PROBABILITY & STATISTICAL DATA ANALYSIS

SECI2143-02

ASSIGNMENT 3

LECTURER'S NAME: DR. CHAN WENG HOWE

Prepared by: -

ADRINA ASYIQIN BINTI MD ADHA	A20EC0174
MADINA SURAYA BINTI ZHARIN	A20EC0203
NAYLI NABIHAH BINTI JASNI	A20EC0105
NUR SYAMALIA FAIQAH BINTI MOHD KAMAL	A20EC0118

Question 1

a) $n = 25$ $\quad \bar{x} = 19.5$ $\quad s = 9.88$

90% CI μ = ?

90% CI μ = $\bar{x} \pm t\left(\frac{s}{\sqrt{n}}\right)$ $\quad t = ?$ $\quad V = 25 - 1 = 24$

$\frac{\alpha}{2} = \frac{1 - 0.9}{2} = 0.05$

$= 19.5 \pm 2.064\left(\frac{9.88}{\sqrt{25}}\right)$ $\quad t_{24, 0.05} = 2.064$

$= 19.5 \pm 4.0785$

$= (15.4215, 23.5785)$

b) $n = 360$ $\quad x = 160$

$\hat{p} = \frac{x}{n} = \frac{160}{360} = 0.4444$

$\hat{q} = 1 - 0.4444 = 0.5556$

$EBP = z\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$= 2.58\sqrt{\frac{(0.4444)(0.5556)}{360}}$

$= (2.58)(0.0262)$

$= 0.0676$

99% CI $= (\hat{p} \pm 0.0676)$

$= (0.3768, 0.5120)$

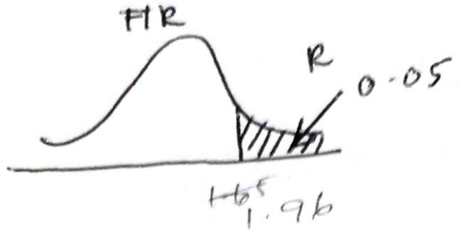
Question 2

a) $n = 350 \quad x = 33$

$\alpha = 0.05$

$H_0 : p < 0.1$

$H_1 : p \geq 0.1$



$Z_{0.05} = 1.645$

$\hat{p} = \frac{x}{n} = \frac{33}{350} = 0.0943$

$p_0 = 0.1 \quad q_0 = 0.9$

$$Z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

$$= \frac{0.0943 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{350}}}$$

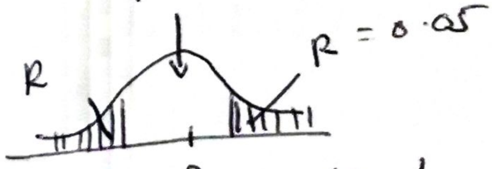
$$= -0.35536$$

→ The claim is correct

→ $-0.35536 < 1.645$ and falls in the fail to reject area.

b) $x = 16 \quad \bar{x} = 58400$

$\sigma = 652 \quad SL = 99\%$



degree of freedom $= 16 - 1$
$= 15$

$H_0 : \mu = 58000$

$H_1 : \mu \neq 58000$

$\alpha = 0.05$

$$Z_c = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{58400 - 58000}{652 / \sqrt{16}}$$

$$= \frac{400}{163} = 2.45$$

Zscore : 1.753

→ Reject hypothesis

→ $2.45 > 1.753$

→ average strength is not 58000 psi

Question 3

a) $n = 10$ $\alpha = 0.01$

$H_0 : \mu = 10$

$H_1 : \mu \neq 10$

$\bar{x} = (10.3 + 9.9 + 10.2 + 10.1 + 9.7 + 9.9 + 9.8 + 10.3 + 10.0 + 10.4) \div 10$

$= 100.6 / 10$

$= 10.06$

x	$x - \bar{x}$		
10.3	0.24		0.0576
9.9	-0.16		0.0256
10.2	0.14		0.0196
10.1	0.04		0.0016
9.7	-0.36		0.1296
9.9	-0.16		0.0256
9.8	-0.26		0.0676
10.3	0.24		0.0576
10.0	-0.06		0.0036
10.4	0.34		0.1156
			Σ 0.504

