Stats 101/101G/108 Workshop

Chi-Square Tests [CST]

2020

by Leila Boyle

Stats 101/101G/108 Workshops

The Statistics Department offers workshops and one-to-one/small group assistance for Stats 101/101G/108 students wanting to improve their statistics skills and understanding of core concepts and topics.

Leila's website for Stats 101/101G/108 workshop hand-outs and information is here: **www.tinyURL.com/stats-10x**

Resources for this workshop, including pdfs of this hand-out and Leila's scanned slides showing her working for each problem are available here: **www.tinyURL.com/stats-CST**

Want to get in touch with Leila?

Leila Boyle
Undergraduate Statistics Assistance, Department of Statistics
Room 303S.288 (second floor of the Science Centre, Building 303S)
l.boyle@auckland.ac.nz; (09) 923-9045; 021 447-018

Want help with Stats?

Stats 101/101G/108 appointments

Book your preferred time with Leila here: **www.tinyURL.com/appt-stats**, or contact her directly (see above for her contact details).

Stats 101/101G/108 Workshops

One computing workshop, four exam prep workshops and four drop-in sessions are held during the second half of the semester.

Workshops are run in a relaxed environment and allow plenty of time for questions. In fact, this is encouraged! ☺

Please make sure you bring your calculator with you to all of these workshops!

No booking is required – just turn up to any workshop! You are also welcome to come along virtually on Zoom if you prefer. Search your emails for "Leila" to find the link – email Leila at l.boyle@auckland.ac.nz if you can't find it.

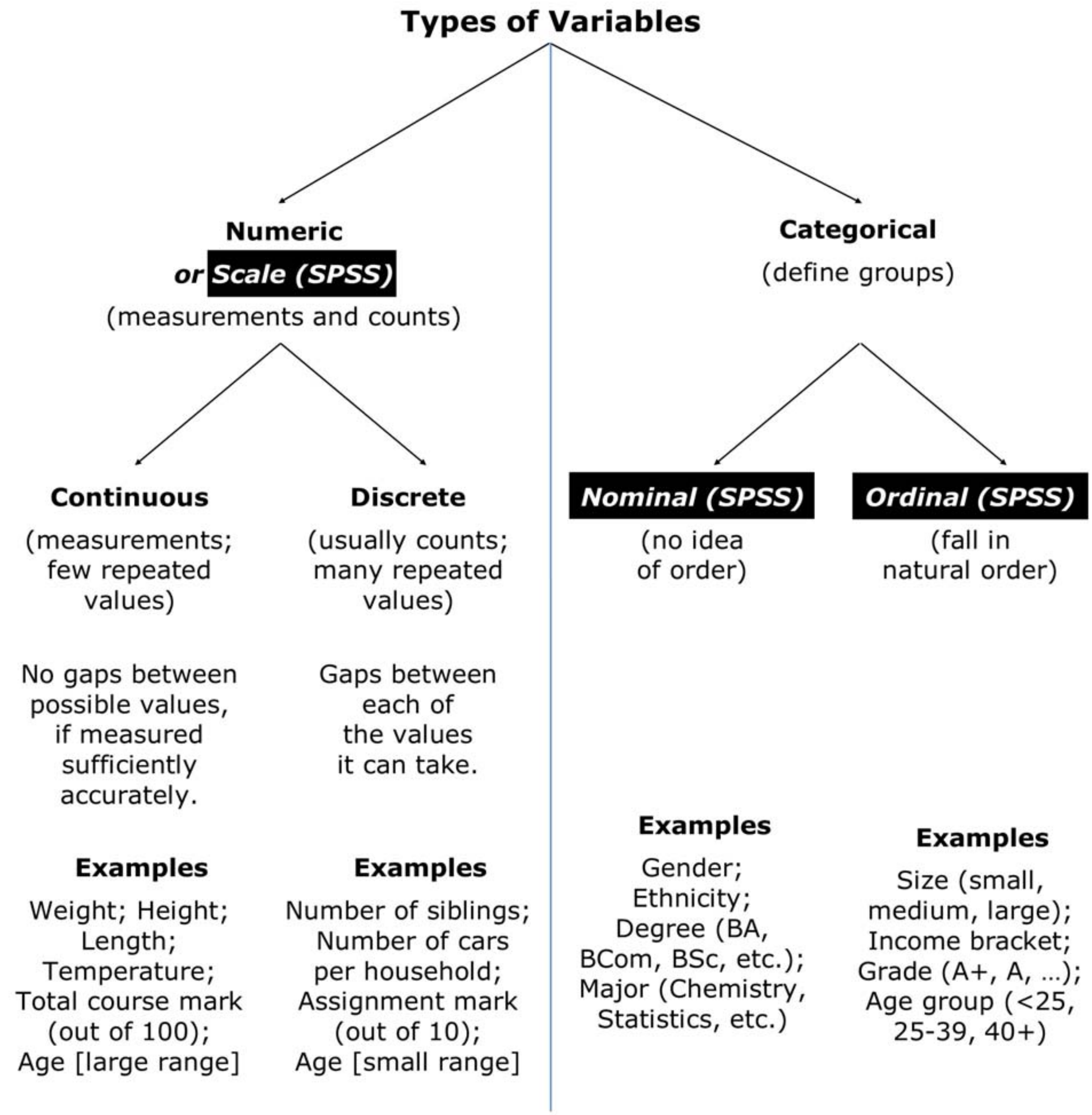
- **Computer workshop: *Hypothesis Tests in SPSS***
  www.tinyURL.com/stats-HTS
  Computing for Assignment 3 **–** covers the **computing** you need to do for **Questions 3 and 4** (iNZight plots & SPSS output). There are **six identical sessions**:
  - Friday 16 October, 3-4pm
  - Monday 19 October, 10-11am
  - Monday 19 October, 2-3pm
  - Tuesday 20 October, 4-5pm
  - Wednesday 21 October, 11am-midday
  - Wednesday 21 October, 3-4pm

- **Exam prep workshops**
  - ***Chi-Square Tests*** **www.tinyURL.com/stats-CST**
    Exam revision for Chapter 9 **–** Saturday 24 October, 1-4pm, LibB15 (useful exam prep and also useful for the **Chapter 9 Quiz** due at 11pm on Wednesday 28 October!)
  - ***Regression and Correlation*** **www.tinyURL.com/stats-RC**
    Exam revision for Chapter 10 **–** Saturday 31 October, 9.30am-12.30pm, LibB10 (useful exam prep and also useful for the **Chapter 10 Quiz** due at 11pm on Wednesday 4 November!)
  - ***Hypothesis Tests: Proportions*** **www.tinyURL.com/stats-HTP**
    Exam revision for Chapters 6 & 7 *(with a focus on proportions)* **–** Tuesday 3 November, 9.30am-12.30pm, LibB10 (useful exam prep)
  - ***Hypothesis Tests: Means*** **www.tinyURL.com/stats-HTM**
    Exam revision for Chapter 6, 7 & 8 **–** Tuesday 3 November, 1-4pm, LibB10 (useful exam prep)

- **Drop-in sessions**
  - Saturday 17 October, 9.30am-4pm, LibB10
  - Saturday 24 October, 9.30am-12.30pm, LibB15
  - Monday 26 October, 9.30am-4pm, LibB10
  - Saturday 31 October, 1-4pm, LibB10

Chi-Square Tests

This material builds on a number of workshops held earlier this semester, which you may or may not have attended. For **extracting a proportion / probability from a two-way table of counts**, see the ***Proportions and Proportional Reasoning*** workshop. For **quantifying the size of a single proportion / difference between two proportions**, see the ***Confidence Intervals: Proportions*** workshop and for **carrying out a hypothesis test for a single proportion / difference between two proportions**, see the ***Hypothesis Tests: Proportions*** workshop.



