Stats 101/101G/108 Workshop

Confidence Intervals:
Means [CIM]

2019

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-CIM**

Leila Boyle
Undergraduate Statistics Assistance, Department of Statistics
Room 303.320 (third floor of the Science Centre, Building 303)
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

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!

- **Preparation at the beginning of the semester:**

  Multiple identical sessions of a preparation workshop are run at the beginning of the semester to get students off to a good start – come along to whichever one suits your schedule!

  - Basic Maths and Calculator skills for Statistics **www.tinyURL.com/stats-BM**

- **First half of the semester**

  Five theory workshops are held during the first half of the semester:

  - Exploratory Data Analysis **www.tinyURL.com/stats-EDA**
  - Proportions and Proportional Reasoning **www.tinyURL.com/stats-PPR**
  - Observational Studies, Experiments, Polls and Surveys **www.tinyURL.com/stats-OSE**
  - Confidence Intervals: *Means* **www.tinyURL.com/stats-CIM**
  - Confidence Intervals: *Proportions* **www.tinyURL.com/stats-CIP**

- **Second half of the semester**

  Five theory workshops and one computing workshop are held during the second half of the semester:

- **Statistics Theory Workshops**

  - Hypothesis Tests: *Proportions* **www.tinyURL.com/stats-HTP**
  - Hypothesis Tests: *Means (part 1)* **www.tinyURL.com/stats-HTM**
  - Hypothesis Tests: *Means (part 2)* **www.tinyURL.com/stats-HTM**
  - Chi-Square Tests **www.tinyURL.com/stats-CST**
  - Regression and Correlation **www.tinyURL.com/stats-RC**

- **Computer Workshop:** Hypothesis Tests in *SPSS* **www.tinyURL.com/stats-HTS**

- **Useful Computer Resource:**

  If you haven't used SPSS before, you may find it useful to work your way through this self-paced workshop: **www.tinyURL.com/stats-IS**

A note about rounding numbers

Often students ask me about rounding numbers – how to do it and how much by.

When you do a calculation, you may end up with an answer that has many (5 or more) decimal places associated with it. People don't deal too well with numbers to this level of accuracy; rounding helps us make a number a little simpler while still keeping its value relatively close to what it was. The result is less accurate, but easier to use, interpret and understand.

How to round numbers

- Decide which is the last digit to keep
- Leave it the same if the next digit is less than 5 (this is called **rounding down**)
- Increase it by 1 if the next digit is 5 or more (this is called **rounding up**)

How many decimal places should I use?

When you are doing calculations in an assignment or test/exam context, you should get some guidance from the question itself. For example, a multi-choice test question may have five answers that are all rounded to one decimal place (1dp for short) so just round your answer when you do the calculation to 1dp.

If you are going to use the value you come up with in a later calculation, go for more accuracy rather than less, as the more you round a number, the more it will affect the results of later calculations that use that value. My default amount of rounding tends to be 4dp – this is reasonably accurate without going over the top!

Recall from Chapter 1:

Types of Variables

Numeric
or Scale (SPSS)
(measurements and counts)

Continuous
(measurements; few repeated values)

No gaps between possible values, if measured sufficiently accurately.

Examples
Weight; Height; Length; Temperature; Total course mark (out of 100); Age [large range]

Discrete
(usually counts; many repeated values)

Gaps between each of the values it can take.

Examples
Number of siblings; Number of cars per household; Assignment mark (out of 10); Age [small range]

Categorical
(define groups)

Nominal (SPSS)
(no idea of order)

Examples
Gender; Ethnicity; Degree (BA, BCom, BSc, etc.); Major (Chemistry, Statistics, etc.)

Ordinal (SPSS)
(fall in natural order)

Examples
Size (small, medium, large); Income bracket; Grade (A+, A, …); Age group (<25, 25-39, 40+)

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

- **Numerical summaries**
  - **Centre:** describes the tendency of the observations to bunch around a particular value – in Chapter 6 we use the **sample mean**, $\bar{x}$
  - **Spread:** describes the dispersion of the observed values – in Chapter 6 we use the **sample standard deviation**, s

Exploring relationships between two variables

Two Numeric | **Numeric versus Categorical** | **Two Categorical**

Exploratory tools to use to explore relationship/s

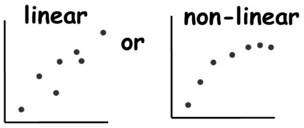
Two Numeric	Numeric versus Categorical	Two Categorical
Scatter plot	Side-by-side plots on the same scale: any n { • Dot plots $n \geq 20$ { • Stem-and-leaf plots • Box plots $n \geq 50$ { • Histograms	Two-way table of counts and/or Bar graphs of proportions (on rows/columns)

Features to look for:

Two Numeric

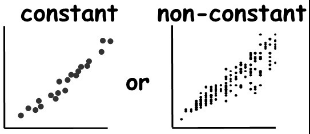
- Trend

linear or **non-linear**



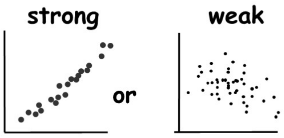
- Scatter

constant or **non-constant**



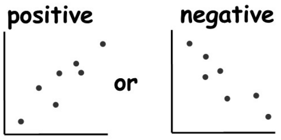
- Strength of relationship

strong or **weak**

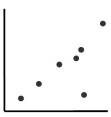


- Association

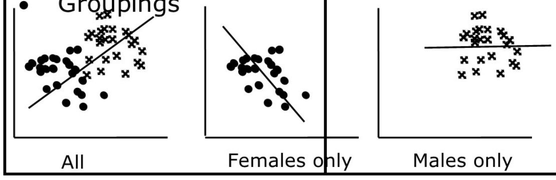
positive or **negative**



- Outliers



- Groupings



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)

Understanding Confidence Intervals

- Statistics is concerned with finding out about the real world and aspects of it specific to our area of interest. Statistical tools allow us to deal with the **uncertainty** present in all samples due to **sampling variation** which occurs because we are unable to survey the entire population of interest.

