Subject: Assignment 7

Date:

Group Members
1) Khor Yong Xin (A19EC0061)
2) Ham Jing Yi (A19EC0048)
3) Tee Hui You (A19EC0170)

Question 1

a. i)

x	y	xy	y^2	x^2
8.80	66.1	581.68	4369.21	77.44
7.70	52.4	403.48	2745.76	59.29
8.10	48.9	396.09	2391.21	65.61
8.40	48.1	404.04	2313.61	70.56
8.50	42.0	357.00	1764.00	72.25
7.35	38.3	281.51	1466.89	54.02
7.95	31.3	248.84	979.69	63.20
56.80	327.10	2672.64	16030.37	462.37

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

$$= \frac{2672.64 - (56.80)(327.10)/7}{\sqrt{[462.37 - (56.80^2)/7][16030.37 - (327.10^2)/7]}}$$

$$= 0.5538$$

b. i)

x	y	xy	x^2
15	2.29	34.35	225
17	3.39	57.63	289
18	3.27	58.86	324
15	2.64	39.60	225
16	2.89	46.24	256
19	3.32	63.08	361
17	2.97	50.49	289
16	2.53	40.48	256
18	3.13	56.34	324
19	3.57	67.83	361
170	30.00	514.90	2910

$n = 10$

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

$$= \frac{514.90 - (170)(30.00)/10}{2910 - 170^2/10}$$

$$= 0.245$$

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

$$= \frac{30.00}{10} - 0.245\left(\frac{170}{10}\right)$$

$$= -1.165$$

$$\therefore \hat{y} = -1.165 + 0.245x$$

ii) At $x = 18$,

$$\hat{y} = -1.165 + 0.245(18)$$

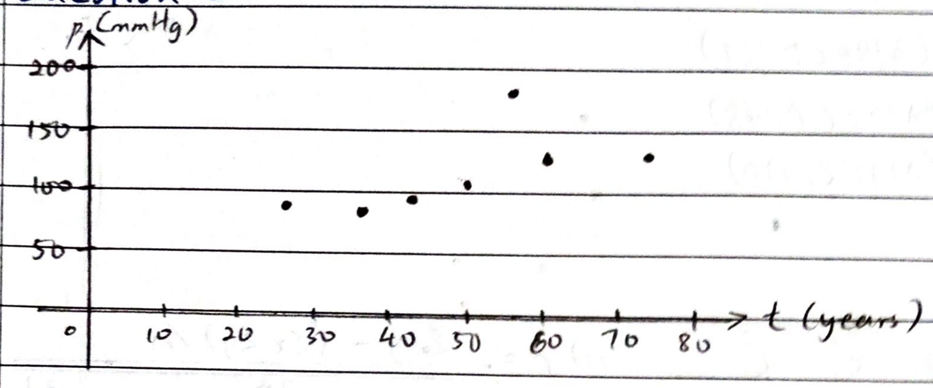
$$= 3.245$$

$\therefore$ Birth weight of baby = 3.245 kg

Subject :   Date:

Question 2

a)



The correlation coefficient calculated is a positive value as there is an obvious positive linear relationship observed between variables. The correlation coefficient calculated is estimated to be a large value as almost all points of the scatter plot lies around a straight line. However, the outlier, (56, 182) may cause severe effect on the line of best fit produced and correlation coefficient calculated.

b)

t	42	74	48	35	56	26	60	$\Sigma t = 341$
p	98	130	120	88	182	80	135	$\Sigma p = 833$
t^2	1764	5476	2304	1225	3136	676	3600	$\Sigma t^2 = 18181$
p^2	9664	16900	14400	7744	33124	6400	18225	$\Sigma p^2 = 106397$
tp	4116	9620	5760	3080	10192	2080	8100	$\Sigma tp = 42948$

$$r = \frac{\Sigma tp - (\Sigma t \Sigma p)/n}{\sqrt{(\Sigma p^2 - (\Sigma p)^2/n)(\Sigma t^2 - (\Sigma t)^2/n)}}$$

$$= \frac{42948 - (341)(833)/7}{\sqrt{(106397 - 833^2/7)(18181 - 341^2/7)}}$$

$$= 0.7013$$

$\therefore$ Since this is a single independent variable case, $R^2 = r^2$. Thus, it fits a regression model to these data.

c)

$$b_1 = \frac{\Sigma tp - (\Sigma t \, \Sigma p)/n}{\Sigma t^2 - (\Sigma t)^2/n}$$

$$= \frac{42948 - (341)(833)/7}{18181 - 341^2/7}$$

$$= 1.5095$$

$$b_0 = \bar{p} - b_1 \bar{t}$$

$$= \frac{833}{7} - 1.5095\left(\frac{341}{7}\right)$$

$$= 45.466$$

d) $\hat{p} = 45.466 + 1.5095t$

e) At $t = 40$, $p = 45.466 + 1.5095(40)$

$= 105.846$

$\therefore$ Blood pressure of a 40-years old patient is 105.846 mm Hg.

