STATS 10x WORKSHOP

SATURDAY IV: CHAPTERS 8-10

SATURDAY 9 OCTOBER 2010

Students **MUST REGISTER** for all workshops with
The Student Learning Centre, 3rd Floor, Information Commons

Student Learning Centre

Topics we teach and can provide advice on include:

- ✓ Essay writing
- ✓ Computer skills
- ✓ Reading and notetaking
- ✓ Memory and concentration
- ✓ Report writing
- ✓ Test and examination skills
- ✓ Thesis and dissertation writing
- ✓ Tutorial skills
- ✓ Research skills
- ✓ Time and stress management
- ✓ Mathematics
- ✓ **Statistics**
- ✓ Oral presentation and seminar skills
- ✓ Language learning
- ✓ Specific learning disabilities
- ✓ Motivation and goal setting
- ✓ Survival skills (in the University system)

Programmes within SLC include:

- Te Puni Wananga
  Maori university tutors committed to enhancing Maori students' success

- Fale Pasifika
  Pacific Island tutors committed to enhancing success for Pacific Island students

- Students with impairments
  Learning assessments are available for students with specific learning disabilities; academic assistance is available for these students and those with mental health impairments.
  If you have any special learning requirements, please feel free to discuss this with Leila in person or via email.

- Academic English Conversation Groups
  Improve your academic English; develop communication skills including critical/creative thinking and clear expression of ideas and opinions. Weekly class held at the SLC on Thursdays, 3-5pm (during semester)

Statistical help available at the SLC

The Student Learning Centre (SLC) offers help for STATS 101/108 by offering:

- one-on-one tutoring help, and
- a number of workshops

One-on-one help

The SLC employs tutors specifically to help students with one-on-one assistance for STATS 101/108. One-on-one tutoring must be booked at SLC reception on the third floor of the Information Commons in person or by calling 373-7599 X 88850. Enquire at the SLC reception for available times.

Note: SLC tutors are not allowed to help students complete their assignments.

SLC STATS 101/108 Workshops

Any questions regarding STATS 101/108 workshops should be forwarded to:

Leila Boyle
SLC Statistics Co-ordinator
l.boyle@auckland.ac.nz

Workshops are run in a relaxed environment, typically set at a pace for those students that find the Statistics Department's tutorials too fast. All workshops allow plenty of time for questions. In fact, this is encouraged ☺

1) Saturday Workshops

These five 3-hour workshops are held on Saturdays throughout the semester to help students with different sections of the course.

2) Computer Workshops: Excel / PASW (SPSS)

These three computer-based workshops introduce students to the skills needed for Excel and PASW (SPSS) use in STATS 101/108 assignments.

3) Pre-test Workshops

These three workshops will cover the basics that you need for the test.

4) Pre-exam Workshops

These six workshops will cover the basics that you need for the exam.

Note: All workshops concentrate on questions reviewing the basic concepts, rather than questions on finer details. They are designed to assist students to achieve a pass; they are not designed to cover all material.

The timetable for these workshops is available with this handout. Currently the SLC website is still partly down so online enrolments are not available until further notice. Please enrol in each of your preferred classes at the Student Learning Centre by:

- **Going to the SLC in person**
- **Emailing slc@auckland.ac.nz with your name, ID number and the workshop/s you wish to attend.**
- **Calling the SLC reception on 373-7599 ext. 88850 and enrol over the phone. Make sure you know which workshop/s you want to enrol in and have your ID number handy.**

Useful Websites

- SLC webpage: **www.slc.auckland.ac.nz** (The SLC website currently has all functionality except online enrolment! Download an undergraduate brochure and enrol in workshops in person or by emailing/phoning the SLC Reception as per above instructions).
- Cecil: **https://cecil.auckland.ac.nz/**
- **Leila's website for STATS 101/108 SLC workshop handouts & information: www.stat.auckland.ac.nz/~leila**

Recall: **Chapter 8 – Confidence Intervals**

Look at blue pages for good notes and test/exam questions for practice

- A **confidence interval** gives a range of plausible values for the parameter of interest that is consistent with the data (at the specified level of confidence). It determines the **size** of the effect or difference.

- You can do all kind of CI's, 90%, 95%, 99%...

- Increasing the confidence level will **increase** the width of the interval.

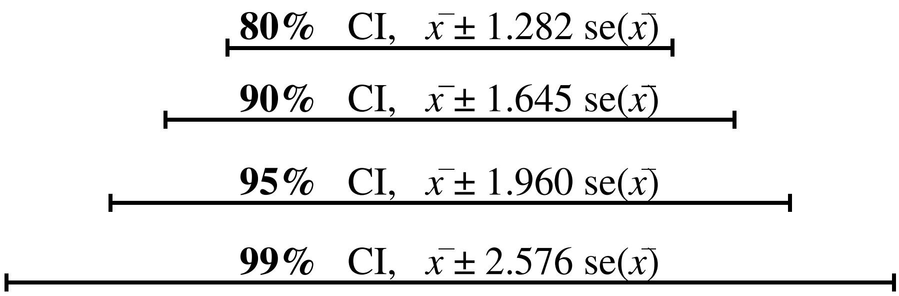


Figure 8.1.3 The greater the confidence level, the wider the interval

- Increasing the sample size will make the confidence interval more precise.

- To double the accuracy of the confidence interval we **need 4 times** as many observations.

- To triple the accuracy of the confidence interval we **need 9 times** as many observations.

- 95% confidence interval

  ✓ Range of plausible values for the parameter of interest that contains the **true value** of our parameter of interest for 95% of samples taken.

  ✓ 5% of samples taken will not have the parameter within the calculated confidence interval.

  ✓ We do not know if the sample we have taken is one of the 95% that contains the true unknown parameter. All we can say is that 95% of the time it will.

- ✓ If you take 1000 samples, based on the same sampling protocol, then you can expect approximately 950 of these samples will contain the true value (e.g. true mean) of the population.

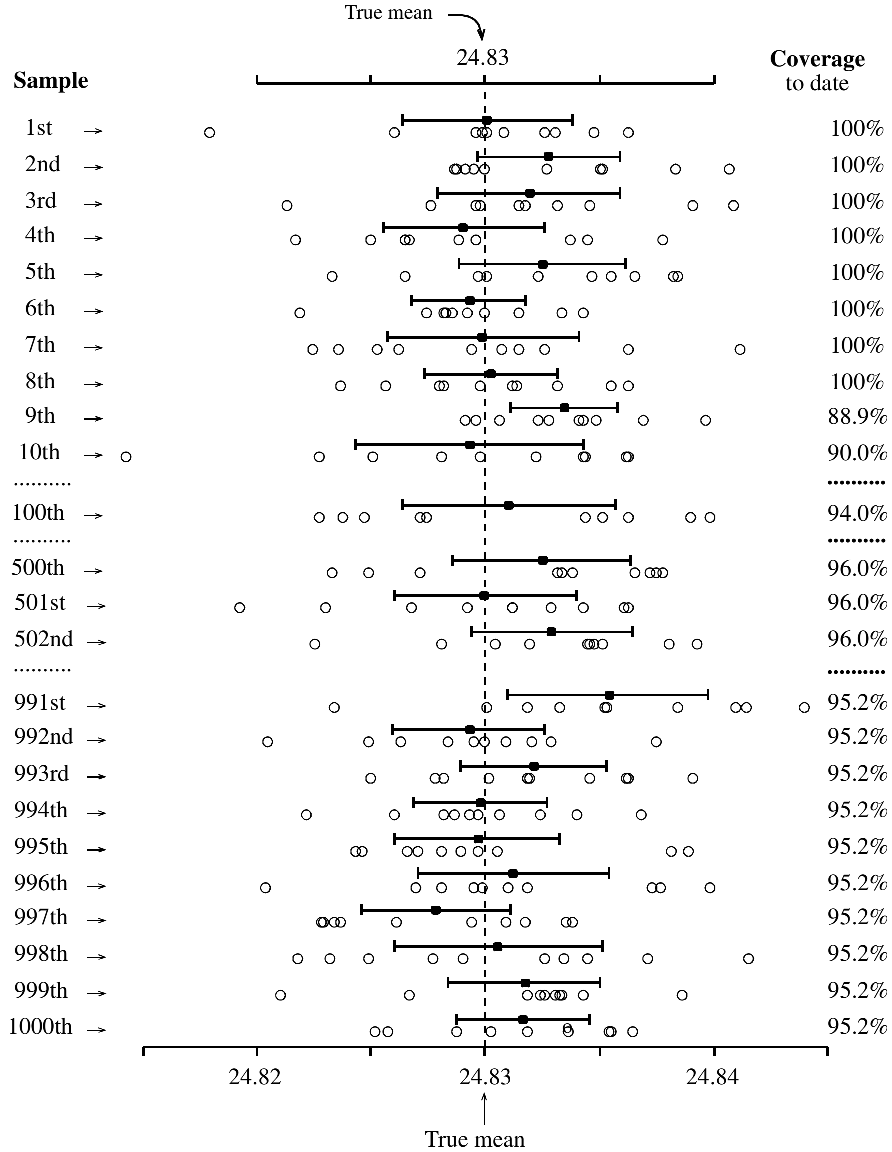


Figure 8.1.2 Samples of size 10 from a Normal(μ=24.83, s=.005) distribution and their 95% confidence intervals for μ..