- We are usually unable to survey the entire population (take a census) as it is too large and/or there are:
  - ✓ budget constraints
  - ✓ time limits
  - ✓ logistical barriers

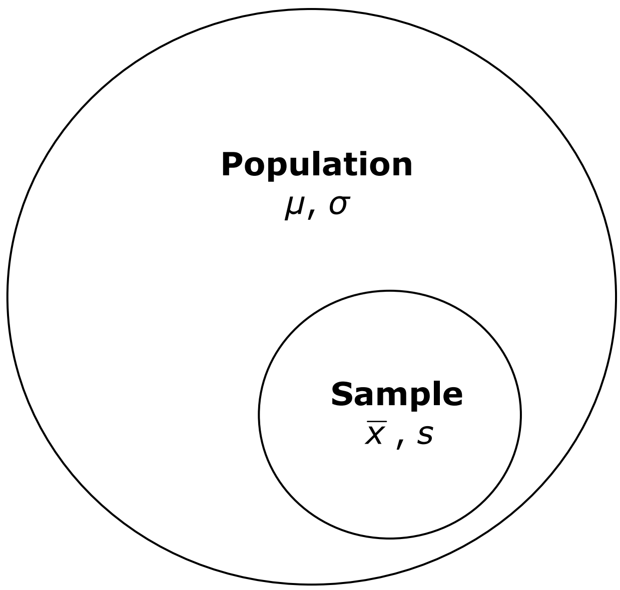
- This means we are unable to establish the **parameters** of interest within our population, such as:
  - ✓ Population mean, μ; and
  - ✓ Population standard deviation, σ

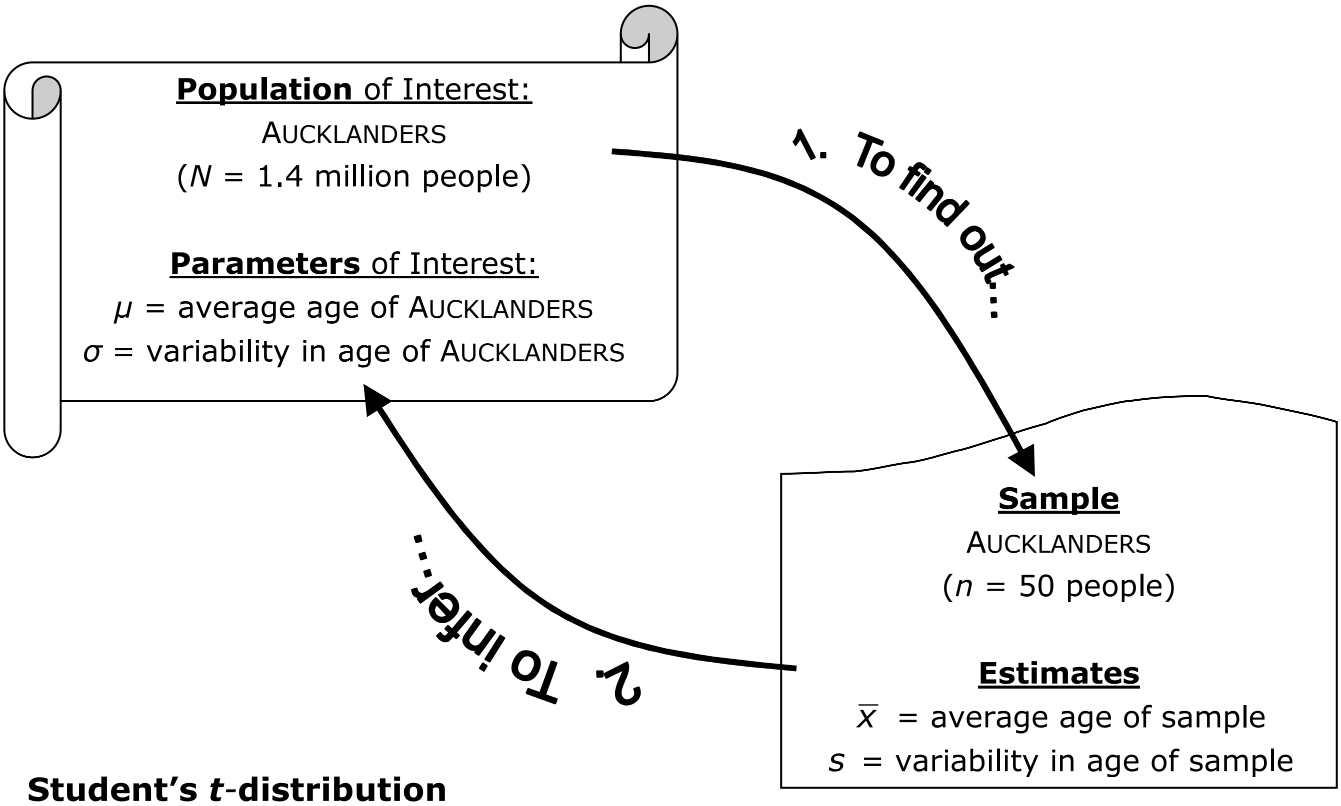
- This means that the **parameter** of interest is an ***unknown* numerical *characteristic*** for that particular population.

- To estimate an ***unknown* numerical characteristic (parameter)** for our population of interest, we take a sample and find a sample **estimate** from it The **sample estimates** of the above **population parameters** are:
  - ✓ Sample mean, $\bar{x}$; and
  - ✓ Sample standard deviation, s

- Usually ^HATS or ¯BARS are used to distinguish between **sample estimates** and **population parameters**.

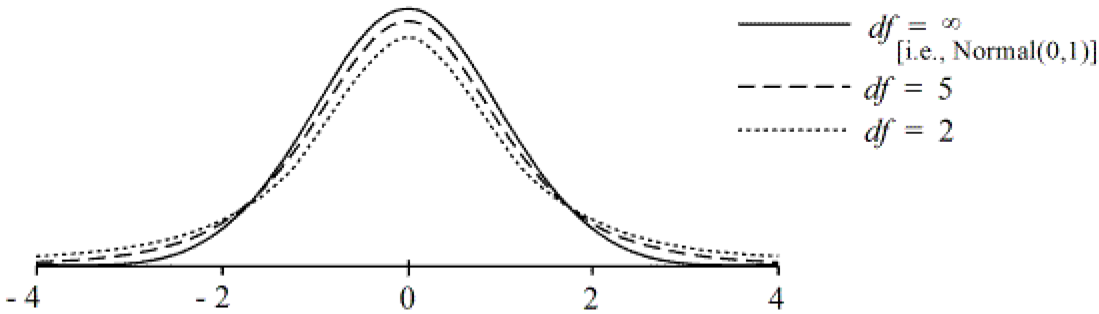
- The process of using sample data to make useful statements about a population is a type of **statistical inference** called **sample-to-population inference**, e.g., when using sample data to estimate an unknown population average.





Student's *t*-distribution

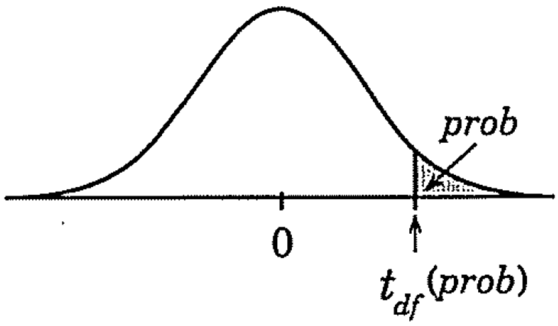
- ✓ Parameter: Degrees of Freedom (*df*).
- ✓ Smooth symmetric, bell-shaped curve centred at 0 like the Standard Normal distribution [Z ~ Normal (μ = 0, σ = 1)] but it's more variable (is more spread out).



- ✓ As *df* becomes larger, the Student (*df*) distribution becomes more and more like the Standard Normal distribution.
- ✓ Student's *t*-distribution ($df = \infty$) and Normal (0,1) are the same distribution.
- ✓ Methods based on this distribution work very well even for small samples that are from very non-Normal distributions.

