Central Limit Theorem

- For a randomly selected sample of size n with a mean μ and a standard deviation σ, the following is true:
  1. The distribution of sample means $\bar{X}$ is approximately normal, regardless of the population distribution. For markedly non-normal populations, the approximation is good enough when the sample size is 25 or more.
  2. The mean of the distribution of sample means is equal to the mean of the population distribution, $\mu_{\bar{x}} = \mu$.
  3. The standard deviation of the distribution of sample means is equal to: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

Examples

- Given a normally distributed population with $\mu_X = 70$ and $\sigma_X = 20$; that is $X \sim N(70, 20)$
- Assume that we take a random sample of size $n = 25$
- Use with examples 1 - 4

Example 1

- What is the probability of obtaining a random sample with a mean of 80 or higher?

$$Pr(X \geq 80 | X \sim N(70, 20))$$

$$Pr(\bar{X} \geq 80 | \bar{X} \sim N(70, 4))$$

Example 2

- What is the probability of obtaining a random sample with a mean that differs from the population mean by more than 10 points?

$$Pr(X \geq 80 or X \leq 60 | X \sim N(70, 20))$$

$$Pr(\bar{X} \geq 80 or \bar{X} \leq 60 | \bar{X} \sim N(70, 4))$$

$$Pr(|\bar{X} - \mu_X| \geq 10 | \bar{X} \sim N(70, 4))$$

Example 3

- What sample mean has a value such that the probability of obtaining one at least that high in random sampling is .05?
- Find $\bar{X}_o$ such that:

$$Pr(\bar{X} \geq \bar{X}_o | \bar{X} \sim N(70, 4)) = 0.05$$

Example 4

- Within what limits would the central 95% of the sample means fall?
- Find $\bar{X}_1$ and $\bar{X}_2$ such that:

$$Pr(\bar{X}_1 \leq \bar{X} \leq \bar{X}_2 | \bar{X} \sim N(70, 4)) = 0.95$$

Example 5

Suppose we are interested in knowing on average how early students show up prior to kickoff for a home UI football game. We know the population has $\mu = 15$, $\sigma = 20$, and the population is severely positively skewed. Use this information to answer the following questions:

- We take a random sample of 10 students and get a sample mean of $\bar{X} = 45$. Can we safely use the central limit theorem here?
- We take a random sample of 50 students where $\bar{X} = 19$. What is the probability of getting a sample mean greater or equal to this value? How many students would we expect to show up 19 minutes or earlier (assume 10000 students attend the game).
- What are the limits in which the central 95% of the sample means will fall when $n = 25$ and when $n = 100$? How do these compare? Why is this pattern occurring?
- Given a random sample of 25 students ($n = 25$), how likely is it for those 25 students to on average show up at kickoff or after?