Naveen Thomas

Statistics Assignment - 2

① 1) $H_0: \mu = 25$, $H_1: \mu \neq 25$

Correctly stated hypothesis

2) $H_0: \sigma > 10$, $H_1: \sigma = 10$

Not correctly stated hypothesis
it should be $H_0: \sigma \geq 10$ $H_1: \sigma < 10$

3) $H_0: \bar{x} = 50$, $H_1: \bar{x} \neq 50$

Not correctly stated hypothesis
It should be population mean instead of sample mean
ie, $H_0: \mu = 50$ $H_1: \mu \neq 50$

4) $H_0: p = 0.1$, $H_1: p = 0.5$

Not correctly stated hypothesis

5) $H_0: s = 30$, $H_1: s > 30$

Not ~~correct~~ correctly stated hypothesis. Instead of of sample standard deviation it should be population variance
and also ~~ie~~ it should be $H_0: \sigma^2 \leq 900$, $H_1: \sigma^2 > 900$

② $\mu = 52$ $\sigma = 4.5$ $n = 100$ $\bar{x} = 52.8$ $\alpha = 0.05$

$H_0: \mu \leq 52$
$H_1: \mu > 52$

$$z_{test} = \frac{SM - PM}{SE} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{52.8 - 52}{4.5/\sqrt{100}} = 1.77$$

$z(0.05)$ critical right tail $=$ $z(0.95)$ left tail

$= 1.65$

z critical z test

$\therefore$ ~~Reject~~ $z_{test} > z_{critical} \Rightarrow$ Reject H_0
$\Rightarrow$ Avg cost is higher //

①

(3) $\mu = 34 \quad \sigma = 8 \quad \alpha = 0.01 \quad n = 50 \quad \bar{x} = 32.5$

$H_0 : \mu \geq 34$

$H_1 : \mu < 34$

$$z_{test} = \frac{SM - PM}{SE} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{32.5 - 34}{8/\sqrt{50}} = -1.325$$

$$z_{critical}(0.01) = -2.32$$

Since $z_{test} > z_{critical}$ (left tail)

$\rightarrow$ Accept H_0 (Failed to reject H_0)

$\rightarrow$ Factory representatives claim is wrong //

(4) $\mu = 1135 \quad n = 22 \quad \alpha = 0.5$

$\bar{x} = 1031.32 \qquad \Sigma(\bar{x} - \mu)^2 = 1213378.77$

$$s = \sqrt{\frac{\Sigma(\bar{x} - \mu)^2}{n-1}} = 240.37$$

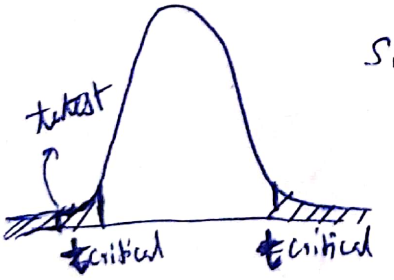
$H_0 : \mu = 1135$

$H_1 : \mu \neq 1135$

$$t(test) = \frac{SM - PM}{SE} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = -2.023$$

$dof = n - 1 = 21$

$$t(0.5)_{critical,\ 21,\ two\ tail} = 0.686$$



Since $t(test) < t_{critical}(-0.686)$

$\rightarrow$ Reject H_0

$\Rightarrow$ The estimation is not accurate for their area

(5) $\mu = 48432$ $n = 400$ $\bar{x} = 48574$ $s = 2000$

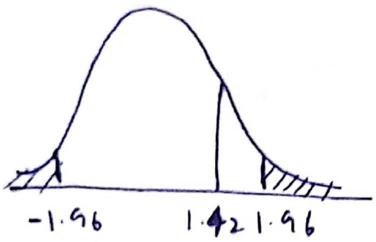
$H_0: \mu = 48432$

$H_1: \mu \neq 48432$

$$z\text{ test} = \frac{SM - PM}{SE} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{48574 - 48432}{2000/\sqrt{400}} = 1.42$$