Student's t–distribution

For fixed $prob$ and df, the tabulated value is the number ($t = t_{df}(prob)$) such that for, $T \sim \text{Student}(df)$, $pr(T \geq t) = prob$.



[e.g. For $prob = 0.025$ and $df = 23, t_{23}(0.025) = 2.069$]

95% CI

df	.20	.15	.10	.05	.025	.01	.005	.001	.0005	.0001
3	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.21	12.92	22.20
4	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610	13.03
5	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869	9.678
6	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959	8.025
7	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7.063
8	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041	6.442
9	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781	6.010
10	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587	5.694
11	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437	5.453
12	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318	5.263
13	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221	5.111
14	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140	4.985
15	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073	4.880
16	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015	4.791
17	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965	4.714
18	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922	4.648
19	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883	4.590
20	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.849	4.539
21	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819	4.493
22	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792	4.452
23	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768	4.415
24	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745	4.382
25	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725	4.352
26	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707	4.324
27	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690	4.299
28	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674	4.275
29	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659	4.254
30	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646	4.234
31	0.853	1.054	1.309	1.696	2.040	2.453	2.744	3.375	3.633	4.215
32	0.853	1.054	1.309	1.694	2.037	2.449	2.738	3.365	3.622	4.198
33	0.853	1.053	1.308	1.692	2.035	2.445	2.733	3.356	3.611	4.182
34	0.852	1.052	1.307	1.691	2.032	2.441	2.728	3.348	3.601	4.167
35	0.852	1.052	1.306	1.690	2.030	2.438	2.724	3.340	3.591	4.153
36	0.852	1.052	1.306	1.688	2.028	2.434	2.719	3.333	3.582	4.140
37	0.851	1.051	1.305	1.687	2.026	2.431	2.715	3.326	3.574	4.127
38	0.851	1.051	1.304	1.686	2.024	2.429	2.712	3.319	3.566	4.116
39	0.851	1.050	1.304	1.685	2.023	2.426	2.708	3.313	3.558	4.105
40	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551	4.094
45	0.850	1.049	1.301	1.679	2.014	2.412	2.690	3.281	3.520	4.049
50	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.261	3.496	4.014
60	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460	3.962
80	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416	3.899
100	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390	3.861
∞	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.719

[Note: If $df \geq 150$ use $df = \infty$]

Confidence Intervals

- 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.

80% CI, $\bar{x} \pm 1.282\ \text{se}(\bar{x})$

90% CI, $\bar{x} \pm 1.645\ \text{se}(\bar{x})$

95% CI, $\bar{x} \pm 1.960\ \text{se}(\bar{x})$

99% CI, $\bar{x} \pm 2.576\ \text{se}(\bar{x})$

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 (precision) of the confidence interval we **need 4 times** as many observations.

- To triple the accuracy (precision) 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.

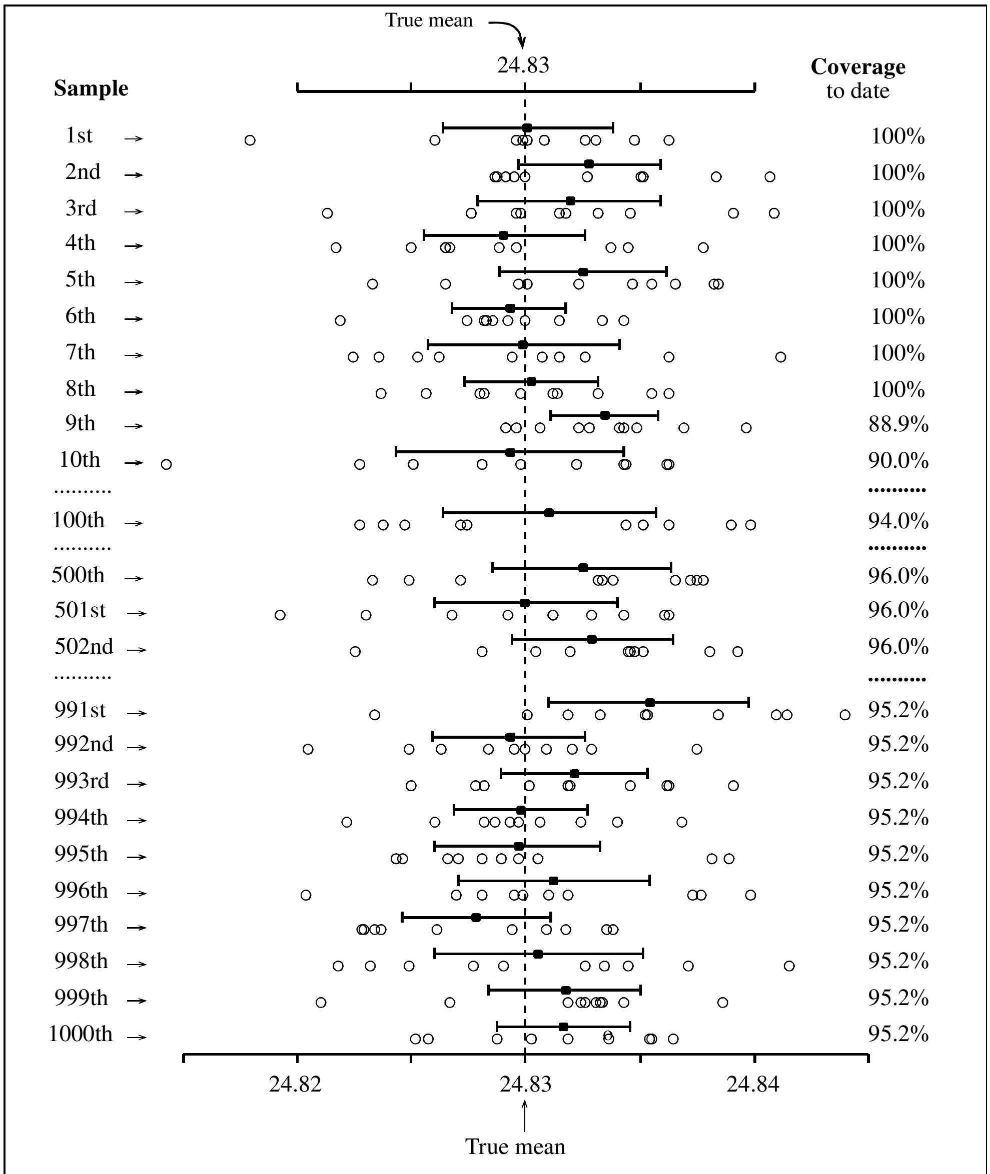


Figure 8.1.2 Samples of size 10 from a Normal(μ=24.83, s=.005) distribution and their 95% confidence intervals for μ..

