Statistics and Data Analysis
Homework 4: Sampling Distributions and Estimations

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. The 1000 bags of candies produced by me follow $\text{ND}(2, \sigma)$. My mom randomly draws n bags and calculate the sample mean $\bar{x}$. If $\bar{x} > 2.2$ or $\bar{x} < 1.8$, I will get punished. I want to know the probability for me to get punished under the following conditions:

   (a) $\sigma = 0.5$ and $n = 5$?

   (b) $\sigma = 0.1$ and $n = 5$?

   (c) $\sigma = 0.5$ and $n = 3$?

   (d) $\sigma = 0.1$ and $n = 3$?

2. A laptop manufacturer produces a type of laptop starting from 2013. At that time, the average battery life is 5 hours. Last month, a new technology was introduced, and the average battery life was expected to increase. The manufacturer wants to know whether the new technology really brought a significant improvement and make the current average battery life $\mu > 6$. For simplicity, let's assume that the current standard deviation is known to be 0.6 hour.[1]

   (a) Suppose we randomly draws 9 laptops and obtains a sample mean $\bar{x}$. What is the sampling distribution of $\overline{X}$?

   (b) If μ is indeed 6, how likely will we get a sample mean that is below 5.7?

   (c) If μ is indeed 6, how likely will we get a sample mean that is below 7.4?

   (d) If μ is indeed 6, find a such that $\Pr(\overline{X} > a) = 0.05$.

   Note. The value a found in this problem provides a decision rule to us: If we obtain $\bar{x} > a$, we may conclude with a 95% confidence level that μ is strictly greater than 6. This is simply because "if μ is still 6 (or smaller than 6), it will be quite unlikely (only with a 5% probability) to see $\bar{x} > a$; but we see $\bar{x} > a$, so it is reasonable to believe that $\mu > 6$ (and the probability for this belief to be wrong is less than 5%)." This is a direct application of sampling distributions. We will investigate this kind of statistical testing more after the midterm exam.

3. Qualitatively answer the following true-and-false questions regarding interval estimation for the population mean:

   (a) After we choose the sample size and confidence level but before we sample from the population, the confidence interval that we will construct after sampling is random.

   (b) Suppose the population variance is unknown. For a given confidence level, increasing the sample size enlarges the confidence interval.

   (c) Suppose the population variance is known. For a given leg length, increasing the confidence level requires a larger sample size.

   (d) Suppose the population follows a normal distribution, the sample size is small, and the population variance is known. In this case, we cannot use the z distribution to do estimation.

   (e) Suppose the population follows a non-normal distribution, the sample size is large, and the population variance is unknown. In this case, we can use the t distribution to do estimation.

   Note. This kind of qualitative problems may appear in the midterm exam with some modifications. However, no problem will ask you to really construct a confidence interval.

[1] Note that even if we do not know the population standard deviation, we may still apply the t distribution to estimate μ and conduct other statistical studies. The assumption of known population standard deviation is simply for educational purpose.