Chapter = 03

Small sample Test

$n < 30$

→ **Student's t - test =**

Test of hypo based on the Student's t-distri → t test. It is used as test of significance in following cases —

1) To test the significance of mean of small sample from normal (P).
2) To test the significance of difference b/w the means of 2 independent samples taken from a normal (P).
3) To test the significance of the difference b/w means of 2 dependent samples taken from a normal (P).
4) To test the significance of an observed correlation (coe).
5) To test the significance of an observed regression (coe).

→ **t - test for (P) mean =**

It is using Student's t-test. The following assumptions must be made.

1) Parent (P) from which the sample is drawn is normal.
2) Sample obsr are independent & random.
3) Sample should be small ($n<30$)
4) (P) S.D σ is unknown.

By testing the mean of normal (P) we are actually testing the significance difference b/w sample mean & the hypothetical of μ.

Here we are testing $H_0 : \mu = \mu_0$ against 1 of the alternatives

$H_1 : \mu < \mu_0$ or $\mu > \mu_0$ or $\mu \neq \mu_0$.

Test statistic,

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n-1}} \quad \text{where } S = \frac{1}{n} \sum (x_i - \bar{x})^2.$$

For the significance level α, BCR are

$\omega = t < -t_\alpha$, $\omega = t > t_\alpha$ & $\omega = |t| \geq t_{\alpha/2}$

where t_α is obtained referring the t-table for $(n-1)$ DF.

Now cal value of t using the sample data if it lies in CR reject H_0 otherwise accept

Q) A sample of 10 obsr. gives a mean $= 38$ & SD $= 4$. Can we conclude that (P) mean is 40

A) $n = 10$. $\bar{x} = 38$, $S = 4$, $\mu_0 = 40$.

we have to test.

$H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$

$t_{\alpha/2} = 2.262$.

let $\alpha = 0.05$,

BCR, $\omega = |t| \geq t_{\alpha/2}$

from t-table $t_{\alpha/2}$ for 9 DF & $\alpha = 0.05$

$\hookrightarrow 10 - 1 = n - 1 = 9$

$|t| \geq 2.262$

$$t = \frac{38 - 40}{4/3} = -1.5$$

$|t| = 1.5 < 2.262$

accept

2) A r.s of size 16 has 53 as mean & sum of □'s of deviation taken from mean is 150. can the sample be regarded as arisen with mean 56.

A) $n = 16$, $\bar{x} = 53$, $\sum (x_i - \bar{x})^2 = 150$, $\mu_0 = 56$

$ns^2 = 150$

$s^2 = \frac{150}{n} = \frac{150}{16} = 9.375$

(SD) → $s = 3.06$

$H_0: \mu_0 = 56$ against $H_1: \mu \neq 56$.

BCR, $w = |t| \geq t_{\alpha/2}$.

$(n-1)$ DF $= 15$ → 2.947

$$t = \frac{53 - 56}{3.06/\sqrt{15}} = \frac{-3}{0.79} = -3.79$$

$|t| = 3.79 \geq 2.947$

reject ~~accept~~

t-test for equality of 2 (p) mean =

By testing the equality of 2 (p) mean we mean the test of significance of 2 sample mean.

Assumptions -

* 2 (p) from which the samples are drawn should follow normal distri—
* sample obs should be independent & random.
* samples should be small. ($n_1 < 30$ & $n_2 < 30$)
* 2 (p) variance σ_1^2 & σ_2^2 must be = but unknown
* Here we are testing $H_0: \mu_1 - \mu_2 = 0$ against 1 of the alternatives $H_1: \mu_1 - \mu_2 < 0$, $\mu_1 - \mu_2 > 0$, $\mu_1 - \mu_2 \neq 0$

~~test statis~~

test statis,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$s_1^2 = \frac{1}{n_1} \sum (x_i - \bar{x}_1)^2$

$s_2^2 = \frac{1}{n_2} \sum (x_j - \bar{x}_2)^2$

For $\alpha = 0.05$, BCR, $w = t < t_\alpha$, $t > t_\alpha$ & $|t| \geq t_{\alpha/2}$ where t_α is obtained by refering t-table for $(n_1 + n_2 - 2)$ DF.

