Practice Questions

1. Suppose the hypothesis test H_0: μ = 100 versus H_1: $\mu \neq$ 100 obtained a *P-value* = 0.001. Which of the following statements is **true**?

F (1) The *P-value* is very small, therefore H_0 is probably false.

T (2) We would reject H_0 at the 1% level of significance. p-val<.01

F (3) A 95% confidence interval for μ does not ~~contains~~ contain the value 100.

F (4) A 99% confidence interval for μ does not ~~contains~~ contain the hypothesised value. 100

F (5) We will accept that H_0 is true.

95% CI: hyp val in $\longleftrightarrow$ p-val > .05
hyp val out $\longleftrightarrow$ p-val < .05

2. In order to study the harmful effects of DDT poisoning, the pesticide was fed to 6 randomly chosen rats out of a group of 12 rats. The other 6 unpoisoned rats comprised of the control group. The following data gives measurements of the amount of tremor detected in the bodies of each rat after the experiment. The more tremor, the more harmful.

Poisoned group:	12.207,	16.869,	25.050,	22.429,	8.456,	20.589
Control group:	11.074,	12.064,	9.351,	6.642,	9.686,	8.182

We wish to test 2 indep samples, story 3

~~(1)~~	$H_0 : \mu_{diff} = 0$	versus	$H_1 : \mu_{diff} \neq 0$
(2)	$H_0 : \mu_1 - \mu_2 = 0$	versus	$H_1 : \mu_1 - \mu_2 \neq 0$
~~(3)~~	$H_0 : \cancel{\bar{x}_1 - \bar{x}_2} = 0$	versus	$H_1 : \cancel{\bar{x}_1 - \bar{x}_2} \neq 0$
~~(4)~~	$H_0 : \cancel{p_1 - p_2} = 0$	versus	$H_1 : \cancel{p_1 - p_2} \neq 0$
~~(5)~~	$H_0 : \cancel{\bar{x}_{Diff}} = 0$	versus	$H_1 : \cancel{\bar{x}_{Diff}} \neq 0$

$\mu_1 - \mu_2$ parameter

3. Which one of the following statements about the one-way analysis of variance *F*-test is **false**?

T (1) The evidence of differences between the true group means comes from comparing the variability between group means with the variability within the groups. p-val s_W^2 s_B^2 $f_0 = s_B^2 / s_W^2$

T (2) It should only be used when comparing independent samples.

T (3) It provides partial protection against multiple comparisons.

T (4) The null hypothesis is that all the true group means are the same.

(5) It ~~is not~~ may be badly affected by the presence of only one or two outliers.

$\rightarrow$ may affect equal std dev assumption

4. Does too much sleep impair intellectual performance? Taub *et al.* (1971) examined this commonly held belief by comparing the performance of 12 subjects on the mornings following (1) two normal nights' sleep and (2) two nights of "extended sleep". In the morning they were given a number of tests of ability to think quickly and clearly. One test was for vigilance where the lower the score, the more vigilant the subject. The following data was collected:

Subject	1	2	3	4	5	6	7	8	9	10	11	12
Normal Sleep	8	9	14	4	12	11	3	26	3	11	10	1
Extended Sleep	8	9	15	2	21	16	9	38	10	11	16	41

To see if the data supports the view that too much sleep can be bad for you, we would test which of the following hypotheses?

paired data comparison

~~(1)~~ $H_0 : \bar{x}_1 - \bar{x}_2 = 0$ versus $H_1 : \bar{x}_1 - \bar{x}_2 < 0$

~~(2)~~ $H_0 : \bar{x}_{Diff} = 0$ versus $H_1 : \bar{x}_{Diff} \neq 0$

(3) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \neq 0$ 1* special case

~~(4)~~ $H_0 : p_1 - p_2 = 0$ versus $H_1 : p_1 - p_2 \neq 0$

(5) $H_0 : \mu_{diff} = 0$ versus $H_1 : \mu_{diff} \neq 0$ parameter $= \mu_{diff}$

$H_0: \mu_S - \mu_M = 0$ vs $H_1: \mu_S - \mu_M \neq 0$

5. Suppose that a 95% confidence interval for the difference in true mean HOSP.RATE level between the *small cars* and *medium cars*, $\mu_{Small} - \mu_{Medium}$, is given by (-0.05, 0.9). Which one of the following statements is **true**?

zero in CI, p-value > 0.05 $\mu_S \approx \mu_M$

F (1) There is ~~a~~ not a significant difference between the true means at the 5% level.

F (2) There is not a significant difference between the ~~sample~~ pop/true means at the 5% level.

F (3) It is likely that mean HOSP.RATE for *small cars* is much smaller than the mean HOSP.RATE for *medium cars*.

T (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*.

F (5) The difference between the ~~sample~~ true/pop means will be outside ~~this~~ the 95% conf interval 5% of the time.

6. Analysis of variance (ANOVA) is (select **one** only):

~~(1)~~ an overall test of no difference between sample variances.

(2) an *F*-test of no difference between population means.

(3) an overall test of no difference between population variances.

(4) an *F*-test for the equality of population variances.

~~(5)~~ an *F*-test of no difference between sample means.

7. Which one of the following statements about significance tests is false?

T (1) Formal tests can help determine whether effects we see in our data may just be due to sampling error.

T (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.

T (3) The *P-value* says nothing about the size of an effect.

