Exercise 18.5

1. A regular dice is rolled 60 times and the results are given in the table. At the level of significance of $\alpha = 0.05$, determine if the dice is fair.

Outcome	1	2	3	4	5	6	Total frequency
Observed	7	10	12	14	8	9	60

2. An upright spinner with 4 equal regions, A, B, C, and D, is suspect due to the unbalanced weight of the spinner needle. The observed frequencies of the outcomes in 40 spins are shown in the table.

Outcome	A	B	C	D	Total frequency
Observed	13	7	9	11	40

Determine at a level of significance of $\alpha = 0.05$ whether or not the four spinner outcomes can be considered to be equally likely.

3. A biased dice is produced to sell as a novelty item with expected outcomes, given as percentages, shown in the table.

Outcome	1	2	3	4	5	6	Total probability
Expected %	20	10	10	10	10	40	100

In 120 rolls of the dice, the observed outcomes are as follows.

Outcome	1	2	3	4	5	6	Total frequency
Observed	27	15	16	18	13	31	120

At the level of significance of $\alpha = 0.05$, test whether the data support the design.

4. A college instructor uses a predetermined grade distribution in all of her classes, regardless of the students who comprise the classes.

Grade	A	B+	B	B−	C	D	E	Total
Planned %	10	10	10	20	20	20	10	100

The actual distribution of marks for a class of 50 students is given below.

Grade	A	B+	B	B−	C	D	E	Total
Actual (observed)	3	4	5	9	10	12	7	50

Calculate the expected values then determine whether the data fit the intended model at a level of significance of $\alpha = 0.05$

5. An experiment consists of flipping a coin 5 times and recording the number of heads. When the experiment is repeated 120 times, the following number of heads is noted in the 160 trials.

Heads	0	1	2	3	4	5	Total frequency
Observed	2	18	41	62	28	9	160

Assuming the coin is fair and using the binomial distribution to find the expected numbers of outcomes, determine at a level of significance of $\alpha = 0.05$ whether or not the observations follow the theoretical binomial distribution.

Exercise 18.6

1. In a random survey of 9th and 10th grade students, their favourite music genre of four choices was tallied. The table below indicates their responses.

Genre	Pop	Classical	Dance	Metal	Total
9th grade	7	8	6	11	32
10th grade	9	11	15	8	43

The claim is made that the choice of music is independent of the student's grade.

(a) State the null and alternative hypothesis, indicating which of the two reflects the claim.

(b) Indicate the degrees of freedom, ν, in this analysis.

(c) At the level of significance of $\alpha = 0.01$, test the claim.

2. The mathematics marks of last year's HL students were compared to the predicted marks for this year's candidates, with the results shown below.

IB mark	7	6	5	4	3	2	Total frequency
Last year	8	14	13	9	8	1	53
This year	11	12	8	5	4	0	40

To test the claim that the marks are independent of the groups:

(a) state the null and alternative hypotheses and identify the claim

(b) identify the number of degrees of freedom, ν

(c) find the χ^2 statistic

(d) determine at the level of significance of $\alpha = 0.05$ whether or not the achievement level distributions and the groups of students are independent.

3. The difference in voting preferences between three areas was found through random sampling by students in the Geography class and is summarised below.

Area	Liberal	Conservative	Socialist	Total
Westside	47	30	10	87
Eastside	25	26	34	85
City centre	30	25	18	73

The claim is that the voting preferences are not independent of the area.

(a) State the null and alternative hypotheses.

(b) Indicate the number of degrees of freedom, ν.

(c) At the level of significance of $\alpha = 0.05$, test the claim.

4. The number of social media postings per week by students and business owners is presented below.

Postings	0–4	5–9	10–14	15–19	20–24	25–30	31+	Total
Students	2	16	20	23	40	35	12	148
Business owners	5	12	16	18	16	8	4	79

The claim by the business owners is that the distributions are independent of the sample groups.

(a) State the null and alternative hypotheses. Identify the claim.

(b) Indicate the number of degrees of freedom, v.

(c) Test the claim at the level of significance of $\alpha = 0.05$

5. Before repainting the interior walls, the school administration sought some colour preferences from the teachers and students. The Maths HL class surveyed random samples of the groups, and their findings are given in the table below.

