STA 501 Week #5 Assignment

Due: 03/03/2021

This assignment focuses on the applications of sampling distributions. You are expected to study the class note, particularly the examples before attempting the following problems. Please check the assumptions of each of the four results and compare them with the information given in each of the problems in the following to determine the correct sampling distribution.

- **Problem 1**

Suppose it is known that in a certain large human population cranial length is approximately **normally distributed** with a mean of 185.6mm and a standard deviation of 12.7 mm.

(1). Find the mean and standard deviation of the sampling distribution respectively.

Solution: Since the population is normally distributed. Using the property of the normal distribution, we have $\bar{x} \to N(185.6, 12.7/\sqrt{n})$. If the underlying sample size is given, we can uniquely determine the sampling distribution of the sample mean $\bar{x}$.

(2). What is the probability that a random sample of size 10 from this population will have a mean greater than 190?

Solution: We given that the sample $n = 10$. The exact sampling distribution of the sample mean is given by $\bar{x} \to N(185.6, 12.7/\sqrt{10})$. The probability $P(\bar{x} > 190)$ can be found in the following R code.

```
1 - pnorm(190, 185.6, 12.7/sqrt(10))   # the right-tail area = 1 - left-tail area
```

```
## [1] 0.1366286
```

Therefore, $P(\bar{x} > 190) \approx 13.66$.

- **Problem 2**

The National Health and Nutrition Examination Survey of 1988–1994 (NHANES III) estimated the mean serum cholesterol level for U.S. females aged 20–74 years to be 204 mg/dl. The estimate of the standard deviation was approximately 44. Using these estimates as the mean μ and standard deviation σ for the U.S. population, consider the sampling distribution of the sample mean based on samples of size 49 drawn from women in this age group.

(1). Find the mean and the standard deviation of the sampling distribution respectively

Solution: Since the sample size $n = 49 > 30$, we can use the Central Limit Theorem to specify the sampling distribution of the sample mean $\bar{x}$ in the following

$$\bar{x} \to N(204, 44/\sqrt{49}) = N(204, 44/7)$$

The next few questions will be based on the above distribution.

(2). Find the probability that the sample mean serum cholesterol level will be between 170 and 195.

Solution: We want to find $P(170 < \bar{x} < 195) = P(\bar{x} < 195) - P(\bar{x} < 170)$, that is, the difference between the two left-tail areas. The probability can be found in the following code.

```
pnorm(195, mean=204, sd = 44/7) - pnorm(170, mean=204, sd = 44/7)
```

```
## [1] 0.0760979
```

Therefore, the probability that the sample mean serum cholesterol level will be between 170 and 195 is about 7.61%.

(3). Find the probability that the sample mean serum cholesterol level will be less than 210.

Solution:

(4). Find the probability that the sample mean serum cholesterol level will be bigger than 195.

- **Problem 3**

Smith et al. performed a retrospective analysis of data on 782 eligible patients admitted with myocardial infarction to a 46-bed cardiac service facility. Of these patients, 248 (32 percent) reported a past myocardial infarction. Use .32 as the population proportion. Suppose 50 subjects are chosen at random from the population.

(1). Find the mean and the standard deviation of the sampling distribution respectively.

(2). What is the probability that over 40 percent would report previous myocardial infarctions?

Solution