SECI2143: PROBABILITY & STATISTICAL DATA ANALYSIS

ASSIGNMENT 3

SECTION 02

DR CHAN WENG HOWE

GROUP MEMBERS:

NAME	MATRIC NO
AMIRAH RAIHANAH BINTI ABDUL RAHIM	A20EC0182
CHLOE RACQUELMAE KENNEDY	A20EC0026
LEE CAI XUAN	A20EC0062
MYZA NAZIFA BINTI NAZRY	A20EC0219

QUESTION 1

No. Date

Question 1

a) $n = 25$

$\bar{x} = 19.5$ seconds

$\sigma = 9.88$ seconds

$\bar{x} \pm (z \text{ critical value})\left(\frac{\sigma}{\sqrt{n}}\right)$

$19.5 \pm (1.645)\left(\frac{9.88}{\sqrt{25}}\right) = 19.5 \pm 3.2505$

The lower and upper confidence limits are 16.2495 seconds and 22.7505 seconds

b) $n = 360$

$\hat{p} = \frac{160}{360}$

$= 0.44$

$SE = \sqrt{\frac{(0.44)(0.56)}{360}}$

$= 0.026$

$\hat{p} \pm 2.58\,(0.026) = 0.44 \pm 0.067$

$= (0.0373,\ 0.507)$

QUESTION 2

Date:

Question 2.

a) The company ABC introduce a new drug test method, 350 subjects were tested and results from 33 subjects were incorrect (either false positive or false negative). The company claims that the incorrect test results will be less than 10%. Test this claim using $\alpha = 0.05$.

i) $H_0 : p = 0.1$

$H_1 : p < 0.1$

ii) $\hat{p} = \frac{33}{350} = 0.094$

iii) $p = 0.1$ $\quad n = 350$, $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

$\hat{p} = 0.094$

$= \frac{0.094 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{350}}}$

$= \frac{-0.006}{0.016} = -0.375$

iv) $\alpha_{0.05} = -1.645$

conclusion : Since $-0.375 > -1.645$, failed to reject H_0.
There is insufficient evidence to show that the company claims of the incorrect test results will be less than 10%.

b) A random sample of 16 steel beams has a mean compressive strength of 58400 psi (pound per square inch) and population standard deviation of 652 psi. Using 99% significance level, test if the true average compressive strength of steel is 58000 psi.

i) $n = 16$ $\qquad H_0 : \mu = 58000$

$\bar{x} = 58400$ $\qquad H_1 : \mu \neq 58000$

$\mu = 58000$

$\sigma = 652$

$\alpha = 0.01$

ii) $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{58400 - 58000}{652/\sqrt{16}}$

$= \dfrac{400}{163}$

$= 2.454$

iii) $C.V = z_{0.01} = 2.576 \mid -2.576$

iv) Conclusion : Since $-2.576 < 2.454 < 2.576$, failed to reject H_0. There is sufficient evidence to show that the true average compressive strength of steel is 58000 psi.

QUESTION 3

Question 3

(a) $H_0 : \mu = 10$ $\alpha = 0.01$

$H_1 : \mu \neq 10$ $n = 10$

$$\bar{x} = \frac{10.3 + 9.9 + 10.2 + 10.1 + 9.7 + 9.9 + 9.8 + 10.3 + 10 + 10.4}{10}$$

$= 10.06$

$$s^2 = \frac{(10.3-10.06)^2 + (9.9-10.06)^2 + (10.2-10.06)^2 + (10.1-10.06)^2 + (9.7-10.06)^2 + (9.9-10.06)^2 + (9.8-10.06)^2 + (10.3-10.06)^2 + (10-10.06)^2 + (10.4-10.06)^2}{10-1}$$

$= 0.056$

$s = 0.237$

$t_{0.005,9} = 3.250$

$$t_{\text{test statistic}} = \frac{10.06 - 10}{\left(\frac{0.237}{\sqrt{10}}\right)}$$

$= 0.801$

since $t_{\text{test statistic}} < t_{0.005,9}$, we fail to reject H_0; there is sufficient evidence to show that the average content value of Brand X car lubricants is 10 l.

No.

