


Chapter 7: Sampling Distributions

Section 7.3
Sample Means

The Practice of Statistics, 4th edition – For AP*
STARNES, YATES, MOORE



Section 7.3
Sample Means

Learning Objectives

After this section, you should be able to...

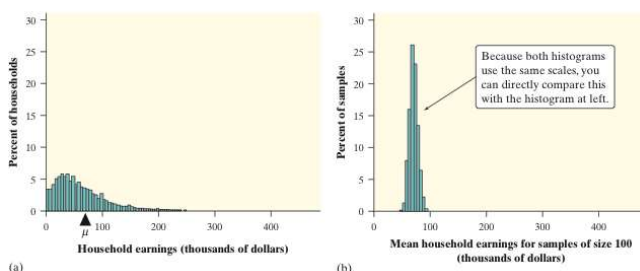
- ✓ FIND the mean and standard deviation of the sampling distribution of a sample mean
- ✓ CALCULATE probabilities involving a sample mean when the population distribution is Normal
- ✓ EXPLAIN how the shape of the sampling distribution of sample means is related to the shape of the population distribution
- ✓ APPLY the central limit theorem to help find probabilities involving a sample mean

■ Sample Means

+ Sample Means

Sample proportions arise most often when we are interested in categorical variables. When we record quantitative variables we are interested in other statistics such as the median or mean or standard deviation of the variable. Sample means are among the most common statistics.

Consider the mean household earnings for samples of size 100. Compare the population distribution on the left with the sampling distribution on the right. What do you notice about the shape, center, and spread of each?



■ The Sampling Distribution of \bar{x}

+ Sample Means

When we choose many SRSs from a population, the sampling distribution of the sample mean is centered at the population mean μ and is less spread out than the population distribution. Here are the facts.

Mean and Standard Deviation of the Sampling Distribution of Sample Means

Suppose that \bar{x} is the mean of an SRS of size n drawn from a large population with mean μ and standard deviation σ . Then:

The **mean** of the sampling distribution of \bar{x} is $\mu_{\bar{x}} = \mu$

The **standard deviation** of the sampling distribution of \bar{x} is

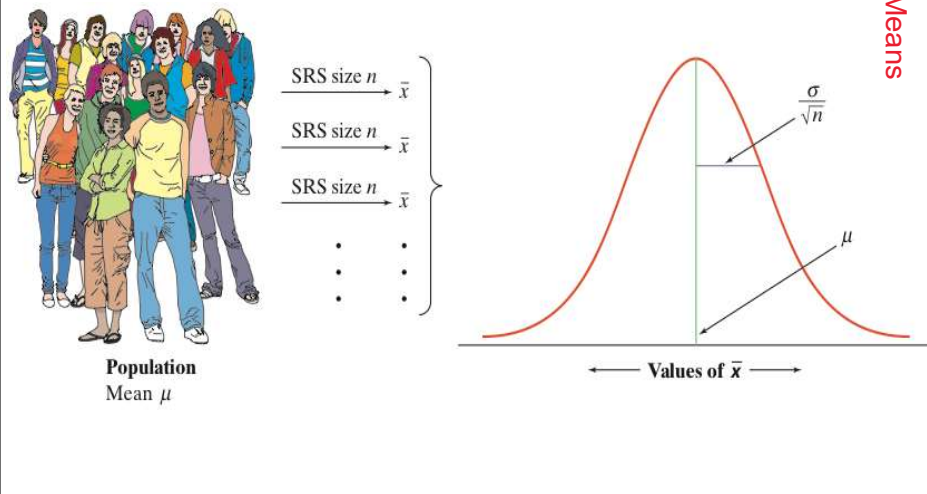
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

as long as the *10% condition* is satisfied: $n \leq (1/10)N$.

Note: These facts about the mean and standard deviation of \bar{x} are true *no matter what shape the population distribution has.*

■ The Sampling Distribution of \bar{x}

When we choose many SRSs from a population, the sampling distribution of the sample mean is centered at the population mean μ and is less spread out than the population distribution. Here are the facts.



