matrix}$$

23 حل: جامعه نوال است، هم نمونه کوچک است و دارای نال ها مجهولند $\sigma_1^2 = \sigma_2^2$

$$\sigma_1^2 \neq \sigma_2^2$$

TAHERIAN

10

$$\sigma_1^2 / \sigma_2^2 = 1$$

Year: Month: Date:



Sa Su Mo Tu We Th Fr

$$H_0: \sigma_1^2 = \sigma_2^2$$

1 پس اولین گام این است که دو جامعه از نظر برابری واریانس شوند:

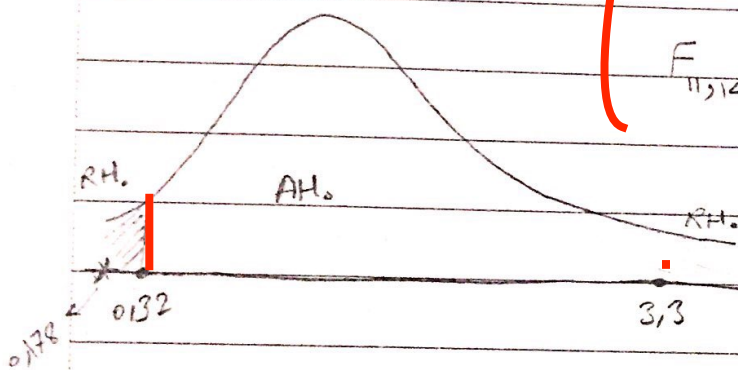
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

$$F_0 = \frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}$$

$$F_0 = \frac{361}{2025} = 0,178$$

$$F_{11,14,0,025} = 3,3$$

$$F_{11,14,0,975} = \frac{1}{3,1} = 0,32$$



$$0,178 < 0,32 \Rightarrow RH_0$$

پس دو جامعه از نظر برابری واریانس نیست.

پس از این حالت دوم استفاده میشود:

$$\mu_1 - \mu_2 = (\bar{X} - \bar{Y}) \pm t^*_{\min\{n_1-1, n_2-1\}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

فاصله اطمینان:

$$= (249 - 233) \pm t_{11,025} \sqrt{\frac{361}{12} + \frac{2025}{15}} \approx (12,28, 44,28)$$

$$\mu_1 > \mu_2$$

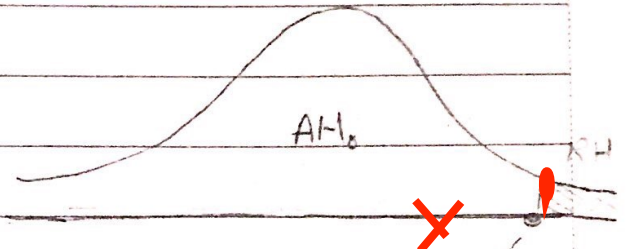
چون هم کران پایین و هم کران بالا میانگین مثبت است و مسائل منفیت پس نتیجه میگیریم

ما اطمینان 95٪ داریم که میانگین اول از میانگین دوم بزرگتر است که از آن این نتیجه میگیریم

$$T = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(249 - 233) - 0}{\sqrt{\frac{19^2}{12} + \frac{45^2}{15}}} \approx 1,25$$

$$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{19^2}{12} + \frac{45^2}{15}}$$

$$\begin{cases} H_0: \mu_1 = \mu_2 \rightarrow \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 > \mu_2 \rightarrow \mu_1 - \mu_2 > 0 \end{cases}$$



$$t_{11, 0,105} = 1,7959$$

$$1,25 > 1,79 \Rightarrow AH_0$$

$$t_{n_1+n_2-2, \alpha} = 1,79$$

این آزمون یک آزمون یک طرفه است و می‌توان طبق فاصله اطمینان آن برداشت کرد.