From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 1999.

Step-by-Step Guide to Producing a Confidence Interval by Hand

1. State the **parameter** to be estimated (symbol and words).
Is it μ, p, $\mu_1 - \mu_2$, or $p_1 - p_2$?

2. State the **estimate** and its value.

3. Write down the **formula** for a CI, ***estimate* ± *t* × se(*estimate*)**, from the Formula Sheet

4. Find the appropriate **standard error** from the Formula Sheet.

5. Find the appropriate ***df***.

6. Find the appropriate ***t*-value.** This is called the '*t*-multiplier'.
Need to know the confidence level and *df*.

7. Calculate the **confidence limits**, the end points of the confidence interval.

8. Interpret the interval using plain English.
Use the confidence limits to construct an answer to the original question in plain English.

- There are four different types of problem covered in Chapter 8:
  1. Single mean 2. Single proportion. 3. Difference between two means
  4. Difference between two proportions:
     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 **estimate** is based on the **parameter** of interest we are investigating:

Parameter	**Estimate**
1. Single mean $\boldsymbol{\mu}$:	*estimate* $= \bar{x}$
2. Single proportion $\boldsymbol{p}$:	*estimate* $= \hat{p}$
3. Difference between two means $\boldsymbol{\mu_1 - \mu_2}$: (independent samples)	*estimate* $= \bar{x}_1 - \bar{x}_2$
4. Difference between two proportions $\boldsymbol{p_1 - p_2}$:	*estimate* $= \hat{p}_1 - \hat{p}_2$

- The ***t*-multiplier** (means) / ***z*-multiplier** (proportions) is based on:
  - ✓ Whether we are investigating means or proportions
  - ✓ The desired level of confidence
  - ✓ The degrees of freedom

Estimate	Degrees of Freedom
1. $estimate = \overline{x}$	$df = n - 1$
2. $estimate = \hat{p}$	$df = \infty$
3. $estimate = \overline{x}_1 - \overline{x}_2$	$df = minimum(n_1 - 1, n_2 - 1)$
4. $estimate = \hat{p}_1 - \hat{p}_2$	$df = \infty$

i.e. for proportions, assume the degrees of freedom is infinity, hence replace t with z score (i.e. the standard Normal).

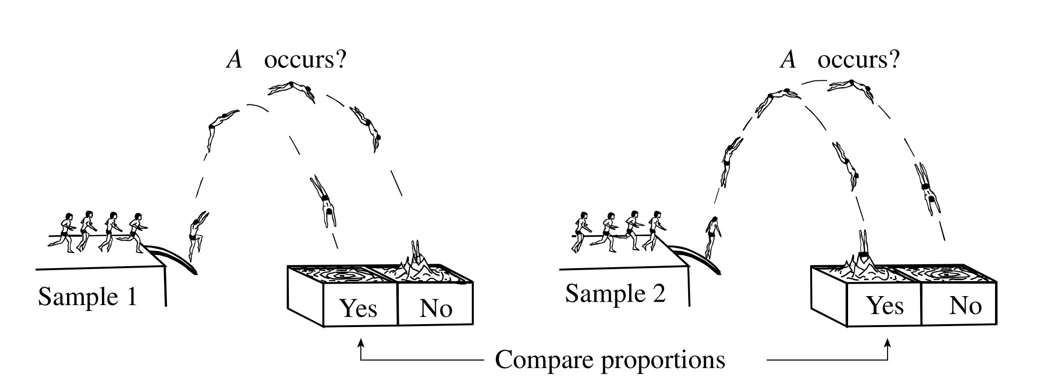
- The **standard error** is based on the estimate, the number of samples and sample size(s).

Estimate	se(estimate)
1. $estimate = \overline{x}$	$se(\overline{x}) = \frac{s}{\sqrt{n}}$
2. $estimate = \hat{p}$	$se(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
3. $estimate = \overline{x}_1 - \overline{x}_2$	$se(\overline{x}_1 - \overline{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
4. $estimate = \hat{p}_1 - \hat{p}_2$	**Situation (a)** ***Proportions from two independent samples*** $se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ **Situation (b)** ***One sample of size n, several response categories*** $se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1 + \hat{p}_2 - (\hat{p}_1 - \hat{p}_2)^2}{n}}$ **Situation (c)** ***One sample of size n, many yes / no items*** $se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\min imum(\hat{p}_1 + \hat{p}_2, \hat{q}_1 + \hat{q}_2) - (\hat{p}_1 - \hat{p}_2)^2}{n}}$ where $\hat{q}_1 = 1 - \hat{p}_1$ and $\hat{q}_2 = 1 - \hat{p}_2$

- CIs for the difference between two means/proportions:
  - ✓ If the CI contains 0 (i.e. one negative and one positive number), there may be no difference between the two means/proportions.
  - ✓ If CI is positive, then μ_1/p_1 is higher/larger than μ_2/p_2.
  - ✓ If CI is negative, then μ_1/p_1 is lower/smaller than μ_2/p_2.

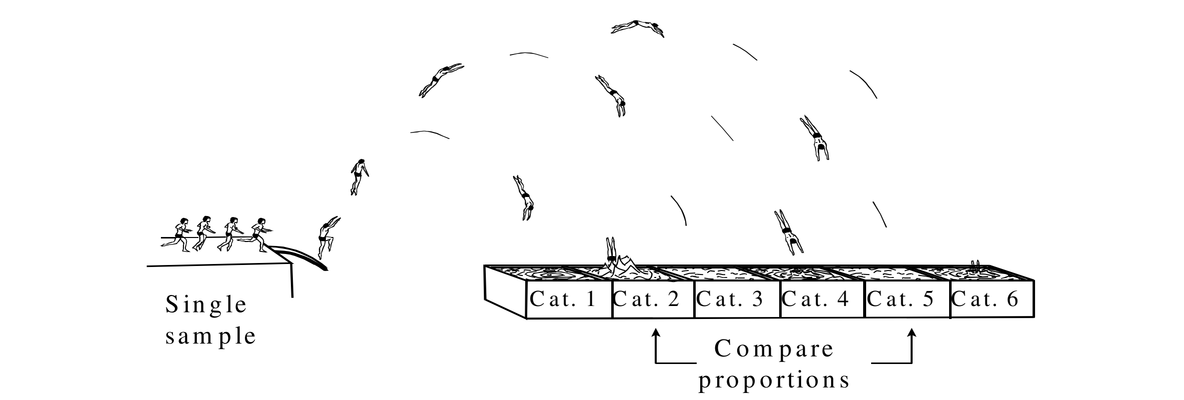
- **3 sampling situations for the difference between two proportions**

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



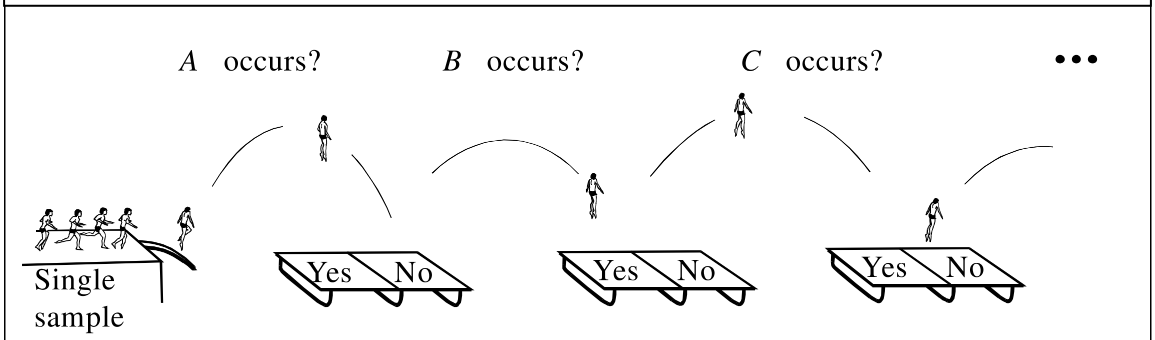
From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 1999.

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



From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

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



From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

Chapter Nine

In hypothesis testing and confidence intervals we use **sample** data to make inferences (draw conclusions) about **population** parameters.