$\alpha = 0.05$

$$z(0.025) = -1.96$$



Since z test is within the acceptance range

$\Rightarrow$ Accept H_0 (Failed to reject H_0)

$\Rightarrow$ Report is valid

(6) $\mu = 32.28$ $n = 19$ $\bar{x} = 31.67$ $s = 1.29$ $\alpha = 0.05$

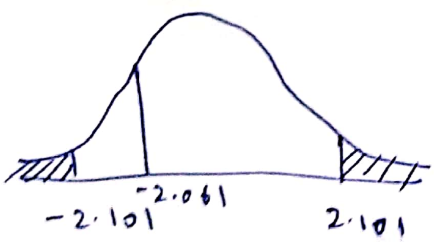
$H_0: \mu = 32.28$

$H_1: \mu \neq 32.28$

$$t\text{ test} = \frac{SM - PM}{s/\sqrt{n}} = \frac{31.67 - 32.28}{1.29/\sqrt{19}} = -2.061$$

$d.f = n - 1 = 18$

$t_{\text{critical } 0.05, \text{ two tail}, 18} = 2.101$



Since t test is within the acceptance range

$\Rightarrow$ Accept H_0 (Failed to reject null hypothesis)

$\Rightarrow$ Avg price per square foot for warehouse in US = 32.28

(7) Since μ & σ is not given we cannot calculate α and β.

(8) $n = 16$ $\mu = 10$ $\bar{x} = 12$ $s = 1.5$

$$t\text{ test} = \frac{SM - PM}{SE} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{12 - 10}{1.5/\sqrt{16}}$$

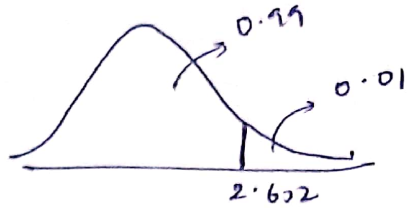
$t_{\text{score}} = 5.33$ //

(9)

$n = 16$

$d.o.f = 15$ $\quad CI = 99\%$

$\alpha = 0.01$



$t(0.01)$ Critical one tail, 15 $= 2.602$

$\therefore$ T score $= 2.602$ //

(10)

$n = 25$ $\quad \bar{x} = 60$ $\quad s = 4$ $\quad CI = 95\%$

$\alpha = 1 - 0.95 = 0.05$

$d.o.f = 24$

$t(0.05)$ 24, one tail $= 1.711$

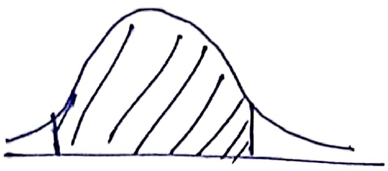
$prob(-t_{0.05} < t < t_{0.10}) =$ ~~0.10 - 0.05~~

$1 - P(t > t_{0.10}) - P(t < -t_{0.05})$

$= 1 - P(t > t_{0.10}) - P(t > t_{0.05})$

$= 1 - 0.10 - 0.05$

$= 0.85$ //



(11)

$n_1 = 1200$ $\quad \bar{x}_1 = 452$ $\quad s_1 = 212$

$n_2 = 800$ $\quad \bar{x}_2 = 523$ $\quad s_2 = 185$

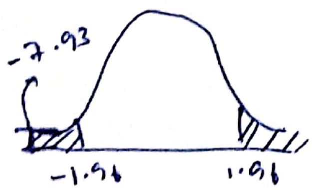
$\alpha = 0.05$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$$z\ test = \frac{S\mu_1 - S\mu_2}{SE} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{452 - 523}{8.95}$$

$= -7.93$



Z Critical 0.025 $= -1.96$

Since z test is not within the acceptance limit

$\Rightarrow$ Reject H_0

$\Rightarrow$ Number of people travelling from Bangalore to Chennai is different from number of people travelling from Bangalore to Hosur in a week.

