A (Very) Brief Introduction to Archaeological Stats

Statistics concerns the organization, analysis, and presentation of data. Within this, we can think of statistical work as a *tool* in archaeological research for illuminating patterns in the archaeological record that can, in turn, provide us with insight on past people's behaviors. Stats should *never* be an "end goal" in archaeological research, they should be used to help answer a research question about the people who created a stone point or built a structure we are measuring. Stats can be helpful in archaeology, for sure. But don't fetishize statistical analysis, archaeological sample sets are small, biased and incomplete – stats can only show you a pattern, you have to tease out what it means.

Archaeologists gather significant amounts of data of varying types (and quality). For the moment, we will put aside the very serious question of data accuracy and fidelity. For the most part data in archaeology consists of measurements taken on artifacts and during excavations. Let's think about the different types of data in archaeology:

Dichotomies measure the presence of absence of a characteristic.

Categorical (nomial) measurements create two or more categories of qualitative difference, such as cord-marked, incised, or burnished.

Rank (ordinal) measurements allow for comparison between two objects based on more or less of a characteristic (like older or younger) but there is no exact quantity of difference.

Numeric measures express amounts of a variable in numerals.

> **Interval** numeric measures do not include an absolute zero (e.g. Fahrenheit, years BP), while **Ratio** numeric measures do (e.g. length, width, Kelvin).
>
> **Discrete** numeric measures only have positive integer values (e.g., number of sites or pots), while **Continuous** can take any real value (e.g., length, thickness).

In archaeology, we never have all of the measurements that have ever existed for a given characteristic, so we say that we are looking at a **sample** from a hypothetical **population**. **Populations** are described by **parameters**, while **samples** are described by **statistics**. In archaeology (and in almost every aspect of life) we lack the population – instead we are looking at samples of a population, estimating the population parameters from the sample statistics.

A few basic statistics that we can look at: **mean -** central value of a discrete set of numbers (sum of numbers divided by the number of values); **median -** the value separating the higher half from the lower half of a sample; **standard deviation -** a measure of the amount of variation or dispersion of a set of values; **quantiles** – a division of data into groups so that the same number of values falls into each group.

Parameters of a population are defined differently depending on the distribution of the values for that characteristic. Archaeological data can be described by a number of distributions, but there are four most common.

> **Binomial distribution** – a discrete distribution where values fall into one of two categories (present/absent, decorated/plain), defined by probability of presence (of one category).
>
> **Poisson distribution** – a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known

constant mean rate and independently of the time since the last event, defined by lambda (average number of specimens in each unit).

Normal distribution – continuous probability distributed in a symmetrical pattern, defined by mean and variance.

Log-normal distribution – a continuous probability distribution whose logarithm is normally distributed (asymmetrical with a long tail extending to the right).

Due to the problems inherent in (archaeological) samples, such as biases and small size, we end up having to use the concept of **probability:** Where in the distribution of sample statistics does the sample I got fit if that distribution is produced by samples sizes of n taken from a population with specific parameter(s)?

If the sample statistic is close to the center we can say there is no evidence **that the sample did not come** from a population with those specific parameters. In archaeology, we often use the **95% confidence interval** to judge this – if the sample statistics fall within the 95 of distributed values occurring in an infinite sampling of a population, then there is no reason to doubt it came from that population.

If the sample statistic is out in the extremes we can say there is **evidence that the sample might not have come** from a population with those specific parameters. See the reverse of above.

This comes into archaeological statistics when we are trying to identify if the differences in two or more samples is a result of sampling bias (meaning they come from the same population with the same parameters) or if the samples come from different populations (i.e. there is some reason for the differences in samples). We might compare whether the different dimensions of stone points from site A to site B are due to some human reason (outside of sampling bias). If we determine that the differences are statistically **significant**, then we can begin to interrogate reasons for this difference. This is called **Hypothesis Testing**.

Principle components analysis (PCA) is a technique to look for a way to simplify the dimensionality of multivariate data. In archaeology we often have a lot of different variables describing a single object, but this makes it difficult to visualize those data. A scatter plot can be use to show two variables (like length and width of a stone point) but what do we do if we have more than three variables? PCA simplifies our data by identifying the direction of maximum covariance (or correlation), then creating two components: direction of that max covariance and the next largest direction of covariance.