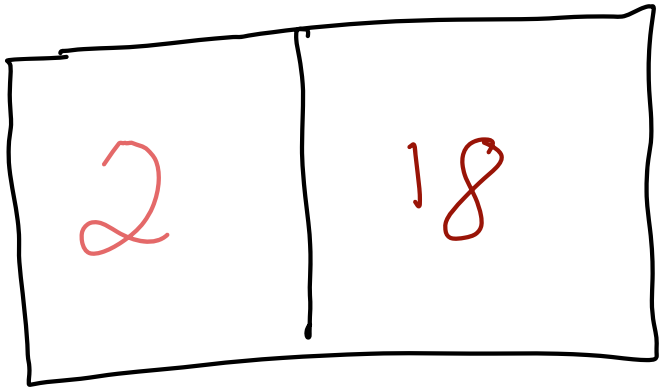
This sampling phenomenan has an analogy in academia. Often, universities want to use metrics of the university to provide evidence of the student experience. An important component is how much a student has access to faculty. Sometimes schools use student:faculty ratios. Othertimes they use average class size. But the estimate that measures "average class size" depends on whether you sample classes as individuals in a population (i.e., the university) or students as individuals in the population.

Here is a simple example- Our university (population) has 2 classes with 20 students total. One class has 2 students and the other has 18 students.

If we take classes as individuals and class size as a characteristic of the individual, here is our population.



Average class size $= \frac{2+18}{2} = 10$

If we take students as individuals and we take their characteristic as the number of students in their class, then we have this picture of the population:

2	2	18	18
18	18	18	18
18	18	18	18
18	18	18	18
18	18	18	18

The average in this case is

$$\frac{2\times 2 + 18\times 18}{20} = 16.4$$

The answers are both correct characterizations of the university. But the answers are different.

They differ because they represent different populations, and therefore relate to different questions. One is about classes and the other is about students. The choice you use should be guided by your question!

Characteristics of Individuals Versus Samples and Populations

- Once we sample individuals, we want to measure and extract information about them that is relevant to our question.
- Information comes in two general classes.

Numeric

- Numeric measurements can be assigned a number.

Examples

Size: 35 mm

Time: 5 h 36 m 25 s

Categorical

- Categorical measurements can not be described by a number.

Size: "Small"

Species Name: E. coli

Color: "Red"

Numeric measurements can be discrete or continuous.

- Discrete means that measures fall within defined bins. Exs: Limb number: 0, 1, 2, 3, ... ; Age (years): 0, 1, 2, 3, ...
- Continuous means that the number one measures about an individual does not fall within a bin.

Exs: Limb length: 26.31, 2.54, 0.25, ...

Biomass: 365.21g, 217.14g, 298.83g, ...

Categorical measures can be ordinal or non-ordinal.

Ordinal measures are not numeric, but can be ordered.

Examples

Size: Small, Medium, Large

Development: Juvenile, Sub-Adult, Adult

Non-ordinal measures have no natural ordering.

