Answers

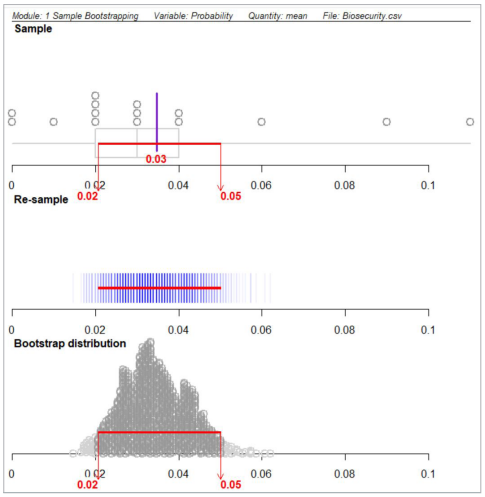
Results in most questions will vary slightly as they rely on bootstrap re-sampling.

Page 22

1. a) 95% percentile confidence interval for the mean by bootstrap re-sampling is 0.021 to 0.052.

b) No, the confidence interval goes beyond 0.05 so they are not justified in saying they are less than 0.05.

c)



d) The sample is only 15 elements and is not a normal distribution

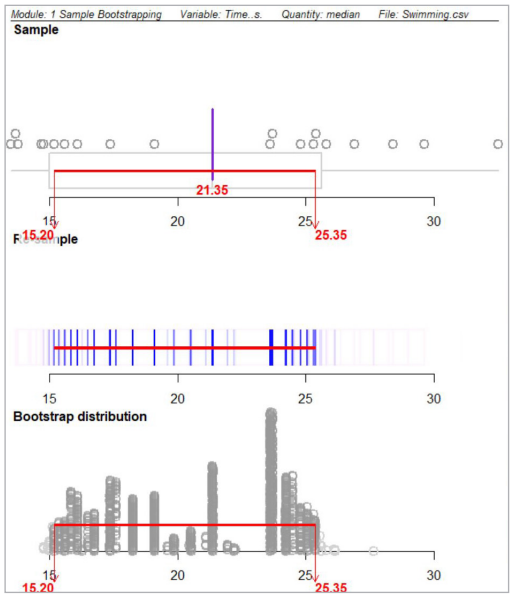
2. a) Unusual distribution, small sample size and information required about the median.

b) 95% percentile confidence interval for the mean by bootstrap re-sampling is 18.5 seconds to 23.6 seconds (1 dp).

95% percentile confidence interval for the median by bootstrap re-sampling is 15.2 seconds to 25.4 seconds (1 dp).

c) The bimodal distribution means the median swings greatly depending on which end of the distribution is over represented in the bootstrap sample.

d)



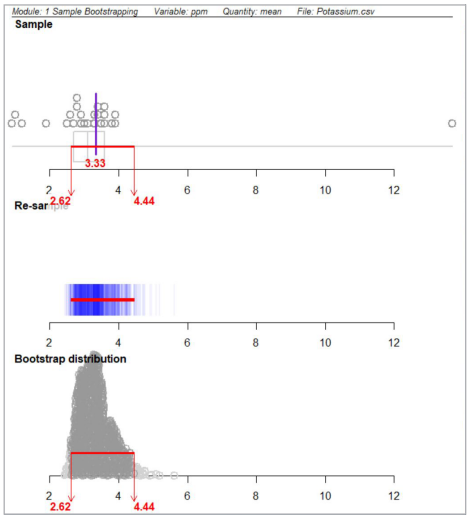
Page 23

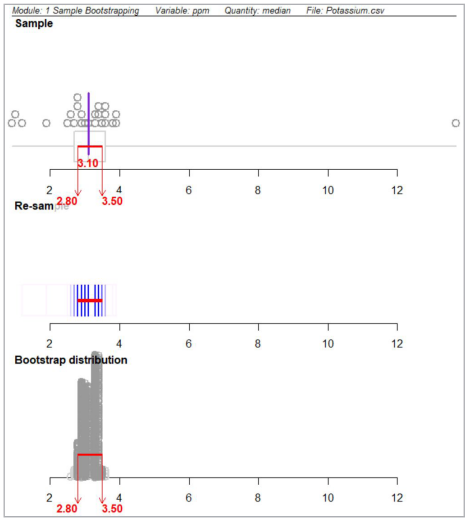
3. a) 95% percentile confidence intervals by bootstrap re-sampling are:
Mean 2.62 ppm to 4.44 ppm
Median 2.80 to 3.50 ppm.

