1a) i) Hypothesis : $H_0 : \mu \geq 30$

$H_A : \mu < 30$

ii) Test statistic : $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{21-30}{11/7} = -5.727$

iii) $\alpha = .05$

iv) Critical value ; $z_\alpha = -1.645$ (one tailed test)

v) Decision : Reject since $-5.727 < -1.645$

We conclude that the pop. mean is < 30

b) i) $P(M \cap P) = .07$ ii) $P(H/S) = \frac{P(H \cap S)}{P(S)} = \frac{300}{500} = .6$

iii) $P(L) = .22$

iv) $P(H \cup S) = P(H) + P(S) - P(H \cap S) =$

$\frac{390}{1000} + \frac{500}{1000} - \frac{300}{1000} = .59$

c) Schaum series - pg. 343 (1 way ANOVA)

d)

x	f	cum.f.	$d = x - \bar{x}$	d^2	fd^2
6	3	3	-3	9	27
7	6	9	-2	4	24
8	9	18	-1	1	9
9	13	31	0	0	0
10	8	39	1	1	8
11	5	44	2	4	20
12	4	48	3	9	36
63	48	$\frac{432}{48} = 9$	0		124

$\bar{x} = \frac{63}{7} = 9$, $\sigma = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{\Sigma f}} = \sqrt{\frac{124}{48}}$

$\bar{x} =$

$= 1.607$

$N = \Sigma f = 48$

$\frac{N}{2} = 24$, cum.f. just greater than 24 is 31

$\therefore$ Median is 9.

9 has maximum freq (13). $\therefore$ Mode = 9

2a) $p_1 = .4238$, $p_2 = .3144$

i) $H_0 : p_1 \le p_2$

$H_A : p_1 > p_2$

ii) $Z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\hat{\sigma}_{\hat{p}_1 - \hat{p}_2}} = \dfrac{.1094}{.0432} = +2.5331$

$\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2} = \dfrac{89 + 94}{210 + 299} = \dfrac{183}{509} = .3595$

$\hat{\sigma}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\dfrac{\bar{p}(1-\bar{p})}{n_1} + \dfrac{\bar{p}(1-\bar{p})}{n_2}} = .0432$

iii) $\alpha = .05$

iv) $Z_\alpha = 1.645$ (one-tailed)

v) $2.533 > 1.645$ $\therefore$ Reject H_0

(ie) alcoholism is more likely to be present in families of subjects with unipolar depression.

b)

x	$X = \log x$	y	$Y = \log y$	XY	X^2
1		2.98			
2		4.26			
3		5.21			
4		6.1			
5		6.8			
6		7.5			
	2.88574		4.3133	2.2671	1.749

$y = a x^b$

$\log y = \log a + b \log x$

$Y = A + bX$

Eqns.

$\Sigma Y = nA + b \Sigma X$

$\Sigma XY = A \Sigma X + b \Sigma X^2$

$A = .474$, $a = 2.9785$, $b = .5143$

$y = 2.9785\, x^{.5143}$

c) Mahajan - pg. 140 (1 way ANOVA)

3a)

Pre	Post	d_i = Post − Pre	d_i	d_i
⋮	⋮	⋮	⋮	⋮
		57	6	84

$\bar{d} = \frac{57}{17} = 3.3529 \simeq 3$

$s_d^2 = \frac{\sum d_i^2}{16} - \left(\frac{\sum d_i}{n-1}\right)^2 = 5.1094$

ii) $t = \frac{3.3529 - 0}{\sqrt{\frac{5.1094}{17}}} = 6.1159 \quad \left(t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}\right)$

i) $H_0 : \mu_d \leq 0$

$H_A : \mu_d > 0$

iii) $\alpha = .01$

iv) $t_{1-\alpha} = 2.583$ (one tailed)

v) Decision: $6.1159 > 2.583$ so reject H_0.

The training was effective or increases knowledge of behaviour modification principl

b)

No.	O_i	Expected rel. freq.	Expected freq. (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
0	8	.1073	10.73	.6946
1	25	.2684	26.84	.1261
2	32	.302	30.2	.1073
3	24	.2013	20.13	.744
4	10	.088	8.8	.1636
5 or more	1	.0339	3.3	1.603
	100	.9934	99.34	3.4386

$P(x) = {}_nC_x\, p^x q^{n-x} = {}_{10}C_x (.2)^x (.8)^{10-}$

i) H_0: The sampled pop. has a binomial dist

H_A: " " does not follow a binomial di

ii) $X^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i}\right] = 3.4386$

iv) d.o.f. $= 6 - 1 = 5$, $\chi^2_{.9} = 9.236$

v) $3.4386 < 9.236$ so cannot reject H_0.

We conclude that the no. of erroneous entries may follow a binomial dist.

c)

	Crop 1	Crop 2	Crop 3	Crop 4	Total	Mea
Fert. A					24.8	6.2
Fert. B					33.2	8.3
Fert. C					23.6	5.9
T					81.6	
						6.8

SST = 23.08 , SSBl = 13.68

SSTr = 2.82 , SSE = 6.58

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

H_A: not all μ's are equal

ii) V.R. = 6.237

iii) $\alpha = 0.01$

iv) $F_{0.01}(2,6) = 10.92$

v) $6.237 < 10.92$, cannot reject H_0.

We conclude that there is no significant difference in yield due to fert.

