

Department of Statistics
STATS 101/101G/108
Mid-semester test: First Semester, 2020

(Time allowed: 24 Hours)

- This test consists of **20** multiple-choice questions. All questions have a single **correct** answer. All questions carry the same mark value.
- No mark is deducted for an incorrect answer.
- Calculators are permitted.
- Formulae are provided (see the attachment Formulae.pdf).

References

Brase, C. and Brase, C. (2016). *Understandable Statistics*. Cengage learning.

Junco, R. (2013). Comparing actual and self-reported measures of facebook use. *Computers in Human Behavior*, **29**, 626–631.

LaFrance, M. and Hecht, M. A. (1995). Why smiles generate leniency. *Personality and Social Psychology Bulletin*, **21**(3), 207–214.

By submitting my test answers on Canvas, I am declaring that the test is my own work. This means that for the 24-hour duration of the test, I confirm that I will not discuss the content of the test with anyone else, I will not give any assistance to another student taking this test and I will not receive any assistance from any person or tutoring service.

Please note: When a question has a box around it, clicking on it will take you to the information you need to look at to help you answer that question.

VERSION 3

Questions 1 to 3 refer to the following information.

Inorganic phosphorus is a naturally occurring element found in plants and animals at different concentrations. For example, the concentration of phosphorus increases in the following order:

fruits → vegetables → cereals → nuts → corpses.

Geochemical surveys take soil samples to determine phosphorus content in parts per million (ppm). A high phosphorus content may indicate an ancient burial site, food storage, or even a rubbish dump. The Hill of Tara is a very important archaeological site in Ireland (Brase and Brase, 2016). Random samples from two regions in Tara ($n_1 = 15$ and $n_2 = 14$) gave the phosphorus measurements shown in Figure 1.

[Link back to Q1](#)

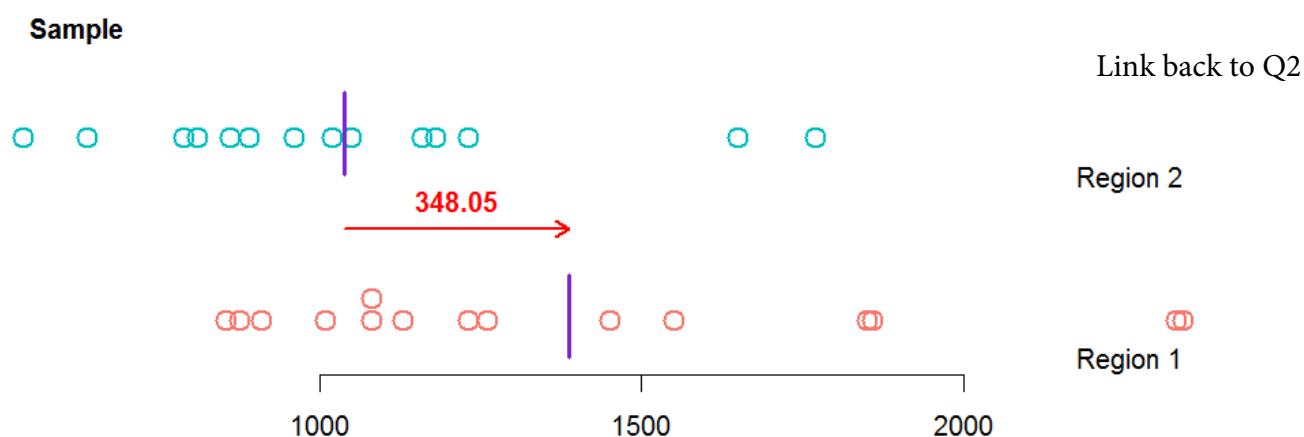


Figure 1: Phosphorus measurements (ppm) from two regions in the Hill of Tara

A bootstrap confidence interval was constructed to estimate the difference between the mean phosphorus measurements for the two regions. Figure 2 shows the bootstrap distribution and the bootstrap confidence interval.