b) The extreme value of 13.7 ppm affects the mean resulting in high means every time it is selected and even higher if it is selected more than once.

c) The extreme value of 13.7 ppm does not affect the median as the median is the middle number.

d) and e)





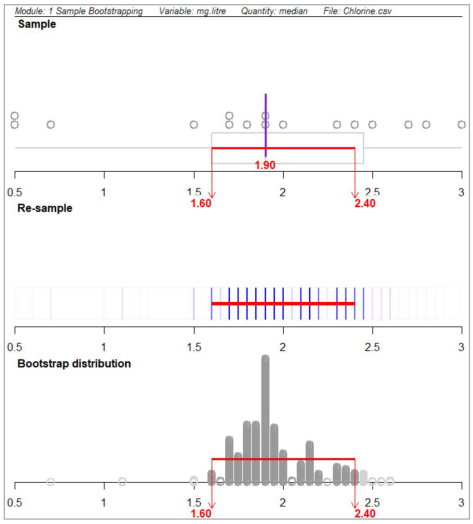
4. a) The 95% percentile confidence interval for the mean by bootstrap re-sampling is 1.49 to 2.21 mg/l.
The 95% percentile confidence interval for the median by bootstrap re-sampling is 1.60 to 2.40 mg/l.

b) Both the mean and median of the population could be over 2 mg/l. The school is not compliant at the 95% percentile level with the Ministry of Education requirement.

c) The distribution is not normal with observations spread erratically between 0.5 and 2.8 mg/l. The population distribution is likely to be similar to the sample distribution.

Page 23 Q4 cont...

d)

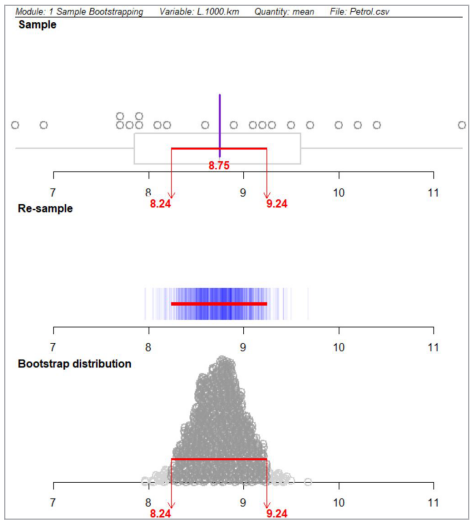


Page 24

5. a) 95% percentile confidence for the mean by bootstrap re-sampling is 8.24 litres / 100 km to 9.24 litres / 100 km.

b) 95% percentile confidence interval for the median by bootstrap re-sampling is 7.90 litres / 100 km to 9.40 litres / 100 km.

c)



d) For petrol consumption we are concerned with the total used so extreme results will be important. In this case the distribution does not include any extreme values.

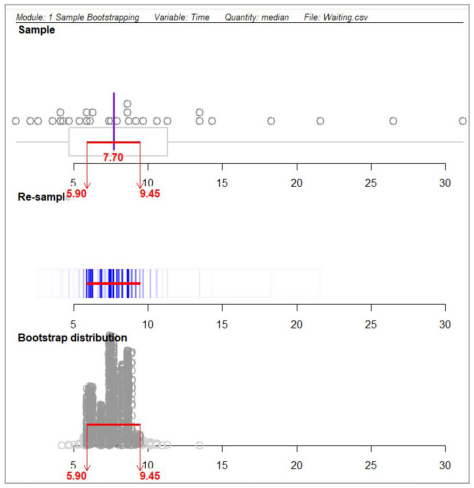
6. a) 95% percentile confidence interval for the mean by bootstrap re-sampling is 7.22 minutes to 12.17 minutes.

b) 95% percentile confidence interval for the median by bootstrap re-sampling is 5.90 minutes to 9.45 minutes.

c) The relatively few long waiting times resulted in the means tending to be higher than the medians. As half the patients would be waiting for less than the median time it would be the best measure of the waiting time.

