Sample Means

Section 9.3

Do You Really "Mean" That?

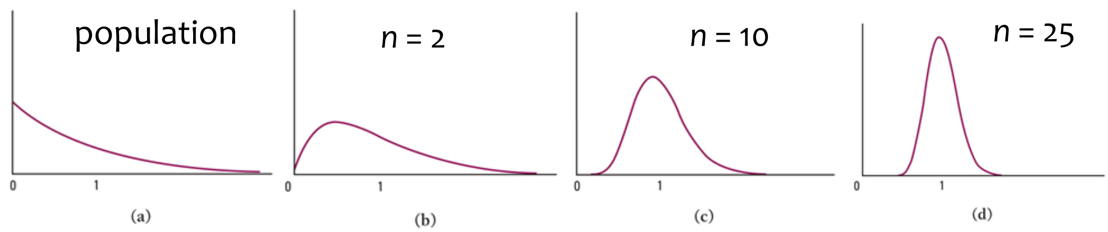
- We use proportions in statistics most often when we are interested in categorical variables.
- When we record quantitative data, we are interested in the mean or median instead.

Central Limit Theorem

Draw an SRS of size n from any population whatsoever with mean μ and finite standard deviation σ. When n is large, the sampling distribution of the sample mean $\bar{x}$ is close to the normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$

$\mu_{\bar{x}} = \mu$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ n is large if $n \geq 30$

The CLT in Action!



The distribution of total points earned by the students taking Calculus I at a large university is slightly skewed left with mean 625 $=\mu$ and standard deviation 44.5. If a random sample of 100 $=n$ students is taken, which statement best describes the sampling distribution of the sample mean?

A. normal with mean 625 $\mu_{\bar{x}}$ and standard deviation 44.5

B. normal with mean 625 $=\mu_{\bar{x}}$ and standard deviation 4.45

~~C.~~ shape unknown with mean 625 and standard deviation 44.5

~~D.~~ shape unknown with mean 625 and standard deviation 4.45

~~E.~~ No conclusion can be drawn because the population is not normally distributed.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{44.5}{\sqrt{100}}$$

A Tall Tale…

- Let's say that young women's heights are ~N (64.5, 2.5). So what is the probability that a randomly selected young woman is taller than 66.5 inches?

$$P(x > 66.5) = P\left(z > \frac{66.5-64.5}{2.5}\right)$$
$$P(z > 0.8) \approx .212$$

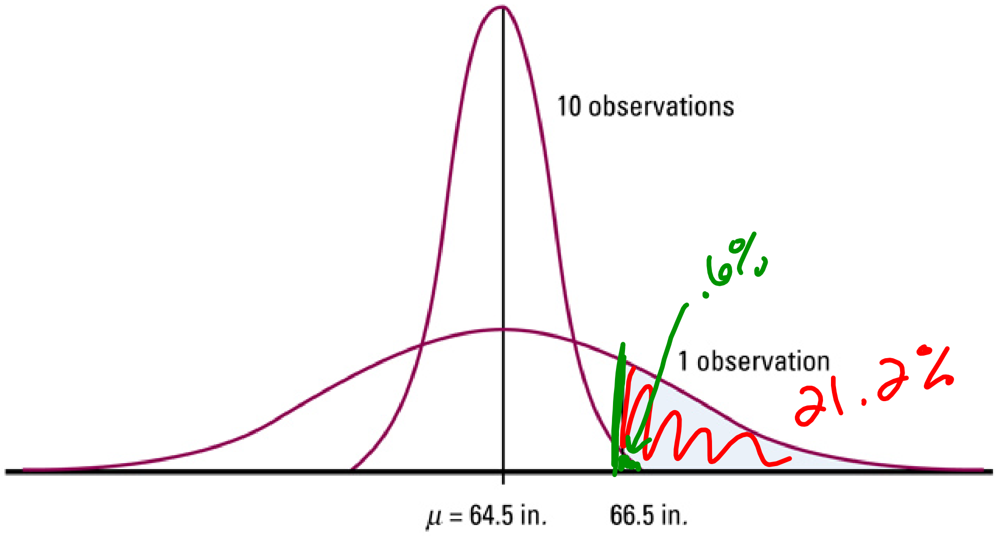
- What is the probability that the mean height of an SRS of 10 young women is greater than 66.5 inches?

The sampling distribution of $\bar{x}$ is ~N b/c the population is ~N and it has $\mu_{\bar{x}} = 64.5$ and $\sigma_{\bar{x}} = \frac{2.5}{\sqrt{10}}$.