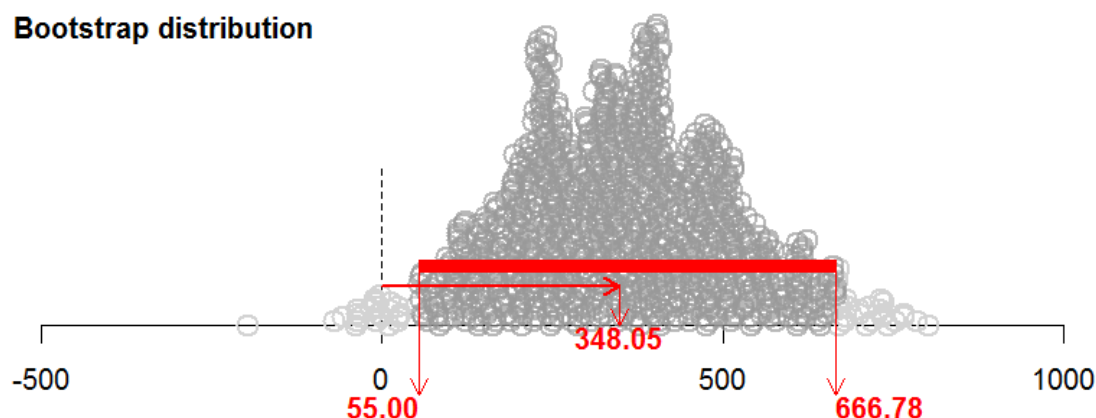


Figure 2: Bootstrap distribution and bootstrap confidence interval

VERSION 3

1. Which **one** of the following statements is **false**?
 - (1) Since the bootstrap distribution includes zero, it is plausible that there is no difference between the true mean phosphorus measurement for Region 1 and that for Region 2.
 - (2) The observed phosphorus measurements for Region 1 were, on average, higher than those for Region 2.
 - (3) We cannot be certain that the difference between the true mean phosphorus measurement for Region 1 and that for Region 2 is between 55 and 667 ppm.
 - (4) It is a fairly safe bet that the true mean phosphorus measurement for Region 1 is somewhere between 55 and 667 ppm higher than that for Region 2.
 - (5) 348.05 is an estimate of the difference between the true mean phosphorus measurement for Region 1 and that for Region 2.

2. Which **one** of the following statements about the bootstrap distribution plot in Figure 2, page 2, is **false**?
 - (1) We would expect another bootstrap distribution calculated from the same sample data to be similar to that shown.
 - (2) The bootstrap distribution plot shows the extent of the variation in the differences between re-sample means of 1000 re-samples (with replacement) from the original sample data.
 - (3) Each dot represents the difference between the re-sample mean of Region 1 and the re-sample mean of Region 2, for one re-sample.
 - (4) For most of the 1000 re-samples, the re-sample mean of Region 1 is higher than the re-sample mean of Region 2.
 - (5) For approximately 5% of the 1000 re-samples, the difference between the re-sample mean of Region 1 and the re-sample mean of Region 2 is greater than 667.

VERSION 3

3. Suppose the way they took the sample in Region 1 was to grid off the area into 94 equally sized units, with each unit being allocated an identification number from 1 to 94. Assume the random digits below were used to take the simple random sample of 15 units. Start at the beginning of the row and use consecutive groups of digits to find the **first five** units that would have been chosen in the sample.

20958 53285 99071 94273 01863 49005

The identification numbers of the **first five** units to be chosen were:

- (1) 20, 85, 32, 85, 7
- (2) 20, 95, 85, 32, 99
- (3) 20, 85, 32, 71, 94
- (4) 20, 85, 32, 7, 19
- (5) 20, 53, 94, 1, 49

VERSION 3

Questions 4 and 5 refer to the following information.

How I Met Your Mother is a ‘situation comedy’ television show from the U.S. that aired on CBS for 9 seasons. Early seasons received generally positive ratings but later seasons’ ratings were more varied. User rating data from *IMDb.com* were collected for all 208 episodes and grouped into three groups: the early (1–3), middle (4–6) and late (7–9) seasons.