- A **confidence interval** gives a range of plausible values for the parameter of interest that is consistent with the data (at the specified level of confidence).
- A **significance test,** tests one possible value for the parameter, called the **hypothesised** value. We determine the strength of evidence provided by the data against the null hypothesis, H_0.

A significance test determines the **strength** of the evidence **against** the **hypothesised** value, while a confidence interval determines the **size** of the effect or difference.

Significance testing is a method used to deal with the **uncertainty** about the true value of a parameter caused by the sampling variation in estimates.

H_0 denotes the **null hypothesis** whereas H_1, denotes the **alternative hypothesis**.

The null hypothesis, H_0

- ✓ It is our best guess as to what we think the parameter of interest is – a single plausible value.
- ✓ General form: H_0: *parameter = hypothesised value (some number)*
- ✓ The hypothesised value is **not** the parameter of interest. Remember that the parameter of interest is an unknown quantity.

The alternative hypothesis, H_1

- ✓ Specifies the type of departure from H_0 that we expect to detect.
- ✓ Corresponds to the research hypothesis.
- ✓ There are three different types:
  - H_1: *parameter ≠ hypothesised value (some number)*
  - H_1: *parameter < hypothesised value (some number)*
  - H_1: *parameter > hypothesised value (some number)*
- ✓ ONLY use a one-sided alternative when you have prior information or a theory.

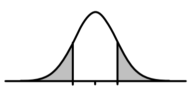
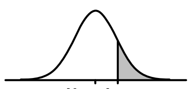
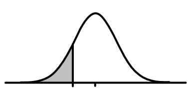
The *t*-test-statistic

- ✓ Is the number of standard errors our estimate is from the hypothesised value.
- ✓ We calculate it using: $t_0 = \dfrac{estimate - hypothesised\ value}{std\ error}$

The *P*-value

- ✓ Is the probability that sampling variation would produce an estimate that is further away from the hypothesised value than the estimate we obtained from our data, assuming that the null hypothesis is true.
- ✓ It is a measure of evidence **against** H_0.
- ✓ We calculate the *P*-value using the *t*-test statistics and the appropriate Student's *t* distribution.
- ✓ We interpret the size of the *P*-value using the following table:

P-value	Evidence against H_0
> 0.12	None
≈ 0.10	Weak
≈ 0.05	Some
≈ 0.01	Strong
< 0.001	Very Strong

Alternative hypothesis	*P-value* = area of shaded region $T \sim Student(df)$
H_1: parameter ≠ hypothesised value (2-sided)	 2-tailed test
H_1: parameter **>** hypothesised value (1-sided)	 1-tailed test
H_1: parameter **<** hypothesised value (1-sided)	 1-tailed test

Step-by-Step Guide to Performing a *t*-test by Hand

1. State the **parameter**, being discussed (symbol/s and words).
   Is it μ, p, $\mu_1 - \mu_2$, or $p_1 - p_2$?

2. State the **null hypothesis**, $\boldsymbol{H_0}$.

   i.e. H_0: parameter = *hyp. val.*

3. State the **alternative hypothesis**, $\boldsymbol{H_1}$.

   i.e. H_1: parameter ≠ *hyp. val.*
   OR H_1: parameter < *hyp. val.*
   OR H_1: parameter > *hyp. val.*

4. State the **estimate**, and its value.

5. Calculate the ***t*-test statistic**:

   - Use: $t_0 = \dfrac{\textit{estimate} - \textit{hypothesised value}}{\textit{std error}}$ (from the Formula Sheet)

     Write down the estimate and the hypothesised value.

     Find the appropriate standard error and the *df* using the Formula Sheet.

   - Calculate t_0.

6. Calculate the ***P-value***.

   See *Finding P-values*, pages 9-12, *Excel* Tutorial, Section F – Lecture Workbook

7. **Interpret** the *P-value*.

   See *Interpreting the P-value*, page 11, Chapter 9 Notes, Section B – Lecture Workbook

8. Calculate the **confidence interval**.

9. **Interpret** the confidence interval using plain English.

Statistical significance

- ✓ Relates to having evidence of the **existence** of an effect or difference.
- ✓ Determined by examining the ***P-value*** of your significance test.

Practical significance

- ✓ Depends on the **size** of the effect or difference.
- ✓ Determined by examining the **confidence interval** in relation to the research in question.

Chapter Ten

In this chapter we dealt with:

- One sample
- Paired data comparisons
- Two independent samples
- More than 2 samples
- Parametric vs Nonparametric issues
- Experimental design

Non-parametric tests

✓ Used when there are strong concerns about the "normality" of the data. You should make sure data don't show separation into clusters or have a multi-modal nature and then apply the 15-40 rule as follows:

Sample Size Guidelines – "15 – 40 Rule"

Small (total $n \le 15$ or so)	Medium ($15 <$ total $n < 40$)	Large (total $n \ge 40$ or so)
no outliers	no outliers	no gross outliers
at most, slight skewness	not strongly skewed	data may be strongly skewed

✓ They test the **median(s)**, $\tilde{\mu}$, NOT the mean(s).

1-sample *t*-test

✓ 1 sample (group) of data

✓ Parameter = μ

✓ **Hypotheses**: $H_0 : \mu = \mu_0$

vs $H_1 : \mu \neq \mu_0$

✓ **Assumptions** for 1-sample *t*-tests

1. **Random sample → Independence** of <u>observations</u> – *CRITICAL!*
2. **Normality** – quite robust, but sensitive to outliers in small-medium samples. No clusters or multi-modes allowed.

✓ Data **not** Normal? Use **Sign Test** <u>*OR*</u> **Wilcoxon Signed-Rank Test**.

- The **Sign Test** is a nonparametric equivalent of the one sample *t*-test.
- Observations above the hypothesised value are given a positive sign (**+**).
- Observations below the hypothesised value are given a negative sign (**−**).
- Observations with the same value as the hypothesised value are ignored.
- Evidence against the null hypothesis is provided by a large imbalance in the number of ' **+** ' and ' **−** ' signs.

Paired data

- ✓ 1 sample (group) of data
- ✓ Two measurements taken on each experimental unit
- ✓ With related or paired data we analyse the <u>differences</u> and use a 1-sample t-test, i.e. we treat the differences as a single sample.
- ✓ Parameter = μ_{diff}
- ✓ **Hypotheses**: H_0: $\mu_{\text{diff}} = 0$
  vs H_1: $\mu_{\text{diff}} \neq 0$
- ✓ **Assumptions** for paired-t-tests
  1. **Random sample → Independence** of <u>pairs of observations</u> – *CRITICAL!*
  2. **Normality** – quite robust, but sensitive to outliers in small-medium samples. No clusters or multi-modes allowed.
- ✓ Data **not** Normal? Use **Sign Test** <u>*OR*</u> **Wilcoxon Signed-Rank Test**.
- ✓ **Note:** You don't need to know how to do **Sign** or **Wilcoxon Signed-Rank** tests by hand – just remember the names and be able to interpret the **Sign** test's *P-value* from SPSS computer output.

2-sample *t*-tests

- ✓ Two independent samples
- ✓ Parameter = $\mu_1 - \mu_2$
- ✓ **Hypotheses**: H_0: $\mu_1 - \mu_2 = 0$
  vs H_1: $\mu_1 - \mu_2 \neq 0$
- ✓ **Assumptions** for 2-sample t-tests
  1. **Independence** of <u>samples</u> – *CRITICAL!*
     <u>**AND**</u>
     Random samples → Independence of <u>observations</u> – *CRITICAL!*
  2. **Normality** – even more robust against non-Normality than one sample t-procedures. No clusters or multi-modes allowed.
- ✓ Data **not** Normal? Use **Mann Whitney Test** <u>*OR*</u> **Wilcoxon Rank-Sum Test**.
- ✓ **Note:** You don't need to know how to do **Mann Whitney** or **Wilcoxon Rank-Sum** tests by hand – just remember the names.