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, from the Formula Sheet — ***estimate* ± *t* × se(*estimate*)**

4. Use the appropriate **standard error**. (Will be provided)

5. Use the appropriate ***t*-multiplier**. (Will be provided)

6. Calculate the **confidence limits**. (End points of the confidence interval)

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

see back page for Formulae Sheet

- There are four different types of problem we will consider:
  1. Single mean 2. Single proportion. 3. Difference between two means
  4. Difference between two proportions

***estimate* ± *t* × se(*estimate*)**

First piece of CI formula / Step 2: ***estimate* ± *t* × se(*estimate*)**

- The **estimate** is based on the **parameter** of interest we are investigating:

Parameter	**Estimate**
1. Single mean $\boldsymbol{\mu}$:	*estimate* = $\overline{x}$
3. Difference between two means $\boldsymbol{\mu_1 - \mu_2}$: (independent samples)	*estimate* = $\overline{x}_1 - \overline{x}_2$

<u>Second piece of CI formula / Step 5:</u> **$estimate \pm t \times \text{se}(estimate)$**

- The ***t*-multiplier** is based on:
  - ✓ The desired level of confidence
  - ✓ The degrees of freedom

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

- The ***t*-multiplier** can be found from the *t*-procedures spreadsheet. In the test/exam situation, the *t*-multiplier will be provided.

<u>Third piece of CI formulae / Step 4:</u> **$estimate \pm t \times \text{se}(estimate)$**

- The **standard error** can be found from the *t*-procedures spreadsheet. In the test/exam situation, the standard error will be provided.

<u>Interpreting the CI limits → Step 7 for story type 3:</u>

- CIs for the difference between two means/proportions:

Examples:

  - ✓ If the CI contains 0 (i.e. one negative and one positive number), there may be no difference between the two means. (-5, 3)
  - ✓ If CI is positive, then μ_1 is higher/larger than μ_2. (3, 5)
  - ✓ If CI is negative, then μ_1 is lower/smaller than μ_2. (-5, -3)

Check you understand! – Practice Questions

1. Printed on every packet of "Yummo" corn chips is a weight of 150g. Let μ be true (population) mean weight of "Yummo" corn chip packets.

   A consumer collects a random sample of 48 packets of "Yummo" corn chips and finds a mean weight of 148.5g and a standard deviation of 2.1g.

   The 95% confidence interval for μ is:

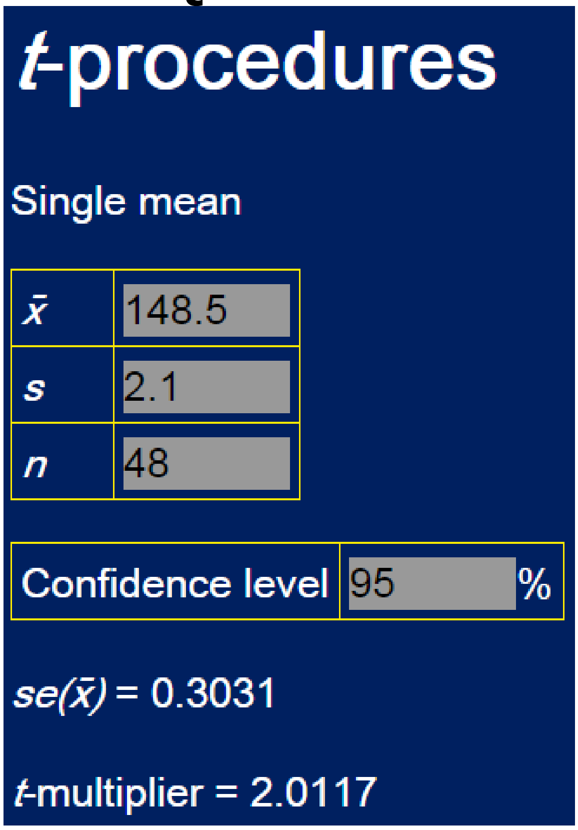
   (1) [147.09, 149.91]

   (2) [148.20, 148.80]

   (3) [147.89, 149.11]

   (4) [147.11, 149.89]

   (5) [147.32, 149.68]



Question 2 refers to the following information.

The heights (in cm) of the carapaces (shells) of a sample of 48 painted turtles were recorded. Shown below is a stem-and-leaf plot of this data set.

Units: 3|5 = 35cm

3 | 55778888999

4 | 0000111222344

4 | 55566789

5 | 0111113

5 | 567

6 | 01233

6 | 7

Figure: Stem-and-leaf plot of carapace height of painted turtles.

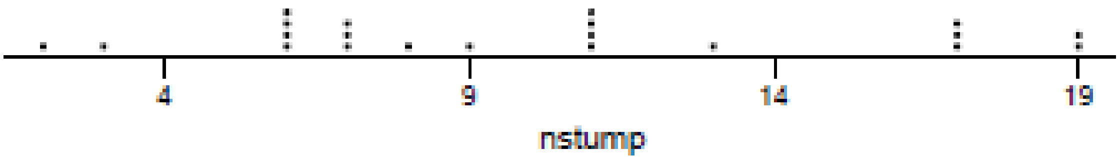
2. Based on this sample of size 48, a 95% confidence interval for the underlying mean carapace height of all painted turtles is 44.0cm to 48.8cm. The number of painted turtle carapace heights we would need to sample in order to halve the width of this interval is, approximately:

   (1) 24
   (2) 192
   (3) 12
   (4) 96
   (5) 72

Questions 3 and **4** refer to the following information.

Data were collected on the number of cricket test match innings in which the batsman was stumped by the wicket keeper. The number of stumping dismissals per year (nstump) were recorded for the years 1970 to 1990 (inclusive). Summary statistics and a dot plot of these data are given below:

	n	$\bar{x}$	s	se($\bar{x}$)	t-multiplier
nstump	21	10.14	5.15	1.1238	2.086



3. A 95% confidence interval for the cricket data was calculated to be (7.80, 12.49). Which one of the following statements is **false**?
   (1) If many such samples are taken and a 95% confidence interval for μ_{nstump} is calculated from each sample, then statements such as "μ_{nstump} is somewhere between the two confidence limits" are true, on average, 19 times out of 20.
   (2) The technique used to calculate the confidence interval generates an interval which contains the true population mean approximately 95% of the time, in the long run.
   (3) With 95% confidence, the sample mean is somewhere between 7.8 and 12.5.
   (4) In light of the data, the plausible values of μ_{nstump} are between 7.8 and 12.5.
   (5) With 95% confidence, the value of μ_{nstump} is estimated to be 10.1, with a margin of error of 2.35.

4. Based on the years 1960 to 1969, the mean of nstump is 7.4 and the standard deviation of nstump is 4.35, while the standard error, se($\bar{x}$) is 1.3756 and the t-multiplier is 2.262. Which one of the following statements about a new 95% confidence interval for μ_{nstump} based on these data compared to the original confidence interval given above is **true**?
   (1) The t-multiplier used in the new confidence interval is larger.
   (2) The standard error used in the new confidence interval is smaller.
   (3) The new confidence interval is centred around a higher value.
   (4) Both confidence intervals would have the same width because they are both 95% confidence intervals for μ_{nstump}.
   (5) The new confidence interval is narrower than the original confidence interval.

