Subject:
Year: Month: Date:

صفحه 398 تا 405 کتاب

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 در هر مسئله اولین گام تشکیل فرضیات است که به طور قراردادی به این صورت است که ادعا را فرض مقابل قرار میگیرد و آنچه به تجربه در مورد مسئله وجود دارد در فرض صفر به صورت زیر

مثال

فرض صفر

$H_0 : \mu = 1200$ → طول عمر لامپ های تولید شده کارخانه

$H_1 : \mu > 1200$ → مهندسی فرایند تولید پیشنهاد میدهد که طول عمر لامپ ها بزرگتر شود

فرض مقابل H_1

$$\equiv \begin{cases} H_0 : \mu \leq 1200 \\ H_1 : \mu > 1200 \end{cases}$$

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 براساس نمونه H_0 رد شود به غلط این خطا، خطای بدی است و همین منظور که خطای نوع اول نامیده میشود به عنوان سطح آزمون در نظر گرفته می شود و قبل از نتیجه‌گیری کنترل می‌شود که با α نشان میدهند و معمولاً عدد کوچکی مثل $\frac{1}{100}$، 0.01، 0.05، 0.1، ... فرض میشود) پس احتمال رد H_0 به شرطی که H_0 درست باشد) سطح آزمون می‌گویند:

$$\text{سطح آزمون} = \alpha = P(\underset{\text{reject}}{R}H_0 \mid H_0 \text{ درست باشد})$$

Sadra

* در عمل دنبال چیزهایی هستیم که توان بالایی دارند *

Subject:
Year: Month: Date:

و در مرحله آخر با فرض درست بودن H_0 و بر اساس نمونه استخراج شده و با توجه به مسئله آماره آزمون "آزمون" تشکیل می شود و تصمیم گیری انجام می شود.

خطای نوع دوم را معمولاً با β نشان میدهیم و به صورت زیر تعریف می شود:

$\beta = P(RH_1 \mid$ درست باشد $H_1)$

توان آزمون = لا $= 1 - \beta$

وضعیت / تصمیم	AH_0 → Accept H_0 پذیرفته می شود	RH_0 → Reject H_0 رد می شود
اگر H_0 درست باشد	تصمیم صحیح است. احتمال $= 1 - \alpha$	خطای نوع اول احتمال $= \alpha$
اگر H_0 غلط باشد	خطای نوع دوم: احتمال $= \beta$	تصمیم صحیح است. توان آزمون احتمال $= 1 - \beta$

مقدار احتمال (P_value): یک روش کلاسیک برای آزمون کردن فرضیه ها محاسبه مقدار احتمال است که به صورت زیر تعریف می شود:

کوچکترین مقداری که برای α میتوان در نظر گرفت = P_value ⇐ مقدار احتمال

تا H_0 رد شود. مثلاً فرض کنید T آماره آزمون تحت H_0 و t مقدار مشاهده شده از آن باشد Sadra

Subject:
Year: Month: Date:

آنگاه:

$$\begin{cases} H_0 = \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} \quad \text{P-Value} = P(T_0 > t)$$

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} \quad \text{P-Value} = P(T_0 < t)$$

برای تصمیم گیری:

if $\text{P-Value} \leq \alpha \Rightarrow RH_0$

اگر فرضیه به صورت $\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$ باشد P-Value عبارت است از:

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

یا

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

نکته: مساوی همیشه در H_0 قرار میگیرد.

- آزمون فرض برای میانگین (μ) جامعه: برای میانگین جامعه سه فرضیه زیر ممکن است مورد نظر باشد.

$$\left.\begin{aligned} &\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases} \equiv \begin{cases} H_0 : \mu \leq \mu_0 \\ H_1 : \mu > \mu_0 \end{cases} \\ &\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases} \equiv \begin{cases} H_0 : \mu \geq \mu_0 \\ H_1 : \mu < \mu_0 \end{cases} \end{aligned}\right\}$$ فرضیه‌های یک طرفه

Sadra

* معادل! if $|Z_0| > Z_{\frac{\alpha}{2}}$

Subject:
Year: Month: Date:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases} \longrightarrow$$ فرضیه دو طرفه

سطح آزمون را مد نظر میگیریم نمونه ای استخراج می شود $X_1 \dots X_n$ و اکنون

برای حالت های مختلف آماره های آزمون بدست آورده می شود

حالت اول: اگر جامعه نرمال باشد و واریانس جامعه معلوم باشد، دیدیم:

(یا n بزرگ باشد) و واریانس معلوم

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

به $Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ آماره آزمون تحت H_0 می گوییم، پس از محاسبه آماره

در مورد فرضیه تصمیم گیری می کنیم:

1 گام اول تعیین فرضیه

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

2 گام دوم تعیین سطح آزمون

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

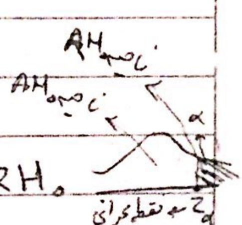
$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases} \quad \alpha$$

3 گام سوم نمونه گیری و محاسبه آماره

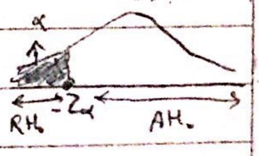
$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

4 گام چهارم تصمیم گیری

if $Z_0 > Z_\alpha \Rightarrow RH_0$



if $Z_0 < -Z_\alpha \Rightarrow RH_0$



*
$\begin{cases} \text{if } Z_0 \geq Z_{\frac{\alpha}{2}} \\ \text{if } Z_0 \leq -Z_{\frac{\alpha}{2}} \end{cases} \Rightarrow RH_0$

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

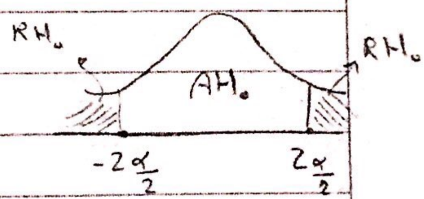
Subject:
Year: Month: Date:

حالت دوم: n بزرگ باشد (جامعه نرمال یا غیرنرمال) و واریانس مجهول باشد در اینصورت

قبلاً دیدیم:

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

①	②	③	④
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$	α	$Z_0 = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$	if $Z_0 \geq Z_\alpha \Rightarrow RH_0$
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$	α	"	if $Z_0 \leq -Z_\alpha \Rightarrow RH_0$
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$	α	"	if $\lvert Z_0 \rvert \geq Z_{\frac{\alpha}{2}} \Rightarrow RH_0$ $\begin{cases} Z_0 \geq Z_{\frac{\alpha}{2}} \\ Z_0 \leq -Z_{\frac{\alpha}{2}} \end{cases}$



حالت سوم: وقتی جامعه نرمال است و n کوچک و واریانس مجهول:

در این نیز فرض کنید نمونه تصادفی $X_1 \dots X_n$ در اختیار است.

①

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

قبلاً دیدیم:

$$\begin{cases} \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i & \text{میانگین نمونه} \\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i-\bar{X})^2 & \text{تغییرات نمونه} \end{cases}$$

Sadra

Subject:
Year: Month: Date:

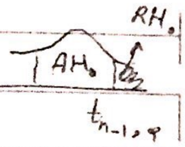
اکنون فرض کنید فرضیه های $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$ ، $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$ ، $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$

مورد نظر هستند. پس آماره ی آزمون با توجه به رابطه ①

$$T_0 \underset{H_0 \text{ تحت}}{=} \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

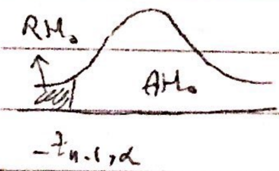
اکنون برای فرضیه اول چون $\bar{X}$ باید به اندازه کافی از μ_0 بزرگتر باشد تا H_0 را رد کنیم پس قاعده ی

فرضیه	سطح آزمون	آماره ی آزمون	تصمیم گیری
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$	α	$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$	if $T_0 > t_{n-1,\alpha} \Rightarrow RH_0$



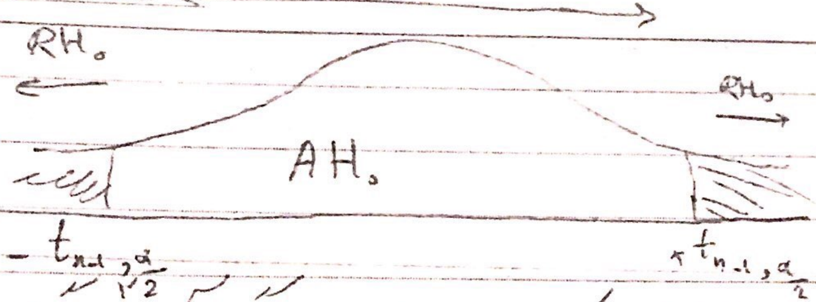
برای فرضیه دوم: برای این فرضیه باید $\bar{X}$ به اندازه کافی از μ_0 کوچکتر باشد که H_0 را رد کنیم.

