Statistics and Interencing

Homework 3

Name:- Basil Khowaja

id: BK08432

Section 9.1)

Q1) a) Population mean of $\bar{x}$

$$\mu_x = 4.1$$

Population mean of $\bar{y}$

$$\mu_y = 4.5$$

The mean of $\bar{x} - \bar{y}$ is

$$\mu_{\bar{x}-\bar{y}} = \mu_x - \mu_y$$

$$= 4.1 - 4.5 = -0.4$$

b) $\sigma_x = 1.8$

$\sigma_y = 2.0$

$n = 100$

$$Var(\bar{x} - \bar{y}) = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{n} = \frac{1.8^2}{100} + \frac{2.0^2}{100}$$

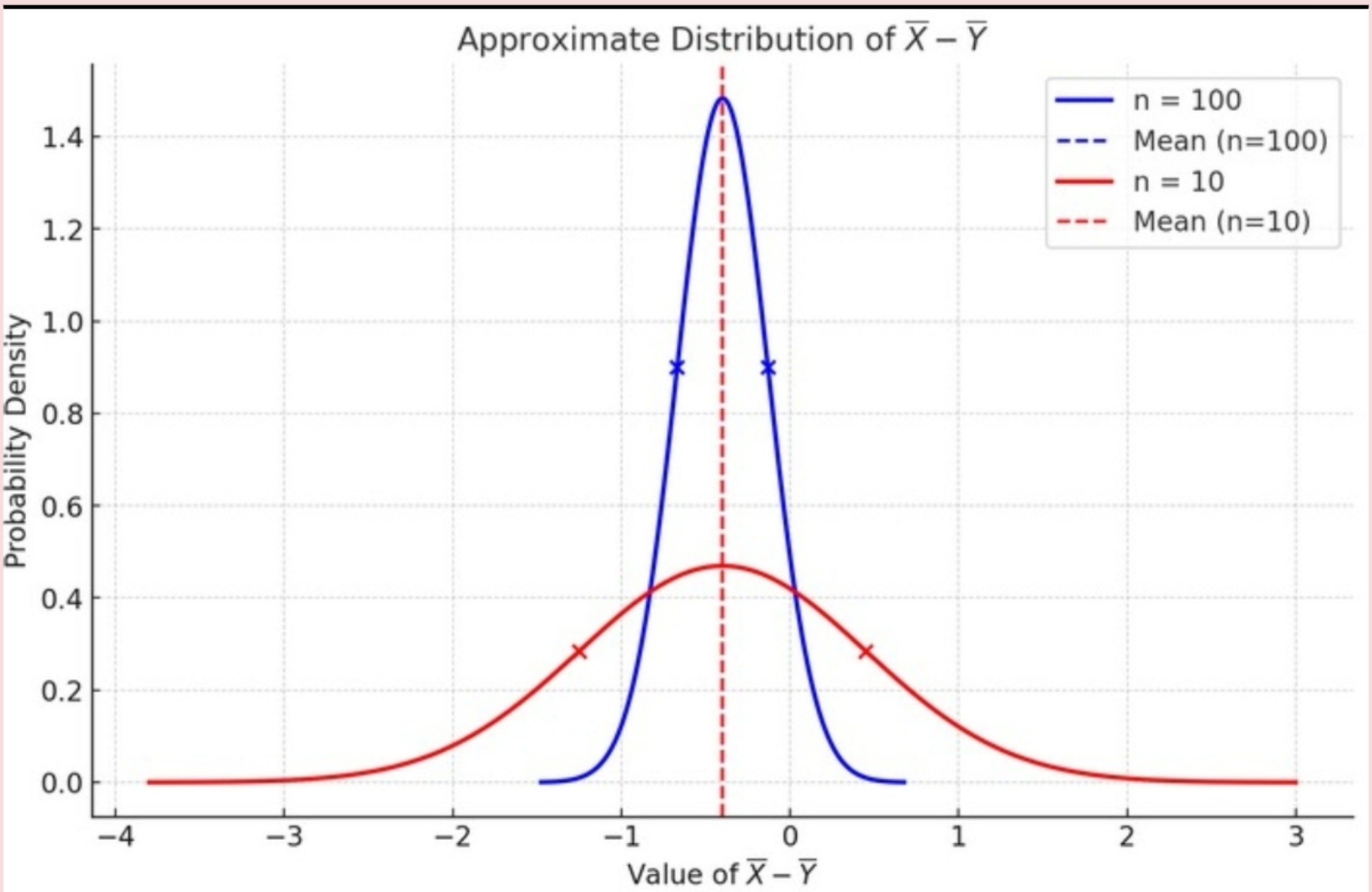
$$= 0.0724$$

Standard deviation of var $(\bar{x} - \bar{y})$

$$\hookrightarrow = \sqrt{0.0724} \approx 0.269$$

c) the Sampling distribution of $\bar{x} - \bar{y}$ is approximately normal (by the central limit theorem) since both $\bar{x}$ and $\bar{y}$ are based on large Sample Size of $n = 100$.