5. Which one of the following statements is **true**?
   (1) A point estimate is preferred to a confidence interval because the interval summarises the uncertainty due to sampling variation.
   (2) The standard error used to construct the interval will be identical for all samples of the same size.
   (3) If I plan to do a study in the future in which I will take a random sample and calculate a 90% confidence interval, there is a 90% chance that I will catch the true value of μ in my interval.
   (4) The size of the t-multiplier depends on only the sample size and not the desired confidence level.
   (5) The process of using a population parameter to construct an interval for the data estimate is an example of statistical inference.

6. Identify which one of the following statements about confidence intervals is **false**?
   (1) If a large number of researchers independently perform studies to estimate μ, about 95% of them will catch the true value of μ in their 95% confidence intervals.
   (2) If I plan to do a study in the future in which I will take a random sample and calculate a 90% confidence interval, there is a 90% chance that I will catch the true value of μ in my interval.
   (3) A two-standard error interval will always capture the true value of μ.
   (4) In the long run, if we repeatedly take samples and calculate a 95% confidence interval from each sample, we expect that 95% of the intervals will contain the true value of μ.
   (5) The estimate is always in the centre of the confidence interval.

7. When using a t-procedure to construct a confidence interval for a population mean, the confidence interval is constructed using the formula:

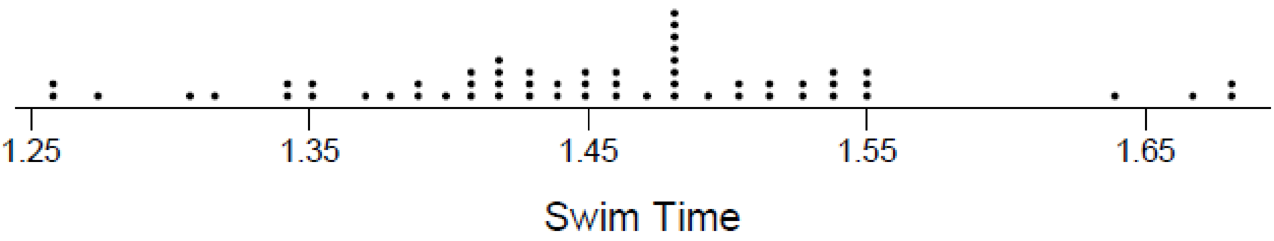
$$estimate \pm t \times se(estimate)$$

   Which one of the following statements is **false**?
   (1) The margin of error is the quantity added to and subtracted from the estimate to construct the interval.
   (2) The standard error used to construct the interval will be identical for all samples of the same size.
   (3) A confidence interval is preferred to a point estimate because the interval summarises the uncertainty due to sampling variation.
   (4) The size of the multiplier, t, depends on both the sample size and the desired confidence level.
   (5) Large samples tend to yield narrower 95% confidence intervals than small samples.

Questions 8 to **11** refer to the following information.

In 2001, a University of Auckland Sports Science student collected swim times from 58 New Zealand development squad swimmers. **Swim Time** is defined to be the number of minutes taken to swim 200 metres freestyle. Summary statistics and a dot plot of these data are given below:

	n	$\bar{x}_{Swim}$	s	se($\bar{x}_{Swim}$)
Swim Time	58	1.4543	0.0933	0.0123



8. A 95% confidence interval for the **Swim Time** data was calculated to be (1.4298, 1.4788). Which **one** of the following statements is **true**?
   (1) There is a 95% chance that a randomly selected development squad swimmer has a swim time in the interval from 1.43 to 1.48 minutes.
   (2) With 95% confidence, μ_{Swim} is somewhere between 1.43 and 1.48 minutes.
   (3) μ_{Swim} is estimated to be approximately 1.4543 minutes with a margin of error of 0.0123.
   (4) If many random samples of 58 development squad swimmers' swim times are taken and a 95% confidence interval calculated for each sample, then approximately 18 out of 20 of these confidence intervals will contain μ_{Swim}.
   (5) No valid statement can be made about the population mean swim time since a different sample would lead to a different mean and different confidence interval.

9. Suppose a random sample of 232 swim times (instead of 58) had been used to form a 95% confidence interval for μ_{Swim}. We would expect this new interval to have a width approximately:
   (1) double the width of the confidence interval formed from the 58 swim times.
   (2) the same width as the confidence interval formed from the 58 swim times.

(3) four times the width of the confidence interval formed from the 58 swim times.
(4) half the width of the confidence interval formed from the 58 swim times.
(5) a quarter of the width of the confidence interval formed from the 58 swim times.

10. A confidence interval for the population mean, μ_{Swim}, is found using the formula:

$$\bar{x}_{Swim} \pm t \times se(\bar{x}_{Swim})$$

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

(1) A confidence interval for μ_{Swim} summarises the uncertainty due to sampling variation.
(2) 95% of the time we carry out such a study, the confidence interval for the population mean, μ_{Swim}, will contain the true sample mean, $\bar{x}_{Swim}$.
(3) The degrees of freedom used when calculating a confidence interval for a single mean always depends on the sample size, *n*.
(4) A type of interval that contains the true value of a parameter for 95% of samples taken, in the long run, is called a 95% confidence interval for that parameter.
(5) The number of swim times in our sample affects both the size of the standard error and the size of the *t*-multiplier.

11. Suppose the Sports Science student realised that the four swim times greater than 1.6 minutes were all errors. After removing these values, the new standard deviation was 0.0754. Suppose a new confidence interval for the remaining 54 observations was calculated using the correct *t*-multiplier of 2.006.

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

(1) The new confidence interval would have a smaller mean and be wider than the original confidence interval.
(2) The new confidence interval would be centred around a smaller mean and be narrower than the original confidence interval.
(3) The original and new confidence intervals could not be compared since they would have two different means.
(4) The new confidence interval would be centred around a larger mean and be wider than the original confidence interval.
(5) The new confidence interval would be the same width as the original confidence interval because they are both 95% confidence intervals.

12. Cyclozocine was an alternative to methadone for treating heroin addiction. The following summary statistics come from a sample of 14 males who were chronic heroin addicts. After cyclozocine had removed the addicts' physical dependence on heroin, they were asked a list of questions designed to assess their psychological dependence. The test scores are called Q-scores and high values represent less psychological dependence.

Sample Size	Mean	Standard Deviation	Standard Error
14	39.93	11.96	3.196

A 95% confidence interval for the true mean Q-score of psychological dependence, where the *t*-multiplier is approximately 2.160 is:

(1) [36.73, 43.13]

(2) [33.03, 46.83]

(3) [38.09, 41.78]

(4) [39.24, 40.62]

(5) [47.32, 49.68]

13. Student Job Search is a nationwide organisation that finds part-time jobs for tertiary students. Data is collected on how many weeks work each student found over the summer break. This variable is called Summer Worker Weeks (SWW). Samples were taken from several different student groups. Below is a summary table of SWW for the students in these samples.