3 or more samples ⇒ *F*-test for 1-way ANOVA

✓ Three or more independent samples

✓ **Hypotheses:** H_0 : all the underlying population means are the same
i.e. $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$
vs H_1 : at least 2 of the underlying population means are not the same
Note: SHOULD NOT be written using symbols

✓ *F*-statistic $= f_0 = \dfrac{s_B^2}{s_W^2}$ where s_B^2 = the **between** group variation and
s_W^2 = the **within** group variation

✓ **Assumptions** for 1-way ANOVA

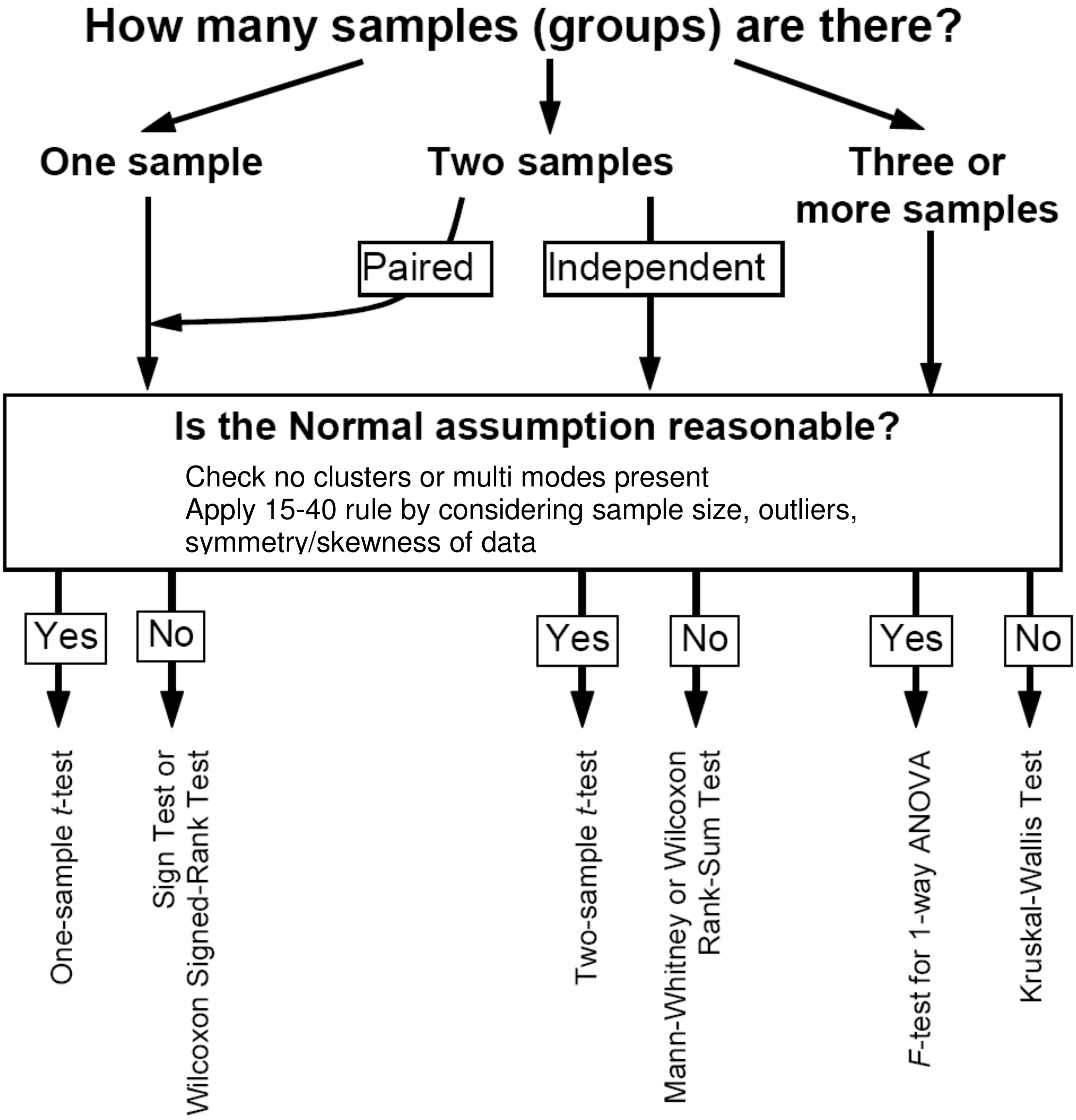
1. **Independence** of <u>samples</u> – *CRITICAL!*
   <u>AND</u>
   Random samples → Independence of <u>observations</u> – *CRITICAL!*
   Check by thinking about the way the data was obtained.
2. **Normality**
   Check by looking at plots of each sample and consider sample sizes. Plots should be unimodal and not too skewed for the n_{tot} you have.
3. **Equality of Standard Deviations**
   Check by using the ratio (fraction) $\dfrac{\text{largest sd}}{\text{smallest sd}} < 2$ as a guide

✓ **One-way ANOVA Table**

	Sum of Squares (SS)	*df*	Mean Sum of Squares (MSS)	*F*-statistic	*P*-value
Between Groups (BG)	BGSS	$df_1 = k - 1$	$\text{MSS}_B = s_B^2$	$f_0 = \dfrac{s_B^2}{s_W^2}$	$\Pr(F \geq f_0)$
Within Groups (WG)	WGSS	$df_2 = n_{\text{tot}} - k$	$\text{MSS}_W = s_W^2$		
Total (T)	TSS	$df_{\text{tot}} = n_{\text{tot}} - 1$			

✓ Data **not** Normal or standard deviations **not** equal? Use **Kruskal-Wallis Test**.

✓ **Note:** You don't need to know how to do **Kruskal-Wallis** test by hand – just remember the name.



Assumptions	Checks
1. Independence – All tests - Single sample assumes indep. between observations. - Paired data assumes indep. between pairs of observations. - 2 or more samples assume independence between observations and between samples.	- The design of the experiment
2. Normality – Parametric tests only - one-sample t-test - two-sample t-test - F-test	- Plot the data and apply 15-40 rule
3. Equal Spread – *F*-test only	- Plot the data - Check the standard deviation ratio: $\frac{\text{largest sd}}{\text{smallest sd}} < 2$

Chapters 8-10 – Questions

1. Which one of the following statements about significance tests is false?
   - (1) Formal tests can help determine whether effects we see in our data may just be due to sampling error.
   - (2) A test statistic is a measure of discrepancy between what we see in our data and what we would expect to see if H_0 was true.
   - (3) The *P-value* says nothing about the size of an effect.
   - (4) The data should be carefully examined in order to determine whether the alternative hypothesis needs to be one-sided or two-sided.
   - (5) The *P-value* describes the strength of evidence against the null hypothesis.

2. Which one of the following statements is **true**?
   - (1) A two-sided test of H_0: *parameter = hypothesised value* has *P-value* less than 0.05 if the *hypothesised value* lies within a 95% confidence interval for the *parameter*.
   - (2) The larger the *P-value,* the stronger the evidence against H_0.
   - (3) The larger the test statistic, $|t_0|$, for a two-sided test, the larger the *P-value* will be.
   - (4) Tests of hypotheses can only deal with random errors and sampling variation. They are ineffective when confronted with data that has systematic bias.
   - (5) An extremely small *P-value* means that the actual effect differs markedly from that claimed in the null hypothesis.

3. Which one of the following statements about hypothesis testing is **false**?
   - (1) The larger the *P-value,* the stronger the evidence against the null hypothesis.
   - (2) The *P-value* is the probability that, if the null hypothesis were true, sampling variation would produce an estimate that is further away from the hypothesised value than our data estimate.
   - (3) We cannot establish an hypothesised value for a parameter, we can only determine whether there is evidence to reject a hypothesised value.
   - (4) H_0 is typically a sceptical reaction to a research hypothesis.
   - (5) The *P-value* measures the strength of evidence against the null hypothesis.

4. Which one of the following statements is **true**?

(1) A small *P-value* provides evidence of the size of an effect.
(2) Statistical significance is the same as practical significance.
(3) Practical significance depends on the size of the effect.
(4) A small *P-value* provides no evidence against H_0.
(5) A confidence interval estimates the strength of an effect.

5. Which one of the following statements about *t*-tests is **false**?

(1) *t*-tests may not be valid if there are outliers present and the sample is not large.
(2) *t*-tests may not be valid when the data show clustering.
(3) In general, *t*-tests are not robust against the Normality assumption.
(4) *t*-tests will generally work well for any large sample.
(5) *t*-tests may not be valid if the data are clearly skewed and the sample is not large.

Questions 6 and **7** refer to the following information.

Printed on every packet of "Yummo" corn chips is a weight of 150g. A consumer collects 48 packets of "Yummo" corn chips and finds a mean weight of 148.5g and a standard deviation of 2.1g.

6. The customer wishes to test H_0: μ = 150 versus H_1: $\mu \neq$ 150. The value of the *t*-test statistic, t_0, and the degrees of freedom *df*, to be used are given by:

(1) t_0 = -4.95, df = 47
(2) t_0 = 1.79, df = 45
(3) t_0 = 2.14, df = 48
(4) t_0 = -1.79, df = 47
(5) t_0 = -2.14, df = 48

7. Suppose the customer finds that the *p-value* for the above test = 0.09 (it is not). The **best** interpretation of this test would be:

(1) With a *p-value* of 0.09 there is no evidence against the null hypothesis.
(2) With a *p-value* of 0.09 there is weak evidence against the null hypothesis.
(3) With a *p-value* of 0.09 there is weak evidence against the null hypothesis that the mean weight of Yummo chips is 150g.
(4) With a *p-value* of 0.09 there is no evidence against the null hypothesis that the mean weight of Yummo chips is 150g.
(5) With a *p-value* of .09% there is weak evidence against the null hypothesis that the mean weight of Yummo chips is 150g.