Now cal value of t, if it lies in the CR, reject H_0 otherwise accept H_0.

Q) The mean life of a sample of 10 electric bulbs was observed to be 1309 hrs with the S.D of 420 hrs. A 2nd sample of 16 of different batch showed a mean life of 1205 hrs with SD of 390 hrs. Test whether there is a significant difference b/w means (use $\alpha = 0.05$)?

A) Assume that the (p) from which the samples are drawn are normal.

$n_1 = 10$, $n_2 = 16$, $\bar{x}_1 = 1309$, $\bar{x}_2 = 1205$

$s_1 = 420$, $s_2 = 390$.

we have to test,

$H_0: \mu_1 - \mu_2 = 0$ against $H_1: \mu_1 - \mu_2 \neq 0$

BCR, $w = |t| \geq t_{\alpha/2}$.

$(n_1 + n_2 - 2)$ DF $= 10 + 16 - 2 = 24$
$\alpha = 0.05$ } table. pg-13

∴ $t_{\alpha/2}$ → 2.064.

test statistics,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{1309 - 1205}{\sqrt{\frac{10 \times 420^2 + 16 \times 390^2}{10 + 16 - 2}\left(\frac{1}{10} + \frac{1}{16}\right)}}$$

$$= \frac{104}{\sqrt{\frac{17640000+2433600}{2}}\sqrt{\frac{1}{10}+\frac{1}{16}}}$$

$$\alpha = \frac{104}{3168.090908 \times 0.40311} = 0.616$$

$0.616 \leq 2.064$ Accept

e) following are samples from 2 independent normal (P). Test the hypo that they have the same. assuming that the V(X) are = by taking the level of significance as 5%

x_1: Sample 1 : 14 18 12 9 16 24 20 21 19 17

x_2: Sample 2 : 20 24 18 16 26 25 18.

A) $\bar{x}_1 = \frac{\sum x_i}{n_1} = \frac{170}{10} = 17$, $\bar{x}_2 = \frac{\sum x_i}{n_2} = 21.$

x_1	x_2	$(x_1-\bar{x}_1)$	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)$	$(x_2-\bar{x}_2)^2$
14	20	−3	9	−1	1
18	24	1	1	3	9
12	18	−5	25	−3	9
9	16	−8	64	−5	25
16	26	−1	1	5	25
24	25	7	49	4	16
20	18	3	9	−3	9
21		4	16		
19		2	4		
17		0	0.		
170	21		178		94

$n_1 s_1^2 = 178$

$n_2 s_2^2 = 94$

$$t = \frac{17-21}{\sqrt{\frac{178+94}{15}\left(\frac{1}{10}+\frac{1}{7}\right)}} = \frac{-4}{\sqrt{18.1333 \times 0.2428}}$$

BCR $w = |t| \geq t_{\alpha/2}$ $= -1.906$

$|t| \geq 2.131.$ $-1.906 \leq 2.131$

accept

→ t-test for dependent samples = (paired t-test)

when we have the data obtained by observing the values of same attribute of different elements of (P), observed "before & after" A typical case could be measure a certain reaction in human scale. before (x_i) & after (y_i) some treatment.

The purpose is to decide whether the treatment has an effect/not.

Imagine that the observation are such that the have the same difference, $d_i = x_i - y_i$, $i = 1, 2 \ldots n$ with mean μ & σ^2 unkn normal distri–

To test null hypo, $H_0 : \mu = \mu_0$ against 1/more 2 sided alternatives $H_1 : \mu > \mu_0$ or $\mu \neq \mu_0$. In most typical appli we set up hypo– as $H_0 : \mu = 0$ against $H_1 : \mu \neq 0$

Test statistic is, $t = \frac{\bar{d}}{s_d / \sqrt{n-1}}$, $\bar{d} = \frac{\sum d_i}{n}$

$s_d \rightarrow$ S.D.

$s_d = \sqrt{\frac{1}{n}\sum d_i^2 - \bar{d}^2}$

$d_i = x_i - y_i$

Table value for n obtained by (n−1) of & (prob) $\alpha = 0.05$.