فرضیه	سطح آزمون	آماره آزمون	قاعده تصمیم
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$	α	$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$	if $T_0 < -t_{n-1,\alpha} \Rightarrow RH_0$



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

فرضیه	سطح آزمون	آماره آزمون	قاعده تصمیم
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$	α	$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$	if $\begin{cases} T_0 > t_{n-1,\frac{\alpha}{2}} \\ T_0 < -t_{n-1,\frac{\alpha}{2}} \end{cases} \Rightarrow RH_0$

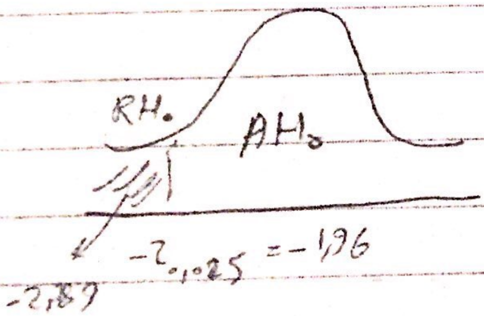


مثال) تعداد زیادی از بیماران مبتلا به یک بیماری بخصوص را گرد آورده و گزارش کرده اند که مدت درمان بیماری به روش استاندارد دارای میانگین 15 روز و انحراف معیار 3 روز است. ادعا شده که یک روش جدید میتواند مدت درمان را کوتاه‌تر کند. برای روش جدید درمان نمونه تصادفی به حجم 70 انتخاب شده است. و پس از اعمال روش جدید درمان بر روی این 70 نفر میانگین مدت درمان 14 روز شده است. آیا در سطح معنی داری (سطح آزمون) 0/025 روش جدید بهتر است؟ $\bar{X}=14$

$\mu_0=15$, $\underbrace{\sigma=3}_{\text{انحراف معیار معلوم}}$, $\underbrace{n=70}_{\text{حجم نمونه بزرگ}}$

$\alpha=0{,}025$

$$\begin{cases} H_0: \mu=15 \\ H_1: \mu<15 \end{cases} \longrightarrow Z_0=\frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}=\frac{14-15}{\frac{3}{\sqrt{70}}}=-2{,}789$$



$-2{,}789<-1{,}96 \Rightarrow RH_0$

مثال2) نمونه تصادفی از پرونده های فراوان شرکتی نشان میدهد که سفارشات برای قطعات معینی از ماشین ها به ترتیب روز بایگانی شده است. اگر تعداد روزهای بایگانی از توزیع نرمال

Subject:
Year: Month: Date:

پیروی کند آیا در سطح آزمون ۰/۰۱ میتوان ادعا کرد که میانگین زمان بایگانی از 10,5 روز بیشتر است.

10 12 19 14 15 18 11 13 $\rightsquigarrow n=8$

$$\begin{cases} H_0: \mu = 10,5 \equiv \mu \leq 10,5 \\ H_1: \mu > 10,5 \end{cases}$$

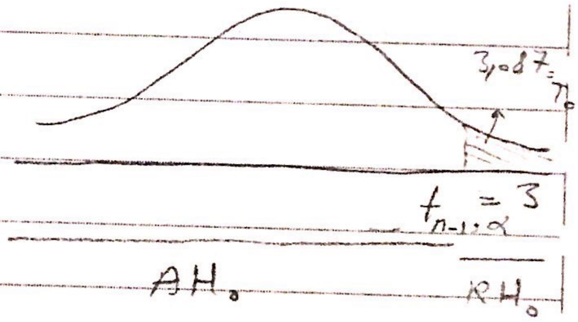
جامعه نرمال، واریانس جامعه مجهول

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

10,286

$$T_0 = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{14-10,5}{\frac{\sqrt{10,286}}{\sqrt{8}}} = 3,087$$

$\alpha = 0,01 \quad t_{n-1,0,01} = t_{8-1,0,01} = 3$



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

$3,087 > 3 \Rightarrow RH_0$ / یعنی میانگین زمان بایگانی قطعات بیشتر از 10,5 روز است.

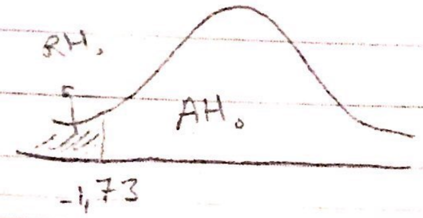
مثال) نمونه شامل ۲۰ رسید فروش کتاب یک فروشگاه کتاب با دارای میانگین 121 دلار و انحراف معیار 10,2 دلار است. از این داده‌ها برای آزمون این فرض که میانگین فروش کتاب بیشتر از 125 دلار است استفاده کنید. سطح آزمون را ۵٪ در نظر بگیرید. (فرض کنید جامعه نرمال است.)

$n = 20 < 30 \quad , \quad \bar{x} = 121 \quad , \quad S = 10,2$

$$T_0 = \frac{121-125}{\frac{10,2}{\sqrt{20}}} = -1,75$$

(اگر α معلوم نبود 0,05 در نظر بگیریم)

Subject:
Year: Month: Date:



$$-t_{n-1,\,0,05} = -t_{19,\,0,05} = -1,73$$

$$-1,75 < -1,73 \Rightarrow RH_0$$

مثال) از یک نمونه تصادفی شامل 36 مشاهده نتایج $S = 16,2$ و $\bar{X} = 80,4$ بدست آمده اند در مورد فرضیه های زیر تصمیم بگیرید. از روش آماره آزمون (مستقیم) و از روش P_Value (مقدار احتمال)

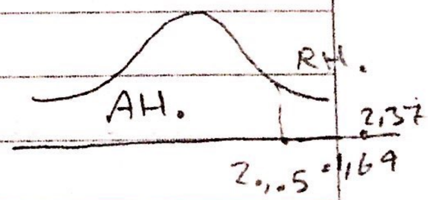
الف)
$$\begin{cases} H_0 : \mu = 74 \\ H_1 : \mu > 74 \end{cases} \qquad \alpha = 0,05$$

ب)
$$\begin{cases} H_0 : \mu = 76 \\ H_1 : \mu \neq 76 \end{cases}$$

الف)

$n = 36 \geq 30$, $\bar{X} = 80,4$, $S = 16,2$

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{80,4 - 74}{\frac{16,2}{\sqrt{36}}} = 2,37$$



$$\text{if } \underbrace{Z_0}_{2,37} > Z_\alpha = Z_{0,05} = 1,64 \Rightarrow RH_0$$

جدول نرمال $P(Z \leq z) = \Phi(z)$, $P\Big(\underbrace{\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}}_{Z_0} \geq 2,37\Big) = 1 - \Phi(2,37)$

$= 0,0089$

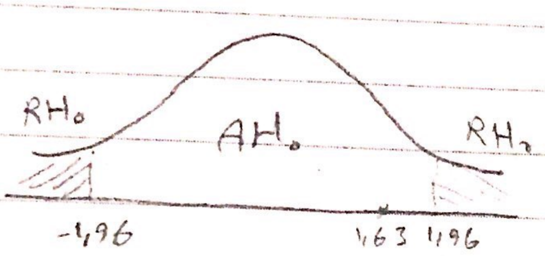
$\text{if } p\text{-value} < \alpha \Rightarrow RH_0$

$$0,0089 < 0,05 \Rightarrow RH_0$$

ب/

$$Z_0=\frac{80.4-76}{\frac{16.2}{\sqrt{36}}}=1.63$$

$$\text{if } Z_0>Z_{\frac{\alpha}{2}} \text{ or } Z<-Z_{\frac{\alpha}{2}} \Rightarrow RH_0$$



$$1.63 \not> 1.96 \Rightarrow AH_0$$

$$\longrightarrow 2\min\{p(Z_0>z_0),\, p(Z_0<-z_0)\}$$

$$P\text{-value}=P(|Z_0|>1.63)=P(Z_0>1.63)+P(Z_0<-1.63)$$

$$=(1-\Phi(1.63))+(\Phi(-1.63))=0.1032 \not< 0.05 \Rightarrow AH_0$$

- آزمون فرض برای (پیرامون) نسبت در جامعه: فرض کنید هدف مطالعه نسبت در جامعه است و در خصوص نسبت جامعه ادعایی شده است که مطالعه‌ی جامعه امکان‌پذیر نیست و از روی نمونه می‌خواهیم ادعا را بررسی کنیم سه فرضیه زیر ممکن است مورد نظر باشد:

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

برای این منظور نمونه‌ای استخراج می‌شود و قبلاً دیدیم:

$$\hat{p}=p \;\underset{\text{بزرگ}}{\overset{n\ \text{آنگاه}}{\approx}}\; N\left(p,\frac{pq}{n}\right)$$

پس تحت H_0:

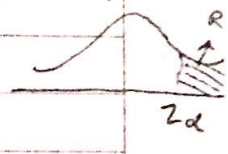
$$\hat{p}\simeq N\left(p_0,\frac{p_0q_0}{n}\right)$$

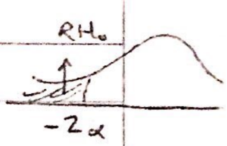
Subject:
Year: Month: Date:

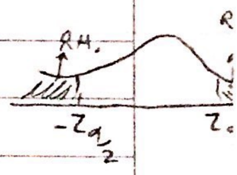
$$Z_0=\frac{\hat{p}-p_0}{\sqrt{\frac{p_0 q_0}{n}}}\sim N(0,1)\longrightarrow$$ آماره‌ی آزمون

پس داریم:

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







مثال) بر مبنای یک سرشماری در 5 سال پیش 20% خانواده های یک شهر زیر خط فقر زندگی می کردند. (جمعیت خانواده ها) برای معلوم شدن این موضوع که آیا این درصد تغییری کرده است یا خیر نمونه ای تصادفی به حجم 400 (نمونه خانواده ها) خانواده مطالعه شد و معلوم شد ۷۰ نفر زیر خط فقرند آیا این یافته ها دلالت بر این دارد که درصد (p) خانواده های زیر خط فقر با 5 سال پیش فرق کرده است؟ این

فرض H_1 را نشان میدهد $\Rightarrow p\neq 0.2$

آزمون را از روش مستقیم و از روش مقدار احتمال انجام دهید.

$$\begin{cases}H_0: p=0.2\\ H_1: p\neq 0.2\end{cases}$$

$n=400$ $\quad \alpha=0.05$

$X=70$ تعداد افرادی که صفت را دارند $\longrightarrow \hat{p}=\frac{70}{400}=0.175$

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

Subject:
Year: Month: Date: در مثال قبل برای آزمون دو طرفه

$\alpha = 0{,}025 \to Z_{0{,}025} = \boxed{1{,}96}$

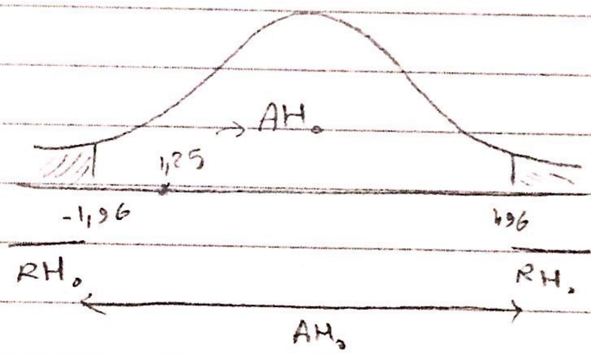
Conclusion $\to$ if $|Z_0| > Z_{\frac{\alpha}{2}} = 1{,}96$

$\Rightarrow 1{,}25 \not> 1{,}96 \Rightarrow AH_0$

$p\text{-value} = p(|Z_0| > -1{,}25) = p(Z_0 > 1{,}25) + p(Z_0 < -1{,}25)$

$= 1 - p(Z_0 < 1{,}25) + p(Z_0 < -1{,}25) = 1 - \phi(1{,}25) + \phi(-1{,}25) = \boxed{0{,}21}$

$p\text{-value} < 0{,}05 \Rightarrow AH_0$

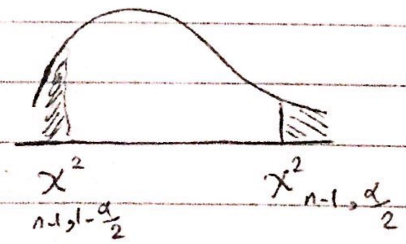


- آزمون برای واریانس جامعه: دیدیم برآورد نقطه‌ای خوب برای واریانس جامعه وقتی μ مجهول است عبارت است از: $\hat{\sigma}^2 = S^2$

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

از طرفی دیدیم وقتی جامعه نرمال است و یا n بزرگ:

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



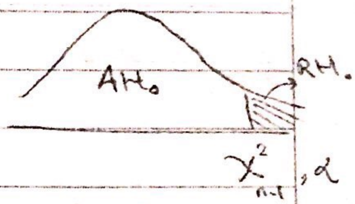
برای آزمون فرض، فرضهای مختلف داریم:

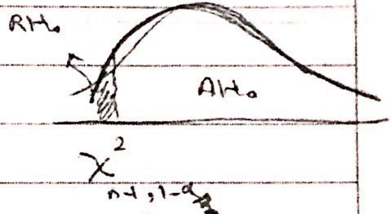
$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \to \mu \text{ مجهول}$$

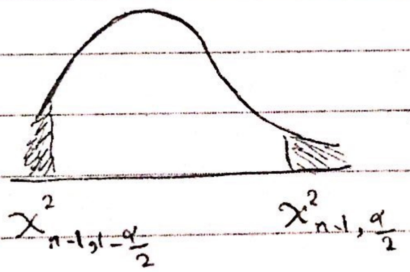
Subject:
Year: Month: Date:

$$S_n^2 = \frac{1}{n} \sum (X_i - \mu)^2 \to \mu \text{ معلوم.}$$

تشکیل فرضیه	سطح آزمون	نمونه‌گیری و محاسبه آماره	نتیجه‌گیری
$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \end{cases}$	α	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$	if $\chi_0^2 > \chi^2_{n-1,\alpha} \Rightarrow RH_0$
$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 < \sigma_0^2 \end{cases}$	α	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$	if $\chi_0^2 < \chi^2_{n-1,1-\alpha}$
$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$	α	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$	if $\chi_0^2 > \chi^2_{n-1,\frac{\alpha}{2}}$ or $\chi_0^2 < \chi^2_{n-1,1-\frac{\alpha}{2}}$







نکته: اگر μ معلوم باشد می‌دانیم برآوردگر خوب برای σ^2، S_n^2 است یعنی:

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

در روابط بالا باید از آماره‌ی آزمون μ استفاده کنید و از توزیع کای دو با درجه آزادی n.

یعنی هرجایی که χ^2_{n-1} است جای (n-1)، n بگذاریم، چون μ معلوم است.

$$\chi_0^2 = \frac{nS_n^2}{\sigma_0^2}$$

Subject:
Year: Month: Date:

مثال) فرض کنید بر مبنای نمونه تصادفی به حجم $n=25$ از توزیع نرمال، فاصله اطمینان

برای میانگین جامعه به صورت $(106,8 \,,\, 115,2)$ باشد از این نتیجه استفاده شده و

الف) مقدار انحراف معیار نمونه را بدست آورید.

هست یا نه؟

ب) در سطح آزمون 0.05 فرض انحراف معیار جامعه برابر 3

الف/ $n=25$ - جامعه نرمال

$(106,8 \,,\, 115,2) = \mu$ - حجم نمونه کوچک

- واریانس مجهول

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma \neq 3 \end{cases}$$

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

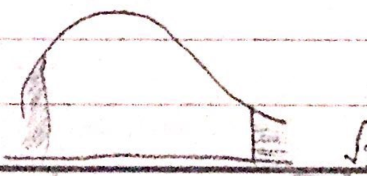
$$t_{24,0,025} = 2,064 \rightsquigarrow \begin{cases} \bar{X} - 2,064 \frac{S}{5} = 106,8 \\ \bar{X} + 2,064 \frac{S}{5} = 115,2 \end{cases}$$

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

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

ب/

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} \xrightarrow{\text{انحراف معیار}} \sqrt{\chi_0^2} = \frac{\sqrt{(n-1)}\,S}{\sigma_0} = \frac{\sqrt{24} \times 10,17}{3} = 16,61$$



Sadra

$\sqrt{39,36} = 6,27$

$3,52 \approx \sqrt{12,40} = \chi^2_{24,0,975}$ $\qquad \chi^2_{24,0,025}$

اگر ما دو جامعه یعنی میزان اثر با دو نوع چیز را ثبت کنیم دارو - گروه - آموزش و ...) *

Subject:
Year: Month: Date:

مبحث پایان ترم ۲

استنباط پیرامون دو جامعه :

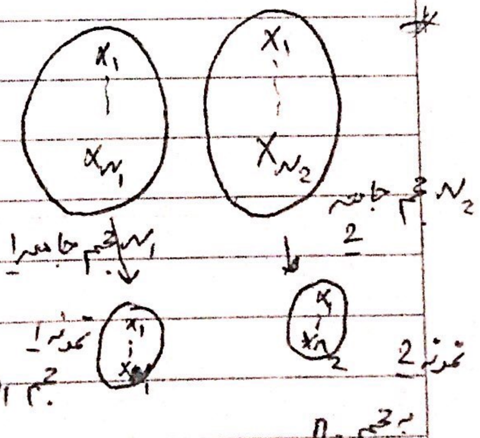
در این قسمت روشهای استنباط آماری برای مقایسه دو جامعه بر مبنای نمونه‌های مستقل را مورد بحث قرار میدهیم دقت کنید حتماً نباید دو جامعه داشته باشیم گاهی منظور دو تیمار (روش درمان) مثل دو نوع لاستیک یا دو نوع دارو مقایسه می‌شوند.

مثلاً فرض کنید 8 بیمار برای مطالعه‌ی اثر نوع دارو در اختیار یک محقق قرار میدهند که هدف مقایسه اثر داروی شماره یک با اثر داروی شماره 2 در مداوای بیماران است.

ابتدا به طور تصادفی 8 بیمار به دو گروه تقسیم میشوند و به گروه 1 داروی شماره 1 و به گروه 2 داروی شماره 2 داده میشود مشاهدات حاصل از داروی شماره یک هیچ ارتباطی با مشاهدات حاصل از داروی شماره 2 ندارد و تقسیم بیمارها به دو گروه کاملاً تصادفی انجام شده است.

