- LEE TONG MING A19EC0069
- SAHARULNIZAM HAKIMY BIN SHOBRI A19EC0205
- AHMAD ZULHAKIM BIN ZAINAL A19EC0179

Ref: Date:

PSDA (Assignment 5)

Question 1

(a) Let μ_1 = mean birth weights for zinc supplement group;
μ_2 = mean birth weights for placebo group

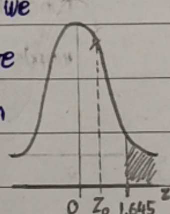
$n_1=10$, $n_2=15$ $\quad H_0: \mu_1=\mu_2$; $H_1: \mu_1>\mu_2$

$\bar{x}_1=3214$, $\bar{x}_2=3088$ $\quad$ Test statistics, $z_0=\frac{\bar{x}_1-\bar{x}_2-0}{\sqrt{s_1^2/n_1+s_2^2/n_2}}$

$s_1=669$, $s_2=728$ $\quad z_0=\frac{3214-3088}{\sqrt{\frac{669^2}{10}+\frac{728^2}{15}}}=0.4452$

$\alpha=0.05$, $z_\alpha=z_{0.05}=1.645$

∴ Since test statistics $z_0=0.4452<z_\alpha=1.645$, we fail to reject null hypothesis $H_0: \mu_1=\mu_2$. There is insufficient evidence to support the claim that variation of birth weights for placebo population is less than the variation for population treated with zinc supplements.



(b) $n=5$, $\alpha=0.01$, $df=n-1=5-1=4$, $t_{0.01,4}=3.747$

$\Sigma D=274+83+29+236+121=743$, $\bar{D}=\frac{\Sigma D}{N}=\frac{743}{5}=148.6$

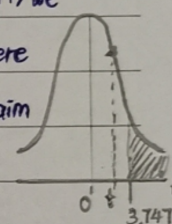
$\Sigma D^2=75076+6889+841+55696+14641=153143$

$H_0: \mu_D=0$; $H_1: \mu_D>0$

$S_D=\sqrt{\frac{\Sigma D^2-(\Sigma D)^2/n}{n-1}}=\sqrt{\frac{153143-(743)^2/5}{5-1}}=\sqrt{\frac{42733.2}{4}}=103.3601$

test statistics, $t=\frac{\bar{D}-\mu_D}{S_D/\sqrt{n}}=\frac{148.6-0}{103.3601/\sqrt{5}}=3.2148$

∴ Since test statistics, $t=3.2148<t_{c.v.}=3.747$, we fail to reject null hypothesis $H_0: \mu_D=0$. There is insufficient evidence to support the claim that the independent workshop has lower repair costs.



Question 2

(a) Let p_1 be population proportion of babies given the medicine and recovered within 2 days;
p_2 be population proportion of babies not given medicine and recovered within 2 days

	Group 1	Group 2
	Given medicine	Not given medicine
Recovered in 2 days	$x_1=29$	$x_2=56$
Babies who had flu	$n_1=120$	$n_2=280$

$\hat{p}_1=\frac{x_1}{n_1}=\frac{29}{120}$; $\hat{p}_2=\frac{x_2}{n_2}=\frac{56}{280}=\frac{1}{5}$, $\alpha=0.01$

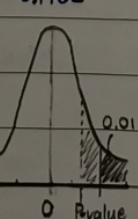
$\bar{p}=\frac{x_1+x_2}{n_1+n_2}=\frac{29+56}{120+280}=0.2125$, $\bar{q}=1-\bar{p}=1-0.2125=0.7875$

$H_0: p_1=p_2$; $H_1: p_1>p_2$

Test statistics, $z=\frac{(\hat{p}_1-\hat{p}_2)(p_1-p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1}+\frac{\bar{p}\bar{q}}{n_2}}}=\frac{\left(\frac{29}{120}-\frac{1}{5}\right)-0}{\sqrt{\frac{(0.2125)(0.7875)}{120}+\frac{(0.2125)(0.7875)}{280}}}$

$z=0.9335\approx0.93$, P-value: $P(z>0.93)=1-0.8238=0.1762$

∴ Since P-value $=0.1762>\alpha=0.01$, we fail to reject null hypothesis $H_0: p_1=p_2$. There is insufficient indication that supports the claim of the effectiveness of the flu medicine.



(b) $\bar{x}_1=\frac{0.48+0.39+0.42+0.52+0.40}{5}=\frac{2.21}{5}=0.442$

$S_1=\sqrt{\frac{(0.48-0.442)^2+(0.39-0.442)^2+(0.42-0.442)^2+(0.52-0.442)^2+(0.4-0.442)^2}{5-1}}=0.0559$

$\bar{x}_2=\frac{0.38+0.37+0.39+0.41+0.38}{5}=\frac{1.93}{5}=0.386$

$S_2=\sqrt{\frac{(0.38-0.386)^2+(0.37-0.386)^2+(0.39-0.386)^2+(0.41-0.386)^2+(0.38-0.386)^2}{5-1}}=0.0152$

