Ch 5 « Sampling Distribution » lecture "24"

If $X_1, X_2, \dots X_n \overset{\text{r.s}}{\sim} n(\mu, \sigma^2)$ (random sample)

Then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$ normal OR $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim n(0, 1)$

convert $\bar{X}$ to z-score

eg) Suppose that the grades in a general examination are normally distributed with mean 68 & standard deviation of 12 points, a sample of 4 grades are to be drawn, what is the prob. that the average of the grades drawn will be. a) more than 71 b) less than 65 c) between 66 and 74.

sol) $x_1, x_2, x_3, x_4 \sim n(68, 12^2)$

$\bar{X} \sim n(68, \frac{12^2}{4})$ ⇒

a) $P(\bar{X} > 71) = P(Z > \frac{71-68}{6}) = P(Z > 0.5) = 1 - P(Z \leq 0.5)$

$1 - 0.6915 = 0.3085$

b) $P(\bar{X} < 65) = P(Z < \frac{65-68}{6}) = P(Z < -0.5) = 0.3085$ (from table)

c) $P(66 < \bar{X} < 74) = P(\frac{66-68}{6} < \bar{X} < \frac{74-68}{6})$

$= P(-0.33 < Z < 1) = P(Z<1) - P(Z < 0.33)$

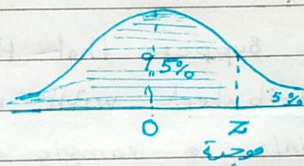
$= 0.8413 - 0.3707 =$

eg) Suppose that the weights of Orange boxes are normally distributed with mean 10 kgs & standard deviation of 1.5 kgs. If a no. of boxes will be loaded in a car with threshold 100 kgs. Find the no. of boxes that will be loaded so that their total weigh doesn't exceed the threshold of the car with prob. about 0.95.

Sol) $x_1, x_2, x_3, \dots, x_n \sim n(10, 1.5^2)$

$$\bar{x} \sim n\left(10, \frac{1.5^2}{n}\right)$$

$$P\left(\frac{\sum_{i=1}^{n} x_i}{n} \leq \frac{1000}{n}\right) = 0.95$$



$$P\left(\bar{x} \leq \frac{1000}{n}\right) = 0.95$$

$$P\left(Z \leq \frac{\frac{1000}{n} - 10}{1.5/\sqrt{n}}\right) = 0.95$$

$$\frac{\frac{1000}{n} - 10}{1.5/\sqrt{n}} = 1.64 \Rightarrow \left(\frac{1000}{n} - 10\right) \cdot \frac{\sqrt{n}}{1.5} = 1.64$$

let $x = \sqrt{n}$ $\therefore x^2 = n$

$$\left(\frac{1000}{x^2} - 10\right) \cdot \frac{x}{1.5} = 1.64$$

$$1.5\left(\frac{1000}{1.5x} - \frac{10x}{1.5}\right) = (1.64)1.5$$

$$\frac{1000}{x} - \frac{x * 10x}{x * 1} = 2.46 \Rightarrow \frac{1000 - 10x^2}{x} = 2.46x$$

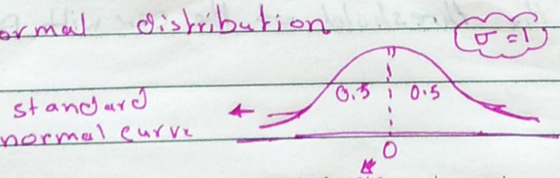
$$\frac{10x^2}{10} + \frac{2.46x}{10} - \frac{1000}{10} = 0 \rightarrow x^2 + .0246x - 100 = 0$$

$$\frac{-b \mp \sqrt{b^2 - 4ac}}{2a} = \frac{-0.246 + \sqrt{(0.246)^2 + 400}}{2} \text{ \& } \frac{-0.246 - \sqrt{(0.246)^2 + 400}}{2}$$

$$\frac{-0.246 + 20}{2} = 9.877 \text{ \& } \text{مرفوض} \; ^2\text{د}$$

$\therefore x = 9.877$ so $x^2 = 97.56 \rightarrow \therefore n = 97.56 \approx 98$ boxes عدد صحيح

* normal distribution



The area under the curve is equal to 1

Chapter 6

T- Distribution

Note :- If $X_1, X_2, \dots X_n \overset{r.s}{\sim} n(\mu, \sigma^2)$, then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$

or $z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim n(0,1)$, provided that σ is <u>Known</u>

If σ is <u>unknown</u>, then we estimate σ by $S = \frac{\sum x_i^2}{n-1} - \frac{(\sum x_i)^2}{n(n-1)}$ & then

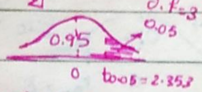
$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$, t- distribution with n-1 degrees of freedom (d.f)

But for $n \geq 30$, then $Z = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim n(0,1)$ اذا كان large(n) بيصير نفس normal

* t- distribution :-

d.f	$t_{0.1}$	$t_{0.05}$	$t_{0.01}$	$t_{0.25}$
1				
2				
8		$\boxed{2.353}$		
⋮				
30				

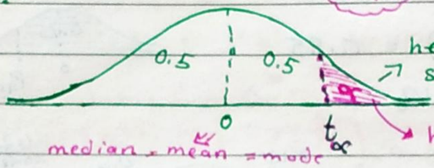
from table



* $P(T \leq 2.353) = 0.95$

* $P(T \geq 2.353) = 0.5$

$\sigma > 1$



hevier tail than the standard normal → meaning that it is more prone to producing values that fall far from its mean

hevier (thiker) اعرض

median = mean = mode

ما معنى (d.f) ؟ يعني مثلاً كان وحطينا ان ال mean ل n=3

مجموعهم يساوي 5 . $\boxed{6}$ $\boxed{5}$ $\boxed{x}$ اذا اخترنا رقمين ، الثالث بيكون مجبور

باختياره مثلاً صون ٤ فالي كنت حُر باختيارهم هم رقمين . ((n-1))

the no. of free choices

eg) Suppose that the weights of new born babies are normally distributed with mean 3 kgs. A random sample of size 10 is taken & ~~showed~~ showed that its ~~standard~~ standard deviation is 2.

a) Find the prob. that the sample average (S وليس σ) is below 4.16 kgs ?

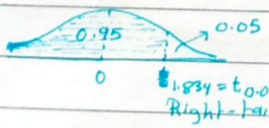
b) What is the 90th percentile of the distribution of the $\bar{X}$?

Sol) $X_1, X_2, \dots X_{10} \sim n(3, \sigma^2)$; $n = 10$, $S = 2$

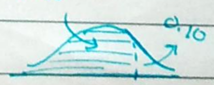
a) $P(\bar{X} < 4.16)$; $P\left(\frac{\bar{x}-\mu}{S/\sqrt{n}} < \frac{4.16-\mu}{S/\sqrt{n}}\right) = P\left(T < \frac{4.16-3}{2/\sqrt{10}}\right) = P(T < 1.834)$

$= 0.95$

d.f = 9



b) $P(\bar{X} < P_{90}) = 0.90 \rightarrow P\left(T < \frac{P_{90}-3}{2/\sqrt{10}}\right) = 0.90$



$\frac{P_{90}-3}{2/\sqrt{10}} = 1.383 \rightarrow P_{90} = 3.87$

$1.383 \overset{\text{من الجدول}}{=} \frac{P_{90}-3}{2/\sqrt{10}} = t_{0.10}$

To Summer up:-

* To sum. up :-

If $x_1, \dots x_n \overset{r.s}{\sim} n(\mu, \sigma^2)$, then

[1] If σ is known, then $\bar{x} \sim n(\mu, \frac{\sigma^2}{n})$ or $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim n(0,1)$

[2] If σ is unknown, if a) $n < 30$, then $T = \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n-1)$

b) $n \geq 30$, then $Z = \frac{\bar{x}-\mu}{s/\sqrt{n}} \overset{\cdot}{\sim} n(0,1)$ ⟵ C.L.T

