Subject:
Date

- میانگین

فرض کنید یک ویژگی برای n آزمودنی اندازه گیری شده است

$$x_1 \dots x_n$$

arithmetic mean

در این صورت میانگین حسابی $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$ می باشد

میانگین هندسی ← $\sqrt[n]{x_1 \cdot x_2 \dots x_n}$

geometric mean

~~کمینه~~ میانه برای محاسبه ی میانه داده ها را مرتب کنید اگر تعداد داده ها فرد باشد عدد وسط میانه است اگر زوج باشد دو عدد وسط را جمع و تقسیم بر ۲ کنید.

واریانس

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

انحراف معیار یا انحراف استاندارد

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

Subject:
Date

مثال ۸ نتایج پنج تیتراسیون توسط چهار دانشجو به شرح زیر است

A	10.08	10.11	10.09	10.10	10.12
B	9.88	10.14	10.02	9.80	10.21
C	10.19	9.79	9.69	10.05	9.78
D	10.04	9.98	10.02	9.97	10.04

میانگین حسابی و واریانس تیتراسیون های دانشجوی A را محاسبه کنید

$$\bar{X}_A = \frac{10.08 + 10.11 + 10.09 + 10.10 + 10.12}{5} = 10.1 \ ml$$

$$S^2 = ?$$

$$\sum_{i=1}^{5}(x_i - \bar{x})^2 = (10.08 - 10.1)^2 + (10.11 - 10.1)^2 + (10.09 - 10.1)^2$$

$$+ (10.10 - 10.1)^2 + (10.12 - 10.1)^2 = 0.0010$$

$$S^2 = \frac{1}{4} \times 0.0010 = 0.00025$$

$$S = \sqrt{0.00025} = 0.0158 \ ml$$

ضریب تغییرات عبارت است از $\frac{S}{\bar{X}}$ که معمولاً به درصد بیان می‌شود

در مثال فوق $\frac{0.0158}{10.1} \cong 0.0016$

$$0.0016 \times 100 = 0.16\%$$

تمرین در مثال ۸ برای دانشجو B میانگین که انحراف معیار و ضریب تغییرات را بدست آورید

P4PCO

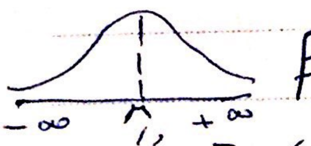
Scanned with CamScanner

```
> xA=c(10.08, 10.11, 10.09, 10.10, 10.12)
> mean(xA)
[1] 10.1
> var(xA)
[1] 0.00025
> sd(xA)
[1] 0.01581139
> sd(xA)/mean(xA)
[1] 0.001565484
>
```

ص ۲۱ کتاب میلر

Normal or Gaussian — توزیع نرمال (گاوسی)

در عمل هدف بررسی یک ویژگی است اگر این ویژگی از آزمودنی به آزمودنی دیگر تغییر کند متغیر تصادفی نامیده می‌شود. مقیاس اندازه گیری یا متغیر تصادفی کمّی باشد یعنی دامنهٔ تعریف آن بازه یا اجتماعی از بازه ها در اعداد حقیقی باشد متغیر تصادفی پیوسته نامیده می‌شود. متغیرهای تصادفی با توجه به ماهیتی که دارند دارای توزیع احتمال هستند. در این درس با چند توزیع مهم آماری آشنا می‌شوید. یکی از مهم ترین این توزیع ها توزیع نرمال است. فرض کنید X یک متغیر تصادفی پیوسته باشد که دامنه تعریف آن $(-\infty, \infty)$ و تابع چگالی احتمال آن به صورت زیر باشد:



شکل

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

و با نماد $X \sim N(\mu, \sigma^2)$ نشان می‌دهیم که به μ میانگین یا پارامتر مکان X می‌گویند (به آن امید ریاضی هم می‌گوییم)

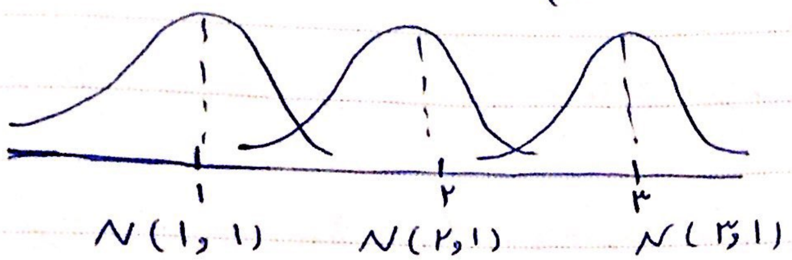
$E(X) = \mu \longrightarrow$ میانگین

و به σ^2 پارامتر شکل یا مقیاس یا واریانس می‌گویند

$V(X) = \sigma^2$

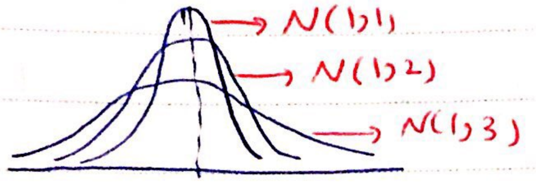
Subject:
Date

پس با تغییر پارامتر مکان، مکان توزیع و با تغییر پارامتر شکل، شکل توزیع تغییر می کند.

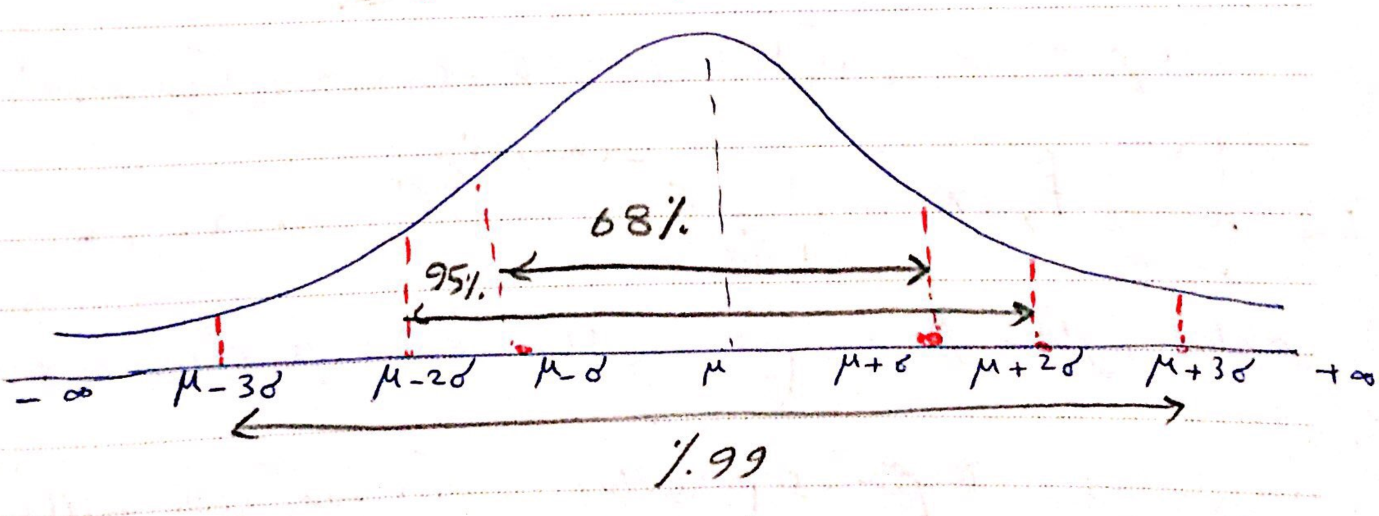


و با تغییر مقیاس شکل تغییر می‌کند

دیر پراکندگی را کنترل می‌کند



همان طور که دیدیم شکل توزیع زنگی شکل و متقارن است. دامنه ی تعریف کل اعداد حقیقی است. خاصیت خوب دیگری که دارد بازه ی بندی توزیع مشخص است.



