Stat 243 – Homework 04

Richmond Yevudza

February 17, 2022

#### 3.14

population mean = **300**

standard error = **5**

#### 3.18a

250 sample mean does not fall in the range of the sampling distribution therefore is **extremely unlikely to occur**.

#### 3.18b

305 sample mean isn’t uncommon in the sampling distribution therefore is **reasonably likely to occur from a sample of this size**.

#### 3.18c

315 sample mean is uncommon in the sampling distribution but the same sample mean takes place quite frequently in this sampling distribution therefore is **unusual but might occur occasionally**.

#### 3.23a

30 dollars shows the average co-payment for month’s supply of ointment for regular users. This is also a parameter and the notation for population mean is

27.90 dollars is the sample mean co-payment of regular users of sample size 75. This is a statistic and the notation for sample statistic is . .

#### 3.23b

The shape of the dot plot would be **bell shaped** because there will be one point of sample mean 27.90. The plot will also be centered at the average of =30.

#### 3.23c

There will be 1000 dots in the dot plot because there are 1000 samples. Each dot shows the sample mean of 75 co-payments.

#### 3.24a

It is given that the average household size in US is 2.61. The sampling distribution which is not centered at the average 2.61 is a biased distribution.

For unbiased distributions, we are looking at distributions A and D since the sampling distribution for these two are centered at 2.61 (the average).

The distribution B has more dots at the right of the center 2.61, which produces a higher average. Therefore, distribution B is a biased distribution. The distribution C has more dots at the left of the center 2.61 which produces a higher average. Therefore, distribution C is also a biased distribution.

#### 3.24b

Variability of distribution decreases with increase in sample size. Given that one distribution has the sample size n=100 and the other has the sample size n=500, looking at the sets of 1000 sample means, the distribution A has more variability and distribution D has less variability. Therefore, according to sample sizes n=100 is of the distribution A and n=500 is of distribution D.

#### 3.42

Interval estimate =

The interval estimate of the difference of population mean is **-3 to 13**.

#### 3.51a

30% is a **statistic**.

The correct notation of the sample proportion is .

#### 3.51b

The estimated proportion p is 0.30 by using the sample information which in this case was estimated with the information, the young people who were arrested by the age of 23 in US.

#### 3.51c

Interval estimate =

The interval estimate of population proportion p is **0.29 to 0.31**.

#### 3.51d

The actual proportion can’t be less than 25% because the range of reasonable values for the population proportion is 0.29 to 0.31.

#### 3.52a

Population: All people in the US whose ages are 18 or above 18.

Sample: The random sample of 147291 adults who were in touch with and inquire whether they got health insurance from an employer or not.

Relevant statistic:

#### 3.52b

Interval estimate =

#### 3.53

It is 95% confident that proportion of all adults who believe the necessity of a car is between 0.83 and 0.89.

#### 3.58a

This is a matched pairs experiment because in the experiment of urinary BPA concentration all the participants were used in both treatments.

In the experiment it is seen that there is a large variability in BPA concentration so the matched pair experiment is used to reduce variability of the data.

#### 3.58b

#### 3.58c

It is 95% confident that the difference in the mean urinary BPA concentration after taking canned soup and fresh soup for five days will be between 19.6g/L and 25.5g/L.

#### 3.58d

The sample mean difference will be close to population mean difference if a large sample is taken. This concludes that the confidence interval will be narrower if the study includes 500 participants instead of 75.

#### 3.60a

The relevant parameter is the **population mean** denoted by . This is the mean effect on weight over two years after one month overeating.

#### 3.60b

To find the actual exact value of the population mean, we have to collect all population who overeat for one month. Then after that, we find out the result after two and a half years.

#### 3.60c

Confidence interval =

Therefore, it is 95% confident that the true population mean effect on weight put on over two years after one month of overeating lies between **4.4 and 9.2 pounds**.

#### 3.60d

Margin of error =

Therefore, it would be expected that estimate of the mean 6.8 pounds lies within 2.4 pounds of true mean.

#### 3.64a

The given interpretation of the 95% confidence interval is not correct because 95% confidence interval is for the population mean pulse rate not for the sample mean pulse rate of the students.

#### 3.64b

The given interpretation of the 95% confidence interval is not correct because 95% confidence interval is for the population mean pulse rate not for the sample mean pulse rate of the students

#### 3.64c

The given interpretation of the 95% confidence interval is not in doubt because it is not sure whether or not the interval considers the population mean.

#### 3.64d

The given interpretation of the 95% confidence interval is not correct because 95% confidence interval for the mean pulse rate is not from the US college students.

#### 3.64e

The given interpretation of the 95% confidence interval is not correct because 95% confidence interval for the mean pulse rate is not from the US college students.

#### 3.64f

The given interpretation of the 95% confidence interval is not correct because the population mean is a particular permanent value.