📖**Useful reference:** Chance Encounters, pages 40 – 42

Exploring relationships between two variables

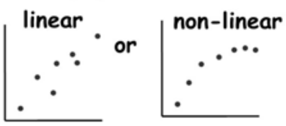
Two Numeric	Numeric versus Categorical	Two Categorical
Exploratory tools to use to explore relationship/s		
Scatter plot	Side-by-side plots on the same scale: any n { • Dot plots $n \geq 20$ { • Stem-and-leaf plots • Box plots $n \geq 50$ { • Histograms	Two-way table of counts and/or Bar graphs of proportions (on rows/columns)

Features to look for:

Two Numeric

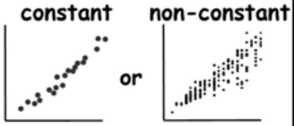
- Trend

  linear or non-linear



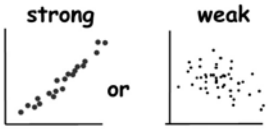
- Scatter

  constant or non-constant



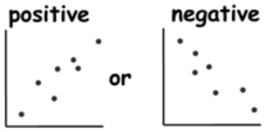
- Strength of relationship

  strong or weak

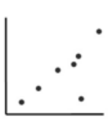


- Association

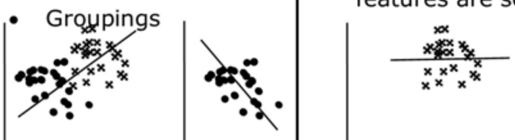
  positive or negative



- Outliers



- Groupings



All | Females only | Males only

Numeric versus Categorical

Compare the groups by looking at:

- Any group differences:
  - averages (centres) → medians, means
  - variability (spread) → IQRs, ranges
  - shapes
    - symmetric/skewed → left/negative, right/positive
    - modes → unimodal, bimodal, trimodal } not box!
- Details of individual groups:
  - outliers, gaps, clusters, groupings

Think about reasons *why* these differences, similarities and features are seen

Two Categorical

- Most common and least common combinations
- Differences in distributions (e.g. row and/or column bar graphs)

Chi-Square Tests

We use the **Chi-Square test for independence** to make inferences about relationships between two categorical variables.

The data for our **Chi-Square test for independence** is best presented in a **two-way table of counts.**

Two-Way Table of Counts – Chi-square test for **Independence**

- Can use when you have either **1 random sample** & **2 factors** of interest.

  or **2 or more independent random samples** & **1 factor** of interest

- **Hypotheses**

  H_0: the two factors are **independent**

  H_1: the two factors are **not independent**

- **Chi-Square Test Statistic**:

$$X_0^2 = \sum_{\text{all cells in the table}} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

see back page for the Formulae Sheet

where **observed** is the **observed count** for a particular cell in the table (comes from the data)

and **expected** is the **expected count** for a particular cell in the table, i.e. what we expect the count to be if the null hypothesis is true. (has to be calculated by hand/SPSS)

$$\text{Expected count in cell } (i, j) = \frac{R_i C_j}{n}$$

where R_i is the cell's row total
and C_j is the cell's column total
and n is the (grand) table total

Note that $\frac{(\text{observed} - \text{expected})^2}{\text{expected}}$ for a particular cell is referred to as the **cell contribution** to the Chi-square test statistic. Hence the Chi-square test statistic is the **sum** of all the cell contributions.

- Under the null hypothesis, we assume the sampling distribution of the test statistic is a **Chi-square(*df*)** distribution where ***df*** (degrees of freedom) are calculated by $df = (I - 1)(J - 1)$ where I is the number of rows in the table and J is the number of columns
- There will be evidence against the null hypothesis if there are relatively large differences between the observed and expected counts in one or more cells.
- As with the t-test and the F-test, the greater the magnitude or size of the test statistic, the stronger the evidence against the null hypothesis.

Requirements for the Chi-Square Test to be Valid

The Chi-square test is a large sample test and requires *n*, the (grand) table total, to be "large".

- For checking the validity of a Chi-square test, i.e., making sure that *n* is "large enough", most of the expected counts should be "large".
- The criteria we require are:
  - ✓ ***At least 80%*** of the **expected counts** must be ***5 or more***, **and**
  - ✓ ***Each* expected count** must be ***greater than 1***.
- SPSS tells you how many cells have an expected count less than five:

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	x_0^2	***df***	***P-value***
Likelihood Ratio			
Linear-by-Linear Association			
N of Valid Cases	***n***		

0 cells (.0%) have expected count less than 5. The minimum expected count is ***some number***.

Another way of looking at it – Chi-Square Test for Homogeneity

The hypotheses can also written as statements of **homogeneity** (sameness):

H_0: the distribution of variable 1 is **the same** for each level of variable 2.

H_1: the distribution of variable 1 is **not the same** at all levels of variable 2.

The sampling situation determines which one of the two variables is variable 1 and which one is variable 2. There are two possibilities:

1. If **2 or more independent samples** are taken from **different populations** and sample members are **categorised by a variable** then the null hypothesis can be a statement of homogeneity among the populations, (i.e., variable 1 categorises the sample members and variable 2 defines the populations):

   H_0: The distribution of variable 1 is the same for each population

   This test for independence is often called a test for homogeneity.

2. If a **single random sample** has been **cross-classified** by variable 1 and variable 2 then the null hypothesis can be a statement of homogeneity (sameness) and it doesn't matter which variable is variable 1 and which variable is variable 2.

We determine whether to use hypotheses of independence or hypotheses of homogeneity by considering the number of samples taken and the number of categorical variables / factors to be tested.

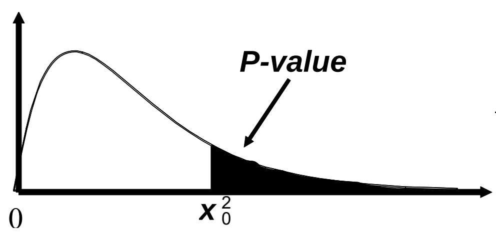
Performing a **Chi-square test** – a step-by-step guide:

1. Identify which **situation** you have (either 1 sample or 2 or more independent samples) and therefore which type of hypotheses are appropriate (independence or homogeneity or either/both).
2. State the **null hypothesis**, H_0.
3. State the **alternative hypothesis**, H_1.
4. Calculate the **expected count** for each cell.
5. Calculate the **cell contribution** for each cell.
6. Determine the Chi-square **test statistic.**
7. Find the **degrees of freedom**.
8. Estimate the ***P*-value.** (Will be provided)
9. **Interpret** the *P*-value.
10. Give an overall **conclusion.**

***Step 8*: - The *P*-value:**

- measures the strength of evidence **against** H_0.
- is calculated using the Chi-square test statistic and the appropriate *df*.