F (4) The data should be carefully examined in order to determine whether the alternative hypothesis needs to be one-sided or two-sided. theory or prev. studies

T (5) The *P-value* describes the strength of evidence against the null hypothesis.

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

F (1) A two-sided test of H_0: *parameter* = *hypothesised value* has *P-value* less than 0.05 if the *hypothesised value* lies ~~within~~ out of a 95% confidence interval for the *parameter*.

F (2) The larger the *P-value,* the ~~stronger~~ weaker the evidence against H_0.

F (3) The larger the test statistic, $|t_0|$, for a two-sided test, the ~~larger~~ smaller the *P-value* will be.

T (4) Tests of hypotheses can only deal with random errors and sampling variation. They are ineffective when confronted with data that has systematic bias. std err formulae calc. chance error.

F (5) An extremely small *P-value* means that the actual effect differs markedly from that claimed in the null hypothesis.

→ CI!

9. Which one of the following statements about hypothesis testing is **false**?

F (1) The larger the *P-value,* the ~~stronger~~ weaker the evidence against the null hypothesis.

T (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.

T (3) We cannot establish an hypothesised value for a parameter, we can only determine whether there is evidence to reject a hypothesised value.

T (4) H_0 is typically a sceptical reaction to a research hypothesis.

T (5) The *P-value* measures the strength of evidence against the null hypothesis.

10. Before conducting formal tests, one should look at plots of the data. Which **one** of the following statements is **false**?

T (1) Plots may highlight strange or interesting features of the data which cannot be seen in a formal test.

T (2) Summaries of the important features of the data can often be obtained from looking at plots.

T (3) Plots are used to check the validity of the assumptions for the formal tests.

F (4) Inferences, i.e. conclusions about the population, drawn from plots do ~~not~~ need to be verified by formal tests.

T (5) Additional points of interest are often suggested by plots.

11. Thirty observations of the relative return of over-the-counter stocks bought in the week of the 9th to the 13th of May, 1994 are given below.

-0.2940	-0.1092	-0.1053	-0.0707	-0.0563
-0.0541	-0.0423	-0.0398	-0.0396	-0.0390
-0.0381	-0.0323	-0.0221	-0.0169	-0.0139
-0.0081	0.0038	0.0057	0.0156	0.0172
0.0182	0.0192	0.0423	0.0459	0.0476
0.0667	0.0714	0.0780	0.1176	0.1224

Note: $\bar{x} = -0.01030$ and $s = 0.0786$

application!

Investors would like to ~~know~~ how the market performed. One measure of market performance is the mean relative return for the week.

A 95% confidence interval for the mean relative return is [-0.040, 0.019]. Which **one** of the following statements is **false**?

T (1) The *P*-value for testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$ is larger than 0.05.

F (2) There is no evidence, at the 5% level, to believe that the mean return is different from zero.

T (3) A 99% confidence interval for the mean return would be wider than the 95% confidence interval.

T (4) It is plausible that the mean relative return is zero.

T (5) An estimate of the mean relative return is −0.01030. $\bar{x}$

zero in CI ∴ μ could be zero.

" " " ∴ p-val > .05

CI!

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

F (1) A small *P-value* provides evidence of the size of an effect.

F (2) Statistical significance is the same as practical significance. No!

T (3) Practical significance depends on the size of the effect.

F (4) A small *P-value* provides ~~no~~ evidence against H_0.

F (5) A confidence interval estimates the ~~strength~~ size of an effect. (or difference)

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

T (1) t-tests may not be valid if there are outliers present and the sample is not large.

T (2) t-tests may not be valid when the data show clustering.

F (3) In general, t-tests are ~~not~~ robust against the Normality assumption.

T (4) t-tests will generally work well for any large sample.

T (5) t-tests may not be valid if the data are clearly skewed and the sample is not large.

Question 14 refers to the following information.

n

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.

14. 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) 7

double the accuracy

$4 \times 48 = 192$

halve the accuracy $48 \div 4 \rightarrow 12$

Questions 15 to **20 refer to the following situation:**

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_x
Egypt (E)	11	4.55	1.83
Kenya (K)	11	3.29	0.851
Mexico (M)	11	2.9	1.8

application 3.
(D)

mean

15. We wish to determine the size of the difference in average weight gain between Egyptian and Kenyan women. Given that the data shows no non-Normal features when plotted, the most appropriate procedure to use here is (select **one** only):

CI   2 indep samples

(1) a confidence interval based on the paired t-test.

(2) a two-independent sample proportion test.

(3) a confidence interval based on the two-independent sample t-test.

(4) a confidence interval based on the paired sign test.

(5) a confidence interval based on the Mann-Whitney test.

16. Suppose a two-independent sample t-test is appropriate here for testing H_0: $\mu_E - \mu_K = 0$, against H_1: $\mu_E - \mu_K \neq 0$. Then the value of the degrees of freedom, df, is (select **one** only):

(1) 21

(2) 10

(3) 11

(4) 20

(5) 22

$$df = \min(n_1 - 1, n_2 - 1)$$
$$= \min(11-1, 11-1)$$
$$= \min(10, 10) = 10$$

17. Suppose the P-*value* for the test in Question 16 is 0.1654 (it is not). Which **one** of the following statements is **correct**?

no ev. against H_0

(1) The data provides no evidence against H_0.

(2) The data provides strong evidence that H_0 is true.

(3) The data provides evidence that H_0 is true.

(4) The data provides evidence in favour of H_1.

