SSN COLLEGE OF ENGINEERING
RECORD SHEET

Sheet No. 1

ASSIGNMENT

DHANYA SRINIVASAN
205001029

1. $n_1 = 6400$ $\quad \bar{x}_1 = 170$ $\quad S_1 = 6.4$
   $n_2 = 1600$ $\quad \bar{x}_2 = 172$ $\quad S_2 = 6.3$

   H_0: There is no significant difference between the average height of English and American men $(\bar{x}_1 = \bar{x}_2)$

   H_1: There is significant difference between the average height of English and American men, i.e, on an average, Americans are taller than English men $(\bar{x}_1 < \bar{x}_2)$

   $\Rightarrow$ Single tailed test (left tailed test)

   Level of significance = 5%

   $Z_\alpha = 1.645$

   $$Z = \left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \right| = \left| \frac{170 - 172}{\sqrt{\frac{6.4^2}{6400} + \frac{6.3^2}{1600}}} \right| = \frac{2}{\sqrt{0.0064 + 0.025}} = \frac{2}{0.1772}$$

   $= 11.287$

   $|Z| > Z_\alpha \Rightarrow$ Reject H_0
   Accept H_1

   Hence, on an average, Americans are taller than English men.

2. $n = 500$ $\quad p = \frac{65}{500} = \frac{13}{100} = 0.13$

   $q = 1 - 0.13 = 0.87$

   <u>To show</u>: $0.085 \leq P \leq 0.175$

   Confidence limits is given by : $p \pm z_\alpha \sqrt{\frac{pq}{n}}$

   $S.E = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.13 \times 0.87}{500}} = \sqrt{0.0002262} = 0.015$

   $Z \leq 3 \Rightarrow Z_{max} = 3$

   Confidence limits $= 0.13 \pm 3(0.015) = 0.13 \pm 0.045$

   $\Rightarrow 0.085 \leq P \leq 0.175$

3. $n=10$, $\mu=100$

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84
972		1833.6

$\bar{x} = \frac{(70+120+110+101+88+83+95+98+107+100)}{10}$

$= \frac{972}{10} = 97.2$

$S^2 = \frac{\Sigma(x_i-\bar{x})^2}{n-1} = \frac{1833.6}{9} = 203.733$

$t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} = \frac{97.2-100}{\frac{\sqrt{203.733}}{\sqrt{10}}} = \frac{-2.8}{4.5137} = -0.6203$

$|t| = 0.6203$

$t_\alpha = 2.262$

$|t| < t_\alpha$ ~~Reject H_0,~~

Accept H_0

$\Rightarrow$ The population mean $\mu = 100$

H_0: The population mean $\mu=100$

H_1: The population mean $\mu \neq 100$

$\Rightarrow$ Two tailed test

Level of significance = ~~2.262~~ 5%

dof $= n-1 = 9$

Confidence limits: $\bar{x} \pm t_{0.05} \frac{S}{\sqrt{n}}$

$= 97.2 \pm \frac{2.262 \times 14.2}{\sqrt{10}} = 97.2 \pm 10.2$

$= 87$ & 107.4

4.

x_1	x_2	$x_1-\bar{x}_1$	$x_2-\bar{x}_2$	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$
25	44	-3.08	14	9.4864	196
32	34	3.92	4	15.3664	16
30	22	1.92	-8	3.6864	64
34	10	5.92	-20	35.0464	400
24	47	-4.08	17	16.6464	289
15	31	-13.08	1	171.0864	1
32	40	3.92	10	15.3664	100
24	30	-4.08	0	16.6464	0
30	32	1.92	2	3.6864	4
31	35	2.92	5	8.5264	25
35	18	6.92	-12	47.8864	144
25	21	-3.08	-9	9.4864	81
	35		5		25
	29		-1		1
	22		-8		64
337	450			352.9168	1410

$n_1 = 12$

$n_2 = 15$

$\bar{x}_1 = \frac{337}{12} = 28.0833$

$\bar{x}_2 = \frac{450}{15} = 30$

H_0: There is no significant difference between the effects on increase in weight between the two diets ($\mu_1 = \mu_2$)

H_1: There is significant difference between the effects on increase in weight between the two diets ($\mu_1 \neq \mu_2$)

$\Rightarrow$ Two tailed test

Level of significance = 5%.

$t_\alpha = 2.060$

$$S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} = \frac{12\times\frac{352.9168}{12} + 15\times\frac{1410}{15}}{12+15-2} = \frac{1762.9168}{25} = 70.518$$

$S = 8.3975$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{28.08-30}{8.3975\sqrt{\frac{1}{12}+\frac{1}{15}}} = \frac{-1.92}{8.3975\times\sqrt{0.15}} = \frac{-1.92}{8.3975\times 0.387}$$

$$= \frac{-1.92}{3.2523} = -0.5903$$

$|t| = 0.5903$

$|t| < t_{\cancel{tab}\,\alpha}$ Accept H_0
Reject H_1

Hence, there is no significant difference between the effects on increase in weight due to the two diets.