Question 3

i)

X	0.27	1.41	2.19	2.83	2.19	1.81	0.85	3.05	$\Sigma X = 14.60$
Y	2	3	3	6	4	2	1	5	$\Sigma Y = 26$
X^2	0.0729	1.9881	4.7961	8.0089	4.7961	3.2761	0.7225	9.3025	$\Sigma X^2 = 32.9632$
Y^2	4	9	9	36	16	4	1	25	$\Sigma Y^2 = 104$
XY	0.54	4.23	6.57	16.98	8.76	3.62	0.85	15.25	$\Sigma XY = 56.80$

$$r = \frac{\Sigma XY - (\Sigma X \Sigma Y)/n}{\sqrt{(\Sigma X^2 - (\Sigma X)^2/n)(\Sigma Y^2 - (\Sigma Y)^2/n)}}$$

$$= \frac{56.80 - (14.60)(26)/8}{\sqrt{(32.9632 - (14.60)^2/8)(104 - (26)^2/8)}}$$

$$= 0.8424$$

ii) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

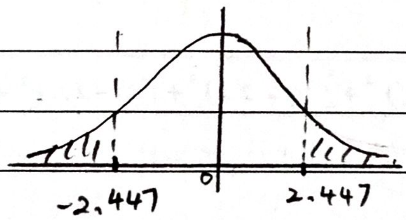
$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

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

$$= 3.8293$$

$\alpha = 0.05$, $df = 8 - 2 = 6$

critical value, $t_{0.025, 6} = 2.447$



Decision:

Since test statistic > 2.447, H_0 is rejected as test statistic lies in critical region.

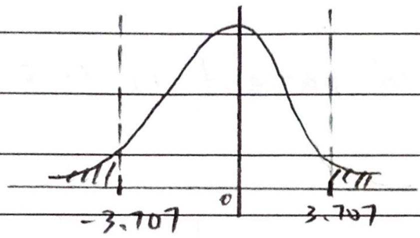
Conclusion:

There is sufficient evidence to conclude that variables X and Y are really correlate at confidence level of 95%

Subject :   Date:

iii) if 99% confidence level, $\alpha = 0.01$

critical value, $t_{0.01/2, 6} = 3.707$



Decision:

Test statistic > 3.707. Test statistic lies in critical region.

Hence, H_0 is rejected.

Conclusion:

There is sufficient evidence to conclude that variable X and Y are really correlate at 99% confidence level.

There is no change in decision in part (ii)

Question 4

a) $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$

H_0 : At least one of the mean is different

b) $$S_B = \sqrt{\frac{2(14-14.8)^2 + (16-14.8)^2 + (13-14.8)^2 + (17-14.8)^2}{5-1}}$$

$= 1.6$

$$S_D = \sqrt{\frac{(24-22)^2 + (19-22)^2 + (21-22)^2 + (26-22)^2 + (20-22)^2}{5-1}}$$

$= 2.9$

c) $$\bar{\bar{x}} = \frac{21.6 + 14.8 + 15.0 + 22.0}{4}$$

$= 18.35$

$$S_{\bar{x}}^2 = \frac{(21.6-18.35)^2 + (14.8-18.35)^2 + (15.0-18.35)^2 + (22.0-18.35)^2}{4-1}$$

$= 15.9$

$nS_{\bar{x}}^2 = 5(15.9) = 79.5$

Subject : Date:

d) $S_p^2 = \dfrac{3.4^2 + 1.6^2 + 3.8^2 + 2.9^2}{4}$

$= 9.2425$

e) test statistic, $F = \dfrac{nS_{\bar{x}}^2}{S_p^2}$

$= \dfrac{79.5}{9.2425}$

$= 8.6016$

f) numerator $= k - 1$

$= 4 - 1$

$= 3$

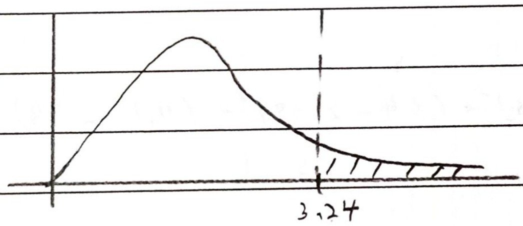
denominator $= k(n-1)$

$= 4(5-1)$

$= 16$

g) $\alpha = 0.05$

$F_{0.05,3,16} = 3.24$



Decision:

Test statistic > 3.24. Test statistic lies in the critical region.