$P(\mu-\sigma < X < \mu+\sigma) = ۰/۶۸$

$P(\mu-۲\sigma < X < \mu+۲\sigma) \doteq /95$

$P(\mu-۳\sigma < X < \mu+۳\sigma) = 0/99$

PAPCO

قضیه حد مرکزی. اگر X متغیر تصادفی از هر توزیعی مثل F باشد با میانگین μ و واریانس σ^2 و $X_1 \dots X_n$ نمونه تصادفی به حد کافی بزرگ از X باشد آن گاه $\bar{X}$ دارای توزیع نرمال است

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

$$\text{if} \quad \underset{\text{هر توزیعی}}{X \sim F}(\mu, \sigma^2) \xrightarrow{X_1 \dots X_n} \bar{X} \approx N(\mu, \frac{\sigma^2}{n})$$

n معمولاً بزرگتر از 30

توزیع نرمال استاندارد. هر توزیع نرمالی را می توان به توزیع نرمال استاندارد تبدیل کرد فرض کنید $X \sim N(\mu, \sigma^2)$ آن گاه

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

Z یک متغیر تصادفی نرمال استاندارد است یعنی میانگین صفر و واریانس یک است. برای توزیع نرمال استاندارد چون میانگین و واریانس آن معلوم است جدول هایی برای مقادیر توزیع آن مشخص شده است

$$P(Z < 1.96) = 0.975$$

$$P(Z < 1.645) = 0.95$$

$F(1.64) = 0.95$

$F(1.96) = 0.975$

مثال ۹ مقادیر تکراری یک تیتر اسیون بلاک میانگین 10.15^{ml} و انحراف معیار (انحراف استاندارد) Sd
توزیع نرمال با
0.02^{ml} است. مطلوبست این که اندازه‌ها بین 10.12^{ml} و 10.2^{ml} باشند
را پیدا کنید. پس $X \sim N(10.15, 0.02)$

$$P(10.12 < X < 10.2) = P(X < 10.2) - P(X < 10.12)$$

$$\overset{\text{استاندارد}}{=} P\Big(\underbrace{\frac{X-10.15}{0.02}}_{Z} < \frac{10.2-10.15}{0.02}\Big)$$

$$- P\Big(Z < \frac{10.12-10.15}{0.02}\Big)$$

$$= \underbrace{F(2.5)}_{pnorm(2.5)} - \underbrace{F(-1.5)}_{pnorm(1.5)}$$

in R ←

$$= 0.9938 - 0.0668 = 0.927$$

مثال ۱۰. در جدول زیر غلظت ۵۰ نمونه از یون نیترات ثبت شده است
۵۰ تکرار ← غلظت یون نیترات در محلول اندازه‌گیری شود

میلی‌لیتر میکروگرم
$\mu g/ml$

0.51	0.51	0.51	0.50	0.51	0.49	0.52	0.53
0.50	0.47	0.51	0.52	0.53	0.48	0.49	0.50
0.52	0.49	0.49	0.50	0.49	0.48	0.46	0.49
0.49	0.48	0.49	0.49	0.51	0.47	0.51	0.51
0.51	0.48	0.50	0.47	0.50	0.51	0.49	0.48
0.51	0.50	0.50	0.53	0.52	0.52	0.50	0.50
0.51	0.51						

- هیستوگرام داده ها را رسم کنید
- فراوانی هر داده را مشخص کنید
- میانگین ۵۰ نمونه را محاسبه کنید
- انحراف استاندارد ۵۰ نمونه را محاسبه کنید

فاصله اطمینان Confidence Interval (یکی از کاربردهای توزیع نرمال)

حالت ۱) فرض کنید X ویژگی مورد بررسی است (مثل غلظت یون نیترات) و می خواهیم براساس نمونه یک بازه ی اطمینان برای میانگین غلظت بدست آوریم.

با فرض این که نمونه بزرگتر از ۳۰ باشد

(طبق ق. ح. م) — قضیه حد مرکزی — می توان فاصله اطمینان را برای میانگین (μ) به صورت زیر بدست آوریم

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

n

$$\begin{cases} S \rightarrow s.d = \sqrt{S^2} \\ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \\ Z_{\alpha/2} \rightarrow \text{جدول} \\ 1-\alpha = \text{سطح اطمینان که ما می دهند} \end{cases}$$

Subject:
Date

مثال ۱۱. در مثال غلظت یون نیترات فاصله اطمینان ۹۵٪ و ۹۹٪ را

برای میانگین غلظت یون نیترات بدست آورید

$1-\alpha = 0/95 \rightarrow \alpha = 0/05 \rightarrow \alpha/2 = 0/025$

$Z_{0/975} = 1.96 \qquad 1-\alpha/2 = 0.975$

$1-\alpha = 0/99 \rightarrow \alpha = 0/01 \rightarrow \alpha/2 = 0.005$

$1-\alpha/2 = 0.995$

$Z_{0.995} = 2.58 \qquad \bar{x} = 0.5 \quad s.d = 0.0165$

$\bar{x} \pm Z_{1-\alpha/2}\frac{S}{\sqrt{n}} = 0.5 \pm 1.96 \cdot \frac{0.0165}{\sqrt{50}}$: 95%

$= 0.5 \pm 0.0046 = (0.4954 \,, 0.5046)$ mg/ml

$\bar{x} \pm Z_{1-\alpha/2}\frac{S}{\sqrt{n}} = 0.5 \pm 2.58 \, \frac{0.0165}{\sqrt{50}}$ %99

$= 0.5 \pm 0.0060 = (0.494 \,, 0.5006)$ mg/ml

0.5006

حالت دوم فرض کنید در جامعه X دارای توزیع نرمال باشد اما $n < 30$ (حجم نمونه

کوچک باشد) در اینصورت فاصله اطمینان برای میانگین X :

$$\boxed{\mu : \bar{x} \pm t_{n-1,\,1-\alpha/2}\frac{S}{\sqrt{n}}}$$

n ↓ نمونه کوچک

جامعه نرمال

که در آن t مقادیر تابع توزیع t_استودنت است که از جدول های

PAPCO

Scanned with CamScanner

Subject :
Date

تی - استودنت قابل استخراج است. ابتدا به معرفی توزیع χ^2 (کای دو یا خی دو) و سپس t- استودنت می پردازیم.