5. $n_1 = 8$, $n_2 = 7$

H_0: There is no significant difference between the ~~two~~ two estimates of population variance

H_1: There is significant difference between the two estimates of population variance

x_i	x_i^2	y_i	y_i^2
9	81	10	100
11	121	12	144
13	169	10	100
11	121	14	196
15	225	9	81
9	81	8	64
12	144	10	100
14	196		
94	1138	73	785

$\bar{x} = \frac{94}{8} = 11.75$ $\quad \bar{y} = \frac{\cancel{785}\,73}{7} = 10.429$

$\Sigma x_i^2 = 1138$ $\quad \Sigma y_i^2 = 785$

$\frac{\Sigma x_i^2}{n_1} = \frac{1138}{8} = 142.25$ $\quad \frac{\Sigma y_i^2}{n_2} = \frac{785}{7} = 112.143$

$s_1^2 = 142.25 - (11.75)^2$
$= 142.25 - 138.0625$
$= 4.1875$

$s_2^2 = 112.143 - (10.429)^2$
$= 112.143 - 108.764$
$= 3.379$

$S_1^2 = \frac{n_1}{n_1-1}(s_1^2) = \frac{8}{7}(4.1875)$
$= 4.786$

$S_2^2 = \frac{n_2}{n_2-1}(s_2^2) = \frac{7}{6}(3.379)$
$= 3.942$

$F = \frac{S_1^2}{S_2^2} = \frac{4.786}{3942} = 1.214$

$F_{tab}(7,6) = 4.21$

$|F| < F_{tab}$ Accept H_0
Reject H_1

The two populations do not differ significantly at 5% level of significance

6. Expected ratio : 9:3:3:1 $n = 1600$

H_0: The experimental results support the theory

H_1: The experimental results do not support the theory.

$E_i(A) = \frac{9}{16} \times 1600 = 900$

$E_i(B) = \frac{3}{16} \times 1600 = 300$

$E_i(C) = \frac{3}{16} \times 1600 = 300$

$E_i(D) = \frac{1}{16} \times 1600 = 100$

Group	A	B	C	D	Total
O_i	882	313	287	118	1600
E_i	900	300	300	100	1600
$O_i - E_i$	−18	13	−13	18	
$(O_i-E_i)^2$	324	169	169	324	
$\frac{(O_i-E_i)^2}{E_i}$	0.36	0.56	0.56	3.24	4.72

$\chi^2 = \Sigma\left(\frac{(O_i-E_i)^2}{E_i}\right) = 4.72$

$\nu = 3$

$\chi^2_{tab} = 7.815$ at 5% significance

$|\chi^2| < \chi^2_{tab} \Rightarrow$ Accept H_0
Reject H_1

Hence, the experimental results support the theory.

7.

X	f
0	5
1	18
2	28
3	12
4	7
5	6
6	4
	80

H_0: Binomial distubution is a good fit
H_1: Binomial distubution is not a good fit

$$\bar{X} = \frac{\sum f_i x_i}{N} = \frac{0+18+56+36+28+30+24}{80} = \frac{192}{80} = 2.4$$

$$P = \frac{\bar{X}}{n} = \frac{2.4}{6} = 0.4 \Rightarrow q = 1-p = 1-0.4 = 0.6$$

$$P(x) = {}^nC_x\, p^x q^{n-x} = {}^6C_x (0.4)^x (0.6)^{6-x}$$

$N(0) = 80P(X=0) = 80 \times {}^6C_0 (0.6)^6 = 3.73248 \simeq 4$
$N(1) = 80P(X=1) = 80 \times {}^6C_1 (0.4)(0.6)^5 = 14.92992 \simeq 15$
$N(2) = 80P(X=2) = 80 \times {}^6C_2 (0.4)^2 (0.6)^4 = 24.8832 \simeq 25$
$N(3) = 80P(X=3) = 80 \times {}^6C_3 (0.4)^3 (0.6)^3 = 22.1184 \simeq 22$
$N(4) = 80P(X=4) = 80 \times {}^6C_4 (0.4)^4 (0.6)^2 = 11.0592 \simeq 11$
$N(5) = 80P(X=5) = 80 \times {}^6C_5 (0.4)^5 (0.6) = 2.94912 \simeq 3$
$N(6) = 80P(X=6) = 80 \times {}^6C_6 (0.4)^6 = 0.32768 \simeq 0$

Grouping $X=0$ and $X=5$ and $X=6$ together.

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
15	7	8	64	9.14
18	15	3	9	0.6
28	25	3	9	0.36
12	22	-10	100	4.55
7	11	-4	16	1.45

$\nu = n-1$
$5-1=4$

$\chi^2 = 9.14 + 0.6 + 0.36 + 4.55 + 1.45 = 16.1$

$\chi^2_{tab} = 9.488$ $\quad \chi^2 > \chi^2_{tab} \Rightarrow$ Reject H_0
Accept H_1

Hence, the binomial distubution is not a good fit

8.

Researchers	Below avg	Average	Above avg	Genius	Total
X	86	60	44	10	200
Y	40	33	25	2	100
Total	126	93	69	12	300

H_0: Data obtained are independent of sampling techniques adopted by the two researchers

H_1: Data obtained are not independent of the sampling techniques adopted by the two researchers

Expected frequency

Researchers	Below avg	Average	Above avg	Genius	Total
X	$\frac{200 \times 126}{300} = 84$	$\frac{200 \times 93}{300} = 62$	$\frac{200 \times 69}{300} = 46$	$\frac{200 \times 12}{300} = 8$	200
Y	$\frac{100 \times 126}{300} = 42$	$\frac{100 \times 93}{300} = 31$	$\frac{100 \times 69}{300} = 23$	$\frac{100 \times 12}{300} = 4$	100
Total	126	93	69	12	300

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
86	84	2	4	0.047619
60	62	-2	4	0.064516
44	46	-2	4	0.086957
10	8	2	4	0.5
40	42	-2	4	0.095238
33	31	2	4	0.129032
25	23	2	4	0.173913
2	4	-2	4	1

$\nu = (2-1)(4-1)$
$= 3$

$$\chi^2 = \Sigma\left(\frac{(O_i - E_i)^2}{E_i}\right) = 2.097275$$

$\chi^2_{tab} = \cancel{5.991}\ 7.815$

$\chi^2 < \chi^2_{tab} \Rightarrow$ Accept H_0
Reject H_1

Hence the data obtained are independent of sampling techniques adopted by the two researchers