Questions 8 to **10** refer to the following information.

DEATH PENALTY SURVEY RESULTS

"Should convicted murderers be put to death?"

	Australia	N.Z.
Yes	46%	42%
No	39%	41%
Can't Say	15%	17%

[Polls of 1307 Australians & 1010 New Zealanders]

8. Based on previous studies, a researcher believes that the proportion of New Zealanders who agree that convicted murderers should be put to death would be more than forty percent. The hypotheses for this test would be:

(1) H_0: p = 0.40; H_1: $p \neq 0.40$

(2) H_0: p < 0.40; H_1: p = 0.40

(3) H_0: p = 0.40; H_1: p > 0.40

(4) H_0: p < 0.40; H_1: $p \neq 0.40$

(5) H_0: p < 0.40; H_1: p > 0.40

9. Let's assume instead that the researcher tested H_0: p = 0.40 versus H_1: $p \neq 0.40$ where p = the proportion of New Zealanders who agree that convicted murderers should be put to death. The value of the t-test statistic, t_0, and the degrees of freedom df, to be used with 95% confidence are given by:

(1) $t_0 = 1.288$, $df = \infty$

(2) $t_0 = 1.288$, $df = 1009$

(3) $t_0 = -1.288$, $df = \infty$

(4) $t_0 = 1.288$, $df = 1.96$

(5) $t_0 = -1.288$, $df = 1009$

Another difference considered by the researcher was between the proportion of Australians supporting the *death penalty for convicted murderers* and the proportion of New Zealanders supporting the *death penalty for convicted murderers*.

10. To test for a difference in the two proportions given above the hypotheses would be:

(1) $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 > 0$

(2) $H_0 : \hat{p}_1 - \hat{p}_2 = 0$ vs $H_1 : \hat{p}_1 - \hat{p}_2 \neq 0$

(3) $H_0 : p_1 - p_2 = 0$ vs $H_1 : p_1 - p_2 \neq 0$

(4) $H_0 : \hat{p}_1 - \hat{p}_2 < 0$ vs $H_1 : \hat{p}_1 - \hat{p}_2 > 0$

(5) $H_0 : p_1 - p_2 > 0$ vs $H_1 : p_1 - p_2 = 0$

Results from two independent polls of voting intentions, taken in July and August 1993, were reported in TIME (20 September, 1993). There were 950 voters sampled in each poll. The results were:

Preferred Party (%)

	Alliance	Labour	National	NZ First
July	18	36	38	-
August	12.5	38	36.5	11.5

There was some interest in whether the NZ First party had a bigger impact on the Labour or the National vote. To investigate this, a test was conducted of the null hypothesis that there was no change in $p_L - p_N$, the Labour-National difference, between the two polls. The resulting *P-value* was 0.76.

11. Which of the following statements is **true**?

(1) There is no evidence at the 5% level that NZ First has had an impact on support for either Labour or National.
(2) The probability that NZ First has had no impact on the difference between National and Labour is 0.76.
(3) A 95% confidence interval for the difference between National and Labour would not contain zero.
(4) There is no evidence at the 5% significance level that the difference between National and Labour has changed.
(5) The probability that NZ First has had an impact on the difference between National and Labour is 0.76.

12. Which one of the following statements about paired data is **false**?

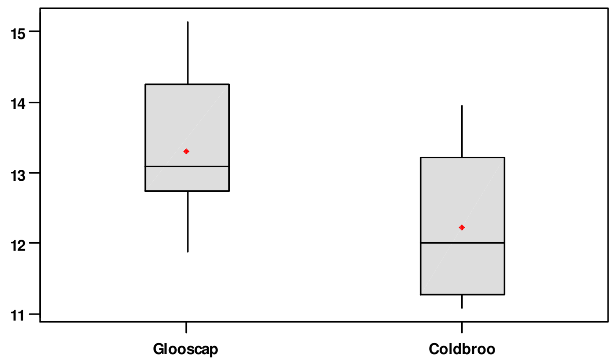
(1) For paired data, we analyse the differences
(2) Pairing is beneficial when the variability within pairs is small compared with the variability between pairs
(3) The Wilcoxon signed rank test can be used to analysis the differences in paired data
(4) Pairing cannot be used in observational studies
(5) The carryover effect occurs when the first treatment alters the effect of the second treatment.

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

(1) A statistically significant result is not always practically significant.
(2) A non-significant hypothesis test does not mean that the null hypothesis is true.
(3) Large positive *t*-test statistics lead to small *P-values* for two-tailed tests.
(4) A small *P-value* from a hypothesis test may result from a very large sample, and the results may be of no practical significance.
(5) A one-tailed *t*-test should be used when the idea for doing the test came about as a result of looking at the data.

Questions 14 to **17** refer to the following information.

As part of a study to compare the physical education programs at two Canadian schools, running times (in seconds) over a set distance were recorded for two independent samples of sixth grade students taken from each school. (Data source courtesy of Chance Encounters).



Group Statistics

	schcode	N	Mean	Std. Deviation	Std. Error Mean
runtime	Glooscap	12	13.3125	.99133	.28617
	Coldbrook	13	12.2285	.99018	.27463

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
runtime	Equal variances assumed	.008	.930	2.733	23	.012	1.08404	.39661	.26359	1.90449
	Equal variances not assumed			2.733	22.836	.012	1.08404	.39663	.26322	1.90485

14. To test for a difference in the physical education programs of the two schools the null and alternative hypotheses would be:

(1) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 > 0$

(2) $H_0 : p_1 - p_2 = 0$ versus $H_1 : p_1 - p_2 \neq 0$

(3) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \neq 0$

(4) $H_0 : \bar{x}_1 - \bar{x}_2 = 0$ versus $H_1 : \bar{x}_1 - \bar{x}_2 \neq 0$

(5) $H_0 : p_D - p_S > 0$ versus $H_1 : p_D - p_S = 0$

15. Which one of the following is false?

(1) There are no gross outliers in the data.
(2) The running times for Glooscap are on average, greater than the running times for Coldbrook.
(3) Two-sample *t*-tests are, in general, more robust to non-Normal features than the one sample *t*-test.
(4) The data taken from Glooscap and Coldbrook show severely non-normal features.
(5) The range of running times for Glooscap is slightly more than the range of running times for Coldbrook.

16. The best interpretation for this test is:

(1) With a *p-value* of 0.12 there is no evidence against the null hypothesis.
(2) With a *p-value* of 0.012 there is no evidence against the null hypothesis that there is a difference in the mean running times.
(3) With a *p-value* of 1.2% there is no evidence against the null hypothesis that there is no difference in the mean running times.
(4) With a *p-value* of 0.012 there is strong evidence against the null hypothesis that there is no difference in the mean running times.
(5) With a *p-value* of 0.012 there is weak evidence against the null hypothesis that there is no difference in the mean running times

17. The 95% confidence interval for the difference between the true mean running times is given in the SPSS output above. Which one of the following interpretations is true?

(1) With a probability of 0.95, the true difference of means $\mu_1 - \mu_2$ lies between 0.26 and 1.91.

(2) In repeated sampling the 95% confidence interval [0.26, 1.90] will contain the true difference in means in 95% of the samples taken.

(3) With 95% confidence, we estimate that the true proportion p_1 will be somewhere between 0.26 larger and 1.90 smaller than p_2.

(4) With 95% confidence, we estimate that the true mean running time from Glooscap μ_1 is somewhere between 0.26 and 1.90 larger than the true mean running time from Coldbrook μ_2.

(5) With 95% confidence the true mean from Glooscap μ_1 is 1.65 larger than the true mean from Coldbrook μ_2.

Questions 18 and **19** refer to the following information.

Factor V is a protein involved in the forming of blood clots. The higher the level of factor V, the faster the blood clots. The Auckland Blood Transfusion Service is interested in the effects of sterilisation of blood plasma because factor V is known to be unstable and may break down during sterilisation. The table below gives measured levels of factor V in blood samples from 16 blood donors. Both pre- and post-sterilisation measurements are given for each blood sample.