#### 3.64g

The given interpretation of the 95% confidence interval is not correct because 95% confidence interval is for the population mean pulse rate not for the sample mean pulse rate of the students.

#### 3.66a

The sample 79, 79, 97, 85, 88 is the possible bootstrap sample from the original sample since all the values are there in the original sample and also the sample size of the given sample and the original sample is same.

#### 3.66b

The sample 72, 79, 85, 88, 97 is the possible bootstrap sample from the original sample since all the values are there in the original sample and also the sample size of the given sample and the original sample is same.

#### 3.66c

The sample 85, 88, 97, 72 is not the possible bootstrap sample from the original sample since the sample size of the bootstrap sample and the original sample should be same. Here, the sample size of the given sample is 4 and the sample size of the original sample is 5.

#### 3.66d

The sample 88, 97, 81, 78, 85 is not the possible bootstrap sample from the original sample since the original sample does not have the value 78.

#### 3.66e

The sample 97, 85, 79, 85, 97 is the possible bootstrap sample from the original sample since all the values are there in the original sample and also the sample size of the given sample and the original sample is same.

#### 3.66f

The sample 72, 72, 79, 72, 79 is the possible bootstrap sample from the original sample since all the values are there in the original sample and also the sample size of the given sample and the original sample is same.

#### 3.75a

Therefore the best estimate of the proportion of all snails of this type to live after being eaten by a bird is **0.149**.

#### 3.75b

SE = 0.028

#### 3.75c

Confidence interval =

Confidence interval =

Therefore, it is 95% confident that the true proportion of all snails that will live after being eaten by a bird lies between **0.093 and 0.205**.

#### 3.76a

s =

s =

Therefore, the standard deviation of the sample is **14.63**.

#### 3.76b

To create the bootstrap shows the eight slips by eight values and mingle those slips. Take one slip and write the number and put that slip back. Do this step for eight times. The eight numbers is the bootstrap samples. The bootstrap statistic is the mean of the eight numbers.

#### 3.76c

We expect the bootstrap distribution to be **bell shaped** and it’s center will be at **34**.

#### 3.76d

The population parameter of interest is population mean . That is the average number of ants that will climb on a piece of a peanut butter sandwich left on the ground near an ant hill. The best point estimate of the population mean is = 34

#### 3.76e

Confidence interval =

Confidence interval =

#### 3.78\*\*

Therefore, the mean of the sample is**605**.

Confidence interval =

Confidence interval =

Therefore, it is 95% confident that the average monthly sales in the United States lies between **463.81 and 746.19**.

#### 3.81a

The correct option is 0.015. This is because the figure of the bootstrap difference in sample proportions of teen and adult cell phone users who text shows that distribution is centered lying on the value of the difference in population proportion. This means that the estimate of the difference in the population proportion is .

The figure of the bootstrap difference in sample proportions of teen and adult cell phone users who text shows that the central 95% of difference in sample proportions show in a range from 0.12 to 0.18. This range should distance about two standard errors below the mean and two standard errors above the mean.

Therefore, the estimate of the standard error is around

#### 3.81b

Confidence interval =

Confidence interval =

Therefore, it is 95% confident that the difference in proportions of teen and adult cell phone users who text will lie between 0.12 and 0.18.

#### 3.84a

**Mean (Difference)**

mean(~Distance, data=CommuteAtlanta)

## [1] 18.156

**Standard Deviation (Difference)**

sd(~Distance, data=CommuteAtlanta)

## [1] 13.79828

#### 3.84b

The shape is bell shaped that is centered at 18.156.

#### 3.84c

Looking at the upper right part of the graph, the standard error for mean commujte distance for 1000 bootstrap sample is **0.606**.

#### 3.84d

Confidence interval =

Confidence interval =

Therefore, it is 95% confident that the mean of the commute distances in Commute Atlanta lies between **16.948 and 19.372 miles**.

#### 3.89a

The maximum confidence level would be 100%. This means that, keeping 95% of the middle value, the remaining percentage left is (100-95)% = 5%. This 5% should be divided equally into two tails. The value that must be chopped off from each tail for 95% confidence level is 2.5%.The bootstrap distribution contains 1,000 bootstrap samples. That is, 2.5% of 1.000 = 25 of the lowest and highest values should be chopped off to create 95% confidence interval

#### 3.89b

The maximum confidence level would be 100%. This means that, keeping 90% of the middle value, the remaining percentage left is (100-90)% =10%. This 10% should be divided equally into two tails. The value that must be chopped off from each tail for 90% confidence level is 5%. The bootstrap distribution contains 1,000 bootstrap samples. That is, 5% of 1.000 = 50 of the lowest and highest values should be chopped off to create 90% confidence interval.

