ANSWERS

Problem W-1

Wilcoxon Rank-sum test

Eton	96	85	112	90	95	99
Rank	4	1	14	2	3	6

Comp.	110	106	103	98	100	101	102	104
Rank	13	12	10	5	7	8	9	11

Sum of ranks (Eton) = 30
$W = 30$, $N_1 = 6$, $N_2 = 8$
Critical value (2-tailed, sig. Level 0.05) = 29

The demonstration has not been effective.

Problem W-2

Wilcoxon Rank-sum test

intact	15	30	11	30	12	47
rank	3	4.5	1	4.5	2	7

removed	90	120	42	382	178	87
rank	9	10	6	12	11	8

Sum of ranks (intact) = 22
Sum of ranks (removed) > 22
$W = 22$, $N_1 = 6$, $N_2 = 6$
Critical value (2-tailed sig. Level 0.05) = 26

2-tailed test appropriate as we cannot be sure removal of hippocampus will make this task harder, rather than easier (no previous scientific data cited in question)

Time to platform significantly slower in rats with their hippocampi removed. There is therefore some support for the neuroscientist's hypothesis.

Problem W-3

Wilcoxon matched-pairs test

Before	18	12	9	15	22	6	9	17
After	22	22	21	15	22	11	11	15

Diff.	-4	-10	-12	~~0~~	~~0~~	-5	-2	2
Rank	3	5	6			4	1.5	1.5

$T+ = 1.5$ $T- > 1.5$ $T = 1.5$
$N = 6$
Critical value (2-tailed sig. Level 0.05) = 1

Exercise does not significantly affect visual acuity.

Problem W-4

Wilcoxon matched-pairs test

-20	-10	+5	+8	+7	-30	-12	+9	-14	+1	-50
11	8	3	5.5	4	13	9	7	10	1	14

-8	-22	+3
5.5	12	2

T- = 82.5 T+ = 22.5 T = 22.5 N = 14
Critical value (1-tailed sig. level 0.05) = 26

- Significant effect of the drug.
- 1-tailed test appropriate on basis of prior experiments.
- However, drug leads to about an 8% DROP in exam performance.
- Therefore, drug does not improve exam performance.
- Critical value (2-tailed sig. level 0.05) = 21
- Does not significantly reduce exam performance on a 2-tailed test.

We are unable to conclude whether the drug has any significant effect, beneficial or otherwise.

Problem V-1

Heterogeneity of variance F-test

This year		Last year	
X	$\left(X-\overline{X}\right)^2$	X	$\left(X-\overline{X}\right)^2$
32	40.107	98	1778.056
24	2.779	12	1921.332
9	277.789	98	1778.056
58	1045.423	25	950.674
12	186.787	3	2791.326
19	44.449	99	1863.390

Mean = 25.667
Var. = 1597.334 / 5
= 319.467

Mean = 55.833
Var. = 11082.834 / 5
= 2216.567

F(5,5) = 2216.567 / 319.467 = 6.94
Critical value (2-tailed sig. level 0.05) = 7.15

The variability has not changed significantly.

The MD should also be made aware that average sales have more than halved since last year. A Wilcoxon or t-test should be performed to check whether this is significant.

Problem V-2

Heterogeneity of variance F-test

Before drug With drug

X	$(X-\overline{X})^2$	X	$(X-\overline{X})^2$
8.5	3.179	7	0.004
0.4	39.904	7.2	0.068
12.9	38.229	8.2	1.588
5.2	2.301	6.9	0.002
4.2	6.335	5.4	2.372
9.1	5.679		

Mean = 6.717	Mean = 6.94
Var. = 95.627 / 5	Var = 4.034 / 4
= 19.125	= 1.009

F(5,4) = 19.125 / 1.009 = 19.0
Critical value (2-tailed sig. level 0.05) = 9.36

The number of hours the child sleeps per night varies less when they are on the drug.

The psychologist should not to over-generalise this result. The effect is reliable for this child – it does not follow that the drug is likely work for all children.

