Chapter 18II

Sampling Distribution for the Sample Mean

p and μ

- p is the population proportion
  - summarizes a categorical variable
    - Ex. What proportion of ISU students smoke?
- take a sample and get $\hat{p}$
- take many samples and get many different $\hat{p}$'s
- the distribution of these $\hat{p}$'s is

$$N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

p and μ

- μ is the population mean
  - summarizes a quantitative variable
    - Ex. What is the mean age of all STAT 101 students?
- take a sample and get $\bar{y}$, the sample mean
- Value of $\bar{y}$ is random
- Changes from sample to sample
- Different from population mean μ

$\bar{y}$

- take many samples of size n and get many different $\bar{y}$'s
- these $\bar{y}$'s are data
- summarize data
  - Shape, Center and Spread
- sampling distribution for $\bar{y}$

Sampling Distribution for $\bar{y}$

- Mean (Center)

$$\mu(\bar{y}) = \mu$$

- Expect to get on average μ
- $\bar{y}$ is unbiased for μ

Sampling Distribution for

- Standard Deviation (Spread) $\bar{y}$

$$\sigma(\bar{y}) = \frac{\sigma}{\sqrt{n}}$$

- As *n* gets larger, $\sigma(\bar{y})$ gets smaller.
- Larger samples are more accurate than smaller samples

Name that distribution

- sampling distribution for $\overline{y}$ is **NORMAL!!!!!!!!**
- only if these three conditions hold:
  1. Sample must be random sample
  2. Sample must be independent values
  3. Sample must be less than 10% of population

  (4. n Is large "enough")

Central Limit Theorem

- As the sample size *n* increases, the mean of n independent values has a sampling distribution that tends toward a normal distribution.

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

How large does *n* need to be?

- Depends on shape of population distribution
  - Symmetric, *n* between 5 and 15.
  - Skewed, *n* at least 30.

Example #1

- Ithaca, New York, gets an average of 35.4 inches of rainfall per year with a standard deviation of 4.2 inches. Assume yearly rainfall follows a normal distribution.

Example #1 – cont.

- What is the probability a single year will have more than 40 inches of rain?
- Y = annual rainfall.
- Y is N(35.4, 4.2)
- P(Y > 40)

Example #1 – cont.

$$P(Y>40)=P\left(Z>\frac{40-35.4}{4.2}\right)$$

$$=P(Z>1.10)$$

$$=1-0.8643$$

$$=0.1357$$

Example #1 – cont.

– What is the probability that over a four year period the mean rainfall will be less than 30 inches?

Example #1 – cont.

- $\bar{Y}$ = mean rainfall over four year period
- Since Y = annual rainfall is N(35.4, 4.2)
- $\bar{Y}$ is

$$N(\mu, \frac{\sigma}{\sqrt{n}}) = N(35.4, \frac{4.2}{\sqrt{4}}) = N(35.4, 2.1)$$

Example #1 – cont.

$$P(\bar{Y} < 30) = P\left(Z < \frac{30 - 35.4}{2.1} \right)$$
$$= P(Z < -2.57)$$
$$= 0.0051$$

Example #2

- Carbon monoxide emissions for a certain kind of car vary with mean 2.9 gm/mi and standard deviation 0.4 gm/mi. A company has 80 cars in its fleet. Estimate the probability that the mean emissions for the fleet is between 2.95 and 3.0 gm/mi.

Example #2 – cont.

- Y = emissions from one car
- $\bar{Y}$ = mean emission from 80 cars
- n = 80 is large
- $\bar{Y}$ is

$$N(\mu, \frac{\sigma}{\sqrt{n}}) = N(2.9, \frac{0.4}{\sqrt{80}}) = N(2.9, 0.045)$$

Example #2 – cont.

$$P(2.95 < \bar{Y} < 3.0) = P\left(\frac{2.95 - 2.9}{0.045} < Z < \frac{3.0 - 2.9}{0.045} \right)$$
$$= P(1.11 < Z < 2.22)$$
$$= 0.9868 - 0.8665$$
$$= 0.1203$$

Example #3

- Grocery store receipts show that customer purchases are skewed to the right with a mean of $32 and a standard deviation of $20.

Example #3 – cont.

- Can you determine the probability the next customer will spend at least $40?
- Y = amount a single customer will spend.
- Y has a skewed distribution
- Given info, we cannot determine this probability.

Example #3 – cont.

- What is the probability the next 50 customers will spend an average of at least $40?
- Y = amount one customer will spend.
- $\overline{Y}$ = mean amount 50 customers will spend.
- n = 50 is large so
- $\overline{Y}$ is

$$N(\mu, \frac{\sigma}{\sqrt{n}}) = N(32, \frac{20}{\sqrt{50}}) = N(32, 2.83)$$

Example #3 – cont.

$$P(\overline{Y} > 40) = P\left(Z > \frac{40-32}{2.83}\right)$$

$$= P(Z > 2.83)$$

$$= 1 - 0.9977$$

$$= 0.0023$$

Example #4

- Suppose there were 312 customers at the grocery store in one day. What is the probability the store's revenues were at least $10,000?

Example #4 – cont.

- Total revenues = total amount spent by all 312 customers
- All 312 customers must spend over $10,000
- On average, each customer must spend $10,000/312 = $32.05

Example #3 – cont.

- What is the probability that 312 customers will spend at least an average of $32.05.
- Y = amount one customer will spend.
- $\overline{Y}$ = mean amount 312 customers will spend.
- n = 312 is large so
- $\overline{Y}$ is

$$N(\mu, \frac{\sigma}{\sqrt{n}}) = N(32, \frac{20}{\sqrt{312}}) = N(32, 1.13)$$

Example #4 – cont.

$$P(\overline{Y} > 32.05) = P\left(Z > \frac{32.05 - 32}{1.13}\right)$$

$$= P(Z > 0.04)$$

$$= 1 - 0.5160$$

$$= 0.4840$$