SWW for Different Institutions and Ethnic Groups

	Polytechnic	**University**	**Maori**	**Pacific Islander**	**Overseas Student**	**Other**
$\bar{x}$	1.38	2.21	1.66	2.34	3.00	2.01
s	0.48	0.53	0.49	2.26	1.70	0.48
n	11	11	11	10	10	11

Suppose we are interested in comparing the SWW for Pacific Island students to the SWW for overseas students. The standard error, $se(\bar{x}_{Overseas} - \bar{x}_{Pacific\ Islanders})$, is approximately 0.8943 while the *t*-multiplier is roughly 2.262. The correct formula to calculate the 95% confidence interval for $\mu_{\text{Overseas}} - \mu_{\text{Pacific Islanders}}$ is:

(1) (3.00 – 2.34) ± 2. 262 x 0.8943

(2) 2.262 x 1.96

(3) (2.34 – 3.00) ± 2. 262 x 0.8943

(4) 2.262 x 0.8943

(5) (3.00 – 2.34) ± 2. 262 x 1.96

Questions 14 to **16** refer to the following information.

The weight gain of women during pregnancy has an important effect on the birth weight of their children. In a study conducted in three countries, weight gains (in kg) of women during the last three months of pregnancy were measured. The results are summarised in the following table:

Country		n	$\bar{x}$	s
Egypt	(E)	11	4.55	1.83
Kenya	(K)	11	3.29	0.851
Mexico	(M)	11	2.9	1.8

We wish to determine the size of the difference in average weight gain between Egyptian and Kenyan women.

14. The value of the estimate, $\bar{x}_E - \bar{x}_K$, is:

    (1) 4.55

    (2) 1.26

    (3) 3.29

    (4) -1.26

    (5) 2.9

15. The value of the degrees of freedom, *df*, is:

    (1) 11

    (2) 2

    (3) 30

    (4) 10

    (5) 33

16. When calculating a 95% confidence interval for $\mu_E - \mu_K$, the appropriate value of the *t*-multiplier is 2.228. The standard error of the estimate, $se(\bar{x}_E - \bar{x}_K)$, is approximately 0.6085, therefore the **margin of error** for the 95% confidence interval is:

    (1) (4.55 - 3.29) ± 2.228 x 0.6085

    (2) (3.29 - 4.55) ± 2. 228 x 0.6085

    (3) (4.55 - 3.29) ± 2.228 x 1.96

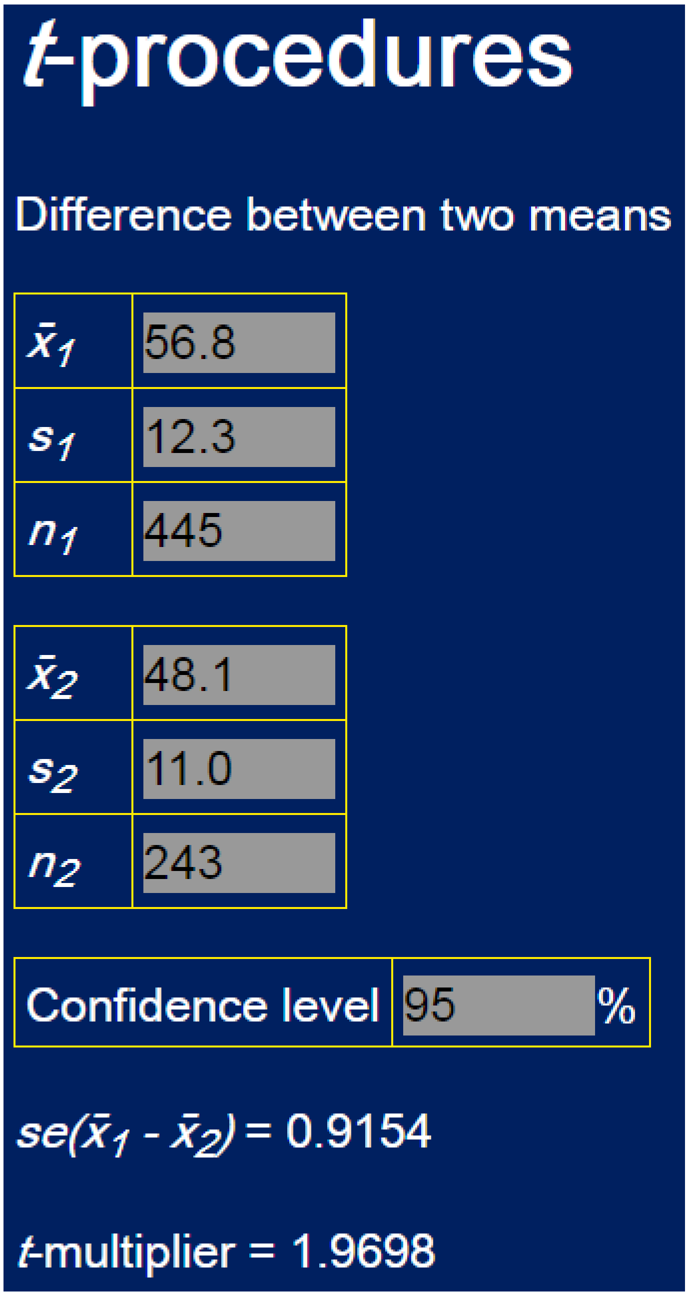
    (4) 2.228 x 1.96

    (5) 2.228 x 0.6085

Questions 17 to **19** refer to the following information.

Urea formaldehyde foam insulation (UFFI) has been used to insulate homes but it can sometimes give off formaldehyde, a gas that some people are allergic to. Sometimes people had an allergic reaction in homes with UFFI so a study was carried out to compare the concentration of formaldehyde in the air in homes with and without UFFI.

Measurements on formaldehyde concentration were made on a random sample of 445 homes with UFFI, and on a random sample of 243 homes without UFFI. Summary statistics from the study were put into the *t*-procedures tool as per the screenshot to the right.



We wish to estimate the difference between the average formaldehyde concentration in homes with UFFI and the average formaldehyde concentration in homes without UFFI.

17. The value of the estimate is:
    (1) 48.1
    (2) 104.9
    (3) 8.7
    (4) 56.8
    (5) 1.3

18. The value for *df* is:
    (1) 444
    (2) 445
    (3) 688
    (4) 242
    (5) 243

19. The 95% confidence interval for $\mu_1 - \mu_2$ is:
    (1) (56.8 – 48.1) ± 1.9698 x 0.9154
    (2) (56.8 – 48.1) ± 1.96 x 0.9154
    (3) (48.1 – 56.8) ± 1. 9698 x 0.9154
    (4) (56.8 – 48.1) ± 0.9154 x 1.970
    (5) (48.1 – 56.8) ± 0.9154 x 1.970

Questions 20 and **21** refer to the following information.

A researcher wanted to compare energy contents of full cream and low fat milk products on sale in Australia.

Let μ_L be the mean energy content (in kJ/100 mL) for **all** low fat milk products.
and
μ_F be the mean energy content (in kJ/100 mL) for **all** full cream milk products

A random sample of 16 full cream milk products and a random sample of 16 low fat milk products are selected. Summary statistics for these samples are shown in Table 4.

