CONTENTS

Z-TEST (CRITICAL/TRADITIONAL METHOD)

Date ______ 20 ___
M T W T F S S

Estimation and hypothesis testing
Book : ES Chap 8.1, 8.2, 8.3.
(Traditional Method.)

α = level of significance (critical Region)
$1-\alpha$ = Non-critical Region.
α = value of Z

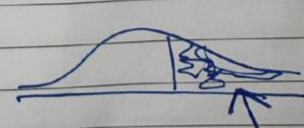
Steps in Solving

1. State Null and Alternative Hypothesis (H_0, H_1)
   $\mu = \mu_0$: H_0
   $\mu \neq \mu_1$: H_1

Two tailed test
$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_1$
$\mu > \mu_0$ or $\mu < \mu_0$

One tailed test
$H_0 : \mu = \mu_0$
$H_1 : \mu < 30$ → Left tailed
$H_1 : \mu > 30$ → Right tailed.

Critical values:
→ Shaded Area = Critical Region.
→ Divides curve into critical and non-critical region.
→ Find value of Z.



· Descision:
→ Accept or Reject H_0.

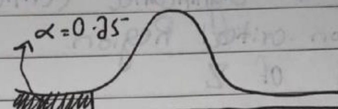
Date ______ 20 ___

Example:

Left tailed test:

$\Rightarrow P(z<Z) = \cancel{P(z<Z)}\ \alpha$

~~= 0.25~~



Right tailed:

$P(z > Z) = 1 - \alpha$

· Two tailed test:

$P(z < Z) = \alpha/2$

$P(z > Z) = 1 - \alpha/2$

· We use Z table to solve when:
Population std σ is known.

Test value formula (central limit):

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

P-VALUE METHOD FOR SINGLE MEAN WHEN DEVIATION IS KNOWN (Z -TEST)

Date ______ 20__
M T W T F S S

P-value method for single μ H.T; when σ is known.

* Steps:

① $H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

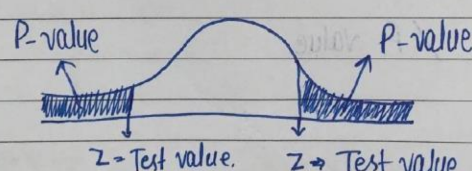
$\mu > \mu_0$

$\mu < \mu_0$

② Test value

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

③ Find P-value:



· Different for, left, right and two-tailed test.

④ Desicsion making:

i) P-value $\leq \alpha$ → Reject H_0

ii) P-value $> \alpha$ → Accept H_0

⑤ Results.

Page #

DIFFERENCE BETWEEN 2 MEANS, WHEN DEVIATIONS IS ARE KNOWN (Z-TEST)

Date ______ 20___

Hypothesis test for difference b/w two mean's; μ_1 and μ_2; when σ_1 and σ_2 are known. (z-test)

* steps *

1) $H_0: \mu_1 = \mu_2 \Rightarrow H_0: \mu_1 - \mu_2 = 0$

   $H_1: \mu_1 \neq \mu_2 \Rightarrow H_1: \mu_1 - \mu_2 \neq 0 \rightarrow$ 2 tail

   $H_1: \mu_1 - \mu_2 > 0 \rightarrow$ Right tail

   $H_1: \mu_1 - \mu_2 < 0 \rightarrow$ Left tail.

2) Test value : σ_1, σ_2 known:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \underbrace{(\mu_1 - \mu_2)}_{\text{Alway equal to } 0}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

3) Critical value / P- value

4) Make descision

5) Result.

T-TEST

Date ____ 20__
M T W T F S S

Hypothesis test for single μ; when σ is unknown. (t-test)

* steps *

- All step are same except:
  for test value we use
- $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$
- C.V is used only.
- No P-value method.
- Two variables required for t-table.
  α , d.f $\Rightarrow n-1$.

d.f	α
⋮	⋮

For two means:

Test value:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Dependent:

Test value:

$$t = \frac{\bar{D} - \mu_D}{\frac{s_D}{\sqrt{n}}}$$

$\bar{D} = \bar{x}_1 - \bar{x}_2$

$\mu_D = \mu_1 - \mu_2 = 0$

$$s_D = \sqrt{\frac{n(\Sigma D^2) - (\Sigma D)^2}{n(n-1)}}$$

CONFIDENCE INTERVAL

Date _____ 20___

Confidence Interval

- Alway 2 tailed test.