Hence, H_0 is rejected.

Conclusion:

At 5% significance level, there is sufficient evidence to conclude that the mean number of customers served per hour by each of these four employees is not the same.

Subject : Date:

Question 5

$H_0 : \mu_A = \mu_B = \mu_C$

H_1: At least 1 of the mean is different

$\bar{x}_A = \frac{5.2+4.7+8.1+6.2+3.0}{5}$

$= 5.44$

$\bar{x}_B = \frac{9.1+7.1+8.2+6.0+9.1}{5}$

$= 7.9$

$\bar{x}_C = \frac{2.4+3.4+4.1+1.0+1.0}{5}$

$= 2.38$

$\bar{\bar{x}} = \frac{5.44+7.9+2.38}{3}$

$= 5.24$

$s_A^2 = \frac{(5.2-5.44)^2+(4.7-5.44)^2+(8.1-5.44)^2+(6.2-5.44)^2+(3.0-5.44)^2}{5-1}$

$= 3.553$

$s_B^2 = \frac{2(9.1-7.9)^2+(7.1-7.9)^2+(8.2-7.9)^2+(6.0-7.9)^2}{5-1}$

$= 1.805$

$s_C^2 = \frac{(2.4-2.38)^2+(3.4-2.38)^2+(4.1-2.38)^2+2(1.0-2.38)^2}{5-1}$

$= 1.952$

$s_{\bar{x}}^2 = \frac{(5.44-5.24)^2+(2.38-5.24)^2+(7.9-5.24)^2}{3-1}$

$= 7.6476$

$n s_{\bar{x}}^2 = 5(7.6476)$

$= 38.238$

$s_p^2 = \frac{3.553+1.805+1.952}{3}$

$= 2.437$

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

$= 38.238 / 2.437$

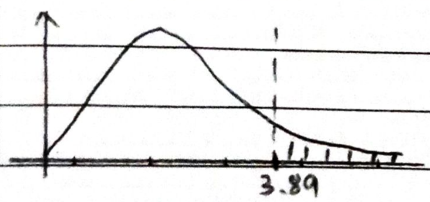
$= 15.69$

numerator $= 3-1 = 2$

denominator $= 3(5-1) = 12$

$\alpha = 0.05$

$F_{0.05,2,12} = 3.89$



Decision:

Test statistic > 3.89. Test statistic lies in critical region. H_0 is rejected.

Conclusion:

There is sufficient evidence to conclude that mean number of hours of relief provided by the tablets is not the same for three brands.

Subject :   Date:

Question 6

$H_0 : \mu_M = \mu_P = \mu_W$

H_1 : At least one mean is different

$\bar{x}_M = \dfrac{70+77+83+90+97}{5} = 83.4$

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

$\bar{x}_W = \dfrac{3+10+17+23+30}{5} = 16.6$

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

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

$= 1115.56$

$ns_{\bar{x}}^2 = 1115.56(5)$

$= 5577.8$

$s_p^2 = \dfrac{112.30+109.00+112.30}{3}$

$= 111.2$

Test statistic, $F = \dfrac{ns_{\bar{x}}^2}{s_p^2}$

$= 5577.8/111.2$

$= 50.16$

numerator $= k-1$

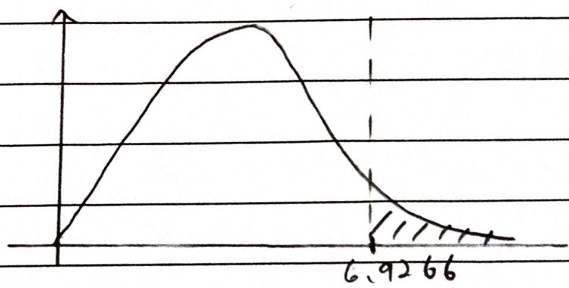
$= 3-1$

$= 2$

denominator $= k(n-1)$

$= 3(5-1)$

$= 12$



$\alpha = 0.01$

critical value, $F_{0.01,2,12} = 6.9266$

Test statistic > critical value. Test statistic lies in the critical region.

H_0 is rejected.

There is sufficient evidence to conclude that treatments will have different effects at significance level of 0.01.