$S = \sqrt{\dfrac{\Sigma (x - \bar{x})^2}{n - 1}}$

$= \sqrt{\dfrac{0.504}{10 - 1}}$

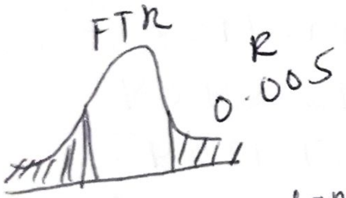
$= 0.2366$

$t = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

$= \dfrac{10.06 - 10}{0.2366 \sqrt{10}}$

$= \dfrac{0.06}{0.7482}$

$= 0.08$



degree of freedom $= 10 - 1$

$= 9$

$\alpha = 0.005$

zsore $= 3.250$

- the hypothesis is correct.
- ~~3.25~~ $0.08 < 3.250$, $0.8 > -3.250$
- is located at the fail to reject area.

Q3 (b)

$n = 14$

$\alpha = \cancel{0.05}\ 0.05$ | $\alpha/2 = \cancel{0.025}\ 0.025$

$1 - \alpha/2 = 0.975$

$H_0 : \sigma = 0.44$

$H_1 : \sigma \neq 0.44$

$$\bar{x} = \frac{3.70 + 4.31 + 3.73 + 4.33 + 3.33 + 2.58 + 4.47 + 3.55 + 4.66 + 3.68 + 3.02 + 4.09 + 2.36 + 3.35}{14}$$

$\bar{x} = 3.654$

$$s^2 = \frac{(3.70-3.654)^2 + (4.31-3.654)^2 + (3.73-3.654)^2 + (4.33-3.654)^2 + (3.33-3.654)^2 + (2.58-3.654)^2 + (4.47-3.654)^2 + (3.55-3.654)^2 + (4.66-3.654)^2 + (3.68-3.654)^2 + (3.02-3.654)^2 + (4.09-3.654)^2 + (2.36-3.654) + (3.35-3.654)^2}{14-1}$$

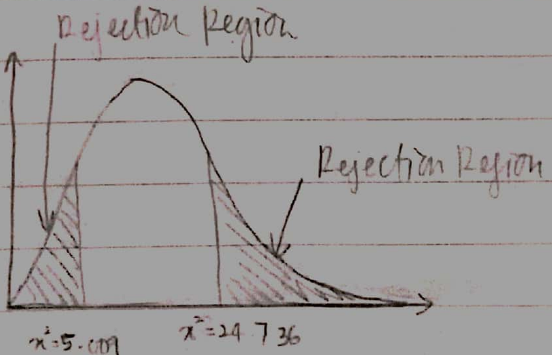
$s^2 = 0.47707$

$$x^2 = \frac{(14-1)(0.47707)}{(0.44)^2}$$

$x^2 = 32.03$

$x^2_{0.025,13} = 24.736$

$x^2_{0.975,13} = 5.009$



$\therefore$ Since $32.03 > 24.736$, so reject the hypothesis. There is not sufficient evidence that support the standard deviation of birth weight is affected by the vitamin supplement.

Question 4

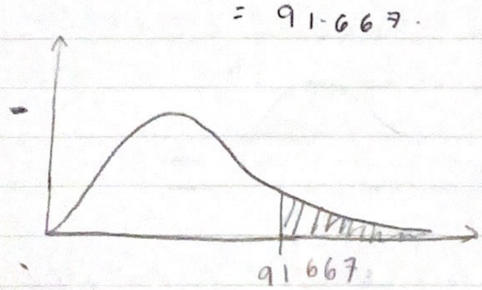
$H_0 : \sigma^2 = 0.18$

$H_1 : \sigma^2 > 0.18$

test statistic, $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(101-1)\,0.165}{0.18}$

$= 91.667$

$\chi^2_{0.05,\,100} = 124.342$



since ~~91.667 < 124.342~~ $124.342 > 91.667$,

reject null hypothesis.

There is no sufficient evidence ~~that~~ to support the ~~standard deviation~~ variance of experienced quality control inspector of measurement of sheet metal stamping is $0.18\ (\text{inch})^2$.

