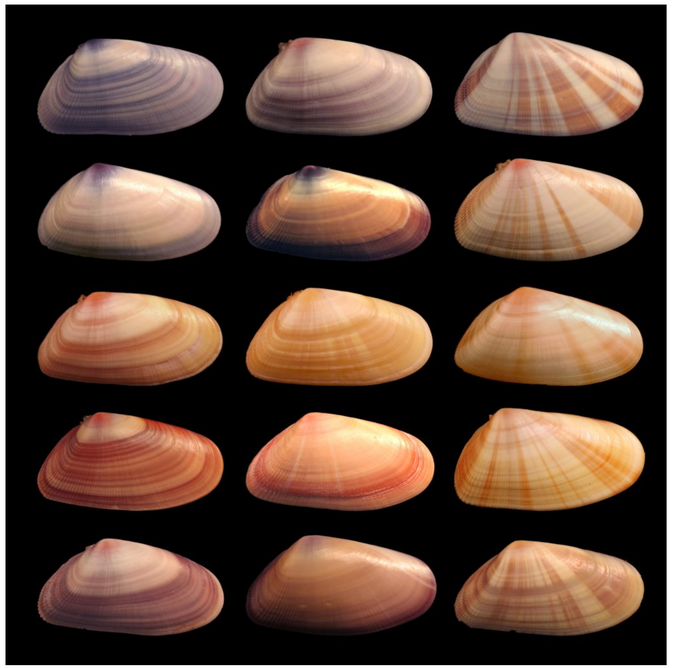
github.com/clay-small/teaching-demo

Quantitative Traits in Nature:

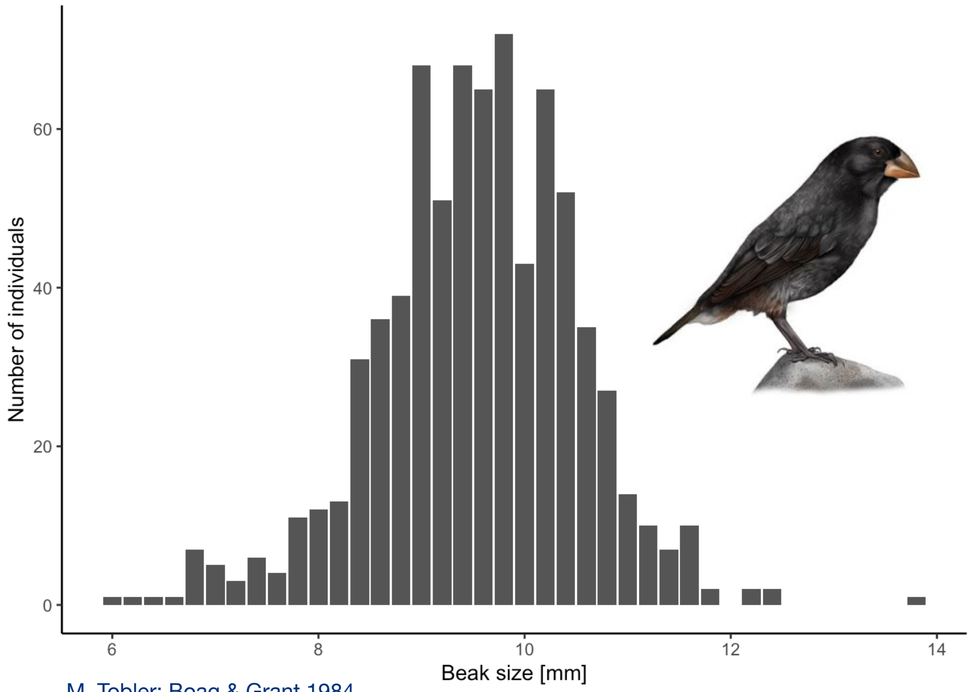
How do first principles of genetic inheritance explain their underlying statistical distributions?