As with the *t*-test and the *F*-test, the *P-value* is the conditional probability of observing a test statistic as extreme as that observed or more so, **given that** the null hypothesis , H_0, is true.



- *P-value* = $\text{pr}(X^2 \geq x_0^2)$, where $X^2 \sim \text{Chi-square}(df)$

The **bigger** the test statistic is, the **stronger** the evidence **against** H_0.

- If you have a **small** *P-value*, note which **cell/s contributed the most** to the **large** Chi-square test statistic, i.e. look for cell/s where there are **large differences** between the **observed** count/s and the **expected** count/s.

In the exam situation, the *P-value* will be provided.

Step 9

- We interpret the *P-value* as a **description** of the **strength of evidence against the null hypothesis, H_0**. The **smaller** the *P-value*, the **stronger** the evidence against H_0:

P-value	Evidence against H_0
> 0.10	None
≈ 0.07	Weak
≈ 0.05	Some
≈ 0.01	Strong
≤ 0.001	Very Strong

- An alternative approach often found in research articles and news items is to describe the test result as (statistically) significant or not significant. A test result is said to be significant when the *P-value* is "small enough"; usually people say a *P-value* is "small enough" if it is less than 0.05 (5%):

Testing at a 5% level of significance:

P-value	Test result	Action
< 0.05	Significant	Reject H_0 in favour of H_1
> 0.05	Nonsignificant	Do not reject H_0

Testing can be done at any level of significance; 1% is common but 5% is what most researchers use.

The level of significance can be thought of as a false alarm error rate, i.e. it is the proportion of times that the null hypothesis will be rejected when it is actually true (which can result in action being taken when really no action should be taken).

Thus, a statistically significant result means that a study has produced a "small" *P-value* (usually < 5%).

Practice Questions

Questions 1 and **2** are about the following information.

The table below was created from a random sample taken from the US population and then cross-classified by **Job Satisfaction** and **Income ($US)**:

	Job Satisfaction				
Income ($US)	Very Dissatisfied	Little Dissatisfied	Moderately Satisfied	Very Satisfied	Totals
< 6,000	20	24	80	82	206
6,000 – 15,000	22	38	104	125	289
15,000 – 25,000	13	28	81	113	235
> 25,000	7	18	54	92	171
Totals	62	108	319	412	901

A Chi-square test was conducted and the SPSS output is given below.

			Age (years)				
			Very Dissatisfied	Little Dissatisfied	Moderately Satisfied	Very Satisfied	Total
Income ($US)	< 6,000	Count	20	24	80	82	206
		Expected Count	14.2	24.7	72.9	94.2	206.0
		Cell contribution	2.393	0.019	0.684	1.579	
	6,000 – 15,000	Count	22	38	104	125	289
		Expected Count	19.9	34.6	102.3	132.2	289.0
		Cell contribution	0.225	0.326	0.028	0.387	
	15,000 – 25,000	Count	13	28	81	113	235
		Expected Count	16.2	28.2	83.2	107.5	235.0
		Cell contribution	0.622	0.001	0.058	0.286	
	> 25,000	Count	7	18	54	92	171
		Expected Count	11.8	20.5	60.5	78.2	171.0
		Cell contribution	1.931	0.304	0.707	2.438	
Total		Count	41	143	186	130	500
		Expected Count	41.0	143.0	186.0	130.0	500.0

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	11.989[a]	++	.214
Likelihood Ratio	12.037	++	.211
Linear-by-Linear Association	39.546	++	.002
N of Valid Cases	500		

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 11.77.

Note: Some values have been replaced by ++.

1. Which one of the following statements is **false**?
   (1) The Chi-square test can be used to test for independence between *Job Satisfaction* and *Income*.
   (2) The degrees of freedom for this Chi-square test are 9.
   (3) When testing for independence between *Job Satisfaction* and *Income*, the null hypothesis tested is that *Income* is independent of *Job Satisfaction*.
   (4) For Chi-square tests, a large value for the Chi-square statistic provides evidence against the null hypothesis.
   (5) When testing for independence between *Job Satisfaction* and *Income*, a large *P*-value provides evidence against independence.

2. On the basis of the information from the Chi-square test, which one of the following statements is **false**?
   (1) Small departures from independence between *Income* and *Job Satisfaction* cannot reliably be detected with small samples.
   (2) The Chi-square test works best when there are small counts in several of the cells.
   (3) The sampling situation for this data is "One sample cross-classified by two factors."
   (4) Large departures from independence between *Income* and *Job Satisfaction* can reliably be detected with large samples.
   (5) This data provides no evidence of a relationship between *Income* and *Job Satisfaction*.

3. Which **one** of the following statements is **false**?
   (1) A Chi-square test for independence is used to carry out formal analyses on data presented in two-way tables of counts.
   (2) The Chi-square test statistic is a measure of the difference between what we see in the data and what we would expect to see if the null hypothesis was true.
   (3) If all of the expected counts in a table are less than 5 then we would have no concerns with the validity of a Chi-square test carried out on these data.
   (4) If, for one or more cells in a table of counts, there are relatively large differences between the observed counts and the expected counts under the null hypothesis, then the data provides evidence against the null hypothesis.
   (5) The *P-value* in a Chi-square test is the probability, assuming that the null hypothesis is true, of observing data at least as unusual as that obtained in the data on which the test is carried out.

Questions 4 to **10** refer to the following **KiwiSaver Data** information.

KiwiSaver is a voluntary, work-based savings initiative that was set up by the New Zealand Government in July 2007. The main purpose of KiwiSaver is to assist New Zealanders with their long-term saving for retirement.

In March 2010 the research company *UMR* conducted a telephone poll of 750 New Zealanders aged 18 years and over on issues related to KiwiSaver membership and knowledge of KiwiSaver schemes. One of the questions asked was:

> *Are you a member of KiwiSaver? — Yes, No*

Those who responded *'Yes'* to the question above were then asked this question:

> *How much would you say you know about the KiwiSaver scheme you are invested in? — a lot, a fair amount, not much or hardly anything?*

The respondents were also categorised into age groups. The results are shown in Table 1.

	Knowledge				
Age Group (years)	A lot	A fair amount	Not much	Hardly anything	**Total**
Under 30	8	17	20	15	60
30–44	7	22	25	21	75
45–59	6	30	31	17	84
60 plus	3	16	13	2	34
Total	24	85	89	55	253

Table 1: Knowledge about KiwiSaver scheme

A Chi-square test was carried out to see if we can claim that the pattern in knowledge about KiwiSaver depends on age group. The results of the Chi-square test are shown in Table 2.

Age Group (years) * Knowledge Crosstabulation

			Knowledge				Total
			A lot	A fair amount	Not much	Hardly anything	
Age Group (years)	Under 30	Count	8	17	20	15	60
		Expected Count	5.7	20.2	21.1	13.0	60.0
	30-44	Count	7	22	25	21	75
		Expected Count	7.1	++	26.4	++	75.0
	45-59	Count	6	30	31	17	84
		Expected Count	8.0	++	29.5	++	84.0
	60 plus	Count	3	16	13	2	34
		Expected Count	3.2	11.4	12.0	7.4	34.0
Total		Count	24	85	89	55	253
		Expected Count	24.0	85.0	89.0	55.0	253.0

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	10.244	++	.331
Likelihood Ratio	11.518	++	.242
Linear-by-Linear Association	1.675	++	.196
N of Valid Cases	253		

Note: Some values have been replaced by ++.