~~would have a vio vari~~

Question 5

$n_1 = 15$, $\bar{x}_1 = 76.40$, $\sigma_1^2 = 25.30$

$n_2 = 18$, $\bar{x}_2 = 71.20$, $\sigma_1^2 = 22.20$

$H_0 : \mu_1 = \mu_2$ $\alpha = 0.95$

$H_1 \ \mu_1 \neq \mu_2$

$$* t_0 = \frac{\bar{x}_1 - \cancel{x} \bar{x}_2 - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{76.40 - 71.20 - 0}{\sqrt{\frac{25.30}{15} + \frac{22.20}{18}}}$$

$$= 3.0431$$

$$v = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{\left[s_1^2/n_1\right]^2}{n_1 - 1} + \frac{\left[s_2^2/n_2\right]^2}{n_2 - 1}} = \frac{\left[\frac{25.30}{15} + \frac{22.20}{18}\right]^2}{\frac{\left[\frac{25.30}{15}\right]^2}{15 - 1} + \frac{\left[\frac{22.20}{18}\right]^2}{18 - 1}}$$

$$= \frac{8.5264}{0.29268} = 29.1321 \approx 29.$$

Therefore ;

$* t_0 > t_{0.95, 29} = 0.0633$

$* t_0 < -t_{0.95, 29} = -0.0633$

Since $t_0 = 3.0431 > t_{0.95, 29} = 0.0633$, reject null hypothesis. There is evidence to conclude that there is difference in the mean spending between the two populations.

Question 6

$H_0 : \sigma_1 = \sigma_2$
$H_1 : \sigma_1 > \sigma_2$

$$F = \frac{s_1^2}{s_2^2} = \frac{\cancel{36}\ 3.6^2}{2.5^2} = 2.0736$$

Degree of freedom : numerator, $n_1 - 1 = 16 - 1 = 15$
denominator, $n_2 - 1 = 21 - 1 = 20$.

$F_{0.95,15,20} = 0.4296$.

Since $F = 2.0736 > F_{0.95,15,20} = 0.4296$, reject null hypothesis. There is sufficient evidence to conclude that processor from batch 1 is greater than batch 2.

Question 7

Before (x_1)	203	390	389	279	333	213	410	364	470	464
After (x_2)	225	410	402	285	355	240	444	370	501	490
$D=(x_1-x_2)$	-22	-20	-13	-6	-22	-27	-34	-6	-31	-26
$D^2=(x_1-x_2)^2$	484	400	169	36	484	729	1156	36	961	676

$\Sigma D = -207$

$\Sigma D^2 = 5131$

$\bar{D} = \frac{-207}{10} = -20.7$ $\quad \alpha = 0.05$

$H_0 = \mu_D = 0$

$H_0 = \mu_D < 0$

d.f. = 9

C.V. = 2.262

$$S_D = \sqrt{\frac{5131 - \frac{(-207)^2}{10}}{9}} = 9.6959$$

$$t = \frac{-20.7 - 0}{\left(\frac{9.6959}{\sqrt{10}}\right)} = -6.7512$$

Conclusion: since $-6.7512 < -2.262$, reject H_0. There is enough evidence to conclude that training can increase ~~then~~ the number of words spelled correctly.

Question 8

									T
X	0.27	1.41	2.19	2.83	2.19	1.81	0.85	3.05	14.6
Y	2	3	3	6	4	2	1	5	26
XY	0.54	4.23	6.57	16.98	8.76	3.62	0.85	15.25	56.8
$(X-\bar{X})^2$	2.418	0.172	0.133	1.010	0.133	0.000	0.951	1.051	6.318
$(Y-\bar{Y})^2$	1.563	0.063	0.063	7.563	0.563	1.563	5.063	3.063	19.504

$\bar{X} = 14.6/8 = 1.825$ $\bar{Y} = 26/8 = 3.25$

i) $r = \dfrac{56.8 - [(14.6)(26)/8]}{\sqrt{(6.318)(\cancel{14.0}\ 19.504)}} = 0.8423$

ii) $H_0 : \rho = 0$ $\quad t = \dfrac{r(\sqrt{n-2})}{\sqrt{1-r^2}} = \dfrac{0.8423(\sqrt{8-2})}{\sqrt{1-(0.8423)^2}} = 3.83$

$H_1 : \rho \neq 0$

$df = 8-2 = 6$ $\quad \alpha = 0.05$ $\quad \cancel{C.v.\ t_6 =}$ $\quad C.V. : t_6 = 2.447$

Conclusion: reject H_0 since $3.83 > 2.447$. There is sufficient evidence of a linear relationship between the size of household and the daily plastic usage consumed by them at the 5% level of significance.

iii) For $\alpha = 0.01$, we still reject H_0 because the $C.V. = 3.707 < 3.83$ and conclude $\rho \neq 0$ which there is sufficient evidence of a linear relationship between the size of household and the daily plastic usage consumed by them at the 1% level of significance.