Date

b) $H_0 : \sigma = 0.44$ $\alpha = 0.05$

$H_1 : \sigma \neq 0.44$ $n = 14$

$$\bar{x} = \frac{3.7 + 4.31 + 3.73 + 4.33 + 3.33 + 2.58 + 4.47 + 3.55 + 4.66 + 3.68 + 3.02 + 4.09 + 2.36 + 3.35}{14}$$

$= 3.65$

$$s^2 = \frac{(3.7-3.65)^2 + (4.31-3.65)^2 + (3.73-3.65)^2 + (4.33-3.65)^2 + (3.33-3.65)^2 + (2.58-3.65)^2 + (2.58-3.65)^2 + (3.55-3.65)^2 + (4.66-3.65)^2 + (3.68-3.65)^2 + (3.02-3.65)^2 + (4.09-3.65)^2 + (2.36-3.65)^2 + (3.35-3.65)^2}{14-1}$$

$= 0.477$

$s = 0.691$

$\chi^2_{0.025, 13} = 24.736$

$$\chi^2 = \frac{(14-1)(0.691)^2}{0.44^2}$$

$= 32.062$

since χ^2 test statistic $> \chi^2$ critical value, we reject H_0; there is sufficient evidence to show that the vitamin supplement has ~~e~~ effects on the standard deviation of the birth weight.

QUESTION 4

QUESTION 4

In general, an experienced quality control inspector of measurement of sheet metal stamping would have a variance of 0.18 (inch)^2. A new joint inspector measures 101 stampings with variance of 0.165 (inch)^2. Assume data is normally distributed, by using 95% significance level, test whether the new joint inspector is making satisfactory measurements.

Hypothesis statement :

$H_0 : \sigma^2 = 0.18$

$H_1 : \sigma^2 < 0.18$

$n = 101$, $\alpha = 0.95$

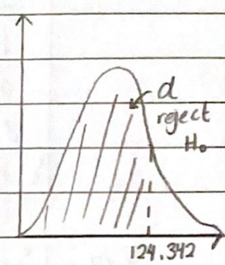
Test significance :

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

$$= 91.6667$$

Degree of freedom, d.f. $= 101 - 1$, Right-tail $0.05\ (1-\alpha)$

$= 100$



$\therefore \chi^2_{0.95,100} = 124.342$

$\therefore$ Conclusion : Since $124.342 > 91.667$, rejects H_0.

There is sufficient evidence that support the variance for new joint inspector is less than 0.18(inch)^2 which proves that the new joint inspector is not making satisfactory measurements at the 95% level of significance.

QUESTION 5

No: Date:

Question 5.

A marketing study was done on consumer from 2 age groups for their spending on 20 categories of consumer items. Based on 15 respondents in the 18-34 age group, the average spending is RM 76.40 with variance of RM 25.30. Meanwhile, based on 18 respondents in the 35+ group, the mean and variance were RM 71.20 and RM 22.20 respectively. By assuming the population variance are not equal, test if there is any difference in the mean spending between these two populations in 95% significance level.

i) $H_0 : \mu_1 = \mu_2$ $\quad \bar{x}_1 = 76.40 \quad \bar{x}_2 = 71.20 \quad \alpha = 0.05$

$H_1 : \mu_1 \neq \mu_2$ $\quad s_1^2 = 25.30 \quad s_2^2 = 22.20$

$n_1 = 15 \quad n_2 = 18$

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

$$= \frac{5.2}{\sqrt{2.92}} = 3.043$$

iii) $$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}} = \frac{\left(\frac{25.30}{15} + \frac{22.20}{18}\right)^2}{\frac{\left(\frac{25.30}{15}\right)^2}{14} + \frac{\left(\frac{22.20}{18}\right)^2}{17}}$$

$$= \frac{2.92^2}{0.293} = 29.10 \approx 29$$

iv) $\alpha = 0.05$, we reject H_0

if $t_0 > t_{0.025, 29} = 2.045$

and $t_0 < -t_{0.025, 29} = -2.045$

v) Conclusion: Since $3.043 > 2.045$, we reject H_0. There is sufficient evidence to conclude that there is any difference in the mean spending between these two populations.