Table 2: SPSS output for knowledge about KiwiSaver scheme

4. Which **one** of the following statements is a correct **null** hypothesis for the Chi-square test described above?

   (1) The underlying distribution of knowledge about KiwiSaver is associated with age group.

   (2) The underlying distribution of age group is not the same for each level of knowledge about KiwiSaver.

   (3) The underlying distribution of knowledge about KiwiSaver is independent of age group.

   (4) The underlying distribution of knowledge about KiwiSaver is related to age group.

   (5) The underlying distribution of knowledge about KiwiSaver is not the same for each age group.

5. Which **one** of the following statements gives the **best** justification concerning the validity of this Chi-square test?

   (1) Because there are 5 cells out of 16 with an observed count of less than 10, the results of this test are **not** valid.

   (2) Because there is 1 cell out of 16 with an expected count less than 5, the results of this test are **not** valid.

   (3) Because there are no cells with an observed count of less than 1, there is **no** cause for any concern with the validity.

   (4) Because there are only 2 cells out of 16 with an observed count of less than 5, there is **no** cause for any concern with the validity.

   (5) Because there is only 1 cell out of 16 with an expected count less than 5 and none less than 1, there is **no** cause for any concern with the validity.

Questions 6 to **10** assume that the use of the Chi-square test is appropriate.

(Note: This assumption may not be true.)

6. Consider the cell in Table 2, page 12, for those **aged 45 to 59** who say they know **hardly anything** about the KiwiSaver scheme they are invested in. Under the null hypothesis, the expected count for this cell is approximately:

   (1) 17.0 (4) 15.8

   (2) 18.3 (5) 21.0

   (3) 13.8

7. Consider the cell in Table 2, page 12, for those **aged 45 to 59** who say they know **a lot** about the KiwiSaver scheme they are invested in. This cell's contribution to the Chi-square test statistic, to 2 decimal places, is:

   (1) 0.25 (4) 0.33

   (2) 0.67 (5) 0.50

   (3) 0.06

8. The *P-value* for this Chi-square test is calculated by:

   (1) pr($0 \le \chi^2 \le 10.244$) where $\chi^2 \sim$ Chi-square(df = 9)

   (2) pr($\chi^2 \ge 10.244$) where $\chi^2 \sim$ Chi-square(df = 16)

   (3) pr($\chi^2 \ge 10.244$) where $\chi^2 \sim$ Chi-square(df = 9)

   (4) 2 × pr($\chi^2 \ge 10.244$) where $\chi^2 \sim$ Chi-square(df = 16)

   (5) 2 × pr($\chi^2 \ge 10.244$) where $\chi^2 \sim$ Chi-square(df = 9)

9. Based on the results of this Chi-square test, which **one** of the following statements is **true**?

   (1) If the test statistic is at least as large as 10.244, then the probability that the null hypothesis is true is approximately 0.331.

   (2) The probability that the alternative hypothesis is true is approximately 0.331.

   (3) If the test statistic is at least as large as 10.244, then the probability that the alternative hypothesis is true is approximately 0.331.

   (4) If the null hypothesis is true, then the probability of getting a test statistic at least as large as 10.244 is approximately 0.331.

   (5) If the alternative hypothesis is true, then the probability of getting a test statistic at least as large as 10.244 is approximately 0.331.

10. Based on the results of this Chi-square test, which **one** of the following statements is **true**?

    (1) There is no evidence against the null hypothesis.

    (2) There is very strong evidence for the alternative hypothesis.

    (3) There is no evidence against the alternative hypothesis.

    (4) There is no evidence for the null hypothesis.

    (5) There is very strong evidence for the null hypothesis.

Questions 11 to **18** refer to the following information:

In October 2010 Research New Zealand conducted a telephone poll of 500 New Zealanders aged 15 years and over to find out whether they expected the outlook for the economy would get better or worse in the following 12 months and whether they felt their own financial circumstances would get better or worse over the same period. One of the questions asked was:

Do you think that your own personal financial circumstances will get better or worse over the next 12 months?

The 500 respondents were asked to choose one of four possible responses: *Better, Stay the same, Worse* or *Don't know*. The 500 respondents were then cross-classified according to their response to the question and their age. The results are shown in Table 3.

	Age (years)				
Response	15–24	25–44	45–64	65+	**Total**
Better	8	42	48	7	105
Stay the same	18	58	76	78	230
Worse	12	37	58	39	146
Don't know	3	6	4	6	19
Total	41	143	186	130	500

Table 3: Response by age group

A Chi-square test is conducted to see whether the pattern of the responses changes with age. Note: Some values have been replaced by ++:

			Age (years)				Total
			15 – 24	25 – 44	45 – 64	65+	
Response	Better	Count	8	42	48	7	105
		Expected Count	8.61	30.03	39.06	27.30	105.0
		Cell contribution	0.043	4.771	2.046	15.095	
	Stay the same	Count	18	58	76	78	230
		Expected Count	+++	+++	85.56	59.80	230.0
		Cell contribution	0.039	0.920	1.068	5.539	
	Worse	Count	12	37	58	39	146
		Expected Count	+++	+++	54.31	37.96	146.0
		Cell contribution	0.000	0.542	+++	+++	
	Don't Know	Count	3	6	4	6	19
		Expected Count	1.56	5.43	7.07	4.94	19.0
		Cell contribution	1.335	0.059	1.332	0.227	
Total		Count	41	143	186	130	500
		Expected Count	41.0	143.0	186.0	130.0	500.0

Table 4: Chi-square test output for Response against Age

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	33.296	++	.000
Likelihood Ratio	38.630	++	.000
Linear-by-Linear Association	3.880	++	.049
N of Valid Cases	500		

Table 5: Chi-square test output for Response against Age

11. Which **one** of following is a correct pair of hypotheses for this Chi-square test?

(1) H_0: Response is not related to age group.
H_1: Response is related to age group.

(2) H_0: The underlying distribution of response is the same for all four age groups.
H_1: The underlying distribution of response is different for all four age groups.

(3) H_0: Response does depend on age group.
H_1: Response does not depend on age group.

(4) H_0: Response is associated with age group.
H_1: Response is not associated with age group.

(5) H_0: The underlying distribution of response is not the same for all four age groups.
H_1: The underlying distribution of response is the same for all four age groups.

12. Which **one** of following statements gives the **best** justification for us to have no concerns about the validity of this Chi-square test?

(1) The sample size, n = 500, is greater than 30.

(2) Two cells have an observed count less than 5 and one of these cells has an expected count that is also less than 5.

(3) None of the cells has an expected count of less than 1.

(4) Only two cells have an expected count of less than 5, both of which are greater than 1.

(5) At least 80% of the cells have an expected count of 5 or more.

Questions 13 to **18** assume that the use of the Chi-square test is appropriate.

13. The number of degrees of freedom, *df*, for this Chi-square test is:

(1) 16
(4) 25
(2) 15
(5) 9
(3) 8

14. Consider the cell in Table 4, page 15, for 25–44 year-olds who thought that their personal financial circumstances will get worse over the following 12 months.

