

Measures of Effect-Size & Association

UM Clinical Scholars Core Competencies

We are often interested in associations (i.e., the relationship between two or more attributes or variables). An association only means that two things are related, but often we are interested in whether one thing affects or influences the other (causally related). Issues of examining causation are discussed in the sections on study design; this section discusses statistical measures of association. Below is a table summarizing the most common statistical tests used to establish the statistical significance of the association. However, always keep in mind the differences between statistical significance and effect size. Statistical significance is a measure of how likely a finding could be expected as a function of chance variation (due to sampling or stochastic variation); measures of association summarize the size of the effect (in order to assess how important the finding would be if true).

Dependent Variable	Independent Variable	Bivariate Test (Stata command)
Continuous	Continuous	Pearson Correlation (pwwcorr) or OLS regression (regress)
	Categorical - Dichotomous - >2 Categories	t-test (ttest) ANOVA (oneway, anova)
Dichotomous	Continuous or Categorical	logit or probit functions (logit)
Categorical	Categorical	Chi-square (tabulate)

Note: The above table includes the more commonly used combinations of independent and dependent variables and the more commonly used tests.

A. Relationships between two continuous variables.

Example 1. Greater age is associated with higher weights (a positive or direct relationship)

Example 2. Greater height is associated with lower IQ (a negative or inverse relationship)

Some relevant **Stata Commands** = *correlate*, *pwwcorr*, *spearman*, *pcorr*

Common Statistical measures:

1. Pearson's Correlation Coefficient: ranges from -1.0 (perfect negative correlation) to 0 (no correlation) to 1.0 (perfect positive correlation). This statistic summarizes information on a scatter plot (therefore visual inspection of the scatter plot is important) and results can be unduly influenced by one or more outliers (i.e., observations with unusually high or low values).
2. Spearman's Rank Correlation Coefficient and Kendall's Tau: These measures also range from 0 to 1.0, but are less effected by outliers and are less dependent upon the variable being truly continuous with equivalent intervals throughout it's scale range (meaning that a 1 point difference has the same meaning whether it's at the low end or the high end of the range). (See C. below)

B. Relationships between two nominal or categorical variables (including dichotomous variables).

Example: People who take beta-blockers post MI are less likely to die during the subsequent year than those who do not.

Some relevant **Stata Commands** = *tabulate, epitab, table, kap*.

Common Statistical Measures [all of the measures below summarize information from cross-tabulations, and you should examine the detailed results in the cross-tabs and not just rely upon the summary statistics]. The Chi-square statistic is the most common statistical test for statistical significance.

1. Relative risk (or Risk Ratio)
2. Odds Ratio (or Cross-Product Ratio)
3. Difference (or delta or absolute risk reduction)
4. Risk Reduction
5. Number Needed to Treat (NNTT)
6. Cohen's Kappa (measure of agreement above chance agreement)
7. Sensitivity, Specificity and likelihood ratios (See section on diagnostic tests and testing)

C. Relationships between two ordinal variables.

Example = Those who are more likely to agree that "country music is one of the great evils of Western society" on a Likert-type scale (5 point scale ranging from strongly disagree to strongly agree) have higher self-ratings of health (measured on a 5 point scale ranging from poor to outstanding).

Some relevant **Stata Commands** = *spearman, tabulate*.

Common Statistical Measures:

1. Spearman's Rank Correlation Coefficient. (see above)
2. Kendall's Tau (see above)

D. Independent relationships between variables (Using Multivariate Analysis)

After adjusting for age and socio-economic status using multivariate analysis (i.e., independent of age and SES), women with diets high in fruits and vegetables and low in animal flesh had fewer pre-term deliveries.

Some relevant **Stata Commands** = *regress, logit, logistic, cox, st, stcox*

Common Statistical Measures:

A variety of measures of association can be generated from multivariate models, many of which are identical or analogous to the measures discussed above (odds ratios are a common summary of effect size derived from multivariable regression). The raw beta coefficient from the regression outputs is a multiplier to be used in the algebraic equation.

Example for Ordinary Least Squares (i.e., Linear) Regression

Formula: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$

- y = dependent variable
- x = independent variables
- β = coefficients that when multiplied by values for x give you the best prediction of y.

Therefore, for the regression output shown on page 117 of the Stata manual on “[5s] regress - Linear Regression,” results for the Stata command “regress mpg weight weightsq foreign” allows you to predict miles per gallon (mpg) by plugging in values for the weight of the car (weight), weight of the car squared (weightsq), whether the care is foreign or domestic (foreign) and a constant (_con) using the following algebraic formula:

$$\text{mpg} = (-0.0166)(\text{weight}) + (1.59 \times 10^{-6})(\text{weightsq}) + (-2.20)(\text{foreign}) + 56.54$$

How well your regression equation will predict values depends upon 2 factors: 1) its discrimination (often measured by the R² [how much of the total variation is explained by your model]), and 2) its calibration (is the model well fitted to the data). The later issue is discussed in the Stata manual under [5s] fit. In order to use a regression equation for prediction it is extra-essential that the calibration of the model is excellent!!

If the dependent variable is:		Type of Multivariable Model
Continuous:	Near normal	Ordinary Least Squares (regress).
	Skewed	Transform the dependent variable
	Positively skewed count data	Negative binomial (nblogit), or Poisson regression (poisson), or two-part models.
Categorical:	Dichotomous	Logistic regression (logit or logistic), Cox-proportional hazards (cox)
	Multiple categories	Multinomial logistic regression (mlogit).
	Ordered categories	Ordered logistic regression (ologit)

Note: The above table should not be seen as complete, but includes the more commonly used techniques.

There are a variety of other measures of association that are less commonly encountered in the medical and health services literature. Some of these include: Tschuprow’s Contingency coefficient, Cramer’s Contingency coefficient, Goodman-Kruskal statistics, Somers’ d and Wilson’s e. Many of these are discussed briefly in the Stata manuals, and for the brave of heart, all can be found in the cryptically written Sage Publication #32.