تقريبي

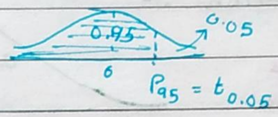
⟹ T distribution

normal تقريبي

* eg) If $T \sim t(10)$, find : a) the 95^{th} percentile of T

b) the 10^{th} percentile of T

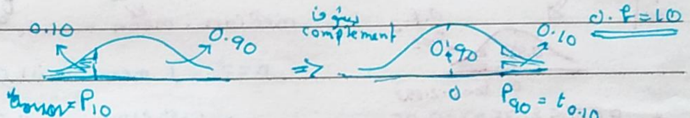
Sol) a) $P(T \leq P_{95}) = 0.95$



from table

$\therefore P_{95} = 1.812$

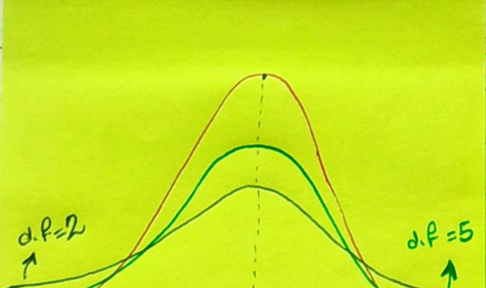
b) $P(T \leq P_{10}) = 0.10$



$P_{90} = t_{0.10} = 1.372 \rightarrow \therefore P_{10} = \boxed{-1.372}$

$P_{10} = -P_{90}$

* As the degrees of freedom increase, the t-distribution approaches the standard normal distribution, For 80 or more degrees of freedom, the t distribution is close to the standard normal distribution.



* لو طلب أقل من P_{50}

منوخذ ال Complement

$P_{10} = -P_{90}$

$P_{20} = -P_{80}$

$P_{30} = -P_{70}$

- - - -

6.4

lecture '25'

* The distribution of the Sample variance, S^2 :-

حطينا أن التوزيع بيكون بالنسبة لـ $\bar{X}$

← اذا الـ σ معروفة $\bar{X} \sim n$

← اذا الـ σ غير معروفة $\bar{X} \sim t$

proportion و variance الـ lecture في جبنا الـ

* If $X_1, X_2, \dots X_n \overset{r.s}{\sim} n(\mu, \sigma^2)$, then

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

chi-squared

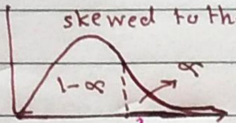
d.f $= n-1$

إله جداول.

بنفس طريقة التعامل مع جدول T...

d.f = n-1

skewed to the right



χ^2_α

*جميع القيم موجبة

All values of χ^2 are greater than or equal to 0

ليس not symmetric و pill-shaped → Positively skewed

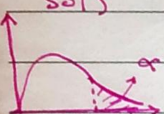
(The total area under the curve is equal to 1)

eg) If a sample of size $n = 6$ is drawn from a population with Variance $\sigma^2 = 10$. Find $P(S^2 > 18.4727)$

اذا سألني عن S^2 منحول لـ chi-squared

sol) $P\left(\frac{(n-1)\,S^2}{\sigma^2} > \frac{(n-1)(18.47277)}{\sigma^2}\right) = P\left(\chi^2 > \frac{(6-1)(18.4721)}{10}\right)$

$\chi^2_{.10}$



$P(\chi^2 > 9.2364)$ from Table →

$= 0.10$

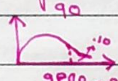
d.f	.10
1	
2	
3	
4	
5	9.2364
⋮	

eg) let $x_1, \dots, X_{10} \sim n(\mu, 25)$

If S^2 is the sample variance, find the 90^{th} percentile of S^2

sol) Need to find P_{90} ⇒ $P(S^2 < P_{90}) = 0.90$

$P\left(\chi^2 > \frac{9(P_{90})}{25}\right)$



$\frac{9P_{90}}{25} = \chi^2_{.10}$

حولناه chi-squared

from table ⇒

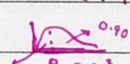
$\frac{9 \cdot P_{90}}{25} = 14.6437 \rightarrow P_{90} = 40.7881$

eg) Let $X \sim \chi^2(10)$. Find : a) The 10^{th} percentile of X

chi-squared

b) The 95^{th} percentile of X c) The 99^{th} percentile of X

sol) a) $P(X < P_{10}) = 0.10$

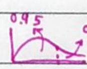


$P_{10} = \chi^2_{0.90}$

d.f = 10 ⇒ ∴ $\chi^2_{0.90} = 4.865180 \approx 4.87$

لاحظ أن القيم زادت

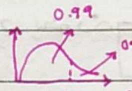
b) $P(X < P_{95}) = 0.95$



$P_{95} = \chi^2_{0.05}$

d.f = 10 ⇒ ∴ $\chi^2_{0.05} = 18.3070 \approx 18.3$

c) $P(X < P_{99}) = 0.99$



d.f = 10

$P_{99} = \chi^2_{0.01} = 23.209$

$\chi^2_{0.01} = 23.209$

To summer up:

Population

Parameters

① mean → μ

② Variance → σ^2

③ Proportion → P

اخذت sample

$x_1, x_2, \dots x_n$

حقا وظروف وقت وجهد ومال.

Sample

Estimators

① mean → $\bar{x}$

② variance → s^2

③ Proportion → $\hat{P}$

$\bar{x}$:

- $\bar{x} \sim n\left(\mu, \frac{\sigma^2}{n}\right)$ → σ known
- $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n-1)$ → σ unknown, $n < 30$
- $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim n(0, 1)$ → σ unknown, $n \geq 30$

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

$\hat{p} \longrightarrow \hat{p} \sim n\left(P, \frac{Pq}{n}\right)$

هدول كلهم لما بكون اخذت عينة واحدة فقط لا أكثر.

* The distribution of the sample proportion $\hat{p} = \frac{x}{n}$:-

مثلاً جبنا عينة لـ 100 شخص وطلعت نسبة المدخنين طلعوا 60 مدخن

فنسبة الـ proportion = $0.60 = \frac{60}{100}$ لحساب الـ $\hat{p}$ اعرف ان n Large

$$\boxed{\hat{p} \sim n\left(p, \frac{pq}{n}\right)} \quad \text{or} \quad Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim n(0,1)$$

eg. Suppose that 10% of a certain production are defective. If 400 items are drawn from the production, what is the prob. that the sample proportion will be:

a) more than 12 % b) between 9% & 11 %

sol) $n = 400$, $p = 0.10$, $q = 0.90$

$$\hat{p} \sim n\left(0.10, \frac{(0.10)(0.90)}{400}\right) \Rightarrow \hat{p} \sim n\left(0.10, \left(\frac{3}{200}\right)^2\right)$$

a) $P(\hat{p} > 0.12) = P\left(Z > \frac{0.12 - 0.10}{3/200}\right) = P(Z > 1.33)$

$= 1 - P(Z \leq 1.33) = 1 - 0.9082 = 0.0918$

b) $P(0.09 < \hat{p} < 0.11) = P\left(\frac{0.09 - 0.10}{3/200} < Z < \frac{0.11 - 0.10}{3/200}\right)$

$P(-0.67 < Z < 0.67) = P(Z < 0.67) - P(Z < -0.67)$

$= 0.7486 - 0.2514 = 0.4972$

p: نسبة الـ defective فيهم => في الـ population

$\hat{p}$: نسبة الـ defective فيهم => في الـ sample eg) Suppose that 90% of the university students pass Calculus 101. In a sample of 200 students taking Calculus 101, What is the prob. that the proportion of those who will pass is less than 85%.

large (n) بملاحظة

$n = 200 \qquad p = 0.90 \qquad q = 0.10$

$$\hat{p} \sim n\left(0.90, \frac{(0.90)(0.10)}{200}\right)$$

$$P(\hat{p} < 0.85) = P\left(Z < \frac{0.85 - 0.90}{\sqrt{(0.90)(0.10)/200}}\right)$$

$$P(Z < -2.36) = 0.0091$$

Sampling distribution. (Sample) بعد ما اخذنا

$\boxed{1}$ * $\bar{x} \sim n\left(\mu, \frac{\sigma^2}{n}\right) \Leftarrow$ normal توزيعه (σ معروفة)

$\boxed{2}$ $\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1) \Leftarrow$ (σ مجهولة) } توزيعه t

$\boxed{3}$ $\frac{\bar{x} - \mu}{s/\sqrt{n}} \dot{\sim} n(0,1)$ اذا كانت (n) كبيرة large يقترب بالnormal

$\boxed{4}$ $\hat{p} \dot{\sim} n\left(p, \frac{pq}{n}\right) \Rightarrow Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim n(0,1)$ توزيعها normal

$\boxed{5}$ $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ Chi-squared توزيعها

Sampling Distribution.