$H_0: \mu_1=\mu_2$; $H_1: \mu_1\neq\mu_2$, $\alpha=0.05$, $\frac{\alpha}{2}=\frac{0.05}{2}=0.025$

Upper tail critical value, $t_{\alpha/2,n_1+n_2-2}=t_{0.025,8}=2.306$

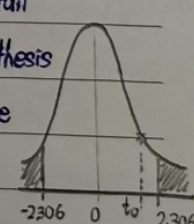
Lower tail critical value, $-t_{\alpha/2,n_1+n_2-2}=-t_{0.025,8}=-2.306$

Pool estimator, $S_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

$S_p^2=\frac{(5-1)(0.0559)^2+(5-1)(0.0152)^2}{5+5-2}=1.6779\times10^{-3}$, $S_p=\sqrt{1.6779\times10^{-3}}=0.041$

Test statistics, $t_0=\frac{\bar{x}_1-\bar{x}_2-0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{0.442-0.386-0}{0.041\sqrt{\frac{1}{5}+\frac{1}{5}}}=2.1596$

∴ Since test statistics $t_0=2.1596<$ upper tail $t_{c.v.}=2.306$, we fail to reject null hypothesis $H_0: \mu_1=\mu_2$. There is insufficient evidence to support the claim that production line 1 is not as consistent as production line 2 in terms of alcohol content.



(c) $H_0: \mu_D = 0$; $H_1: \mu_D < 0$

$\alpha = 0.05$, $n = 10$, $df = 10-1 = 9$, $t_{c.v.} = -t_{0.05,9} = -1.833$

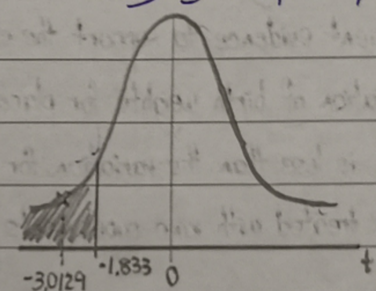
Before(x_1)	After(x_2)	$D = x_1 - x_2$	$D^2 = (x_1-x_2)^2$
204	223	-19	361
393	412	-19	361
391	402	-11	121
265	285	-20	400
326	353	-27	729
220	243	-23	529
423	443	-20	400
342	340	2	4
480	582	-102	10404
464	490	-26	676
	Total	$\Sigma D = -265$	$\Sigma D^2 = 13985$

$$\bar{D} = \frac{\Sigma D}{n} = \frac{-265}{10} = -26.5$$

$$S_D = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n-1}} = \sqrt{\frac{13985 - \frac{(-265)^2}{10}}{10-1}} = \sqrt{\frac{6962.5}{9}} = 27.8139$$

$$\text{Test statistics, } t = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} = \frac{-26.5-0}{\frac{27.8139}{\sqrt{10}}} = -3.0129$$

$\therefore$ Since test statistics $t = -3.0129 < t_{c.v.} = -1.833$, we reject null hypothesis $H_0: \mu_D = 0$. There is sufficient evidence to conclude that training can increase the number of words spelled correctly by the participants.



Question 3

(a) Let Method A be first population,
Method B be second population.

$n_1 = 10$, $n_2 = 10$, $\alpha = 0.05$, $df = n_1 + n_2 - 2 = 10 + 10 - 2 = 18$

$$\bar{x}_1 = \frac{71+75+65+69+73+66+68+71+74+68}{10} = \frac{700}{10} = 70$$

$$S_1 = \sqrt{\frac{(65-70)^2+(66-70)^2+2(68-70)^2+(69-70)^2+2(71-70)^2+(73-70)^2+(74-70)^2+(75-70)^2}{10-1}} = \sqrt{\frac{102}{9}} = 3.3665$$

$$\bar{x}_2 = \frac{72+77+84+78+69+70+77+73+65+75}{10} = \frac{740}{10} = 74$$

$$S_2 = \sqrt{\frac{(65-74)^2+(69-74)^2+(70-74)^2+(72-74)^2+(73-74)^2+(75-74)^2+2(77-74)^2+(78-74)^2+(84-74)^2}{10-1}} = \sqrt{\frac{262}{9}} = 5.3955$$

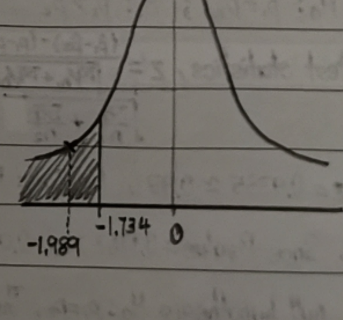
$H_0: \mu_1 = \mu_2$; $H_1: \mu_1 < \mu_2$, critical value $t_{c.v.} = -t_{0.05,18} = -1.734$

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(10-1)(3.3665)^2 + (10-1)(5.3955)^2}{10+10-2} = 20.2224$$

$$S_p = \sqrt{S_p^2} = \sqrt{20.2224} = 4.4969$$

$$\text{Test statistics, } t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{70-74-0}{4.4969\sqrt{\frac{1}{10}+\frac{1}{10}}} = -1.989$$

$\therefore$ Since test statistics $t_0 = -1.989 < t_{c.v.} = -1.734$, we reject null hypothesis $H_0: \mu_1 = \mu_2$. This shows that mean score for Method B is higher, there is sufficient evidence to support the claim that Method B is more effective.