Clay Small

Data Science Department - University of Oregon



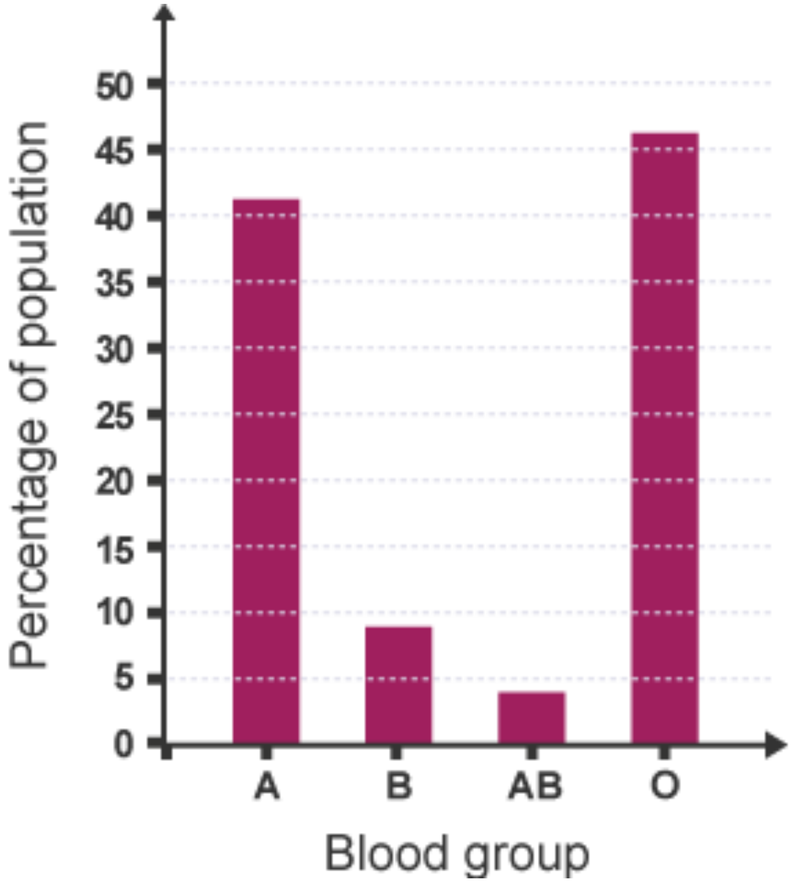
B. de Bivort



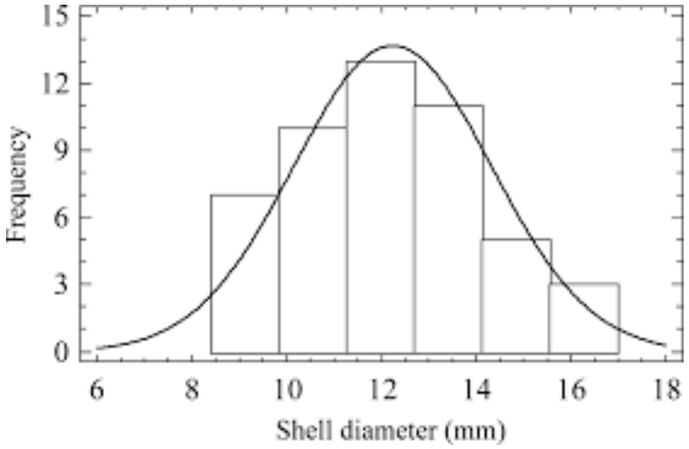
M. Tobler; Boag & Grant 1984

Measuring and quantifying traits

Discrete



Continuous



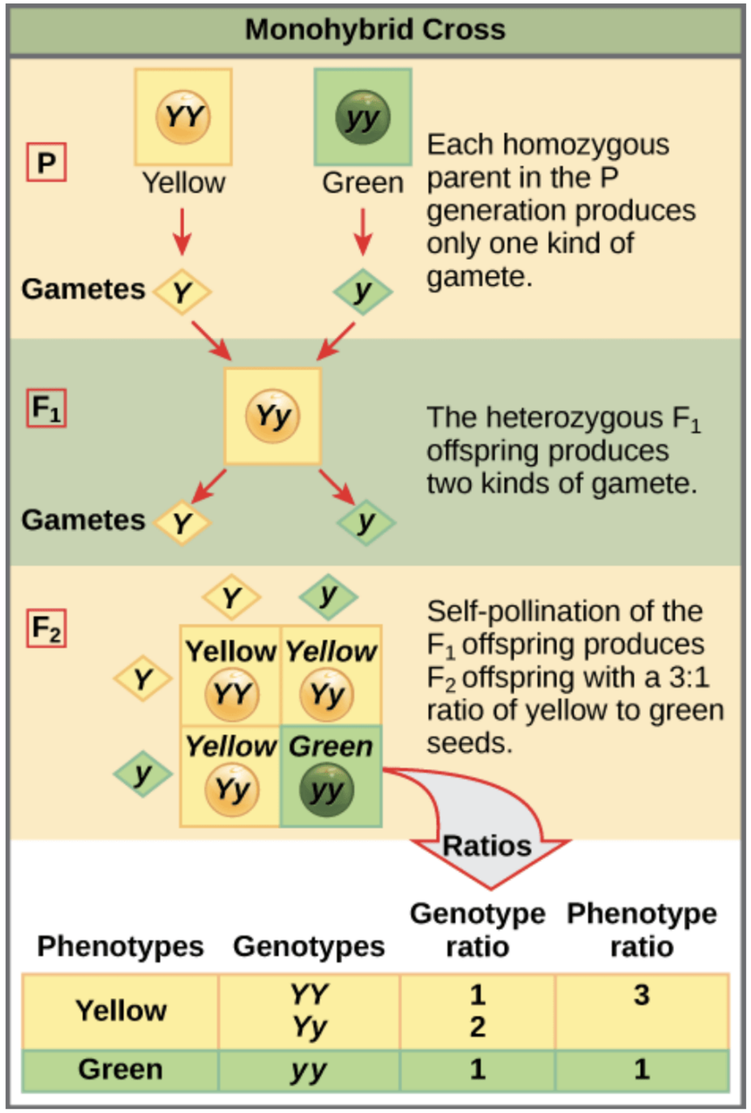
Measuring and quantifying traits

Importance of understanding quantitative trait variation

- Plant and animal breeding
- Predicting risk for human disease traits
- Phenotypic evolution in natural populations

Genes -> Phenotypes

A historical controversy

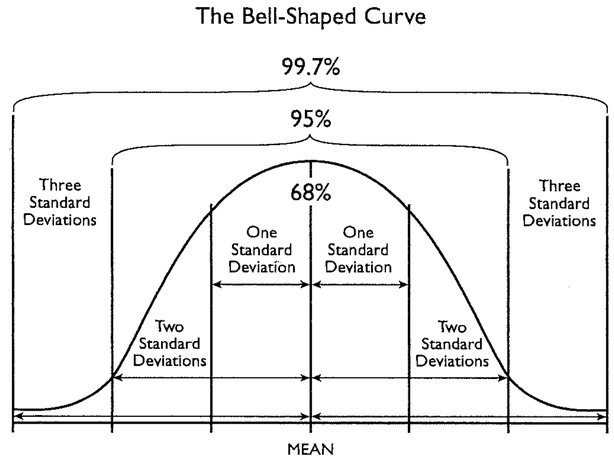


OpenStax Biology

The "Mendelians"