Under the null hypothesis, the expected count for this cell is approximately:

(1) 23.56
(4) 37
(2) 52.34
(5) 29.37
(3) 41.76

15. Consider the cell in Table 4, page 15, for 45–64 year-olds who thought that their personal financial circumstances will get worse over the following 12 months.

This cell's contribution to the Chi-square test statistic is:

(1) 0.235
(4) 0.064
(2) 0.213
(5) 0.251
(3) 0.068

16. The *P-value* for this Chi-square test is calculated by:

(1) $2 \times \text{pr}(\chi^2 \geq 33.296)$ where $\chi^2 \sim$ Chi-square(*df*)

(2) $\text{pr}(0 \leq \chi^2 \leq 33.296)$ where $\chi^2 \sim$ Chi-square(*df*)

(3) $2 \times \text{pr}(\chi^2 \leq 33.296)$ where $\chi^2 \sim$ Chi-square(*df*)

(4) $\text{pr}(\chi^2 \geq 33.296)$ where $\chi^2 \sim$ Chi-square(*df*)

(5) $\text{pr}(\chi^2 \leq 33.296)$ where $\chi^2 \sim$ Chi-square(*df*)

17. Which **one** of the following statements gives the **best** explanation for the small *P-value* of 0.000?

(1) There are far more people aged 15–24 who felt their personal financial circumstances will get better in the following 12 months than would have been expected when the null hypothesis is true.

(2) There are far fewer people aged 15–24 who felt their personal financial circumstances will get better in the following 12 months than would have been expected when the null hypothesis is true.

(3) There are far fewer people aged 65 years or older who felt their personal financial circumstances will get better in the following 12 months than would have been expected when the null hypothesis is true.

(4) The number of people aged 15–24 years who felt their personal financial circumstances will stay the same in the following 12 months and the number of people expected to be in this category when the null hypothesis is true, are quite similar.

(5) There is almost no difference between the number of people aged 15–24 years who felt their personal financial circumstances will get worse in the following 12 months and the number of people expected to be in this category when the null hypothesis is true.

18. Which **one** of the following statements is a **valid** conclusion at the 5% level of significance?

(1) The underlying pattern of response does not change for different age groups.

(2) It would be surprising to obtain the observed differences in the patterns of response for different age groups if there is a difference in the underlying patterns of response.

(3) The observed differences between age groups in the patterns of response could be just due to sampling variation alone.

(4) The observed differences between age groups in the patterns of response are more than just sampling variation alone.

(5) It would not be surprising to obtain the observed differences in the patterns of response for different age groups if there is no difference in the underlying patterns of response.

Commuter Altruism Study

Questions 19 to **25** refer to the following information.

In a study from the University of Queensland[1], researchers were interested in altruism (unselfishness) in society. They conducted a study into who is unselfish and who is selfish at selected traffic intersections, analysing data on 959 commuters at intersections where drivers could choose to let another commuter from a side road enter the main road (they were unselfish) or drivers could choose to keep going and save themselves a few seconds (they were selfish).

Two of the variables recorded are defined as follows:

Status Based on an estimate of the worth of the car:

- Low ($25,000 or less)
- Medium ($25,001 to $60,000)
- High (Over $60,000)

Altruistic choice A measure of altruism according to whether or not the driver chooses to let another commuter in:

- Unselfish (Driver lets another commuter in)
- Selfish (Driver does not let another commuter in)

The 959 drivers were cross-classified according to **Altruistic choice** and **Status**.

The results are shown in Table 6 and Figure 1.

	Status			
Altruistic choice	Low	Average	High	**Total**
Unselfish	116	192	71	379
Selfish	247	217	116	580
Total	363	409	187	959

Table 6: Altruistic choice of commuters

[1] Institute for the Study of Labour (IZA), April 2011, http://ftp.iza.org/dp5648

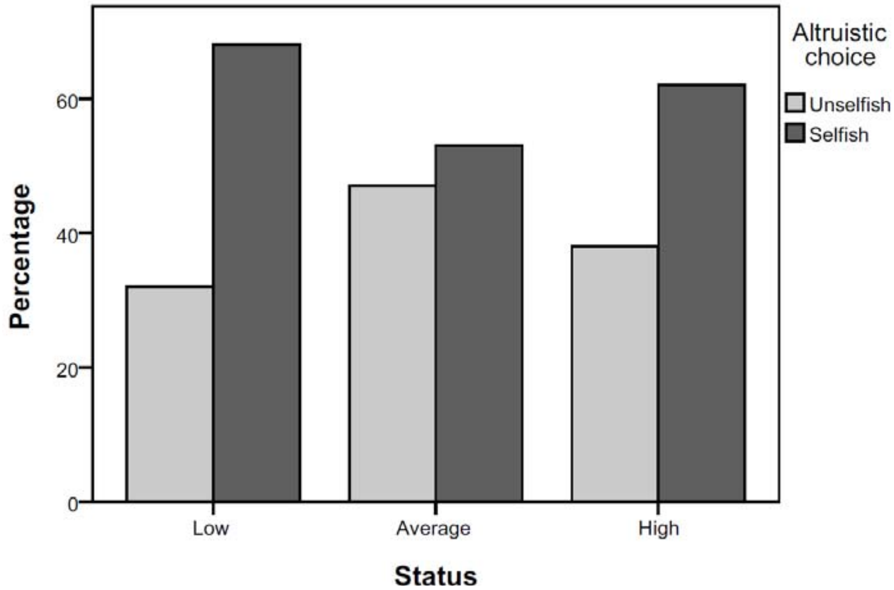


Figure 1: Bar graph of choice on letting another commuter in

A Chi-square test was conducted to see if the distribution of altruistic choice depends on social status (as measured by the worth of their car). The results of the Chi-square test are shown in Table 7.

			Status			
			Low	Average	High	Total
Altruistic choice	Unselfish	Count	116	192	71	379
		Expected Count	143.5	161.6	73.9	379.0
		Cell contribution	5.256	5.703	0.114	
	Selfish	Count	247	217	116	580
		Expected Count	219.5	247.4	113.1	580.0
		Cell contribution	3.434	3.727	++	
Total		Count	363	409	187	959
		Expected Count	363.0	409.0	187.0	959.0

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	18.308	++	.000
Likelihood Ratio	18.382	++	.000
Linear-by-Linear Association	4.868	++	.027
N of Valid Cases	959		

Note: Some values have been replaced by ++.

Table 7: Chi-square test output

Questions 19 to **25** refer to the **Commuter Altruism Study** information, given above (on pages 19 and 20).

19. The variable **Status** is **best** described as:

    (1) an ordinal categorical variable.

    (2) a categorical numeric variable.

    (3) a continuous numeric variable.

    (4) a dependent variable.

    (5) a discrete numeric variable.

Note: Questions 20 to **25** assume that the 959 drivers form a simple random sample of Australian drivers.

Questions 20 to **25** refer to the Chi-square test described on page 20.

20. Which **one** of following give **correct** hypotheses for this Chi-square test?

    (1) H_0: The underlying distribution of altruistic choice is the same for all three status groups.
    H_1: The underlying distribution of altruistic choice is different for each status group.

    (2) H_0: The underlying distribution of altruistic choice is the same for all three status groups.
    H_1: The underlying distribution of altruistic choice is not the same for all three status groups.

