

An Expanded Glossary of Terms for the BUS 302L Statistics Exam

wayne.smith@csun.edu

[updated: Thursday, October 12, 2006]

“The first step in solving any problem is to call things by the correct name.”

---Confucius (c. 500 B.C.)

The purpose of this document is to expand on the existing Glossary for the BUS302L Statistics exam. No terms have been added, changed or deleted. I simply organized each term by the Top Ten Concept which it is most closely associated. Within each Top Ten Concept, I grouped related terms together to aid in memory recall and use. The terms within each group are alphabetical. Pages 1-3 show the groupings by topic *without* descriptive names while pages 4-7 show the groupings by topic *with* descriptive names.

Glossary (by Top Ten Concept, *without* group descriptions)

Top Ten #1 – Descriptive Statistics

Frequency Distribution
Symmetric
Skewness

Mean
Median
Mode

Range
Interquartile Range
Variance
Standard Deviation

Percentile

Joint Frequency Table

Bar Chart
Boxplot
Pie Chart
Line Chart
Scatter Diagram

Top Ten #2 - Hypothesis Testing

Null Hypothesis (H_0)

Alternative Hypothesis (H_A)

Confidence Level

Significance Level = $\alpha = P(\text{Type I error})$

Beta = $P(\text{Type II error})$

One-Tail Test

Two-Tail Test

Critical Value

Rejection Region

Test Statistic (such as z-statistic or t-statistic)

Top Ten #3 - Confidence Intervals

Confidence Interval

Point Estimate

Interval Estimate

Length of Interval

Narrow Interval

Wide Interval

Top Ten #4 - Linear Regression

Coefficient of Correlation

Coefficient of Determination

Dependent Variable

Independent Variable

Total Variation

Explained Variation

Linear Regression Line

Slope = Regression Coefficient (b_1)

Y Intercept (b_0)

Predicted Value

Standard Error of Estimate

Top Ten #5 - Expected Value

Expected Value
Weighted Average

Top Ten #6 - Probability Distributions

Binomial Random Variable (success or failure outcomes)
Normal Random Variable
Student-t Random Variable
Degrees of Freedom

Top Ten #7 - p -values

p -value
Empirical Rule for Bell/Mound-Shaped Distributions

Top Ten #8 - Variation and Uncertainty

Standard Error of the Mean (Standard Deviation of the Sample Mean)

Top Ten #9 - Populations and Samples

Population
Population Parameter
Population Proportion (π)

Sample
Sample Size
Sample Statistic
Sample Proportion (p)

Top Ten #10 - Qualitative and Quantitative Variables

Qualitative Data
Quantitative Data

Discrete
Continuous

Glossary (by Top Ten Concept, *with* group descriptions)

Top Ten #1 – Descriptive Statistics

(the “shape” of the distribution)

Frequency Distribution
Symmetric
Skewness

(measures of central tendency)

Mean
Median
Mode

(measures of dispersion)

Range
Interquartile Range
Variance
Standard Deviation

(ordered ranking)

Percentile

(typical presentation of qualitative-only data)

Joint Frequency Table

(typical presentation of quantitative data)

Bar Chart
Boxplot
Pie Chart
Line Chart
Scatter Diagram

Top Ten #2 - Hypothesis Testing

(opposing, statistical tests for confirmation/disconfirmation)

Null Hypothesis (H_0)
Alternative Hypothesis (H_A)

(your tolerance to “be wrong”)

Confidence Level

(the two types of errors in decision making)

Significance Level = $\alpha = P(\text{Type I error})$
Beta = $P(\text{Type II error})$

(directional or non-directional tests)

One-Tail Test

Two-Tail Test

(calculated numbers and area under the curve to test the hypothesis)

Critical Value

Rejection Region

Test Statistic (such as z-statistic or t-statistic)

Top Ten #3 - Confidence Intervals

(your tolerance for statistical error)

Confidence Interval

(the two types of statistical estimates)

Point Estimate

Interval Estimate

(the size of the interval on both sides of the center value)

Length of Interval

(narrower intervals are better—the true value is more likely to be in the interval)

Narrow Interval

Wide Interval

Top Ten #4 - Linear Regression

(measures of “fit” between the independent variable and the dependent variable)

Coefficient of Correlation

Coefficient of Determination

(the two types of variables of interest)

Dependent Variable

Independent Variable

(the fraction of the total variance that our regression formula can help explain)

Total Variation

Explained Variation

(component parts of the linear equation line and graph)

Linear Regression Line

Slope = Regression Coefficient (b_1)

Y Intercept (b_0)

Predicted Value
Standard Error of Estimate

Top Ten #5 - Expected Value

(the values that we are likely to get in the long run)

Expected Value
Weighted Average

Top Ten #6 - Probability Distributions

(the main types of random variables often seen in sampling situations)

Binomial Random Variable (success or failure outcomes)
Normal Random Variable
Student-t Random Variable
Degrees of Freedom

Top Ten #7 - p -values

(cumulative probability—area under the normal curve)

p -value
Empirical Rule for Bell/Mound-Shaped Distributions

Top Ten #8 - Variation and Uncertainty

(the dispersion of the data is the uncertainty of the situation)

Standard Error of the Mean (Standard Deviation of the Sample Mean)

Top Ten #9 - Populations and Samples

(“we know the whole thing”—in practice, relatively uncommon)

Population
Population Parameter
Population Proportion (π)

(“we take one or more draws from the population”— in practice, quite common)

Sample
Sample Size
Sample Statistic
Sample Proportion (p)

Top Ten #10 - Qualitative and Quantitative Variables

(two types of data--“lower”-level and “higher”-level)

Qualitative Data

Quantitative Data

(two types of numbers—cannot and can be “broken down” further)

Discrete

Continuous