

## Chi-squared tests 6C

- 1  $H_0$ : The observed data can be modelled by a discrete uniform distribution. (The dice is not biased.)  
 $H_1$ : The observed data cannot be modelled by a discrete uniform distribution. (The dice is biased.)  
 The number of degrees of freedom  $\nu = 5$  (six data cells with a single constraint on the total)  
 From the tables:  $\chi_5^2(5\%) = 11.070$

$$\sum \frac{(O_i - E_i)^2}{E_i} = \frac{4^2 + 1^2 + 1^2 + 3^2 + 4^2 + 3^2}{12} = 4.333\dots$$

As 4.333 is less than 11.070, there is not enough evidence to reject  $H_0$  at the 5% level and to suggest that the dice is not fair.

- 2  $H_0$ : Probability of drawing a winning ticket is 0.2  
 $H_1$ : Probability of drawing a winning is not 0.2

The number of degrees of freedom  $\nu = 1$  (2 data cells with a single constraint on the total)  
 From the tables:  $\chi_1^2(5\%) = 3.841$

The observed and expected results are:

Result	Winning	Losing	Total
Observed ( $O_i$ )	15	105	120
Expected ( $E_i$ )	24	96	120
$\frac{(O_i - E_i)^2}{E_i}$	3.375	0.84375	4.21875

As 4.21875 is greater than 3.841, reject  $H_0$ ; there is evidence at the 5% significance level that the probability of drawing a winning ticket is not 20%

- 3  $H_0$ : The observed data is drawn from the travel agent's expected distribution.  
 $H_1$ : The observed data is not drawn from the travel agent's distribution.

The number of degrees of freedom  $\nu = 2$  (three data cells with a single constraint on the total)  
 From the tables:  $\chi_2^2(2.5\%) = 7.378$

$$\sum \frac{(O_i - E_i)^2}{E_i} = \frac{6^2}{10} + \frac{13^2}{60} + \frac{7^2}{30} = 8.05$$

As 8.05 is greater than 7.378, reject  $H_0$ ; there is evidence at the 2.5% significance level that the expected distribution does not fit the data.

- 4 a The expected values in the final three data columns are all less than 5, so these categories must be merged. The adjusted table has five columns (0, 1, 2, 3,  $\geq 4$ ) with a single constraint on the total, and therefore there are four degrees of freedom.
- b  $H_0$ : Data is drawn from the expected distribution.  
 $H_1$ : Data is not drawn from the expected distribution.

From the tables:  $\chi_4^2(5\%) = 9.488$

The observed and expected results are:

Dogs	0	1	2	3	$\geq 4$	Total
Observed ( $O_i$ )	45	19	11	8	17	100
Expected ( $E_i$ )	55	20	10	7	8	100
$\frac{(O_i - E_i)^2}{E_i}$	1.818	0.05	0.1	0.143	10.125	12.236

As 12.236 is greater than 9.488, reject  $H_0$ ; there is evidence at the 5% significance level that the expected distribution does not fit the data.

- 5  $H_0$ : Birth weights from 2000 can be used as a model for birth weights in 2015.  
 $H_1$ : Birth weights from 2000 cannot be used as a model for birth weights in 2015.

The number of degrees of freedom  $\nu = 5$  (six data cells with a single constraint on the total)

From the tables  $\chi_5^2(5\%) = 11.070$

Calculate the expected results by multiplying the total number of observations (687660) by the percentage in each weight band in the year 2000. The observed and expected results are:

Weight (g)	<1500	1500 –1999	2000 –2499	2500 –2999	3000 –3499	$\geq 3500$	Total
Observed ( $O_i$ )	7286	9304	32121	112535	244472	281942	687660
Expected ( $E_i$ )	8939.58	10314.9	34383	113464	245495	275064	687660
$\frac{(O_i - E_i)^2}{E_i}$	305.9	99.1	148.8	7.6	4.3	172.0	737.6

As 737.6 is greater than 11.070, reject  $H_0$ ; there is evidence at the 5% significance level that distribution seen in the 2000 data does not provide a good model for the 2015 data.