Approximate distribution of $\bar{x} - \bar{y}$



↳ the shape of the curve depends on the sample size ;

1) larger samples $n = 100$ result in narrower, taller curves because variability decreases.

2) Smaller Samples $(n=10)$ lead to wider, flatter curves due to increased variability

Q2)

a) $\bar{x}_1 - \bar{x}_2 = 64.9 - 63.1 = 1.8$

Standard error of the difference is:

$$SE_{diff} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{0.0081 + 0.0121}$$

$$SE_{diff} \approx 0.142$$

$$CI = (\bar{x}_1 - \bar{x}_2) \pm z^* . SE_{diff}$$

$$CI = 1.8 \pm 1.96\,(0.142)$$

$$CI = (1.522,\ 2.078)$$

b) $H_0 = \mu_1 - \mu_2 = 1$

$H_a : \mu_1 - \mu_2 > 1$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2 \text{ under } H_0)}{SE \text{ diff}}$$

$$z = \frac{1.8 - 1}{0.142} = \frac{0.8}{0.142} \approx 5.63$$

for $\alpha = 0.001$ the critical z values from the Standard normal distribution is approx 3.09

Since $z = 5.63 > 3.09$ we reject the null hypothesis

c) from Part b, the calculated test statistic $z = 5.63$

using the Standard normal table

$$P(z > 5.63) \approx 0$$

Since the P-value is nearly 0 we reject the null hypothesis

d) H_0: the difference in means is exactly 1 inch;

$$H_0: \mu_2 - \mu_1 = 1$$

$$H_a: \mu_2 - \mu_1 > 1$$

$\mu_2 \rightarrow$ refers to the Population mean height of younger women

$\mu_1 \rightarrow$ refers to the Population mean height of older woman.

$\Rightarrow$ the test is one tailed

Q3)

a) $H_0 : \mu_C = \mu_T$ (the mean pain levels are equal)

$H_a : \mu_C > \mu_T$ (the control group mean is greater)

$$t = \frac{(\bar{x}_C - \bar{x}_T)}{\sqrt{\frac{s_C^2}{n_C} + \frac{s_T^2}{n_T}}}$$

$$t = \frac{5.2 - 3.1}{\sqrt{\frac{2.3^2}{43} + \frac{2.3^2}{43}}}$$

$$t = \frac{2.1}{\sqrt{\frac{5.29}{43} + \frac{5.29}{43}}} = \frac{2.1}{\sqrt{0.246}} \approx 4.23$$

$df = 43 + 43 - 2 = 84$

for $\alpha = 0.01$

$$t_\alpha = 2.37$$

Since $t = 4.23 > 2.37$

we reject the null hypothesis

b) $H_0: \mu_C - \mu_T = 1$

$H_a: \mu_C - \mu_T > 1$

$$t = \frac{(\bar{x}_C - \bar{x}_T) - 1}{\sqrt{\frac{s_C^2}{n_C} + \frac{s_T^2}{n_T}}}$$ $\rightarrow$ found earlier

$$t = \frac{2.1 - 1}{0.496} \approx 2.22$$

Since $t = 2.22 < 2.37$

we fail to reject

the null hypothesis

Q 6)

Part a)

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 > 0$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

$$z = \frac{18.12 - 16.87}{\sqrt{\frac{1.6^2}{40} + \frac{1.4^2}{32}}}$$

$$z = \frac{1.25}{\sqrt{\frac{2.56}{40} + \frac{1.96}{32}}} \approx 3.53$$

critical z value $\simeq 2.33$

for $\alpha = 0.01$

Since $z = 3.53 > 2.33$ we reject H_0

b)

$$z = \frac{(\bar{m} - \bar{u}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

critical value at $\alpha = 0.01$ is

$z = 2.33$ (right tail)