QUESTION 6

Date

Question 6

$H_0 : \sigma_1 = \sigma_2$

$H_1 : \sigma_1 > \sigma_2$

$n_1 = 16$, $s_1 = 3.6$

$n_2 = 21$, $s_2 = 2.5$

$\alpha = 0.05$

numerator $= 16 - 1 = 15$

denominator $= 21 - 1 = 20$

$F_{0.05, 15, 20} = 2.203$

$$F_{\text{test statistic}} = \frac{3.6^2}{2.5^2} = 2.074$$

since $F_{\text{test statistic}} < F_{\text{critical value}}$, we fail to reject H_0; there is sufficient evidence to claim that the standard deviation of processors from batch 1 is greater than the standard deviation of batch 2.

QUESTION 7

Question 7

$H_0 : \mu_P = 0$ $\alpha = 0.05$ $df = 10-1 = 9$ $t_{0.05,9} = -1.833$

$H_1 : \mu_P < 0$

Before training	After training	d	d^2
203	225	-22	484
390	410	-20	400
389	402	-13	169
279	285	-6	36
333	355	-22	484
213	240	-27	729
410	444	-34	1156
364	370	-6	36
470	501	-31	961
464	490	-26	676
		$\Sigma d = 207$	$\Sigma d^2 = 5131$

$\bar{d} = \frac{-207}{10} = -20.7$

$S_d = \sqrt{\frac{5131 - \frac{(-207)^2}{10}}{10-1}} = 9.696$

$t = \frac{-20.7 - 0}{\frac{9.696}{\sqrt{10}}} = -6.751$

Since $t = -6.751 < -1.833$, H_0 is rejected. There is sufficient evidence to conclude that training can increase the number of words spelled correctly.

QUESTION 8

QUESTION 8

A local authority is going to study the relationship between the size of household and the daily plastic usage consumed by them. The results of a sampling given below in Table 5 indicate some values of X (weight of plastic usage) and Y (the size of household).

TABLE 5 : Weight of plastic usage (X) and size of household (Y)

X	0.27	1.41	2.19	2.83	2.19	1.81	0.85	3.05
Y	2	3	3	6	4	2	1	5

i- From the above sampling, find the value of correlation coefficient r.

X	Y	XY	Y^2	X^2
0.27	2	0.54	4	0.0729
1.41	3	4.23	9	1.9881
2.19	3	6.57	9	4.7961
2.83	6	16.98	36	8.0089
2.19	4	8.76	16	4.7961
1.81	2	3.62	4	3.2761
0.85	1	0.85	1	0.7225
3.05	5	15.25	25	9.3025
$\Sigma = 14.6$	$\Sigma = 26$	$\Sigma = 56.8$	$\Sigma = 104$	$\Sigma = 32.9632$

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

$$= \frac{56.8 - [14.6(26)]/8}{[32.9632 - 14.6^2/8][104 - 26^2/8]}$$

$= 0.0759_{\#}$

ii - Using the same sample data set above, conduct the hypothesis testing to know whether the variable X and Y are really correlates using 95% confidence level.

Hypothesis :

$H_0 : \rho = 0$ (no linear correlation)

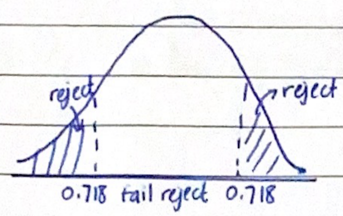
$H_1 : \rho \neq 0$ (linear correlation exists)

$\alpha = 0.95$

critical value ; $n-2 = 8-2$, $\alpha/2$

$= 6$ $= 0.5$

$\therefore$ c.v = 0.718



Test statistics :

$$t = \frac{0.0759}{\sqrt{\frac{1-0.0759^2}{6}}}$$

$= 0.1865$

$\therefore$ Conclusion : Fail to reject H_0. There is sufficient proof that there are no linear correlation between variable X and Y at the 95% confidence level.

iii - Find out if the decision in (ii) may change if you increase the confidence level to 99%.