توزیع کای دو. اگر X: یک متغیر پیوسته باشه که دارای توزیع نرمال است $X \sim N(\mu, \sigma^2)$ آن گاه $X_1 \dots X_n$ هم نمونه ای از X باشد در اینصورت هرکدام از $X_i \sim N(\mu, \sigma^2)$ و $Z_i = \frac{X_i - \mu}{\sigma}$ دارای نرمال استاندارد است توزیع کای از به توان دو رساندن Z_i بدست می آید $Z_i^2 \sim \chi^2_1$ ← درجه آزادی

با جمع چندکای دو بازهم توزیع کای دو است فقط درجه آزادی جمع می شود

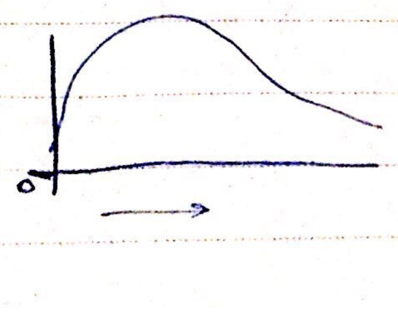
$$Y = \sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$$

نکته. اگر در مسئله ای به جای مقدار واقعی یک پارامتر (مثلاً μ) از مقدار نمونه ($\bar{X}$ (برآورد μ)) استفاده شود از درجه آزادی یک واحد کم می شود و به هر تعداد که به جای مقدار واقعی از مقدار نمونه استفاده کنید از درجه آزادی کم کنید.

$E(Y) = n \qquad V(Y) = 2n$

میانگین یا امید ریاضی

چون توان ۲ است پس Y ها مثبت هستند و شکل توزیع ما



Subject:
Date

نکته مهم: ثابت می شود که

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

$X \sim N(\mu, \sigma^2)$ $\quad$ $S^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2$

$\bar{X}$ به جای μ

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

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

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

توزیع t ـ استودنت

مثل نرمال متقارن است

$$T = \frac{Z}{\sqrt{Y/df}} \sim t_{df}$$

درجه آزادی

$Y \sim \chi^2_{df}$ درجه آزادی

$Z \sim N(0,1)$

نکته مهم: ثابت می شود که $\underbrace{\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}}$ $n \downarrow$ و اگر n بزرگ

باشد $\underbrace{\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim Z}$ نرمال $n \uparrow$

مثال ۱۲. میزان جذب یک طیف سنج در طول موج خاصی با یک محلول استاندارد که دارای جذب 0.47 است آزمایش شود است ۱۰ بار تکرار و میانگین جذب

طیف سنج $\bar{X} = 0.461$ و $S = 0.003$ بدست آمد. اکنون

یک فاصله اطمینان ۹۵٪ برای میانگین جذب توسط طیف سنج بیابید آیا با توجه به این نمونه طیف سنج دارای خطا است؟

$n = 10 \downarrow \qquad \mu : \bar{x} \pm t_{n-1} \frac{s}{\sqrt{n}}$

$\mu : 0.461 \pm 2.26 \times \frac{0.003}{\sqrt{10}} = 0.461 \pm 0.002$

$1-\alpha = 0.95$

$\Rightarrow 1-\alpha/2 = 0.975$

$= (0.459, 0.463)$

$t_{9, 0.975} = 2.26$

پس فاصله اطمینان شامل 0.47 نیست و خطا رخ داده است.

نکته: گاهی اوقات ویژگی مورد بررسی نرمال نیست اما لگاریتم ویژگی دارای توزیع نرمال است در این صورت ابتدا از نمونه های بدست آمده لگاریتم بگیرید فاصله اطمینان را بدست و مجدداً فاصله اطمینان را برای ویژگی اصلی تبدیل کنید به عبارت دیگر وقتی لگاریتم گرفته می شود فاصله اطمینان بدست آمده برای لگاریتم میانگین هندسی است

$\log \sqrt[n]{x_1 \cdots x_n} = \frac{1}{n}(\log x_1 + \cdots + \log x_n)$

مثال ۱۲: مقادیر زیر غلظت آنتی بادی موجود در سرم خون انسان به درصد برای هشت نفر است فاصله اطمینان ۹۵٪ را برای میانگین هندسی بدست آورید

2.15 1.13 2.04 1.45 1.35 1.09 0.99 2.07

Date

ابتدا لگاریتم (در مبنای ۱۰) بگیرید

0.332 0.053 0.310 0.161

0.130 0.037 -0.004 0.316

میانگین بگیرید

$$\frac{0.332+\dots+0.316}{8} = 0.1669$$

انحراف معیار بگیرید: $\frac{(0.332-0.1669)^2 + \dots}{7} = 0.1365$

فاصله برای لگاریتم میانگین هندسی

$$: \cancel{0.17}^{\,0.1669} \pm 2.36 \frac{0.1365}{\sqrt{8}} = (0.053, 0.2808)$$

$$t_{7, 0.975} = 2.36$$

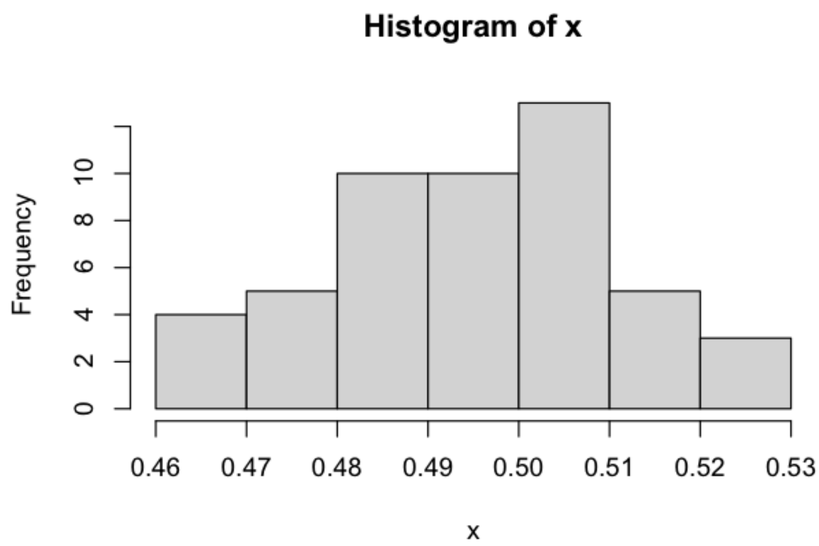
اکنون برای میانگین هندسی مقادیر اصلی

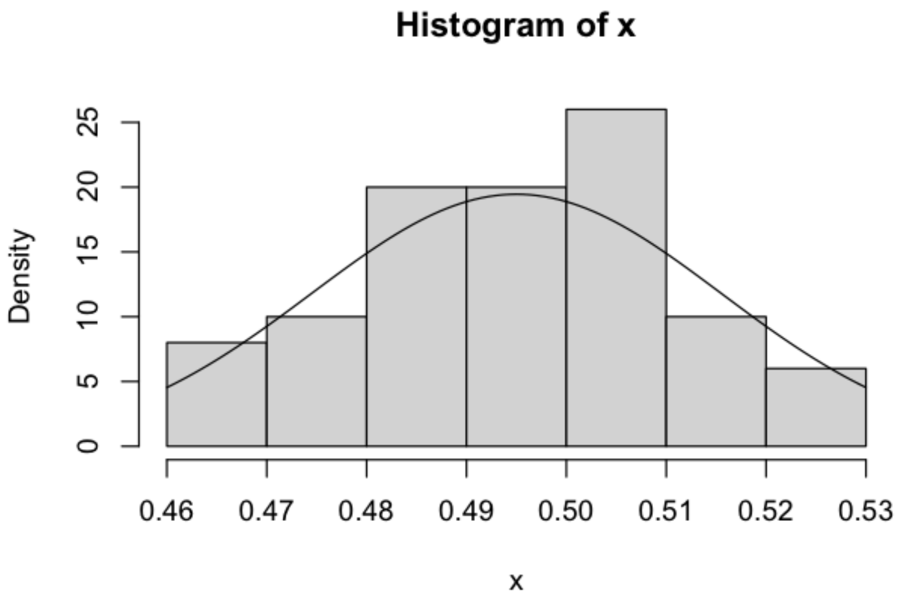
$$10^{0.053} = 1.13 \qquad 10^{0.2808} = 1.91$$

پس فاصله اطمینان برای میانگین هندسی $(1.13, 1.91)$ است.