پس در این بخش می‌خواهیم دو تیمار یا دو جامعه را بر مبنای نمونه‌های مستقل مقایسه کنیم.

برای استنباط پیرامون دو جامعه با فرض اینکه به هر دلیلی سرشماری امکان پذیر نیست ابتدا دو نمونه از هر جامعه انتخاب می‌کنیم. استخراج دو نمونه کاملاً مستقل است.



از جامعه اول $\rightarrow$ $\underbrace{X_1 \dots X_{n_1}}_{\text{متغیرها از هم مستقل}}$

از هم مستقل

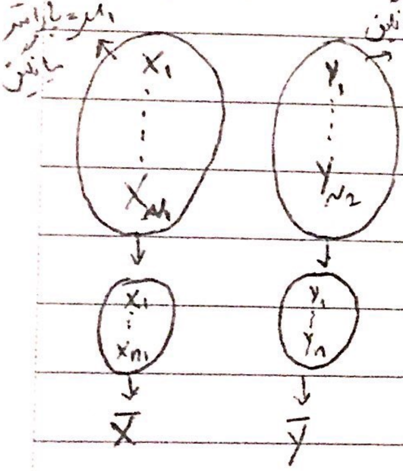
از جامعه دوم $\rightarrow$ $\underbrace{Y_1 \dots Y_{n_1}}_{\text{متغیرها از هم مستقل}}$

تا به حال استنباط پیرامون یک جامعه. می‌خواهیم از روی نمونه‌های تصادفی دو جامعه را باهم

مقایسه کنیم و پارامترهای مهم مانند میانگین و واریانس نسبت به هم را مورد بررسی قرار می‌دهیم.

مقایسه میانگین دو جامعه ($\mu_2-\mu_1$) یا ($\mu_1-\mu_2$ فرقی ندارد): استقلال دو نمونه

مستقل از دو جامعه به ترتیب با میانگین μ_1 و μ_2 استخراج می‌شوند. می‌دانیم میانگین نمونه



برآوردگر خوبی برای میانگین‌های جامعه است.

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

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

از خواص امید ریاضی:

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

پس $\bar{Y}-\bar{X}$ یک برآوردکننده نااریب برای $\mu_2-\mu_1$ است.

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

$$\boxed{\widehat{\mu_2-\mu_1}=\bar{Y}-\bar{X}}$$ → «برآورد نقطه‌ای»

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

برای بدست آوردن برآورد فاصله‌ای حالت‌های مختلف را بررسی می‌کنیم:

حالت اول: نمونه‌ها به اندازه کافی بزرگ باشند و واریانس‌ها معلوم یا جامعه نرمال باشد.

$$V(ax+by) = a^2V(X) + b^2V(Y) - 2ab\,cov(X,Y)$$

واریانس ها معلوم (یا جامعه نرمال باشد و واریانس معلوم).

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

چون از هم مستقل اند

$$V(\bar{Y}-\bar{X}) = V(\bar{Y}) + V(\bar{X}) - \cancel{2cov}(0) = V(\bar{Y}) + V(\bar{X})$$

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

طبق خواص نرمال

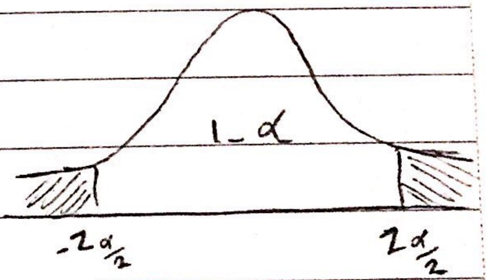
$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

$$\longrightarrow \boxed{\bar{Y}-\bar{X} \sim N\left(\mu_2-\mu_1, \frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}\right)}$$

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

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



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

با جایگذاری Z :

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

با نگه داشتن $\mu_2-\mu_1$ در وسط و انتقال سایر جمله ها به طرفین داریم:

Year : Month : Date :

Sa | Su | Mo | Tu | We | Th | Fr

$$P\left((\bar{y}-\bar{x}) - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}} < \mu_2-\mu_1 < (\bar{y}-\bar{x}) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}}\right)$$

بنابراین فاصله اطمینان $(1-\alpha)100$ ٪ برای $\mu_2-\mu_1$ وقتی n_1 و n_2 بزرگ

باشند و واریانس ها معلوم هستند بصورت زیر است:

$$\boxed{\mu_2-\mu_1 : \left((\bar{y}-\bar{x}) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}}\right)}$$

نکته: فقط در حالتیکه n_1 و n_2 هر دو بزرگ هستند و واریانس ها مجهول در اینصورت میتوان

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

$$\boxed{\mu_2-\mu_1 : \left((\bar{y}-\bar{x}) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{S_2^2}{n_2}+\frac{S_1^2}{n_1}}\right)}$$

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

مثال) برای مقایسه سن ازدواج زنان در دو ایل A و B از هر ایل نمونه‌ای شامل 100

زن استخراج شده است و سن ازدواج آنها را ثبت کرده‌ایم و خلاصه آماره‌ها به شرح زیر بدست

آمده است.

	A	B
حجم نمونه	$n_1=100$	$n_2=100$
میانگین	$\bar{x}=18,5$	$\bar{y}=20,7$
واریانس	$S_A=6,2$	$S_B=5,8$

$\mu_B-\mu_A=$

برآورد نقطه‌ای $\Leftarrow$ $\bar{y}-\bar{x}=20,7-18,5=2,2$

سن ازدواج در ایل A کمتر از ایل B است.

برآورد فاصله‌ای :

$\alpha=0,05 \Rightarrow 1-\alpha=0,95$ و $Z_{0,025}=1,96$

$$\mu_B-\mu_A:\left((\bar{y}-\bar{x})\pm 1{,}96\sqrt{\frac{S_B^2}{n_2}+\frac{S_A^2}{n_1}}\right)$$

$$=\left((20{,}7-18{,}5)\pm 1{,}96\sqrt{\frac{(5{,}8)^2}{100}+\frac{(6{,}2)^2}{100}}\right)$$

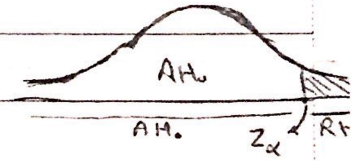
$= (0{,}52\ ,\ 3{,}88) \rightarrow$ چون کران پایین و کران بالا

فاصله اطمینان مثبت است پس با اطمینان 95% سن ازدواج زنان ایل B از ایل A بیشتر است.

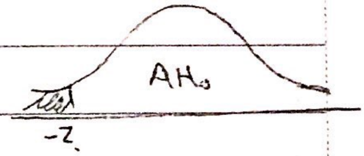
- آزمون فرض در این حالت: (n_1 و n_2 بزرگند و واریانس ها مجهول (معلوم)) سه فرضیه مورد توافق است به شرح زیر:

تشکیل فرضیه — سطح آزمون — آماره ی آزمون — تصمیم گیری (قاعده تصمیم)

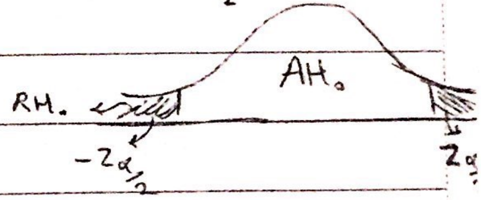
$$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ H_1: \mu_2-\mu_1>\delta_0 \end{cases} \qquad \alpha \qquad Z_0=\frac{(\bar{y}-\bar{x})-\overbrace{(\mu_2-\mu_1)}^{\delta_0}}{\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}}} \qquad \text{if } Z_0>Z_\alpha \Rightarrow RH_0$$



$$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ H_1: \mu_2-\mu_1<\delta_0 \end{cases} \qquad \alpha \qquad // \qquad \text{if } Z_0<-Z_\alpha \Rightarrow RH_0$$



$$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ H_1: \mu_2-\mu_1\neq\delta_0 \end{cases} \qquad \alpha \qquad // \qquad \text{if } |Z_0|>Z_{\frac{\alpha}{2}} \Rightarrow RH_0 \quad *$$



$\ast\ Z_0>Z_{\frac{\alpha}{2}}$ or $Z_0<-Z_{\frac{\alpha}{2}} \Rightarrow RH_0$

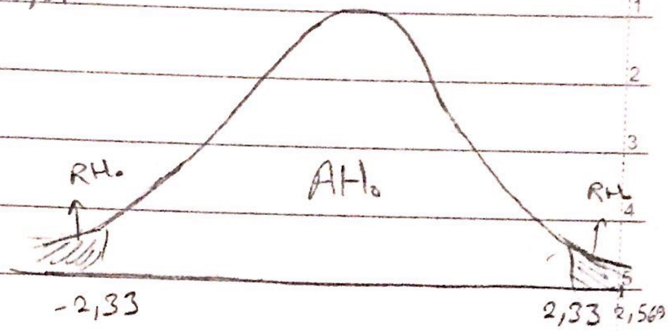
مثال) آیا در مثال قبل شواهد قوی برای قبول این ادعا که متوسط سن ازدواج در دو ایل متفاوت است وجود دارد؟ سطح آزمون را 0,02 در نظر بگیرید.