Page 24 Q6 cont...

d)

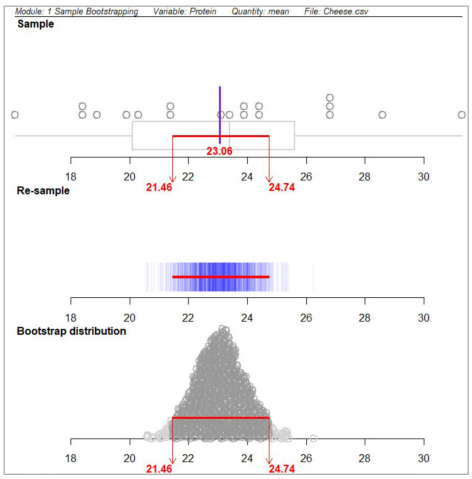


e) The surgery should say that patients could expect to wait between 6 and 10 minutes but in unusual circumstances longer waiting times are possible.

Page 25

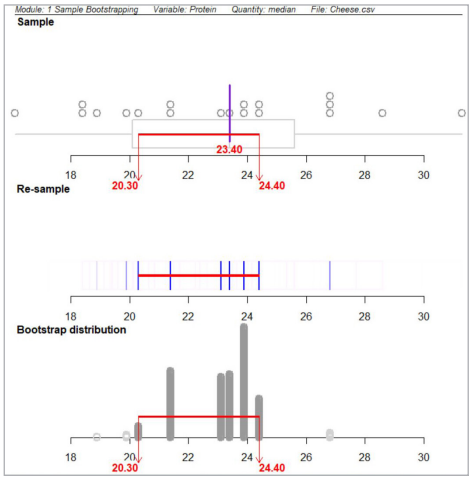
7. a) The 95% percentile confidence interval for the mean weight of protein by bootstrap re-sampling is from 21.46 g to 24.74 g per 100 g of cheese.

b)



8. a) The 95% percentile confidence interval for the median weight of protein by bootstrap re-sampling is from 20.3 g to 24.4 g per 100 g of cheese.

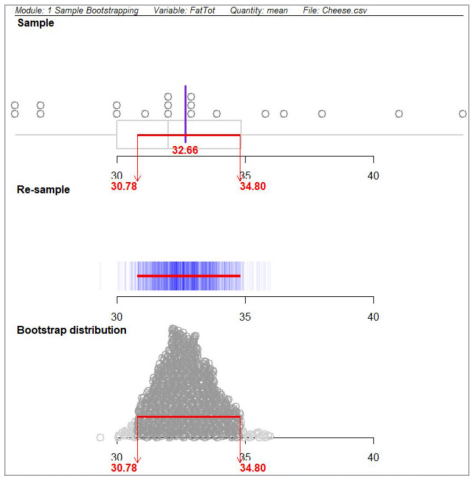
b)



Page 25 cont...

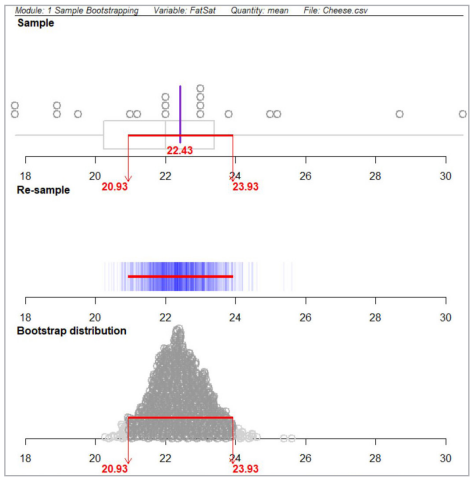
9. a) The 95% percentile confidence interval for the mean weight of total fats by bootstrap re-sampling is from 30.78 g to 34.80 g per 100 g of cheese.

b)



10. a) The 95% percentile confidence interval for the mean weight of saturated fat by bootstrap re-sampling is from 20.93 g to 23.93 g per 100 g of cheese.

