Chapter 14: Do my data fit the expected frequencies? Chi squared.

Full answers to study questions

1. Chi square goodness of fit with one variable (type of television programme) that has three categories (documentary, drama, reality TV).
2. 2 (sex) x 3 (programme type) chi square test of association.
3. The variable that has already been designed has two levels: original recipe or new and improved recipe. You could add any relevant variable, as long as the data collected provides frequencies within categories. This would be analysed using a 2 (recipe) x X (your variable) chi square test of association. For example, you could look into how frequently participants tended to eat the sweets made by that company: never, occasionally or frequently. This would be analysed using a 2 (recipe) x 3 (frequency of consumption) chi square test of association.

Full answers for additional dataset

		Teddy held?		
		Not holding teddy	Holding teddy	Row total
Help offered?	No help offered	10	3	13
	Delayed help offered	17	14	31
	Immediate help offered	3	13	16
	Column total	30	30	*Grand total* = 60

First, calculate the row, column and grand totals. You will need these to calculate the expected values in the next table.

	No help offered	Not holding teddy	Holding teddy
Step 1:	Observed data (O)	10	3
Step 2:	Expected data (E)	(13*30)/60 = 6.5	(13*30)/60 = 6.5
Step 3:	$O - E$	10 – 6.5 = 3.5	3 – 6.5 = -3.5
Step 4:	$(O - E)^2$	3.5^2 = 12.25	-3.5^2 = 12.25
Step 5:	$\frac{(O-E)^2}{E}$	12.25/6.5 = 1.885	12.25/6.5 = 1.885
	Delayed help offered	Not holding teddy	Holding teddy
Step 1:	Observed data (O)	17	14
Step 2:	Expected data (E)	(31*30)/60 = 15.5	(31*30)/60 = 15.5
Step 3:	$O - E$	17 – 15.5 = 1.5	14 – 15.5 = -1.5
Step 4:	$(O - E)^2$	1.5^2 = 2.25	-1.5^2 = 2.25
Step 5:	$\frac{(O-E)^2}{E}$	2.25/15.5 = 0.145	2.25/15.5 = 0.145
	Immediate help offered	Not holding teddy	Holding teddy
Step 1:	Observed data (O)	3	13
Step 2:	Expected data (E)	(16*30)/60 = 8	(16*30)/60 = 8
Step 3:	$O - E$	3 – 8 = -5	13 – 8 = 5
Step 4:	$(O - E)^2$	-5^2 = 25	5^2 = 25
Step 5:	$\frac{(O-E)^2}{E}$	25/8 = 3.125	25/8 = 3.125

Step 6: χ^2 = 1.885 + 1.885 + 0.145 + 0.145 + 3.125 + 3.125
χ^2 = 10.31

1. What method of analysis will you use to analyse this dataset? Fully explain your answer.

A 2x3 chi square test of association. The first variable is whether a teddy is held or not, with two categories: not holding or holding a teddy. The second variable is whether the participant offered no help, delayed help or immediate help, which means there are three categories.

2. Suggest a suitable hypothesis for this analysis.

The hypothesis needs to be two tailed, so it could be something like: The observed frequencies of offering help, in association with whether participants are holding a teddy or not, will be different from the frequencies expected by random chance.

3. Calculate the χ^2 statistic and determine whether it is significant.

$\chi^2 = 10.31$ (see above)

df = (number of categories in rows - 1) * (number of categories in columns - 1)
df = (3 - 1) * (2 - 1)
df = 2 * 1
df = 2

Critical value at α = .050 is 5.991
Critical value at α = .025 is 7.378
Critical value at α = .010 is 9.210
Critical value at α = .005 is 10.597

The calculated value of 10.31 is more extreme than the critical value at the smallest α of .010, therefore the analysis is significant.

4. Interpret and write up your findings using APA standards.

The chi square analysis was significant (χ2 (2, N = 60) = 10.31, p < .010). In the no teddy condition, no help was offered more often than would be expected by chance, whereas immediate help was offered less often than chance. In the teddy condition, no help was offered less often than would be expected by chance, whereas immediate help was offered more often than chance. For those offering delayed help, observed frequencies were similar to expected frequencies.

5. Suggest how you could redesign the "teddy" variable so that the study would have a 3x3 categorical design.

The key thing here is to add one more category to the "teddy" variable. There are lots of ways this could be done. For example, it might matter whether a participant holds their own childhood teddy, so we could have the following three categories: no teddy, new teddy provided by the researchers, own childhood teddy.