PROBABILITY & STATISTICAL DATA ANALYSIS SECTION-02

SECI2143 20202021/2

GROUP/ASSIGNMENT:

GROUP 4(HIV)/ASSIGNMENT 3

LECTURER'S NAME:

DR CHAN WENG HOWE

GROUP MEMBERS:

NAME	MATIRC NO.
HONG PEI GEOK	A20EC0044
LOW JUNYI	A20EC0071
TAN YONG SHENG	A20EC0157
VINCENT BOO EE KHAI	A20EC0231

Q1

a) $n = 25$ $\quad \bar{x} \pm (t_{0.1,24}) \frac{s}{\sqrt{n}} = (19.5) \pm (1.711)\left(\frac{9.88}{\sqrt{25}}\right)$

$\bar{x} = 19.5$

$s = 9.88$ $\quad = (19.5) \pm (3.381)$

$df = 25 - 1 = 24$ $\quad = (16.12, 22.88)$

$t_{0.1,24} = 1.711$

I'm 90% confidence that the mean time for woman's 100-performance is between 16.12 and 22.88.

b) $n = 360$ $\quad SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $\quad \hat{p} \pm (z)(SE) = (0.444) \pm (2.575)(0.02619)$

$z = 2.575$ $\quad = (0.377, 0.512)$

$\hat{p} = \frac{160}{360}$ $\quad = \sqrt{\frac{0.444(0.556)}{360}}$

$= 0.444$ $\quad = 0.02619$

I'm 99% confidence that the population of food stores that offers promotion is between 0.377 and 0.512.

Q2

a) $H_0: p = 0.1$

$H_1: p < 0.1$

$\hat{p} = \frac{33}{350}$

$= 0.0943$

$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

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

$= -0.36$

$P(z < -0.36) = 0.5 - 0.1406$

$= 0.3594$

Since $0.3594 > 0.05$, fail to reject H_0. There is sufficient evidence to shows that the incorrect result is less than 10 %.

b) $n = 16$

$\bar{x} = 58400$

$\sigma = 652$

$\alpha = 0.01$

$C.V = Z_{0.05} = 2.575$

$H_0: \mu = 58000$

$H_1: \mu \neq 58000$

$z = \frac{\bar{x} - \mu_x}{\sigma / \sqrt{n}}$

$= \frac{58400 - 58000}{652/\sqrt{16}}$

$= 2.454$

Conclusion: Since $2.454 < 2.575$, fail to reject H_0. There is sufficient evidence to conclude that the average compressive strength of steel is 58000 psi.

Question 3

a.

$\bar{x} = 10.06$

$n = 10$

$\mu = 10$

$\alpha = 0.01$

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

$= 0.056$

$s = 0.2366$

$H_0 : \mu = 10$

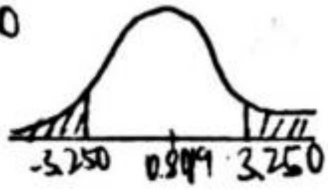
$H_1 : \mu \neq 10$

$t_{0.005,9} = 3.250$

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

$= \dfrac{10.6 - 10}{0.2366/\sqrt{10}}$

$= 0.8019$



∴ H_0 failed to rejected. There is sufficient evidence to conclude that the everage content volume of Brand X car lubricants is 10l

b.

$n = 14$

$\sigma = 0.44$

$s^2 = 0.4771$

$s = 0.6907$

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

$\chi^2_{0.975,13} = 5.009$

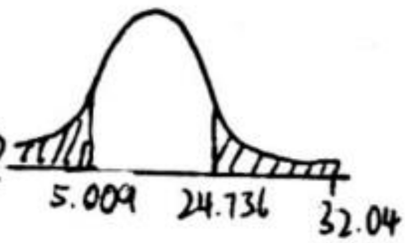
$H_0 : \sigma = 0.44$

$H_1 : \sigma \neq 0.44$

$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

$= \dfrac{13(0.4771)}{(0.44)^2}$

$= 32.04$



∴ Reject H_0. There is sufficient evidence to conclude that the vitamin supplement has effect of birth on standard deviation of the birth wight.

