SCHOOL OF COMPUTING

SEMESTER II 2020/2021

SECI2143 PROBABILITY & STATISTICAL DATA ANALYSIS

SECTION: 07

LECTURER: DR. NIES HUI WEN

ASSIGNMENT 3

GROUP 6: SISTHREEHOOD

GROUP MEMBERS:

1. EUNICE LIM XIAN NI (A20EC0034)
2. SHABRINA SALSABILA SAKRONI (A20EC0334)
3. TEH JING LING (A20EC0228)

Assigment 3

Question 1

ⓐ • $n = 25$ • $s = 9.88$ • 90% C.I

• $\bar{x} = 19.5$ second • $\alpha = 0.1$ %

• degree of freedom $= 25 - 1$
$= 24$

→ $\bar{x} \pm$ (t critical value) $\left(\frac{s}{\sqrt{n}}\right)$ $= 19.5 \pm (1.711)\left(\frac{9.88}{\sqrt{25}}\right)$

→ t 90% with df = 24

$t = 1.711$

$= 19.5 \pm (1.711)(1.976)$

$= 19.5 \pm 3.380936$

$= 19.5 \pm 3.381$

$19.5 + 3.381 = 22.881$

$19.5 - 3.381 = 16.119$

→ (16.119 , 22.881)

→ for a 90 % confidence interval, it can be assumed the value of mean is between 16.119 and 22.881.

ⓑ • $n = 360$ and 160 (special) • 99 % C.I

• $\alpha = 0.01$

• $\hat{p} = \frac{160}{360} = 0.4444...$

$\hat{p} = 0.444$

→ $\hat{p} \pm$ (t critical value) $\left(\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$ $= 0.444 \pm (2.575)\sqrt{\frac{0.444(1-0.444)}{360}}$

$= 0.444 \pm (2.575)\sqrt{0.0261865}$

$= 0.444 \pm 0.0674^{303}$

$= 0.444 \pm 0.0674$

$0.444 + 0.0674$
$= 0.5114$
$= 0.511$

$0.444 - 0.0674$
$= 0.3766$
$= 0.377$

→ (0.377 , 0.511)

∴ the population interval

((0.377 x 360) , (0.511 x 360))

(135.72 , 183.96) = (135 , 183)

→ we are 99 % confidant that the estimate population of food stores that offers population is between 135 and 183.