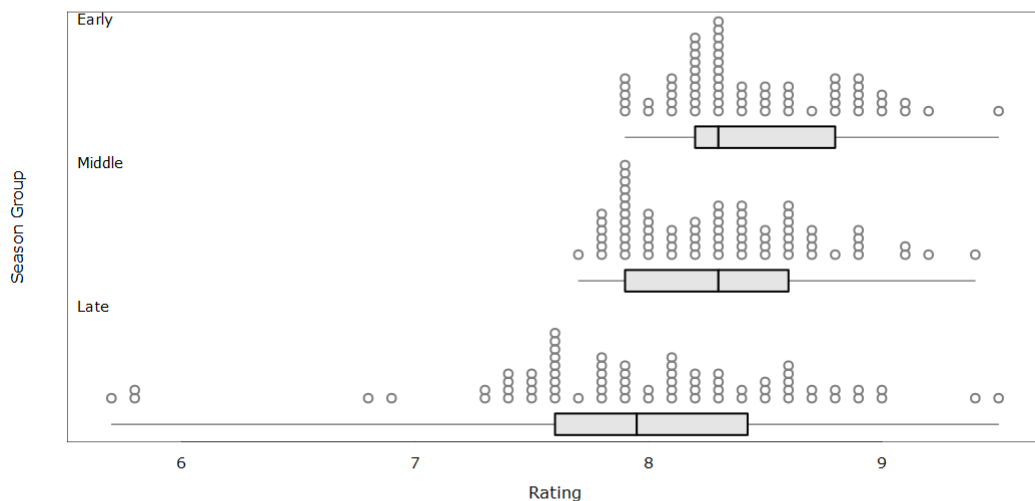


Figure 3: **Rating** by **Season Group**

4. Which **one** of the following statements about Figure 3 is **false**?
- (1) The early and middle seasons have the same median rating.
 - (2) The mean rating is higher than the median rating for the early seasons.
 - (3) There is overlap in the ratings between all three levels of **Season Group**.
 - (4) These plots are an example of a trimodal distribution.
 - (5) The early season ratings are positively skewed.
5. Consider only the **Late** seasons (seasons 7–9).

Which **one** of the following is the only possible **correct pair** of values for the mean and standard deviation of the episode ratings?

- (1) mean = 8.0 standard deviation = 0.2
- (2) mean = 7.2 standard deviation = 0.7
- (3) mean = 8.5 standard deviation = 0.2
- (4) mean = 8.5 standard deviation = 1.5
- (5) mean = 8.0 standard deviation = 0.7

VERSION 3

Questions 6 to 9 refer to the following information.

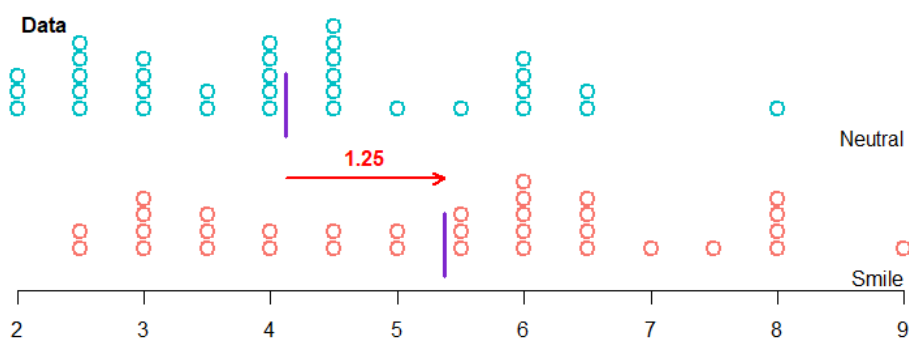
LaFrance and Hecht (1995) investigated whether a person suspected of cheating is judged more leniently when the evidence used to make the judgement has a photograph of them smiling compared to when it has a photograph of them with a neutral facial expression.

68 volunteers were randomly assigned to one of two equal-sized groups. The study participants were asked to play the role of a university disciplinary panel judging a possible case of academic misconduct. Each participant was given the summary of evidence of the case in which a student was suspected of cheating in an examination.

Participants were also given background information on the suspect which was identical except for a photograph of the face of the suspect. Each participant in one group was given a photograph of the suspect smiling and each participant in the other group was given a photograph of the suspect with a neutral expression. The participants independently reviewed the information they received and were then asked several questions.

Based on participants' answers to the questions, a leniency score was calculated for each participant. Higher scores indicate a more lenient punishment for the suspected student.

Dot plots of the data and the difference between the two group means are displayed in Figure 4.

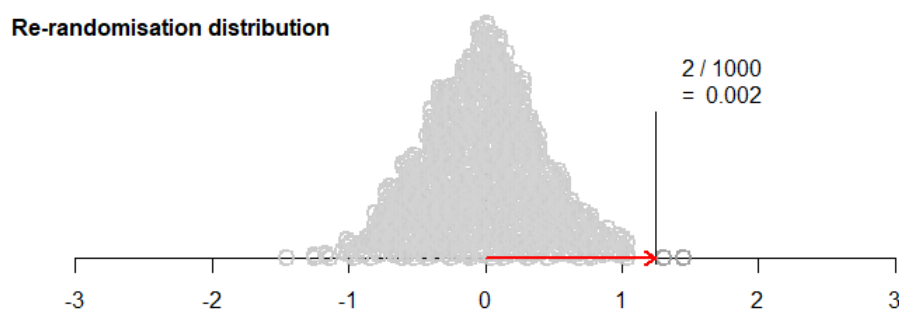


[Link back to Q6](#)

[Link back to Q7](#)

Figure 4: Leniency score

A randomisation test was performed on the data and the resulting output is displayed in Figure 5.