مثال ١٠

```
x=c(0.51,0.51,0.51,0.50,0.51,0.49,0.52,0.53,0.50,0.47,0.51,0.52,0.53,0.48,0.49,0.50,0.52,0.49,0.49,0.50,0.49,0.48,0.46,0.49,0.49,0.48,0.49,0.49,0.51,0.47,0.51,0.51,0.51,0.48,0.50,0.47,0.50,0.51,0.49,0.48,0.51,0.50,0.50,0.53,0.52,0.52,0.50,0.50,0.51,0.51)
> hist(x)
> hist(x,prob=TRUE)
> curve(dnorm(x, mean=mean(x), sd=sd(x)), add=TRUE)
> table(x)
x
0.46 0.47 0.48 0.49  0.5 0.51 0.52 0.53
   1    3    5   10   10   13    5    3
> mean(x)
[1] 0.4998
> sd(x)
[1] 0.01647385
> n=length(x)
>
```





مثال ١١

```
> CIU=mean(x)+qnorm(0.975)*sd(x)/sqrt(n)
> CIU
[1] 0.5043662
> CIL=mean(x)-qnorm(0.975)*sd(x)/sqrt(n)
> CIL
[1] 0.4952338
>
> CIU=mean(x)+qnorm(0.995)*sd(x)/sqrt(n)
> CIU
[1] 0.505801
> CIL=mean(x)-qnorm(0.995)*sd(x)/sqrt(n)
> CIL
[1] 0.493799
>
> |
```

مثال ١٢

```
> meanx=0.461
> sdx=0.003
> n=10
> CIU=meanx+qt(0.975,9)*sdx/sqrt(n)
> CIU
[1] 0.4631461
> CIL=meanx-qt(0.975,9)*sdx/sqrt(n)
> CIL
[1] 0.4588539
```

مثال ١٣

```
> x=c(2.15,1.13,2.04,1.45,1.35,1.09,0.99,2.07)
> y=log10(x)
> n=length(x)
> logCIU=mean(y)+qt(0.975,7)*sd(y)/sqrt(n)
> logCIU
[1] 0.2810972
> logCIL=mean(y)-qt(0.975,7)*sd(y)/sqrt(n)
> logCIL
[1] 0.05287302
> CIU=10^(logCIU)
> CIU
[1] 1.910281
> CIL=10^(logCIL)
> CIL
[1] 1.129466
```

Subject:
Date

ترکیبات خطی :

گاهی اوقات ویژگی مورد بررسی به صورت ترکیب خطی از چند ویژگی دیگر

بیان می شود مثلاً $y = k + k_a a + k_b b + \dots$

به طوری که k ، k_a ، k_b ... ثابت هستند و

$var(a) = \sigma_a^2$

$var(b) = \sigma_b^2$ ....

در این صورت

با شرط استقلال a و b ... $\rightarrow var(y) = k_a^2 \sigma_a^2 + k_b^2 \sigma_b^2 + \dots$

و انحراف استاندارد $sd \rightarrow \sigma_y = \sqrt{k_a^2 \sigma_a^2 + \dots}$

مثال ۱۴! در یک تیتراسیون مقدار اولیه روی بورت 3.51^{ml} و مقدار نهایی 15.67^{ml}

ثبت شده است اگر هر دو دارای انحراف معیار 0.02^{ml} باشند حجم تیترانت به کار

رفته چقدر است و انحراف معیار آن را بدست آورید.

$$\text{حجم} = 15.67 - 3.51 = 12.16$$

$$S.d = \sqrt{(0.02)^2 + (-0.02)^2} = 0.028^{ml}$$

نکته: اگر ویژگی مورد بررسی از رابطه ای به صورت

$$y = k\frac{ab}{cd}$$

بدست بیاید که K ثابت و a، b، c و d مستقل باشند در اینصورت

انحراف معیار Y

$$\frac{\sigma_Y}{Y} = \sqrt{\left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_b}{b}\right)^2 + \left(\frac{\sigma_c}{c}\right)^2 + \left(\frac{\sigma_d}{d}\right)^2}$$

انحراف معیار نسبی

relative s.d = RSD

مثال ۱۵: بازده کوانتومی فلورسانس Φ از رابطه ی زیر بدست می آید

$$\Phi = I_l / Kcl\, I_0\, \varepsilon$$

که در آن (داخل پرانتز انحراف معیار نسبی) ←

I_0 = شدت نور تابشی (0.5 %)

I_l = شدت نور فلورسانس (2 %)

ε = ضریب جذب مولی (۱ %)

c = غلظت (0.2 %)

l = طول مسیر (0.2 %)

K = ثابت ساز

انحراف معیار نسبی Φ را بدست آورید.

$$RSD = \sqrt{0.5^2 + 2^2 + 1^2 + 0.2^2 + 0.2^2} = 2.3\%$$

نکته: اگر $Y = b^n$ آنگاه $\left|\frac{\sigma_Y}{Y}\right| = \left|n\frac{\sigma_b}{b}\right|$

Subject:
Date

نکته اگر Y تابعی از X باشد ، $Y = f(x)$ آنگاه انحراف استاندارد X و

Y به صورت زیر ارتباط دارند $\sigma_Y = \left| \sigma_X \frac{dY}{dx} \right|$

مثال ۱۶ جذب A یک محلول از $A = -\log T$ به دست می آید که در آن

T بیانگر مقدار عبور است اگر مقدار $T = 0.501$ باشد با انحراف استاندارد 0.001

انحراف استاندارد A را بیابید.

$$\frac{dA}{dT} = \frac{-\log e}{T}$$

$$\sigma_A = \left| \sigma_T (-\log e / T) \right| = \left| 0.001 \times (-0.434 / 0.501) \right|$$

$$\log e = \log 2{,}72 = 0.434 \qquad = 0.0008$$

انحراف معیار نسبی $RSD = \frac{\sigma_A}{A} = \frac{\sigma_T \log e}{T \log T} \xrightarrow[\text{با مشتق گیری}]{\min} T = \frac{1}{e} = 0.368$

و مساوی صفر نشان داد که مینیمم در ۳۶٪ رخ می دهد.

«آزمون فرض ها»

گاهی در عمل می خواهیم آزمون کنیم که میانگین مقدار واقعی ویژگی مورد بررسی

برابر مقدار خاصی است یا خیر. معمولاً آن چیزی که به تجربه در جامعه برقرار

است را فرضیه صفر و آنچه که در مورد ویژگی ادعا می شود فرض مقابل می گویند

فرض صفر را با H_0 و فرض مقابل را با H_1 نشان می دهند.