(5) The data provides strong evidence that H_1 is true.

t

18. When calculating a 95% confidence interval for $\mu_E - \mu_K$, the value of the t-multiplier obtained from a *Student's* t-table is 2.2281. The standard error, $se(\bar{x}_E - \bar{x}_K) = 0.6085$, then the **margin of error** for the 95% confidence interval is:

moe

(1) $(4.55 - 3.29) \pm 2.2281 \times \dfrac{0.6085}{\sqrt{11}}$

moe: $\pm t \times se(est)$

(2) $\pm 2.2281 \times \dfrac{0.6085}{\sqrt{11}}$

$\rightarrow \pm t \times se(\bar{x}_E - \bar{x}_K)$

(3) $(3.29 - 4.55) \pm 2.2281 \times 0.6085$

(4) $(3.29 - 4.55) \pm 2.2281 \times \dfrac{0.6085}{\sqrt{11}}$

$\rightarrow \pm 2.2281 \times .6085$

(5) $\pm 2.2281 \times 0.6085$

19. We wish to determine if there are differences in average weight gain in any of the three countries. The most appropriate procedure to use here is (select **one** only):

(1) a confidence interval for the highest-lowest average weight gain: $\mu_{high} - \mu_{low}$.
(2) an F-test for H_0: $\mu_E = \mu_K = \mu_M$.
(3) a paired t-test for H_0: $\mu_{diff} = 0$, where $\mu_{diff} = \mu_{high} - \mu_{low}$.
(4) three paired t-tests.
(5) a Tukey interval for the highest-lowest average weight gain: $\mu_{high} - \mu_{low}$.

20. We wish to perform an F-test for the weight gain data. The appropriate degrees of freedom are (select **one** only):

(1) $df_1 = 2$ and $df_2 = 30$
(2) $df_1 = 3$ and $df_2 = 33$
(3) $df_1 = 33$ and $df_2 = 3$
(4) $df_1 = 30$ and $df_2 = 2$
(5) $df_1 = 2$ and $df_2 = 33$

$df_1 = k - 1 = 3 - 1 = 2$

$df_2 = n_{tot} - k = 33 - 3 = 30$

Questions 21 to **24** refer to the following information.

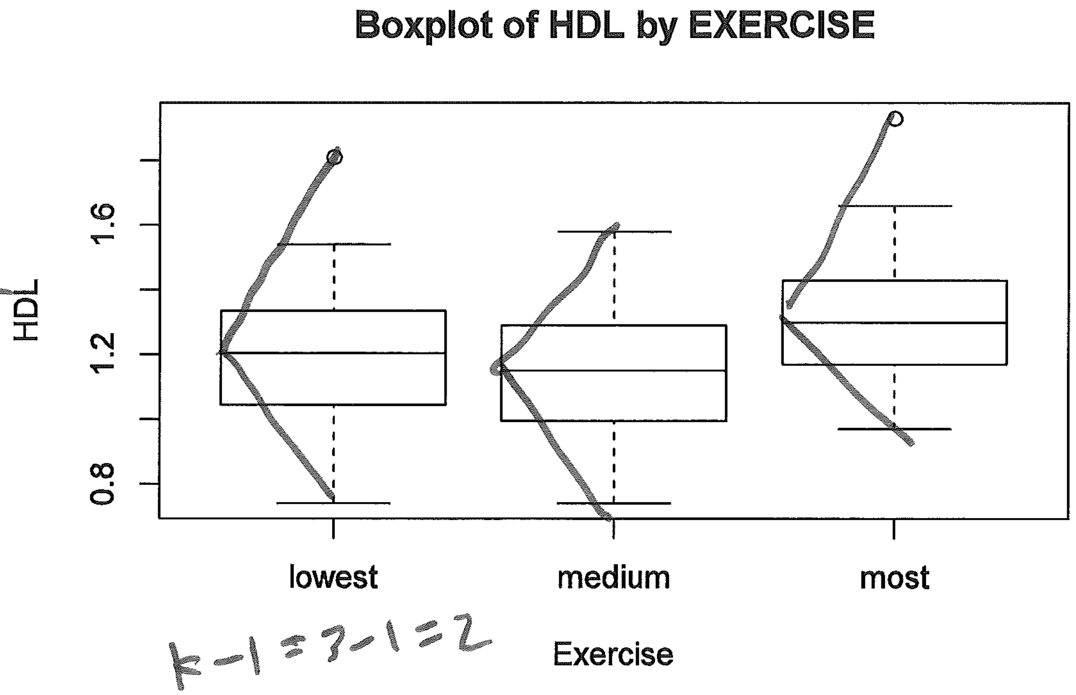
HDL cholesterol is known as the "good cholesterol" as it is associated with lower risks of problems like heart disease. The following data were collected on men working in New Zealand companies.

The men were divided into exercise groups by the amount of exercise they reported. categorical, ordinal

Levels of exercise were classified as lowest, medium, and most. Sixty men were randomly sampled from each exercise group, and their HDL cholesterol level was measured. $n_{tot} = 3 \times 60 = 180$

numeric

Box plots and some SPSS output analysing the data follow.



$k - 1 = 3 - 1 = 2$

Figure 3: Box plot of HDL by exercise level.

v. st. ev. against H_0

Analysis of Variance for HDL

Source	DF	SS	MS	F	P
EXERCISE	***	0.7013	0.3507	***	0.000
Error	***	6.7632	0.0382		
Total	179	7.4645			

$n_{tot} - k = 180 - 3 = 177$

$\frac{.3507}{.0382} = 9.181$

	n	sample mean	sample std dev
lowest	60	1.1998	0.1926
medium	60	1.1497	0.2301
most	60	1.2998	0.1906

Table 3: SPSS output for the cholesterol data.

