Hypothesis Testing of Single Mean and Single Proportion: Distribution Needed for Hypothesis Testing*

Susan Dean
Barbara Illowsky, Ph.D.

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License †

Earlier in the course, we discussed sampling distributions. **Particular distributions are associated with hypothesis testing.** Perform tests of a population mean using a **normal distribution** or a **student's-t distribution.** (Remember, use a student's-t distribution when the population **standard deviation** is unknown and the distribution of the sample mean is approximately normal.) In this chapter we perform tests of a population proportion using a normal distribution (usually n is large or the sample size is large).

If you are testing a **single population mean**, the distribution for the test is for **means**:

$\overline{X} \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$ or t_{df}

The population parameter is μ. The estimated value (point estimate) for μ is $\overline{x}$, the sample mean.

If you are testing a **single population proportion**, the distribution for the test is for proportions or percentages:

P' $\sim N\left(p, \sqrt{\frac{p \cdot q}{n}}\right)$

The population parameter is p. The estimated value (point estimate) for p is p'. p' $= \frac{x}{n}$ where x is the number of successes and n is the sample size.

Glossary

Definition 1: Normal Distribution
A continuous random variable (RV) with pdf f(x) $= \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)2}/2\sigma^2$, where μ is the mean of the distribution and σ is the standard deviation. Notation: $X \sim N(\mu, \sigma)$. If $\mu = 0$ and $\sigma = 1$, the RV is called **the standard normal distribution**.

Definition 2: Standard Deviation
A number that is equal to the square root of the variance and measures how far data values are from their mean. Notation: s for sample standard deviation and σ for population standard deviation.

Definition 3: Student's-t Distribution
Investigated and reported by William S. Gossett in 1908 and published under the pseudonym Student. The major characteristics of the random variable (RV) are:

*Version 1.13: Jun 12, 2012 8:21 am -0500
†http://creativecommons.org/licenses/by/3.0/

- It is continuous and assumes any real values.
- The pdf is symmetrical about its mean of zero. However, it is more spread out and flatter at the apex than the normal distribution.
- It approaches the standard normal distribution as n gets larger.
- There is a "family" of t distributions: every representative of the family is completely defined by the number of degrees of freedom which is one less than the number of data.