Colour	blue	green	beige	pink	yellow	Total frequency
Teachers	9	8	11	6	6	40
12th grade	12	7	5	6	9	39
11th grade	11	10	8	7	9	45
10th grade	10	9	5	8	9	41

Test the claim that the colour preferences are independent of the school grade or group at the level of significance of $\alpha = 0.05$

(a) State the null and alternative hypotheses. Identify the claim.

(b) Indicate the number of degrees of freedom, v.

(c) Test the claim at the level of significance of $\alpha = 0.05$

6. The school counsellor would like to test whether mark distribution in IB subjects are independent in the following subjects: Mathematics HL, Biology HL, Chemistry HL, and Physics HL. Data from the previous graduating class are given in the table.

IB mark awarded	7	6	5	4	3	2	1	Total
Mathematics	4	6	3	5	3	2	2	25
Biology	2	4	7	9	4	2	1	29
Chemistry	2	7	10	7	3	1	1	31
Physics	6	4	9	10	3	3	0	35

Test her claim at the level of significance of $\alpha = 0.05$

Exercise 18.4

1. (a) $H_0: d < 0$, $H_1: d > 0$ (claim: 'before' minus 'after' is positive).
   (b) The p-value ≈ 0.0137. Since the p-value is less than α, reject H_0 and do not reject the claim at the level of significance of $\alpha = 0.05$.
2. (a) $H_0: d \geqslant 10$, $H_1: d < 10$ (claim)
   (b) The p-value ≈ 0.00871. Since the p-value is less than α, reject H_0 and reject the claim at the level of significance of $\alpha = 0.05$.
3. (a) $H_0: d \geqslant 0$, $H_1: d < 0$ (claim: 'before' – 'after' should be negative)
   (b) The p-value ≈ 0.0431. Since the p-value is less than α, reject H_0 and do not reject the claim at the level of significance of $\alpha = 0.05$.
4. (a) $H_0: d \geqslant 0$, $H_1: d < 0$ (claim: 'before' – 'after' should be negative)
   (b) The p-value ≈ 0.0292. Since the p-value is less than α, reject H_0 and do not reject the claim at the level of significance of $\alpha = 0.05$.
5. $H_0: \mu_{girls} < \mu_{boys}$ (claim)
   $H_1: \mu_{girls} > \mu_{boys}$
   The p-value ≈ 0.222. Since the p-value is greater than α, we do not reject H_0 but reject the claim at the level of significance of $\alpha = 0.05$.
6. $H_0: \mu_1 = \mu_2$ (claim)
   $H_1: \mu_1 \neq \mu_2$
   The p-value ≈ 0.0310. Since the p-value is less than α, reject H_0 and reject the claim at the level of significance of $\alpha = 0.05$.

Exercise 18.5

1.

Outcome	1	2	3	4	5	6	Total frequency
Observed	7	10	12	14	8	9	60
Expected	10	10	10	10	10	10	60

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.639. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. (The data support the notion that the die is fair.)

2.

Outcome	A	B	C	D	Total frequency
Observed	13	7	9	11	40
Expected	10	10	10	10	40

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.572. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. (The data support the intended probability distribution.)

3.

Outcome	1	2	3	4	5	6	Total frequency
Observed	27	15	16	18	13	31	120
Expected	24	12	12	12	12	48	120

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.0413. Since the p-value is less than α, reject H_0 at the level of significance of $\alpha = 0.05$. (The data do not support the intended probability model.)

4.

Grade	A	B+	B	B−	C	D	E	Total
Observed	3	4	5	9	10	12	7	50
Expected	5	5	5	10	10	10	5	50

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.890. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$.

5. The expected values are found using the binomial probability feature on your GDC:
   $160 \times$ binomPdf(5, 0.5) = {5, 25, 50, 50, 25, 5}

Heads	0	1	2	3	4	5	Total frequency
Observed	2	18	41	62	28	9	160
Expected	5	25	50	50	25	5	160

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.0373. Since the p-value is less than α, reject H_0 at the level of significance of $\alpha = 0.05$. (The observations do not fit the binomial model of a fair die tossed 5 times.)

6. (a)

Day	Mon	Tue	Wed	Thu	Fri	Total absences
Observed	24	14	12	14	26	90
Expected	18	18	18	18	18	90

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.0533. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. (There is insufficient evidence that absenteeism is dependent on the day of the week.)