Question 2

ⓐ. $n = 350$; 33 result is incorrect

- $\hat{P} = \frac{33}{350} = 0.0942857$

  $= 0.094$

- $H_0 : P = 10\% = 0.10$
- $H_1 : P < 0.10$
- $\alpha = 0.05.$

- $q = 1 - P$

  $= 1 - 0.10$

  $= 0.90$ //

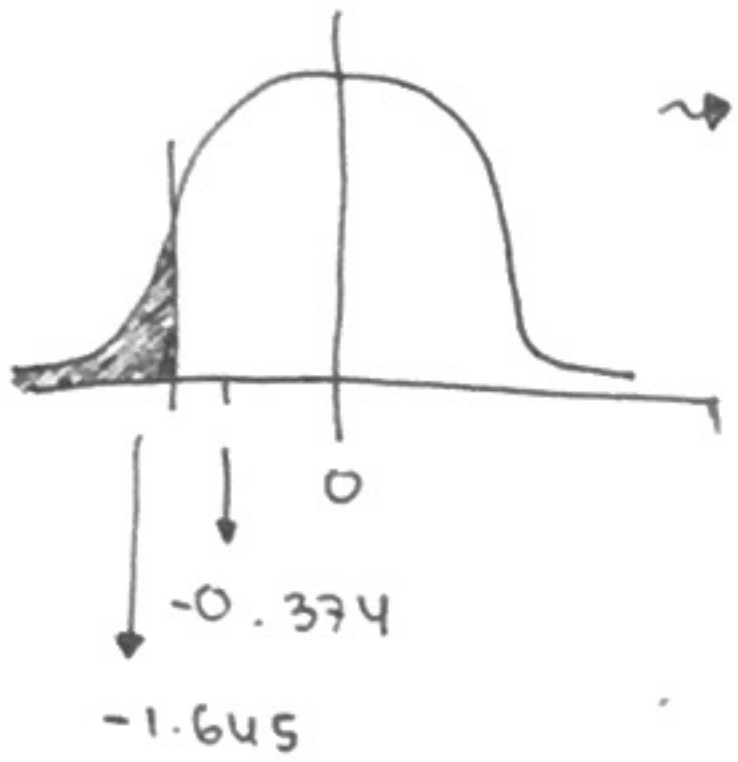
→ statistic test

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}} = \frac{0.094 - 0.10}{\sqrt{\frac{(0.10)(0.90)}{350}}} = \frac{-0.006}{\sqrt{\frac{0.09}{350}}} = \frac{-0.006}{0.0160357}$$

-proportion-one population

$= -0.374166$

$Z = -0.374$ //



→ critical value $Z_\alpha \rightarrow Z_{0.05} = -1.645$

→ $-0.374 > -1.645$, we fail to reject H_0. we do not have sufficient evidence to claim that the incorrect test result have to be less than 10% or 0.10 //.

ⓑ. $n = 16$

- $s = 652$ psi
- $df = 16 - 1$

  $= 15$

- $\bar{x} = 58400$ psi
- 99% significance level
- $\alpha = 0.01$

  $\rightarrow \frac{\alpha}{2} = 0.05$

- $H_0 : \mu = 58000$
- $H_1 : \mu \neq 58.000$

→ statistic test

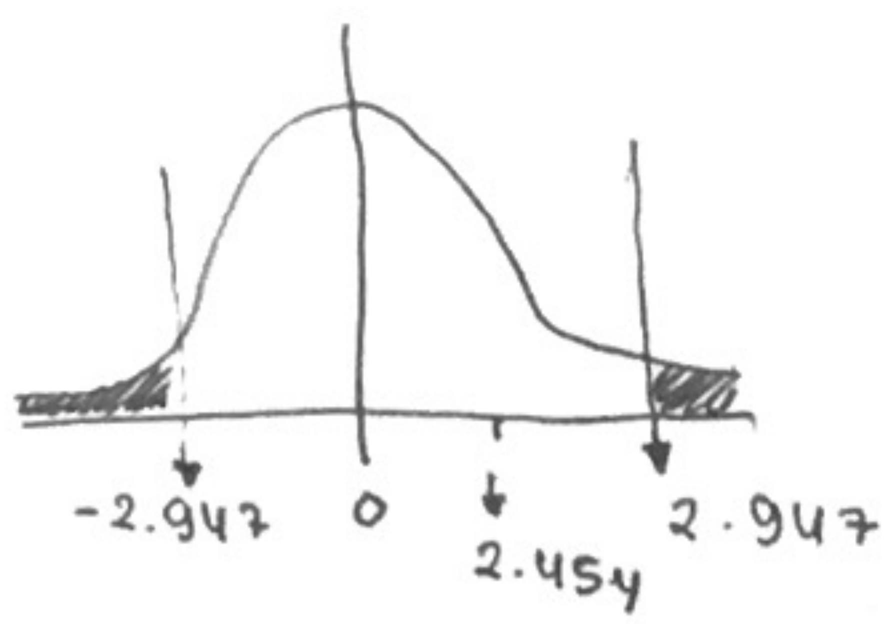
$$t = \frac{\bar{x} - \mu}{(s/\sqrt{n})} = \frac{58400 - 58000}{(652/\sqrt{16})}$$

$$= \frac{400}{\left(\frac{652}{4}\right)} = 2.45399$$

$= 2.454$

→ Critical value $\pm t_\alpha \rightarrow t_{0.005, df=15} = 2.947$

$-t_{0.005, 15} = -2.947$



→ $2.454 < 2.947$, we fail to reject null Hypothesis (H_0). there is sufficient evidence to conclude that the true average compressive strength of steel is 58000 psi.

Question 3

ⓐ $n = 10 \rightsquigarrow n < 30$ use t test.