$\frac{.2301}{.1906} = 1.2 < 2$

21. Which one of the following statements about the F–test shown in Table 3 is **false**?

F (1) The outside values in the lowest and most groups show that the Normality assumption of the F–test is violated. ntot =180!

T (2) The alternative hypothesis states that at least one of the exercise groups has a different underlying mean HDL level from another.

T (3) The differences in the sample standard deviations of the lowest, medium and most exercise groups do not affect the validity of the F–test in practice. 1.2 < 2

T (4) The box plots in Figure 3 give us no information on the independence of the exercise groups.

T (5) The null hypothesis states that the underlying mean HDL level is the same for each exercise group.

22. The values for the degrees of freedom, df_1 and df_2, at the top of Table 3 are:

(1) df_1=2, df_2=177

~~(2)~~ df_1=2, ~~df_2=178~~

~~(3)~~ df_1=2, ~~df_2=59~~

~~(4) df_1=3,~~ df_2=177

~~(5) df_1=3, df_2=176~~

23. The value of the F–test statistic, f_0, at the top of Table 3 is nearest to:

(1) 9.643

(2) 10.643

(3) 9.181

(4) 0.104

(5) 0.109

24. Which one of the following statements is the **best** interpretation of the *P-value* in the analysis of variance for HDL shown in Table 3?

(1) There is no evidence that all of the exercise groups have different underlying mean HDL cholesterol levels.

(2) There is extremely strong evidence that at least one of the exercise groups has a different underlying mean HDL cholesterol level from another. good

(3) There is no evidence that any of the exercise groups have different underlying mean HDL cholesterol levels.

(4) There is extremely strong evidence that all of the exercise groups have different underlying mean HDL cholesterol levels. too far!

(5) There is some evidence that at least one of the exercise groups has a different underlying mean HDL cholesterol level from another.

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

paired data

⇓

analyse differences!

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) Paired *t*-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. → use non-parametric test instead

(4) Two sample *t*-test on the two samples if plots of the two samples are severely non-Normal.

(5) *F*-test on the differences.

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

(1) T A statistically significant result is not always practically significant.

(2) T A non-significant hypothesis test does not mean that the null hypothesis is true.

(3) T Large positive *t*-test statistics lead to small *P-values* for two-tailed tests.

(4) T 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) F 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.

theory or prior info only!

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

(1) T In hypothesis testing, statistical significance does not imply practical significance.

CI!

(2) F In a hypothesis test for no difference between two means, a very small *P-value* indicates a very large difference in the means.

(3) T In hypothesis testing, a non-significant test result does not imply that H_0 is true.

(4) T In hypothesis testing, large samples can lead to small *P-values* without the results having any practical significance.

(5) T In a hypothesis test for no difference between two means, a two-sided test should be used when the idea of doing the test has been triggered as a result of looking at the data.

same!

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

(1) T Tests of hypotheses can only deal with random errors and sampling variation. They are ineffective when confronted with data that has systematic bias.

(2) T The larger the *P-value,* the weaker the evidence against H_0.

(3) T A two-sided test of H_0: *parameter = hypothesised value* has *P-value* greater than 0.05 if the *hypothesised value* lies within a 95% confidence interval for the *parameter*.

(4) F An extremely small *P-value* means that the actual effect differs markedly from that claimed in the null hypothesis. CI!

(5) T The larger the test statistic, $|t_0|$, for a two-sided test, the smaller the *P-value* will be.

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

T (1) For paired data, we analyse the differences.

T (2) Pairing is beneficial when the variability within pairs is small compared with the variability between pairs.

T (3) The one sample *t*-test can be used to analyse the differences in paired data.

F (4) Pairing ~~cannot~~ be used in observational studies.

T (5) The carryover effect occurs when the first treatment alters the effect of the second treatment.

Questions 30 and **31** refer to the following information. application 1. (A)

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. s

30. The customer wishes to test H_0: μ = 150 versus H_1: $\mu \neq 150$. The standard error, se($\bar{x}$), is 0.3031. 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 = 4.95, *df* = 47

(3) t_0 = -4.95, *df* = 48

(4) t_0 = -4.95, *df* = 49

(5) t_0 = 4.95, *df* = 48

$df = n-1 = 48-1 = 47$

$t_0 = \frac{\text{est} - \text{hyp value}}{se(\text{est})}$

$= \frac{\bar{x} - \mu_0}{se(\bar{x})} = \frac{148.5-150}{.3031}$

$= -4.95$ (2dp)

31. 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: weak ev. against H_0

F (1) With a *p-value* of 0.09 there is ~~no~~ evidence against the null hypothesis.

T (2) With a *p-value* of 0.09 there is weak evidence against the null hypothesis.

T (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.

F (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.

F (5) With a *p-value* of .09% there is weak evidence against the null hypothesis that the mean weight of Yummo chips is 150g.

.09 = 9%

32. Which **one** of the following statements about an F-test is **false**?

T (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).

T (2) The larger the value of the F-test statistic, f_0, the smaller the *P-value*.

T (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.

F (4) An increase in the spread of the data within each group will result in a decrease 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).

T (5) The value of the F-test statistic, f_0, is the ratio of the between-mean variation and the within-group variation.