Q
1) To test the efficiency of sleeping pills, a drug company uses a sample of insomniacs the time in min until falling asleep is observed for each of them. few days later the same persons are given a sleeping pill & time until falling asleep is measured again. The measurement

persons : A B C D E

No pill (x_i) : 65 35 80 40 50 } dependent

with pill (y_i) : 45 15 61 31 20. } < 30. Paired sample test.

Test whether sleeping pills are effective.

Assume that ~~sleeping pills are not effective.~~

$H_0 : \mu = 0$ against $H_1 : \mu \neq 0$

$\alpha = 0.05$, $w = |t| \geq t_{\alpha/2}$.

$n = 5$

$(n-1)$ Df

table $\left\{ \begin{array}{l} 4 \text{ Df.} \\ \alpha = 0.05 \end{array} \right.$ $t_{\alpha/2} = 2.775$

x_i	y_i	$d_i = y_i - x_i$
65	45	-20
35	15	-20
80	61	-19
40	31	-9
50	20	-30

$\bar{D} = \dfrac{-20-20-19-9-30}{5} = -19.6.$

$SD = 10.77$ $n = 5$

$t = \dfrac{\bar{D}}{Sd/\sqrt{n-1}} = \dfrac{-19.6}{10.77/\sqrt{5}} = -3.39$

$(n-1)$ DF $= 4$, $\alpha = 0.05$ } $t_{\alpha/2} = 2.776.$

$2.776 \ngtr 3.39$

Reject

Sleeping pills are effective in reducing the time to fall asleep

A)

x_i	y_i	$d_i = x_i - y_i$	d_i^2	~~$d_i^2 - \bar{d}^2$~~	$d_i^2 - \bar{d}^2$
65	45	20	400	~~379.6~~	15.84
35	15	20	400	~~379.6~~	15.84.
80	61	19	361	~~420.8~~	-23.16
40	31	9	81	~~100.6~~	-303.16
50	20	30	900	~~896.6~~	515.84
		~~98~~	2142	~~2160~~	221.2

$S_d = \sqrt{\dfrac{1}{n} \sum d_i^2 - \bar{d}^2}$

$= \sqrt{\dfrac{1}{5} \times 221.2} = \sqrt{44.24} = 6.65$

$t = \dfrac{\bar{d}}{Sd/\sqrt{n-1}}$

$= \dfrac{19.6}{6.65/2} = 5.89$

$5.89 \geq 2.775$

Reject

$\alpha = 0.05$

$w = |t| \geq t_{\alpha/2}$

$n = 5$ $t_{\alpha/2} = 2.775$

$df = 4$

$\bar{d} = 19.6$

$\bar{d}^2 = 384.16$

⇒ F-Test in testing of hypothesis :

* used for testing the quality of V(x) of 2 normal (P) / whether 1 (P) V(x) is > the other.

* Test procedure as follows –
Draw a random sample of size n_1 from $N(\mu_1, \sigma_1)$ & a sample of size n_2 from $N(\mu_2, \sigma_2)$.
Let S_1^2 & S_2^2 be their sample V(x).

$$\boxed{F = \frac{\sigma_1^{*2}}{\sigma_2^{*2}} = \frac{\frac{n_1 S_1^2}{n_1 - 1}}{\frac{n_2 S_2^2}{n_2 - 1}}}$$

* This statistic follows Snedecor's F distri with $(n_1 - 1, n_2 - 1)$ DF.
the F-test is performed as a right tailed test, so if $S_1^2 > S_2^2$, the hypo,
$H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 > \sigma_2^2$.
Let $\alpha = 0.05$ or 0.01.
test statistic,

$$F = \frac{n_1 S_1^2 / (n_1 - 1)}{n_2 S_2^2 / (n_2 - 1)}.$$

BCR is gvn by $\underline{\omega \equiv F > f_\alpha}$ where f_α is obtained by referring F table for (n_1-1, n_2-1) df with $P(F > f_\alpha) = \alpha$. cal value of F & if it exceeds the table value, H_0 is rejected otherwise H_0 is accepted.

Q
1) Sample size 10 & 18 taken from 2 normal (p) give $S_1 = 14$, $S_2 = 20$. Test the hypo. that the samples have come from (P)s with same S.D?