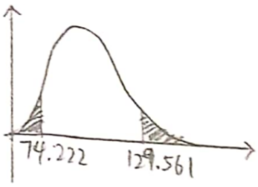
Q4

$\sigma^2 = 0.18$ $H_0: \sigma^2 = 0.18$

$n = 101$ $H_1: \sigma^2 \neq 0.18$

$s^2 = 0.165$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(101-1)(0.165)}{(0.18)} = 91.67$$



$\chi^2_{0.025,100} = 129.561$

$\chi^2_{0.975,100} = 74.222$

Since $74.222 < 91.67 < 129.561$, fail to reject H_0. There is sufficient evidence to show that the new joint inspector is making satisfactory measurements.

QUESTION 5

let μ_1 = mean spending of 18-34 age group

μ_2 = mean spending of 35+ group

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$\alpha = 0.05$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - D}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{76.40 - 71.20}{\sqrt{\frac{25.30}{15} + \frac{22.20}{18}}}$$

$$= 3.043$$

$$V = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

$$= \frac{\left(\frac{25.30}{15} + \frac{22.20}{18}\right)^2}{\frac{\left(\frac{25.30}{15}\right)^2}{14} + \frac{\left(\frac{22.20}{18}\right)^2}{17}}$$

$$= 29.13$$

$$\approx 29$$

$t_{0.025,29} = 2.045$, reject H_0 if $t_0 > 2.045$ or $t_0 < -2.045$

$\therefore$ Since $t_0 = 3.043 > 2.045$, thus reject H_0 and there is sufficient evidence to conclude that the mean spending between 18-34 age and 35+ group are different.

Question 6

$H_0 : \sigma_1 = \sigma_2$

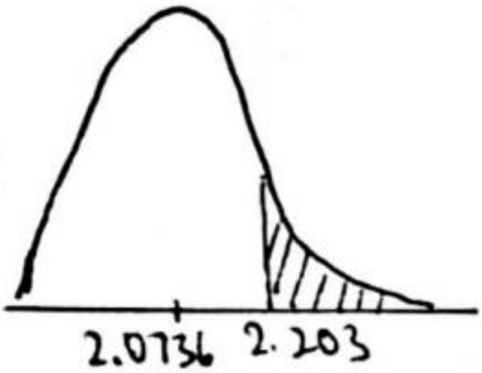
$H_1 : \sigma_1 > \sigma_2$

$\alpha = 0.05$

$$F = \frac{\sigma_1^2}{\sigma_2^2}$$

$$= \frac{(3.6)^2}{(2.5)^2}$$

$$= 2.0736$$



$F_{0.05, 15, 20} = 2.2033$

$\therefore H_0$ fail to reject. There is no sufficent evidents to conclude that the standard variation of processors from batch 1 is greater than batch 2.

Question 7

$H_0 : \mu_D = 0 \quad \alpha = 0.05$
$H_1 : \mu_D < 0$

Let x_1 = Number of words spelled correctly before training,
x_2 = Number of words spelled correctly after training
D = Differences of the values of the pairs

Degree of freedom, $df = n-1$
$= 10-1$
$= 9$

Critical value, C.V : $t_{0.05,9} = -1.833$ (left-tailed test)

Test value :

Before (x_1)	After (x_2)	$D = x_1 - x_2$	$D^2 = (x_1 - x_2)^2$
203	225	-22	484
390	410	-20	400
389	402	-13	169
271	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
		$\Sigma D = -207$	$\Sigma D^2 = 5131$

Standard deviation difference, S_D

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

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

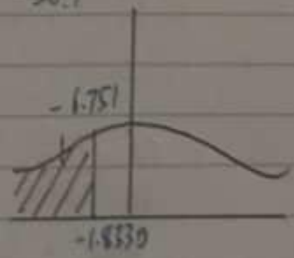
$$= 9.696$$

Mean of Differences, $\bar{D} = \frac{\Sigma D}{n}$

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

T-test value, $t = \frac{\bar{D} - \mu_0}{\left(\frac{S_D}{\sqrt{n}}\right)}$

$$= \frac{-20.7 - 0}{\left(\frac{9.696}{\sqrt{10}}\right)} = -6.751$$ #



$\therefore -6.751 < -1.833$, the null hypothesis is rejected. Hence, there is enough evidence to conclude that training can increase the number of words spelled correctly.

