Statistical applications

Answers

Skills check

1 a mean = 3.61 standard deviation = 1.21
The small standard deviation implies that the data are close to the mean
b mean = 14 standard deviation = 0.643
The mean is the middle data value (14) since the frequencies are symmetrical about this value. The standard deviation is very small since most of the data values equal the mean and the rest are close to it

2 a
\[ y = -3x + 4 \]

b
\[ y = 2x - 3 \]

Exercise 5A

1 a

b 0.16 c 0.815 d 200 \times 0.16 = 32

2 a

b 81.5% c 100 \times 0.5 = 50

Exercise 5B

1 a

b 0.0766

c 365 \times 0.0766 = 27.959 or 28 days

2 a i P(IQ < 90) = 0.159
ii P(IQ > 120) = 0.0228
iii P(80 < IQ < 110) = 0.819

b P(IQ > 115) = 0.0668, 2000 \times 0.0668 = 134

3 a 0.0668 b 0.00621 c 300 \times 0.927 = 278

4 a

b 0.0401

c 80 \times 0.0122 = 0.976 or 1

5 a 78.9%
b 0.00621
c 100 \times 0.0304 = 3.04 or 3

6 0.106

7 a 86.4%
b 30 \times 0.0304 = 0.912 or 1
Worked solutions: Chapter 5

Exercise 5C

1  \( p = 4.93 \)
2  \( h = 183 \)
3  \( k = 20.8 \)
4  \( w = 222 \)
5  a  3.47 to 4.99 kg
    b  \( 180 \times 0.683 = 123 \)
    c  0.0685
    d  87.7%
    e  \( w = 5.48 \)
6  a  \( a = 29, b = 30, c = 31 \)
    b  0.919
    c  \( d = 32.8 \)
    d  \( 5000 \times 0.6246\ldots = 3123 \) (accept 3120 to 3125)
7  a  0.000429
    b  0.854
    c  \( t = 5885 \)
8  a  \( f(x) \)
    b  62.5%
    c  \( p = 106 \)
    d  \( 800 \times 0.911 = 729 \)
    e  \( 800 \times 0.159 = 127 \)
9  a  \( f(x) \)
    b  0.0228
    c  0.0668
    d  \( 400 \times 0.0668 = 26.7 \) or 27
    e  \( p = 1006 \)
10 a  0.466%
    b  A baby weighing 2.34 kg (2.34 is nearer the mean than 5.5).
    c  \( 300 \times 0.0808 = 24.2 \) or 24
    d  \( w = 3.16 \)

Exercise 5D

1  a  strong positive linear
    b  moderate negative linear
    c  moderate positive linear

2  a

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
10 & 40 \\
20 & 30 \\
30 & 20 \\
40 & 10 \\
\hline
\end{array} \]

moderate positive linear correlation

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
10 & 40 \\
20 & 30 \\
30 & 20 \\
40 & 10 \\
\hline
\end{array} \]

moderate negative linear correlation

Exercise 5E

1  i  a, c

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
90 & 0 \\
100 & 10 \\
110 & 20 \\
120 & 30 \\
130 & 40 \\
\hline
\end{array} \]

strong positive correlation

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
10 & 40 \\
20 & 30 \\
30 & 20 \\
40 & 10 \\
\hline
\end{array} \]

strong negative correlation

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
20 & 200 \\
25 & 250 \\
\hline
\end{array} \]

mean of \( x = 13, \) mean of \( y = 25.75 \)

ii a, c

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
997 & 0 \\
999 & 10 \\
1001 & 20 \\
1003 & 30 \\
1005 & 40 \\
1007 & 50 \\
1009 & 60 \\
\hline
\end{array} \]

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 25 \\
10 & 50 \\
15 & 75 \\
20 & 100 \\
25 & 125 \\
\hline
\end{array} \]

mean of \( x = 16.5, \) mean of \( y = 20.2 \)

2 a, c

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
20 & 200 \\
\hline
\end{array} \]

moderate positive correlation

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
20 & 200 \\
\hline
\end{array} \]

mean height = 4.78 m

mean weight = 896 kg

\[ \begin{array}{|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
\hline
\end{array} \]

mean of \( x = 5.5, \) mean of \( y = 25.5 \)