Hypothesis :

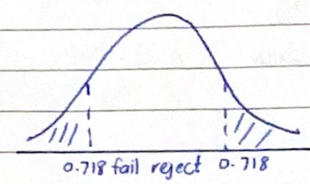
$H_0 : \rho = 0$ (no linear correlation)

$H_1 : \rho \neq 0$ (linear correlation exists)

$\alpha = 0.99$

critical value : $n-2 = 8-2 = 6$, $\alpha/2 = 0.5$

$\therefore$ c.v $= 0.718$



Test statistics :

$$t = \frac{0.0759}{\sqrt{\frac{1-0.0759^2}{6}}}$$

$= 0.1865$

Decision : Fail to reject H_0.

$\therefore$ Conclusion : The decision does not change if the confidence level increases to 99%.

QUESTION 9

Question 9

i)

x	y	xy	x^2
97	24	2328	9409
85	29	2465	7225
98	26	2548	9604
105	24	2520	11025
120	24	2880	14400
151	22	3322	22801
140	23	3220	19600
134	23	3082	17956
146	21	3066	21316
$\sum x = 1076$	$\sum y = 216$	$\sum xy = 25431$	$\sum x^2 = 133336$

$$b_1 = \frac{25431 - \frac{(1076)(216)}{9}}{133336 - \frac{(1076)^2}{9}} = -0.08372$$

$$\bar{y} = \frac{216}{9} = 24$$

$$b_0 = 24 - (-0.08372)\left(\frac{1076}{9}\right) = 34.0092$$

$$\hat{y} = 34.0092 - 0.08372x$$

ii)

$$y = 34.0092 - 0.08372(125)$$

$$y = 34.0092 - 10.465$$

$$y = 23.5442\text{km}$$

QUESTION 10

DATE

QUESTION 10

Table 7 show a test scores (%) for three different groups of athletes that received particular treatment before they begin the performance test.

TABLE 7 : Performance Test Scores

Memory booster	Placebo	Without treatment
70	37	3
77	43	10
83	50	17
90	57	23
97	63	30

Conduct an analysis using a one-way ANOVA to test whether these data provide evidence to support the claim that the treatments will have different effects. Use a significance level of 0.01.

Hypothesis :

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

$H_1 :$ at least one mean is different

Category 1 (Memory booster)

$n = 5$

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

$= 83.4$

$$S = \sqrt{\frac{(70-83.4)^2 + (77-83.4)^2 + (83-83.4)^2 + (90-83.4)^2 + (97-83.4)^2}{4}}$$

$= 10.5971$

Category 2 (Placebo)

$n = 5$

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

$= 50$

$$S = \sqrt{\frac{(37-50)^2 + (43-50)^2 + (50-50)^2 + (57-50)^2 + (63-50)^2}{4}}$$

$= 10.4403$

DATE

Category 3 (No treatment)

$n = 5$

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

$= 16.6$

$$S = \sqrt{\frac{(3-16.6)^2+(10-16.6)^2+(17-16.6)^2+(23-16.6)^2+(30-16.6)^2}{4}}$$

$= 10.5972$

Mean between samples :

$$\bar{\bar{x}} = \frac{83.4+50+16.6}{2}$$

$= 50$

$$S^2 = \frac{(83.4-50)^2+(50-50)^2+(16.6-50)^2}{2}$$

$= 1115.56$

Variance between samples $= ns^2$

$= 5(1115.56)$

$= 5577.8$

Variance within samples :

$$S_p^2 = \frac{(10.5971)^2+(10.4403)^2+(10.5972)^2}{3}$$

$= 111.2$

Test statistic, F:

$$F = \frac{ns_{\bar{x}}^2}{S_p^2} = \frac{5577.8}{111.2}$$

$= 50.16$

Degree of freedom ;

Numerator $= k-1 = 3-1$

$= 2$

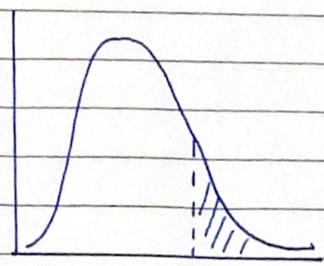
Denominator $= k(n-1) = 3(5-1)$

$= 12$

DATE

Critical value of F with $\alpha = 0.01$ from F-distribution table

$\therefore$ F critical value
$= 6.927$



$\therefore$ Conclusion : Since $F_{test\ statistic} > F_{critical\ value}$ $(50.16 > 6.927)$, we reject the null hypothesis. There is sufficient evidence to claim that the treatments will have different effects.