    (3) H_0: The underlying distribution of altruistic choice is not the same for all three status groups.
    H_1: The underlying distribution of altruistic choice is the same for all three status groups.

    (4) H_0: The underlying distribution of altruistic choice is different for each status group.
    H_1: The underlying distribution of altruistic choice is the same for all three status groups.

    (5) H_0: The underlying distribution of altruistic choice depends on the status group.
    H_1: The underlying distribution of altruistic choice does not depend on the status group.

21. Which **one** of the following gives the **best** justification concerning the validity of this Chi-square test?

    (1) Because none of the cells have an expected count less than 5, there are no concerns about the validity of this Chi-square test.

    (2) Because none of the cells have an expected count less than 1, there are no concerns about the validity of this Chi-square test.

    (3) Because fewer than 80% of the cells have cell contributions greater than 5 there are concerns about the validity of this Chi-square test.

    (4) Because none of the cells have an observed count less than 5, there are no concerns about the validity of this Chi-square test.

    (5) Because at least one of the cells has a cell contribution less than 1 there are concerns about the validity of this Chi-square test.

Questions 22 to **25** assume that the use of the Chi-square test is appropriate.

(Note: This assumption may not be correct.)

22. Consider the cell in Table 7, page 20, for those who are categorised as **selfish** and **high** status.

    This cell's contribution to the Chi-square test statistic, to 3 decimal places, is:

    (1) 2.900

    (2) −2.900

    (3) 0.074

    (4) 0.073

    (5) 3.100

23. The number of degrees of freedom, *df*, for this Chi-square test is:

(1) 6 (4) 2

(2) 958 (5) 959

(3) 5

24. The *P-value* for this Chi-square test is calculated by:

(1) pr($\chi^2 \geq 959.0$) where $\chi^2 \sim$ Chi-square(*df*)

(2) pr($\chi^2 \leq 959.0$) where $\chi^2 \sim$ Chi-square(*df*)

(3) pr($\chi^2 \leq 18.308$) where $\chi^2 \sim$ Chi-square(*df*)

(4) pr($0 \leq \chi^2 \leq 959.0$) where $\chi^2 \sim$ Chi-square(*df*)

(5) pr($\chi^2 \geq 18.308$) where $\chi^2 \sim$ Chi-square(*df*)

25. Which **one** of the following gives the **best** reason for the *P-value* of 0.000?

(1) For all categories there are roughly the number of drivers you would expect if **Altruistic choice** and **Status** were unrelated.

(2) For those drivers whose choice was unselfish, there are far fewer in the low status category and far more in the average status category than would be expected if **Altruistic choice** and **Status** were related.

(3) For those drivers whose choice was unselfish, there are far more in the low status category and far fewer in the average status category than would be expected if **Altruistic choice** and **Status** were unrelated.

(4) For those drivers in the high status category there are roughly the number of drivers that you would expect to get for both the unselfish and the selfish choice, if **Altruistic choice** and **Status** were related.

(5) For those drivers whose choice was unselfish, there are far fewer in the low status category and far more in the average status category than would be expected if **Altruistic choice** and **Status** were unrelated.

Questions 26 to **32** refer to the following information.

The University of Otago Injury Prevention Research Unit published a report titled *Road traffic practices among a cohort of young adults in New Zealand*. The aim of the study was to describe the road safety practices of young adults in New Zealand. Face-to-face interviews were conducted with 21-year-olds who were born in Dunedin. The report concluded that unsafe road practices, especially among males, were unacceptably high.

One area of the study investigated the wearing of seat belts. Some results are given in Table 8 below, a two-way table of counts for seat belt usage by rear seat passengers:

	Usage				
Gender	Always	Nearly Always	Sometimes	Never	**Total**
Female	138	79	139	107	463
Male	103	66	152	161	482
Total	241	145	291	268	945

Table 8: Self-reported seat belt usage by rear seat passengers

A Chi-square test was conducted to investigate whether there was a difference in **Usage** distribution between females and males. SPSS output is given in Table 9 below.

			Usage				Total
			Always	Nearly Always	Sometimes	Never	
Gender	Female	Count	138	79	139	107	463
		Expected Count	118.1	++	142.6	++	++
		Cell contribution	++	++	0.091	4.497	
	Male	Count	103	66	152	161	482
		Expected Count	122.9	++	148.4	++	++
		Cell contribution	3.222	0.865	0.087	4.320	
Total		Count	241	145	291	268	945
		Expected Count	241.0	145.0	291.0	268.0	945.0

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	17.335[a]	++	.001
Likelihood Ratio	17.422	++	.001
Linear-by-Linear Association	16.754	++	.001
N of Valid Cases	945		

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 71.0.

Note: Some values have been replaced by ++.

Table 9: Self-reported seat belt usage by rear seat passengers

26. For this investigation, the null hypothesis is:

    (1) H_0: The distribution of Usage is different for females and males.

    (2) H_0: The factors Gender and Usage are associated.

    (3) H_0: $p_1 = p_2 = p_3 = p_4$ where p_i is the proportion of 21-year-olds in each Usage group.

    (4) H_0: The distribution of Usage is the same for females and males.

    (5) H_0: The factors Female and Male are dependent.

27. The expected cell count, under the null hypothesis, for those 21-year-old **males** who **never** wear a rear seat belt is:

    (1) 137.1

    (2) 130.3

    (3) 136.8

    (4) 131.3

    (5) 136.7

28. The **degrees of freedom** for this Chi-square test is:

    (1) 6

    (2) 4

    (3) 8

    (4) 3

    (5) 2

29. Consider the cell for **Female** and **Always**. This cell's **contribution** to the Chi-square test statistic value of 17.335 is:

    (1) 2.9

    (2) 0.1

    (3) 1.6

    (4) 0.2

    (5) 3.4

30. The *P-value* for this Chi-square test is calculated by:

(1) $2 \times \text{pr}(\chi^2 \geq 17.335)$ where $\chi^2 \sim \text{Chi-square}(df)$

(2) $2 \times \text{pr}(\chi^2 \leq 17.335)$ where $\chi^2 \sim \text{Chi-square}(df)$

(3) $\text{pr}(\chi^2 \leq 17.335)$ where $\chi^2 \sim \text{Chi-square}(df)$

(4) $\text{pr}(0 \leq \chi^2 \leq 17.335)$ where $\chi^2 \sim \text{Chi-square}(df)$

(5) $\text{pr}(\chi^2 \geq 17.335)$ where $\chi^2 \sim \text{Chi-square}(df)$

31. Which **one** of the following statements regarding the *P-value* of 0.001 is **true**?

(1) Such a small *P-value* indicates that there must be a big difference between the **Female** and **Male Usage** distributions.

(2) Such a small *P-value* indicates that the alternative hypothesis must be true.

(3) The probability that the null hypothesis is false is 0.001.

(4) If the null hypothesis for this test is true, then the probability of getting a test statistic at least as large as 17.335 is 0.001.

(5) The probability that the null hypothesis is true is 0.001.

32. Which **one** of the following statements is **false**?

(1) One of the main reasons for such a small *P-value* in this test is because of the relatively small number of **Males** who said that they were **Always** users of rear seat belts.

