

The Description, Display, and Measurement of Variation

One of the fundamental characteristics of life is that individuals are not identical. Nevertheless, they belong to biologically meaningful categories that we term species, populations, colonies, foraging groups etc. Moreover, in the course of scientific experiments, we place individuals into groups of our own devising -- for example, control and experimental treatment groups.

Given that there is always some variation between individuals, how do we describe groups without describing every individual? The answer is that we rely on what are known as **descriptive statistics** and certain related graphical **techniques**. You will use both approaches extensively in lab; what is provided below is a brief introduction into a few measures we will commonly use in class.

In describing a group, first we must decide on what attribute(s) that we intend to use and how to measure them. There are two general types of variables and it is important that you understand their differences (these have been mentioned previously so what is below should be review):

- **Discrete variables** are those that have a limited number of defined states. To a first approximation, the sex of an organism is a discrete variable -- female or male. (You will see later in other courses that individual's sex is may in fact be a far more complicated question, but hopefully you get the idea of what a discrete variable is). Discrete variables are also sometimes (but not always) referred to as **categorical variables** since they assign individuals to separate categories or groups.
- **Continuous variables** seem to vary smoothly or continuously and, given enough sensitivity in measurement, there are potentially an infinite number of values between any two measurements. Height and weight are commonly cited measurements of continuous variables.

Discrete variables are normally used "as is". Individuals are simply characterized because there are a limited number of choices. But what about continuous variables? We normally round them off -- we do what is commonly referred to as assigning them to bins or pigeonholes. For example, it is typically meaningless to give a human's weight as 153.936783 lbs. Instead, you would say 154 or 155. In doing this you have made a continuous variable into a discrete one (values 153, 154, 155 etc or 150, 155, 160) that follows a certain definition (closest pound or closest 5 lbs). Lumping and approximating continuous variables makes them far easier to work with and it is the typical result of any measurement. However, one should always be aware that approximation is involved and some error is induced in "binning" a continuous variable and this must be taken into account when drawing conclusions about a

group (this is one of the purposes of inferential statistics mentioned in the "Methods of Science" handout).

The Histogram

Histograms are graphs that are used to display information about groups. Graphs are useful because, once reading them is mastered, they allow for the presentation of a large amount of data in short form.

Histograms are always graphed with:

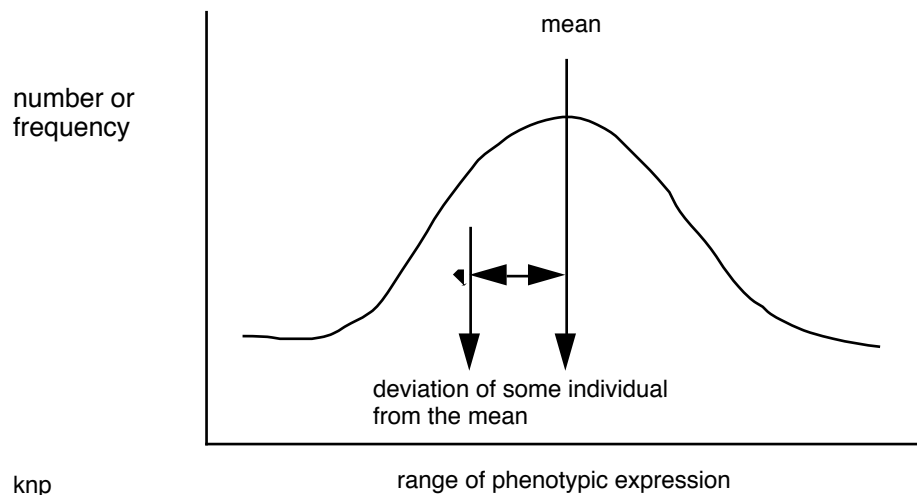
- the number or frequency of a particular value plotted on the Y-axis (for example, the **number or frequency** of subjects whose heart rate at rest is between 71 and 75 beats min^{-1} between 76 and 80 etc.

What is a frequency? How would you calculate a frequency?

- some measure of the trait on the x-axis -- in the example above, that would be the "bins" 71 to 75, 76 to 80 etc.

Here is an example that shows a special type of distribution called a normal distribution. You will learn about it in lab:

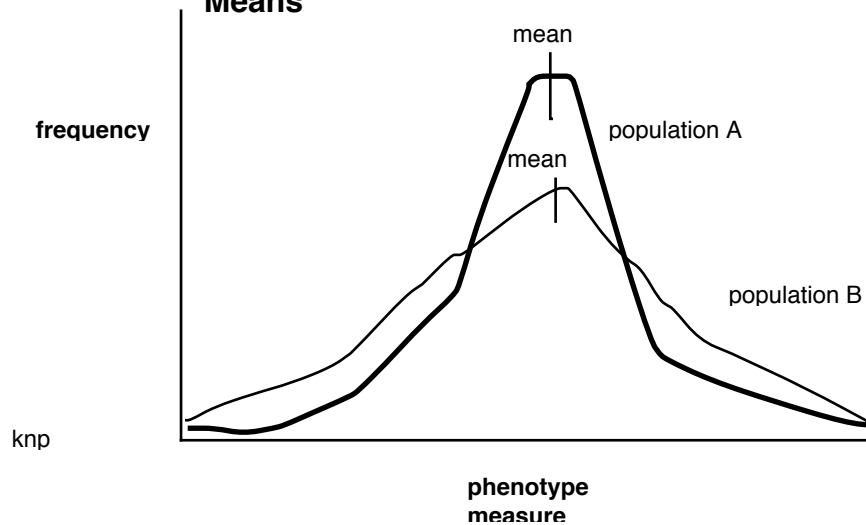
Normally Distributed Variation



We will see other types of distributions in class and lab throughout the semester.

Any population can also be abstracted to a description consisting of its **average or mean** and some type of **measure of variability**. Here is an example of histograms of two populations with the same means (averages) but with some differences in their variability:

Histograms of Two Populations with Different Degree of Variation and the Same Means



Describe the differences in the variation between the two populations. How are they similar?

You will learn measurements of variation in lab. Examples include range, variance, and standard deviation.

Be sure you know how to construct and read a histogram -- there is an example below.

Questions

1. Be able to distinguish between discrete and continuous variables. Be sure you understand what categorizing or binning continuous variables means.

Incidentally, binning is what happens when a computer digitizes any signal -- for example, when a sound (a continuously varying pressure/displacement wave) is digitized to make a digital recording. Old style analog tape measurements essentially represent the sound as a continuous variable.

2. To some degree is the differentiation between continuous and discrete variation artificial? Explain.

3. Suppose that you measure the following weights for a group of experimental animals:

1.2	2.0	0.5	3.0	2.5
1.4	1.8	0.9	1.1	2.2
1.4	1.5	1.0	0.9	2.8
1.5	1.8	2.4	1.9	2.0

Assign all values to 0.5 unit bins (0 to 0.49, 0.50 to 0.99 etc. up to 3.0 to 3.49). Plot two histograms -- one using the number of values in each bin and the other using the frequencies. Plot each bin using its central number (0.25, 0.75 et). Bring your graphs to class.