[Link back to Q8](#)

Figure 5: Re-randomisation distribution

VERSION 3

6. Which **one** of the following statements is **false**?
- (1) The response variable for this study is the leniency score for each participant.
 - (2) This study uses a randomised block design because the participants were blocked on whether they were given a photograph of the suspect smiling or a photograph of the suspect with a neutral expression.
 - (3) The units were the 68 participants.
 - (4) The randomisation process does not guarantee that the same number of female and male participants were assigned to each group.
 - (5) The 68 participants were randomly assigned to one of the two groups in order to try and make the two groups as similar as possible apart from the photograph they were given.
7. Which **one** of the following statements about the randomisation test output in Figure 5 is **true**?
- (1) At least 2 of 1000 re-randomisations under chance alone produced a difference between group means of exactly 1.25.
 - (2) 2 of the 1000 re-samples with replacement from the original data produced a difference between group means of 1.25 or more.
 - (3) 2 of the 1000 re-randomisations under chance alone produced a difference between group means of 1.25 or more.
 - (4) If this study was repeated 1000 times then 2 of those 1000 studies would produce an observed difference between group means of 1.25 or more.
 - (5) 2% of the 1000 re-randomisations under chance alone produced a difference between group means of 1.25 or more.
8. Which **one** of the following statements is **not** a valid interpretation of the randomisation test result?
- (1) We may claim that a photograph of the suspect smiling caused higher average leniency scores.
 - (2) An observed difference between the group means of 1.25 or more is unlikely when chance is acting alone.
 - (3) The observed difference between the group means of 1.25 is large enough for us to claim that the facial expression of the suspect in the photograph had an effect on the leniency scores.
 - (4) It is not believable that chance played any part in the observed difference between the two group means of 1.25.
 - (5) We may claim there is a difference between average leniency scores when the photograph was of the suspect smiling and when the photograph was of the suspect with a neutral expression.

VERSION 3

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

In this experiment **chance is acting alone** means that:

- (1) the leniency score for each participant was randomly selected by the participant.
- (2) it would not have mattered to which group the participant had been randomly assigned, the leniency score for each participant would be the same.
- (3) the observed difference between the mean leniency score for the group that was given a photograph of the suspect smiling and the corresponding mean for the group that was given a photograph of the suspect with a neutral expression is not due to the photograph that was given.
- (4) the observed difference between the mean leniency scores for the two groups is due solely to the random assignment of the participants to the two groups.
- (5) the leniency score for each participant has nothing to do with whether the participant was given a photograph of the suspect smiling or with a neutral expression.

10. Which **one** of the following statements about Normality-based confidence intervals is **false**?

- (1) A confidence interval for a population parameter is a set of plausible values for that parameter.
- (2) A 95% confidence interval is used to estimate the value of a population parameter.
- (3) The level of confidence is the long-run success rate for a method which aims at producing confidence intervals which include the unknown value of the population parameter.
- (4) Producing a 95% confidence interval is a way of accounting for the uncertainty of an estimate due to sampling variability.
- (5) Once the confidence limits have been found, then we need to determine whether or not a confidence interval actually includes the value of the population parameter.

VERSION 3

Questions 11 to 13 refer to the following information.

Junco (2013) conducted a study about the use of Facebook by students. A random sample of students from a small U.S. university was contacted via email and asked to complete a survey.

One of the questions asked in the survey was: *On average, how many times per day do you check Facebook? Please give your answer as a number.*

11. Which **one** of the following types of error would **not** be a potential source of nonsampling error in this survey?
- (1) Survey-format effects
 - (2) Selection bias
 - (3) Question effects
 - (4) Behavioural considerations
 - (5) Self-selection bias

Questions 12 to 13 refer to the following additional information.