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
\hline
\end{array} \]

mean height = 5.5 m

\[ \begin{array}{|c|c|c|}
\hline
x & y \\
\hline
0 & 0 \\
5 & 50 \\
10 & 100 \\
15 & 150 \\
\hline
\end{array} \]

mean of \( x = 13, \) mean of \( y = 25.75 \)
Worked solutions: Chapter 5

Exercise 5G

1 a. \( r = 0.994 \), strong positive correlation
   b. \( y = 1.47x + 116 \)
   c. \( y = 1.47(1000) + 116 = 1586, £1590 \) (3 s.f)

2 a. \( r = 0.974 \)
   b. \( y = 0.483x + 15.6 \)
   c. \( y = 0.483(8) + 15.6 = 19.464, 19.5 \) cm

3 a. \( \bar{x} = 68.6 \) s
   \( s_x = 6.55 \)
   \( \bar{y} = 138 \) s
   \( s_y = 5.97 \)
   b. \( r = -0.860 \)
   c. strong negative correlation
   d. \( y = -0.784x + 192 \)
   e. \( y = -0.784(70) + 192 = 137.12, 137 \) seconds

Exercise 5H

1 a. \( H_0: \) Genre of book is independent of age
   \( H_1: \) Genre of book is dependent on age
   b. \( 300 \times 300 = 42.0 \)
   c. \( (3 - 1) (3 - 1) = 4 \)
   d. \( \chi^2_{calc} = 26.9 \)
   e. \( 26.9 > 9.488, \) therefore we reject the null hypothesis. There is enough evidence to conclude that genre of book is dependent on age. \( (p\text{-value} = 0.0000207 < 0.05) \)

2 a. \( H_0: \) Hair color and eye color are independent
   \( H_1: \) Hair color and eye color are dependent.
   b. \( 90 \times 227 \times 227 = 33.7 \)
   c. \( (3 - 1) (3 - 1) = 4 \)
   d. \( \chi^2_{calc} = 44.3 \)
   e. \( 44.3 > 7.779, \) therefore we reject the null hypothesis. There is enough evidence to conclude that hair colour and eye color are dependent. \( (p\text{-value} = 0.0000000556 < 0.1) \)
3  a  \(H_0\): Favorite flavor is independent of race.
\(H_1\): Favorite flavor is dependent on race.

\[ \chi^2 = 0.675 \]

\(0.675 < 12.59\), therefore we do not reject the null hypothesis. There is enough evidence to conclude that favorite flavor is independent of race. (\(p\)-value = 0.995 > 0.05)

4  a  \(H_0\): Film genre is independent of gender
\(H_1\): Film genre is dependent on gender

\[ \chi^2 = 19.0 \]

\(19.0 > 11.345\), therefore we reject the null hypothesis. There is enough evidence to conclude that film genre is dependent on gender. (\(p\)-value = 0.000276 < 0.01)

5  a  \(H_0\): Grade is independent of the number of hours
\(H_1\): Grade is dependent on the number of hours

\[ \chi^2 = 42.1 \]

\(42.1 > 9.488\), therefore we reject the null hypothesis. There is enough evidence to conclude that grade is dependent on number of hours spent playing computer games. (\(p\)-value = 0.0000000159 < 0.05)

6  a  \(H_0\): Employment grade is independent of gender
\(H_1\): Employment grade is dependent on gender

\[ \chi^2 = 51.6 \]

\(51.6 > 4.605\), therefore we reject the null hypothesis. There is enough evidence to conclude that employment grade is dependent on gender. (\(p\)-value = 6.23 × 10^{-12} < 0.1)

7  a  \(H_0\): Amount of sushi sold is independent the day of the week
\(H_1\): Amount of sushi sold is dependent on the day of the week.

\[ \chi^2 = 84.0 \]

\(84.0 < 9.488\), therefore we do not reject the null hypothesis. There is enough evidence to conclude that the amount of sushi sold is independent of the day of the week. (\(p\)-value 0.933 > 0.05)