$\alpha = 0.01$.

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

$$= \frac{1006}{10} = 10.06$$

- $H_0 : \mu = 10$
- $H_1 : \mu \neq 10$

$$\rightsquigarrow s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}}$$

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

$$\sqrt{\frac{\ldots + (9.8-10.06)^2 + (10.3-10.06)^2 + (10.0-10.06)^2 + (10.4-10.06)^2)}{10-1}}$$

$$= \sqrt{\frac{0.0576 + 0.0256 + 0.0196 + 0.0016 + 0.1296 + 0.0256 + 0.0676 + 0.0576 + 0.0036}{9}}$$

$$\sqrt{\frac{\ldots + 0.1156}{9}} = \sqrt{\frac{0.504}{9}}$$

$$= \sqrt{0.056} = 0.237$$

$$\rightsquigarrow t = \frac{\bar{x} - \mu}{(s/\sqrt{n})} = \frac{10.06 - 10}{\left(\frac{0.237}{\sqrt{10}}\right)} = \frac{0.06}{\frac{\frac{237}{1000}}{\sqrt{10}}} = \frac{20\sqrt{10}}{79} = 0.800577$$

$$= 0.801$$

$\rightsquigarrow$ c.v. $t_{0.005, df=9} = 3.250$

$\rightsquigarrow$ since $t = 0.801 < t_{0.005,9} = 3.250$; we fail to reject H_0.

there is sufficient evidence to conclude that the average content of Brand x car lubricates is 10 liters.

(6) • $n = 14$ (mothers taking special vitamins)

• $\sigma = 0.44$ kg

• $\alpha = 0.05 = 5\%$

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

$$\frac{\ldots + 2.36 + 3.35)}{14} = \frac{51.16}{14} = \frac{\frac{1279}{25}}{14}$$

$$= \frac{1279}{350} = 3.654$$

• $$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$s = \sqrt{\frac{(3.70 - 3.654)^2 + (4.31-3.654)^2 + (3.73-3.654)^2 + (4.33-3.654)^2 + (3.33-3.654)^2 + (2.58-3.654)^2 + (3.35-3.654)^2 + (4.447-3.654)^2 + (3.55-3.654)^2 + (4.66-3.654)^2 + (3.68-3.654)^2 + (3.02-3.654)^2 + (4.09-3.654)^2 + (2.36-3.654)^2}{14-1}}$$

$$s = \sqrt{\frac{0.002116 + 0.430336 + 0.005776 + 0.456976 + 0.104976 + 1.153476 + 0.092416 + 0.628849 + 0.010816 + 1.012036 + 0.000676 + 0.401956 + 0.190096 + 1.674436}{13}}$$

$s = 0.691$

• $H_0 = \sigma = 0.44$ kg

• $H_1 = \sigma \neq 0.44$ kg

→ $$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(14-1)\ (0.691)^2}{(0.44)^2} = 32.06226$$

$$= 32.062$$

$X^2_{0.025}$. $df = 13 = 24.736$

$X^2_{1-\frac{\alpha}{2}} = X^2_{0.975,13} = 5.009$

→ since $X^2 = 32.062 > X^2_{0.025,13} = 24.736$; we reject Null Hypothesis (H_0)

→ there is sufficient evidence to conclude that the vitamin suplement has any effects on the standard deviation of the bith weight.

Q4

$n = 101$

$s^2 = 0.165$

$\alpha = 1 - 0.95$
$= 0.05$

$df = 101 - 1$
$= 100$

$H_0 : \sigma^2 = 0.18$

$H_1 : \sigma^2 \neq 0.18$

Critical value :

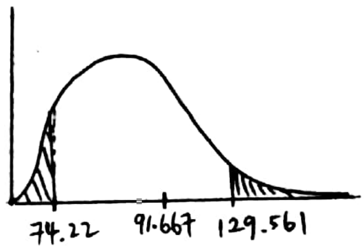
$\chi^2_{\frac{\alpha}{2}, 100} = \chi^2_{0.025, 100}$
$= 129.561$

$\chi^2_{1-\frac{\alpha}{2}, 100} = \chi^2_{0.975, 100}$
$= 74.222$

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

$= \dfrac{(101-1)\,0.165}{0.18}$

$= 91.667$



$\therefore$ Since $91.667 < 129.561$, fail to reject H_0. There is sufficient evidence to conclude that the new joint inspector is making satisfactory measurements where variance is $0.18 (\text{inch})^2$.