Summary statistics were calculated for the numerical responses to this question and the required information was entered into the STATS10x online t -procedures tool, as shown in Figure 6.

t-procedures

Single mean

$\bar{x}$	7
s	11
n	110

Confidence level	95 %
------------------	------

$se(\bar{x}) = 1.0488$

$t\text{-multiplier} = 1.982$

Figure 6: A screenshot of the t -procedures tool

VERSION 3

Let μ be the mean number of times per day Facebook was checked by students at this small U.S. university.

12. The 95% confidence interval for μ is:

- (1) (1.05, 1.98)
- (2) (4.92, 9.08)
- (3) (−4.00, 18.00)
- (4) (5.02, 8.98)
- (5) (5.95, 8.05)

13. Suppose instead that 440 students had responded to the survey and that for this sample the number of times per day Facebook is checked had a standard deviation of 11.

Comparing the 95% confidence interval constructed in **Question 12**, to the 95% confidence interval constructed from this sample, which **one** of the following statements is **false**?

- (1) The standard error of the sample mean calculated from this sample will be smaller.
- (2) The margin of error for the new confidence interval will be smaller.
- (3) The new confidence interval will be more precise.
- (4) The new confidence interval will be wider.
- (5) The t -multiplier for the new confidence interval will be slightly smaller.

VERSION 3

Questions 14 to 20 refer to the following information.

Taste.com.au is a food website that provides free recipes. In January 2016, a random sample of the main meal recipes available was obtained by scraping data from the website.

Some of the variables selected were:

Total time	The total time to prepare and cook the main meal, in minutes
Total ingredients	The total number of individual ingredients required to make the main meal
Difficulty	The difficulty rating of the recipe (Easy, Difficult)
Season	The season of the year the main meal recipe has been listed under (Spring-Summer, Autumn-Winter)

14. Which **one** of these displays would be the **most suitable** to explore the relationship between the variables **Total ingredients** and **Total time**?

- (1) Scatter plot
- (2) Pie chart
- (3) Stacked bar chart
- (4) Two-way table of counts
- (5) Side-by-side bar chart

15. 37% of the recipes in this sample were listed under the season Autumn-Winter, the rest were listed under Spring-Summer. 80% of the recipes listed under the season Autumn-Winter were rated easy. 87% of the recipes listed under the season Spring-Summer were rated easy.

Of the recipes rated as easy, the proportion that were listed under Autumn-Winter is approximately:

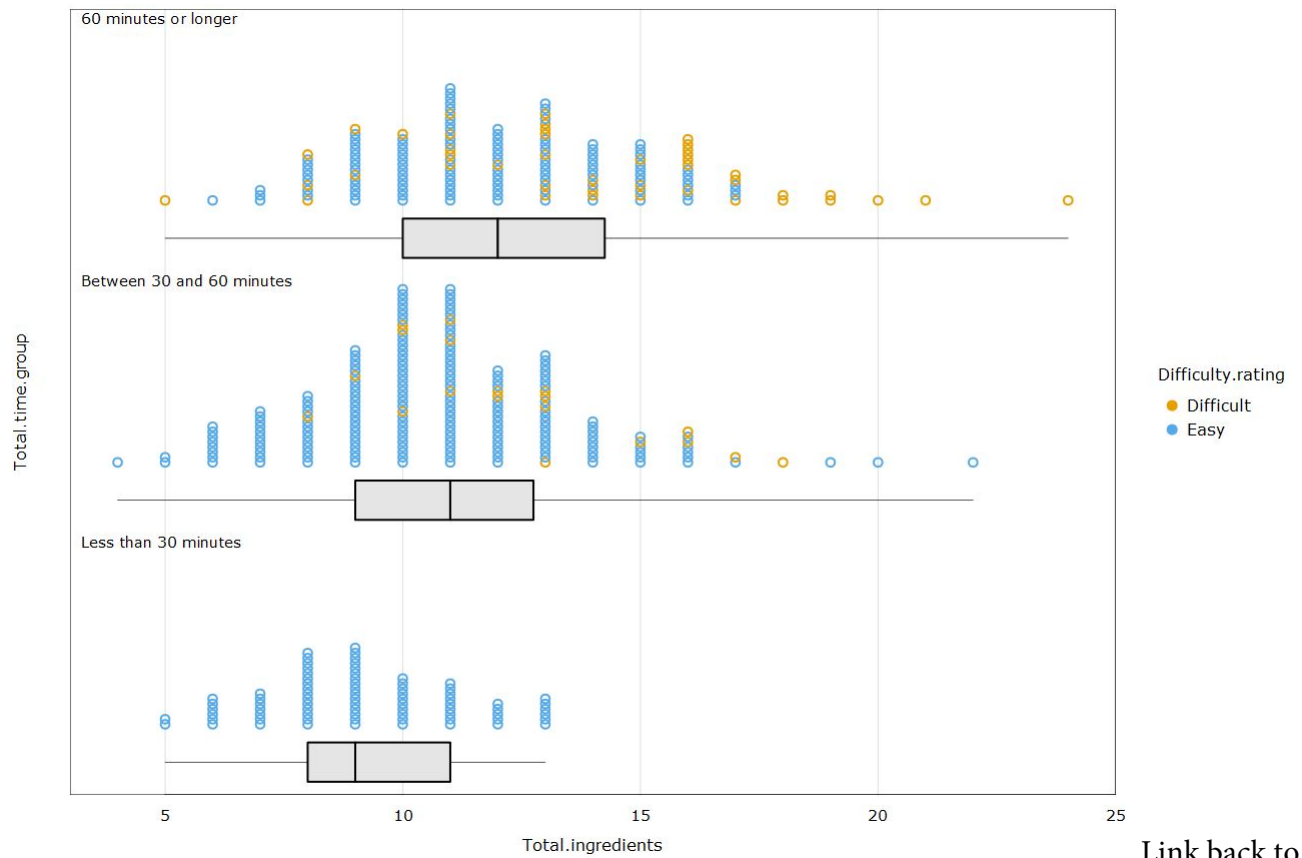
- (1) 0.370
- (2) 0.463
- (3) 0.351
- (4) 0.800
- (5) 0.479