Milk product	Sample size	Sample mean (kJ/100 mL)	Sample std. dev. (kJ/100 mL)
Full cream	16	$\bar{x}_F = 279.4$	$s_F = 21.8$
Low fat	16	$\bar{x}_L = 207.9$	$s_L = 27.8$

Table 4: Energy content for milk product samples

Based on the data, a 95% confidence interval for $\mu_F - \mu_L$ is constructed.

20. The value for *df* is:
    (1) 16
    (2) 32
    (3) 31
    (4) 30
    (5) 15

21. Based on the data, a 95% confidence interval for $\mu_F - \mu_L$ is (53.4, 89.6). Which **one** of the following statements is **false**?
    (1) It is hard to believe that the mean energy content for all full cream milk products has the same value as the mean energy content for all low fat milk products.
    (2) With 95% confidence, the mean energy content for all full cream milk products is estimated to be 71.5 kJ/100mL greater than that for all low fat milk products, with a margin of error of 18.1 kJ/100 mL.
    (3) With 95% confidence, the mean energy content for all full cream milk products is estimated to be between 53.4 kJ/100mL and 89.6 kJ/100mL greater than that for all low fat milk products.
    (4) A claim that there is no difference between the mean energy content for all full cream milk products and the mean energy content for all low fat milk products is consistent with the data.
    (5) It cannot be stated with certainty that the interval (53.4, 89.6) contains the true value of $\mu_F - \mu_L$.

Questions 22 to **24** refer to the following information.

Table 2 is part of a table containing information about a random sample of 158 cities with subway systems. The cities are ordered by length of track.

ID No.	City	Age	Length (km)	Stations	km per station
1	London	0	421	275	1.53
2	New York	0	370	468	0.79
3	Tokyo	0	292	202	1.45
4	Seoul	1	287	298	0.96
5	Moscow	0	278	138	2.01
⋮	⋮	⋮	⋮	⋮	⋮
157	Rouen	1	2	5	0.44
158	Haifa	0	2	6	0.29

Table 2: Subway data Source: *Metro Bits*

Some of the variables referred to are:

Age The age of the subway:
0 – opened before 1980
1 – opened in 1980 or later

Length The length of track to the nearest kilometre

Stations The number of stations

km per station The average length of track per station in kilometres

We are interested in the relationship between the length of track and the age of the subway.

Let: μ_B be the underlying mean length of track for subways opened **before 1980**

and μ_L be the underlying mean length of track for subways opened **in 1980 or later**.

Summary statistics of the data are given below:

Group Statistics

	Age	N	Mean	Std. Deviation
Length (km)	Before 1980	72	80.439	87.484
	1980 or later	86	25.835	21.219

22. The value for *df* is:
    (1) 86
    (2) 71
    (3) 158
    (4) 85
    (5) 72

23. A 95% confidence intervals for $\mu_B - \mu_L$ is (33.546, 75.662).

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

    (1) With 95% confidence, we estimate the underlying mean length of track for subways opened before 1980 is 54.6km more than that of subways opened in 1980 or later, with a margin of error of about 21.1 km.
    (2) It is plausible that the mean length of track for subways opened before 1980 is greater than that of subways opened in 1980 or later.
    (3) We can claim that there is a difference between the mean length of track for subways opened before 1980 and that of subways opened in 1980 or later.
    (4) With 95% confidence, we estimate the underlying mean length of track for subways opened before 1980 is somewhere between 33.5km and 75.7 km greater than that of subways opened in 1980 or later.
    (5) With 95% confidence, we estimate the underlying mean length of track for subways opened before 1980 is 54.6km more than that of subways opened in 1980 or later, with a margin of error of about 42.1 km.

24. Suppose that the sample size for subways opened before 1980 was 32 rather than 72. Also suppose that all other summary statistics are the same as those given above. A new 95% confidence interval for the difference in the underlying means was constructed.

    Which one of the following statements **best** describes this new confidence interval relative to the original confidence interval?

    It would be centred around:

    (1) a smaller mean and would be wider.
    (2) the same mean and would be wider.
    (3) a smaller mean and would be narrower.
    (4) the same mean and would be narrower.
    (5) the same mean and would be the same width.

Question 25 refers to the following information.

A researcher wants to investigate the differences in the price of car models made by Australian and USA companies for the New Zealand market. Figure 8 below shows a dot plot of the data with n_{Aust} = 5 and n_{USA} = 7.

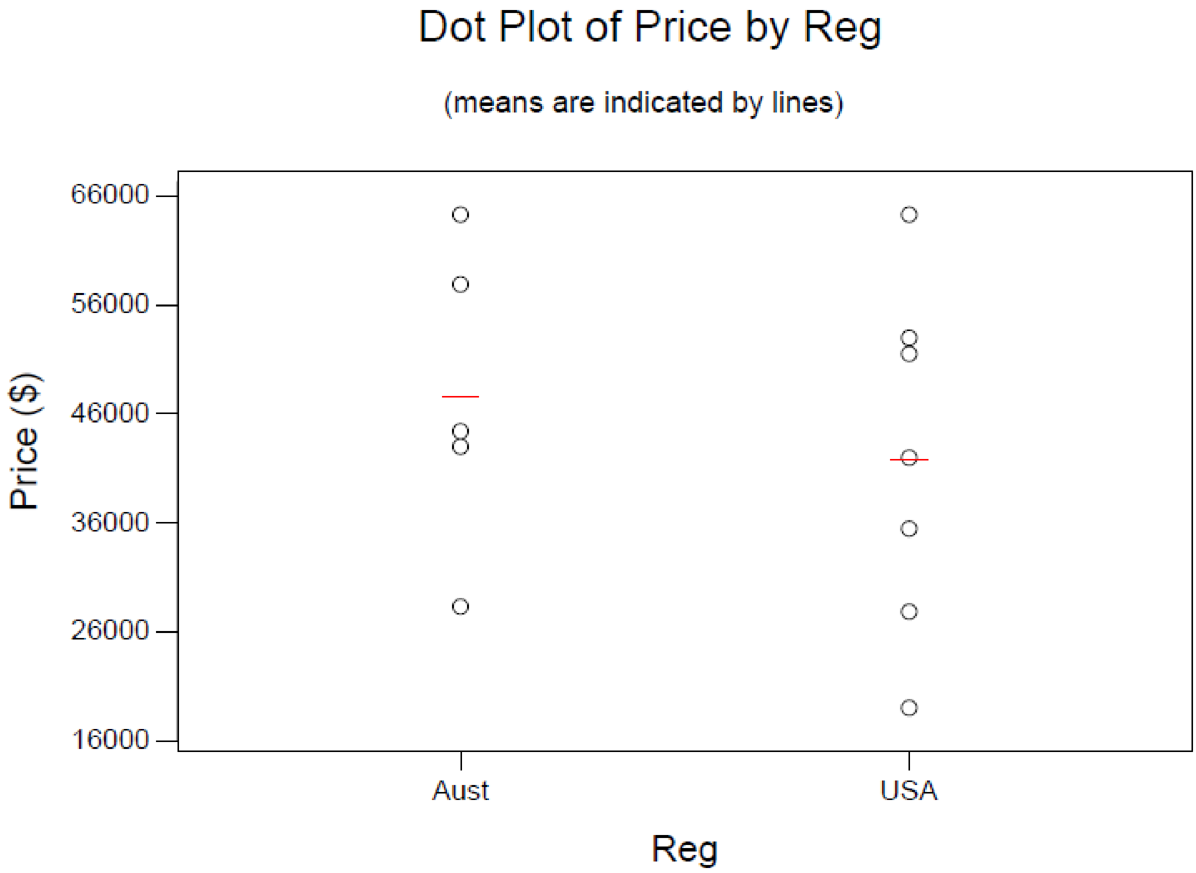


Figure 8: Dot plot of prices of new Australian and USA cars

Let μ_{Aust} and μ_{USA} be the underlying means for the price of Australian and USA models, respectively.