Q5

18-34 age group

$n_1 = 15$

$\bar{x}_1 = 76.4$

$s_1^2 = 25.30$

35+ age group

$n_2 = 18$

$\bar{x}_2 = 71.2$

$s_2^2 = 22.20$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$\alpha = 1 - 0.95 = 0.05$

$\frac{\alpha}{2} = 0.025$

Degree of freedom : $V = \dfrac{\left(\frac{25.30}{15} + \frac{22.20}{18}\right)^2}{\frac{\left(\frac{25.30}{15}\right)^2}{14} + \frac{\left(\frac{22.20}{18}\right)^2}{17}} = 29.132 \approx 29$

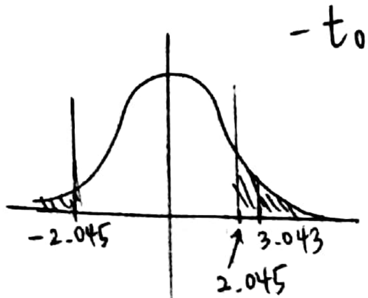
Critical value : $t_{0.025, 29} = 2.045$

$-t_{0.025, 29} = -2.045$

Test statistics :

$t = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

$= \dfrac{76.4 - 71.2}{\sqrt{\frac{25.30}{15} + \frac{22.20}{18}}} = 3.043$



$\therefore$ 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 in 95% significance level.

Q6

Batch 1 : $n_1 = 16$
$\bar{x}_1 = 93.6$
$s_1 = 3.6$
$s_1^2 = 12.96$

Batch 2 : $n_2 = 21$
$\bar{x}_2 = 98.5$
$s_2 = 2.5$
$s_2^2 = 6.25$

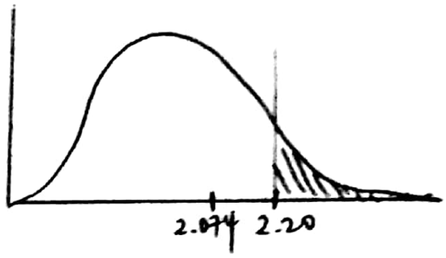
$\alpha = 1 - 0.95 = 0.05$

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

Test statistics : $F = \dfrac{s_1^2}{s_2^2} = \dfrac{12.96}{6.25} = 2.074$

df : Numerator $= n_1 - 1 = 16 - 1 = 15$
Denominator $= n_2 - 1 = 21 - 1 = 20$

Critical value : $F_{0.05, 15, 20} = 2.20$



$\therefore$ Since $2.074 < 2.20$, we fail to reject H_0.
There is insufficient evidence to conclude that standard deviation of processors from batch 1 is greater than standard deviation of processors of batch 2.

Q7

$n = 10$, $\alpha = 0.05$
$H_0 : \mu_D = 0$
$H_1 : \mu_D < 0$

$df = 10 - 1 = 9$

critical value : $-t_{0.05, 9} = -1.833$

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
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

$\Sigma D = -207$ $\Sigma D^2 = 5131$

Test statistics :

$\bar{D} = \dfrac{\Sigma D}{n} = \dfrac{-207}{10} = -20.7$

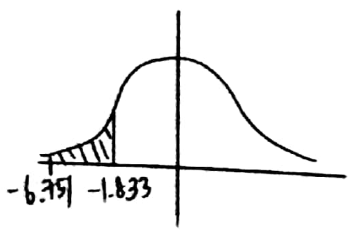
$s_D = \sqrt{\dfrac{5131 - \frac{(-207)^2}{10}}{10 - 1}}$

$= 9.696$

$t = \dfrac{\bar{D} - \mu_D}{\left(\frac{s_D}{\sqrt{n}}\right)}$

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

$= -6.751$



$\therefore$ Since $-6.751 < -1.833$, we reject H_0. There is sufficient evidence to conclude that training can increase the number of words spelled correctly.