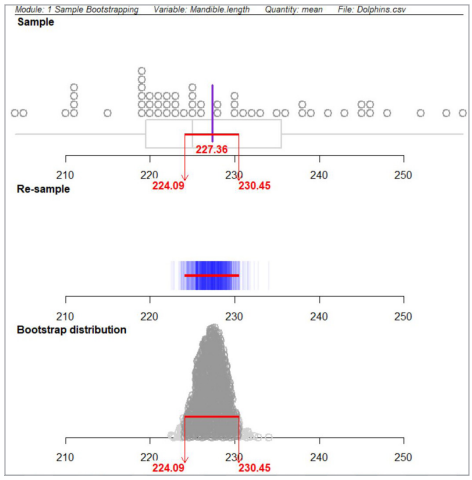
b)



Page 26

11. a) The 95% percentile confidence interval for the mean mandible length by bootstrap re-sampling is from 224.1 mm to 230.5 mm.

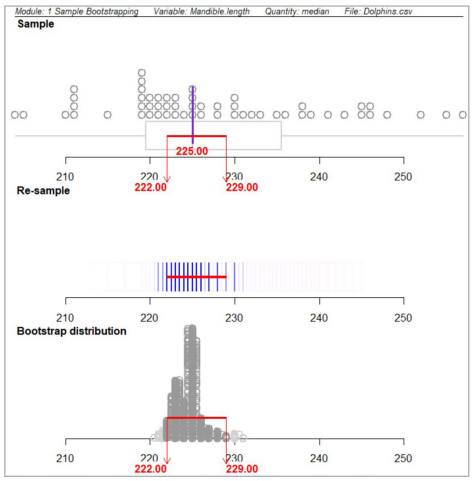
b)



Page 26 cont...

12. a) The 95% percentile confidence interval for the median mandible length by bootstrap re-sampling is from 222 mm to 229 mm.

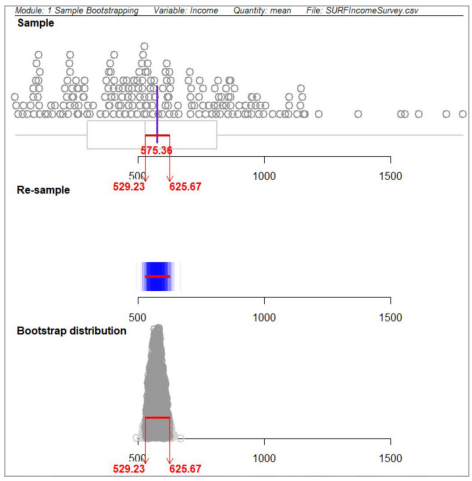
b)



c) The distribution of medians is made up of columns (discrete results). The interval range is the same but a few large results have the mean 2 mm higher at both ends.

13. a) The 95% percentile confidence interval for the mean income per week by bootstrap re-sampling is from $529.23 to $625.67.

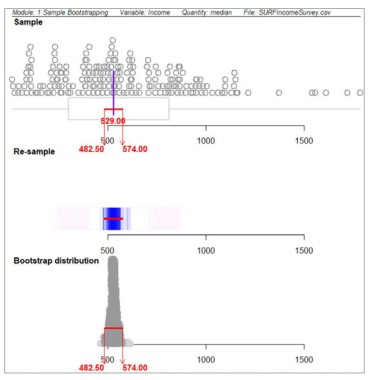
b)



c) The sample size of 200.

14. a) The 95% percentile confidence interval for the median income per week by bootstrap re-sampling is from $482.50 to $574.00.

b)



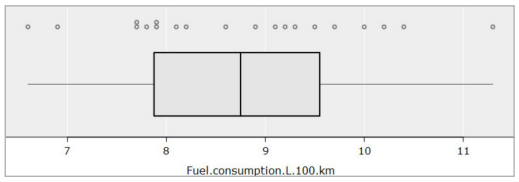
Page 26 Q14 cont...

14. c) You would expect the mean to be higher as there were some very high weekly incomes (e.g. $1789) that would affect the mean but would have had no effect on the median.