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

T (1) In a t-test for no difference between two means, being able to demonstrate that the difference was of practical significance (importance) would almost always imply statistical significance.

T (2) In hypothesis testing, large samples can lead to small *P-values* without the results having any practical significance (importance).

T (3) In hypothesis testing, statistical significance does not imply practical significance (importance).

F (4) In a hypothesis test for no difference between two means, a very small *P-value* always indicates a very large difference in the means. CI

T (5) In hypothesis testing, a non-significant test result does not imply that the null hypothesis is true.

34. Which one of the following statements about the validity of confidence intervals of the form sample mean $\pm$ t standard errors is **false**?

T (1) It is critical that the sample is random. → not if n is moderate to large

F (2) It is critical that the distribution being sampled is Normal.

T (3) It is critical that the observations come from the same distribution.

T (4) Outliers and clusters of data can invalidate confidence intervals.

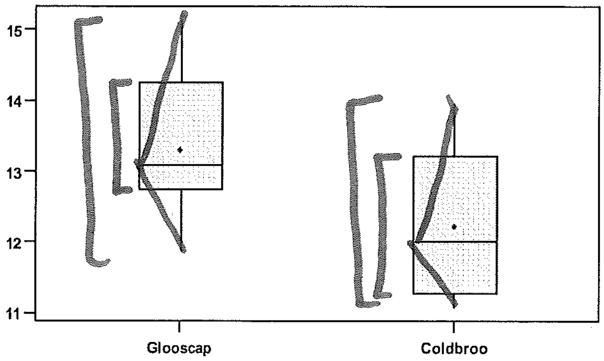
T (5) It is critical that observations are independent.

Questions 35 to **38** 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).

Boxplots of Glooscap and Coldbroo

(means are indicated by solid circles)



Group Statistics

	schcode	N	Mean	Std. Deviation	Std. Error Mean
runtime	Glooscap	12 n_1	13.3125 $\bar{x}_1$	.99133 s_1	.28617
	Coldbrook	13 n_2	12.2285 $\bar{x}_2$	.99018 s_2	.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	df 22.836	p-val .012	$\bar{x}_1-\bar{x}_2$ 1.08404	.39663	(.26322	, 1.90485)

$t_0 = 1.08404/.39663$ $se(\bar{x}_1-\bar{x}_2)$ 95% CI for $\mu_1-\mu_2$

35. 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)~~ $H_0: \cancel{p_1 - p_2} = 0$ versus $H_1: \cancel{p_1 - p_2} \neq 0$

(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)~~ $H_0: \cancel{p_D - p_S} > 0$ versus $H_1: \cancel{p_D - p_S} = 0$

36. Which one of the following is false?

T (1) There are no gross outliers in the data.

T (2) The running times for Glooscap are on average, greater than the running times for Coldbrook.

T (3) Two-sample t-tests are, in general, more robust to non-Normal features than the one sample t-test. $n_1 + n_2 = 12 + 13 = 25$

F (4) The data taken from Glooscap and Coldbrook show ~~severely~~ slightly non-normal features.

T (5) The range of running times for Glooscap is slightly more than the range of running times for Coldbrook.

37. The best interpretation for this test is: p-val = .012 = 1.2%

(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

38. 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? 95% CI: [.26, 1.90]

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

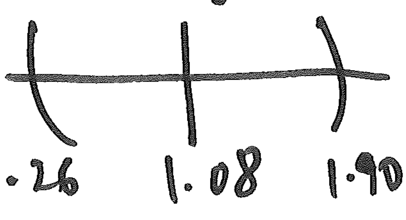
won't get same CI!

F (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.

F (3) With 95% confidence, we estimate that the true proportion p_1 will be somewhere between 0.26 ~~larger~~ and 1.90 ~~smaller~~ larger than p_2. not p's!

T (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.

F (5) With 95% confidence the true mean from Glooscap μ_1 is ~~1.65~~ 1.08 larger than the true mean from Coldbrook μ_2. with a margin of error of .82.



width = 1.90 − .26
= 1.64
moe = 1.64/2 = .82

Questions 39 and **40** 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.

$n=16$

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

diffs

analyse as a single sample!

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)	$\bar{x}_{diff}$ = 157.562	s_{diff} = 140.132	n_{diff} = 16

Summary Statistics

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

F-test (1) H_0: all of the μ's are equal
H_1: none of the μ's are equal ✗

data is paired

t-test for 2 indep. samples (2) H_0: $\mu_1 - \mu_2 = 0$
H_1: $\mu_1 - \mu_2 \neq 0$ ✗

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

∴ need a dot plot of differences

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

t-test for single proportion (5) H_0: $p = 0$
H_1: $p \neq 0$ ✗

weird combination of 2 indep samples & paired!

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

(1) $\dfrac{\overline{x}_{\text{diff}}}{se\left(\overline{x}_{\text{diff}}\right)}$

(2) $\dfrac{\overline{x}_{\text{diff}}}{se\left(\overline{x}_{\text{Post}} - \overline{x}_{\text{Pre}}\right)}$

(3) $\dfrac{\overline{x}_{\text{Post}} - \overline{x}_{\text{Pre}}}{se\left(\overline{x}_{\text{Post}} - \overline{x}_{\text{Pre}}\right)}$

(4) $\dfrac{\overline{x}_{\text{Post}} - \overline{x}_{\text{Pre}}}{se\left(\overline{x}_{\text{Post}} + \overline{x}_{\text{Pre}}\right)}$

(5) $\dfrac{\overline{x}_{\text{Post}}}{se\left(\overline{x}_{\text{Post}}\right)} - \dfrac{\overline{x}_{\text{Pre}}}{se\left(\overline{x}_{\text{Pre}}\right)}$

$t_0 = \dfrac{\text{est} - \text{hyp val}}{\text{std err}} = \dfrac{\overline{x}_{\text{diff}} - 0}{se(\overline{x}_{\text{diff}})}$

41. 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. H_0!
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

42. 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 too far!