Scanned with CamScanner

Question 8

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
$\Sigma X = 14.6$	$\Sigma Y = 26$	$\Sigma XY = 56.8$	$\Sigma Y^2 = 104$	$\Sigma X^2 = 32.9632$

i) $r = \dfrac{56.8 - (14.6)(26)/8}{\sqrt{[32.9632 - (14.6)^2/8][104 - 26^2/8]}} = 0.8424$

ii) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

$\alpha = 0.05$

$df = 8-2 = 6$

$t = \dfrac{0.8424}{\sqrt{\dfrac{1-0.8424^2}{8-2}}} = 3.829$

$t_{0.025, 6} = 2.447$

Since $t = 3.829 > t_{0.025,6} = 2.447$, we reject H_0. There is sufficient evidence of a linear relationship between weight of plastic usage and size of household at 5% level of significance.

iii) $\alpha = 0.01$ $\quad t_{0.005, 6} = 3.707$

Since $t = 3.829 > t_{0.005,6} = 3.707$, we reject H_0. There is sufficient evidence of a linear relationship between weight of plastic usage and size of household at 1% level of significance.

Question 9

X	Y	$(X-\bar{X})$	$(Y-\bar{Y})$	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$
97	24	-22.5556	0	0	508.7551
85	29	-34.5556	5	-172.778	1194.0895
98	26	-21.5556	2	-43.1112	464.6439
105	24	-14.5556	0	0	211.8655
120	24	0.4444	0	0	0.1975
151	22	31.4444	-2	-62.8888	988.7503
140	23	20.4444	-1	-20.4444	417.9735
134	23	14.4444	-1	-14.4444	208.6407
146	21	26.4444	-3	-79.3332	699.3063
$\Sigma X=1076$	$\Sigma Y=216$			$\Sigma(X-\bar{X})(Y-\bar{Y})=-393$	$\Sigma(X-\bar{X})^2=4694.2223$

a) $\bar{X} = \frac{1076}{9} = 119.556$

$b_1 = \frac{-393}{4694.2223} = -0.0837$

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

$b_0 = 24 - (-0.0837)(119.556) = 34.007$

$\hat{y} = 34.007 - 0.0837x$

b) $\hat{y} = 34.007 - 0.0837(125) = 23.5445 \approx 23.545$

Scanned with CamScanner

-Difference

Question 10

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

H_1: at least one mean is different

$n=5$

$\alpha = 0.01$

Memory booster

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

$$s = \sqrt{\frac{(70-83.4)^2 + (77-83.4)^2 + (83-83.4)^2 + (90-83.4)^2 + (97-83.4)^2}{5-1}} = 10.5972$$

Placebo

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

$$s = \sqrt{\frac{(37-50)^2 + (43-50)^2 + (50-50)^2 + (57-50)^2 + (63-50)^2}{5-1}} = 10.4403$$

Without Treatment

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

$$s = \sqrt{\frac{(3-16.6)^2 + (10-16.6)^2 + (17-16.6)^2 + (23-16.6)^2 + (30-16.6)^2}{5-1}} = 10.5972$$

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

$$s_{\bar{x}} = \sqrt{\frac{(83.4-50)^2 + (50-50)^2 + (16.6-50)^2}{3-1}} = 33.4$$

$$n s_{\bar{x}}^2 = 5(33.4)^2 = 5577.8$$

$$s_p^2 = \frac{10.5972^2 + 10.4403^2 + 10.5972^2}{3} = 111.2004$$

$$F = \frac{5577.8}{111.2004} = 50.16$$

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

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

$F_{0.01,2,12} = 6.93$

Since $F = 50.16 > F_{0.01,2,12} = 6.93$, we reject H_0. There is sufficient evidence to claim that the treatments will have different affects.