VERSION 3

A new categorical variable called **Total time group** was created by grouping times from the **Total time** variable. iNZight was used to explore the variable **Total time group** and its relationship with some of the other variables in the data set. The two different plots created are shown in Figure 7, page 13.

16. Which **one** of the following statements about the sample data shown in Figure 7 is **false**?
- (1) The distribution of total ingredients for recipes that had a total time of less than 30 minutes looks approximately symmetric.
 - (2) The majority of the recipes rated easy had a total time of less than 60 minutes.
 - (3) None of the recipes rated difficult had a total time of less than 30 minutes.
 - (4) There were more recipes rated easy than difficult.
 - (5) A higher proportion of the recipes with a total time of 60 minutes or longer were rated difficult than were rated easy.

VERSION 3



[Link back to Q16](#)

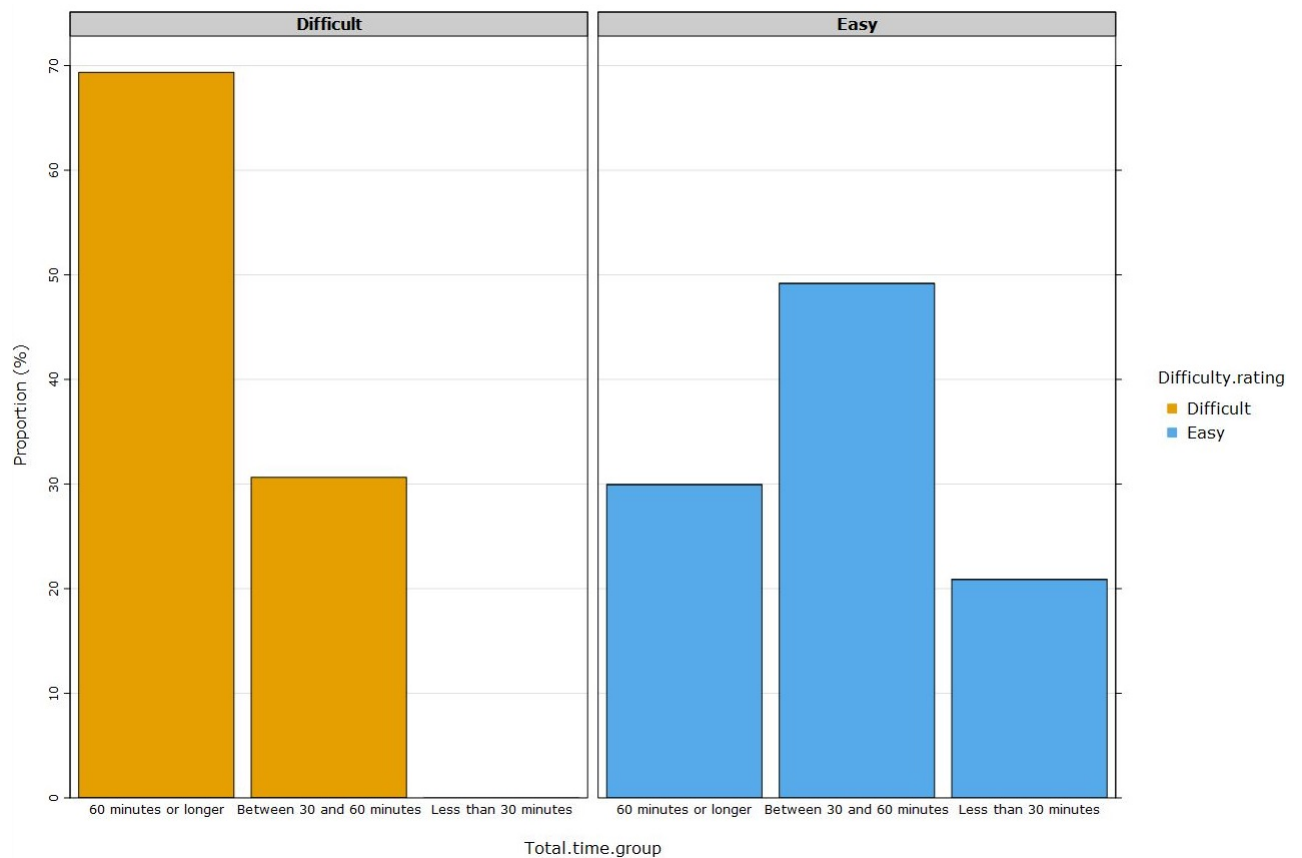


Figure 7: Exploratory data analysis for **Total time group**

VERSION 3

Questions 17 to 20 refer to the following additional information.

A bootstrap confidence interval was constructed to estimate the difference between the median number of ingredients required for recipes rated as difficult and the median number of ingredients required for recipes rated as easy, for all main meal recipes on *taste.com.au* as of January 2016. Two plots from the associated VIT output are shown in Figures 8 and 9. Figure 8 shows the sample data, sample medians and the observed difference between the medians. Figure 9 shows the bootstrap distribution and the bootstrap confidence interval.