■ Sampling from a Normal Population

We have described the mean and standard deviation of the sampling distribution of the sample mean \bar{x} but not its shape. That's because the shape of the distribution of \bar{x} depends on the shape of the population distribution.

In one important case, there is a simple relationship between the two distributions. If the population distribution is Normal, then so is the sampling distribution of \bar{x} . *This is true no matter what the sample size is.*

Sampling Distribution of a Sample Mean from a Normal Population

Suppose that a population is Normally distributed with mean μ and standard deviation σ . Then the sampling distribution of \bar{x} has the Normal distribution with mean μ and standard deviation σ/\sqrt{n} , provided that the 10% condition is met.

Example: Young Women's Heights

The height of young women follows a Normal distribution with mean $\mu = 64.5$ inches and standard deviation $\sigma = 2.5$ inches.

Find the probability that a randomly selected young woman is taller than 66.5 inches.

Let X = the height of a randomly selected young woman. X is $N(64.5, 2.5)$

$$z = \frac{66.5 - 64.5}{2.5} = 0.80 \quad P(X > 66.5) = P(Z > 0.80) = 1 - 0.7881 = 0.2119$$

The probability of choosing a young woman at random whose height exceeds 66.5 inches is about 0.21.

Find the probability that the mean height of an SRS of 10 young women exceeds 66.5 inches.

For an SRS of 10 young women, the sampling distribution of their sample mean height will have a mean and standard deviation

$$\mu_{\bar{x}} = \mu = 64.5 \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.5}{\sqrt{10}} = 0.79$$

Since the population distribution is Normal, the sampling distribution will follow an $N(64.5, 0.79)$ distribution.

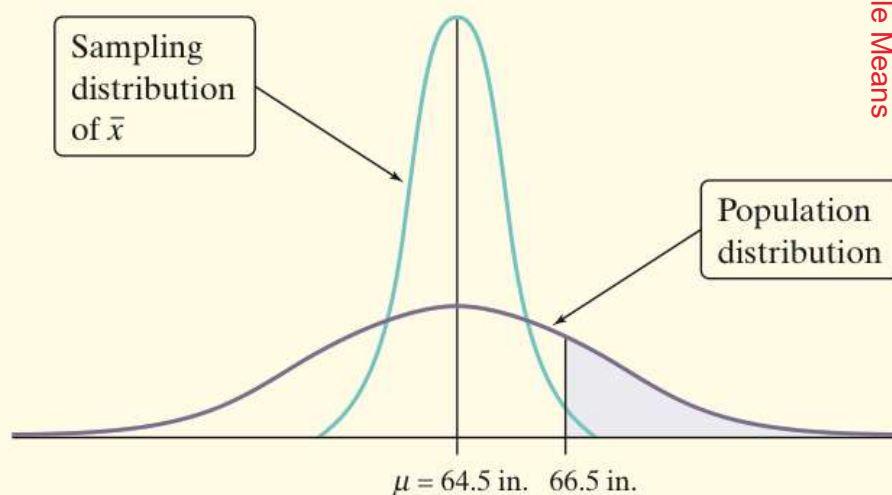
$$z = \frac{66.5 - 64.5}{0.79} = 2.53 \quad P(\bar{x} > 66.5) = P(Z > 2.53) = 1 - 0.9943 = 0.0057$$

It is very unlikely (less than a 1% chance) that we would choose an SRS of 10 young women whose average height exceeds 66.5 inches.

Sample Means

Example: Young Women's Heights

The height of young women follows a Normal distribution with mean $\mu = 64.5$ inches and standard deviation $\sigma = 2.5$ inches.



Sample Means

■ The Central Limit Theorem

+ Sample Means

Most population distributions are not Normal. What is the shape of the sampling distribution of sample means when the population distribution isn't Normal?

It is a remarkable fact that as the sample size increases, the distribution of sample means changes its shape: it looks less like that of the population and more like a Normal distribution! When the sample is large enough, the distribution of sample means is very close to Normal, *no matter what shape the population distribution has*, as long as the population has a finite standard deviation.