- $a \le \mu \le b$
- If $\mu = \mu_0$ lies in C.I we accept H_0 else we reject it.

★ Formula's for Confidence Interval ★

1) For single mean, when σ is known. (z-test)

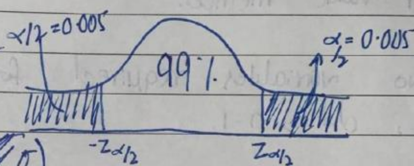
- $\bar{x} - Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

Example: C.I $\Rightarrow$ 99% ; σ is known.

$\alpha = 0.01$ (1%)

$\alpha/2 = 0.005$.

$z = -2.58$



$\Rightarrow \bar{x} - 2.58\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + 2.58\left(\frac{\sigma}{\sqrt{n}}\right)$

2) For single mean when, σ is unknown. (t-test)

- $\bar{x} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$.

3) For diff b/w μ_1 and μ_2, when σ_1 and σ_2 are known.

- $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ $\Rightarrow (\bar{x}_1 - \bar{x}_2) - Z_{\alpha/2}\left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + Z_{\alpha/2}\left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$

4) For difference b/w μ_1 and μ_2 ; when σ_1 and σ_2 are unknown. (t-test.)

i) Independent:

- $(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right) < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$.

ii) Dependent:

$\bar{D} - t_{\alpha/2}\left(\frac{S_D}{\sqrt{n}}\right) < 0 < \bar{D} + t_{\alpha/2}\left(\frac{S_D}{\sqrt{n}}\right)$

LINEAR REGRESSION

FORMULA OF REGRESSION LINE

Linear Regression

→ Regression is a method that shows relation between a and b.

→ +ve /-ve ; linear /non-linear.

- $y = mx + C$ → equation of a straight line.
  m = slope.
  C = y-intercept.

- slope (m) can be +ve, -ve, 0, or ∞
  +ve = line upward. → $y = mx + c$
  -ve = line downwards. → $y = -mx + c$
  0 = y-intercept → $y = c$
  ∞ = vertical line.

* Eqn of Regression line

$$y = b_0 + b_1 x$$

b_0 = y-intercept
b_1 = slope.

$$b_0 \Rightarrow \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b_1 \Rightarrow \frac{(n)(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

~~3.39~~

- Scatter Plot with Regression line:

Example: Walpole pg No: 393

x	3	7	11	15	18	27	29	30	30	31	31	32	33	33
y	5	11	21	16	16	28	27	25	35	30	40	32	34	32

x	34	36	36	36	37	38	39	39	39	40	41	42
y	34	37	38	34	36	38	37	36	45	39	41	40

x	42	43	44	45	46	47	50
y	44	37	44	46	46	49	51

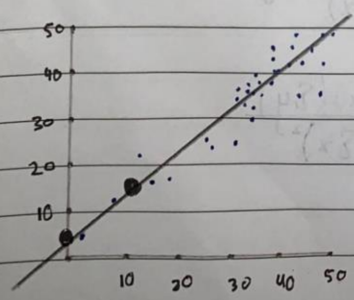
Sol: $n = 33$ $\quad \Sigma x : 1104$ $\quad \Sigma y : 1124$ $\quad \Sigma x^2 : 41086$ $\quad \Sigma xy : 41355$

$$b_0 = \frac{(\Sigma y) \cdot (\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b_1 = \frac{(n)(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b_0 = \frac{(1124)(41086) - (1104)(41355)}{33(41086) - (1104)^2} \Rightarrow 3.8296$$

$$b_1 = \frac{(33)(41355) - (1104)(1124)}{(33)(41086) - (1104)^2} \Rightarrow 0.9036$$