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

TAHERIAN

$\alpha = 0{,}02 \rightarrow 1-\alpha = .98 \Rightarrow \frac{\alpha}{2} = 0{,}01$

$$Z_0 = \frac{(20{,}7 - 18{,}5) - (0)}{\sqrt{\frac{(5{,}8)^2}{100} + \frac{(6{,}3)^2}{100}}} = 2{,}569$$



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

$2{,}569 > 2{,}33 \Rightarrow RH_0$

تفسیر: چون H_0 رد می‌شود ⇐ متوسط سن ازدواج دو ایل متفاوت است.

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

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

$\sigma_1^2 = \sigma_2^2 = \sigma^2$

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

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

برای زیر حالت اول داریم:

چون جامعه نرمال $X_1 \dots X_n \rightarrow \bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1} = \frac{\sigma^2}{n_1}\right)$

نمونه از جامعه نرمال اول

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

نمونه‌ای تصادفی از جامعه نرمال دوم

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

$\frac{(n_1 - 1) S_1^2}{\sigma^2} \sim \chi^2_{n_1 - 1}$

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

$\frac{(n_2 - 1) s^2}{\sigma^2} \sim \chi^2_{n_2 - 1}$

جمع دو توزیع کای دو که ای دو مستقل شده (فقط درجه آزادی ها باهم جمع می‌شوند)

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

$$\left.\begin{array}{l} E(S_1^2) = \sigma_1^2 = \sigma^2 \\ E(S_2^2) = \sigma_2^2 = \sigma^2 \end{array}\right\} \rightarrow S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

pulld ترکیبی

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

از طرفی طبق خواص توزیع نرمال:

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

$$\begin{cases} E(\bar{y} - \bar{x}) = \mu_2 - \mu_1 \\ V(\bar{y} - \bar{x}) = V(\bar{y}) + V(\bar{x}) - 2\operatorname{cov}(\bar{x}, \bar{y}) = \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \end{cases}$$

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:

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

بنابراین:

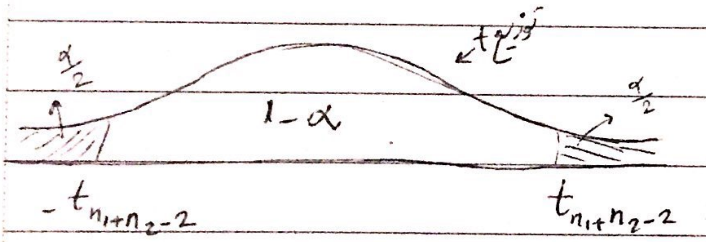
$$\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{\underbrace{(n_2-1)+(n_1-1)}_{n_1+n_2-2}} \quad (2)$$

اکنون یک توزیع نرمال استاندارد (1) و یک توزیع کای دو (2) داریم و می‌دانیم:

$$T=\frac{\dfrac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\cancel{\sigma}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}{\sqrt{\dfrac{\left(\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\cancel{\sigma^2}}\right)}{n_1+n_2-2}}}\sim t_{n_1+n_2-2}$$

$$\Longrightarrow \boxed{T=\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{n_1+n_2-2}} \quad ☆$$

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



$$P\left(-t_{n_1+n_2-2,\frac{\alpha}{2}}<T<T_{n_1+n_2-2,\frac{\alpha}{2}}\right)=1-\alpha$$

$$\Rightarrow P\left(-t_{n_1+n_2-2,\frac{\alpha}{2}}<\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}<t_{n_1+n_2-2,\frac{\alpha}{2}}\right)=1-\alpha$$

$$\overset{\text{پس}}{\Longrightarrow} P\left((\bar{Y}-\bar{X})-t_{n_1+n_2-2,\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_2-\mu_1<(\bar{Y}-\bar{X})+t_{n_1+n_2-2,\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

پس فاصله اطمینان $100(1-\alpha)\%$ برای حالتی که: جوامع نرمال باشند σ_1^2 و σ_2^2 مجهول ولی برابر

و n_1 و n_2 یا هر دو کوچک باشند.

$\delta_0 = 0$

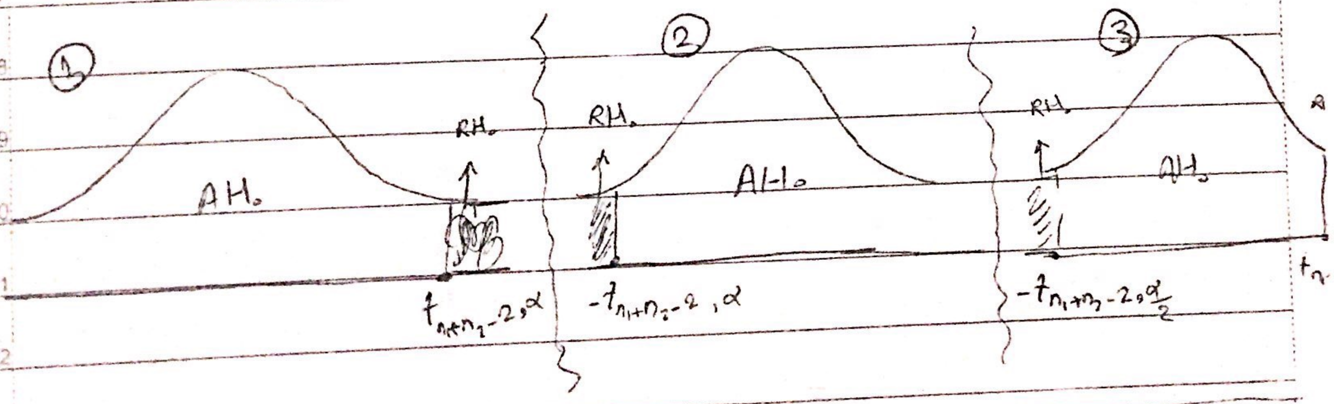
Year: Month: Date:

Sa Su Mo Tu We Th Fr

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

برای آزمون فرضیه های زیر هم از آماره ی (*) استفاده می کنیم:

فرضیه ها	سطح آزمون	آماره آزمون	قاعده تصمیم
(1) $\equiv H_0: \mu_2-\mu_1 \leq \delta_0$ $H_0: \mu_2-\mu_1 = \delta_0$ $H_1: \mu_2-\mu_1 > \delta_0$	α	$T_0 = \frac{(\bar{y}-\bar{x})-\delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	if $T_0 > t_{n_1+n_2-2,\alpha}$ $\Rightarrow RH_0$
(2) $\equiv H_0: \mu_2-\mu_1 \geq \delta_0$ $H_0: \mu_1-\mu_1 = \delta_0$ $H_1: \mu_2-\mu_1 < \delta_0$	α	//	if $T_0 < t_{n_1+n_2-2,\alpha} \Rightarrow RH_0$
(3) $H_0: \mu_2-\mu_1 = \delta_0$ $H_0: \mu_2-\mu_1 \neq \delta_0$	α	//	if $\lvert T_0\rvert < t_{n_1+n_2-2,\frac{\alpha}{2}} \Rightarrow R$



برای زیر حالت دوم داریم:

جوامع نرمال و نمونه کوچک و واریانس ها مجهول و نا برابر:

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

در این زیر حالت از آماره‌ی زیر استفاده می‌شود:

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

یا df را به صورت زیر بدست می‌آوریم که دقیق‌تر است:

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

و برای آزمون کردن فرضیه‌های بالا از آماره آزمون:

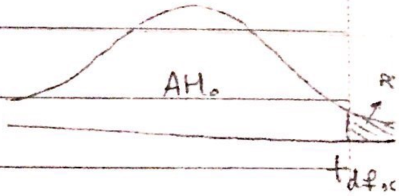
$$T_0^* = \frac{(\bar{Y}-\bar{X})-\delta_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \quad (**)$$

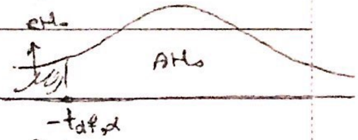
استفاده می‌کنیم و فاصله اطمینان را در این حالت:

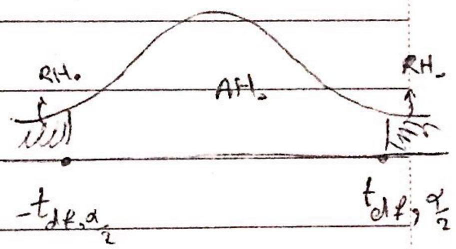
$$\mu_2-\mu_1 : \left((\bar{Y}-\bar{X}) \pm t^*_{df,\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)$$

و برای آزمون فرضیه‌های یکطرفه هم از آماره‌ی ** استفاده می‌کنیم:

فرضیه‌ها	سطح آزمون	آماره آزمون	قاعده تصمیم
$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ \quad\; = \mu_2-\mu_1 \leq \delta_0 \\ H_1: \mu_2-\mu_1 > \delta_0 \end{cases}$	α	$T_0^* = \dfrac{(\bar{y}-\bar{x})-\delta_0}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}}$	if $T_0^* > t^*_{df,\alpha} \Rightarrow RH_0$
$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ \quad\; = \mu_2-\mu_1 \geq \delta_0 \\ H_1: \mu_2-\mu_1 < \delta_0 \end{cases}$	α	//	if $T_0^* < -t_{df,\alpha} \Rightarrow RH_0$
$\begin{cases} H_0: \mu_2-\mu_1=\delta_0 \\ H_1: \mu_2-\mu_1 \neq \delta_0 \end{cases}$	α	//	if $\lvert T_0^* \rvert > t^*_{df,\frac{\alpha}{2}} \Rightarrow RH_0$







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

است. قبلاً دیدیم: (با فرض نرمال بودن دو جامعه یا n_1 و n_2 بزرگ):

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

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

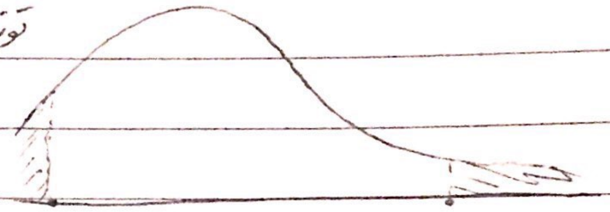
$$\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$$

6

(توزیع F از کای دو هر دو مثبت هستند مقدار منفی معنا ندارد)

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

توزیع



$F_{n_1-1, n_2-1, 1-\frac{\alpha}{2}}$ $\quad$ $F_{n_1-1, n_2-1, \frac{\alpha}{2}}$

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

بنابراین فاصله اطمینان $100(1-\alpha)\%$ برای نسبت واریانس ها:

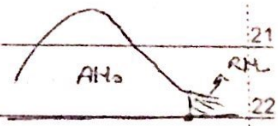
$$\frac{\sigma_2^2}{\sigma_1^2} : \left(\frac{S_2^2}{S_1^2} F_{n_1-1,n_2-1,1-\frac{\alpha}{2}} \,,\, \frac{S_2^2}{S_1^2} F_{n_1-1,n_2-1,\frac{\alpha}{2}}\right)$$

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

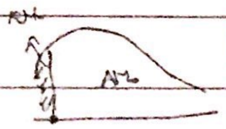
برای آزمون کردن فرضیه های مربوط به واریانس دو جامعه داریم:

فرضیه ها $\quad$ سطح آزمون $\quad$ از *** داریم

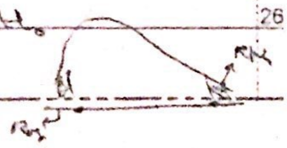
$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases}$ $\quad \alpha \quad$ $F_0 = \dfrac{S_1^2}{S_2^2}$ $\quad$ if $F_0 > F_{n_1-1,n_2-1,\alpha} \Rightarrow RH_0$



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



$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$ $\quad \alpha \quad$ // $\quad$ if $\begin{cases} F_0 < F_{n_1-1,n_2-1,1-\frac{\alpha}{2}} \\ F_0 > F_{n_1-1,n_2-1,\frac{\alpha}{2}} \end{cases} \Rightarrow RH_0$



TAHERIAN

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

مثلاً مقایسه نرخ بیکاری در شهرستان ساری با نسبت افراد دارای ماشین ظرفشویی

در منطقه یک یا یازده تهران.

با فرض بزرگ بودن نمونه ها (n_1 و n_2) پس دو نمونه با اندازه کافی بزرگ از دو جامعه استخراج

می کنیم. قبلاً دیدیم نسبت نمونه برآورد خوبی برای نسبت جامعه است یعنی:

نسبت نمونه برای جامعه 2

نسبت نمونه برای جامعه 1 $\hat{p}_1 = \frac{X}{n_1}$, $\hat{p}_2 = \frac{Y}{n_2}$ $\Leftarrow$ نمونه ها مستقل

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

طبق خواص امید ریاضی:

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

پس برآورد نقطه ای برای $p_1 - p_2$ ، $\hat{p}_1 - \hat{p}_2$ است.

و چون نمونه ها به اندازه کافی بزرگ هستند پس می دانیم:

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

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

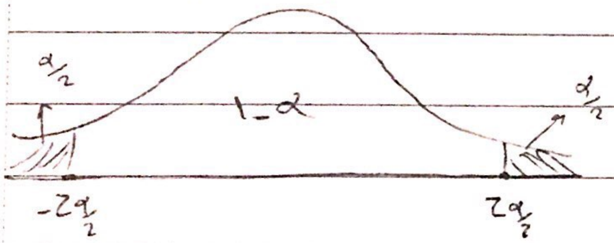
از خواص توزیع نرمال:

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

با استاندارد کردن:

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}} \quad ①$$

نکته: چون $\frac{p_1q_1}{n_1}$ ، $\frac{p_2q_2}{n_2}$ مجهول‌اند به‌دلیل از برآوردگرها یعنی $\frac{\hat{p}_1\hat{q}_1}{n_1}$ ، $\frac{\hat{p}_2\hat{q}_2}{n_2}$ استفاده می‌کنیم. از رابطه ①



$$P\left(-Z_{\frac{\alpha}{2}}<\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}<Z_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$\Rightarrow P\left((\hat{p}_1-\hat{p}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}<p_1-p_2<(\hat{p}_1-\hat{p}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}\right)$$

$$=1-\alpha$$

پس فاصله اطمینان $100(1-\alpha)\%$ برای p_1-p_2:

$$p_1-p_2:(\hat{p}_1-\hat{p}_2)\pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}$$

اگر خروجی نبود از

برآوردهاشون استفاده می‌کنیم یعنی $\frac{\hat{p}_1\hat{q}_1}{n_1}$ ، $\frac{\hat{p}_2\hat{q}_2}{n_2}$

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

یک مقدار

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

پس آماره آزمون از رابطه ①

$$Z_0 = \frac{(\hat{p}_1 - \hat{q}_1) - 0}{\sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}}} \quad ②$$

H_0 تحقق

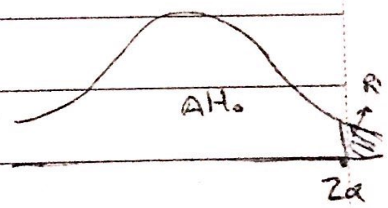
$\rightsquigarrow p_1 - p_2 = 0 \Rightarrow p_1 = p_2 = p$ این مجهول است

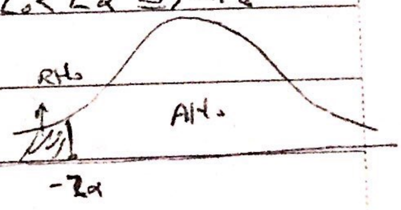
چون معمولاً p و q مجهول‌اند پس یک برآورد ترکیبی از هر دو نمونه برای p در نظر می‌گیرند

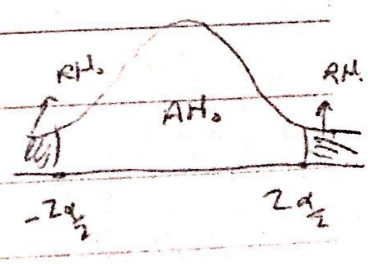
$$\hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$$

بطور کلی:

فرضیه	سطح آزمون	آماره آزمون ②	ناحیه تصمیم
$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_1: p_1 - p_2 > 0 \end{cases}$	α	$Z_0 = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$	if $Z_0 > Z_\alpha \Rightarrow RH_0$
$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_1: p_1 - p_2 < 0 \end{cases}$	α	//	if $Z_0 < Z_\alpha \Rightarrow RH_0$
$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_1: p_1 - p_2 \neq 0 \end{cases}$	α	//	if $\lvert Z_0 \rvert > Z_{\frac{\alpha}{2}} \Rightarrow RH_0$







مقایسات زوجی: در این مقایسات نمونه‌ها وابسته هستند. برای مثال فرض کنید هدف مطالعه (مقایسه) باشد دو محیط متفاوت بر روی ظرفیت یادگیری کودکان قبل از مدرسه است. در این گونه مسائل بهتر است از دوقلوهای یکسان استفاده کرد که از نظر سنی و عوامل ژنتیکی یکسان هستند یا فرض کنید هدف مقایسه نمرات پاییز و بهار دانشجویان است که از یک گروه دانشجویی برای هر دو نمونه استفاده می‌شود. پس دو نمونه از دو جامعه مستقل نداریم به عبارتی دو نمونه وابسته هستند و یک جامعه یا گروه داریم که دو بار مورد آزمایش قرار گرفته‌اند.

دانشجو	نمره‌های بهار	نمره‌های پاییز	
1	X_1	Y_1	$\to (X_1, Y_1)$
$\vdots$	$\vdots$	$\vdots$	
n	X_n	Y_n	$\to (X_n, Y_n)$

برای تحلیل این نوع داده‌ها ابتدا تفاضل‌ها محاسبه می‌شوند.

$$\begin{array}{lll} X_1 & Y_1 & D_1 = X_1 - Y_1 \\ \vdots & \vdots & \vdots \\ X_n & Y_n & D_n = X_n - Y_n \end{array}$$

$X_1 \ldots X_n$ ، $Y_1 \ldots Y_n$ ، $D_i = X_i - Y_i$

بطوریکه $D_1, D_2, \ldots, D_n$ مستقل هستند. فرض کنید δ و σ_D^2 به ترتیب میانگین تفاضل‌های جامعه و واریانس آن‌ها باشد.