8  a  \(H_0\): A puppy’s weight is independent of its parent’s weight.
\(H_1\): A puppy’s weight is dependent on its parent’s weight

\[ \chi^2 = 13.1 \]

\(13.7 > 13.277\), therefore we reject the null hypothesis. There is enough evidence to conclude that a puppy’s weight is dependent on its parent’s weight. (\(p\)-value = 8.08 × 10^{-40} < 0.05)

9  a  \(H_0\): Music preference is independent of age
\(H_1\): Music preference is dependent on age

\[ \chi^2 = 31.5 \]

\(31.5 > 12.59\), therefore we reject the null hypothesis. There is enough evidence to conclude that music preference is dependent on age. (\(p\)-value = 0.0000204 < 0.05)

10 a  \(H_0\): Age at which a baby is potty trained is independent of gender.
\(H_1\): Age at which a baby is potty trained is dependent on gender.

\[ \chi^2 = 51.6 \]

\(51.6 > 4.605\), therefore we reject the null hypothesis. There is enough evidence to conclude that the age at which a baby in potty trained is dependent on gender. (\(p\)-value = 6.23 × 10^{-12} < 0.1).
11 a \( H_0: \) Grade is independent of gender  
\( H_1: \) Grade is dependent on gender

<table>
<thead>
<tr>
<th></th>
<th>5,6 or 7</th>
<th>3 or 4</th>
<th>1 or 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>25.2</td>
<td>66.6</td>
<td>15.1</td>
</tr>
<tr>
<td>Female</td>
<td>24.8</td>
<td>65.4</td>
<td>14.9</td>
</tr>
</tbody>
</table>

c \((2 - 1)(3 - 1) = 2\)
d \(\chi^2 = 0.467\)
e \(0.467 < 5.991, \) therefore we do not reject the null hypothesis. There is enough evidence to conclude that grade is independent of gender. \((p = 0.792 > 0.05)\)

Review exercise

Paper 1 style questions

1 a  
\[ f(x) \]

b 0.0548

c 0.731

d \(60 \times 0.0599 = 3.59, \) 4 men

e 166

2 a 32.2%
b 6000 \times 0.00982 = 58.9, 59 people

3 a 93.3%b \( p = 1.01 \)

4 a strong positive correlation
b no correlation
c moderate negative correlation

5 a strong positive correlation
b \( \bar{x} = 11 \)
c \( \bar{y} = 25 \)
d 23

6 a \( r = 0.980, \) strong positive correlation
b \( y = 0.801x - 77.4 \)
c \( y = 0.801(170) - 77.4 = 58.77, \) 58.8 cm

7 a \( r = 0.810, \) strong positive correlation
b \( y = 0.215x + 14.3 \)
c \( y = 0.215x(40) + 14.3 = 22.9 \) seconds

8 H\(_0\): Flavor of ice creams is independent of age  
H\(_1\): Flavor of ice creams is dependent on age

<table>
<thead>
<tr>
<th></th>
<th>( x &lt; 25 )</th>
<th>( 25 \leq x &lt; 45 )</th>
<th>( x \geq 45 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla</td>
<td>14.06</td>
<td>11.84</td>
<td>11.1</td>
</tr>
<tr>
<td>Strawberry</td>
<td>10.64</td>
<td>8.96</td>
<td>8.4</td>
</tr>
<tr>
<td>Chocolate</td>
<td>13.3</td>
<td>11.2</td>
<td>10.5</td>
</tr>
</tbody>
</table>

degrees of freedom \((3 - 1)(3 - 1) = 4\)
p-value = 0.963 \(> 0.05, \) \(\chi^2 = 0.604\)

We do not reject the null hypothesis. There is enough evidence to conclude that flavor of ice cream is independent of age. 
(critical value = 9.488, \(\chi^2 = 0.604 < 9.488\))

9 a \( H_0: \) The number of pins knocked down is independent of which hand is used.
b \((2 - 1)(3 - 1) = 2\)
c \(20 \times \frac{60}{120} \times 120 = 10\)
d \(p\)-value = 0.422 \(> 0.1\) (significance value).
Therefore we do not reject the null hypothesis. There is enough evidence to conclude that the number of pins knocked down is independent of which hand is used.