Donor Number	Pre-Sterilisation	Post-Sterilisation
1	1073	916
2	1064	1030
3	967	923
4	849	892
5	810	628
6	855	759
7	1047	828
8	1008	784
9	957	809
10	829	773
11	821	786
12	1257	1106
13	1095	832
14	1098	863
15	932	783
16	1440	869

Factor V and Blood Sterilisation

	Sample mean	Standard deviation	Sample size
Pre - Sterilisation	1006.37	170.899	16
Post - Sterilisation	848.812	112.195	16
Difference (Pre-Post)	157.562	140.132	16

Summary Statistics

18. Which **one** of the following statements gives the **correct** hypotheses for this test?

(1) H_0: all of the μ 's are equal
H_1: none of the μ's are equal

(2) H_0: $\mu_1 - \mu_2 = 0$
H_1: $\mu_1 - \mu_2 \neq 0$

(3) H_0: $\mu_{\text{diff1}} - \mu_{\text{diff2}} = 0$
H_1: $\mu_{\text{diff1}} - \mu_{\text{diff2}} \neq 0$

(4) H_0: $\mu_{\text{diff}} = 0$
H_1: $\mu_{\text{diff}} \neq 0$

(5) H_0: $\tilde{\mu}_{\text{diff}} = 0$
H_1: $\tilde{\mu}_{\text{diff}} \neq 0$

19. The t-test statistic for testing whether there is any evidence of an effect of sterilisation is given by:

(1) $\dfrac{157.562}{\dfrac{140.132}{\sqrt{16}}}$

(2) $\dfrac{157.562}{\dfrac{140.132}{16}}$

(3) $\dfrac{1006.37 - 848.812}{\sqrt{\dfrac{170.899^2}{16} + \dfrac{112.195^2}{16}}}$

(4) $\dfrac{1006.37 - 848.812}{\sqrt{\dfrac{170.899}{16} + \dfrac{112.195}{16}}}$

(5) $\dfrac{1006.37 - 848.812}{\dfrac{170.899^2}{16} + \dfrac{112.195^2}{16}}$

Questions 20 to **22** refer to the following information.

A certain drug was claimed to have a side effect of increasing the heart beat rate. An experiment was performed on 8 rats. The number of heartbeats was recorded over a fixed time period immediately before and immediately after each rat received the drug. The data is given below.

20. It would be **inappropriate** to use a two independent sample t-test to test the hypothesis that $\mu_{\text{after}} - \mu_{\text{before}} = 0$ mainly because the:
    (1) Population standard deviations are unknown.
    (2) Sample sizes are small.
    (3) Data are related.
    (4) Samples are independent.
    (5) Population means are unknown.

21. The value of the t-test statistic, t_0, to test the hypothesis that $\mu_{\text{diff}} = 0$, is:

(1) $\dfrac{\overline{x}_{\text{after}}}{se(\overline{x}_{\text{after}})} - \dfrac{\overline{x}_{\text{before}}}{se(\overline{x}_{\text{before}})}$

(2) $\dfrac{\overline{x}_{\text{diff}}}{se(\overline{x}_{\text{after}} - \overline{x}_{\text{before}})}$

(3) $\dfrac{\overline{x}_{\text{after}} - \overline{x}_{\text{before}}}{se(\overline{x}_{\text{after}} - \overline{x}_{\text{before}})}$

(4) $\dfrac{\overline{x}_{\text{after}} - \overline{x}_{\text{before}}}{se(\overline{x}_{\text{after}} + \overline{x}_{\text{before}})}$

(5) $\dfrac{\overline{x}_{\text{diff}}}{se(\overline{x}_{\text{diff}})}$

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 after - before	-69.25	14.60	5.16	-81.46	-57.04	-13.42	7	.000

22. The best description for the paired test on the heartbeats of the rats would be:

(1) With 95% confidence we estimate that on average the heartbeats of the rats is 69.25 beats

(2) With 95% confidence we estimate that on average the heartbeat of the rats before taking the drug was between 57.04 and 81.46 beats lower than the heartbeat of the rats after taking the drug.

(3) With 95% confidence we estimate that on average the heartbeat of the rats before taking the drug was between 57.04 and 81.46 beats higher than the heartbeat of the rats after taking the drug.

(4) With 95% confidence we estimate that the difference in the rats was between 57.04 and 81.46 heartbeats.

(5) With 95% confidence we estimate that on average the heartbeat of the rats after taking the drug was between 57.04 and 81.46 beats higher than the heartbeat of the rats before taking the drug.

23. For an *F*-test to be valid, which of the assumptions listed below are **required**?

I The samples are independent.
II The underlying means (i.e. the population means) are equal.
III The underlying level of variability is the same for each of the groups.
IV The sample sizes are equal.
V The underlying distribution of each group is Normal.

(1) II, III and V
(2) I, III and V
(3) II, III and IV
(4) I, II and V
(5) I, IV and V

24. Which **one** of the following statements gives the **correct** hypotheses for an *F*-test?

(1) H_0: all of the μ's are equal
H_1: none of the μ's are equal
(2) H_0: not all of the μ's are equal
H_1: all of the μ's are equal
(3) H_0: all of the μ's are equal
H_1: at least one of the μ's is different
(4) H_0: none of the μ's are equal
H_1: some of the μ's are equal
(5) H_0: some of the μ's are equal
H_1: not all of the μ's are equal

Question 25 refers to the following additional information.

A medical study was carried out to test the effectiveness of a new sleeping drug. It was hoped that the drug would be suitable for the New Zealand market. Ten people who had recently been diagnosed with having a particular type of sleeping disorder were used as subjects in the study. They were all given the drug on one night and then, on the next night, they were all given a placebo. The subjects could not tell beforehand, and nor were they told, which was the drug and which was the placebo. On each of the two nights, each subject was measured for the number of hours of sleep. The results are shown below.

	Hours of Sleep		
Patient	**Drug**	**Placebo**	**Difference**
1	6.1	5.2	0.9
2	7.0	7.9	-0.9
3	8.2	3.9	4.3
4	7.6	4.7	2.9
5	6.5	5.3	1.2
6	8.4	4.1	3.0
7	6.9	4.2	2.7
8	6.7	6.1	0.6
9	7.4	3.8	3.6
10	5.8	6.3	-0.5

Patient	Drug	Placebo	Difference
Sample Mean	7.06	5.28	1.78
Sample Std. Dev.	0.85	1.26	1.77

To determine the efficacy of the drug, the researcher wanted to see if there was a difference between the average number of hours of sleep when the drug is taken and the average number of hours of sleep when the placebo is taken.

25. An appropriate test to perform is a:

(1) Sign test on the differences.

(2) Two sample *t*-test if plots of the two samples are **not** severely non-Normal.

(3) *t*-test on the differences if a plot of the differences displays severe non-Normality.

(4) Wilcoxon Rank-Sum test on the two samples if plots of the two samples are severely non-Normal.

(5) Mann-Whitney test on the differences.

Questions 26 and **27** refer to the following information.

It has already been established that increased reproduction decreases longevity of female fruitflies. Therefore, an experiment was designed to test whether increased reproduction also reduces longevity for male fruitflies. Longevity is the life span (i.e. how long they live). Each male fruitfly was randomly assigned to one of five groups. There were **twenty-five** male fruitflies in each group. This is the variable GP.

The five groups are:

GP1: Male forced to live alone
GP2 Male lives with one receptive female, i.e. the female is willing to mate.
GP3 Male lives with one non-receptive female.
GP4 Male lives with 8 non-receptive females.
GP5 Male lives with 8 receptive females.

One-Way Analysis of Variance (ANOVA) for Longevity

	Deg. of freedom	Sum of Squares	Mean Sum of Squares	*F*-statistic	*p*-value
Between groups	**	11939.28	2984.82	***	0
Within groups	**	26313.52	219.28		
Total		38252.80			

26. The degrees of freedom for the test statistic, f_0, for this F-test are:

(1) $df_1 = 5$, $df_2 = 125$
(2) $df_1 = 4$, $df_2 = 120$
(3) $df_1 = 120$, $df_2 = 4$
(4) $df_1 = 4$, $df_2 = 124$
(5) $df_1 = 125$, $df_2 = 5$

27. The value of the test statistic, f_0, for this F-test is:

(1) 79.20
(2) 654,500
(3) 13.61
(4) 0.0735
(5) 0.4537

Questions 28 and **29** refer to the following information.

The *NZ Herald*, 14 August 2001, reported the results of a two-year study at Hong Kong's Kwong Wah Hospital. The study comprised 5450 impotent men who were given Viagra at the hospital. Of these men, 3651 were smokers. Of the 926 impotent men for whom Viagra did not work, 840 were smokers.

Assume this sample of 5450 impotent men is a random sample of all impotent Chinese men and recall that Viagra did **not** work for 926 of the men in this sample. Suppose it is known that Viagra works for 78% of impotent Western males.

