31. 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~~ cannot 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.

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 ~~an increase~~ de 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. $f_0 = \frac{s_B^2}{s_W^2}$

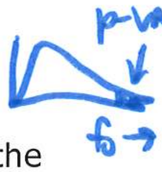
33. Which **one** of the following statements about one-way analysis of variance *F*-tests is **false**?

T (1) The greater the variability between the sample or group means relative to the variability within the samples or groups then the smaller the *P-value*.

F (2) A very small *P-value* suggests that there is ~~a large~~ difference between at least two of the underlying means. size? → CIs!

T (3) If the *P-value* is very small, then observed differences between the sample or group means could be explained as being a result of differences between the underlying means.

T (4) The larger the value of the *F*-test statistic, f_0, the smaller the *P-value*.



T (5) If the *P-value* is very large, then observed differences between the sample or group means could be explained as being just due to chance alone.

Questions 34 to 42 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

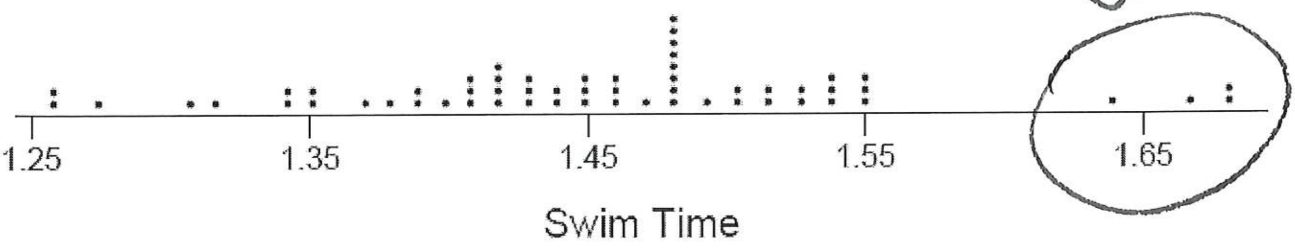


Figure 5: Dotplot of Swim Time (in minutes) for the New Zealand development squad

Summary Statistics and Confidence Interval

	N	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval	
					Lower	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

1/2

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.

2 indep samples

$H_0: \mu_1 - \mu_2 = 0$

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

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 34 **to** 42 refer to the **Swim Performance Study** information given above, on page 55. 95% CI : (1.4298, 1.4788)

34. 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

35. 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 x n → double accuracy & halve the width...

36 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 un related 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!

37. 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 37 and 39 refer to the **Swim Performance Study** information given above, on pages 55 and 60.

38. Assuming the student interpreted the correct *t*-test, which **one** of the following statements is **false**? (Use Tables 8 and 9 on pages 56 and 57 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. $\bar{x}_{diff} = 3.175$

39 Suppose Table 8, page 56 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 40 to **42** refer to the **Swim Performance Study** information given above, on page 58.

40. In a two-sample t-test on the **Differences** for the **Control** and **Rest** treatment groups (Table 10, page 58), 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.

41. Using Table 10, page 58, 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
(2) 2.45
(3) 0.99
(4) 1.56
(5) 1.87

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

(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: $\mu_{\text{Control}} - \mu_{\text{Rest}} \cancel{=} 0$

(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: $\cancel{\bar{x}}_{\text{Control}} - \cancel{\bar{x}}_{\text{Rest}} \cancel{=} 0$

(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

A. **(1)**	B. **(3)**	C. **(3)**	D. **(5)**	E. **(2)**	F. **(4)**
G. **(4)**	H. **(4)**	I. **(4)**	J. **(3)**	K. **(1)**	L. **(4)**
M. **(2)**	N. **(4)**	O. ~~**(3)**~~ (4)	P. **(3)**	Q. **(2)**	R. **(3)**
S. **(4)**	T. **(2)**	U. **(5)**	V. **(2)**	W. **(2)**	X. **(5)**
Y. **(2)**	Z. **(1)**	AA. **(2)**	BB. **(3)**	CC. **(2)**	DD. **(5)**

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

What should I do next?

- Do Assignment 3!
- Go through the Chapter 6, 7 and 8 blue pages. For each chapter, this includes *notes*, a *glossary*, *true/false statements*, *Sample Exam Questions*, and *tutorial* material.
- Attend the optional Chapters 7 & 8 tutorials.
- Do all the problems in this workshop handout and mark them. If you get a question wrong, have a look at the working on Leila's scanned slides at www.tinyURL.com/stats-HTM to see how she did it.
- Try Chapter 6, 7 and 8 questions from three of the past five exams on Canvas (get them from *Modules → Past Tests and Exams (with answers)* and use the *Exam questions index* document from there to identify the questions from Chapters 6, 7 and 8!)
- If you get anything wrong and don't know why, get some help. You can post a question on Piazza (search first as it may have already been asked!), or talk to someone about it (your lecturer, an Assistance Room tutor or Leila).

FORMULAE

Confidence intervals and t-tests

Confidence interval: $\quad estimate \pm t \times \mathrm{se}(estimate)$ Ch 6,7,8,10

t-test statistic: $\quad t_0 = \dfrac{estimate - hypothesised\ value}{standard\ error}$ Ch 7,8,10

Applications:

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

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

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

   $estimate = \overline{x}_1 - \overline{x}_2; \quad df = \min(n_1 - 1, n_2 - 1)$ Ch 6/7

~~4. Difference between two proportions $\boldsymbol{p_1 - p_2}$:~~

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

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

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

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

The F-test (ANOVA)

F-test statistic: $f_0 = \dfrac{s_B^2}{s_W^2}; \quad df_1 = k - 1, \; df_2 = n_{\text{tot}} - k$ 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: $\quad \widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$ Ch 10

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