10 a \( H_0: \) The outcome is independent of the time spent preparing for a test.
b \((3 - 1)(2 - 1) = 2\)
c \(p\)-value = 0.069 \(> 0.05, \) therefore we do not reject the null hypothesis. There is enough evidence to conclude that the outcome is independent of the time spent preparing for a test.

Paper 2 style questions

1 a  
\[ f(x) \]

b 0.252
c 0.731
d \(60 \times 0.0599 = 3.59, \) 4 men
e \(k = 166\)

2 a  
\[ f(x) \]

b 0.159
c \(300 \times 0.252 = 75.6, \) 76 sweets
3 a, d

\[ x = 166.9 \]
\[ y = 67.3 \]
\[ y = 0.719x - 52.8 \]
\[ 69 \text{ kg} \]

4 a \[ r = 0.823 \]
\[ \text{strong positive correlation} \]
\[ y = 0.219x + 3.85 \]
\[ y = 0.29(35) + 3.85 = 11.515 \]
\[ 12 \text{ hours (nearest hr)} \]

5 a \[ r = 0.866 = 0.9 \text{ (1 d.p.)} \]
\[ \text{strong positive correlation} \]
\[ y = 0.0666x - 2.36 \]

6 a \[ r = 0.887 = 0.89 \text{ (2 d.p.)} \]
\[ \text{strong positive correlation} \]
\[ y = 0.015x + 0.229 \]
\[ y = 0.0151(80) + 0.229 = 1.437, 1.44 \text{ euros} \]

7 a \[ y = 0.163x - 15.0 \]
\[ y = 0.163(170) - 15.0 = 12.71, \text{ dress size 13} \]
\[ r = 0.741 \]
\[ \text{moderate positive correlation} \]

8 \[ H_0: \text{Choice of game is independent of gender} \]
\[ H_1: \text{Choice of game depends on gender} \]

<table>
<thead>
<tr>
<th>Expected values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badminton</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
</tbody>
</table>

degrees of freedom = (2 − 1)(3 − 1) = 2
\[ \chi^2_{\text{calc}} = 0.667 \]
\[ p\text{-value} = 0.717 > 0.05 \]
We do not reject the null hypothesis. There is enough evidence to conclude that choice of game is independent of gender.
(critical value = 5.991, \[ \chi^2_{\text{calc}} = 0.667 < 5.991 \])

9 a \[ p = 21.6 \]
\[ q = 14.4 \]
\[ r = 13.6 \]

b i \[ H_0: \text{The extra-curricular activity is independent of gender} \]
\[ (2 − 1)(3 − 1) = 2 \]

b c \[ \chi^2_{\text{calc}} = 4.613 \]
\[ 4.613 > 4.605, \text{ therefore we reject the null hypothesis. There is enough evidence to conclude that extra-curricular activity is dependent on gender.} \]

10 a \[ b \times c \times d = 108 \]
\[ b \]
\[ c = 132 \]
\[ d = 88 \]

b \[ H_0: \text{position in upper management is independent of gender} \]
\[ H_1: \text{position in upper management is dependent on gender} \]

\[ c i \chi^2_{\text{calc}} = 54.9 \]
\[ (2 − 1)(3 − 1) = 2 \]
\[ (2 − 1)(3 − 1) = 2 \]
\[ 54.9 > 5.991, \text{ therefore we reject the null hypothesis. There is enough evidence to conclude that position in upper management is dependent on gender.} \]

11 a \[ H_0: \text{The choice of candidate is independent on where the voter lives.} \]
\[ H_1: \text{The choice of candidate is dependent on gender.} \]

\[ c i \chi^2_{\text{calc}} = 58.4 \]
\[ (3 − 1)(2 − 1) = 2 \]
\[ 58.4 > 9.21, \text{ therefore we reject the null hypothesis. There is enough evidence to conclude that choice of candidate is dependent on gender.} \]

12 a \[ b \times c \times d = 49.5 \]
\[ b \]

\[ b i \chi^2_{\text{calc}} = 0.400 \]
\[ 0.400 < 5.991, \text{ therefore we do not reject the null hypothesis. There is enough evidence to conclude that grade is independent of gender.} \]