Central Limit Theorem

- **Parameter.** A *parameter* is a numerical descriptive measure of a population. *This quantity is usually unknown but it is constant at a given time.*

  Examples: $\mu, \sigma^2, \sigma, \rho$.

- **Statistic.** A *sample statistic* is a numerical descriptive measure of a sample. *Its value is calculated from the observations in the sample.*

  Examples: $\bar{x}, s^2, s, \hat{p}$.

- Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution (or population) with mean μ and variance σ^2. Then

  1. $E(\bar{X}) = \mu_{\bar{X}} = \mu$

  2. $V(\bar{X}) = \sigma^2_{\bar{X}} = \sigma^2/n$

- **Sampling Distribution.** The *sampling distribution* of a sample statistic calculated from a sample of n measurements is the probability distribution of the statistic.

- **Sampling Distribution of $\bar{X}$.** If a population has the $N(\mu, \sigma)$ distribution, then the sample mean $\bar{X}$ of n independent observations has the $N(\mu, \sigma/\sqrt{n})$ distribution.That is,

$$\bar{X} \sim N(\mu, \sigma/\sqrt{n}). \tag{1}$$

- **Central Limit Theorem:** Draw an **SRS** of size n from any population with mean μ and finite standard deviation σ. When n is large (*rule of thumb*: $n \geq 30$), the sampling distribution of the sample mean $\bar{X}$ is *approximately* normal with mean μ and standard deviation $\sigma/\sqrt{n}$. That is,

$$\bar{X} \approx N(\mu, \sigma/\sqrt{n}). \tag{2}$$

- **Practice:**

  1. A soft-drink machine is regulated so that it discharges an average of 200 ml per cup. If the amount of drink discharged is normally distributed with a standard deviation equal to 10 ml, what is the probability that a cup will contain less than 188 ml?

  2. In the previous problem, suppose we sample 25 cups from this machine and take their average content. What is the probability that the average content of these 25 cups is less than 188 ml?

3. In the previous problem, if all the contents of the sample of 25 cups are combined in a big container, what is the probability that the combined amount is at least 5100 ml?

4. The average time it takes a senior high school student to complete a certain achievement test is 46.2 minutes with standard deviation of 8 minutes. If a random sample of 100 senior high school students who took the test was selected, find the probability that the mean time it takes the group to complete the test will be

a. less than 44 minutes.

b. more than 48 minutes.

c. between 44 to 48 minutes.

- **Point Estimator.** A *point estimator* of a population parameter is a rule or formula that tells us how to use the sample data to calculate a single number that can be used as an *estimate* of the population parameter.
- **Unbiased Estimator.** If the sampling distribution of a sample statistic has a mean equal to the population parameter the statistic is intended to estimate, the statistic is said to be an *unbiased estimator* of the parameter. Otherwise, the statistic is said to be *biased estimator* of the parameter.
- **Homework problems:**
  Section 5.4: pp. 229-230; # 47, 49, 53, 55.