A) $n_1 = 10$, $n_2 = 18$, $S_1 = 14$, $S_2 = 20$

① $\underline{S_1^2 > S_2^2}$ $\rightarrow 14^2 > 20^2$ ✗ $\rightarrow S_1 < S_2$

So interchange S_1 & S_2.

also interchange n_1 & n_2.

∴ $n_1 = 18$, $n_2 = 10$, $S_1 = 20$, $S_2 = 14$ $\Rightarrow S_1^2 > S_2^2$ ✓

② we have tst,

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

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

Let $\alpha = 0.05$,

BCR, $\omega = F > F_\alpha$.

table value

$((n_1-1)(n_2-1)) = (18-1, 10-1) = (17, 9)$ df

(prob) $\rightarrow 0.05$

$\rightarrow 3 = F_\alpha$.

pg-5

③ $F = \dfrac{\frac{n_1 S_1^2}{n_1-1}}{\frac{n_2 S_2^2}{n_2-1}}$

$= \dfrac{18 \times 20^2 / 17}{10 \times 14^2 / 9}$

$= \dfrac{423.529}{217.777}$

$= 1.944$

$1.944 < 3$

$\underline{\text{accept}}$

2) 2 r.s 8 & 11 drawn from 2 normal (p) are characterised as follows.

Sample size	Sum of obs^r (x_i)	Sum of O's of obs^r (x_i^2)
8	9.6	61.52
11	16.5	73.26

Examine whether 2 samples come from (P) having the same v(x).

A) $$\boxed{\begin{aligned} S_1^2 &= \frac{\Sigma x_1^2}{n_1} - \bar{x}_1^2 \\ S_2^2 &= \frac{\Sigma x_2^2}{n_2} - \bar{x}_2^2 \end{aligned}}$$

$\bar{x}_1 = \frac{9.6}{8} = 1.2$

$\bar{x}_2 = \frac{16.5}{11} = 1.5$

$S_1^2 = \frac{61.52}{8} - 1.2^2$

$= \underline{6.25}$

$S_2^2 = \frac{73.26}{11} - 1.5^2 = \underline{4.41}$

$H_0: \sigma_1^2 = \sigma_2^2$ against $H_1: \sigma_1^2 > \sigma_2^2$

Let $\alpha = 0.01$

$\omega = F > F_\alpha$.

$S_1^2 > S_2^2$ ✓

table $((n_1-1)(n_2-1)) = (7, 10)$ df

Prb $\rightarrow 0.01$

$F_\alpha = 5.21$

$F = \dfrac{\frac{n_1 S_1^2}{n_1-1}}{\frac{n_2 S_2^2}{n_2-1}} = \dfrac{8 \times 6.25 / 7}{11 \times 4.41 / 10}$

$= 1.472$

pg-6

$1.47 < 5.21$ accept

⇒ χ^2. test =

* An imp non-paramtric test is χ^2 test.
* χ^2 test is 1 of the simplest & most commonly used non-paramtric test of significnce.
- Imp appli- are –

a) To test the v(x) of normal (p).
b) " the goodness of fit.
c) " the indepndnce of attributes.

I χ^2 test for (p) v(x) =

* This test is conducted whn we want to test if the gvn normal (p) has a specified v(x)., $\sigma^2 = \sigma_0^2$
* χ^2 test for v(x) is generally a right paired test, so we are testing $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma_1^2 > \sigma_0^2$
* Test statistic is,

$$\boxed{\chi^2 = \frac{ns^2}{\sigma_0^2}}$$

For $\alpha = 0.05$

* The BCR, w $\equiv \chi^2 > \chi^2_\alpha$

where χ^2_α obtained by refering χ^2 table for (n-1) DF with (prob) $P(\chi^2 > \chi^2_\alpha) = \alpha$.

cal the value of χ^2 & if it lies in the CR reject H_0 otherwise accept.

Q 1) A manufacturer process is expected to produce good's in a specified weight with varient < 5 units. A RS of 10 was found to hve variance 6.2 units is there reason to suspect that the process variet's has rised $\alpha = 0.05$?

A) $\sigma_0^2 = 5$, $s^2 = 6.2$, $n = 10$.