~~(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~~ all of the μ's are equal
H_1: some of the μ's are ~~equal~~ diff

(5) H_0: ~~some~~ all of the μ's are equal
H_1: not all of the μ's are equal

n

Questions 43 to **45** 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. paired data! analyse diffs...

43. 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.

$$t_0 = \frac{\text{est} - \text{hyp val}}{\text{std err}} = \frac{\bar{x}_{diff} - 0}{se(\bar{x}_{diff})}$$

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

(1) $\dfrac{\bar{x}_{\text{after}}}{se\left(\bar{x}_{\text{after}}\right)} - \dfrac{\bar{x}_{\text{before}}}{se\left(\bar{x}_{\text{before}}\right)}$

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

(3) $\dfrac{\bar{x}_{\text{after}} - \bar{x}_{\text{before}}}{se\left(\bar{x}_{\text{after}} - \bar{x}_{\text{before}}\right)}$

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

(5) $\dfrac{\bar{x}_{\text{diff}}}{se\left(\bar{x}_{\text{diff}}\right)}$

$$t_0 = \frac{-69.25}{5.16}$$

Paired Samples Test

diff s_{diff}

		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

$\bar{x}_{diff}$ $se(\bar{x}_{diff})$ p-val $df = n-1$

45. 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.

same!

Questions 46 and **47** 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.

k

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			

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

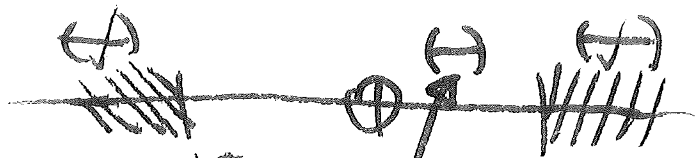
~~(1) $df_1 = 5$, $df_2 = 125$~~
(2) $df_1 = 4$, $df_2 = 120$
~~(3) $df_1 = 120$, $df_2 = 4$~~
(4) $df_1 = 4$, ~~$df_2 = 124$~~
~~(5) $df_1 = 125$, $df_2 = 5$~~

$df_1 = k - 1 = 5 - 1 = 4$
$df_2 = n_{tot} - k = 125 - 5 = 120$

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

(1) 79.20
(2) 654,500
(3) 13.61
(4) 0.073S5
(5) 0.4537

$$f_0 = \frac{2984.82}{219.28} = 13.61 \text{ (2dp)}$$



48. Which one of the following statements is **false**? −10 10

T (1) In a *t*-test for no difference between two means, being able to demonstrate that the difference was of practical significance (importance) would almost always imply statistical significance.

T (2) In hypothesis testing, large samples can lead to small *P-values* without the results having any practical significance (importance).

T (3) In hypothesis testing, statistical significance does not imply practical significance (importance).

F (4) In a hypothesis test for no difference between two means, a very small *P-value* always indicates a very large difference in the means. prac. sig.

T (5) In hypothesis testing, a nonsignificant test result does not imply that the null hypothesis is true.

49. Which one of the following is **false**?

T (1) A *P-value* calculated for a hypothesis formulated after looking at the data provides less convincing evidence than if the study had been designed to investigate the hypothesis.

T (2) Formulae for the standard errors of data estimates do not take into account systematic biases in the experiment or survey.

F (3) The fact that multiple comparisons have been made from a single set of data can be ignored when reporting the results.

T (4) If 100 people independently collect data and calculate a 95% confidence interval for a population mean we expect approximately 95 people to capture the true mean in their interval and 5 to miss it.

T (5) If 100 people independently collect data and test a true hypothesis, then just by chance, we expect about 5 to obtain results, which were significant at the 5% level.

Questions **50** to **58** refer to the **Swim Performance Study** information given below.

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. Figure 5 below shows a dotplot of these swim times. A confidence interval for the population mean swim time (μ_{Swim}) for the New Zealand development squad is given in Table 7 below.

error

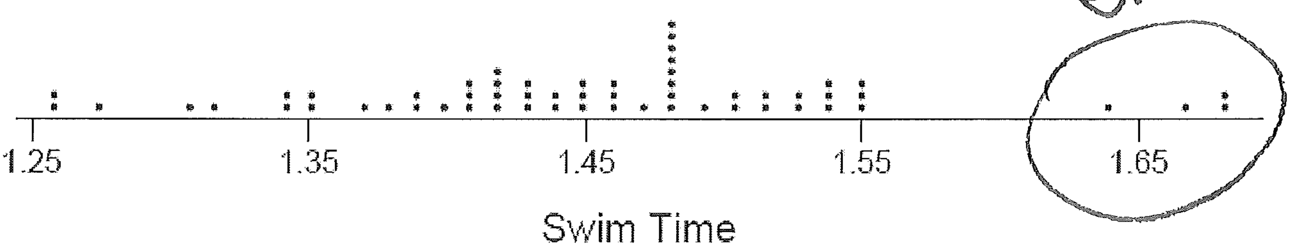


Figure 5: Dotplot of Swim Time (in minutes) for the New Zealand development squad

Summary Statistics and Confidence Interval

	N	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval Lower	95% Confidence Interval Upper
Swim Time	58	1.4543	0.0933	0.0123	1.4298	1.4788

Table 7: Summary statistics and confidence interval for the population mean Swim Time, μ_{Swim} (in minutes) for the New Zealand development squad

½

The Sports Science student monitored the swim performance for a subsample of 15 New Zealand development squad swimmers. Swim performance was measured by calculating their swim speed in the 200m freestyle as a percentage of the world record swim speed. For example, a swim performance of 100% would mean that the swimmer was as fast as the world record.