شکل کلی نوشتن فرضیات به صورت

$$\begin{cases} H_0: \mu = \mu_0 & \mu \geq \mu_0 & \mu \leq \mu_0 \\ H_1: \underbrace{\mu \neq \mu_0}_{\text{دوطرفه}} \text{ یا } \overbrace{\mu < \mu_0}^{\text{یکطرفه}} \text{ یا } \overbrace{\mu > \mu_0}^{\text{یکطرفه}} \end{cases}$$

است. اگر n بزرگ باشد

$$Z_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad \text{(توزیع نرمال استاندارد)}$$

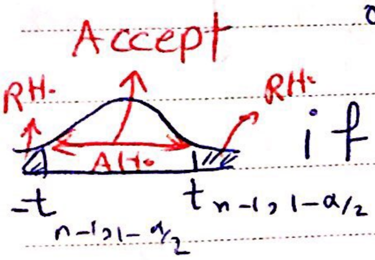
آماره‌ی آزمون است و اگر n کوچک و جامعه نرمال باشد

$$t_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad \text{(توزیع t استودنت)}$$

آماره‌ی آزمون است.

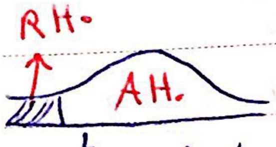
(μ میانگین جامعه ، $\bar{X}$ میانگین نمونه ، s انحراف معیار نمونه ، n حجم نمونه)

برای تصمیم گیری اگر فرض مقابل $\overbrace{\mu \neq \mu_0}^{\text{دوطرفه}}$ باشد آن گاه



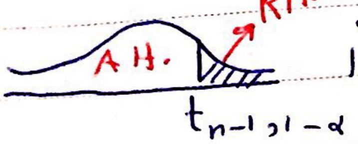
if $|t_0| > \underbrace{t_{n-1,1-\alpha/2}}_{\text{مقدار بحرانی}} \Rightarrow \underbrace{\text{Reject } H_0}_{RH_0}$

اگر فرض مقابل $\mu < \mu_0$ باشد:



if $t_0 < -t_{n-1,1-\alpha} \Rightarrow RH_0$

اگر فرض مقابل $\mu > \mu_0$ باشد:



if $t_0 > t_{n-1,1-\alpha} \Rightarrow RH_0$

(α سطح آزمون است که در مسئله می دهند)

Subject:
Date

مثال ۱۷ در یک روش جدید برای تعیین سلنو آوود در آب ، مقادیر زیر در آب لوله کشی

با $50 \ ng/ml$ سلنو آوده بدست آمده است . آیا شواهدی مبنی بر خطای سیستماتیک

وجود دارد. (یعنی $\begin{cases} H_0: \mu = 50 \\ H_1: \mu \neq 50 \end{cases}$)

50.4 50.7 49.1 49 51.1

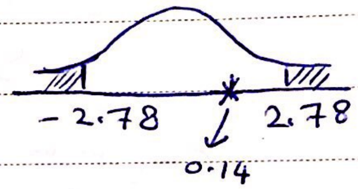
$$t_0 = \frac{50.6 - 50}{\frac{0.956}{\sqrt{5}}} = \boxed{0.14}$$

$$\bar{X} = \frac{50.4 + \dots + 51.1}{5} = 50.6$$

$$S = 0.956$$

$$\alpha = 0.05 \rightarrow 1 - \alpha/2 = 0.975$$

$$t_{4, 0.975} = \boxed{2.78}$$



p-value = ~~pt...~~ $p(t > 0.14) = 0.851 \not< 0.05$

$0.14 \not> 2.78 \Rightarrow A H_0$

خطای سیستماتیکی وجود ندارد.

in R

```
X = c(50.4, ..., 51.1)
t.test(x, mu=50, alternative="two.sided", conf.level=0.95)
                           "less" or "greater"
```

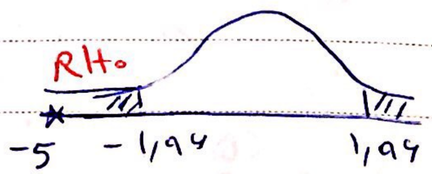
مثال ۱۸ در مثال ۱۰ (غلظت یون نیترات) ، $\begin{cases} H_0: \mu = 0.51 \\ H_1: \mu \neq 0.51 \end{cases}$ را آزمون کنید

$$\bar{X} = 0.5$$

$$s.d = 0.0165$$

$$Z_0 = \frac{0.5 - 0.51}{\frac{0.0165}{\sqrt{50}}} = \frac{-0.01}{0.002} \simeq -5 \rightarrow RH_0$$

$$Z_{0.975} = 1.96$$



(در نرم افزار همان t.test انجام می شود چون وقتی n بزرگ است توزیع t استودنت همان نرمال می شود)

PAPCO

مقایسه‌ی دو گروه با هم (با استفاده از نمونه‌های مستقل)

گاهی در مسئله‌ی آزمون فرض هدف مقایسه‌ی میانگین دو گروه (دو جامعه) (دو تیمار) ... با هم است.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \quad \text{or } \mu_1 > \mu_2,\ \mu_1 < \mu_2 \end{cases}$$

حالت اول

در این گونه مسائل دو حالت داریم اگر دو گروه از نظر پراکندگی یکسان باشند $\sigma_1^2 = \sigma_2^2$

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$ → آماره آزمون (n_1 و n_2 لزوماً برابر نیستند)

که در آن $\bar{X}_1$ میانگین نمونه گروه اول، $\bar{X}_2$ میانگین نمونه گروه دوم،

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

n_1 حجم نمونه گروه اول و n_2 حجم نمونه گروه دوم است.

فرض دوطرفه $|t_0| > t_{n_1+n_2-2,\,1-\alpha/2} \Rightarrow RH_0$

برای بزرگتری $t_0 > t_{n_1+n_2-2,\,1-\alpha} \Rightarrow RH_0$

برای کوچکتری $\mu_1<\mu_2$: $t_0 < -t_{n_1+n_2-2,\,1-\alpha} \Rightarrow RH_0$

مثال ۱۹) در مقایسه‌ی دو روش برای تعیین مقدار کلرم در گلخه، نتایج زیر بدست آمد $mg/kg\ cr$

$n_1 = 5$ روش ۱: $\bar{X} = 1.48$ $mg/kg\ cr$ $S.d = 0.28$

$n_2 = 5$ روش ۲: $\bar{X} = 2.33$ $S.d = 0.31$

سطح آزمون $\boxed{\alpha = 0.05}$

آیا این دو روش دارای میانگین متفاوتی هستند.

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

$$S_p^2 = \frac{(5-1)\,0.28^2 + (5-1)(0.31)^2}{5+5-2} = 0.0873$$

$$S_p \simeq 0.295$$

$$t_0 = \frac{1.48 - 2.33}{0.295\sqrt{1/5+1/5}} = -4.54$$

$t_{8,\,0.025} = -2.31$ $\quad t_{8,\,0.975} = 2.31$

$|-4.54| > 2.31 \Rightarrow RH_0$

پس تفاوت معنی‌داری بین دو روش وجود دارد.

