Introduction to Statistics Tutorial: Sampling Distributions

INCOGEN, Inc.
2008

INCOGEN

Outline

- Sampling Distribution
- The Central Limit Theorem
- The Sampling Distribution of the Sample Mean
- The Sampling Distribution of the Sample Proportion

INCOGEN

The Sampling Distribution

The sampling distribution of a statistic is the probability distribution for all possible values of the statistic that results when random samples of size *n* are repeatedly drawn from the population.

INCOGEN

The Sampling Distribution of the Sample Mean

If a random sample of *n* measurements is selected from a population with mean *μ* and standard deviation *σ*, the sampling distribution of the sample mean $\bar{x}$ will have a mean *μ* and standard deviation of $\sigma/\sqrt{n}$.

Also, for large *n* ($n \geq 25$), the sampling distribution will be approximately normally distributed.

INCOGEN

Problem Solving Help

Before trying to calculate the probability that a statistic $\bar{x}$ falls in some interval, complete the following:

1. Calculate the mean and standard deviation
2. Sketch the sampling distribution with the mean and standard deviation
3. Locate the interval of interest and shade the area that you wish to calculate
4. Find the z-score and use the standard normal table to find the probability
5. Look at your sketch to see that the probability makes sense

INCOGEN

Example: Sample Mean

Suppose you select a random sample of 25 observations from a population with mean $\mu = 9$ and $\sigma = 0.2$. Find the approximate probability that the sample mean will lie within 0.1 of the population mean $\mu = 9$.

$P(8.9 \leq \bar{x} \leq 9.1) = ?$

Example: Sample Mean

1. Calculate the mean and standard deviation of the sampling distribution

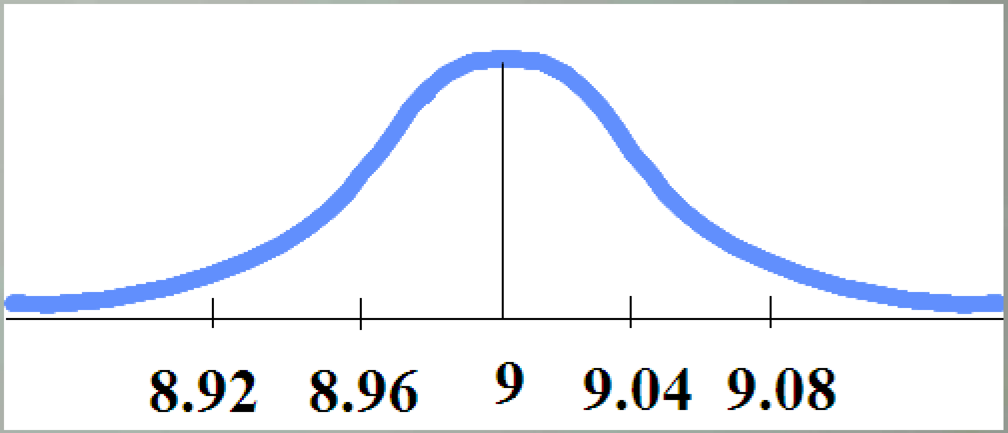
$$\mu_{\bar{x}} = \mu = 9$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.2}{\sqrt{25}} = \frac{0.2}{5} = 0.04$$

INCOGEN

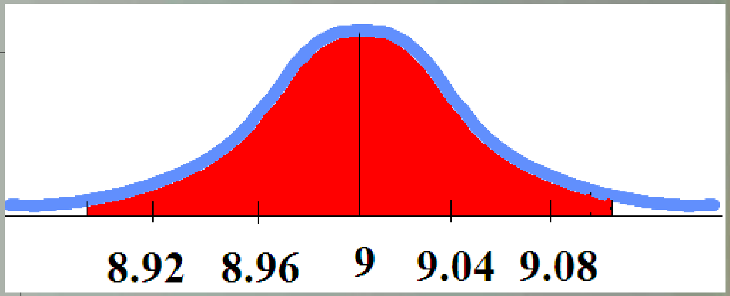
Example: Sample Mean

2. Sketch the sampling distribution with the mean and standard deviation



Example: Sample Mean

3. Locate the interval of interest and shade the area that you wish to calculate:
   $8.9 \le x \le 9.1$



4. Find the z-scores and use the standard normal table to find the probability

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{8.9 - 9}{0.04} = \frac{-0.1}{0.04} = -2.5$$

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{9.1 - 9}{0.04} = \frac{0.1}{0.04} = 2.5$$

$P(8.9 \le x \le 9.1)$
$= P(-2.5 \le z \le 2.5)$
$= 0.4938 + 0.4938$
$= 0.9876$

5. Look at your sketch to see that the probability makes sense

The Sampling Distribution of the Sample Proportion

If a random sample of n observations is selected from a binomial population with parameter p, the sampling distribution of the sample proportion

$$\hat{p} = x/n$$

will have a mean $$\mu_{\hat{p}} = p$$

and a standard deviation $$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

where $q = 1 - p$

INCOGEN

The Sampling Distribution of the Sample Proportion

Also, when the sample size *n* is large, the sampling distribution of the sample proportion can be approximated by a normal distribution.

INCOGEN

The Sampling Distribution of the Sample Proportion: Example

A random sample of 1002 people walking on the streets in New York City were asked if they had the Flu this past season. 31% said yes, confirmed by a doctor, and 35% said that that thought they may have had the Flu.

Describe the sampling distribution of $\hat{p}$, the proportion of NYC people who had confirmed cases of the Flu.

Answer: Approximately normal, with $\mu_{\hat{p}} = p$=0.31

and

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.31(0.69)}{1002}} = \sqrt{0.000213473} = 0.0146$$

INCOGEN

Example Continued

What is the probability that $\hat{p}$ will differ from *p* by more than 0.04?

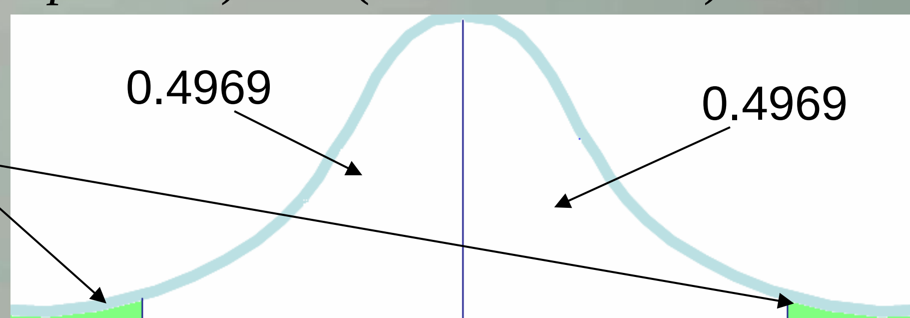
Answer: =

$$1 - P(\hat{p} - 0.04 < p < \hat{p} + 0.04) = 1 - P(0.27 < p < 0.35)$$

$$= 1 - P\left(\frac{0.31 - 0.27}{0.0146} < p < \frac{0.31 - 0.35}{0.0146}\right)$$

$$= 1 - P(-2.74 < p < 2.74) = 1 - (0.4969 + 0.4969) = 1 - 0.9938 = 0.0062$$

Area of interest



References

This tutorial is comprised of materials from the following sources:

Introduction to Probability and Statistics by Mendenhall and Beaver. ITP/Duxbury.

The Cartoon Guide to Statistics by Gonick and Smith. HarperCollins.

Basic Statistics: an abbreviated overview by Ackerman, Bartz, and Deville. 2006 Accountability Conference

INCOGEN