* The distribution of the difference between 2 sample means: Lecture 26

If $X_1, \dots, X_n \overset{r.s}{\sim} n(\mu_1, \sigma_1^2)$ إذا اخذت عينتين

& $Y_1, \dots, Y_m \overset{r.s}{\sim} n(\mu_2, \sigma_2^2)$, then

$\bar{X} \sim n(\mu_1, \frac{\sigma_1^2}{n})$ (σ معروفة.)

& $\bar{Y} \sim n(\mu_2, \frac{\sigma_2^2}{m})$

$$\therefore \bar{X} - \bar{Y} \sim n\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

Provided that σ_1 & σ_2 are known.

$$\text{or} \quad Z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim n(0,1)$$

If $\sigma_1 = \sigma_2 = \sigma$ (unknown)

$$T = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{\cancel{S_p}}{n} + \frac{\cancel{S_p}}{m}}} \sim t(n+m-2)$$

d.f for $\bar{x} = n-1$

" " $\bar{y} = m-1$

$\boxed{n+m-2}$ اجمع

Where $S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$ is the Pooled variance.

Note: If $n, m \geq 30$, then $\frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}} \sim n(0,1)$

e.g.) Suppose that the grades of female & male students in Calculus 101 are normally distributed with means 70 & 65; respectively & standard deviation 8 & 10. respectively. In samples of 15 female & 20 male students, find the prob. that the female students will have an average more than male students average?

Sol.) $X_1, X_2, \ldots X_{15} \overset{r.s}{\approx} n(70, 8^2)$ ← female

$Y_1, Y_2, \ldots X_{20} \overset{r.s}{\approx} n(65, 10^2)$ ← male

$$\bar{X} \sim n\left(70, \frac{8^2}{15}\right)$$

$$\bar{Y} \sim n\left(65, \frac{10^2}{20}\right)$$

$$\bar{X} - \bar{y} \sim n\left(5, \frac{8^2}{15} + \frac{10^2}{20}\right)$$

$P(\bar{X} > \bar{y}) = P(\bar{X} - \bar{y} > 0) \longrightarrow$ نحولها Z

$$P\left(Z > \frac{0-5}{\sqrt{\frac{64}{15} + \frac{100}{20}}}\right) = P(Z > -1.64)$$

$1 - P(Z \leq -1.64) = 1 - 0.0505 = 0.9495$

من الجدول

* The distribution of the difference between 2 sample proportions:

$\hat{p} \sim n(P, \frac{pq}{n})$ كما نعلم سابقاً

$$\hat{P}_1 - \hat{P}_2 \sim n\left(P_1 - P_2, \frac{P_1 q_1}{n} + \frac{P_2 q_2}{m}\right)$$

or $$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n} + \frac{P_2 q_2}{m}}} \sim n(0, 1)$$

* eg) Suppose that 50% of population A own cars, while 35% of population B own cars. If a sample of size 100 is drawn from population A & a sample of size 80 is drawn from population B, What is the prob. that the difference between the sample proportions $\hat{P}_A - \hat{P}_B$ will be between 0.1 & 0.2?

Sol)

A

$P_1 = 0.50$

$n = 100$

$q_1 = 0.50$

B

$P_2 = 0.35$

$m = 80$

$q_2 = 0.65$

$$\hat{P}_1 - \hat{P}_2 \sim n\left(0.5 - 0.35 \, , \, \frac{(0.50)(0.50)}{100} + \frac{(0.35)(0.65)}{80}\right)$$

$$\hat{P}_1 - \hat{P}_2 \sim n\,(0.15 \, , \, 0.073^2)$$

$$P(0.1 < \hat{P}_1 - \hat{P}_2 < 0.2) = P\left(\frac{0.1 - 0.15}{0.073} < Z < \frac{0.2 - 0.15}{0.073}\right)$$

$$P(-0.68 < Z < 0.68) = P(Z < 0.68) - P(Z < 0.68)$$

$$= 0.7517 - 0.2483 = 0.5034$$

Lecture 2

Chapter 6 (concept of statistical Inference) Lecture 27

Confidence Interval (C.I) | Tests of hypotheses

* Interval Estimation : [1] C.I for μ [2] C.I for P [3] C.I for σ^2

* Estimation by C.I :-

Def) Let L & U be functions of $x_1, x_2, \dots x_n$.

(L, U) is $\boxed{(1-\alpha)100\%}$

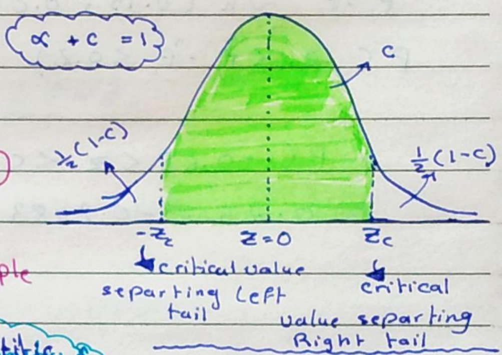
C.I . $(0<\alpha<1)$ for θ if: $P(L<\theta<U)=1-\alpha$

$1-\alpha$: confidence coefficient : C

✓ α : significance level : $(1-C)$

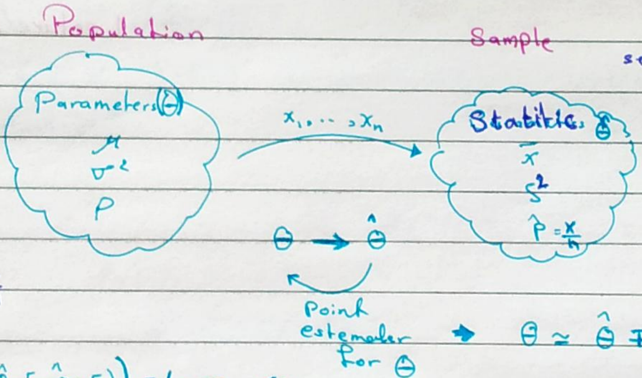
L: Lower confidence Limit (L.C.L)

U: upper " " (U.C.L)



* point estimate of σ^2 is S^2

* point estimate of P is $\hat{P} = \frac{x}{n}$



* The point unbiased estimate of μ is $\bar{X}$

$\rightarrow \theta \simeq \hat{\theta} \mp$ Error $\rightarrow$ sampling distribution نعوض عن

$P(\theta \in (\hat{\theta}-E, \hat{\theta}+E)) = 1-\alpha \leftarrow$

$\hat{\theta}-E \quad \hat{\theta} \quad \hat{\theta}+E$ ↳ margin of error

$\theta \in (\hat{\theta}-E, \hat{\theta}+E)$ ↳ Lower CL ↳ upper CL

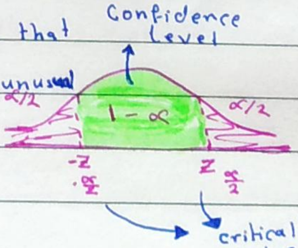
$\%100(1-\alpha) = 0.90 * 100\%$

$\%100(1-\alpha) = 90\%$

Note :- If $x_1, x_2, \dots x_n \overset{r.s}{\sim} n(\mu, \sigma^2)$, then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$

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

* critical values are values that separate sample statistics that are probable from sample statistic that are improbable or unusual



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

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

* الكتاب اعتمد $-Z_c$ & Z_c

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

* غير مطالبين بهذا الاشتقاق.