(b) $H_0: \mu_D = 0$; $H_1: \mu_D > 0$

$n=12$, $df = 12-1 = 11$, $\alpha = 0.01$, $t_{c.v.} = t_{0.01,11} = 2.718$

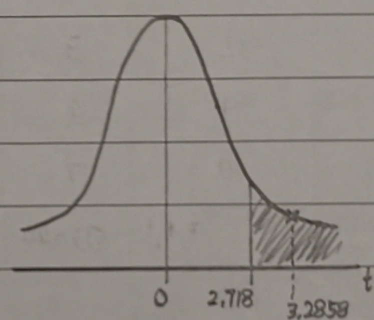
Before(x_1)	After(x_2)	$D = x_1 - x_2$	$D^2 = (x_1 - x_2)^2$
95	89	6	36
81	77	4	16
76	77	-1	1
96	94	2	4
82	80	2	4
87	86	1	1
71	72	-1	1
82	82	0	0
77	74	3	9
83	81	2	4
81	78	3	9
65	63	2	4
	Total	$\sum D = 23$	$\sum D^2 = 89$

$$\bar{D} = \frac{\sum D}{n} = \frac{23}{12} = 1.9167$$

$$S_D = \sqrt{\frac{\sum D^2 - (\sum D)^2/n}{n-1}} = \sqrt{\frac{89 - (23)^2/12}{12-1}} = 2.0207$$

Test statistics, $t = \dfrac{\bar{D} - \mu_0}{S_D/\sqrt{n}} = \dfrac{1.9167 - 0}{2.0207/\sqrt{12}} = 3.2858$

$\therefore$ Since test statistics $t = 3.2858 > t_{c.v.} = 2.718$, we reject null hypothesis $H_0: \mu_D = 0$. This means that the weight before is larger than weight after, there is sufficient evidence to conclude that the prescribed program of exercise is effective in weight reduction.



Question 4

(a) Let μ_1 be mean from (18-34) age group

μ_2 be mean from 35+ age group

$n_1 = 150$, $n_2 = 250$, $\bar{x}_1 = 75.5$, $\bar{x}_2 = 65.2$, $S_1 = 53.2$, $S_2 = 46.1$

$H_0: \mu_1 = \mu_2$; $H_1: \mu_1 \neq \mu_2$

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(\frac{53.2^2}{150} + \frac{46.1^2}{250}\right)^2}{\frac{(53.2^2/150)^2}{149} + \frac{(46.1^2/250)^2}{249}} = 279.549 \approx 280$$

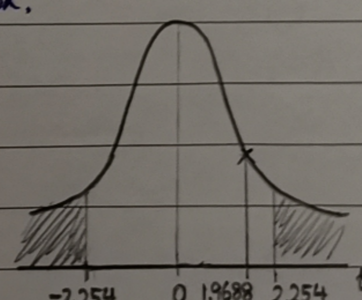
$\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$

Upper tail $t_{c.v.} = t_{0.025,280} = 2.254$

Lower tail $t_{c.v.} = -t_{0.025,280} = -2.254$

Test statistics, $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \dfrac{75.5 - 65.2 - 0}{\sqrt{\frac{53.2^2}{150} + \frac{46.1^2}{250}}} = 1.9688$

$\therefore$ Since test statistics $t_0 = 1.9688 <$ upper tail $t_{c.v.} = 2.254$, we fail to reject null hypothesis, $H_0: \mu_1 = \mu_2$. There is sufficient evidence to support the hypothesis that there is no difference in the mean spending between the two population.



(b) $H_0: \mu_D = 3$; $H_1: \mu_D > 3$

$n=10$, $df = 10-1 = 9$, $\alpha = 0.05$, $t_{c.v.} = t_{0.05,9} = 1.833$

1 hour later (x_1)	24 hour later (x_2)	$D = x_1 - x_2$	$D^2 = (x_1-x_2)^2$
14	10	4	16
12	4	8	64
18	14	4	16
7	5	2	4
11	8	3	9
9	5	4	16
16	11	5	25
15	12	3	9
13	9	4	16
17	10	7	49
	Total	$\Sigma D = 44$	$\Sigma D^2 = 224$

$$\bar{D} = \frac{\Sigma D}{n} = \frac{44}{10} = 4.4$$

$$S_D = \sqrt{\frac{\Sigma D^2 - (\Sigma D)^2/n}{n-1}} = \sqrt{\frac{224 - (44)^2/10}{10-1}} = \sqrt{\frac{30.4}{9}} = 1.8379$$

$$\text{Test statistics, } t = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} = \frac{4.4-3}{1.8379/\sqrt{10}} = 2.4088$$

$\therefore$ Since test statistics, $t = 2.4088 > t_{c.v.} = 1.833$, we reject null hypothesis $H_0: \mu_D = 3$. There is sufficient evidence to support that the mean number of words recalled after 1 hour exceeds the mean recall after 24 hours by more than 3, $\mu_D > 3$.