28. Which **one** of the following statements is **true**?
    (1) The proportion of Chinese men in the sample for whom Viagra works is **both** an **estimate** value and a **parameter** value.
    (2) The proportion of Chinese men in the sample for whom Viagra works is an **estimate** value whereas the '78%' of impotent Western males for whom Viagra works is a **parameter** value.
    (3) The proportion of Chinese men in the sample for whom Viagra works and the '78%' of impotent Western males for whom Viagra works are **both parameter** values.
    (4) The proportion of Chinese men in the sample for whom Viagra works and the '78%' of impotent Western males for whom Viagra works are **neither estimates nor parameter** values.
    (5) The proportion of Chinese men in the sample for whom Viagra works is a **parameter** value whereas the '78%' of impotent Western males for whom Viagra works is an **estimate** value.

29. Suppose the results of this study are used to conduct a *t*-test to see if the percentage of impotent Chinese men for whom Viagra works has the same value as the percentage of impotent Western men for whom it is known to work. The formula for calculating the approximate standard error of the estimate in this test is:

(1) $\sqrt{\dfrac{0.67(1 - 0.67)}{5450} + \dfrac{0.78(1 - 0.78)}{5450}}$

(2) $\sqrt{\dfrac{0.83(1 - 0.83)}{5450} + \dfrac{0.78(1 - 0.78)}{5450}}$

(3) $\sqrt{\dfrac{(0.83 + 0.78) - (0.83 - 0.78)^2}{5450}}$

(4) $\sqrt{\dfrac{0.83(1 - 0.83)}{5450}}$

(5) $\sqrt{\dfrac{0.67(1 - 0.67)}{5450}}$

30. Which **one** of the following statements about an *F*-test is **false**?
    (1) A decrease in the size of the differences between the group means will result in a decrease in evidence against the hypothesis that the underlying true group means are the same (given the variability/spread within each group remains unchanged).
    (2) The larger the value of the *F*-test statistic, f_0, the smaller the *P-value*.
    (3) An increase in the size of the differences between the group means will result in an increase in evidence against the hypothesis that the underlying true group means are the same.
    (4) An increase in the spread of the data within each group will result in an increase in evidence against the hypothesis that the underlying true group means are the same (given the size of the differences between the group means are the same).
    (5) The value of the *F*-test statistic, f_0, is the ratio of the between-mean variation and the within-group variation.

Questions 31 to **36** refer to the following information.

New Zealand drug survey results obtained in 1998 and 2001 were compared by the Alcohol and Public Health Research Unit (NZ). In 1998, the nationwide study contained 5,475 randomly selected respondents and in 2001 the study contained 5,504 randomly selected respondents.

In both surveys, drug users were asked if they had wanted rehabilitation for their drug use but had not received it. Of those respondents who felt they had not received the rehabilitation they needed, they were then asked the reasons why. Each person could give more than one reason and their responses are listed in Table 12.

Drug rehabilitation barriers	**1998** n = 146	**2001** n = 166
Didn't know where to go	33%	32%
Social pressure	28%	8%
Fear of what would happen on contacting service	23%	23%
No time/too busy	20%	19%
Services too expensive	17%	14%
Fear of losing friends	15%	9%
Fear of law/Police	14%	29%
No local services available	11%	10%
Transport problems	7%	6%
Services weren't ongoing	4%	1%
Others	18%	24%

Table 12: Barriers said to have prevented drug users obtaining rehabilitation

31. Let p_{pressure} be the proportion of people who felt social pressure prevented them receiving drug rehabilitation in **1998**, and p_{friends} the proportion of people who felt the fear of losing friends prevented them receiving drug rehabilitation in **1998**. An estimate of the difference between p_{pressure} and p_{friends} is:
    (1) 0.19
    (2) 0.07
    (3) 0.01
    (4) 0.13
    (5) 0.00

32. In order to calculate the standard error of the difference between $\hat{p}_{\text{pressure}}$ and $\hat{p}_{\text{friends}}$, the sampling situation can be described as:
    (1) Two independent samples of size 146 and 166.
    (2) One sample of size 146, many yes/no items.
    (3) One sample of size 146, several response categories.
    (4) One sample of size 166, several response categories.
    (5) Two independent samples, each of size 146.

Questions 33 to **36** refer to the following additional information.

Let p_{98} be the proportion of people who said social pressure prevented them receiving drug rehabilitation in **1998** and p_{01} be the proportion of people who said social pressure prevented them receiving drug rehabilitation in **2001**.

33. The **correct** standard error of the difference between $\hat{p}_{98}$ and $\hat{p}_{01}$, is:

(1) $\sqrt{\dfrac{\hat{p}_{98}(1-\hat{p}_{98})}{146} - \dfrac{\hat{p}_{01}(1-\hat{p}_{01})}{166}}$

(2) $\sqrt{\dfrac{\hat{p}_{98}(1-\hat{p}_{98})}{146} + \dfrac{\hat{p}_{01}(1-\hat{p}_{01})}{166}}$

(3) $\sqrt{\dfrac{(\hat{p}_{98}+\hat{p}_{01})-(\hat{p}_{98}-\hat{p}_{01})^2}{146}}$

(4) $\sqrt{\dfrac{(\hat{p}_{98}+\hat{p}_{01})-(\hat{p}_{98}+\hat{p}_{01})^2}{166}}$

(5) $\sqrt{\dfrac{(1-\hat{p}_{98})+(1-\hat{p}_{01})-(\hat{p}_{98}-\hat{p}_{01})^2}{146}}$

34. In a test for no difference between p_{98} and p_{01} (i.e. $p_{98} - p_{01} = 0$), the estimate of the difference is 0.2, the test statistic is 4.68 and the *P-value* is 0.000. Which **one** of the following statements is **false**?
(1) It is impossible for sampling variability alone to result in an estimate of a difference larger than 0.2.
(2) There is very strong evidence of a difference between p_{98} and p_{01}.
(3) Social pressure was stated significantly less often as a reason for not receiving drug rehabilitation in 2001 compared to 1998.
(4) The *P-value* indicates that the *t*-test is significant at the 1% level of significance.
(5) The estimate of the difference is approximately 4.7 standard errors above the hypothesised difference.

35. The 95% confidence interval for the difference $p_{98} - p_{01}$ is (0.0972, 0.3028). The **best** interpretation of this interval is:
With 95% confidence, the proportion of drug users who felt social pressure prevented them receiving drug rehabilitation:
(1) fell by somewhere between 10 and 30 percentage points between 1998 and 2001.
(2) was somewhere between 10 percentage points higher and 30 percentage points lower in 2001 compared with 1998.
(3) increased by somewhere between 10 and 30 percentage points between 1998 and 2001.
(4) was 10 percentage points in 1998 and 30 percentage points in 2001.
(5) was somewhere between 10 percentage points lower and 30 percentage points higher in 2001 compared with 1998.

36. The formula for the *t*-test statistic for $H_0 : p_{98} - p_{01} = 0$, is:

(1) $\dfrac{\hat{p}_{98}-\hat{p}_{01}}{\sqrt{se(\hat{p}_{98})^2 - se(\hat{p}_{01})^2}}$

(2) $\dfrac{p_{98}-p_{01}}{se(p_{98}-p_{01})}$

(3) $\dfrac{\hat{p}_{98}-\hat{p}_{01}}{se(\hat{p}_{98}-\hat{p}_{01})}$

(4) $\dfrac{\hat{p}_{98}-\hat{p}_{01}}{se(\hat{p}_{98})+se(\hat{p}_{01})}$

(5) $\dfrac{p_{98}-p_{01}}{se(\hat{p}_{98})+se(\hat{p}_{01})}$

Questions 37 to **41** refer to the following information.

The Washington Post, The Henry J Kaiser Family Foundation and *Harvard University* conducted a poll (8 March – 22 April, 2001) 'to gauge the racial attitudes of American adults'. The telephone poll surveyed 1709 adults including 779 whites, 323 African Americans, 315 Hispanics and 254 Asian Americans. Assume this sample of 1709 adults is a random sample of American adults. Two of the questions in the survey were:

Question 1:

Do you feel that African Americans have more, less or about the same opportunities in life as whites have?

and

Question 2:

Do you feel that Asian Americans have more, less or about the same opportunities in life as whites have?

The percentage results for these two questions are shown in Table 7 below.

	Response				Sample size
	More %	Less %	Same %	Unsure %	
Question 1					
White	13	27	58	2	779
African American	1	74	23	2	323
Hispanic	8	46	44	2	315
Asian American	10	44	39	7	254
Total Sample	11	35	51	2	1709
Question 2					
White	13	14	70	4	779
African American	15	38	39	8	323
Hispanic	18	24	55	3	315
Asian American	7	34	53	5	254
Total Sample	14	18	63	4	1709

Table 7: Americans' responses to racial attitudes survey

Let p_{more} be the proportion of **whites** who feel that African Americans have **more** opportunities in life than whites have and p_{less} be the proportion of **whites** who feel that African Americans have **less** opportunities in life than whites have. (Ie, whites' responses to Question 1.)