Page 28

15. a) The 95% percentile confidence interval for the mean petrol consumption, by bootstrap re-sampling, is $8.3 < \mu < 9.3$ L/100 km.

b)



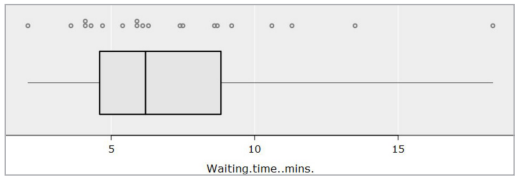
16. a) The 95% percentile confidence interval for the mean waiting time, by bootstrap re-sampling, is $5.8 < \mu < 9.1$ minutes.

b)

17. a) The 95% percentile confidence interval for the mean driving time, by bootstrap re-sampling, is $51.6 < \mu < 55.5$ minutes.

He should quote 52 to 56 minutes.

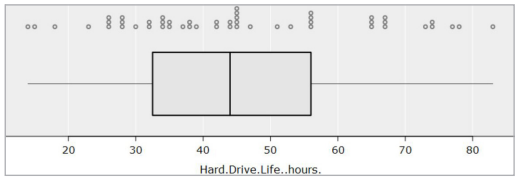
b)



18. a) The 95% percentile confidence interval for the mean hard drive life, by bootstrap re-sampling, is $40\,540 < \mu < 50\,310$ hours.

Minimum average life is 41 000 (2 sf).

b)

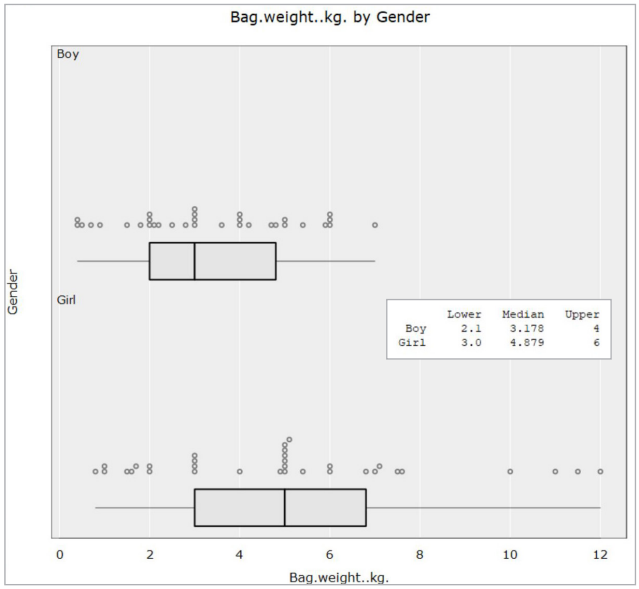


Page 29

19. a) The 95% percentile confidence interval for the median bag weight, by bootstrap re-sampling, is
Boys $2.2 <$ median < 4.0 kg
Girls $3.0 <$ median < 5.4 kg

Page 29 Q19 cont...

19. b)



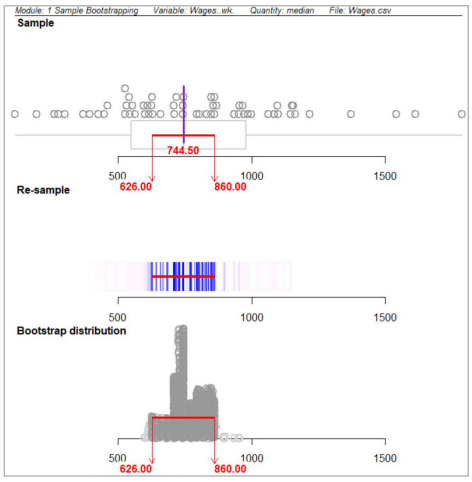
20. a) The 95% percentile confidence interval for the median income, by bootstrap re-sampling, is $628 < \text{median} < $859 per week

b)

21. a) The 95% percentile confidence interval for the median income, by bootstrap re-sampling, is $626 < \text{median} < $860 per week

b) As expected these are very similar. Bootstrap uses chance to create the distribution so every time the confidence interval will vary slightly.

c)