Scanned with CamScanner

مثال ۲۰ برای تعیین قلع در مواد غذایی، نمونه‌ها با اسید کلریدریک تحت رفلاکس در زمان‌های متفاوت جوشانده شده (فرض کنید واریانس در زمان‌های متفاوت یکسان است)

زمان رفلاکس	قلع پیدا شده mg/kg					
نمونه ← ۳۰	۵۹	۵۶	۵۶	۵۹	۵۷	۵۵
نمونه‌ی ← ۷۵	۵۹	۵۹	۵۹	۵۸	۵۵	۵۷

آیا مقادیر میانگین قلع در دو زمان رفلاکس با هم متفاوتند

۳۰ دقیقه — ۷۵ دقیقه

$\bar{X}_1 = 57 \qquad \bar{X}_2 = 57.83$

$S_1^2 = 2.80 \qquad S_2^2 = 2.57$

$$S_p^2 = \frac{(5\times 2.8 + 5\times 2.57)}{6+6-2} = 2.685$$

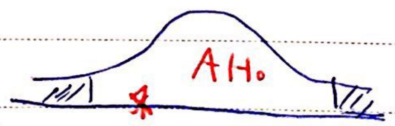
$S_p = 1.64$

$$t_0 = \frac{57 - 57.83}{1.64\sqrt{1/6 + 1/6}} = -0.88$$

$|0.88| \ngtr 2.23 \Rightarrow AH_0$

نکته: فاصله اطمینان

$$\mu_1-\mu_2: \quad (\bar{X}_1-\bar{X}_2) \pm S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \cdot t_{n_1+n_2-2,\,1-\alpha/2}$$



$-t_{10,\,0.975} = -2.23 \qquad 2.23$

$= (-2.941 , 1.274)$

↑ تمکیه

پس تفاوت معنی داری بین دو زمان وجود ندارد.

حالت دوم. اگر دو گروه از نظر پراکندگی یکسان نباشند در این صورت

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

از آماره‌ی

استفاده می‌شود که ثابت شده است دارای توزیع t - استودنت با

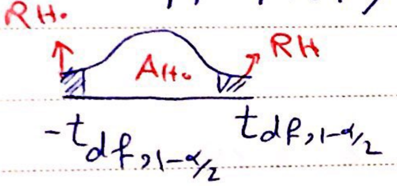
$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\left(\frac{S_1^4}{n_1^2(n_1-1)} + \frac{S_2^4}{n_2^2(n_2-1)}\right)}$$

درجه آزادی

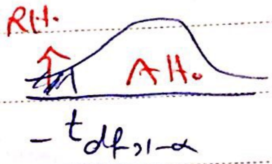
است و برای فرضیه‌های زیر:

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

$$\text{if } |t_0| > t_{df, 1-\alpha/2} \Rightarrow RH_0$$

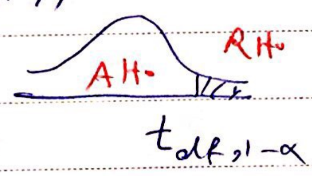


$$\begin{cases} H_0: \mu_1 \geq \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$$



$$\text{if } t_0 < -t_{df, 1-\alpha} \Rightarrow RH_0$$

$$\begin{cases} H_0: \mu_1 \leq \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases} \rightarrow \text{if } t_0 > t_{df, 1-\alpha} \Rightarrow RH_0$$



مثال ۲۱. غلظت تیول موجود در خون همولیز شده‌ی دو گروه (دو نمونه مستقل) اندازه‌گیری شد که گروه اول طبیعی و گروه دوم روماتوئید بودند. آیا از روی این داده‌ها می‌توان نتیجه گرفت که غلظت تیول برای دو گروه طبیعی و روماتوئید متفاوت است.

طبیعی	1.84	1.92	1.94	1.92	1.85	1.91	2.07
روماتوئید	2.81	4.06	3.62	3.27	3.27	3.76	

$n_1 = 7$, $n_2 = 6$

$$\alpha = 0.01 \rightarrow \alpha/2 = 0.005$$

$\bar{X}_1 = 1.921 \quad S_1 = 0.076 \quad n_1 = 7$

$\bar{X}_2 = 3.465 \quad S_2 = 0.440 \quad n_2 = 6$

$$df = \frac{\left(\frac{0.076^2}{7} + \frac{0.44^2}{6}\right)^2}{\left(\frac{0.076^4}{7^2(7-1)} + \frac{0.440^4}{6^2(6-1)}\right)} \cong 5.3 \simeq 5$$

(در برخی کتب $min(n_1-1, n_2-1)$ پیشنهاد می‌شود

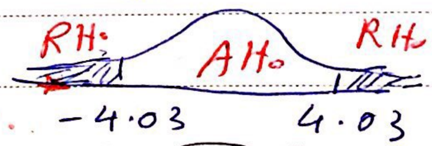
$min(\underset{6-1}{5}, \underset{7-1}{6}) = 5$

$$t_0 = \frac{1.921 - 3.465}{\sqrt{\frac{0.076^2}{7} + \frac{0.44^2}{6}}} = -8.48$$

$t_{5, 0.995} = 4.03$

or $t_{0.005}$

$|-8.48| > 4.03 \Rightarrow RH_0$



پس تفاوت معنی‌داری بین غلظت تیول موجود در خون دو گروه طبیعی و روماتوئید وجود دارد.

فاصله اطمینان → $\mu_1 - \mu_2 : (\bar{X}_1 - \bar{X}_2) \pm \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \cdot t_{df}$

$(1.921, 3.46)$

آزمون جفتی (زوجی) (paired t-test) ((نمونه‌های وابسته))

گاهی در عمل هدف مقایسه نتیجه آزمایش قبل و بعد روی یک آزمودنی است به عبارت دیگر مثل دو حالت قبل نسبت به دو گروه مجزا با دو نمونه مستقل مقایسه می‌شود. برای مثال فرض کنید اثر یک نوع دارو روی بیماران دارای فشارخون

مورد بررسی قرار می‌گیرد. در این مطالعه ابتدا فشار خون افراد نمونه اندازه‌گیری می‌شود سپس دارو تجویز و سپس از مدتی مجدد فشار خون همان افراد اندازه‌گیری می‌شود سپس دقت کنید نمونه‌ها عوض نشدند و برای همین می‌توانید تحلیل نمونه‌ای وابسته (زوجی) ابتدا تفاضل‌ها حساب می‌شوند

فشار خون قبل از مصرف دارو	فشار خون بعد از مصرف دارو	
x_1	y_1	$d_1 = x_1 - y_1$
$\vdots$	$\vdots$	$\vdots$
x_n	y_n	$d_n = x_n - y_n$

$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ میانگین تفاضل‌ها

واریانس تفاضل‌ها حساب می‌شود:

$$S_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

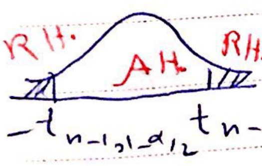
و در نهایت آماره آزمون:

$$t_0 = \frac{\bar{d}}{S_d/\sqrt{n}} \sim t_{n-1}$$

$H_0: \mu_d = 0$ → اختلاف معنی‌دار وجود ندارد

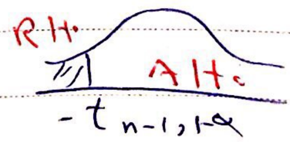
$H_1: \mu_d \neq 0$ → اختلاف معنی‌دار وجود دارد.