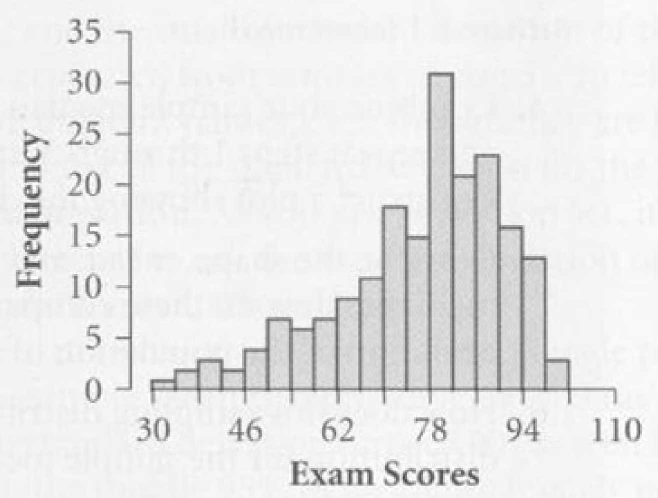
$$P(\bar{x} > 66.5) = P\left(z > \frac{66.5-64.5}{\frac{2.5}{\sqrt{10}}}\right) \approx .006$$

— OR —

$$P\left(z > \frac{66.5-64.5}{.791}\right) = .006$$



The histogram in Display 5.9 shows the distribution of exam scores for 192 students in an introductory statistics course at the University of Florida.



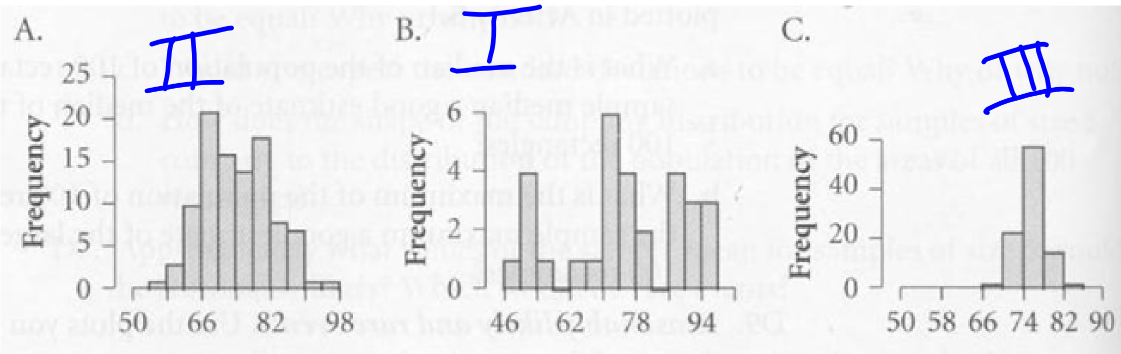
Display 5.9 A distribution of exam scores.

Match the histograms A, B, and C in Display 5.10 to their descriptions, I, II, or III.

I. The individual scores for one random sample of 30 students.

II. A simulated sampling distribution of the mean of the scores of 100 random samples of 4 students.

III. A simulated sampling distribution of the mean of the scores of 100 random samples of 30 students.



Display 5.10 Histograms for exam scores.

After repeated observations, it has been determined that the waiting time at the drive-thru window of a local bank is skewed left, with a mean of 3.5 minutes and a standard deviation of 1.9 minutes. A random sample of 100 $=n$ customers is to be taken. What is the probability that the mean of the sample will exceed 4 minutes?

The sampling distribution is ~N with

$\mu_{\bar{x}} = 3.5$ and $\sigma_{\bar{x}} = \frac{1.9}{\sqrt{100}}$ b/c $n \geq 30$.

$100 \geq 30$

$$P(\bar{x} > 4) = P\left(z > \frac{4-3.5}{.19}\right) \approx P(z > 2.632)$$

$$\approx .00425$$

The distribution of the number of television sets per household in the United States is approximately normal with mean 2.37 and standard deviation 1.16. What is the interval of reasonably likely sample means for samples of size 400?

invNorm(.025) ≈ 1.96

$$-1.96 = \frac{\bar{x} - 2.37}{\frac{1.16}{\sqrt{400}}} \approx 2.256$$

$$1.96 = \frac{x - 2.37}{\frac{1.16}{20}} \approx 2.484$$

btwn 2.256 and 2.484 TVs

Assignment:
pg. 595 #31-35, 37, 39