<u>Problem Z-1</u>

Z test

Z = (4-15.3) / 6.6 = -1.712
1-tailed probability = 0.04

The patient is significantly impaired on the task.

A one-tailed test is appropriate given the weight of clinical evidence that head injuries do not improve performance on this task.

<u>Problem Z-2</u>

Z-test

Z = (7.6-7) / 0.316 = 1.90
2-tailed probability = 0.0287 x 2 = 0.0574

The deviation from the mean is not significant at the 0.05 level. However, it is still pretty likely that the machine is not working properly, so a check may be in order in any case.

<u>Problem t-1</u>

Related samples t-test

Diff.	$(X-\overline{X})^2$
30	6.250
20	156.250
80	2256.250
40	56.250
30	6.250
-5	1406.250

Mean = 32.5
Var = 3887.5 / 5 = 777.5
Std. Dev. = 27.884
Std. Err. = 27.884 / 2.449 = 11.386
t(5) = 32.5 / 11.386 = 2.854

Critical value (2-tailed sig. level 0.05) = 2.571

Words are read significantly faster.

Problem t-2

Related samples t-test

Diff.	$\left(X - \overline{X}\right)^2$
1	0.16
1	0.16
1	0.16
1	0.16
0	0.36
0	0.36
2	1.96
-1	2.56
0	0.36
1	0.16

Mean = 0.6
Var = 6.4 / 9 = 0.711
Std. Dev. = 0.843
Std. Err. = 0.843 / 3.162 = 0.267
t(9) = 0.6 / 0.267 = 2.247

Critical value (1-tailed sig. level 0.05) = 1.833

A one-tailed test is appropriate given the weight of prior psychophysical evidence.

The dark background significantly increases the perceived brightness of the towels.

Problem t-3

Unrelated samples t-test

Three		Individual	
X	$\left(X - \overline{X}\right)^2$	X	$\left(X - \overline{X}\right)^2$
20	42.25	10	0.391
18	20.25	8	6.891
15	2.25	9	2.641
17	12.25	6	21.391
8	30.25	10	0.391
3	110.25	13	5.641
		9	2.641
		20	87.891

Mean = 13.5 | Mean = 10.625
Var = 217.5 / 5 = 43.5 | Var = 127.878 / 7 = 18.269

F(5, 7) = 43.5 / 18.269 = 2.38
Critical value (2-tailed sig. level 0.05) = 5.29

Variances are not significantly different.

Pooled variance estimate
= (5 x 43.5 + 7 x 18.269) / 12 = 28.782

t (12) = (13.5-10.625) / √(28.782/6 + 28.782/8)
= 2.875 / 2.897

= 0.99

Critical value (2-tailed sig. level 0.05) = 2.179

Being in a group does not significantly affect the level of risk you are prepared to accept.

Problem t-4

Unrelated samples t-test on the differences

Placebo		Drug	
X	$\left(X-\overline{X}\right)^2$	X	$\left(X-\overline{X}\right)^2$
10	0.16	46	187.69
5	29.16	30	5.29
-3	179.56	40	59.29
15	21.16	20	151.29
4	40.96	-3	1246.09
-12	501.76	24	68.89
40	876.16	40	59.29
12	2.56	66	1135.69
15	21.16	40	59.29
18	57.76	20	151.29

Mean = 10.4
Var = 1730.4 / 9
= 192.267

Mean = 32.3
Var = 3124.1 / 9
= 347.122

F(9, 9) = 347.122 / 192.267 = 1.81
Critical value (2-tailed sig level 0.05) = 4.03

Variances are not significantly different.

t(18) = 21.9 / √ (539.389 / 10) = 21.9 / 7.344 = 2.98
Critical value (2-tailed sig. 0.05) = 2.101

The drug reduces alcohol consumption significantly more than the placebo control. We can therefore conclude that the drug does indeed help.

t(9) = 10.4 / (√192.267 / √ 10) = 10.4 / 4.385 = 2.37
Critical value, 2-tailed = 2.262

Related-samples t-test on placebo reveals that it also significantly reduces consumption, illustrating the importance of that control.