(μ_d میانگین تفاضل‌های جامعه)

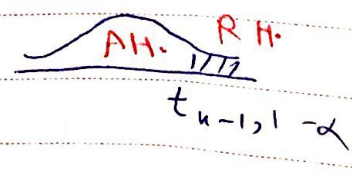


if $|t_0| > t_{n-1,1-\alpha/2} \Rightarrow RH_0$

if $t_0 < -t_{n-1,1-\alpha} \Rightarrow RH_0$: برای $\delta < 0$



if $t_0 > t_{n-1,1-\alpha} \Rightarrow RH_0$: برای $\delta > 0$



مثال ۲۲. فرض کنید می‌خواهیم غلظت داروی پاراستامول را با دو روش اندازه‌گیری و نتایج را مقایسه کنیم. ده قرص (از تولیدهای متفاوت) دو بار مورد اندازه‌گیری قرار گرفتند (یعنی نمونه عوض شده بلکه نمونه دو بار مورد آزمایش قرار گرفته است)

تولید	1	2	3	4	5	6	7	8	9	10
آزمایش طیف‌سنجی فرابنفش	84.63	84.38	84.08	84.41	83.82	83.55	83.92	83.69	84.06	84.03
طیف‌سنجی انعکاسی نزدیک مادون قرمز	83.15	83.72	8384	84.2	83.92	84.16	84.02	83.6	84.13	84.24

(نمونه‌ها برابر)

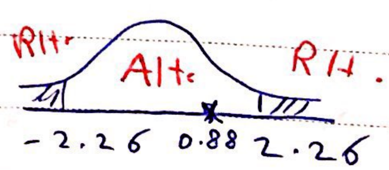
آیا اختلاف معنی‌داری بین نتایج بدست آمده از دو روش وجود دارد.

حل. ابتدا تست فرض‌ها را مرتب کنید $\begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d \neq 0 \end{cases}$ (فرضیه)

d_i: 1.48 0.66 0.24 −0.1 −0.61 −0.1 0.09 −0.07 −0.21

$\bar{d} = 0.159 \qquad S_d = 0.570 \to t_0 = \dfrac{0.159}{0.57/\sqrt{10}} = 0.88$

$t_{9, 0.975} = 2.26$



$|0.88| \not> 2.26 \Rightarrow AH_0$

پس بنابراین دو روش غلظت یکسانی گزارش می‌دهند.

$\mu_d : \bar{d} \pm t_{n-1, 1-\alpha/2} \dfrac{S_d}{\sqrt{n}} = (-0.249, 0.567)$

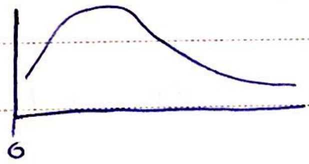
Subject:
Date

- آزمون برابر بودن پراکندگی دو جامعه

دیدیم در مقایسه میانگین دو جامعه، (گروه) دو حالت داشتیم یکی این که دو جامعه از نظر پراکندگی یکسان است و دو جامعه از نظر پراکندگی یکسان نیست. برای این که این مطلب را آزمون کنید از آماره $F_0 = \frac{S_1^2}{S_2^2}$ استفاده کنید. این آماره دارای توزیع فیشر با n_1-1 و n_2-1 درجه آزادی است. (توزیع فیشر یکی از توزیع های پیوسته ی آماری است که از تقسیم دو متغیر کای دو مستقل بر درجات آزادی شان بدست می آید.

$$u_1 \sim \chi^2_{df_1} \qquad u_2 \sim \chi^2_{df_2} \qquad \frac{u_1/df_1}{u_2/df_2} \sim F_{df_1,df_2}$$

این توزیع مثبت است و چوله ی



و جداول وجود دارد)

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

if $F_0 < F_{n_1-1,n_2-1,1-\alpha/2}$

or

$F_0 > F_{n_1-1,n_2-1,1-\alpha/2}$

$\Downarrow$

RH_0

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

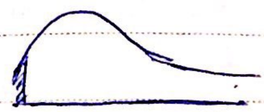
$F_{n_1-1,n_2-1,1-\alpha}$

$$\begin{cases} H_0: \sigma_1^2 \geq \sigma_2^2 \\ H_1: \sigma_1^2 < \sigma_2^2 \end{cases}$$

if $F_0 < F_{n_1-1,n_2-1,1-\alpha}$

$\Downarrow$

RH_0



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

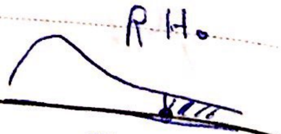
$$\begin{cases} H_0: \sigma_1^2 \leq \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases}$$

$\downarrow$

if $F_0 > F_{n_1-1,n_2-1,\alpha}$

$\Downarrow$

RH_0



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

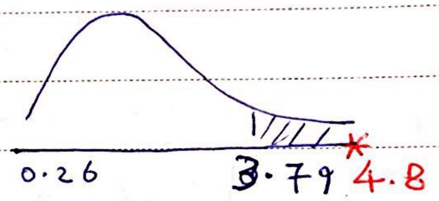
Chemical oxygen Demand

مثال ۲۳ برای تعیین مقدار اکسیژن شیمیایی مورد نیاز فاضلاب یک روش پیشنهاد شده است می خواهیم پراکندگی آن را با روش استاندارد مقایسه کنیم. آیا دقت روش استاندارد «بیشتر» از روش پیشنهادی است

استاندارد $n_1 = 8 \quad \bar{x}_1 = 72 \quad S_1 = 3.31$

پیشنهادی $n_2 = 8 \quad \bar{x}_2 = 72 \quad S_2 = 1.51 \rightarrow F_0 = \frac{3.31^2}{1.51^2} = 4.8$

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases}$$



$4.8 > 3.79 \rightarrow RH_0$

$F_{7,7,0.05} = 1/3.79 = 0.26$

$F_{7,7,0.95} = 3.79$

نکته
$$F_{n_1-1, n_2-1, \alpha} = \frac{1}{F_{n_2-1, n_1-1, 1-\alpha}}$$
$$F_{7,7,0.05} = \frac{1}{F_{7,7,0.95}}$$

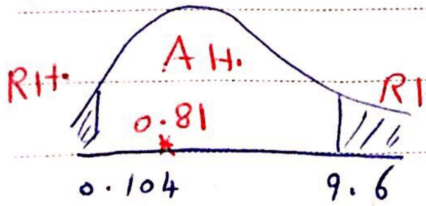
مثال ۲۴. در مثال ۱۹ آیا دقت (واریانس های) دو روش یکسان اند.

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

$F_0 = \frac{۰/۱۲۸^2}{۰/۱۳۱^2} \simeq 0.81 \ \not> 9.6, \ \not< 0.104 \Rightarrow AH_0$

$F_{4,4,0.025} = \frac{1}{9.6} = 0.104$

$F_{4,4,0.975} = 9.6$



دو جامعه از نظر پراکندگی یکسان اند

تمرین. در مثال ۲۰ و ۲۱ آزمون برابری واریانس ها (انجام شود)

Subject:
Date

آزمون دیکسون و آزمون گراببس (داده پرت) outliers

Dixon's Q test　　Grabbs test

H_0: داده پرت نیست

H_1: داده پرت است

گاهی نتایج آزمایش که جمع آوری می شود مشاهداتی به نظر انحرافی به نظر می آیند (پرت)

12.12^{ml}　12.15^{ml}　12.13^{ml}　13.14^{ml}　12.12^{ml}

برای نمونه با تعداد کم معمولاً آزمون دیکسون انجام می شود آماره به صورت زیر (۳ تا ۷ نمونه) است

$$Q = \frac{|\text{نزدیکترین جواب به آن} - \text{جواب مشکوک}|}{|\text{کوچکترین جواب} - \text{بزرگترین جواب}|}$$

که مقادیر بحرانی:

تعداد نمونه	4	5	6	7
$\alpha = 0.05$ q:	0.831	0.717	0.621	0.570

$\text{if } Q > q \Rightarrow$ داده مشکوک حذف RH_0 است.