مثال: آماره‌های خلاصه‌شده برای نمونه‌های تصادفی مستقل دو جامعه ثبت شده‌اند با فرض نول

بودن جوامع قاصد اطمینان 95٪ برابر $\mu_1 = \mu_2$ و فرض $H_1: \mu_1 > \mu_2$

رایج سطح 0,105 آزمون کنید

$$\begin{aligned} n_1 &= 11 & \bar{X} &= 107 & S_1^2 &= 1,36 \\ n_2 &= 13 & \bar{Y} &= 96 & S_2^2 &= 2,17 \end{aligned}$$

ابتدا فرضیه برای بارش‌ها بررسی می‌شود:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

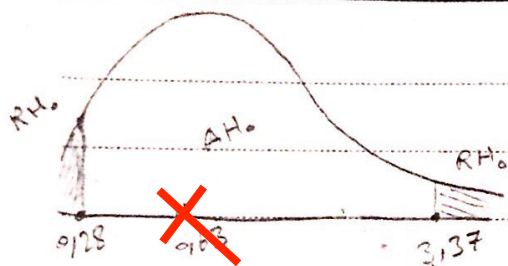
$$F_0 = \frac{S_1^2}{S_2^2} = \frac{1,36}{2,17} = 0,63$$

$$F_{n_1-1, n_2-1, \alpha/2} = F_{10, 12, 0,0525} = 3,37 \approx F_{10, 12, 0,0525} = \frac{1}{F_{12, 10, 0,0525}} = \frac{1}{3,62} = 0,276$$

11

Subject:

Date



$$3.37 > 0.63 > 0.28 \Rightarrow \text{Accept } H_0$$

این از زیر حالت اول استفاده می کنیم یعنی (جواب صحیح نهال، n ها و n_1 ها برابر است):

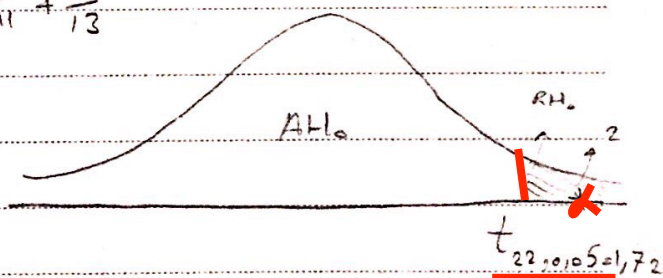
$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{10 \times 1.36 + 12 \times 2.17}{11 + 13 - 2} \approx 1.8$$

$$\mu_1 - \mu_2: (\bar{X} - \bar{Y}) \pm t_{n_1 + n_2 - 2, \frac{\alpha}{2}} \cdot S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (10.7, 9.6) \pm 2.074 \sqrt{1.8} \sqrt{\frac{1}{11} + \frac{1}{13}} = (-1.4, 2.24)$$

$$T_0 = \frac{(\bar{X} - \bar{Y}) - \delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(10.7 - 9.6) - 0}{\sqrt{1.8} \sqrt{\frac{1}{11} + \frac{1}{13}}} \approx 2$$

$$2 > 1.72 \Rightarrow \text{Reject } H_0$$



$$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases} \quad , \quad \begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 < P_2 \end{cases}$$

نشان 1 با توجه به آماره های زیر فرضیه

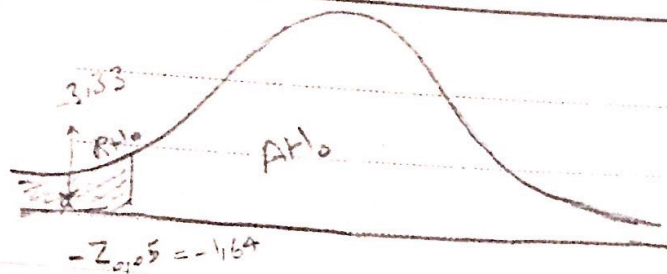
فرض اول را از روش p -value نیز می آید.

$$n_1 = 100 \quad \hat{P}_1 = 0.15$$

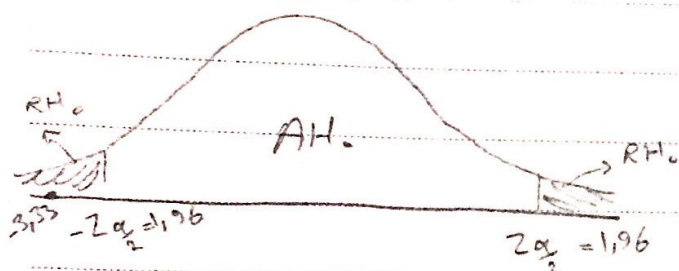
$$n_2 = 200 \quad \hat{P}_2 = 0.17$$

$$\hat{P} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2} = \frac{100 \times 0.15 + 200 \times 0.17}{100 + 200} \approx 0.163$$