Problem PCB-1: $0.25^5 = 0.00098$

Problem PCB-2: C = N! / [X! (N-X)!] = 7! / [3! 4!] = (7 x 6 x 5) / (3 x 2) = 35

Problem PCB-3

Binomial test

2 increase, 7 decrease p = q = 0.5 N = 9 X = 7

$P(7) = [\ 9! / (\ 7! \times 2)\ ] \times 0.5^7 \times 0.5^2$
$= [9 \times 8 / 2\ ] \times 0.008 \times 0.25 = 0.072$

$P(8) > 0$ $P(9) > 0$

Probability of at least 7 decreases out of 9 is >0.072.

The drug does not significantly reduce blood pressure.

Problem PCB-4
Binomial test

$N = 6, p = q = 0.5$
$P(5) = (\ 6! \times 0.5^5 \times 0.5\ ) / 5! = 6 \times 0.03125 \times 0.5 = 0.09375$
$P(6) = = (\ 6! \times 0.5^6 \times 1\ ) / 6! = 0.5^6 = 0.0156$
$P(5) + P(6) = 0.11$

Not significant at .05 level. Cannot conclude that rats have learned from experience.

Problem C-1

Chi-square test

	Heinz	Daddy's	Own	Value	
Obs	20	12	9	3	Total = 44
Ex	11	11	11	11	

$\chi^2 = (81 + 1 + 4 + 64) / 11 = 13.62$
d.f. = N –1 = 3
Critical value (0.05) = 7.815

There is a significant effect of brand.

Problem C-2

Chi-square test

	High-energy	*Low-energy*	*TOTAL*
near	67 (69.125)	31 (28.875)	98
far	12 (9.875)	2 (4.125)	14
TOTAL	79	33	112

$\chi^2 = 0.065 + 0.156 + 0.457 + 1.095 = 1.773$
d.f. = 1 Critical value (at 0.05) = 3.841

Energy level of the food has no significant effect on the number of rodents arriving at near and far feeders.

The hypothesis is not supported.

Problem C-3

Chi-square test

	Success	Failure	TOTAL
Phillip	17 (20.52)	17 (13.48)	34

Ricki	18 (14.48)	6 (9.52)	24
TOTAL	35	23	58

$\chi^2 = 0.604 + 0.919 + 0.856 + 1.302 = 3.681$
d.f. = 1 Critical value (at 0.05) = 3.841

Ricki is not significantly more famous on this measure.

Problem CLR-1

Pearson correlation

										Mean
Introversion	2	4	12	11	10	9	12	18	20	10.889
Anxiety	1	6	2	5	5	6	5	10	9	5.4444
X -mean	-8.89	-6.89	1.11	0.11	-0.89	-1.89	1.11	7.11	9.11	
Y - mean	-4.44	0.56	-3.44	-0.44	-0.44	0.56	-0.44	4.56	3.56	Sum
Product	39.5	-3.83	-3.83	-0.05	0.4	-1.05	-0.49	32.4	32.4	95.444

$cov_{xy} = 95.444 / 8 = 11.931$

For x
$s^2 = (79.03 + 47.47 + 1.232 + 0.012 + 0.792 + 3.572 + 1.232 + 50.552 + 82.992) / 8$
$= 33.361$
$s = 5.776$

For y
$s^2 = (19.714 + 0.314 + 11.834 + 0.194 + 0.194 + 0.314 + 0.194 + 20.794 + 12.674)/8$
$= 8.278$
$s = 2.877$

$r = 11.931 / 16.618 = 0.718$
N = 9
Critical value = 0.657

The two measures are significantly correlated. However, correlation does not imply causation, so there is no direct evidence for the student's hypothesis as stated.

Problem CLR-2

Spearman correlation as X-axis is experimenter defined and so unlikely to be normally distributed.