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

$$P\left(\bar{x}+Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}} > \mu > \bar{x}-Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\rightarrow P\left(\bar{x}-Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}} < \mu < \bar{x}+Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}\right)$$

$\therefore$ The $(1-\alpha)100\%$ C.I. for μ is: $\left(\bar{x}-\underbrace{Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}}_{\text{Error}}, \bar{x}+\underbrace{Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}}_{\text{Error}}\right)$

Interval estimate $\Leftarrow$

Add and substract a margin of error.

~~6.1~~

$\boxed{1}$ The C.I. for μ:-

The $(1-\alpha)100\%$ C.I. for μ is:

i) $\boxed{\bar{x} \mp Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}}$ if σ is Known.

ii) $\boxed{\bar{x} \pm t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}}}$ if σ is unknown.

iii) $\boxed{\bar{x} \pm Z_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}}}$ if σ is unknown but $n \geq 30$

Not: $Z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$: Error = E $\Leftarrow$ Margin of error

$\frac{\sigma}{\sqrt{n}}$: standard error (S.E) ~~(margin error)~~ $\Leftarrow$

يجب الانتباه الى أمثلة الكتاب $\Leftarrow$ (Page 321)

eg) The Salaries of teachers in Jordan for 1990-2000 are normally distributed with standard deviation 50 JD. The average salary based on sample of 400 teachers for 1990-2000 was 215 JD per month.

a) What is the point estimate for the mean salaries & its S.E.?

b) Give a 90% C.I. for the mean salaries

c) Give a 95% C.I for the mean salaries.

Sol) $x_1, x_2, \dots x_{400} \sim n(\mu, 50^2)$ $\boxed{\bar{x}=215}$

a) Point estimate for μ is $\bar{x}=215$ $\rightarrow$ $S.E = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{400}} = \frac{50}{20} = 2.5$

b) $\overset{C}{1-\alpha} = 0.90 \Rightarrow \overset{1-C}{\alpha = 0.10} \Rightarrow \overset{\frac{1}{2}(1-C)}{\frac{\alpha}{2} = 0.05}$

$-1.64 = -Z_{0.05}$ 0 $Z_{0.05} = 1.64$

Table من

The 90% C.I for μ is:

$\bar{x} \mp Z_{0.05} \cdot \frac{\sigma}{\sqrt{n}}$

$215 \pm (1.64) \cdot \left(\frac{50}{\sqrt{400}}\right)$

$L = 215 - (1.64)(2.5) = 210.90$

$U = 215 + (1.64)(2.5) = 219.10$

$\therefore$ The 90% C.I for μ is: $(\overset{\text{left endpoint}}{210.90}, \overset{\text{Right endpoint}}{219.10})$

والعطمس طبعا $\bar{x}$

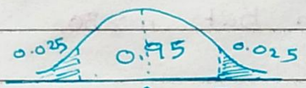
point estimate μ آو $= \frac{210.90 + 219.10}{2} = \boxed{215}$

Error لو بدنا نوجد

215 - 210.90 بطرح
او 219.10 - 215
لان (215) بالوسط

standard Error بقدر اطلع Error اذا عرفت ←

c) $1-\alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \frac{\alpha}{2} = 0.025$



$-1.96 = -Z_{0.025}$ 0 $Z_{0.025} = +1.96$

لو من جدول السالب بجيبه

	0.06
-1.9	$\boxed{0.025}$

from table

$L = \bar{x} - Z_{0.025} \frac{\sigma}{\sqrt{n}} = 215 - (1.96)(2.5) = 210.10$

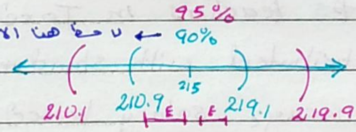
$U = \bar{x} + Z_{0.025} \frac{\sigma}{\sqrt{n}} = 215 + (1.96)(2.5) = 219.90$

$\therefore$ The 95% C.I. for μ is: $(210.10, 219.90)$

* you are 95% confident that the margin of error for the population mean is about 4.9

لاحظ أنه

لاحظ هنا الارقام اقرب على بعضها (أضيق). ← 90%



كلما زاد الشيء $(1-\alpha)$ كانت الفترة المحسوبة $(C.I)$ width أوسع.

* C increases
بالتالي E increases

يعني كلما زادت $(1-\alpha)$ كلما قلت (α)

فالعلاقة بين C.I و ال significance level عكسية

كلما زادت (n) فترة ال (E) أضيق .: يعني لو كانت $n = \infty$ بيصير $\frac{\sigma}{\sqrt{\infty}} = \cdot$

يعني اعلى اشي لو أخذنا كل ال population . فبصير $\bar{x} = \mu$

فبتصير أضيق الفترة ، C.I becomes narrower

* Sample size to estimate μ.

$$E = Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow \frac{E}{Z_{\frac{\alpha}{2}}} = \frac{\sigma}{\sqrt{n}}$$

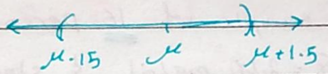
$$\Rightarrow \frac{\sqrt{n}}{\sigma} = \frac{Z_{\frac{\alpha}{2}}}{E} \rightarrow \sqrt{n} = \frac{Z_{\frac{\alpha}{2}}}{E} \cdot \sigma \Rightarrow \boxed{n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot \sigma^2}$$

eg) A researcher wants to estimate the average weight loss of people who are on a new diet plan. In a previous study, the population standard deviation σ of weight losses is about 5 kgs. How large a sample should be to estimate the mean weight loss by a 95% C.I to within 1.5 kgs?

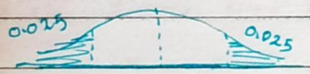
sol) $\sigma = 5$, $n = ??$, $1-\alpha = 0.95 \rightarrow \alpha = 0.05$

$\rightarrow \frac{\alpha}{2} = 0.025$

within 1.5 $\Rightarrow$



$\Rightarrow \bar{X} \pm E \rightarrow$ Error = 1.5



$-1.96 = -Z_{0.025}$ 0 $Z_{0.025} = 1.96$

من الجدول.

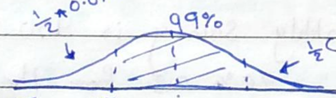
$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot \sigma^2 = \left(\frac{1.96}{1.5}\right)^2 * 5^2 = 42.68$$

$\boxed{n \approx 43}$

سؤال اختبر نفسك :- ←

eg) Suppose that Jordan Bureau of Census in 2004 wants to estimate the mean size μ of all Jordan families by 99% C.I. It is known that the standard deviation $\sigma = 1.5$. How large a sample size should be the bereau select to (estimate μ) within 0.02 of the population mean.

$c = 99\%$, $\sigma = 1.5$, $E = 0.02$



$-2.575 = -Z_{0.005}$ 0 $Z_{0.005} = 2.575$

from table

$$n = \left(\frac{Z_c}{E}\right)^2 \cdot \sigma^2 = \left(\frac{2.575}{0.02}\right)^2 \cdot (1.5)^2 = 37.297\ldots$$

$n \approx \boxed{38}$

* Because n is not a whole number, round up ← to the next whole number. 38

دائماً منقرب round up

ملاحظة اذا ذكر بالسؤال from past studies ← يكون σ وليس S

* Interval Estimation

(Confidence Interval C.I)

① Interval estimation of μ:

ملاحظة :- اذا كانت التوزيع ليس normal ولا $n \geq 30$ فما بنقدر نطلع C.I لـ μ

The $(1-\alpha)100\%$ C.I for μ is:

a) $\bar{x} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ if σ is known.

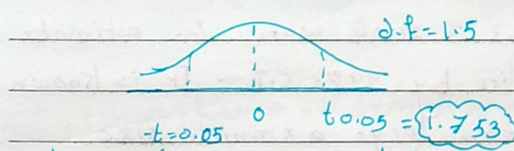
b) $\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ if σ is unknown & $n < 30$ — من جدول t.distribution بدلاً من جدول Z standard normal

c) $\bar{x} \pm Z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ if σ is unknown & $n \geq 30$

Note $n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot \sigma^2$. (sample size)

eg) The mean cholestrol levels in general population are normally distributed. A sample of 16 persons is taken under a test with sample mean $\bar{x} = 220$ mg/dL & standard deviation $S = 25$ mg/dL. Give a 90% C.I for the Population mean μ.

