Statistical basics for Biology: p's, alphas, and measurement scales.

Catherine Teare Ketter

School of Marine Programs
University of Georgia
Athens Georgia 30602-3636
(706) 583-0862
cmscatk@uga.edu

Catherine Teare Ketter received her B.S. and M.S. in Biology from the University of Alabama. She received her Ph.D. in research methodology and applied statistics in 1990 from the University of Alabama with an emphasis in the application of nonparametric multivariate techniques to biomedical and health behavioral data. Since 2000, Catherine has been responsible for developing the undergraduate marine sciences laboratory courses at the University of Georgia. During the summer Catherine teaches a nonscience major marine biology field course and a Marine Science for Teachers course. She holds a Master's level certificate in comprehensive science.

Note: This is a summary of the original PowerPoint presentation made at the UNLV 2003 ABLE meeting.

Reprinted From: Teare Ketter, C. 2004. Statistical basics for biology: p's, alphas, and measurement scales.. Pages 334-339, *in* Tested studies for laboratory teaching, Volume 25 (M. A. O'Donnell, Editor). Proceedings of the 25th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 414 pages.

Statistical basics for Biology: p's, alphas, and measurement scales.

- Purpose: To address common misconceptions regarding and errors in the use of statistical procedures by basic science students.
- Overview of presentation
  1. Scales of measurement
  2. Nonparametric and parametric techniques
  3. Descriptive versus inferential statistics
  4. Population parameters versus sample statistics
  5. Type I and Type II error
  6. Tests of inference and modeling
  7. Data transformation

Scientists have long been concerned with the observation and quantification of relationships between objects and processes in biotic and abiotic systems. There has been controversy in the assignment of numbers to objects, behavioral and physical. In practice, this assignment of numbers to objects should be applied in a systematic and meaningful way.

Scales of Measurement – Data Types

On the theory of scales and measurement. Stevens, S.S. 1946. *Science*. 103, 677-680.

1. *Nominal*
    - Simplest level of measurement
    - Discrete data which represent group membership to a category which does not have an underlying numerical value
    - Data transformations which preserve group membership are acceptable
    - Examples: ethnicity, color, pattern, soil type, media type, license plate numbers, football jersey numbers, etc.
    - May also be dichotomous: present/absent, male/female, live/dead
2. *Ordinal*
    - Next level of measurement
    - From the word “order”
    - Includes variables that can be ordered but for which there is no zero point and no exact numerical value
    - Data transformations that preserve group membership and order are acceptable
    - Often used for affective instruments which assess feelings or preferences
    - Likert scale used in surveys – strongly agree, agree, undecided, disagree, strongly disagree
    - Distances between each ordered category are not necessarily the same
    - Examples: shoe size, preference ranks (Thurstone rating scale)
3. *Interval*
    - Includes group membership, order, and equal distance between points on the measurement scale
    - Arbitrary zero point on measurement scale
    - Data transformations which preserve group membership, rank-order, and interval equality are permissible
    - Most common example of interval scale is temperature
    - Celsius and Fahrenheit scales have arbitrary zero points, rank-order, an interval equality – can convert from one scale to another
    - Behavioral data – semantic differential, composite subscale scores for intelligence, aptitude, creativity, etc.
4. *Ratio*
    - Measurements in which group membership, rank-order, interval equality, and a true zero point are present
    - Frequently encountered in the physical sciences
    - Examples: acceleration (m/hr^2 or km/hr^2), velocity (m/hr or km/hr), density (mass/unit volume), and force (Newtons=kg m/s)
    - Behavioral data does not produce ratio data
        - Intelligence has no theoretical or observable zero point, for example.
        - Only when a ratio of two known interval values on behavioral constructs is calculated are ratio level behavioral variables produced

- Permissible operations for ratio data include the multiplication by a positive constant
- Permissible transformations preserve group membership, rank-order, interval equality, and the existence of a zero point

Summary of Measurement Scales

(after S.S. Stevens, 1946)

Scale	**Basic Empirical Operations**	**Mathematical Group Structure**	**Permissible Statistics**
Nominal	Determination equality	Permutation group x' = f(x) f(x) means any one-to-one substitution	Number of cases Mode Contingency correlation
Ordinal	Determination greater or lesser	Isotonic group x' = f(x) f(x) means any monotonic increasing function	Median Percentile Ranks Rank-order Correlation
Interval	Determination of interval equality	General linear group x' = ax + b or differences	Mean Standard Deviation Pearson product-moment Correlation
Ratio	Determination of equality	Similarity of group of ratios x' = ax	Coefficient of Variation

Table of Correlation Coefficients for Variables Measured with Different Scale Properties

Adapted from Glass & Stanley (1970)