												Mean
Degrees	1	2	3	4	5	6	7	8	9	10	11	6
RT (s)	1	2	3	4	5	6	7	9	8	10	11	6
(X- mean of X)	-5	-4	-3	-2	-1	0	1	2	3	4	5	
(Y - mean of Y)	-5	-4	-3	-2	-1	0	1	3	2	4	5	Sum
Product	25	16	9	4	1	0	1	6	6	16	25	109

$cov_{xy} = 109 / 10 = 10.9$

For x, and for y
$s^2 = (25+16+9+4+1+0+1+4+9+16+25) / 10 = 11$

$r = 10.9 / 11 = 0.99$

N = 11
Critical value = 0.596
Reaction time and degree of rotation are significantly correlated

Linear regression

												Mean
Degrees	0	20	40	60	80	100	120	140	160	180	200	100
RT (s)	1.1	1.3	1.5	1.7	1.9	2.1	2.3	2.5	2.4	2.9	3.2	2.0818
(X- mean)	-100	-80	-60	-40	-20	0	20	40	60	80	100	
(Y – mean)	-0.98	-0.78	-0.58	-0.38	-0.18	0.02	0.22	0.42	0.32	0.8182	1.1182	
												Sum
Product	98.2	62.5	34.9	15.3	3.64	0	4.36	16.7	19.1	65.455	111.82	432

$cov_{xy} = 432 / 10 = 43.2$

For x
$s^2 = (10000+6400+3600 + 1600 + 400 + 0 + 400 + 1600 + 3600 + 6400 + 10000)/10$
$= 4400$

b = 43.2 / 4400 = 0.0098
a = 2.0818 – 0.0098x100 = 1.102

RT = 0.0098 x degrees +1.102

Problem E-1

0 - 0.33	]]	This distribution seems rather asymetric, raising concerns over its normality. Given the small sample size, non-parametric statistics might be advisable.
0.34 - 0.67	]]]	
0.68 - 1.01	]]]]]]	

86-90	]]	With the sample size given, this is about as close to normality as you could possibly expect. The sample is also reasonably large and so no real concerns over using parametric statistics here.
91-95	]]]	
96-100	]]]]]]]]	
101-105	]]]]]]]	
106-110	]]	
111-115	]	

Note: Non-parametric means statistics that do not make an assumption of normality, such as the Wilcoxon tests or Spearman's r_s.

Problem E-2

Differences
-30 -8 -6 -1 2 2 5 5

Median = 0.5 LQ = -7 UQ = 3.5 IQR = 10.5 Whisker = 15.75

There are no outliers.

Question F-1
EDA - Boxplot of each group. State there are no outliers.

Wilcoxon

Bike	50 (13), 49(12), 25(1), 27(2), 29(5),45(11),52(14)
Car	28(3), 30(7.5), 31(9.5), 30 (7.5), 31 (9.5), 29 (5), 29 (5)

$W_{bike} = 58$ $W_{car} = 47$ $W = 47$

Not significant at .05 level. There is no evidence for the police argument

Question F-2

	None	1 cup	6 cups	TOTAL
1^{st}	0	2	2	4
2:1	14	14	4	32
2:2	6	4	14	24
TOTAL	20	20	20	60

Expected values for 1^{st} will be below 5 (1.33) so a chi-square will be invalid.
Combine 1^{st} and 2:1 rows to give valid expected values.

	None	1 cup	6 cups	TOTAL
2:1 or 1st	14 (12)	16 (12)	6 (12)	36
2:2	6 (8)	4 (8)	14 (8)	24
TOTAL	20	20	20	60

Chi-square = 1/3 + 4/3 + 3 + 1/2 + 2 + 4.5 = 11.7
df = 2

Caffeine intake and exam performance are related. Nature of relationship is that low caffeine level seem to slightly improve performance whilst high caffeine levels reduce it.

Question F-3

$F(9,11) = 1.62$

Pooled variance = 37
$t(20) = 6.7 / 2.60 = 2.58$

Attempt to suppress significantly increases number of thoughts, so no support for experimental hypothesis.

Question F-4
Two histograms showing approximately normal distributions for RT and IQ

Covariance	-371.962
Std Dev X	34.97391
Std Dev Y	11.51646
Mean X	239.5833
Mean Y	99.08333

r = -0.92

Significant relationship between RT and IQ

IQ = 171.9 - 0.304 x RT