(b)

Day	Mon/Fri	Tue	Wed	Thu	Total absences
Observed	24	14	12	14	90
Expected	18	18	18	18	90

H_0: The observed outcomes fit the intended model.
H_1: The observed outcomes do not fit the intended model.
The p-value ≈ 0.0265. Since the p-value is less than α, reject H_0 at the level of significance of $\alpha = 0.05$. (Friday and Monday absences combined are statistically higher than during the rest of the week.)

Exercise 18.6

1. (a) H_0: There is no difference in the preferences of the 9th and 10th graders. (claim)
   H_1: There is a difference that cannot be attributed to chance alone.
   (b) 3
   (c) The p-value ≈ 0.319. Since the p-value is greater than α, we do not reject H_0 at the level of significance of $\alpha = 0.05$.

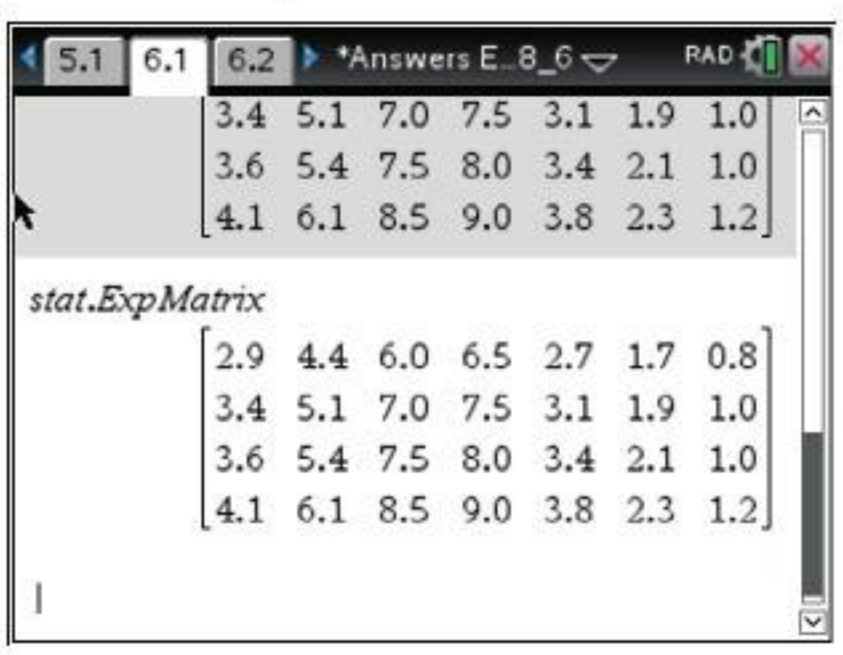
2. (a) H_0: There is no difference in the marks between last year's and this year's class. (claim)
H_1: There is a difference that cannot be attributed to chance alone.
(b) 5
(c) $\chi^2 \approx 3.55$
(d) The p-value ≈ 0.616. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. The achievement level distributions and the groups of students are independent.

3. (a) H_0: There is no difference between the voting preferences between areas.
H_1: There is a difference that cannot be attributed to chance alone. (claim)
(b) 4
(c) p-value ≈ 0.00043. Since the p-value is less than α, reject H_0 at the level of significance of $\alpha = 0.05$. The voting preferences are not independent of the ridings.

4. (a) H_0: There is no difference in the posting distributions between the groups. (claim)
H_1: There is a difference that cannot be attributed to chance alone.
(b) The expected values, shown below to 1 decimal place, contain one value that is less than 5.
$$\begin{bmatrix} 4.6 & 18.3 & 23.5 & 26.7 & 36.5 & 28.0 & 10.4 \\ 2.4 & 9.7 & 12.5 & 14.3 & 19.5 & 15.0 & 5.6 \end{bmatrix}$$
The values in the first and second columns will be combined. The observed matrix will become
$$\begin{bmatrix} 18 & 20 & 23 & 40 & 35 & 12 \\ 17 & 16 & 18 & 16 & 8 & 4 \end{bmatrix}$$ and the expected values to 1 decimal place will become
$$\begin{bmatrix} 22.8 & 23.5 & 26.7 & 36.5 & 28.0 & 10.4 \\ 12.2 & 12.5 & 14.3 & 19.5 & 15.0 & 5.6 \end{bmatrix}.$$
So $\upsilon = 5$.
(c) The p-value ≈ 0.0285. Since the p-value is less than α, reject H_0 at the level of significance of $\alpha = 0.05$. The distributions of postings by the groups are not independent.