SE from Part (a) $= 0.354$

under the alternative $\mu_1 - \mu_2 = 1$, the null value shifts and the test Statistic becomes

$$z' = \frac{2.33 - 1}{0.354}$$

$$z' \approx 3.76$$

$\Rightarrow$ the Probability of type 2 is the Probability that z-value fall below the critical value $z = 2.33$ under the Shifted distribution

$\Rightarrow$ to compute this, we find the left tail probability of $z = 2.33$ under the shifted distribution

$$N(1, 0.354)$$

$$P(z < 2.33 - 3.76)$$

$$P(z < -1.43) \approx 0.0764$$

Ans

c) $$n = \frac{(z_\alpha + z_\beta)^2 \cdot \left(\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right)}{\Delta^2}$$

$$SE^2 = \frac{\sigma_1^2}{m} = \frac{1.6^2}{40} = 0.064$$

$$n = \frac{(1.645 + 1.28)^2 \cdot \left(0.064 + \frac{1.4^2}{n}\right)}{1}$$

$$n = 8.56\left(0.064 + \frac{1.96}{n}\right)$$

Starting with $n = 50$ we

find $n \approx 50.5$

hence $\boxed{n = 51}$

d) $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

$$t = \frac{18.12 - 16.87}{0.354} \approx 3.53$$

$$df = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{\left(\frac{s_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n}\right)^2}{n-1}}$$

$df \approx 69.5$ (round down)

$df = 69$

for one tailed t-test

with $\alpha = 0.01$ with $df = 68$

the critical t-value from the table is approx

$t \approx 2.38$

Since $t = 3.53 > 2.38$

we reject H_0.

Q10)

a) $H_0 : \mu_1 - \mu_2 = 5$

$H_a : \mu_1 - \mu_2 > 5$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - 5}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

$$SE = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} = \sqrt{\frac{1.2^2}{32} + \frac{1.1^2}{38}}$$

$$SE = \sqrt{0.0768} \approx 0.277$$

$$Z = \frac{5.8-5}{0.277} \approx 2.89$$

$$Z_{crit} = 3.09$$

Since $z = 2.89 < 3.09$

we fail to reject H_0

b) $$z' = \frac{(\bar{x}_1 - \bar{x}_2) - 5}{SE}$$

under the alternative hypothesis $\mu_1 - \mu_2 = 6$

$$\text{Shift} = \frac{6-5}{SE} \approx 3.61$$

$$z' = z^* - \text{Shift}$$

$$= 3.09 - 3.61 = -0.52$$

$$P(z < -0.52) \approx 0.3015$$

Ans

Section 9.2)

Q 19)

unpooled

$$t = \frac{(\bar{x}-\bar{y})-\Delta_0}{\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}}$$

where $\Delta_0 = -10$

$$SE = \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

$$SE = \sqrt{9.04122} \approx 3.01$$

$$t = \frac{-13.6+10}{3.01} = \frac{-3.6}{3.01}$$

$$t \approx -1.20$$

$\alpha = 0.01$, $df = m+n-2 = 10$

$$t_{0.01, 10} = -2.763$$

Since $t > -2.763$ we fail to reject the H_0

Pooled:-

$$S_p^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}$$

$$S_p^2 = \frac{(6-1)(5.03)^2 + (6-1)(5.38^2)}{6+6-2}$$

$$S_p^2 = \frac{271.2365}{10} \approx 27.12$$

$$S_p = \sqrt{Sp_2} \approx 5.21$$

$$t = \frac{(\bar{x} - \bar{y}) - \Delta_0}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

$$t = \frac{-13.6 + 10}{3.01} \approx -1.20$$

since $t > -2.763$

we fail to reject H_0

Q21)

$H_0: \mu_1 = \mu_2$

$H_a: \mu_2 > \mu_1$

unPooled:

$$SE = \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

$$SE = \sqrt{\frac{0.53^2}{8} + \frac{0.87^2}{10}}$$

$$SE \approx 0.333$$

$$t = \frac{2.53 - 1.71}{0.333} \approx 2.46$$

$$df = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{\left(\frac{s_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n}\right)^2}{n-1}}$$

$$df = \frac{0.0123}{0.000813} \approx 15.13$$

from the t_table for $\alpha = 0.01$

and $df = 15$

critical t value $= 2.602$

Since $t = 2.46 > 2.602$ we fail

to reject H_0

Pooled:

$$Sp^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}$$

$$Sp^2 = \frac{(8-1)(0.53)^2 + (10-1)(0.87)^2}{8+10-2}$$

$$Sp^2 = \frac{8.7784}{16} \approx 0.5486$$

$$Sp = \sqrt{0.5486} \approx 0.74$$

$$t = \frac{2.53 - 1.71}{0.74\sqrt{\frac{1}{8} + \frac{1}{10}}}$$

$$t \approx 2.34$$

$df = 16$ and $\alpha = 0.01$

$$t_{crit} = 2.583$$

since $t = 2.34 < 2.583$

we fail to reject H_0

Section 9.3)