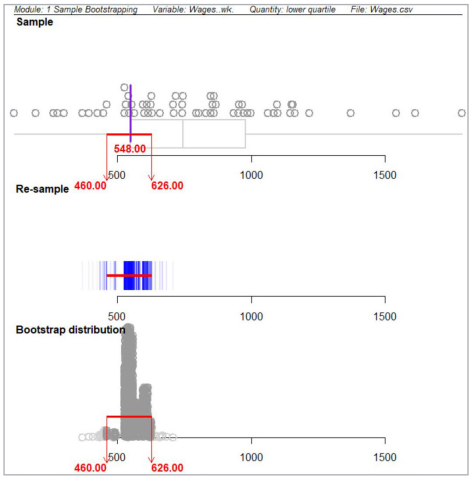


Page 29 cont...

22. a) The 95% percentile confidence interval for the lower quartile income, by bootstrap re-sampling, is \$460 < lower quartile < \$626 per week.

b) \$600 is within the confidence interval so the union does not have enough information to refute or challenge the company's' claim.

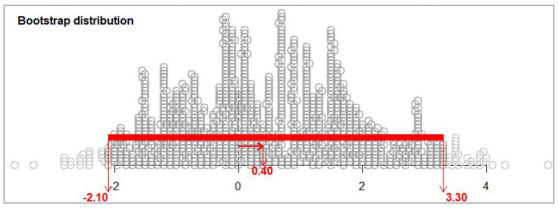
c)



Page 33

23. a) $-2.10 < \text{Median}_1 - \text{Median}_2 < 3.30$

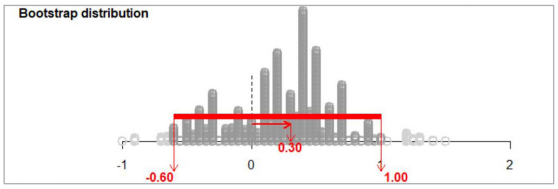
b)



c) There is insufficient evidence to indicate a difference in the assembly method as 0 is included in the interval. The median difference back in the population could be positive or negative.

24. a) $-0.60 < \text{Median}_1 - \text{Median}_2 < 1.00$

b)



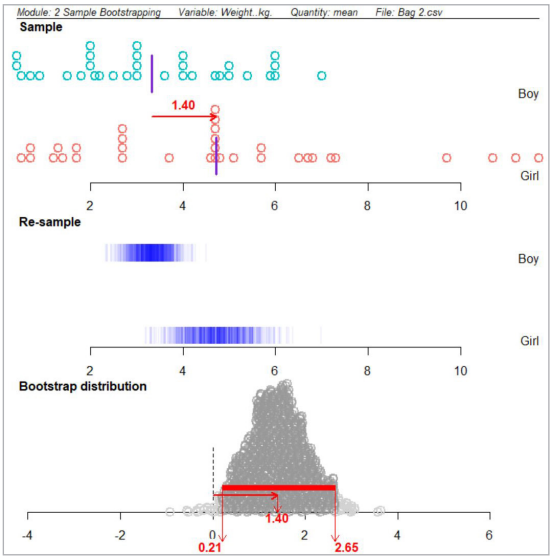
c) There is insufficient evidence to indicate a difference made as a result of the coaching technique as 0 is included in the interval. The median difference back in the population could be positive or negative.

Page 34

25. a) The 95% percentile confidence interval for the difference between mean weights is 0.21 kg to 2.65 kg.

Page 34 Q25 cont...

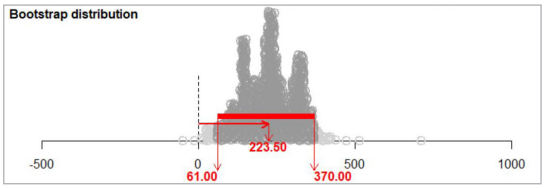
25. b)



c) This implies a difference exists back in the population as this interval is always positive. We expect the mean difference to be that the boy's bags are between 0.21 kg to 2.65 kg lighter.

26. a) The 95% percentile confidence interval of differences in median income is \$61 to \$370 per week.