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}\hat{Q}(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{0.15 - 0.17}{\sqrt{0.163(1 - 0.163)(\frac{1}{100} + \frac{1}{200})}} \approx -3.33$$



$$-3.33 < -1.64 \Rightarrow RH_0$$



برای فرضیه دوم:

$$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 \neq \mu_2$$

$$P\text{-Value} = p(Z \leq -3.33) = \Phi(-3.33) = 0.0004 < 0.05 \Rightarrow RH_0$$

المانی

مثال) برای تأثیرات تبلیغات صدجایه دانشگاه گیتی از دانشجوین و درن اینک را قبل از

تبلیغات و بعد از تبلیغات اندازه گرفت، یک فاصله اطمینان 95٪ برای تغییر وزن میابید آیا تبلیغات

تأثیر روی وزن داشته است؟ چون نمونه تغییر نکرد پس دو نمونه مستقل نداریم پس از استیلا
زوجی استفاده میکنیم.

قبل از تبلیغات 84 97 78 91 85
بعد 80 98 75 90 82

$$D_i = \text{قبل} - \text{بعد} \quad -4 \quad 1 \quad -3 \quad -1 \quad -3 \rightarrow \bar{D} = \frac{1}{5} (-4 + 1 - 3 - 1 - 3) = -2$$

$$(D_i - \bar{D}) = (-4+2)^2 + (-1)^2 + (-3+2)^2 + (-1+2)^2 + (-3+2)^2 = 4 + 1 + 1 + 1 + 1 = 8 \rightarrow S_D^2 = \frac{1}{4} (8) = 2$$

$$\delta: \bar{D} \pm t_{4, 0.025} \frac{S_D}{\sqrt{n}} = -2 \pm 2.776 \frac{\sqrt{2}}{\sqrt{5}} = (-4.48, 0.48)$$

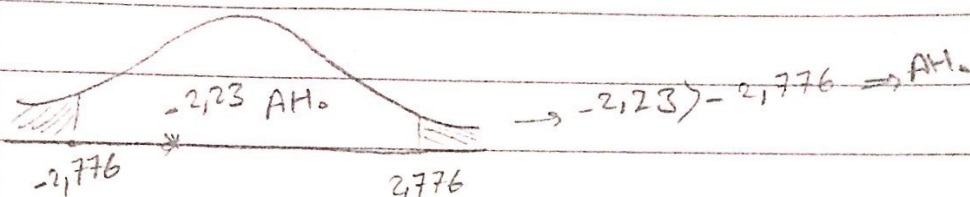
PAPCO

صفر را از میان اطمینان هست پس $\delta = 0$ نپذیرفته میشود



$$T_o = \frac{\bar{D}}{\frac{S_p}{\sqrt{n}}} = \frac{-2}{\frac{2}{\sqrt{5}}} = -2,23$$

$$\left\{ \begin{array}{l} H_0: \delta = 0 \\ H_1: \delta \neq 0 \end{array} \right.$$



طایفه‌های داده‌ها معلوم

$$\left\{ \begin{array}{l} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{array} \right.$$

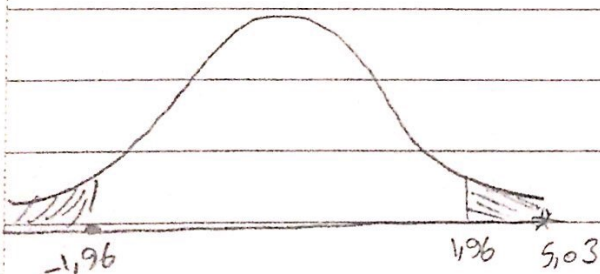
مثال) با توجه به اطلاعات زیر فرض کنید

$$n_1 = 36 \quad \bar{X} = 81 \quad \sigma_1 = 5,2$$

$$n_2 = 49 \quad \bar{Y} = 76 \quad \sigma_2 = 3,4$$

چون طایفه‌های داده‌ها معلوم است از Z استفاده می‌کنیم:

$$Z_o = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_o}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{81 - 76 - 0}{\sqrt{\frac{(5,2)^2}{36} + \frac{(3,4)^2}{49}}} = 5,023$$



پس چون $5,03 > 1,96$ فرض می‌کنیم رد می‌کنیم