$(X_1, Y_1) \to D_1 = X_1 - Y_1$

$(X_n, Y_n) \to D_n = X_n - Y_n$

$(X_1, Y_1) \to D_1 = X_1 - Y_1$
$\vdots$
$(X_n, Y_n) \to D_n = X_n - Y_n$

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

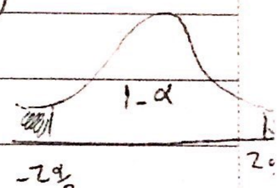
$$\sigma_D^2 = \frac{1}{N}\sum_{i=1}^{N} (D_i - \delta)^2$$ واریانس تفاضل های جامعه

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$$ میانگین تفاضل نمونه و $E(\bar{D}) = \delta$ برآورد نااریب δ

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n} (D_i - \bar{D})$$ واریانس تفاضل نمونه و $E(S_D^2) = \sigma_D^2$ برآورد نااریب σ_D^2

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

$$\bar{D} \sim N\left(\delta, \frac{\sigma_D^2}{n}\right) \longrightarrow Z = \frac{\bar{D} - \delta}{\frac{\sigma_D}{\sqrt{n}}} \sim N(0,1)$$



و مشابه یک جامعه فاصله اطمینان برای میانگین تفاضل جامعه:

$$\underbrace{\delta : \bar{D} \pm Z_{\frac{\alpha}{2}} \frac{\sigma_D}{\sqrt{n}}}_{\text{جامعه نرمال یا n بزرگ و } \sigma_D \text{ معلوم}} \xrightarrow[\text{مجهول باشد یا n بزرگ}]{\sigma_D} \bar{D} \pm Z_{\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$$

و مشابه یک جامعه اگر جامعه نرمال، n کوچک و σ_D مجهول باشد از رابطه زیر برای محاسبه فاصله اطمینان استفاده می‌کنیم:

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

برای آزمون فرض فرضیه های زیر:

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

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

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

! در صورتیکه n بزرگ باشد یا جامعه نرمال و σ_D معلوم:

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

۲) درصورتیکه n بزرگ باشد و σ_D مجهول:

$$Z_0 = \frac{\bar{D} - \delta_0}{\frac{S_D}{\sqrt{n}}} \longrightarrow$$ آماره‌ی آزمون

۳) اگر n کوچک و جامعه نرمال و σ_D مجهول:

$$T_0 = \frac{\bar{D} - \delta_0}{\frac{S_D}{\sqrt{n}}} \longrightarrow$$ آماره‌ی آزمون

برای حالت 1 و 2 فرضیه‌های I و II و III:

فرضیه I : قاعده تصمیم $\quad$ if $Z_0 > Z_\alpha \Rightarrow RH_0$

II " : " $\quad$ if $Z_0 < -Z_\alpha \Rightarrow RH_0$

III " : " $\quad$ if $|Z_0| > Z_{\frac{\alpha}{2}} \Rightarrow RH_0$

برای حالت سوم:

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

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

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

مثال) در تحقیقی به منظور مقایسه سیگار طبی با فیلتر، تعداد افراد ترک سیگار در افرادی که سیگار را ترک نکرده اند به صورت زیر است.

	ترک سیگار	ادامه سیگار	
سیگار طبی	21	36	57
سیگار فیلتر	$\underbrace{11}_{\text{صفت خاص}}$	44	55

یک فاصله اطمینان ٪95 برای تفاضل بین نسبت افرادی که سیگار را ترک کرده اند در دو گروه به دست آورید.

سیگار فیلتر: گروه دوم $\hat{p} = \frac{11}{55}$ — سیگار طبی: گروه اول $\hat{p} = \frac{21}{57}$

$$P_1 - P_2 = (\hat{P}_1 - \hat{P}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}$$

$P_1 - P_2 > 0 \Rightarrow P_1 > P_2$

$$\Rightarrow P_1 - P_2 : \left(\frac{21}{57} - \frac{11}{55}\right) \pm 1{,}96 \sqrt{\frac{\frac{21}{57} \times \frac{36}{57}}{57} + \frac{\frac{11}{55} \times \frac{44}{55}}{55}} \simeq (\underbrace{0{,}01}_{>0}, \underbrace{0{,}33}_{>0})$$

با توجه به فاصله اطمینان به دست آمده چون صفر در بازه اطمینان نیست نتیجه می گیریم موفقیت سیگار طبی از سیگار فیلتر بیشتر است. (چون بازه اطمینان مثبت هم هست)

مثال) آماره های خلاصه زیر از دو نمونه تصادفی مستقل از دو جامعه ثبت شده است. با توجه به این اطلاعات فرضیه را آزمون کنید همچنین یک فاصله اطمینان ٪95 برای $\mu_1 - \mu_2$ به دست آورید (با فرض نرمال)

$n_1 = 12 \quad \bar{x} = 249 \quad S_1 = 19$

$n_2 = 15 \quad \bar{y} = 233 \quad S_2 = 45$

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

حل: جامعه نرمال است، حجم نمونه کوچک است و واریانس ها مجهول اند ←

$\sigma_1^2 = \sigma_2^2$

$\sigma_1^2 \neq \sigma_2^2$ ✓

TAHERIAN

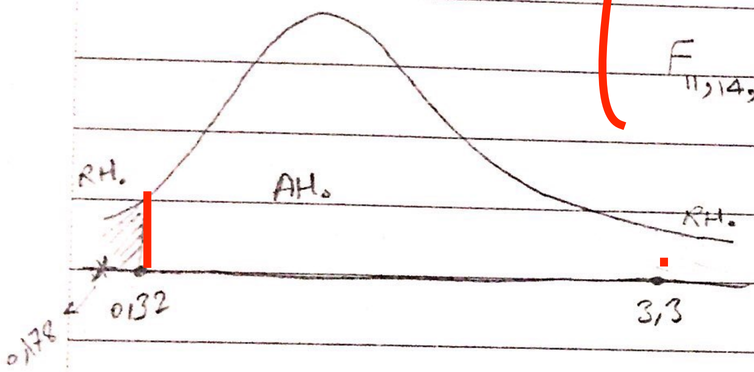
پس اولین گام این است که دو جامعه از نظر پراکندگی بررسی شوند:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases} \qquad F_0 = \frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}$$

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

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

$$F_{11,14,0{,}975} = \frac{1}{\underbrace{F_{14,11,0{,}025}}_{3{,}1}} = 0{,}32$$



$$0{,}178 < 0{,}32 \Longrightarrow RH_0$$

پس دو جامعه از نظر پراکندگی یکسان نیست.

پس از زیر حالت دوم استفاده می‌شود:

فاصله اطمینان:

$$\mu_1 - \mu_2 : (\bar{X} - \bar{Y}) \pm t^*_{\min\{\overset{11}{(n_1-1)}, \overset{14}{(n_2-1)}\}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$= (249 - 233) \pm t_{11, 0{,}025} \sqrt{\frac{361}{12} + \frac{2025}{15}} \simeq (12{,}28 \,,\, 44{,}28)$$

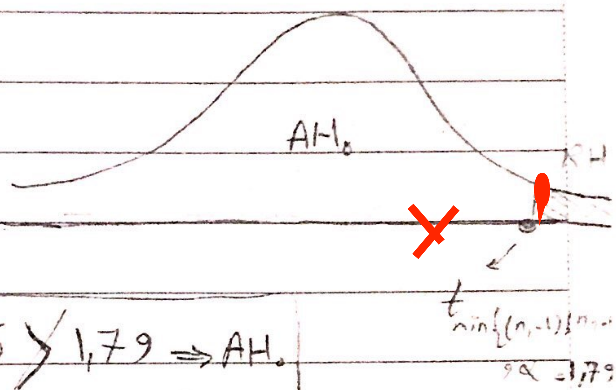
$\mu_1 - \mu_2 > 0$

چون هم کران پایین و هم کران بالای فاصله مثبت است و شامل صفر نیست پس نتیجه میگیریم

با اطمینان ۹۵٪ میانگین جامعه اول از میانگین جامعه دوم بزرگتر است که از اول این را نتیجه‌گیری کنیم

$$T_0 = \frac{(\bar{X}-\bar{Y})-\delta_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} = \frac{(249-233)-0}{\sqrt{\frac{19^2}{12}+\frac{45^2}{15}}} \simeq 1{,}25$$

$$\begin{cases} H_0: \mu_1 = \mu_2 \rightarrow \mu_1-\mu_2 = \overset{\delta_0}{0} \\ H_1: \mu_1 > \mu_2 \rightarrow \mu_1-\mu_2 > 0 \end{cases}$$



$$t_{11,0{,}05} = 1{,}7959$$

$$\boxed{1{,}25 \not> 1{,}79 \Rightarrow AH_0}$$

در این آزمون یک آزمون یک‌طرفه است و نمی‌توان طبق فاصله اطمینان آن را بررسی کرد.