Q 36)

$H_0: \mu_D = 0$

$H_a: \mu_D > 0$

mean of differences

$$\bar{D} = \frac{58.9}{8} = 7.3625$$

$$S_D = \sqrt{\frac{\sum(D_i - \bar{D})^2}{n-1}}$$

$$s_D^2 = \frac{964.4730}{7} \approx 137.7819$$

$$s_D \approx 11.74$$

$$t = \frac{\bar{D}}{s_D / \sqrt{n}}$$

$$t = \frac{7.3625}{11.74 / \sqrt{8}} = \frac{7.3625}{4.15} \approx 1.77$$

for one tailed test

with $\alpha = 0.01$

and $df = n-1 = 7$

$$t_{crit} = 3.499$$

Since $t = 1.77 < 3.499$

we fail to reject H_0

Q37)

$$D_i = [0.29 - 0.07,\ 0.68 - 0.08,$$
$$0.47 - 0.09,\ \ldots\ 0.36 - 0.62]$$

$$\bar{D} = \frac{\sum D_i}{n} \approx 0.51$$

$$S_D = 0.40$$

$$SE_{\bar{D}} = \frac{S_D}{\sqrt{n}} = \frac{0.40}{\sqrt{33}} \approx 0.0696$$

for a 95% CI, $df = 33 - 1 = 32$

$$t_{crit} = 2.037$$

$$CI = \bar{D} \pm t_{crit} \cdot SE_{\bar{D}}$$

$$CI = 0.51 \pm 2.037\,(0.0696)$$

$$CI = (0.368,\ 0.652)$$

$$PI = \bar{D} \pm t_{crit}\left(\sqrt{s_D^2 + SE^2_{\bar{D}}}\right)$$

$$PI = 0.51 \pm 2.037\left(\sqrt{0.40^2 + 0.0696^2}\right)$$

$$PI = 0.51 \pm 0.826$$

$$PI = (-0.316, 1.336)$$

Q41)

a) $CI = \bar{x} \pm t^* \cdot SE$

for 95% CI and df = 40

$$t_{crit} \approx 2.021$$

$$CI = 3.75 \pm 2.021(1.15)$$

$$CI = 3.75 \pm 2.324$$

$$CI \Rightarrow (1.426, 6.074)$$

b) Ho : true average cholestrol level change = 0

Ha: true average cholestrol level change > 0

$$t = \frac{\bar{x} - 0}{SE} = \frac{9.05}{4.256} \approx 2.13$$

for one tailed test with $df = 36 - 1 = 35$

$$P(t > 2.13) \approx 0.02$$

⇒ at $\alpha = 0.05$ the P value is 0.02 is less than α, so we reject Ho

⇒ if $\alpha = 0.01$, the P_value of 0.02 is greater than α so we fail to reject Ho

c) $SE = \dfrac{\text{margin of error}}{t_{crit}}$

where $t_{crit} = 2.014$

$df = 45 - 1 = 44$

$$SE = \frac{1.155}{2.014} \approx 0.5736$$

for a 99% CI, $t_{crit} = 2.69$

↳ from t table (for df = 44)

Margin of error $= t_{crit} \cdot SE$

Margin of error $= 2.69 \times 0.5736$

" ≈ 1.54

$$\bar{x} = \frac{7.38 + 9.69}{2} = 8.535$$

$CI = 8.535 \pm 1.54$

$CI = (6.995,\ 10.075)$

Section 9.4)

Q 49)

