Sri Lanka Institute of Information Technology

Faculty of Computing

IT2120 - Probability and Statistics

Ms. K. G. M. Lakmali

Year 02 and Semester 01

Lecture 7

SAMPLING DISTRIBUTIONS

Simple Random Sampling (SRS)

- This is starting point for all probability sampling designs.
- Each unit in the population (N) has **same chance** of being selected to the sample (n).
- Sampling may be done with or without replacement (SRSWR or SRSWOR).
- Generally, SRSWOR yields more precise results and is operationally more convenient.
- SRS is a set of **independent and identically distributed (iid)** observable random variables.

Statistic

A function of observable r.v.s that does not depend on any unknown parameters is called a statistic.

Example: Sample Mean

Sampling Distributions

The probability distribution of a statistic is known as a **sampling distribution**.

Sampling Distribution of the Mean

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a population with mean μ and variance σ^2. Then,

$$E(\overline{X}) = \mu$$
$$V(\overline{X}) = \sigma^2/n$$

The standard deviation of a sampling distribution of a statistic is called the **Standard Error (SE)**.
Although the r.v.s are identically distributed, a specific distribution type is not needed.

Sampling Distribution of the Mean

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Normal population with mean μ and variance σ^2. Then,

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Examples

1) A particular brand of drink has an average of 12 ounces per can. As a result of randomness, there will be small variations in how much liquid each bottle really contains. It has been observed that the amount of liquid in these bottles is normally distributed with $\sigma = 0.8$ ounce. A sample of 10 bottles of this brand of soda is randomly selected from a large lot of bottles, and the amount of liquid, in ounces, is measured in each. Find the probability that the sample mean will be within 0.5 ounce of 12 ounces.

Examples

2) A company that manufactures cars claims that the gas mileage for its new line of hybrid cars, on the average, is 60 miles per gallon (mpg) with a standard deviation of 4 mpg. A random sample of 16 cars yielded a mean of 57 miles per gallon. If the company's claim is correct, what is the probability that the sample mean is less than or equal to 57 mpg? What assumptions did you make?

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be a large random sample of size n from a population with mean μ and variance σ^2. Then,

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

- A rule of thumb is that the sample size n must be **at least 30.**
- Central Limit Theorem can be applied regardless of the distribution of the population.

Thanks!

Any questions?