5. (a) H_0: There is no difference in the colour preferences between the groups. (claim)
H_1: There is a difference that cannot be attributed to chance alone.
(b) $\upsilon = 12$
(c) The p-value ≈ 0.9. Since the p-value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. The colour preferences are independent of the groups.

6. The expected values include a few whose values are less than 5, notably in columns 1, 6, and 7.



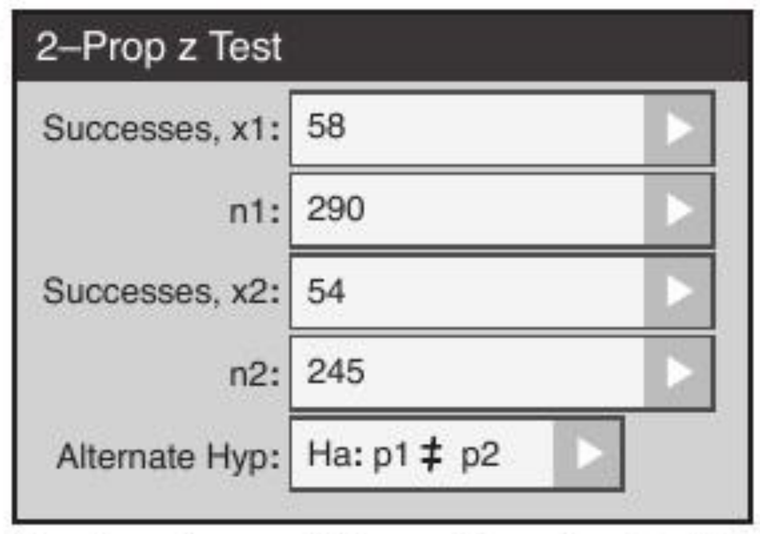
The first two columns as well as the last three columns will be combined so that the χ^2 onditions are met.

The new matrix of observed values is $\begin{bmatrix} 10 & 3 & 5 & 7 \\ 6 & 7 & 9 & 7 \\ 9 & 10 & 7 & 5 \\ 10 & 9 & 10 & 6 \end{bmatrix}$

and the expected values will become $\begin{bmatrix} 7.3 & 6.0 & 6.5 & 5.2 \\ 8.5 & 7.0 & 7.5 & 6.0 \\ 9.0 & 7.5 & 8.0 & 6.5 \\ 10.2 & 8.5 & 9.0 & 7.3 \end{bmatrix}$.

The value $\approx 0.7$and $\upsilon = 9$. Since the value is greater than α, do not reject H_0 at the level of significance of $\alpha = 0.05$. The mark distributions are independent of the groups.

7. The p-value $\approx 0.2$and $\upsilon = 2$. Since the p-value is greater than α, do not reject H_0. The league outcomes are independent of the teams.

8. 

The p-value ≈ 0.6, we do not reject the null hypothesis. There is no difference to the proportion of wins.

Chapter 18 Practice questions

1. (a) 17 (b) 5 (c) 8
2. (a) 1.71 (b) -1.71
(c) ± 2.10 (d) ± 2.90
3. (a) No (b) Yes
(c) Yes (d) No
4. (59.5, 68.5)
5. Since the 95% confidence interval (94.8, 105) does not contain any value equal to or greater than 110, we reject the claim.
6. The 95% confidence interval (12.5, 27.5) does not contain the average of 30 minutes. The statistics do not support the average given.
7. (a) H_0: $\mu = 62.4$ (the claim)
H_1: $\mu \neq 62.4$
(b) H_0: $\mu \geqslant 13$ (the claim)
H_1: $\mu < 13$
(c) H_0: $\mu < 72.3$
H_1: $\mu > 72.3$ (the claim)
(d) H_0: $\mu \geqslant 102$
H_1: $\mu < 102$ (the claim)
8. H_0: $\mu = 28.6$ (the claim)
H_1: $\mu \neq 28.6$
The p-value ≈ 0.0514 , so we do not reject the null hypothesis and claim at the level of significance of $\alpha = 0.05$. The average American BMI is 28.6.
9. H_0: $\mu = 79°$ C (the claim)
H_1: $\mu \neq 79°$ C
The p-value ≈ 0.0000952 , so we reject the null hypothesis and the claim at the level of significance of $\alpha = 0.05$. The coffee is not being served at 79° C