#### 3.89c

The maximum confidence level would be 100%. This means that, keeping 98% of the middle value, the remaining percentage left is (100-98)% =2%. This 2% should be divided equally into two tails. The value that must be chopped off from each tail for 98% confidence level is 1%. The bootstrap distribution contains 1,000 bootstrap samples. That is, 1% of 1,000 = 10 of the lowest and highest values should be chopped off to create 98% confidence interval.

#### 3.89d

The maximum confidence level would be 100%. This means that, keeping 99% of the middle value, the remaining percentage left is (100-99)% =1%. This 1% should be divided equally into two tails. The value that must be chopped off from each tail for 99% confidence level is 0.5%. The bootstrap distribution contains 1,000 bootstrap samples. That is, 0.5% of 1.000 = 5 of the lowest and highest values should be chopped off to create 99% confidence interval.

#### 3.90

**Option A: “66 to 74” is the most likely result after the change**. This is because the length of the confidence interval increases with the increment of the confidence level.

Therefore, 99% confidence interval should be 66 to 74.

#### 3.91

**Option C: “67.5 to 72.5” is the most likely result after the change**. This is because the length of the confidence interval decreases with the decrease of the confidence level.

Therefore, 90% confidence interval should be 67.5 to 72.5.

#### 3.92

**Option C: “67.5 to 72.5” is the most likely result after the change**. This is because the standard error decreases with the increase of the sample size. Thus, the bootstrap distribution will be less widened, so the confidence interval will be of small width.

Therefore, using the sample size n = 45, the 95% confidence interval should be 67.5 to 72.5

#### 3.93

**Option A: “66 to 74” is the most likely result after the change**. This is because the standard error increases with the decrease of the sample size. Thus, the bootstrap distribution will be more widened, so the confidence interval will be of larger width.

Therefore, using the sample size n= 16, the 95% confidence interval should be 66 to 74.

#### 3.94

**Option B: “67 to 73” is the most likely result**. This is because the width of the confidence interval will not change by adding or subtracting bootstrap samples

#### 3.95

**Option B: “67 to 73” is most likely result**. This is because the width of the confidence interval will not change by adding or subtracting bootstrap samples.

#### 3.100a

From the Figure of the bootstrap distribution of sample means of IQ scores it is clear that the distribution has its center at 100.

Therefore, the mean of the original sample of IQ scores is .

#### 3.100b

The maximum confidence level would be 100%. This means that keeping 99% of the middle value, the remaining percentage left is (100-99)% = 1%. This 1% should be divided equally into two tails. The values to be removed from each tail for 99% confidence level are 0.5%. The bootstrap distribution contains 1,000 bootstrap samples. That is, 0.5% of 1,000 = 5 of the lowest and highest values should be removed to create 99% confidence interval.

Therefore, the 99% confidence interval is **(88,112)**.

#### 3.102

Here for the 90% confidence interval,z valueis 1.645 Margin of error =

New Interval = (0.753 -0.022， 0.753+ 0.022）= **(0.731，0.775）**.

#### 3.105

Confidence interval =

Confidence interval =

Therefore, the 95% confidence interval for the average tip left at the restaurant is **(3.48, 4.23)**.

#### 6.4a

SE =

Therefore the standard error of the distribution of sample proportion is **0.081**.

#### 6.4b

Since the value of np is less than 10 and n(1-p) is greater than 10, the distribution sample proportion isn’t reasonably normal.

#### 6.10a

SE =

Therefore the standard error of the distribution of sample proportion is **0.046**.

#### 6.10b

SE =

Therefore the standard error of the distribution of sample proportion is **0.015**.

#### 6.10c

SE =

Therefore the standard error of the distribution of sample proportion is **0.043**.

#### 6.10d

SE =

Therefore the standard error of the distribution of sample proportion is **0.014**.

#### 6.12

**p = 0.8**

SE =

Therefore the standard error of the distribution of sample proportion is **0.040**.

**p = 0.5**

SE =

Therefore the standard error of the distribution of sample proportion is **0.050**.

**p = 0.3**

SE =

Therefore the standard error of the distribution of sample proportion is **0.046**.

**p = 0.1**

SE =

Therefore the standard error of the distribution of sample proportion is **0.030**.

The largest standard error is **0.050 for p = 0.5** and the smallest standard error is **0.030 for p = 0.1**.

#### 6.16a

#### 6.16b

SE =

Therefore the standard error of the distribution of sample proportion is **0.049**. This shows us that the results obtained from part (a) is similar to the results obtained in part (b).

#### 6.18a

#### 6.18b

Since the value of np is less than 10 and n(1-p) is greater than 10, the distribution sample proportion doesn’t follow normal distribution.

SE =

Therefore the standard error of the distribution of sample proportion is **0.126**. This shows us that the results obtained from part (a) is slightly closer to the results obtained in part (b).