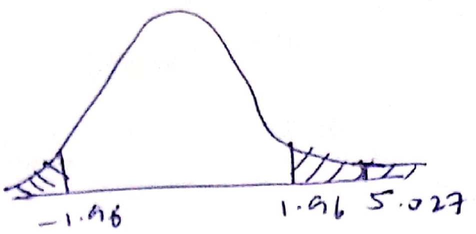
(12)

$n_1 = 100 \quad \bar{x}_1 = 308 \quad s_1 = 84$

$n_2 = 100 \quad \bar{x}_2 = 254 \quad s_2 = 67$

$H_0: \mu_1 = \mu_2 \qquad \alpha = 0.05$

$H_1: \mu_1 \neq \mu_2$



$$z\ test = \frac{SM_1 - SM_2}{SE} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{308 - 254}{10.74}$$

$= 5.027$

$z(0.025) = -1.96$

Since z test is not within the acceptance region

Reject Ho

⟹ Number of people preferring Duracell battery is different from no. of people preferring Energizer battery.

(13)

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2 \qquad \alpha = 0.05$

$\bar{x}_1 = 0.317 \qquad \bar{x}_2 = 0.21$

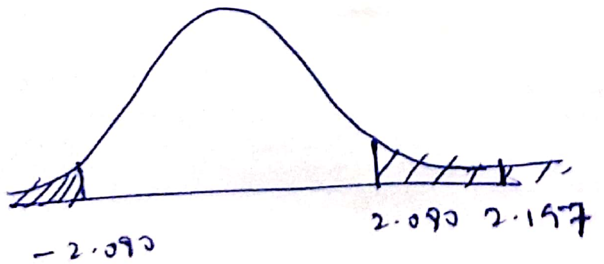
$s_1 = 0.12 \qquad s_2 = 0.11$

$n_1 = 14 \qquad n_2 = 9$

$$t_{test} = \frac{SM_1 - SM_2}{SE} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{0.317 - 0.21}{0.0487}$$

$d.o.f = 14 + 9 - 2 = 21 \qquad = 2.197$

$t_{critical}$ 0.05, two tail, 21 $= 2.080$



Since t test is not within the acceptance region

Reject Ho

⟹ Avg percentage increase in the price of sugar differs when it is sold at two different prices.

(14)

$n_1 = 15$ $\qquad$ $n_2 = 12$ $\qquad$ $H_0: \mu_1 \geq \mu_2$

$\bar{x}_1 = 6598$ $\qquad$ $\bar{x}_2 = 6870$ $\qquad$ $H_1: \mu_1 < \mu_2$

$s_1 = 844$ $\qquad$ $s_2 = 669$

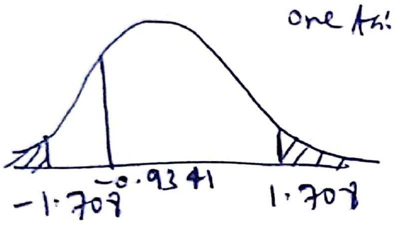
$$t_{test} = \frac{S_{M_1} - S_{M_2}}{SE} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{6598 - 6870}{291.18}$$

$\alpha = 0.05$

$$= -0.9341$$

$dof = 15 + 12 - 2 = 25$

$t_{critical}$ one tail, 0.05, 25 $= 1.708$



Since ~~t~~ $t_{critical}$ is within the acceptance region

$\Rightarrow$ Accept H_0 (Failed to reject H_0)

$\Rightarrow$ Small price redn is enough to increase sales of compact disk players.