$H_0 : \sigma^2 = 5$ against $H_1 : \sigma^2 > 5$

$\alpha = 0.5$ BCR $\chi^2 > \chi^2_\alpha$

9 DF → $\chi^2_\alpha = 16.92$ (table value)

$\therefore \chi^2 = \frac{ns^2}{\sigma_0^2} = \frac{10 \times 6.2}{5} = 12.4$

$12.4 < \chi^2_\alpha$

→ Chi Square test of independnce =

Data classified into grps according to 2/more classification systms.
The data will be gvn in a 2x2 contingency table.
Let there be 2 attributes A & B.

A ÷ into m classes $A_1, A_2 \cdots A_m$ & B ÷ into n classes $B_1, B_2 \cdots B_n$.

A \ B	B_1	B_2 …	… B_n	Total
A_1	f_{11}	f_{12} …	… f_{1n}	f_1
A_2	f_{21}	f_{22}	… f_{2n}	f_2
⋮	⋮	⋮	⋮	⋮
A_m	f_{m1}	f_{m2}	… f_{mn}	f_m
Total	f_1	f_2	… f_n	N

The expected freq, $f_{ij} = \frac{f_i \times f_j}{N}$

test statistics,

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i}$$

table value of χ^2 is obtained by referring the table $(m-1)(n-1)$ DF. & at a gvn L of significant.

Q) from the follow table showing the no. of plants hving certain characters. Test the hypo that the flwr clr is indpndnt of flatness of leaves.

Flwr \ Leave	Flat leaves	curled leaves	Total
white flwr	99	36	135
Red flwr	20	5	25
Total	119	41	160

Let H_0: the 2 attributes are independent against H_1: the 2 attributes are not indpndnt.

	Flat	curle	Total
white	100	35	135
red	19	6	25
Total	119	41	160

$f_{11} = \frac{135 \times 119}{160} = 100$

$f_{12} = \frac{135 \times 41}{160} = 35$

$f_{21} = \frac{25 \times 119}{160} = 19$

$f_{22} = \frac{25 \times 41}{160} = 6$

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i}$$

$$= \frac{(99-100)^2}{100} + \frac{(36-35)^2}{35} + \frac{(20-19)^2}{19} + \frac{(5-6)^2}{6}$$

$$= \frac{1}{100} + \frac{1}{35} + \frac{1}{19} + \frac{1}{6}$$

$$= 0.249$$

$\alpha = 0.05$ & $(2-1)(2-1) = 1$.

Then $\chi^2_\alpha = 3.841$

$0.249 < 3.841$ Accept

independent

→ Yate's correction =

if any cell freq < 5 we add 0.5 to that cell & adjst for remaining cell freq accordingly. Then the χ^2 test of

goodness of fit is applied.
consider 2x2 contingency table,

a	b
c	d

we have,

$$\chi^2 = \frac{N(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$

According to yates correction we add $1/2$ to a & b & subtract $1/2$ from b & c so that the totals & the marginal freq are not at all disturbed.
Thus with yet's correction χ^2 value become,

I

$$\chi^2 = \frac{N\left(|ad-bc| - 1/2 N\right)^2}{(a+b)(a+c)(b+d)(c+d)}.$$

Q) consider following 2x2 contingency table.

	A_1	A_2	Total
B_1	7	1	8
B_2	6	8	14
Total	13	9	22

test the independent of attribute after applying yates correction.

A) set up the hypo H_0: 2 attributes A & B are independent against H_1: 2 attri- are not independent.
$\therefore$ 1 of the cell freq is 1 (<5) we apply yate's c.

$$\chi^2 = \frac{N\left(|ad-bc| - 1/2 N\right)^2}{(a+b)(a+c)(b+d).(c+d)}$$

$$= \frac{22\left(|7\times8 - 1\times6| - 1/2\times22\right)^2}{(7+1)(7+6)(1+8)(6+8)}$$

$$= \frac{33462}{13104} = 2.554$$

df : 1 , $\alpha = 0.05$ $\chi^2_\alpha = 3.841$.

$\chi^2 > \chi^2_\alpha$ Accept

2 attributes are independent.