(2) If the Chi-square test statistic had been 27.000 instead of 17.335, then the resulting *P-value* would have been smaller than 0.001.

(3) One of the main reasons for such a small *P-value* in this test is because of the relatively large number of **Males** who said that they were **Sometimes** users of rear seat belts.

(4) If one of the cells had an expected count of less than 1, then it would have been unwise to interpret the output from this test.

(5) The sum of the expected counts for **Males** is 482 and the sum of the expected count for **Females** is 463.

Questions 33 to **38** are about the following information.

A market research company interviewed 299 randomly selected car owners in Auckland. Each car owner filled out a questionnaire and this information was used to classify each person as cautious conservative (CC), middle-of-the-road (MR), or confident explorer (CE). At the same time, each person was asked to give an overall opinion of small cars.

	Self-Perception			
Opinion of Small Cars	CC	MR	CE	Totals
Favourable	79	58	49	186
Neutral	10	8	9	27
Unfavourable	10	34	42	86
Totals	99	100	100	299

Table 10: Opinion of small cars by self-perception

The market research company would like to investigate whether a person's opinion of small cars was the same regardless of their self-perception as a driver.

33. The **correct hypotheses** for the market research company's investigation are:

(1) H_0: Self-perception and opinion of small cars are related.
H_1: Self-perception and opinion of small cars are independent.

(2) H_0: For every level of opinion of small cars the distribution of self-perception is the same.
H_1: The distribution of self-perception differs for some levels of opinion of small cars.

(3) H_0: The distribution of self-perception differs for some levels of opinion of small cars.
H_1: For every level of opinion of small cars the distribution of self-perception is the same.

(4) H_0: Self-perception and opinion of small cars are associated.
H_1: Self-perception and opinion of small cars are not associated.

(5) H_0: All treatment means are the same.
H_1: There is at least one treatment mean that differs from the remaining population means.

Suppose that the market research company was also interested in investigating if there was an association between the opinion of small cars and a driver's self-perception. A Chi-square test was conducted and the results obtained from SPSS are given in Table 11 below. Some values have been replaced with ++.

			Self-Perception			Total
			CC	MR	CE	
Opinion of Small Cars	Favourable	Count	79	58	49	186
		Expected Count	61.6	62.2	62.2	186.0
		Cell contribution	4.924	0.285	++	
	Neutral	Count	10	8	9	27
		Expected Count	++	++	++	27.0
		Cell contribution	0.126	++	0.000	
	Unfavourable	Count	10	34	42	86
		Expected Count	++	++	++	86.0
		Cell contribution	++	0.954	6.092	
Total		Count	99	100	100	299
		Expected Count	99.0	100.0	100.0	299.0

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	27.289[a]	++	.000
Likelihood Ratio	30.327	++	.000
Linear-by-Linear Association	24.391	++	.000
N of Valid Cases	299		

Table 11: Opinion of small cars by self-perception

34. The **expected count** for ***middle-of-the-roaders*** who have an ***unfavourable*** *opinion of small cars* is:
    (1) 285.2
    (2) 27.4
    (3) 0.95
    (4) 28.8
    (5) 34

35. Suppose the expected count for *cautious conservatives* who have an *unfavourable opinion of small cars* is 28.5. The **cell contribution** for ***cautious conservatives*** who have an ***unfavourable*** *opinion of small cars* is:

    (1) 12.0

    (2) 0.42

    (3) −0.65

    (4) 34.1

    (5) 1.14

36. The **degrees of freedom** for the above test would be:

    (1) 9

    (2) 6

    (3) 4

    (4) 5

    (5) 1

37. The *P-value* for this Chi-square test is calculated by:

    (1) pr($\chi^2 \geq 27.289$) where $\chi^2 \sim$ Chi-square(*df*)

    (2) 2 × pr($\chi^2 \leq 27.289$) where $\chi^2 \sim$ Chi-square(*df*)

    (3) pr($\chi^2 \leq 27.289$) where $\chi^2 \sim$ Chi-square(*df*)

    (4) 2 × pr($\chi^2 \geq 27.289$) where $\chi^2 \sim$ Chi-square(*df*)

    (5) pr($0 \leq \chi^2 \leq 27.289$) where $\chi^2 \sim$ Chi-square(*df*)

38. From the results of the above test the **best** interpretation would be:

    (1) There is very strong evidence that the distribution of self-perception of a driver and their opinion of a small car are the same.

    (2) There is no evidence that the distribution of self-perception of a driver and their opinion of a small car are not the same.

    (3) There is very strong evidence that the self-perception of a driver and their opinion of a small car are not related.

    (4) There is no evidence that the self-perception of a driver and their opinion of a small car are not related.

    (5) There is very strong evidence that the self-perception of a driver and their opinion of a small car are related.

39. Which **one** of the following statements is **true**:

    (1) If, for all cells in a table of counts, there are relatively small differences between the observed counts and the expected counts under the null hypothesis, then the data provides evidence against the null hypothesis.

    (2) The greater the value of the Chi-square test statistic, the larger the *P-value*.

    (3) For a Chi-square test to be valid the total count in the table, *n*, is required to be small.

    (4) The *P-value* in a Chi-square test is the probability, given that the null hypothesis is true, of obtaining a test statistic as extreme, or less so, as that observed.

    (5) If the *P-value* is small, then the cells with the largest contributions to the test statistic show which cells have observed counts that are far different (relatively) from those expected under the null hypothesis.

40. The **most appropriate** plot to use for analysing proportions that are to be tested using Chi-square tests is:

    (1) Scatterplot

    (2) Side-by-side dot plot

    (3) Side-by-side box plot

    (4) Table of counts

    (5) Histogram

Questions 41 to **45** are about the following information.

The data in the following table came from a study of predictors of social distress among 245 American third and fifth grade children by Crick and Ladd reported in *Developmental Psychology*, 1993. One aim of the study was to determine whether there was a relationship between the level of social distress and the peer status of the child.

Level of Social Distress by Peer Status

Level of Social Distress	**Peer Status**					Total
	Popular	Average	Neglected	Rejected	Controversial	
High	8	19	10	26	2	65
Low	41	57	32	33	17	180
Total	49	76	42	59	19	245

41. Suppose it is appropriate to conduct a Chi-square test for independence (**Note**: This may **not** be correct). The null and alternative hypotheses for this test are:

(1) H_0: The level of social distress for the child is not independent of the peer status of the child.
H_1: The level of social distress for the child is independent of the peer status of the child.

(2) H_0: The level of social distress for the child is independent of the peer status of the child.
H_1: The level of social distress for the child is not independent of the peer status of the child.

(3) H_0: The means of the social distress levels for the child are the same for each peer status factor.
H_1: The means of the social distress levels for the child are not the same for each peer status factor.

(4) H_0: The level of social distress for the child is related to the peer status of the child.
H_1: The level of social distress for the child is not related to the peer status of the child.

(5) H_0: The distribution of the social distress levels for the child are not the same for each peer status factor.
H_1: The distribution of the social distress levels for the child are the same for each peer status factor.

The Chi-square test for independence is conducted and SPSS output is shown below. Some values have been replaced with ******.

