Sullivan Section 7.5

Gene Quinn

The Sampling Distributions

The bell-shaped curve of the normal distribution is completely characterized by two values known as **parameters**:

- The mean μ

The Sampling Distributions

The bell-shaped curve of the normal distribution is completely characterized by two values known as **parameters**:

- The mean μ
- The standard deviation σ

The Sampling Distributions

Recall that a **random sample of size** n consists of n individuals chosen in a way that makes every random sample of size n equally likely.

The Sampling Distributions

Recall that a **random sample of size** n consists of n individuals chosen in a way that makes every random sample of size n equally likely.

Usually, a sampling scheme that consists of n draws without replacement in which each member of the population has an equal chance of being chosen at each draw will make all possible random samples equally likely.

The Sampling Distribution of the Mean

When a random sample of size n is drawn from a population, we obtain n values which are considered random variables.

The Sampling Distribution of the Mean

When a random sample of size n is drawn from a population, we obtain n values which are considered random variables.

Suppose the underlying population has mean μ and standard deviation σ.

The **sample mean** $\overline{x}$ for our random sample of size n is calculated as the sum of the n values in the random sample, divided by n:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

The Sampling Distribution of the Mean

When a random sample of size n is drawn from a population, we obtain n values which are considered random variables.

Suppose the underlying population has mean μ and standard deviation σ.

The **sample mean** $\overline{x}$ for our random sample of size n is calculated as the sum of the n values in the random sample, divided by n:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

When we add random variables or multiply them by constants, the result is a new random variable.

The Sampling Distribution of the Mean

So, the sample mean

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

can be considered a new random variable.

The Sampling Distribution of the Mean

So, the sample mean

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

can be considered a new random variable.

As such, it is associated with a probability distribution, which is known as the **sampling distribution of the sample mean**.

The Sampling Distribution of the Mean

So, the sample mean

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

can be considered a new random variable.

As such, it is associated with a probability distribution, which is known as the **sampling distribution of the sample mean**.

It is an important fact that the sampling distribution of the sample mean nearly always differs from the probability distribution of the population from which the sample is drawn.

The Sampling Distribution of the Mean

If a simple random sample of size n is drawn from a population with :

- mean=μ
- standard deviation=σ

then the **sampling distribution** of $\overline{x}$ has:

$$\text{mean } \mu_{\overline{x}} = \mu$$

and

$$\text{standard deviation } \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

The Sampling Distribution of the Mean

If a simple random sample of size n is drawn from a population with :

- mean=μ
- standard deviation=σ

then the **sampling distribution** of $\overline{x}$ has:

$$\text{mean } \mu_{\overline{x}} = \mu$$

and

$$\text{standard deviation } \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

The quantity $\sigma_{\overline{x}}$ is called the **standard error of the mean**.

The Sampling Normal Populations

If a simple random sample of size n is drawn from a **normal** population with :

- mean=μ
- standard deviation=σ

then the **sampling distribution** of $\overline{x}$ is a **normal distribution** with:

$$\text{mean } \mu_{\overline{x}} = \mu$$

and

$$\text{standard deviation } \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

The Sampling Normal Populations

This is a stronger conclusion than the previous slide, where we only stated what the mean and standard deviation of the sampling distribution are, without specifying form of the sampling distribution.

The Central Limit Theorem

If a simple random sample of size n is drawn from a population with :

- mean=μ
- standard deviation=σ

then as the sample size n increases, the **sampling distribution** of $\overline{x}$ becomes **approximately normal** with:

$$\text{mean } \mu_{\overline{x}} = \mu$$

and

$$\text{standard deviation } \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

The Central Limit Theorem

This is actually a special case of a more general result which states that sums of independent random variables tend to a normal distribution as the number of independent variables in the sum increases (under some very mild assumptions).

The Law of Large Numbers

As the size n of a simple random sample increases, the difference

$$\overline{x} - \mu$$

approaches zero.

The Normal Approximation to Binomia

Recall that a **binomial experiment** consists of a fixed number n of independent trials, each of which has two possible outcomes.

The Normal Approximation to Binomia

Recall that a **binomial experiment** consists of a fixed number n of independent trials, each of which has two possible outcomes.

If we call the two outcomes success and failure, and each trial has probability of success p, then we characterize the distribution as:

$$\text{binomial}(n, p)$$

The Normal Approximation to Binomia

Recall that a **binomial experiment** consists of a fixed number n of independent trials, each of which has two possible outcomes.

If we call the two outcomes success and failure, and each trial has probability of success p, then we characterize the distribution as:

$$\text{binomial}(n, p)$$

When n is large, accurate computation of binomial probabilities becomes difficult.

The Normal Approximation to Binomia

Recall that a **binomial experiment** consists of a fixed number n of independent trials, each of which has two possible outcomes.

If we call the two outcomes success and failure, and each trial has probability of success p, then we characterize the distribution as:

$$\text{binomial}(n, p)$$

When n is large, accurate computation of binomial probabilities becomes difficult.

In this situation, the binomial distribution can be *approximated* by a normal distribution.

The Normal Approximation to Binomia

Suppose a random variable is binomial with parameters (n, p)

The Normal Approximation to Binomia

Suppose a random variable is binomial with parameters (n, p)

When n is large, the distribution is approximately normal with:

$$\text{mean} = np$$

and

$$\text{standard deviation} = \sqrt{np(1-p)}$$