Sol) $n = 16$, $\bar{x} = 220$, $S = 25$, $1-\alpha = 0.90 \rightarrow \alpha = 0.10 \Rightarrow \frac{\alpha}{2} = 0.05$



$\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$

$\Rightarrow 220 \pm (1.753)\left(\frac{25}{\sqrt{16}}\right)$

L.C.L $= 220 - (1.753)\left(\frac{25}{4}\right) = 209.04$

U.C.L $= 220 + (1.753)\left(\frac{25}{4}\right) = 230.96$

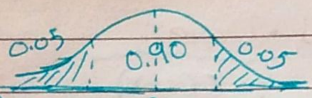
The 90% C.I for μ is: $(209.04, 230.96)$

eg) A random sample of 400 people with a professional degree taken showed that their mean monthly salary is 450 JD with a standard deviation of 100 JD. Give a 90% C.I for the mean monthly salary.

Sol)

n ⇒ very larg $(n = 400)$, $\bar{X} = 450$, $S = 100$, σ is unknown

$1-\alpha = 0.90 \rightarrow \alpha = 0.10 \rightarrow \therefore \frac{\alpha}{2} = 0.05$



$$\bar{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

$-1.64 = -Z_{0.05}$ 0 $Z_{0.05} = 1.64$

الأسهل من السالب من جدول Z

$L.C.I = 450 - (1.64)\left(\frac{100}{\sqrt{400}}\right) = 441.8$

$U.C.I = 450 + (1.64)\left(\frac{100}{20}\right) = 458.2$

The 90% C.I for μ is: $(441.8 , 458.2)$

② Interval estimation for P : proportion

The $(1-\alpha)100\%$ C.I for P is $\hat{P} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P} \cdot \hat{q}}{n}}$ $S.E = \sqrt{\frac{P \cdot q}{n}}$

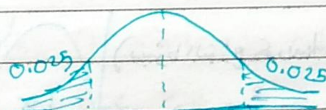
* The sample size is .... $n\hat{p} \geq 5$ & $n\hat{q} \geq 5$ بشرط ان

اذا ما تحقق الشرط ما منقدر نحسب (C.I) لـ $\hat{p}$

eg) It was believed in the Arab World that 50% of persons are smoking. During the year 200 , a sample of 100 persons showed that the number of smokers is 620 Establish 95% C.I. for the proportion of smokers , P.

Sol) $n = 1000$, $X = 620$, $\hat{P} = \frac{X}{n} = \frac{620}{1000} = 0.62$

$\hat{q} = 1 - 0.62 = 0.38 \Rightarrow 1-\alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \therefore \frac{\alpha}{2} = 0.025$



$$\hat{P} \pm Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

✓ $n\hat{q} \geq 5$ & $n\hat{p} > 5$ ولكن

$-1.96 = -Z_{0.025}$ 0 $Z_{0.025} = 1.96$

من الجدول

$L.C.I \rightarrow 0.62 - (1.96)\sqrt{\frac{(0.38)(0.62)}{1000}} = 0.59$

$U.C.I \rightarrow 0.62 + (1.96) = \sqrt{\frac{(0.38)(0.62)}{1000}} = 0.65$

The 95% C.I for P is : $(0.59 , 0.65)$

* Determination of the sample size :-

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot \overset{*}{p} \cdot (1-\overset{*}{p})$$

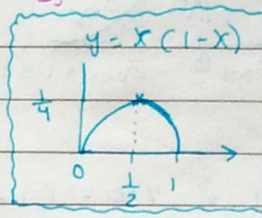
لأني مو عارفة $\hat{p}$ بعطيها للان $\overset{*}{p}$
حتى أميزها

$\overset{*}{p}$

Known from previous study

or

$\overset{*}{p} = \frac{1}{2}$ if it's not Known from previous study
$\hat{p} = \hat{q} = 0.5$



ولأنه ←

eg) Assume that it is required to estimate the proportion of patients suffering a bad reaction from taking a certain medication P by 95% C.I. Determine the sample size needed if the error of estimation is about 0.10 in the following cases;

a) no prior information about P.

b) Previous study showed that P is approximately 0.20.

sol) $n = ??$, $E = 0.10$, $1-\alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \frac{\alpha}{2} = 0.025$

$Z_{0.025} = 1.96$

a) $\overset{*}{p} = 0.5$

$$n = \left(\frac{1.96}{0.10}\right)^2 \cdot (0.5)(1-0.5) \simeq 97$$

احنا منعمل round up ·
فهو ..96، بوشت اشي ·

b) $\overset{*}{p} = 0.20$

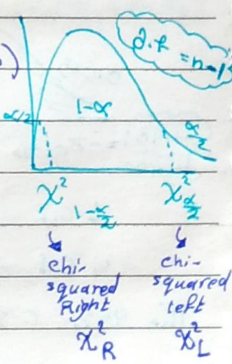
$$n = \left(\frac{1.96}{0.10}\right)^2 \cdot (0.20)(1-0.20) \simeq 62$$

* وجود عينة سابقة بقلل الـ n وبالتالي بقلل cost
أوفر طبعا أوخذ عينة اصغر

* Interval Estimation for σ^2 :- (normal distribution بشرط التوزيع يكون)

The $(1-\alpha)100\%$ C.I for σ^2 is:

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



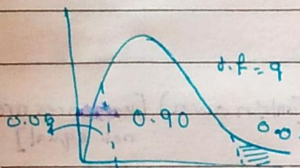
C.I for σ^2 ⇒ $\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} , \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} \right)$

هون طالع symetric لأنه مو معتبر
نجمع ونقسم على اثنين لا يجاد $\bar{X}$

C.I for σ → $\left(\sqrt{\frac{(n-1)s^2}{\chi^2_R}} , \sqrt{\frac{(n-1)s^2}{\chi^2_L}} \right)$

eg) Quality - control engineer wishes to study the weight variation of a new product. A sample of 10 items is taken & provided $\bar{x} = 0.60$ kg & $s = 0.4$ Kgs. Find a 90% C.I for the variances of all items. (Assume that the distribution of the weights can be modeled as a normal distribution).

Sol) $n = 10$, $\bar{x} = 0.60$, $S = 0.40$, $1-\alpha = 0.90 \rightarrow \alpha = 0.10 \rightarrow \therefore \frac{\alpha}{2} = 0.05$



The 90% C.I for σ^2 is :

$$\frac{(n-1)\,S^2}{\chi^2_{0.05}} \;,\; \frac{(n-1)\,S^2}{\chi^2_{0.95}} = \left(\frac{(10-1)*(0.40)^2}{16.199} , \frac{(10-1)(0.40)^2}{3.3251} \right)$$

$3.3251 = \chi^2_{0.95}$ $\qquad$ $\chi^2_{0.05} = 16.919$

من الجدول $\qquad$ المساحة على اليمين $\qquad$ المساحة على اليمين $\qquad$ من الجدول

$= (0.09 , 0.43)$

$\therefore$ The confidence interval for σ is

$(\sqrt{0.09} < \sigma < \sqrt{0.043})$

* To sum up :-

The $(1-\alpha)$ 100% C.I for μ is

① $\bar{x} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, if σ is Known.

② $\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$, if σ is unknown but $n<30$

③ $\bar{x} \pm Z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$, if σ is = = $n \geq 30$

$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E} \right)^2 \cdot \sigma^2$

$\mu \in (\bar{x} \mp E)$

$p \in (\hat{p} \pm E)$

لماذا ,, نفعل

$\sigma^2 \in (S^2 \mp E)$

؟ الجواب ,, لأنها ليست

symmetrice. doesn't have S as center.

* The $(1-\alpha)$ 100% C.I for p is :

$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$

$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E} \right)^2 \cdot \overset{*}{p} \cdot \overset{*}{q}$ < $\overset{*}{p}$ is given ; $\overset{*}{p} = \frac{1}{2}$

The $(1-\alpha)$ 100% C.I for σ^2 is :

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

Chapter "8"

C.I for two sampling distribution.