37. An estimate of the difference between p_{more} and p_{less} is:
    (1) 0.018
    (2) -0.01
    (3) -0.14
    (4) 0.01
    (5) -0.10

38. Information from Table 7 above, is used to construct a 95% confidence interval for the difference p_{more} - p_{less}. For the purpose of calculating $se(\hat{p}_{more} - \hat{p}_{less})$, the sampling situation can be described as:
    (1) two independent samples of sizes 779 and 254.
    (2) one sample of size 779, several response categories.
    (3) one sample of size 1709, many yes/no items.
    (4) one sample of size 1709, several response categories.
    (5) one sample of size 779, many yes/no items.

39. A 95% confidence interval for the difference p_{more} - p_{less} is (-0.1833, -0.09668). The **best** interpretation of this interval is:

    With 95% confidence, the percentage of whites who feel that African Americans have more opportunities in life than whites have is somewhere between:

    (1) 10% higher than and 18% lower than the percentage who feel that African Americans have less opportunities in life than whites have.
    (2) 10% and 18%.
    (3) 10% and 18% higher than the percentage who feel that African Americans have less opportunities in life than whites have.
    (4) 10% lower than and 18% higher than the percentage who feel that African Americans have less opportunities in life than whites have.
    (5) 10% and 18% lower than the percentage who feel that African Americans have less opportunities in life than whites have.

Questions 40 and **41** refer to the following additional information.

Let:

$p_{question1}$ be the proportion of Asian Americans who feel that African Americans have more opportunities in life than whites have

and

$p_{question2}$ be the proportion of Asian Americans who feel that Asian Americans have more opportunities in life than whites have.

Information from Table 7, is used to conduct a 2-tailed test for no difference between $p_{question1}$ and $p_{question2}$.

40. The expression for evaluating the test statistic for the null hypothesis, H_0: $p_{question1} - p_{question2} = 0$, is:

(1) $\dfrac{\hat{p}_{question1} - \hat{p}_{question2}}{se(\hat{p}_{question1}) + se(\hat{p}_{question2})}$

(2) $\dfrac{\hat{p}_{question1} - \hat{p}_{question2}}{\sqrt{se(\hat{p}_{question1})^2 - se(\hat{p}_{question2})^2}}$

(3) $\dfrac{p_{question1} - p_{question2}}{se(\hat{p}_{question1}) + se(\hat{p}_{question2})}$

(4) $\dfrac{\hat{p}_{question1} - \hat{p}_{question2}}{se(\hat{p}_{question1} - \hat{p}_{question2})}$

(5) $\dfrac{p_{question1} - p_{question2}}{se(\hat{p}_{question1} - \hat{p}_{question2})}$

41. The formula for the standard error of the estimate, $se(\hat{p}_{question1} - \hat{p}_{question2})$, is:

(1) $\sqrt{\dfrac{(1 - \hat{p}_{question1}) + (1 - \hat{p}_{question2}) - (\hat{p}_{question1} - \hat{p}_{question2})^2}{254}}$

(2) $\sqrt{\dfrac{(\hat{p}_{question1} + \hat{p}_{question2}) - (\hat{p}_{question1} + \hat{p}_{question2})^2}{254}}$

(3) $\sqrt{\dfrac{\hat{p}_{question1}(1 - \hat{p}_{question1})}{254} + \dfrac{\hat{p}_{question2}(1 - \hat{p}_{question2})}{254}}$

(4) $\sqrt{\dfrac{\hat{p}^2_{question1}}{254} - \dfrac{\hat{p}^2_{question2}}{254}}$

(5) $\sqrt{\dfrac{(\hat{p}_{question1} + \hat{p}_{question2}) - (\hat{p}_{question1} - \hat{p}_{question2})^2}{254}}$

Questions 42 to **44** refer to the following information.

During 1999, students enrolled in stage one statistics at the University of Auckland were surveyed regarding their access to, and experience with, computers. The survey was included as a question in an assignment, and students were given marks for completing it (irrespective of the answers they gave). Staff administering the courses wished to use the results of this survey to draw conclusions about future stage one statistics students.

One question asked: 'At the start of the course, how would you describe your Excel experience?'. A total of 918 students answered this question. Each of the 918 answers were classified according to the response given by the student, and the stream the student attended. The results are given in the table below, where 107, 108 and 101 refer to the various streams.

	Stream			
Response	107	108	101	Total
None	15	36	102	153
Very Little	44	89	119	252
Some	74	150	200	424
Lots	9	29	51	89
Total	142	304	472	918

42. A statistical test is performed on the data for no difference between the proportion of 107 students who responded None and the proportion of 101 students who responded None. Which one of the following statements about this test is **true**?

(1) The degrees of freedom used depends on the number of 107 and 101 students in the sample.

(2) The Wilcoxon rank-sum test could be used.

(3) The test should be two–tailed.
(4) The test could only be used to show a difference existed in the sample proportions.
(5) An appropriate null hypothesis is that the difference between the proportion of 107 students who responded None and the proportion of 101 students who responded None, is not zero.

43. The standard error for the difference in the proportions tested in question 42 is:
(1) 0.033
(2) 0.022
(3) 0.023
(4) 0.032
(5) 0.184

44. The *P-value* for the statistical test mentioned in question 42 is 0.004. Which one of the following statements gives the best interpretation of this *P–value*?
(1) There is some evidence that the underlying proportions of 107 students and 101 students that would respond None are different.
(2) There is strong evidence that the sample proportions of 107 students and 101 students that responded None are different.
(3) There is strong evidence that the underlying proportions of 107 students and 101 students that would respond None are different.
(4) There is no evidence that the underlying proportions of 107 students and 101 students that would respond None are different.
(5) There is weak evidence that the underlying proportions of 107 students and 101 students that would respond None are different.

45. Which one of the following statements is false?
(1) In a *t*-test for no difference between two proportions, being able to demonstrate that the difference was of practical significance (importance) would almost always imply statistical significance.
(2) In hypothesis testing, large samples can lead to small *P-values* without the results having any practical significance (importance).
(3) In hypothesis testing, statistical significance does not imply practical significance (importance).
(4) In a hypothesis test for no difference between two proportions, a very small *P-value* always indicates a very large difference in the proportions.
(5) In hypothesis testing, a nonsignificant test result does not imply that the null hypothesis is true.

Answers

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

Review the Basics, S1, 2010

Code	Workshop Title	Day	Date	Time	Tutor	Room
1a	Surveys, Experiments & Exploring Data (Ch 1 ~ 3)	Sat	2-Oct	9.30am-midday	Leila	LibB15
1b	Surveys, Experiments & Exploring Data (Ch 1 ~ 3) [repeat]	Wed	27-Oct	1-3.30pm	Leila	LibB15
2a	Probabilities, Normal & Sampling Distributions (Ch 4, 6 & 7)	Sat	2-Oct	1-3.30pm	Leila	LibB15
2b	Probabilities, Normal & Sampling Distributions (Ch 4, 6 & 7) [repeat]	Wed	27-Oct	9.30am-midday	Leila	LibB15
3a	CIs & Tests for Proportions (Ch 8 & 9 for Proportions)	Sat	16-Oct	9.30am-midday	Leila	LibB15
3b	CIs & Tests for Proportions (Ch 8 & 9 for Proportions) [repeat]	Tue	26-Oct	9.30am-midday	Leila	LibB15
4a	CIs & Tests for Means (Ch 8 ~ 10 for Means)	Sat	16-Oct	1-3.30pm	Leila	LibB15
4b	CIs & Tests for Means (Ch 8 ~ 10 for Means) [repeat]	Tue	26-Oct	1-3.30pm	Leila	LibB15
5a	Chi-Square Tests (Ch 11)	Mon	25-Oct	9.30am-midday	Leila	LibB15
5b	Chi-Square Tests (Ch 11) [repeat]	Mon	1-Nov	1-3.30pm	Leila	Case Room 3
6a	Regression (Ch 12)	Mon	25-Oct	1-3.30pm	Leila	LibB15
6b	Regression (Ch 12) [repeat]	Mon	1-Nov	9.30am-midday	Leila	Case Room 3

Also available for enrolments:

Skills in Stats IV - Ch 8-10	Sat	9-Oct	9.30am-12.30pm	Leila	LibB15
Skills in Stats V - Ch 11 & 12	Sat	23-Oct	1-4pm	Leila	LibB15