Definition:

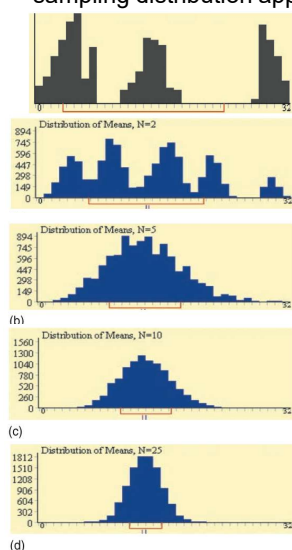
Draw an SRS of size n from any population with mean μ and finite standard deviation σ . The **central limit theorem (CLT)** says that when n is large, the sampling distribution of the sample mean \bar{x} is approximately Normal.

Note: How large a sample size n is needed for the sampling distribution to be close to Normal depends on the shape of the population distribution. More observations are required if the population distribution is far from Normal.

■ The Central Limit Theorem

+ Sample Means

Consider the strange population distribution from the Rice University sampling distribution applet.



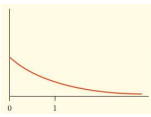
Describe the shape of the sampling distributions as n increases. What do you notice?

Normal Condition for Sample Means

If the population distribution is Normal, then so is the sampling distribution of \bar{x} . This is true no matter what the sample size n is.

If the population distribution is not Normal, the central limit theorem tells us that the sampling distribution of \bar{x} will be approximately Normal in most cases if $n \geq 30$.

Example: Servicing Air Conditioners



Based on service records from the past year, the time (in hours) that a technician requires to complete preventative maintenance on an air conditioner follows the distribution that is strongly right-skewed, and whose most likely outcomes are close to 0. The mean time is $\mu = 1$ hour and the standard deviation is $\sigma = 1$

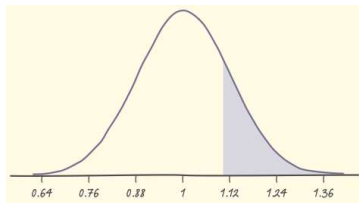
Sample Means

Your company will service an SRS of 70 air conditioners. You have budgeted 1.1 hours per unit. Will this be enough?

Since the 10% condition is met (there are more than $10(70)=700$ air conditioners in the population), the sampling distribution of the mean time spent working on the 70 units has

$$\mu_{\bar{x}} = \mu = 1 \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{70}} = 0.12$$

The sampling distribution of the mean time spent working is approximately $N(1, 0.12)$ since $n = 70 \geq 30$.



We need to find $P(\text{mean time} > 1.1 \text{ hours})$

$$z = \frac{1.1 - 1}{0.12} = 0.83 \quad P(\bar{x} > 1.1) = P(Z > 0.83) \\ = 1 - 0.7967 = 0.2033$$

If you budget 1.1 hours per unit, there is a 20% chance the technicians will not complete the work within the budgeted time.

+ Section 7.3 Sample Means

Summary

In this section, we learned that...

- When we want information about the population mean μ for some variable,
 - ✓ we often take an SRS and use the sample mean \bar{x} to estimate the unknown parameter μ . The **sampling distribution** of \bar{x} describes how the statistic varies in all possible samples of the same size from the population.
 - ✓ The mean of the sampling distribution is μ , so that \bar{x} is an unbiased estimator of μ .
 - ✓ The standard deviation of the sampling distribution of \bar{x} is σ/\sqrt{n} for an SRS of size n if the population has standard deviation σ . This formula can be used if the population is at least 10 times as large as the sample (10% condition).

+ Section 7.3 Sample Means

Summary

In this section, we learned that...

- ✓ Choose an SRS of size n from a population with mean μ and standard deviation σ . If the population is Normal, then so is the sampling distribution of the sample mean \bar{x} . If the population distribution is not Normal the **central limit theorem (CLT)** states that when n is large, the sampling distribution of \bar{x} is approximately Normal.
- ✓ We can use a Normal distribution to calculate approximate probabilities for events involving \bar{x} whenever the Normal condition is met:

If the population distribution is Normal, so is the sampling distribution of \bar{x} .

If $n \geq 30$, the CLT tells us that the sampling distribution of \bar{x} will be approximately Normal in most cases.