If $x_1, \ldots, x_{n_1} \overset{r.s}{\sim} n(\mu_1, \sigma_1^2)$ & $y_1, \ldots, y_{n_2} \overset{r.s}{\sim} n(\mu_2, \sigma_2^2)$

μ_1 is $\bar{x}$ $\qquad$ μ_2 is $\bar{y}$

① C.I for $\mu_1 - \mu_2$:

- $(\bar{x} - \bar{y}) \mp Z_c \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$, if σ_1, σ_2 is known.

→ The sample must be randomly selected and independent.

- $(\bar{x} - \bar{y}) \mp t_c \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ if $\sigma_1 \neq \sigma_2$ (unknown) [variances are not equal]

يعني بنقص الأول والثاني ونختار الأصغر $\quad$ * $d.f = \min\{n_1 - 1, n_2 - 1\}$

- $(\bar{x} - \bar{y}) \mp t_c \cdot sp \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, if $\sigma_1 = \sigma_2$ (unknown)

* $d.f = n_1 + n_2 - 2$

⟹ when $\hat{\sigma}^2 = S^2p = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$

. pooled variance لازم نحسب وبعدين

② C.I for $P_1 - P_2$: $\qquad$ Page 476

$(\hat{P}_1 - \hat{P}_2) \pm Z_c \cdot \sqrt{\frac{\hat{P}_1 \cdot \hat{q}_1}{n_1} + \frac{\hat{P}_2 \cdot \hat{q}_2}{n_2}}$

Assume the samples are random and independent

Estimation $\qquad$ Test of hypotheses

point $\qquad$ C.I

Test of hypotheses : اختبار الفرضيات

* Hypotheses Testing :

* Null hypothesis , H_0 : فرضية العدم

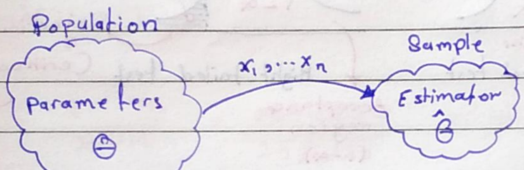
a statement about a population parameter that is assumed to be true until it is declared false. [contains equality $=, \leq, \geq$]

بديلة

* Alternative Hypothesis , H_a : نقيض فرضية العدم [$\neq, >, <$]

a statement about a population parameter that will be true if the null hypothesis is false.

(Null and Alternative Hypothesis الفرق بين) Example 1 لحل Page 372 يجب الرجوع الى الكتاب



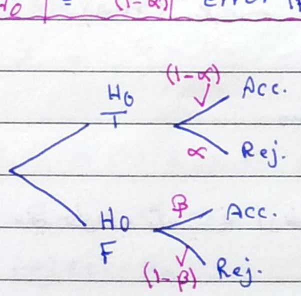
$H_0 : \Theta = \hat{\Theta}$ vs $H_1 : \underbrace{\Theta > \hat{\Theta} \text{ or } \Theta < \hat{\Theta}}_{\text{one tailed test}}$ or $\underbrace{\Theta \neq \hat{\Theta}}_{\text{2 tailed test}}$

* Types of errors :

	H_0 True	H_0 False
Reject H_0	Type I error (α)	رفض فرضية خاطئة ✓ $(1-\beta)$
(Accept) Don't Reject H_0 (Fail to Reject)	قبول فرضية صحيحة ✓ $(1-\alpha)$	Type II error (β)

* Type I error :- to reject a true statement
* Type II error :- to accept a false statement.

- α = P(Type I error)
  = P(Reject H_0 | H_0 is true)
  = significance level.
- β = P(Type II error)
  = P(Accept H_0 | H_0 false)



- $1-\beta$: power of the test.

- لا يمكن السيطرة على كلا α و β فنثبتوا احدهما (α)
ويقللوا β ← فنزيد $(1-\beta)$ ← فيكون Test أفضل.

- $1-\alpha$: confedence level

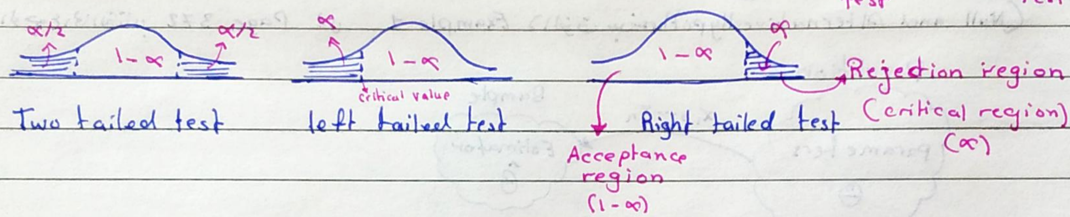
$\alpha + C = 1 \rightarrow C = 1-\alpha \rightarrow \alpha = 1 - C$

رح نقل Test :-

Test for μ	Test for P	Test for σ^2
$H_0: \boxed{\mu = \mu_0}$ vs $H_a: \mu \neq \mu_0$	$H_0: \boxed{P = P_0}$ vs $H_a: P \neq P_0$	$H_0: \boxed{\sigma^2 = \sigma_0^2}$ vs $H_a: \sigma^2 \neq \sigma^2$ } Two tailed test
or $H_0: \mu \leq \mu_0$ vs $H_a: \mu > \mu_0$	or $H_0: P \leq P_0$ vs $H_a: P > P_0$	$H_0: \sigma^2 \leq \sigma_0^2$ vs $H_a: \sigma^2 > \sigma_0^2$ } own tailed test
or $H_0: \mu \geq \mu_0$ vs $H_a: \mu < \mu_0$	or $H_0: P \geq P_0$ vs $H_a: P < P_0$	$H_0: \sigma^2 \geq \sigma_0^2$ vs $H_a: \sigma^2 < \sigma_0^2$

Right tailed test

left tailed test



Test statistic :- من خلال ← كيف بدي أخذ قرار :. الاست / standardized test statistic.

بس نحسبها ⇙ a function of $x_1, x_2 \ldots x_n$

اذا بكون ضمن . reject region بكون القرار reject H_0

واذا كان ضمن ال Accept region يكون القرار Acceptance.

* $\mu \to \bar{x} \rightsquigarrow \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim n(0,1)$, σ is known

$\dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$, σ is unknown $n < 30$

* $P \to \hat{P} \rightsquigarrow \dfrac{\hat{P} - P_0}{\sqrt{\dfrac{P_0 q_0}{n}}} \sim n(0,1)$
  * $n P_0 \geq 5$ and
  * $n q_0 \geq 5$ } بشرط

* $\sigma^2 \to S^2 \rightsquigarrow \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

Null Hypoth... من مؤخذهم عند المساواة.

⇐ في طريقة ثانية لاتخاذ القرار وهي من خلال :. P- Value .

Test for μ

eg) The mean cholestrol levels in a general population are normally distributed. A sample of 16 persons is taken under a test with mean $\bar{x} = 220$ mg/dL & standard deviation $s = 25$ mg/dL. Test at 1% significance level that the mean cholestrol level is less than 230 mg/dL.

sol) $n = 16$, $\bar{x} = 220$, $s = 25$, $\alpha = 0.01$

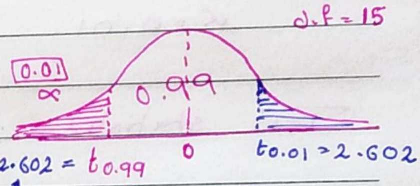
$H_0: \mu \geq 230$ vs $H_a: \mu < 230$ — بعض السؤال اعطا ال Alternative

left taild test

σ is unknown, $n < 30 \rightarrow$ t-distribution

Test statistic is :-

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{220 - 230}{25/\sqrt{16}} = -1.6$$



since $t = -1.6 > t_\alpha = -2.602$

$\therefore$ Don't reject H_0.