	Dichotomous Nominal	**Dichotomous Nominal with Underlying Normal Distribution**	**Ordinal**	**Interval or Ratio**
Dichotomous Nominal	phi	*	rank biserial	point biserial
Dichotomous Nominal with Underlying Normal Distribution	*	tetrachoric r	*	biserial
Ordinal	rank biserial	*	Spearman's or Kendall's tau	*
Interval or Ratio	point biserial	biserial r	*	Pearson product moment correlation

**** No coefficient is specifically available.***

Parametric versus nonparametric methods

- *Nonparametric statistical methods*
  - Used with discrete, non-continuous data (nominal or ordinal variables)
  - Used when the sample data are not normally distributed or when the underlying variable of interest is not normally distributed

- *Parametric statistical methods*
  - Used with interval and ratio level data
  - Used when the sample distribution is normally distributed
  - Used when the variable of interest is normally distributed in the larger population from which the sample was taken

Descriptive versus inferential statistics

- *Descriptive statistics*
  - Operations performed on data which constitute a sample from some larger population
  - Describe the properties of the sample population only
  - Results are not generalizable to the larger population
  - Limited sample size, restriction in range for the variable of interest in the sample, other factors limit the generalizability of study results

- *Inferential statistics*
  - Operations performed sample data extracted from a larger parent population
  - Sample statistics relate directly to the larger parent population from which the sample was taken
  - Purpose of inferential statistical procedures is to evaluate hypotheses relating to a sample and generalize these relationships to the larger parent population

Population parameters versus sample statistics
scribed by a set of numerical measures called parameters

  - Population mean, μ
  - Population variance, σ

- Sample: A subset of measurements taken from a population
  - The distribution of the parameter/variable in the sample should approximate the distribution of the parameter in the population – problems: restriction of range
  - Increasing sample size increases precision of the sample estimates of population parameters
  - Sample mean, C (x-bar)
  - Sample variance, s

Type I and Type II error

- A ***type I error*** occurs when the null hypothesis is accepted when it is true. The probability of type I error is α.
- A ***type II error*** occurs when the null hypothesis is accepted when it is false. The probability of a type II error is β.
- For a given sample size, α and β are inversely related. As α increases, β decreases.
- A priori error rate that is acceptable for a test of hypothesis is α.
- α is usually set as 0.05 or 0.01 depending upon the purpose of the analyses.

Decision	**Null Hypothesis**	
	False	**True**
Reject H_0	Correct	Type I error
Fail to reject H_0	Type II error	Correct

Distinction between p and α

- *p* is the probability associated with the acceptance of the null hypothesis, H_0, when it is true. Prior to a test of hypothesis, the researcher needs to set an a priori α, error rate.
- The probability associated with the calculated value of the test statistic is *p*.
- The level of statistical significance is determined by α, a test of hypothesis is significant if $p \leq \alpha$.
- Calculated values of the test statistic may have greater or lesser-associated probabilities, but it is not correct to say that one is "more significant" than another!

Inferences and modeling

- The purpose of inferential statistical procedures is to evaluate hypotheses relating to a sample, and generalize these relationships to the larger parent population.
- Modeling generates an equation that can be used to predict future events using data collected over time. Modeling usually involves some type of regression analysis (can be linear, quadratic, polynomial, logistic, etc).

Data transformation

- Data transformation is useful if:
    - Data are not normally distributed
    - Modeling
  - Results of data analysis using transformed sample data are generalizable only to the population of transformed data.
  - Transformation is NOT useful in the purpose of the analysis is to make inferences from the sample to the population from which the sample was derived.

Summary

- Four levels of measurement: nominal, ordinal, interval, and ratio
- Nonparametric statistical methods are used with nominal and ordinal data.
- Parametric statistical methods are used with interval and ratio data.
- Descriptive statistics produce results that describe the properties of the sample population only.
- Descriptive statistics cannot be used to make inferences about the larger population from which the sample was drawn.
- Descriptive statistics are appropriate when sampling is not random (restriction in range) or the underlying variable of interest is not normally distributed.
- The purpose of inferential statistical procedures is to evaluate hypotheses relating to a sample and to generalize these relationships to the larger parent population.
- A ***type I error*** occurs when the null hypothesis is accepted when it is true. The probability of type I error is α.
- A ***type II error*** occurs when the null hypothesis is accepted when it is false. The probability of a type II error is β.
- The level of statistical significance is determined by α, a test of hypothesis is significant if $p \leq \alpha$.
- Calculated values of the test statistic may have greater or lesser-associated probabilities, but it is not correct to say that one is “more significant” than another!
- Don’t transform non-normally distributed data unless you are using a modeling technique – if the purpose of the analyses is inferential, use the appropriate nonparametric technique!

Take home message

- Take the time to talk with your class about measurement levels. The lowest level of measurement in your data set dictates the types of procedures you can employ in data analysis.
- Make a distinction between p and α.
- Statistics are a very useful tool but should not substitute for good experimental design. Design should dictate the analyses used.
- Have fun!