Question 9 (i)

Car Manufacturer	Engine Vol, x (cubic km)	Mileage, y (km)	xy	x^2
Honda	97	24	2328	9409
Datsun	85	29	2465	7225
Chevrolet	98	26	2548	9604
Mazda	105	24	2520	11025
Proton	120	24	2880	14400
Perodua	151	22	3322	22801
Toyota	140	23	3220	19600
Nissan	134	23	3082	17956
Kia	146	21	3066	21316
	$\Sigma x = 1076$	$\Sigma y = 216$	$\Sigma xy = 25435$	$\Sigma x^2 = 133336$

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

$$= \frac{25435 - \frac{(1076)(216)}{9}}{133336 - \frac{(1076)^2}{9}}$$

$$= -0.0828678$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$= 24 - (-0.0828678)(119.55556)$$

$$b_0 = 33.9073059$$

$$\hat{y}_i = b_0 + b_1 x$$

$$\hat{y} = 33.9073 - 0.0829x$$

Scanned with CamScanner

No.: Date:

Question 9 (ii)

$\hat{y} = 33.9073 - 0.0829x$

$x = 125$

$\hat{y} = 33.9073 - 0.0829(125)$

$= 23.5448$

So, from the least-square equation $\hat{y} = 33.9073 - 0.0829x$, it is found that the estimated mileage for a car that has 125 cubic km engine volume is 23.5448 km.

No. Date:

Question 10

Step 1 : Define Hypothesis

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_1 : at least one of the mean is different

Step 2 : For each category, find n, $\bar{x}$ and s

Category 1 : (Memory Booster)

$n = 5$

$$\bar{x} = \frac{70 + 77 + 83 + 90 + 97}{5}$$

$= 83.4$

$$s = \sqrt{\frac{(70-83.4)^2 + (77-83.4)^2 + (83-83.4)^2 + (90-83.4)^2 + (97-83.4)^2}{(5-1)}}$$

$= 10.597$

Category 2 : (Placebo)

$n = 5$

$$\bar{x} = \frac{37 + 43 + 50 + 57 + 63}{5}$$

$= 50$

$$s = \sqrt{\frac{(37-50)^2 + (43-50)^2 + (50-50)^2 + (57-50)^2 + (63-50)^2}{(5-1)}}$$

$= 10.440$

No. Date

Category 3 : (Without Treatment)

$n = 5$

$$\bar{x} = \frac{3 + 10 + 17 + 23 + 30}{5}$$

$$= 16.6$$

$$s = \sqrt{\frac{(3-16.6)^2 + (10-16.6)^2 + (17-16.6)^2 + (23-16.6)^2 + (30-16.6)^2}{(5-1)}}$$

$$= 10.597$$

Step 3: Find variance between samples

Step 3a: Find mean between samples

$$\bar{\bar{x}} = \frac{83.4 + 50 + 16.6}{3}$$

$$= 50$$

Step 3b : Find standard deviation between samples

$$s_{\bar{x}} = \sqrt{\frac{(83.4-50)^2 + (50-50)^2 + (16.6-50)^2}{(3-1)}}$$

$$= 33.4$$

Step 3c : Find variance between samples

$$n s_{\bar{x}}^2 = 5(33.4)^2$$

$$= 5577.8$$

Step 4 : Find variance within samples

$$s_p^2 = \frac{(10.597)^2 + (10.440)^2 + (10.597)^2}{3}$$

$= 111.195$

Step 5 : Calculate test statistic, F

$$F = \frac{n s_{\bar{x}}^2}{s_p^2}$$

$$= \frac{5577.8}{111.195}$$

$= 50.162$

Step 6 : Calculate numerator & denominator degree of freedom

Numerator $= k-1 = 3-1 = 2$

Denominator $= k(n-1) = 3(5-1) = 12$

Step 7 : Find $F_{critical\ value}$ in $\alpha = 0.01$

$F_{critical\ value} = 6.927$

Step 8 : Test the claim & conclude

Since $F_{test\ statistic} > F_{critical\ value}$ $(50.162 > 6.927)$, we reject the null hypothesis.

There is not sufficient evidence to claim that the treatments recieved has the same effects.

Therefore, the data provided has sufficient evidence to support the claim that the treatments will have different effects.