Level of Social Distress * Peer Status Crosstabulation

			Peer Status					Total
			Popular	Average	Neglected	Rejected	Controversial	
Level of Social Distress	High	Count	8	19	10	26	2	65
		Expected Count	13.0	20.2	11.1	**	5.0	65.0
	Low	Count	41	57	32	33	17	180
		Expected Count	36.0	55.8	30.9	43.3	14.0	180.0
Total		Count	49	76	42	59	19	245
		Expected Count	49.0	76.0	42.0	59.0	19.0	245.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14.674[a]	**	.005
Likelihood Ratio	14.541	**	.006
N of Valid Cases	245		

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.04.

The cell contributions to the Chi-square test statistic are shown in the table below:

Level of Social Distress	**Peer Status**				
	Popular	Average	Neglected	Rejected	Controversial
High	1.923	0.071	0.109	6.757	1.800
Low	0.694	***	0.039	2.450	0.643

Cell contributions to the Chi-square test statistic

42. The **degrees of freedom** for this test are:

(1) 5
(2) 10
(3) 8
(4) 4
(5) 2

43. The **expected count** for those children who had a **high** level of social distress in relation to a **rejected** peer status is:

    (1) 14.0

    (2) 15.7

    (3) 5.0

    (4) 11.1

    (5) 43.3

44. The **cell contribution** for those children who had a **low** level of social distress in relation to an **average** peer status is:

    (1) 0.109

    (2) 0.694

    (3) 0.039

    (4) 0.026

    (5) 0.071

45. The *P-value* for the test described in Question 41 is 0.005. Which one of the following statements gives the **best** interpretation of this *P-value*?

    (1) There is strong evidence that a child's level of social distress is not independent of their peer status.

    (2) There is strong evidence that a child's level of social distress is independent of their peer status.

    (3) There is evidence that a child's level of social distress is not independent of their peer status.

    (4) There is weak evidence that a child's level of social distress is not independent of their peer status.

    (5) There is no evidence that a child's level of social distress is not independent of their peer status.

46. Which **one** of the following statements is **false**:
    (1) If one or more of the expected counts in a table is less than 1 then we would have concerns with the validity of a Chi-square test carried out on these data.
    (2) If, for several cells in a table of counts, there are relatively large differences between the observed counts and the expected counts under the null hypothesis, then the *P-value* for a Chi-square test will be small.
    (3) The greater the value of the Chi-square test statistic, the weaker the evidence against the null hypothesis.
    (4) A Chi-square test for independence is used to carry out a formal analysis on data presented in a two-way table of counts.
    (5) The Chi-square test statistic is a measure of the difference, over all cells in the table, between the counts observed from the sample and the counts that would have been expected under the null hypothesis.

47. When is it **not** appropriate to conduct a Chi-square test?
    (1) 1 sample cross-classified by two factors of interest.
    (2) Only 20% of expected counts are greater than 5.
    (3) 2 samples classified on the factor of interest.
    (4) No expected counts are less than 1.
    (5) Testing proportions from tables of counts.

48. Which one of the following statements about data in tables of counts is **false**?
    (1) The *P-value* in a Chi-square test is the probability, given that the null hypothesis is true, of obtaining a test statistic as extreme, or more so, as that observed.
    (2) A Chi-square test of homogeneity on the column distributions can be used on a single random sample cross-classified by two response factors.
    (3) A Chi-square test of homogeneity on the row distributions can be used on a single random sample cross-classified by two response factors.
    (4) A Chi-square test of independence can be used on a single random sample cross-classified by two response factors.
    (5) The greater the value of the Chi-square test statistic, the greater the *P-value*.

ANSWERS

1. **(5)**	2. **(2)**	3. **(3)**	4. **(3)**	5. **(5)**	6. **(2)**
7. **(5)**	8. **(3)**	9. **(4)**	10. **(1)**	11. **(1)**	12. **(4)**
13. **(5)**	14. **(3)**	15. **(5)**	16. **(4)**	17. **(3)**	18. **(4)**
19. **(1)**	20. **(2)**	21. **(1)**	22. **(3)**	23. **(4)**	24. **(5)**
25. **(5)**	26. **(4)**	27. **(5)**	28. **(4)**	29. **(5)**	30. **(5)**
31. **(4)**	32. **(3)**	33. **(2)**	34. **(4)**	35. **(1)**	36. **(3)**
37. **(1)**	38. **(5)**	39. **(5)**	40. **(4)**	41. **(2)**	42. **(4)**
43. **(2)**	44. **(4)**	45. **(1)**	46. **(3)**	47. **(2)**	48. **(5)**

WHAT SHOULD I DO NEXT?

- Do all the problems in this workshop handout and mark them. If you get a question wrong, have a look at the working on Leila's scanned slides at www.tinyURL.com/stats-CST to see how she did it.
- Go through the Chapter 9 blue pages. This includes:
    - the *notes* on page 13,
    - the *glossary* on page 14,
    - the *true/false statements* on page 15,
    - the *Sample Exam Questions* on pages 16-18, and
    - the *tutorial* material on pages 19 & 20.
- Attend the optional Chapter 9 tutorial.
- Try the **PRACTICE Ch9 Quiz**.
- Do three attempts of the **Chapter 9 Quiz** (due at 11pm on Wednesday 28 October 2020).
- Do the Chapter 9 parts of the Revision Assignment. Get this from Canvas under *Assignments*. **Note that this assignment is not one of the formal assessments for your final mark & grade so it is not to be handed in!**
- Try Chapter 9 questions from three of the past five exams on Canvas (get them from *Modules* → *Past Tests and Exams (with answers)* and use the *Exam questions index* document from there to identify the Chapter 9 questions!)
- If you get anything wrong and don't know why, get some help. You can post a question on Piazza (search first as it may have already been asked!), or talk to someone about it (your lecturer, an Assistance Room tutor or Leila).

FORMULAE

Confidence intervals and t-tests

Confidence interval: $estimate \pm t \times \text{se}(estimate)$

t-test statistic: $t_0 = \dfrac{estimate - hypothesised\ value}{standard\ error}$

Applications:

1. Single mean $\boldsymbol{\mu}$: $estimate = \overline{x}$; $df = n - 1$
2. Single proportion $\boldsymbol{p}$: $estimate = \widehat{p}$; $df = \infty$
3. Difference between two means $\boldsymbol{\mu_1 - \mu_2}$: (independent samples)

   $estimate = \overline{x}_1 - \overline{x}_2$; $df = \min(n_1 - 1, n_2 - 1)$
4. Difference between two proportions $\boldsymbol{p_1 - p_2}$:

   $estimate = \widehat{p}_1 - \widehat{p}_2$; $df = \infty$

   Situation (a): *Proportions from two independent samples*

   Situation (b): *One sample of size n, several response categories*

   Situation (c): *One sample of size n, many yes/no items*

The F-test (ANOVA)

F-test statistic: $f_0 = \dfrac{s_B^2}{s_W^2}$; $df_1 = k - 1$, $df_2 = n_{\text{tot}} - k$

The Chi-square test

Chi-square test statistic: $\chi_0^2 = \displaystyle\sum_{\text{all cells in the table}} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$

Expected count in cell $(i, j) = \dfrac{R_i C_j}{n}$

$df = (I - 1)(J - 1)$

Regression

Fitted least-squares regression line: $\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$

Inference about the intercept, β_0, and the slope, β_1: $df = n - 2$