مثال ۲۵. مقادیر زیر برای غلظت نیترات در یک نمونه از آب رودخانه بدست آمده اند

0.403　0.410　0.401　0.380

آخرین مشاهده به نظر مشکوک است آیا آن را حذف کنیم

$$Q = \frac{|0.380 - 0.401|}{|0.410 - 0.380|} = 0.7$$

مقدار بحرانی $= 0.831$　　$0.7 \not> 0.831$

داده مشکوک نگه داشته شود.

Subject:
Date

آزمون گراببس انحراف جواب مشکوک از میانگین نمونه را با انحراف معیار نمونه مقایسه می کند

$$G = |\text{مقدار مشکوک} - \bar{x}| / S$$

(فرض نرمال بودن حاکم است)

مقادیر بحرانی G ←

$\alpha = 0.05$

مقدار بحرانی g:	1.155	1.481	1.715	1.887	2.02	2.126	2.215	2.290
تعداد نمونه n:	3	4	5	6	7	8	9	10

if $G > g \Rightarrow$ داده پرت و حذف میشود

مثال ۲۶. فرض کنید به مثال ۲۵ سه اندازه گیری دیگر اضافه شده است آیا ۰٫۳۸ را حذف می کنید (هم Q هم G را انجام دهید)

0.403 0.41 0.401 0.380 0.4 0.413 0.408

$$Q = \frac{|0.38 - 0.41|}{|0.413 - 0.381|} = 0.606$$

in R - p_value = 0.03 ⇒ RH_0

$q_{7, 0.05} = 0.570$ $0.606 > 0.570$

داده مشکوک حذف شود (دیکسون)

$\bar{x} = 0.4026$ $S = 0.01088$

$$G = |0.380 - 0.4026| / 0.01088 = 2.031$$

$g_{0.05, 7} = 2.02$ $2.031 > 2.02$

داده مشکوک حذف شود (گراببس)

in R: p-value = 0.04 → RH_0

P4PCO

Scanned with CamScanner

در این مواقع بهتر است داده مشکوک را کنار بگذارید و انحراف معیار را حساب کنید

اگر انحراف معیار بعد از حذف داده پرت خیلی کمتر شد بهتر است حذف کنید

پیشنهاد شده که چون گاهی اوقات در بعضی کارهای تجزیه ای اندازه گیری ها

فقط تعداد محدود تکرار می شود و از کنار گذاشتن داده مشکوک خودداری شود می توان

به جای مقدار مشکوک میانه اعداد را گذاشت. گاهی اوقات با یک تبدیل

مثلاً لگاریتم گرفتن می توان ~~شکل داده ها~~ دید اگر داده پرت از بین می رود.

یک روش دیگر برای تشخیص داده پرت، نمودار جعبه ای است.

در minitab : (گرافیکی)

stat / basic statistics / outlier test

~~dixon~~Test

(~~پکیج را نصب کنید~~)

در R

```
chisq.out.test(x, variance = var(x))
dixonTest(x)
grubbs.test(x)
dixon.test(x)
```

outliers ← پکیج

مثال ۱۷

```
> ####example 17
> x=c(50.4,50.7,49.1,49,51.1)
> t.test(x,mu=50,alternative="two.sided",conf.level=0.95)

	One Sample t-test

data:  x
t = 0.14041, df = 4, p-value = 0.8951
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
 48.87358 51.24642
sample estimates:
mean of x 
    50.06
```

مثال ۱۸

```
> 
> ####example 18
> 
x=c(0.51,0.51,0.51,0.50,0.51,0.49,0.52,0.53,0.50,0.47,0.51,0.52,0.53,0.48,0.49,0.50,0.52,0.49,0.49,0.50,0.49,0.
48,0.46,0.49,0.49,0.48,0.49,0.49,0.51,0.47,0.51,0.51,0.51,0.48,0.50,0.47,0.50,0.51,0.49,0.48,0.51,0.50,0.50,0.5
3,0.52,0.52,0.50,0.50,0.51,0.51)
> t.test(x,mu=0.51,alternative="two.sided",conf.level=0.95)

	One Sample t-test

data:  x
t = -4.3781, df = 49, p-value = 6.274e-05
alternative hypothesis: true mean is not equal to 0.51
95 percent confidence interval:
 0.4951182 0.5044818
sample estimates:
mean of x 
   0.4998
```

مثال ۲۰

```
> ####example 20
> x=c(59,56,56,59,57,55)
> y=c(59,59,59,58,55,57)
> t.test(x,y,alternative="two.sided",conf.level=0.95,var.equal = TRUE)

	Two Sample t-test

data:  x and y
t = -0.88113, df = 10, p-value = 0.3989
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.940597  1.273931
sample estimates:
mean of x mean of y 
 57.00000  57.83333
```

```
> var.test(x, y, alternative = "two.sided")

	F test to compare two variances

data:  x and y
F = 1.0909, num df = 5, denom df = 5, p-value = 0.9263
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1526519 7.7960529
sample estimates:
ratio of variances 
          1.090909 
```

مثال۲۱

```
> ####example 21
> x=c(1.84,1.92,1.94,1.92,1.85,1.91,2.07)
> y=c(2.81,4.06,3.62,3.27,3.27,3.76)
> t.test(x,y,alternative="two.sided",conf.level=0.95,var.equal = FALSE)

	Welch Two Sample t-test

data:  x and y
t = -8.4772, df = 5.2528, p-value = 0.0002937
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.004938 -1.082205
sample estimates:
mean of x mean of y 
 1.921429  3.465000 

> var.test(x, y, alternative = "two.sided")

	F test to compare two variances

data:  x and y
F = 0.029451, num df = 6, denom df = 5, p-value = 0.0005021
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.004220663 0.176336947
sample estimates:
ratio of variances 
        0.02945053 
```

مثال ۲۲

```
> ####example 22
> x=c(84.63,84.38,84.08,84.41,83.82,83.55,83.92,83.69,84.06,84.03)
> y=c(83.15,83.72,83.84,84.2,83.92,84.16,84.02,83.6,84.13,84.24)
> t.test(x,y,paired=TRUE, alternative="two.sided",conf.level=0.95,var.equal = FALSE)

	Paired t-test

data:  x and y
t = 0.88211, df = 9, p-value = 0.4007
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2487527  0.5667527
sample estimates:
mean of the differences 
                  0.159 
```

مثال۲۵ و ۲۶

```
>
> ###example25,26
> library(outliers)
> x=c(0.403,0.41,0.401,0.380)
> dixon.test(x)

	Dixon test for outliers

data:  x
Q = 0.7, p-value = 0.1721
alternative hypothesis: lowest value 0.38 is an outlier

>
> x=c(0.403,0.41,0.401,0.380,0.4,0.413,0.408)
> grubbs.test(x)

	Grubbs test for one outlier

data:  x
G = 2.0343, U = 0.1953, p-value = 0.02161
alternative hypothesis: lowest value 0.38 is an outlier

> dixon.test(x)

	Dixon test for outliers

data:  x
Q = 0.60606, p-value = 0.0308
alternative hypothesis: lowest value 0.38 is an outlier

>
>
```