$y = 3.8296 + 0.9036\,x$

$y = b_0 + b_1 x$

+ve relation

x :	0	10
y :	3.8296	12.86

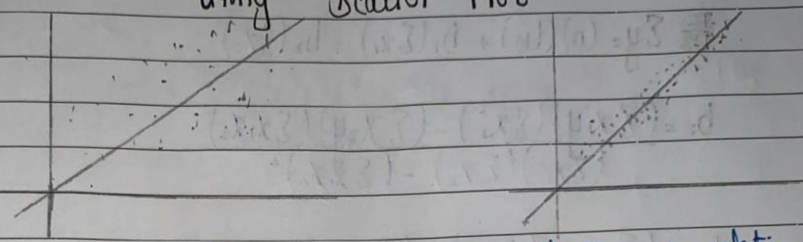
Relation is +ve.

CO-RELATION COEFFICIENT (R) AND COEFFICIENT OF DETERMINATION (R^2)

Co-relation coefficient (r)

It shows the direction and strength of relation.

Using Scatter Plot



weak +ve relation

strong +ve relation.

Formula of linear regression (r)

$$r = \frac{(n)(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[(n)(\Sigma x^2) - (\Sigma x)^2][(n)(\Sigma y^2) - (\Sigma y)^2]}}$$

$r = [-1, 1]$

1 = strong +ve relation.

0 = no Relation.

-1 = strong -ve relation.

-0.5 = Moderate -ve

-0.3 = weak -ve

+0.8 = strong +ve

+0.5 = Moderate +ve

Co-efficient of Determination (r^2)

C. of. D = $(r^2) \times 100\%$

MULTIPLE REGRESSION

FORMULA FOR MULTIPLE REGRESSION

* Equation:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_kx_k$$

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2$$

$$\sum X_1Y = a\sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1X_2$$

$$\sum X_2Y = a\sum X_2 + b_1 \sum X_1X_2 + b_2 \sum X_2^2$$

Solve of values using calculator

MULTIPLE CO-RELATION (R) AND CO-EFFICIENT OF MULTIPLE DETERMINATION

Multiple Co-relation (R)

* Formula:

$$\sqrt{\frac{r^2(yx_1) + r^2(yx_2) - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r^2(x_1x_2)}}$$

- $R \Rightarrow [0,1]$
- R can never be negative.

Co-efficient of Multiple Determination (R^2)

C. of. MD $\Rightarrow R^2 \times 100\%$

ANOVA

Chapter no # 18 : Elementary statistics
ANOVA

- Analysis of Variance.

$$\Rightarrow \sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

- ANOVA always right tailed test.
- More than 2 samples are taken from the given population.

x_1	x_2	x_3	x_4
⋮	⋮	⋮	⋮
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$

$H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$

H_1 = Atleast one mean is different from others.

- ANOVA is solved by F-test, (f-table)

*· Steps *

1) $H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$

   H_1 : Atleast one mean is different.

2) Test value

$$F = \frac{S_B^2}{S_W^2} \quad \because S^2 = \text{sum of squares.}$$

a) Find mean of all data sets.

$\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, $\bar{x}_4$

b) Find sample variance of each

S_1^2, S_2^2, S_3^2, S_4^2

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

c) $\bar{X}_{GM} \Rightarrow \frac{\text{Sum of All values}}{\text{Total values.}}$

d) $$S_B^2 = \frac{\sum n_i (\bar{X}_i - \bar{X}_{GM})^2}{K-1}$$

e) $$S_W^2 = \frac{\sum (n_i - 1) S_i^2}{N-K}$$

f) $$F = \frac{S_B^2}{S_W^2}$$

3) Critical value from (F-table)

α, v_1, v_2

$v_1 = K-1$; $v_2 = N-K$

K = No of Samples

N = Total No of objects.

	α
v_2	v_1

4) Make descision

5) Result.