---

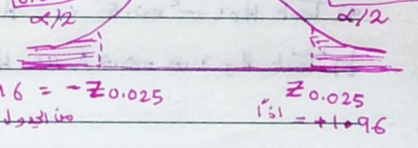
eg) A random sample of 400 people with professional degree taken showed that their mean monthly salary is 450JD with a standard deviation of 100 JD. Test at 5% significance level that the mean monthly salary is different from 460 JD.

sol) $n = 400$ (very large), $\bar{x} = 450$, $s = 100$, $\alpha = 0.05$

$H_0: \mu = 460$ vs : $H_a: \mu \neq 460$ → two tailed test

Test statistic is :

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{450 - 460}{100/\sqrt{400}} = -2$$



$\frac{\alpha}{2} = 0.05/2 = 0.025$

since $Z = -2 < -Z_{0.025} = -1.96$, then reject H_0

Test for $\hat{p}$:-

eg) It was believed in the Arab World that 50% of Persons are smoking. During the year 2000, a sample of 1000 persons showed that the no. of smokers is 620. Can you conclude that the proportion of smokers is different from 50% ? use $\alpha = 0.01$

sol) $n = 1000$, $x = 620$ $\rightarrow$ $\hat{p} = \frac{x}{n} = \frac{620}{1000} = 0.62$

$H_0: P = 0.50$ vs $H_a: P \neq 0.50$

$\alpha = 0.01$.

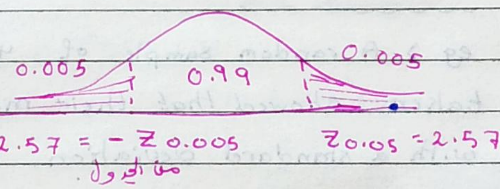
Test statistic is :-

$$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.62 - 0.50}{\sqrt{\frac{0.50(0.50)}{1000}}} = 7.59$$

$\alpha = 0.01 \rightarrow \frac{\alpha}{2} = 0.005$

since $Z = 7.59 > Z_{0.005} = 2.57$

then reject H_0



Test for σ^2 :-

eg) Quality-control enginer wishes to study the weight variation of a new product. A sample of 10 items is taken & provided $\bar{X} = 0.6$ kgs & $s = 0.4$ kgs. Assume that the distribution of the weights can be modeled as a normal distribution

a) Test $H_0: \sigma^2 = 0.5$ vs $H_1: \sigma^2 > 0.5$ (use $\alpha = 0.025$)

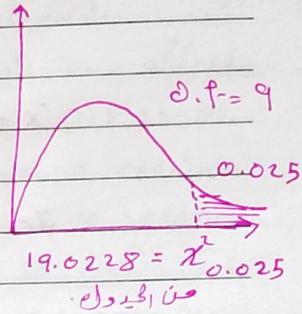
b) Test $H_0: \sigma = 0.7$ vs $H_1: \sigma \neq 0.74$ (use $\alpha = 0.10$)

sol) $n = 10$, $\bar{x} = 0.6$, $S = 0.4$

(a) $H_0: \sigma^2 = 0.5$ vs $H_a: \sigma^2 > 0.5$ $(\alpha = 0.025)$

Test statistic is:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(10-1)(0.4)^2}{(0.5)} = 2.88$$



since $\chi^2 = 2.88 < \chi^2_{0.025} = 19.0228$,
then don't reject H_0.

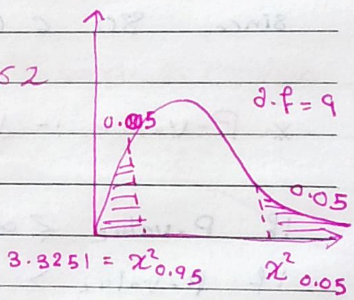
(b) $H_0: \sigma = 0.74$ vs $H_a: \sigma \neq 0.74$

or

$H_0: \sigma^2 = 0.55$ vs $H_a: \sigma \neq 0.55$

Test statistic is:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(10-1)(0.4)^2}{(0.55)} = 2.62$$



$\alpha = 0.10 \rightarrow \frac{\alpha}{2} = 0.05$

since $\chi^2 = 2.62 < \chi^2 = 3.3251$
then we ~~don't~~ reject H_0.

* Relationship between C.I & tests. → two tailed

Let (L, U) be $(1-\alpha)100\%$ C.I for unknown parameter θ.
the null hypothesis $H_0: \theta = \theta_0$ is rejected againts $H_1: \theta \neq \theta_0$ at significance level α if θ_0 doesn't belong to (L, U).

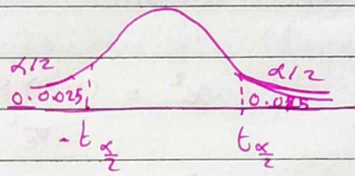
eg) A random sample of 8 observations was taken from a normal population. The sample mean & standard deviation are $\bar{x} = 70$ & $S = 20$.
Find a 95% C.I for μ & test at 5% significance level
$H_0: \mu = 80$ vs $H_a: \mu \neq 80$

COLLEGE

Sol) $n = 8$, $\bar{x} = 70$, $s = 20$

$\boxed{c = 1 - \alpha = 0.95}$ $\rightarrow \alpha = 0.05 \rightarrow \alpha/2 = 0.025 = \frac{1}{2}(1-c)$

$\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$



$70 \pm (2.365)\left(\frac{20}{\sqrt{8}}\right)$ $= -t_{(0.025+0.95)}$ $t_{0.025} = 2.365$

the 95 % C.I for μ is $(53.29 , 86.71)$

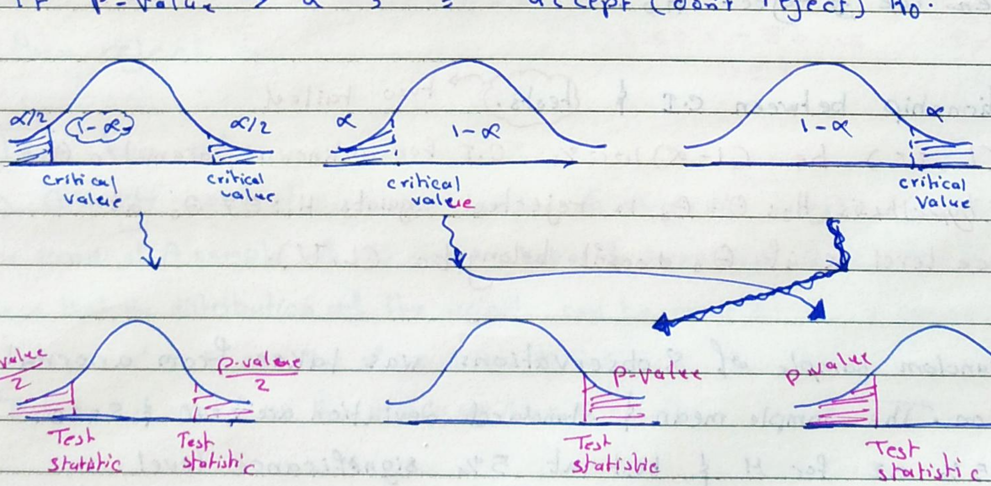
المطلوب التالي جا انه two tailed test و $H_a(\neq)$ وعند النسبة 5%
فمنطل إلى 80 اذا داخل الفترة او خارجها ← خبر انها ضمن C.I

since $80 \in (53.29, 86.71) \rightarrow \therefore$ don't reject H_0.

* P-value :- طريقة ثانية للاختبار (Test) وعادةً ما يُعطى بالسؤال

يجب الرجوع الى الكتاب (Page 372)

If P-value $\leq \alpha$, then reject H_0.
If P-value $> \alpha$, " accept (don't reject) H_0.



Example 4 Page 389

في سباق

اللي يوقفوا عندها لعشان يعني حتى يتأكدوا من عجلاته، ليا النقطة

In auto racing, $\boxed{\text{a pit stop}}$ is where a racing vehicle stops for new tires, fuel, repairs, and other mechanical adjustments. The efficiency of a pit crew that makes these adjustments can affect the outcome of a race. A pit crew claims (الادعاء) that its mean pit stop time (for 4 new tires and fuel) is less than 13 seconds. A random sample of 32 pit stop times has a sample mean of 12.9 seconds. Assume the population standard deviation is 0.19 second. Is there enough evidence to support the claim at $\alpha = 0.01$? [use a P-value].