$n_1 = 200$

$x_1 = 164$

$$\hat{P}_1 = \frac{x_1}{n_1} = \frac{164}{200} = 0.82$$

$n_2 = 200$

$x_2 = 140$

$$\hat{P}_2 = \frac{x_2}{n_2} = \frac{140}{200} = 0.70$$

$H_0 : P_1 = P_2$

$H_a : P_1 > P_2$

$$z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\hat{p} = \frac{164 + 140}{200 + 200} = \frac{304}{400} = 0.76$$

$$SE = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$SE = \sqrt{0.001824} \approx 0.0427$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{SE} = \frac{0.12}{0.0427} \approx 2.81$$

for one tailed test at

$\alpha = 0.05$

z crit $= 1.645$

Since $z = 2.81 > 1.645$ we reject H_0

Q 50)

for Perdue:

a) $\hat{P}_1 = \frac{u_1}{n_1} = \frac{45}{80} = 0.5625$

for tyson:

$n_2 = 80$

$\hat{P}_2 = \frac{u_2}{n_2} = \frac{14}{80} = 0.175$

$H_0: P_1 = P_2$

$H_a: P_1 \neq P_2$

$$z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\hat{p} = \frac{u_1 + u_2}{n_1 + n_2}$$

$$\hat{p} = \frac{45+14}{80+80} = \frac{59}{160} = 0.36875$$

$$SE = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$SE = \sqrt{0.005824} \approx 0.0763$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{SE} = \frac{0.3875}{0.0763} \approx 5.08$$

for 2 tailed test at $\alpha = 0.01$

the critical z value is

$$z_{crit} = 2.576$$

Since $z = 5.08 > 2.576$ we reject H_0

b) when $P_1 = 0.50$ and $P_2 = 0.25$

$$SE = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

$$SE = \sqrt{\frac{0.25}{80} + \frac{0.1875}{80}}$$

$$P_1 - P_2 = 0.50 - 0.25 = 0.25$$

$$z = \frac{P_1 - P_2}{SE} = \frac{0.25}{0.0739} \approx 3.38$$

for a significance level

$\alpha = 0.01$ 2 tailed

the critical z-values

are

$$z = \pm 2.576$$

$$P(z > 2.576) + P(z < -2.576)$$

the shifted z value is $z = 3.38$

$$P(z > 2.576) \approx 0.0049$$

Q60)

a) $v_1 = 5$, $v_2 = 10$

upper tailed test

$f = 4.75$

using an f - table

$P(f > 4.75) \approx 0.015$

b) $v_1 = 5$, $v_2 = 10$, $f = 2.00$

$P(f > 2.00) \approx 0.165$

c) $v_1 = 5$, $v_2 = 10$, 2 tailed

$f = 5.64$

$P(f > 5.64) \approx 0.01$

$P = 2 \times 0.01 = 0.02$

d) lower tailed test, $v_1 = 5$

$v_2 = 10$, $f = 0.200$

$P(f < 0.200) \approx 0.020$

e) $v_1 = 35$, $v_2 = 20$, upper tailed test, $f = 3.24$

$P(f > 3.24) \approx 0.001$

Q 61) Data for younger females:

YF : 29, 34, 33, 27, 28, 32, 31, 34, 32, 27

$n_1 = 10$

$$S_1^2 = \frac{\sum (x_i - \bar{x})^2}{n_1 - 1}$$

Data for older females

OF : 18, 15, 23, 13, 12

$n_2 = 5$

$H_0: \sigma_1^2 = \sigma_2^2$

$H_a: \sigma_1^2 \neq \sigma_2^2$

test statistic:

$$f = \frac{s_1^2}{s_2^2}$$

for younger females (s_1^2)

$$\bar{x}_1 = \frac{29+34+33+27+28+32+31+34+32+27}{10}$$

$$\bar{x}_1 = 30.7$$

sum of squared deviation $(x_i - \bar{x}_1)^2$

$$\sum (x_i - \bar{x}_1)^2 = 67.1$$

$$s_1^2 = \frac{67.1}{10-1} = \frac{67.1}{9} = 7.456$$

for older females :

$$\overline{x}_2 = \frac{18+15+23+13+12}{5}$$

$$\overline{x}_2 = 16.2$$

$$\sum (x_i - \overline{x}_2)^2 = 78.8$$

$$s_2^2 = \frac{78.8}{5-1} = \frac{78.8}{4} = 19.7$$

Test statistic:

$$f = \frac{s_1^2}{s_2^2} = \frac{7.456}{19.7} \approx 0.378$$

for $\alpha = 0.10$

$df_1 = 9, \quad df_2 = 4$

$f_{lower} = 0.275$

$f_{upper} = 5.999$

→ from f table

here $f = 0.378$ lies within $(0.275, 5.999)$

hence we fail to reject H_0.