Question 8

i) the value of correlation coefficient, r

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

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

$$= \frac{9.35}{\sqrt{(6.3182)(19.5)}}$$

$r = 0.8424$ #

Weight of Plastic Usage, x	Value of Size Household, y	xy	x^2	y^2
0.27	2	0.54	0.0729	4
1.41	3	4.23	1.9881	9
2.19	3	6.57	4.7961	9
2.83	6	16.98	8.0089	36
2.19	4	8.76	4.7961	16
1.81	2	3.62	3.2761	4
0.85	1	0.85	0.7225	1
3.05	5	15.25	9.3025	25
$\sum x = 14.6$	$\sum y = 26$	$\sum xy = 56.8$	$\sum x^2 = 32.9632$	$\sum y^2 = 104$

ii) Significance level : $95\% = (1-a)\,100\%$

$1 - a = 0.95$

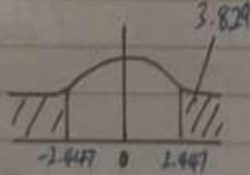
$a = 0.05$

Critical value of $t_{0.025,6} = -2.447,\ 2.447$

$H_0 : p = 0$

$H_1 : p \neq 0$

Degree of freedom, df $= n-2$
$= 8-2$
$= 6$



test, $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$

$$= \frac{0.8424}{\sqrt{\frac{1-(0.8424)^2}{8-2}}}$$

$= 3.829$ #

$\therefore 3.829 > 2.447$, the null hypothesis is rejected. The variable x and y are really correlates using 95% confidence level.

8) iii) Significance level : 99% = $(1-a)100\%$

$1-a = 0.99$

$a = 0.01$

$\frac{a}{2} = 0.005$

$H_0 : p = 0$

$H_1 : p \neq 0$

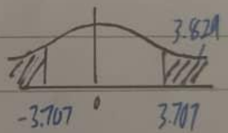
Degree of freedom, $df = n-2$

$= 8-2$

$= 6$

Critical value of $t_{0.05,6} = -3.707, 3.707$

test, $t = 3.829$



$\therefore$ $3.829 > 3.707$, the null hypothesis is rejected. The decision in (ii) will not change if the confidence level rised to 99%.

QUESTION 9

i)

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

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

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

$$= -0.08372$$

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

$$= \frac{216}{9} - (-0.08372)\left(\frac{1076}{9}\right)$$

$$= 34.00919$$

$$\hat{y} = 34.00919 - 0.08372x$$

ii) $x = 125$

$$\hat{y} = 34.00919 - 0.08372(125)$$

$$= 23.54$$

$\therefore$ Mileage for a car which has 125 cubic km engine volume is 23.54 km.

Question 10

Memory Booster, $\bar{x}_1 = (70+77+83+90+97)/5 = 83.4$
Placebo, $\bar{x}_2 = (37+43+50+57+63)/5 = 50$
Without treatment, $\bar{x}_3 = (3+10+17+23+30)/5 = 16.6$

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1: at least one mean is different

Memory Booster, $S_1 = \sqrt{\dfrac{(70-83.4)^2+(77-83.4)^2+(83-83.4)^2+(90-83.4)^2+(97-83.4)^2}{5-1}}$

$= 10.5972$

Placebo, $S_2 = \sqrt{\dfrac{(37-50)^2+(43-50)^2+(50-50)^2+(57-50)^2+(63-50)^2}{5-1}}$

$= 10.4403$

Without treatment, $S_3 = \sqrt{\dfrac{(3-16.6)^2+(10-16.6)^2+(17-16.6)^2+(23-16.6)^2+(30-16.6)^2}{5-1}}$

$= 10.5972$

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

$= 50$

$S_{\bar{x}}^2 = \dfrac{(83.4-50)^2+(50-50)^2+(16.6-50)^2}{3}$

$= 1115.5601$

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

$= 111.2004$

Test statistic, $F = \dfrac{n S_{\bar{x}}^2}{S_p^2} = \dfrac{5(1115.5601)}{111.2004} = 50.1599$

Numerator $= k-1 = 3-1 = 2$
Denominator $= k(n-1) = 3(5-1) = 12$ Critical value, $F_{0.01,2,12} = 6.927 < 50.1599$

Since $F_{\text{test statistic}} > F_{\text{critical value}}$, thus reject H_0.
There is sufficient evidence to support the claim that the treatments will have different effects.