$H_0: \mu \geq 13$

$H_a: \mu < 13$ (claim)

Alternative

(equality)

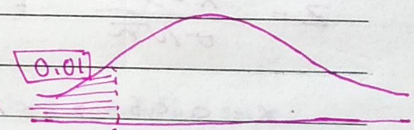
$n = 32$, $\bar{x} = 12.9$, $\sigma = 0.19$

$\alpha = 0.01$

Test statistic :- Because σ is known → the sample is random and $n \geq 30$

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{12.9 - 13}{0.19/\sqrt{32}} = -2.98$$

⇒ Critical value اذا لانه يدنا كل ده ⇒



Left tailed test

$-Z_c = -Z_{0.01} = -2.33$

standard normal

من جدول المعالم

0.03

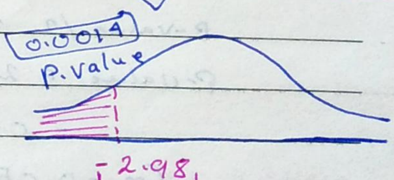
$-2.3 \ldots \boxed{0.01}$

since $-2.98 < -Z_c = -2.33$ ⇒

reject H_0 → Accept H_a (the claim)

⇒ P. value $= 0.0014 < \alpha = 0.01$

then reject H_0 ⇒ Accept H_a



0.08

$-2.9 \ldots \boxed{0.0014}$

Hypothesis Testing Using a P-Value.

Test for μ:-

eg) According to a study of U.S homes that use heating equipment the mean indoor temprature at night during witer is 68.3°F. You think this information is incorrect. You randomly select 25 U.S. homes that use heating equipment at night is 67.2°F. From Past studies, the population standard deviation is known to be 3.5°F and the Population is normally distributed. Is there enough evidence to support your claim at $\alpha = 0.05$? [Use a P-value]

Solution: $n = 25$, $\sigma = 3.5°F$ (is known) $\Rightarrow$ Z-test

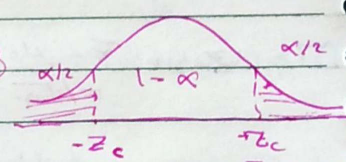
$\bar{x} = 67.2°F$, $\alpha = 0.05$

claim $\rightarrow$ the mean is different from 68.3°F

$H_0: \mu = 68.3°F$ $\qquad H_a: \mu \neq 68.3°F$ (claim)

The standardized test statistic is:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{67.2 - 68.3}{3.5/\sqrt{25}} = -1.57$$



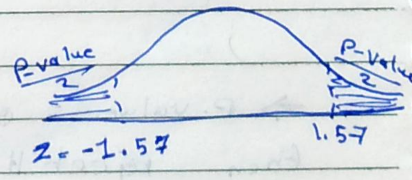
$\alpha = 0.05 \rightarrow \alpha/2 = 0.025$

$-1.96 = -Z_{0.025}$ من الجدول

$Z_{0.025} = 1.96$

since $-Z_{0.025} = -1.96 < -1.57 < Z_{0.025} = +1.96$

$\rightarrow$ Accept H_0 $\rightarrow$ reject the claim



$P\text{-value}/2 = 0.0582$

$P\text{-value} = 2(0.0582)$

$= 0.1164$

$(\alpha = 0.05)$

	0.07
-1.5	.0582

لاحظ ان P-value $> \alpha$ $\rightarrow$ Accept H_0 $\rightarrow$ reject the claim eg) A used car dealer says that the mean price of used cars sold in the last 12 months is at least (claim) 21.000\$. You suspect this claim is incorrect and find that a random sample of 14 (n) used cars sold in the last 12 months has a mean price of 19,189 \$ ($\bar{x}$) and a standard deviation of 2950 \$ (s). Is there enough evidence to reject the dealer's claim at $\alpha = 0.05$? Assume the population is normally distributed.

$n = 14 \rightarrow d.f = 13$, $\bar{x} = 19,189$ & $s = 2950$ (σ is unknown) → t-test

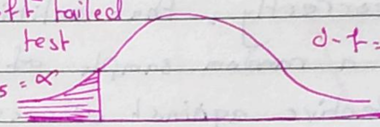
$\alpha = 0.05$

$H_0: \mu \geq 21,000$ (claim) & $H_a: \mu < 21,000$

at least

left tailed test

Test statistic:-



$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{19,189 - 21,000}{2950/\sqrt{14}}$$

لان لا يوجد قيم سالبة t

$t = -2.297$

since $-2.297 < -1.771$ → reject H_0 (claim)

Test for $\hat{p}$:-

eg) A researcher claims that less than 45% of U.S. adults use passwords that are less secure because complicated ones are too hard to remember. In a random sample of 100 adults, 41% say they use passwords that are less secure because complicated ones are too hard to remember. At $\alpha = 0.01$, is there enough evidence to support the researcher's claim?

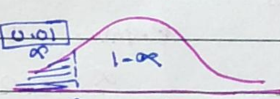
$H_0: p \geq 0.45$ & $H_a: p < 0.45$ (claim) (left)

$n = 100$, $\hat{p} = 0.41$, $\alpha = 0.01$

$nq_0 \rightarrow 100(0.55) = \boxed{55 > 5}$ and $\boxed{5 < 45} = 100(0.45) \leftarrow np_0$ لازم نتأكد من ان

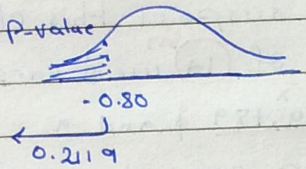
Test statistic:-

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.41 - 0.45}{\sqrt{\frac{(0.45)(0.55)}{100}}} = -0.80$$



$-2.33 = -z_{0.01}$ Accept H_0 → reject the claim.

حل المثال السابق باستخدام P-value



P-value = 0.2119

$\alpha = 0.01$

P-value > α → ∴ Accept H_0

→ Reject the claim.

There is not enough evidence at the 1% level of significanse to support the claim that less than 45% of U.S. adults us Passwords ...

eg) A researcher claims that 51% of U.S. adults believe, incorrectly, that antibiotics are effective against viruses. In a random sample of 2202 adults, 1161 say antibiotics are effective against viruses. At $\alpha = 0.10$, is there enough evidence to support the researcher's claim? (there is enough evidence to support the claim.)

$p_0 = 51\% \rightarrow q_0 = (0.49)$, $n = 2202$

$np_0 = 1123$ $nq_0 = 1079$ ← are both greater than 5.

The Claim → "51% of U.S. adults ... viruses"

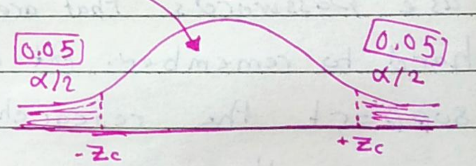
$H_0 : P = 0.51$ (claim) & $H_a : P \neq 0.51$

↳ two tailed test

$$z = \frac{\hat{P} - P_0}{\sqrt{P_0 q_0 / n}} = \frac{0.527 - 0.51}{\sqrt{(0.51)(0.49)/2202}} \approx 1.60 \quad , \quad \hat{P} = \frac{x}{n} = \frac{1161}{2202} = 0.527$$

Accept H_0

→ Accept H_0 → Accept the claim.

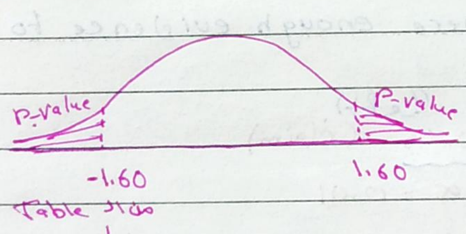


$\boxed{-1.645} = \frac{-3.29}{2} = -Z_{0.05}$ $+Z_{0.05}$

من جدول النابل ← بين 0.0495 و 0.0505

$\boxed{+1.645}$

$\boxed{\frac{-1.65 + -1.64}{2}}$



P-Value = 0.0548

$\alpha = 0.10$

P-value > α → Accept H_0 → Accept the claim.