(15)

$n_1 = 100$ (not 1000 error in question) $\qquad \alpha = 0.05$

$x_1 = 53$ $\qquad \hat{p}_1 = \frac{53}{100} = 0.53$

$H_0: p_1 = p_2$

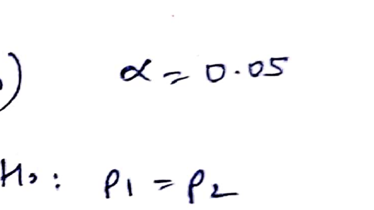
$n_2 = 100$

$x_2 = 43$ $\qquad \hat{p}_2 = \frac{43}{100} = 0.43$ $\qquad H_1: p_1 \neq p_2$

$$z_{test} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \cancel{= \frac{0.53 - 0.43}{\sqrt{(0.48)(0.52)}}}$$

$$= \frac{0.53 - 0.43}{\sqrt{\frac{(0.53)(1-0.53)}{100} + \frac{0.43(1-0.43)}{100}}} = 1.415$$

$z_{critical}$ 0.025 $= -1.96$



Since zcritical is within the acceptance region

$\Rightarrow$ Accept H_0 (Failed to reject H_0)

$\Rightarrow$ Population proportion of ~~quality~~ bank's share of the car loan market in 1980 and 1995 is equal

(16)

$H_0 : p_1 - p_2 \leq 0.10$

$H_1 : p_1 - p_2 > 0.10$

$n_1 = 300 \quad x_1 = 120 \quad \hat{p}_1 = \frac{x_1}{n_1} = \frac{120}{300} = 0.4$

$n_2 = 700 \quad x_2 = 140$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{140}{700} = 0.20$

$\alpha = 0.05$

$$z_{test} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

$$= \frac{(0.4 - 0.2) - 0.10}{\sqrt{\frac{0.4 \times 0.6}{300} + \frac{0.2 \times 0.8}{700}}} = 3.118$$

$Z_{critical\ 0.025} = -1.96$

Since zcritical is not within the acceptable region
→ Reject Null hypothesis
→ The population proportion of traveller's check buyers who buy atleast \$2500 in checks when sweepstakes prizes are offered are atleast 10% higher than the proportion of such buyers when no sweepstakes are on.

(17)

Number	Observed (O)	Expected (E)	$(O-E)^2$
1	16	22	36
2	20	22	4
3	25	22	9
4	14	22	64
5	29	22	49
6	28	22	36

H_0: Die is unbiased

H_1: Die is biased

$\alpha = 0.05$

$E = \frac{132}{6} = 22$

$d.o.f = n-1 = 6-1 = 5$

$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{198}{22} = 9$

$\chi^2_{critical\ 0.05,5} = 11.070$

Since $\chi^2_{test} < \chi^2_{critical}$

→ Accept H_0 (Failed to reject H_0)

→ Die is unbiased

(18)

H_0: Gender and voting are independent
H_1: Gender and voting are dependent

	Men	Women	Total
Voted	2792 O / E 2731	3591 O / E 3652	6383
Not Voted	1486 O / E 1547	2131 O / E 2070	3617
Total	4278	5722	10000

E of men voted $= \frac{6383 \times 4278}{10000} = 2731$

E of men not voted $= \frac{3617 \times 4278}{10000} = 1547$

$E_{\text{women voted}} = 3652$

$E_{\text{women not voted}} = 2070$

$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{(2792-2731)^2}{2731} + \cdots + \frac{(2131-2070)^2}{2070} = 6.584$ //

$d.o.f = (2-1)(2-1) = 1$ //

$\alpha = 0.05$

$\chi^2_{critical,\,0.05,\,1} = 3.841$

Since $\chi^2_{test} > \chi^2_{critical}$

$\Rightarrow$ Reject H_0

$\Rightarrow$ Gender and voting are dependent

(19)

Candidate	Observed (O)	Expected (E)	$(O-E)^2$
Higgins	41	25	256
Reardon	19	25	36
White	24	25	1
Charlton	16	25	81

H_0: All candidates are equally popular

H_1: All candidates are not equally popular

$E = \frac{100}{4} = 25$ //

$\alpha = 0.05$

$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{374}{25} = 14.96$

$d.o.f = 4-1 = 3$

$\chi^2_{critical\,3,\,0.05} = 7.815$

Since $\chi^2_{test} > \chi^2_{critical}$
reject H_0
$\Rightarrow$ All candidates are not equally popular