مثال) آماره‌های خلاصه زیر برای نمونه‌های تصادفی مستقل دو جامعه ثبت شده‌اند با فرض نرمال بودن جوامع فاصله اطمینان ۹۵٪ را برای $\mu_1-\mu_2$ بدست آورید و فرضیه

$$\begin{cases} H_0: \mu_1 \le \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$$

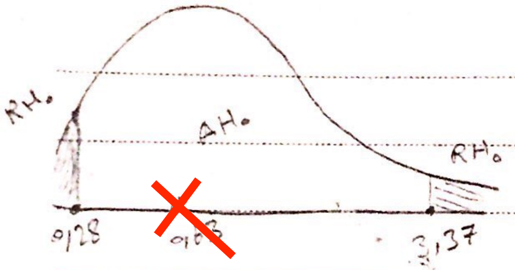
را در سطح ۰٫۰۵ آزمون کنید

$$\left.\begin{matrix} n_1 = 11 & \bar{X} = 10{,}7 & S_1^2 = 1{,}36 \\ n_2 = 13 & \bar{Y} = 9{,}6 & S_2^2 = 2{,}17 \end{matrix}\right\}$$

ابتدا فرضیه برابری واریانس‌ها بررسی می‌شود:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases} \qquad F_0 = \frac{S_1^2}{S_2^2} = \frac{1{,}36}{2{,}17} = 0{,}63$$

$$F_{n_1-1, n_2-1, \frac{\alpha}{2}} = F_{10,12,0{,}025} = 3{,}37 \rightarrow F_{10,12,0{,}975} = \frac{1}{F_{12,10,0{,}025}} = \frac{1}{3{,}62} = 0{,}2$$



$3,37 > 0,63 > 0,28 \Rightarrow AH_0$

پس از زیر حالت اول استفاده می‌کنیم یعنی (جوامع نرمال، n ها کوچک، واریانس ها مجهول ولی برابر):

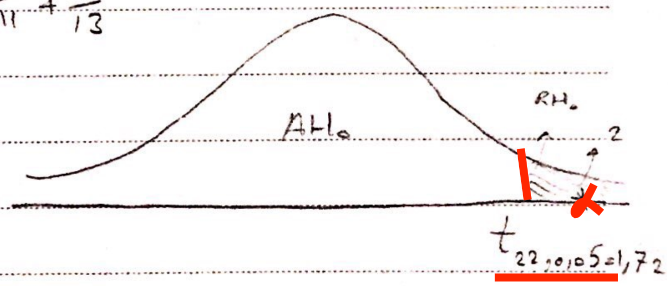
$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{10 \times 1,36 + 12 \times 2,17}{11+13-2} \simeq 1,8$$

$$\mu_1-\mu_2 : (\bar{x}-\bar{y}) \pm t_{\underbrace{n_1+n_2-2}_{22},\ \underbrace{\frac{\alpha}{2}}_{0,025}} \cdot S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$= (10,7-9,6) \pm 2,074 \sqrt{1,8}\ \sqrt{\frac{1}{11}+\frac{1}{13}} = \cancel{(-0,04\ ,\ 2,24)}$$

$$T_0 = \frac{(\bar{x}-\bar{y})-\delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(10,7-9,6)-0}{\sqrt{1,8}\ \sqrt{\frac{1}{11}+\frac{1}{13}}} \simeq 2$$

$2 > 1,72 \Rightarrow RH_0$



مثال ۱) با توجه به آماره های زیر فرضیه $\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 < P_2 \end{cases}$ و $\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases}$ را آزمون کنید.

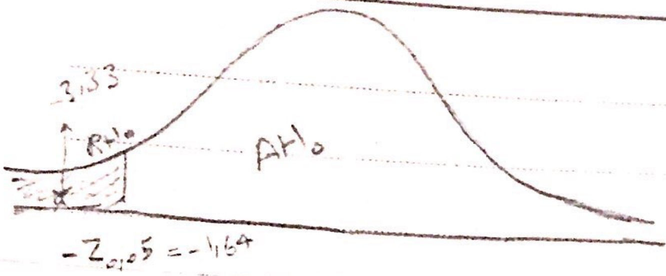
فرض اول را از روش p_value نیز آزمون کنید.

$n_1 = 100 \quad \hat{P}_1 = 0,5$

$n_2 = 200 \quad \hat{P}_2 = 0,7$

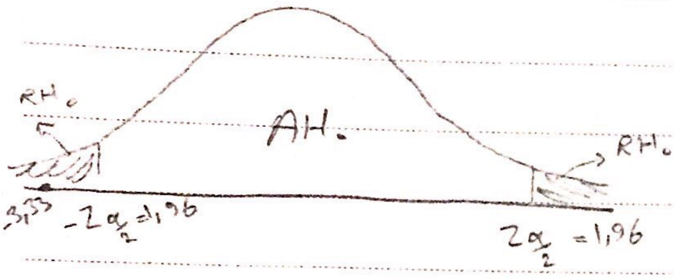
$$\hat{P} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1+n_2} = \frac{(100 \times 0,5 + 200 \times 0,7)}{100+200} \simeq 0,63$$

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0,5-0,7}{\sqrt{0,63(1-0,63)\left(\frac{1}{100}+\frac{1}{200}\right)}} \simeq -3,33$$



$-3.33 < -1.64 \Rightarrow RH_0$

برای فرضیه دوم :



پس برای فرضیه $\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases}$ هم H_0 رد می‌شود

$P\text{-Value} = P(Z < -3.33) = \Phi(-3.33) = 0.0004 < 0.05 \Rightarrow RH_0$

مثال) برای تاثیرات (اثرپذیری) تبلیغات صدحاجی دانشگاه 5 تن از کارمندان را در نظر گرفت و وزن ایشان را قبل از تبلیغات و بعد از تبلیغات اندازه گرفت، یک فاصله اطمینان 95% برای تغییر وزن بیابید. آیا تبلیغات تاثیری روی وزن داشته است؟ چون نمونه تغییری نکرده پس دو نمونه مستقل نداریم پس از استنباط زوجی استفاده می‌کنیم.

قبل از تبلیغات	84	97	78	91	85
بعد " "	80	98	75	90	82

$D_i =$ قبل - بعد $\quad -4 \quad 1 \quad -3 \quad -1 \quad -3 \rightarrow \bar{D} = \frac{1}{5}(-4+1-3-1-3) = -2$

$(D_i - \bar{D})^2 = (-4+2)^2 = 4 \quad 9 \quad 1 \quad 1 \quad 1 \rightarrow S_D^2 = \frac{1}{4}(4+9+1+1+1) = 4$

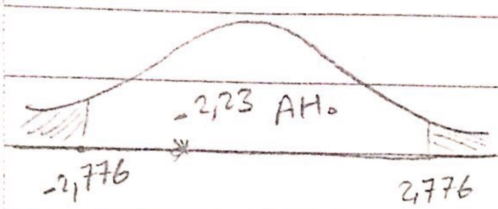
$\delta : \bar{D} \pm t_{4, 0.025} \frac{S_D}{\sqrt{n}} \quad : -2 \pm 2.776 \frac{2}{\sqrt{5}} = (-4.48, 0.48)$

صفر در بازه اطمینان هست پس $\delta = 0$ پذیرفته می‌شود

Year : Month : Date :

Sa Su Mo Tu We Th Fr

$$T_0 = \frac{\bar{D}}{\frac{S_D}{\sqrt{n}}} = \frac{-2}{\frac{2}{\sqrt{5}}} = -2,23 \qquad \begin{cases} H_0: \delta = 0 \\ H_1: \delta \neq 0 \end{cases}$$



$\rightarrow -2,23 > -2,776 \Rightarrow AH_0$

مثال) باتوجه به اطلاعات زیر فرضیه $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$ دا اگر برابریم واریانس ها معلوم

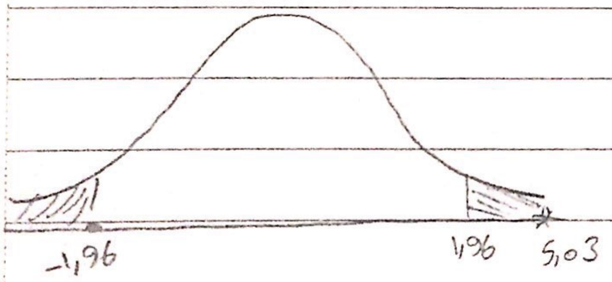
را آزمون کنید.

$n_1 = 36 \quad \bar{X} = 81 \quad \sigma_1 = 5,2$

$n_2 = 49 \quad \bar{Y} = 76 \quad \sigma_2 = 3,4$

چون واریانس ها معلوم است از Z استفاده می کنیم:

$$Z_0 = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{81 - 76 - 0}{\sqrt{\frac{(5,2)^2}{36} + \frac{(3,4)^2}{49}}} = 5,023$$



پس چون $5,03 > 1,96$ فرض صفر را رد می کنیم