25. A 95% confidence interval for $\mu_{Aust} - \mu_{USA}$ is calculated to be (−\$13,844, \$25,258). Based on this confidence interval, which **one** of the following statements is **true**?

(1) With 95% confidence, μ_{Aust} is somewhere between \$13,844 higher than and \$25,258 lower than μ_{USA}.

(2) With 95% confidence, μ_{Aust} is somewhere between \$13,844 lower than and \$25,258 higher than μ_{USA}.

(3) With 95% confidence, μ_{Aust} is either \$13,844 lower than μ_{USA} or \$25,258 higher than μ_{USA}.

(4) With 95% confidence, μ_{Aust} is either \$13,844 higher than μ_{USA} or \$25,258 lower than μ_{USA}.

(5) No statement can be made about the relative sizes of μ_{Aust} and μ_{USA} because another sample of cars would give different estimates of μ_{Aust} and μ_{USA}.

26. Which **one** of the following statements about confidence intervals is **false**?
    (1) For a given level of confidence, increasing the sample size generally decreases the width of a confidence interval.
    (2) For a given level of confidence, decreasing the standard error decreases the width of a confidence interval.
    (3) For a given sample size, increasing the confidence level increases the width of a confidence interval.
    (4) For a given sample size, increasing the confidence level increases the precision of a confidence interval.
    (5) For a given level of confidence, decreasing the sample size generally decreases the precision of a confidence interval.

27. Which **one** of the following statements is **true**?
    (1) The Student's t-distribution has tails which become fatter as the degrees of freedom increase.
    (2) An estimate is more precise if it has more variability.
    (3) Student($df = \infty$) and Normal($\mu = 0$, $\sigma = 1$) are identical distributions.
    (4) If greater confidence in a confidence interval calculated from our data is desired, then a narrower interval needs to be used.
    (5) A parameter is a numerical characteristic which can be calculated from a sample.

28. Suppose that a 95% confidence interval for the difference in true mean HOSP.RATE level between the *small cars* and *medium cars*, $\mu_{\text{Small}} - \mu_{\text{Medium}}$, is given by (-0.05, 0.9). Which one of the following statements is **true**?
    (1) A 99% confidence interval for $\mu_{\text{Small}} - \mu_{\text{Medium}}$, is narrower than (-0.05, 0.9).
    (2) With 95% confidence the true mean HOSP.RATE for *small cars* is somewhere between 0.05 units and 0.9 units bigger than the mean HOSP.RATE for *medium cars*.
    (3) It is likely that mean HOSP.RATE for *small cars* is much smaller than the mean HOSP.RATE for *medium cars*.
    (4) With 95% confidence the true mean HOSP.RATE for *small cars* is somewhere between 0.05 units smaller and 0.9 units bigger than the mean HOSP.RATE for *medium cars*.
    (5) The difference between the sample means will be outside this interval 5% of the time.

Questions 29 and **30** refer to the following information.

Myers *et al.* (2010) conducted a study with 270 customers of a Boston (USA) coffee shop.

Some variables used in the study were:

Gender	The gender of the customer – Female – Male
Coffee type	The type of coffee ordered – Fancy – Not fancy
Waiting time	The time between ordering and receiving coffee, in seconds

Assume the study involves a random sample from the population of all customers of this Boston coffee shop.

Let:

μ_F be the mean waiting time for a **fancy** coffee for all customers of this Boston coffee shop

and

μ_N be the mean waiting time for a **not fancy** coffee for all customers of this Boston coffee shop.

The summary statistics for the waiting times of the 270 customers in the study are given in Table 1.

Name	Group size n	Mean $\overline{x}$	Std. deviation s
Fancy	175	124.0	24.2
Not fancy	95	51.9	11.8

Table 1: Summary statistics for waiting times

29. The standard error of the estimate, se($\bar{x}_F - \bar{x}_N$), is approximately 2.194 and the value of the t-multiplier for constructing a 95% confidence interval for $\mu_F - \mu_N$ is approximately 1.986.

    The 95% confidence interval for $\mu_F - \mu_N$ is:

    (1) (70.1, 74.1)
    (2) (119.6, 128.4)
    (3) (69.9, 74.3)
    (4) (67.7, 76.5)
    (5) (−76.5, −67.7)

30. Suppose that both a 99% confidence interval and a 95% confidence interval for $\mu_F - \mu_N$ are to be constructed using the summary statistics in Table 1.

    When comparing the 99% confidence interval with the 95% confidence interval, which **one** of the following statements is **false**?

    (1) The margin of error for the 99% confidence interval will be greater than that for the 95% confidence interval.
    (2) The 99% confidence interval will be wider than the 95% confidence interval.
    (3) The interval statement for the 99% confidence interval will be more precise than that for the 95% confidence interval.
    (4) The value of the standard error, se($\bar{x}_F - \bar{x}_N$), will be the same for both confidence intervals.
    (5) The value of the t-multiplier for the 99% confidence interval will be greater than that for the 95% confidence interval.

ANSWERS

1.	**(3)**	2.	**(2)**	3.	**(3)**	4.	**(1)**	5.	**(3)**	6.	**(3)**
7.	**(2)**	8.	**(2)**	9.	**(4)**	10.	**(2)**	11.	**(2)**	12.	**(2)**
13.	**(1)**	14.	**(2)**	15.	**(4)**	16.	**(5)**	17.	**(3)**	18.	**(4)**
19.	**(1)**	20.	**(5)**	21.	**(4)**	22.	**(2)**	23.	**(5)**	24.	**(2)**
25.	**(2)**	26.	**(4)**	27.	**(3)**	28.	**(4)**	29.	**(4)**	30.	**(3)**

WHAT SHOULD I DO NEXT?

- **Go through the Chapter 6 blue pages. The blue pages relevant to the material in this workshop are the notes on pages 15 to 17, the glossary on page 20, the true/false statements on page 21, the questions on page 22 and the tutorial material on pages 23-25.**
- **Try Chapter 6 questions from three of the past five tests that are relevant to this workshop.**

FORMULAE

Confidence intervals and t-tests

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

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

Applications:

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

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

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

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

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: $\quad \widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$

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