Swim performance for each of the 15 swimmers was recorded at the beginning of the study (referred to as **Before**), and at the end of the study (referred to as **After**). The **Differences** in performance for each swimmer were calculated as **After – Before**.

The Sports Science student wished to formally test for no difference between the mean **Before** and the mean **After** swim performance. Results for a two-sample *t*-test testing for no difference between swim performances **Before** and **After** the study are shown below in Table 8, while results for a paired sample *t*-test on the **Differences** are shown in Table 9.

paired data comparison!

T-Test

Group Statistics

	Swim performance	N	Mean	Std. Deviation	Std. Error Mean
Swim performance	After	15	83.94	3.53	0.91
	Before	15	80.76	4.23	1.10

inappropriate!

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Swim performance	Equal variances assumed	.079	.780	2.23	28	0.034	3.18	1.423	0.25	6.1
	Equal variances not assumed			2.23	27.56	0.034	3.18	1.423	0.25	6.1

Table 8: SPSS output: confidence interval and two-sample *t*-test comparing swim performance **After** with swim performance **Before**

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	After	83.94	15	3.53	0.91
	Before	80.76	15	4.23	1.09

$\bar{x}$ diff

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper			
Pair 1 After - Before	3.175	3.507	0.905	1.233	5.117	3.51	14	0.003

Table 9: SPSS output: confidence interval and paired *t*-test for swim performance **Differences**

t_0

$< .05$

Δ appropriate!

The effect of a resting treatment on swim performance for the same subsample of 15 swimmers was also investigated. The resting treatment involved suspending the swimmers in a heated bath in the dark for a number of hours. After recording each swimmer's performance at the beginning of the study (referred to as **Before**) each swimmer was randomly allocated into either the **Control** group (who received no treatment), or the **Rest** group (who received the resting treatment). Each swimmer's performance was also recorded at the end of the study (referred to as **After**). The **Differences** in each swimmer's performance were calculated as **After − Before**.

Two-sample *t*-test results comparing swim performance **Differences** for the **Control** and **Rest** groups are shown in Table 10 below.

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

2 indep samples

T-Test

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
Swim performance	Control	9	1.76	2.19	0.73
	Rest	6	5.29	4.22	1.72

Independent Samples Test

$se(\bar{x}_C - \bar{x}_R)$

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Swim performance	~~Equal variances assumed~~	~~.079~~	~~.780~~	~~-1.88~~	~~13~~	~~0.11~~	~~-3.53~~	~~1.871~~	~~-8.11~~	~~1.1~~
	Equal variances not assumed			-1.88	12.536	0.11	-3.53	1.871	-8.11	1.1

Table 10: Two sample *t*-test comparing swim performance **Differences** between treatment groups

no ev. against Ho

not sig @ 5% level

zero in CI

not prac sig.

Questions 50 to 58 refer to the **Swim Performance Study** information given above, on page 48. 95% CI: (1.4298, 1.4788)

50. Which **one** of the following statements is **true**? (Use Table 7, page 48.)

F (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.

T (2) With 95% confidence, μ_{Swim} is somewhere between 1.43 and 1.48 minutes.

F (3) μ_{Swim} is estimated to be approximately 1.4543 minutes with a margin of error of 0.0123.

F (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~~ 19 out of 20 of these confidence intervals will contain μ_{Swim}.

F (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.

moe: width of CI = 1.4788 − 1.4298 = .049

moe = .049/2 = .0245

51. 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.

$58 \xrightarrow{\times 4} 232$

$4 \times n \rightarrow$ double accuracy & halve the width...

52. 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 **true**?

T (1) A confidence interval for μ_{Swim} summarises the uncertainty due to sampling variation.

F (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, $\bar{x}_{Swim}$.

F (3) A sample of 58 swim times is large enough to allow the sample to consist of unrelated observations.

F (4) It is critical that the swim times come from a Normal distribution.

F (5) The number of swim times in our sample affects the size of the standard error ~~but does not affects~~ and the size of the t-multiplier.

Not with n=58!

53. Suppose the Sports Science student realised that the four swim times greater than 1.6 minutes were all errors. (See Figure 5, page 48.) 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.

	n	$\bar{x}$	s	$se(\bar{x})$	t	moe
Original	58	1.4543	.0933	.0123	< 2.006	bigger
new	54	< 1.4543	.0754	.0103	2.006	smaller

smaller

$$\frac{.0754}{\sqrt{54}}$$

Questions 54 and **55** refer to the **Swim Performance Study** information given above, on pages 49 and 50.

54. Assuming the student interpreted the correct *t*-test, which **one** of the following statements is **false**? (Use Tables 8 and 9 on pages 49 and 50 to answer this question.)

T (1) The test is comparing swim performance at the beginning of the study with swim performance at the end of the study.