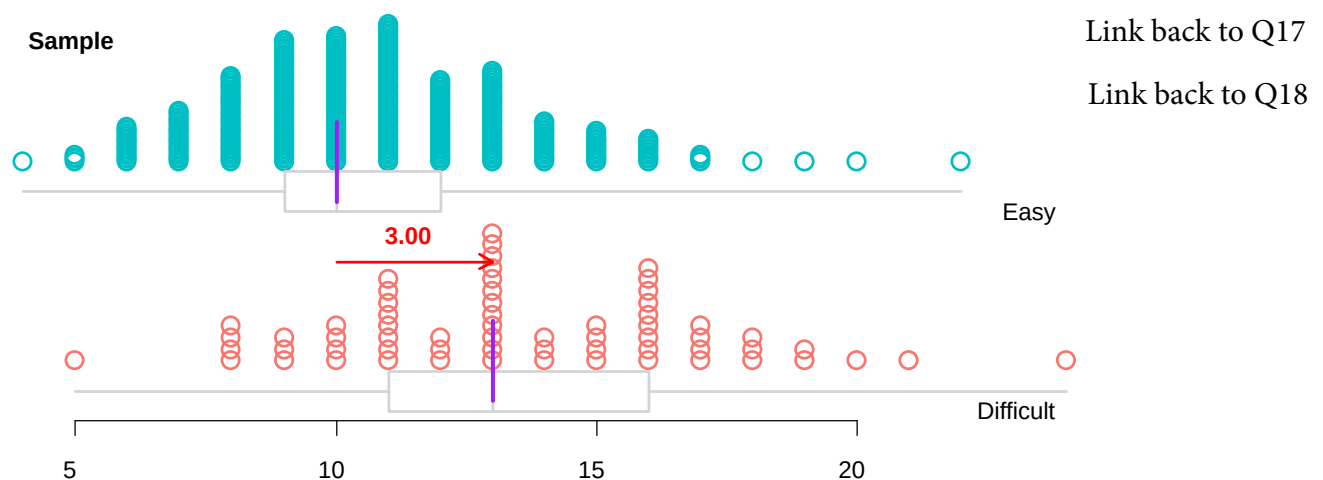


Figure 8: Number of ingredients required

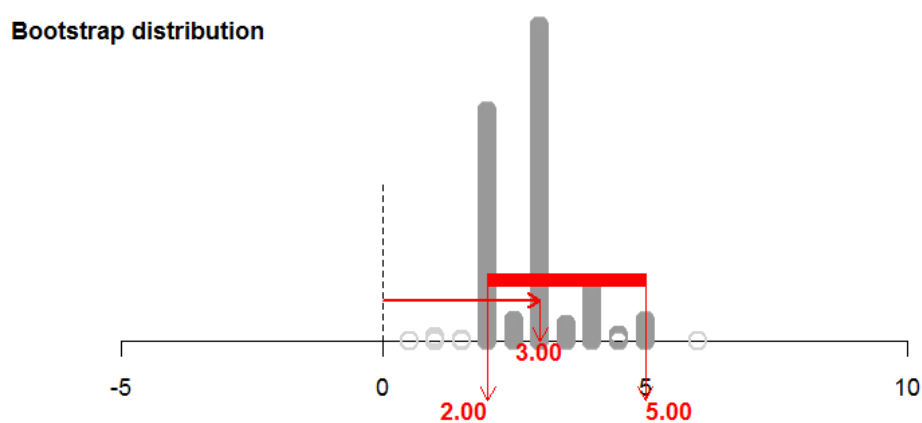


Figure 9: Bootstrap distribution and bootstrap confidence interval

VERSION 3

17. Which **one** of the following statements about the **sample data** shown in Figure 8, page 14, is **false**?

- (1) The recipes that were rated difficult require more ingredients on average than those that were rated easy.
- (2) The recipes that were rated easy had a lower mean and median number of ingredients than the recipes rated difficult.
- (3) Knowing whether a recipe is rated easy or difficult does not appear to be related to the number of ingredients.
- (4) The recipe that required the most number of ingredients was rated difficult and the recipe that required the least number of ingredients was rated easy.
- (5) There were fewer recipes rated difficult than were rated easy.

18. Suppose it is claimed that, for the main meal recipes on *taste.com.au* as of January 2016, the median number of ingredients required for recipes rated difficult is 4.5 higher than the median number of ingredients required for recipes rated easy.

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

- (1) This claim is plausible because the lower limit of the bootstrap confidence interval is 2.
- (2) This claim is plausible because the entire bootstrap confidence interval is positive.
- (3) This claim is plausible because 4.5 is in the bootstrap confidence interval.
- (4) This claim is not plausible because most of the differences in the bootstrap distribution are less than 4.5.
- (5) This claim is not plausible because the bootstrap distribution does not contain many re-sample differences of 4.5.

VERSION 3

Questions 19 and 20 refer to the following additional information.

Let:

p_{60} be the proportion of all **Autumn-Winter** main meal recipes on *taste.com.au* as of January 2016 that have a total preparation time of 60 minutes or longer

and

p_{30} be the proportion of all **Autumn-Winter** main meal recipes on *taste.com.au* as of January 2016 that have a total preparation time of between 30 and 60 minutes.

A 95% confidence interval for $p_{60} - p_{30}$ is given by $(-0.115, 0.183)$.

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

- (1) The point estimate of $p_{60} - p_{30}$ is 0.034.
- (2) The sampling situation for calculating the standard error of this estimate, $se(\hat{p}_{60} - \hat{p}_{30})$, is two independent samples.
- (3) The width of the confidence interval is 0.298.
- (4) Sample size is taken into account when calculating the standard error of this estimate, $se(\hat{p}_{60} - \hat{p}_{30})$.
- (5) The margin of error of the confidence interval is 0.149.

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

With 95% confidence, the proportion of all Autumn-Winter main meal recipes on *taste.com.au* as of January 2016 that have a total time to prepare and cook of 60 minutes or longer is estimated to be:

- (1) 18.3 percentage points greater than the proportion of those that have a total time to prepare and cook of between 30 and 60 minutes.
- (2) 11.5 percentage points less than the proportion of those that have a total time to prepare and cook of between 30 and 60 minutes.
- (3) the same as the proportion of those that have a total time to prepare and cook of between 30 and 60 minutes.
- (4) somewhere between 11.5 percentage points less than and 18.3 percentage points greater than the proportion of those that have a total time to prepare and cook of between 30 and 60 minutes.
- (5) somewhere between 18.3 percentage points less than and 11.5 percentage points greater than the proportion of those that have a total time to prepare and cook of between 30 and 60 minutes.