(20)

Age of child	Photograph A	Photograph B	Photograph C	Total
5-6 years	O: 18, E: 12	O: 22, E: 18	O: 20, E: 30	60
7-8 years	O: 2, E: 14	O: 28, E: 21	O: 40, E: 35	70
9-10 years	O: 20, E: 14	O: 10, E: 21	O: 40, E: 35	70
Total	40	60	100	200

H_0: No relationship between age and photograph preference

H_1: There is a significant relationship between age and photograph preference

$$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{(18-12)^2}{12} + \cdots + \frac{(40-35)^2}{35}$$

$= 29.60$

d.o.f $= (3-1) \times (3-1) = 4$

$\chi^2_{critical_{0.05,4}} = 9.488$

Since $\chi^2_{test} > \chi^2_{critical}$

$\Rightarrow$ Reject H_0

$\Rightarrow$ There is a significant relationship between age and photograph preference.

(21)

	Support	No Support	Total
Conform	O: 18, E: 29	O: 40, E: 29	58
Not conform	O: 32, E: 21	O: 10, E: 21	42
Total	50	50	100

$\alpha = 0.05$

H_0: There is no significant diff between the support and no support condn in the frequency with which individuals are likely to conform

H_1: There is significant diff b/w support and no support condn with conform & not conform

$$\chi^2_{test} = \sum \left[\frac{(O-E)^2}{E} \right] = \frac{(18-29)^2}{29} + \cdots \frac{(10-21)^2}{21}$$

$= 19.87$ //

d.o.f $= (2-1) \times (2-1) = 1$

$\chi^2_{critical_{0.05,1}} = 3.841$

Since $\chi^2_{test} > \chi^2_{critical}$

$\Rightarrow$ Reject H_0

$\Rightarrow$ There is significant difference between the "support" and "no support" conditions in the frequency with which individuals are likely to conform.

(22)

	Height: Short	Height: Tall	Total
Leader	O 12, E 20	O 32, E 24	44
Follower	O 22, E 16	O 14, E 20	36
Unclassifiable	O 9, E 7	O 6, E 8	15
Total	43	52	95

H_0: There is no relationship between height and leadership qualities

H_1: There is some relationship between height and leadership qualities

$\alpha = 0.01$

$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{(12-20)^2}{20} + \cdots + \frac{(6-8)^2}{8} = 10.99$

$d.o.f = (3-1) \times (2-1) = 2$ //

$\chi^2_{critical\ 0.01, 2} = 9.210$

Since $\chi^2_{test} > \chi^2_{critical}$

$\Rightarrow$ ~~Accept Ho~~ Reject H_0

$\Rightarrow$ There is some relationship between height and leadership qualities

(23)

	Married	Widowed, divorced, separated	Never married	Total
Employed	O 679, E 654	O 103, E 109	O 114, E 133	896
Unemployed	O 63, E 68	O 10, E 11	O 20, E 14	93
Not in laborforce	O 42, E 62	O 18, E 11	O 25, E 12	85
Total	784	131	159	1074

H_0: Men of different marital status do not seem to have different distributions of labor force status.

H_1: Men of different marital status seem to have different distribution of labor force status

$\alpha = 0.05$

$\chi^2_{test} = \sum \frac{(O-E)^2}{E} = \frac{(679-654)^2}{654} + \frac{(103-109)^2}{109} + \frac{(114-133)^2}{133} + \cdots \frac{(25-12)^2}{12} = 32.02$

$d.o.f = (3-1) \times (3-1) = 4$

$\chi^2_{critical\ 0.05, 4} = 9.488$

Since $\chi^2_{test} > \chi^2_{critical}$

$\Rightarrow$ Reject H_0

$\Rightarrow$ Men of different marital status seem to have different distributions of labor force status. //