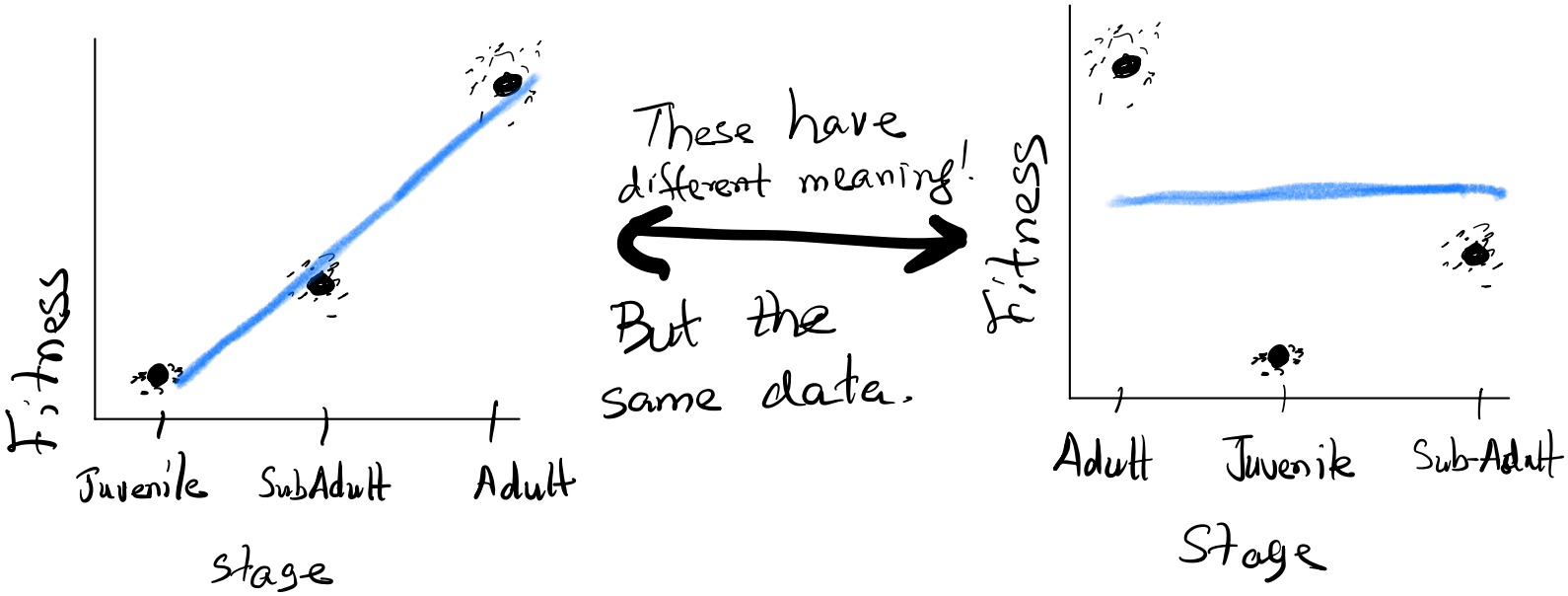
Examples

Color: Blue, Yellow, Green, etc.

Sex: Male/Female

Genus: Astragulus, Poa, Eriogonum, ... ← These are plant genera.

We distinguish ordinal measures because it changes the prentation of data and the kinds of analyses we cand do.



Properties of populations

Properties are anything you can measure about individuals that are aggregated at a population level.

- Population properties are measurements about sets of individuals. They summarize information about a collection.

Measures of Average Tendency

Means
Medians
Quantiles
Proportions

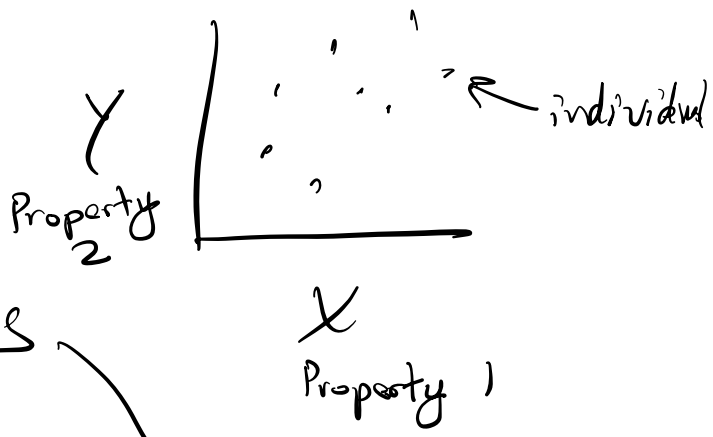
Measures of Variability

Variance
Standard Deviation (= square root of variance)
Skew
Kurtosis
Differences

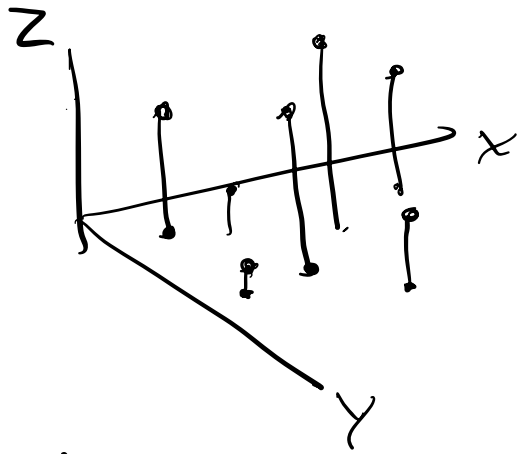
Measures of Association

Correlations

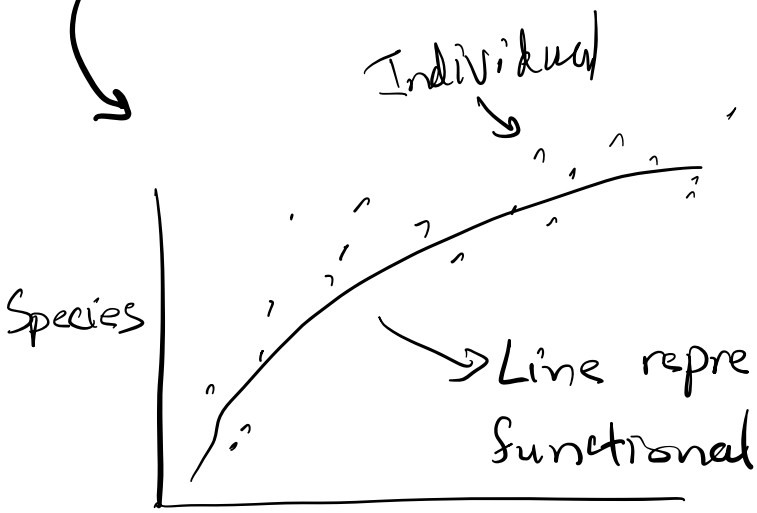
Covariances



Multivariate dissimilarities



Functional Relationships



Area