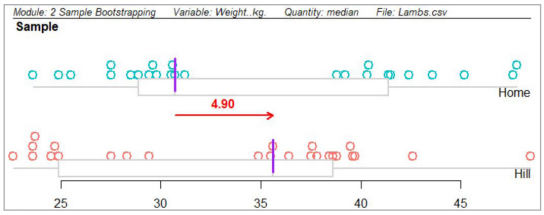
b)



c) At the 95% percentile level the population difference in median income could be as low as \$61 to as high as \$370 and as 0 is not in this interval the difference is significant so the assertion is supported.

Page 35

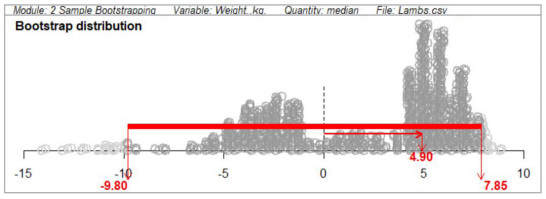
27. a)



b) The distributions are both bimodal. Very few data values in the middle which means that predictions for the median will be unreliable.

Page 41 Q27 cont...

27. c) The 95% percentile confidence interval for the difference between the medians is ⁻9.80 to 7.85.



d) As the confidence interval is both positive and negative (0 is in the interval), no inference can be made about differences in the population.

e) Possible lurking variables could be the locations had different water supply, different ages of sheep etc.

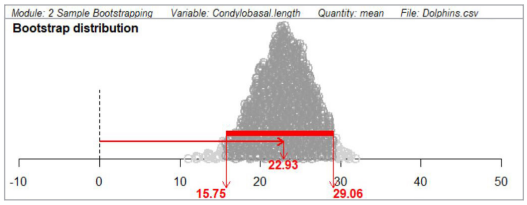
Page 36

28. a) **Results will vary slightly.**

	North Is.	South Is.
Count	13	46
Mean	304.5	281.6
Median	304.0	282.5
Confidence Interval		
Min CI	298.5	278.7
Max CI	309.8	284.5

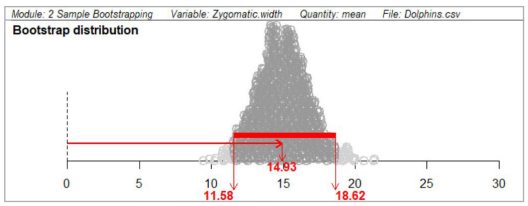
b) The 95% percentile confidence interval for the difference between the mean lengths, by bootstrap re-sampling, is 15.8 mm to 29.1 mm.

c)



d) At the 95% percentile level the population difference in mean condylobasal length is significant as 0 is not in the interval. The differences are always positive.

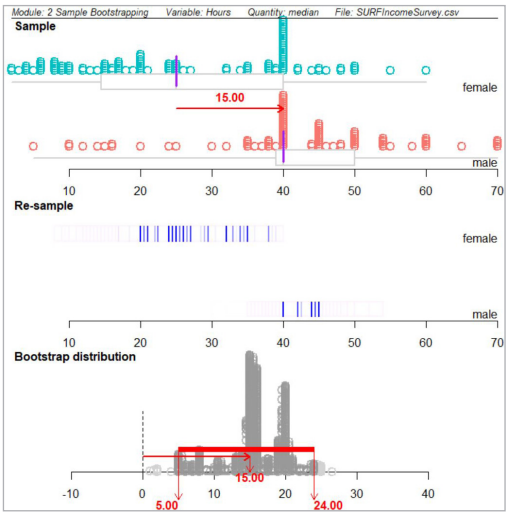
e) At the 95% percentile level using bootstrap re-sampling the population difference in mean Zygomatic width is between 11.6 mm to 18.6 mm from North to South Island.



Page 37

29. a) In the sample the median number of hours worked for males was 40 and the females' median was 15 hours less. Back in the population the differences in the median hours of paid work is between 5 hours to 24 hours.

b)



c) The 95% percentile for the population difference in median hours of paid work is significant as 0 is not in the interval. The differences are always positive. There is a difference back in the population between males and females median hours worked.

d) Many females would be working part time or be at home working as they often have responsibility for the children in a family. The SURF survey only looked at paid hours not total hours worked.