T (2) The test is significant at the 5% level of significance. p-val = .003

(3) The *t*-test statistic is ~~2.23.~~ 3.51

T (4) The test is two-tailed.

T (5) The difference in the means is about 3.2. $\bar{x}_{diff}$ = 3.175

55. Suppose Table 8, page 49 shows the correct analysis for the Before/After swim performance comparisons. Note: this may **not** be true. it's not!
How would **one** best explain the results of this SPSS output to someone **unfamiliar** with statistics?

(1) There is a statistically significant difference between the sample average swim performance before and after the study.

(2) A 95% confidence interval states that the population mean swim performance of the swimmers in our sample dropped somewhere between 0.25 and 6.1 percentage points during the study.

(3) We can be 95% confident that the population average swim performance improved somewhere between 0.25 and 6.1 percentage points during the study.

(4) There is very strong evidence of a difference in population average swim performance at the beginning and end of the study.

(5) It is a reasonable bet that the population average swim performance at the end of the study was between 0.25 and 6.1 percentage points higher than at the beginning of the study.

- p-val = .034 → sig @ 5% level
- CI: (.25, 6.1)
- unfamiliar with stats → don't use stats language! (1), (2), (3), (4) all use stats language. Also some have other issues!

Questions 56 to **58** refer to the **Swim Performance Study** information given above, on page 51.

56. In a two-sample t-test on the **Differences** for the **Control** and **Rest** treatment groups (Table 10, page 51), which **one** of the following statements is **true**?

F (1) The *P-value* would be ~~smaller~~ larger if the standard errors of the **Control** and **Rest** groups were larger.

T (2) There is no evidence that the underlying means of the **Control** and **Rest** groups are different.

F (3) The *P-value* is not significant at the 5% level, ~~but~~ and the results are not practically significant.

F (4) The test is not significant at the 5% level of significance.

F (5) The average of the differences was ~~higher~~ lower for the **Control** group.

57. Using Table 10, page 51, the standard error of the difference between the two independent sample means, se($\bar{x}_{\text{Control}} - \bar{x}_{\text{Rest}}$), is approximately:

(1) 2.43	(4) 1.56
(2) 2.45	(5) 1.87
(3) 0.99	

58. Which **one** of the following statements gives the null and alternative hypotheses for the t-test shown in Table 10, page 51?

(1)	H_0: $\mu_{\text{Control}} - \mu_{\text{Rest}} = 0$	H_1: $\mu_{\text{Control}} - \mu_{\text{Rest}} \cancel{>} 0$
~~(2)~~	H_0: $\mu_{\text{Control}} - \mu_{\text{Rest}} \cancel{\neq} 0$	H_1: $\mu_{\text{Control}} - \mu_{\text{Rest}} = 0$
(3)	H_0 : $\mu_{\text{Control}} - \mu_{\text{Rest}} = 0$	H_1: $\mu_{\text{Control}} - \mu_{\text{Rest}} \neq 0$
~~(4)~~	H_0: $\cancel{\bar{x}}_{\text{Control}} - \cancel{\bar{x}}_{\text{Rest}} \cancel{\neq} 0$	H_1: $\cancel{\bar{x}}_{\text{Control}} - \cancel{\bar{x}}_{\text{Rest}} = 0$
~~(5)~~	H_0: $\cancel{\bar{x}}_{\text{Control}} - \cancel{\bar{x}}_{\text{Rest}} = 0$	H_1: $\cancel{\bar{x}}_{\text{Control}} - \cancel{\bar{x}}_{\text{Rest}} \neq 0$

ANSWERS

1. **(2)**	2. **(2)**	3. **(5)**	4. **(5)**	5. **(4)**	6. **(2)**
7. **(4)**	8. **(4)**	9. **(1)**	10. **(4)**	11. **(2)**	12. **(3)**
13. **(3)**	14. **(2)**	15. **(3)**	16. **(2)**	17. **(1)**	18. **(5)**
19. **(2)**	20. **(1)**	21. **(1)**	22. **(1)**	23. **(3)**	24. **(2)**
25. **(1)**	26. **(5)**	27. **(2)**	28. **(4)**	29. **(4)**	30. **(1)**
31. **(3)**	32. **(4)**	33. **(4)**	34. **(2)**	35. **(3)**	36. **(4)**
37. **(4)**	38. **(4)**	39. **(4)**	40. **(1)**	41. **(2)**	42. **(3)**
43. **(3)**	44. **(5)**	45. **(3)**	46. **(2)**	47. **(3)**	48. **(4)**
49. **(3)**	50. **(2)**	51. **(4)**	52. **(1)**	53. **(2)**	54. **(3)**
55. **(5)**	56. **(2)**	57. **(5)**	58. **(3)**		

FORMULAE

Confidence intervals and t-tests

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

Ch 6/7 (8, 9, 10)

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

Ch 7, 8, 10

Applications:

1. Single mean μ: $\quad estimate = \overline{x}; \qquad df = n - 1$

~~2. Single proportion p: $estimate = \widehat{p}; \quad df = \infty$~~

3. Difference between two means $\mu_1 - \mu_2$: (independent samples)

   $estimate = \overline{x}_1 - \overline{x}_2; \qquad df = \min(n_1 - 1, n_2 - 1)$

~~4. Difference between two proportions $p_1 - p_2$:~~

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

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

Ch 8

The Chi-square test

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

Ch 9

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: $\quad df = n - 2$

Ch 10