ANOVA table for the Randomised Complete Block Design

Source	SS	d.f	MS	V.R.
Treatments	2.82	3	.94	.8571
Blocks	13.68	2	6.84	6.2371
Residual	6.58	6	1.0966	
Total	23.08	11		

4a). $w_1 = \frac{s_1^2}{n_1} = .2$, $w_2 = \frac{s_2^2}{n_2} = 4.1667$

$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.0897$

$100(1-\alpha)\%$ Confidence interval for $\mu_1 - \mu_2$ is given by $\bar{x}_2 - \bar{x}_1 \pm t'_{1-\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

90% conf. int. $= (25.417, 32.583)$

$t_1 = 1.7291$, $t_2 = 1.7139$

$t'_{1-\alpha/2} = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} = 1.7146$

95% $\Rightarrow t_1 = 2.093$, $t_2 = 2.0687$

$t'_{1-\alpha/2} = 2.0698$

conf. int. is $(24.6747, 33.3253)$

99% $\Rightarrow t_1 = 2.8609$, $t_2 = 2.8073$,

$t'_{1-\alpha/2} = 2.8098$

conf. int. is $(23.1284, 34.8716)$

4b) i) H_0: The two variables are independent.

H_A: " " " " not ".

ii) $X^2 = \sum\left[\frac{(O_i - E_i)^2}{E_i}\right] = 42.5786$

iii) $\alpha = 0.05$

iv) $(r-1)(c-1) = (2-1)(2-1) = 1$ d.o.f.

$\chi^2_{.95} = 3.841$

v) $42.5786 > 3.841$. We reject H_0.

There is a relation between knowledge of diseases and major field of study.

O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
31	12.2	28.9705
91	109.8	3.2189
19	37.8	9.3503
359	340.2	1.0389
500	500	42.5786

5a)

x	$d = x - 80$	d^2
⋮	⋮	⋮
	-2	394

$$s^2 = \frac{\Sigma d^2}{n-1} - \left(\frac{\Sigma d}{n-1}\right)^2 = 28.1224$$

$s = 5.30$, $t_{1-\alpha/2} = 2.1448$

$$\bar{x} \pm t_{(1-\alpha/2)} \frac{s}{\sqrt{n}} = 76.193, 82.807$$

b) i) H_0: The two communities do not differ with respect to the variable of interest

H_A: differ with " " " " " "

ii) $\chi^2 = 16.8$

iii) $\alpha = 0.05$

iv) $\chi^2_{.95} = 3.841$ (1 d.o.f)

v) We reject H_0 since $16.8 > 3.841$

The two communities differ with respect to the variable of interest.

O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
43	62	5.8
81	62	5.8
157	138	2.6
119	138	2.6
		16.8

6 a) $p = .9$, $q = .1$ $n = 3$ $P(x) = 3C_x (.9)^x (.1)^{3-x}$

i) $P\{X = 0\} = .001$

ii) $P\{X = 1\} = .027$

iii) $P\{X > 1\} = P\{X = 2\} + P\{X = 3\} = .972$

iv) $P\{X \leq 2\} = P\{X = 0\} + P\{X = 1\} + P\{X = 2\} = .271$

v) $P\{X = 2 \text{ or } 3\} = .972$

b) Biostatistics (seventh edition) – Wayne W. Daniel – pg. 671

c) i) $H_0 : \sigma^2 = 1$

$H_A : \sigma^2 \neq 1$

Test statistic : $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = 6.75$

$\alpha = .05$

Critical value : Values of χ^2 are 2.7 and 19.023

Decision: $2.7 < 6.75 < 19.023$

So do not reject H_0. Unable to conclude that the pop. variance is not 1.

ii) $df_1 = 10$, $df_2 = 12$

$F_{.975} = 3.37$, $F_{.025} = \frac{1}{3.62} = .2762$

$$\frac{s_1^2/s_2^2}{F_{1-\alpha/2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2/s_2^2}{F_{\alpha/2}}$$

$$\Rightarrow \frac{.7095}{3.37} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{.7095}{.2762}$$

$$\Rightarrow .2105 < \frac{\sigma_1^2}{\sigma_2^2} < 2.5686$$

7a)

Age	no.	cum. freq.	rel. freq.	cum. rel. freq.
5-14	5	5	.0286 ~~.0667~~	~~.0667~~ .0286
15-24	10	15	.0571 ~~.1333~~	~~.2~~ .0857
25-34	120	135	.6857 ~~.2667~~	~~.4667~~ .7714
35-44	22	157	.1257 ~~.2933~~	~~.76~~ .8971
45-54	13	170	.0743 ~~.0667~~	~~.9333~~ .9714
55-64	5	175	.0286	1 1
	175		1	

Also histogram & frequency polygon.

7b) i) $H_0: M_A = M_B$

$H_A: M_A \neq M_B$

ii) $\chi^2 = 16.13$

iii) $\alpha = .05$ (1 d.o.f)

iv) $\chi^2 = 3.841$

& v) since $16.13 > 3.841$, reject H_0.

v) we conclude that the two pop. medians are different.

Median $= \frac{79+80}{2} = 79.5$ (common median)

	Hospital A	Hospital B	Total
No. above median	13	2	15
No. below median	2	13	15
	15	15	30

$$\chi^2 = \frac{n(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)} = \frac{30(169-4)^2}{15\times15\times15\times15}$$

$$= 16.13$$

c) i) $H_0: \mu_1 \leq \mu_2$

$H_A: \mu_1 > \mu_2$

ii) $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = 1.75$

$$s = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = .443$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s} = 3.95$$

iii) $\alpha = .01$

iv) $t_{1-\alpha} = 2.438$, d.o.f. $= n_1 + n_2 - 2$
$= 35$

v) ~~8~~ 9 $3.95 > 2.438$ we reject H_0
Attenders score higher than nonattenders.

95% conf. interval for $\mu_1 - \mu_2$ is given by $(\bar{x}_1 - \bar{x}_2) \pm t_{(1-\alpha/2)} \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$

$= 1.75 \pm 2.7239(.443)$

$= 1.75 \pm 1.2067$

$= (2.9567, .5433)$