discrete genetic factors ->
discrete phenotypes

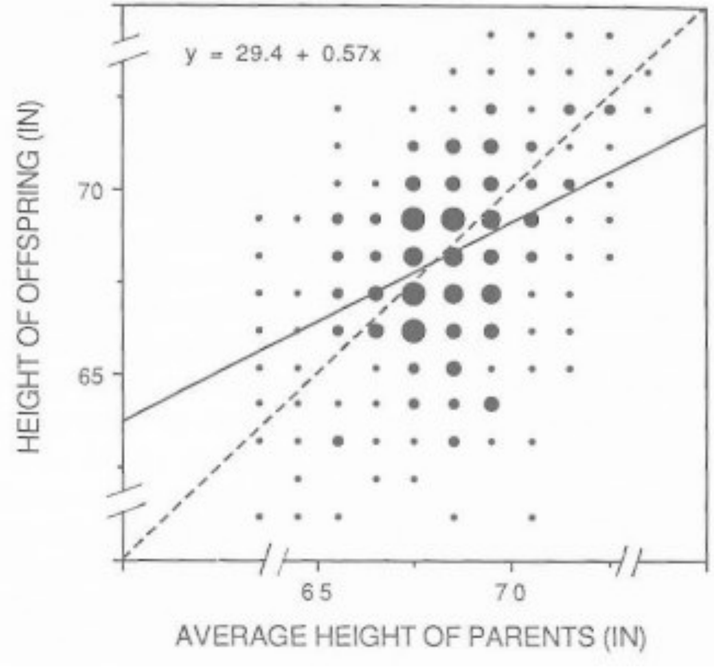
evolution: large steps



? genetic factors ->
continuous phenotypes

evolution: gradual

The "Biometricians"



A historical controversy (reconciled)

Moving toward synthesis: “The multiple-factor hypothesis”

1 (diallelic) locus -> **3 genotypes**: AA Aa aa

10 (diallelic) loci -> $3^{10} \approx$ **60,000 genotypes**

<u>Also</u>: Non-genetic (environmental) factors add continuous variation

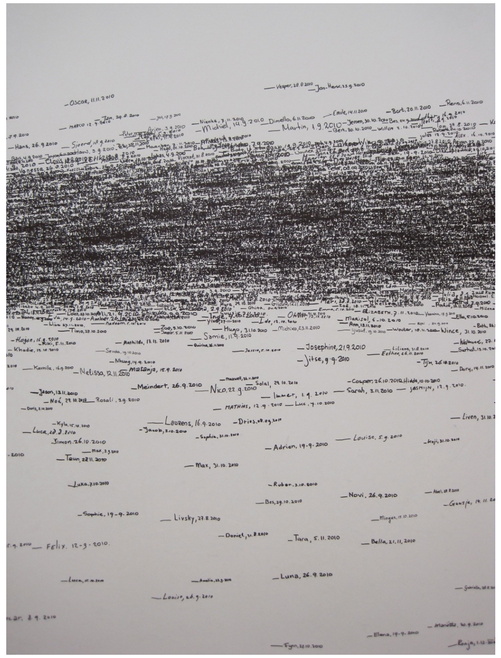
Simulating the "multiple-factor hypothesis" in R

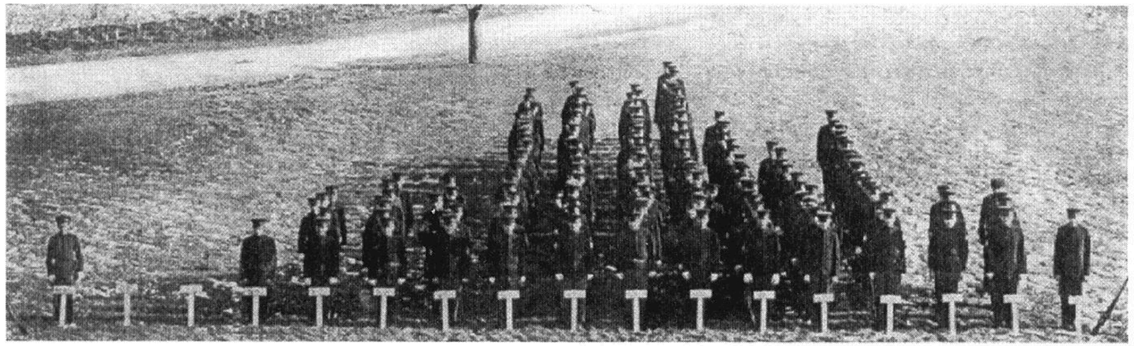
Can we simulate a ≈ Normal distribution for 500 Idaho giant salamander tail length measurements, from discrete genotypes at 5 loci?

J. Cossel Jr.

What “ingredients” do we need for our simulation?

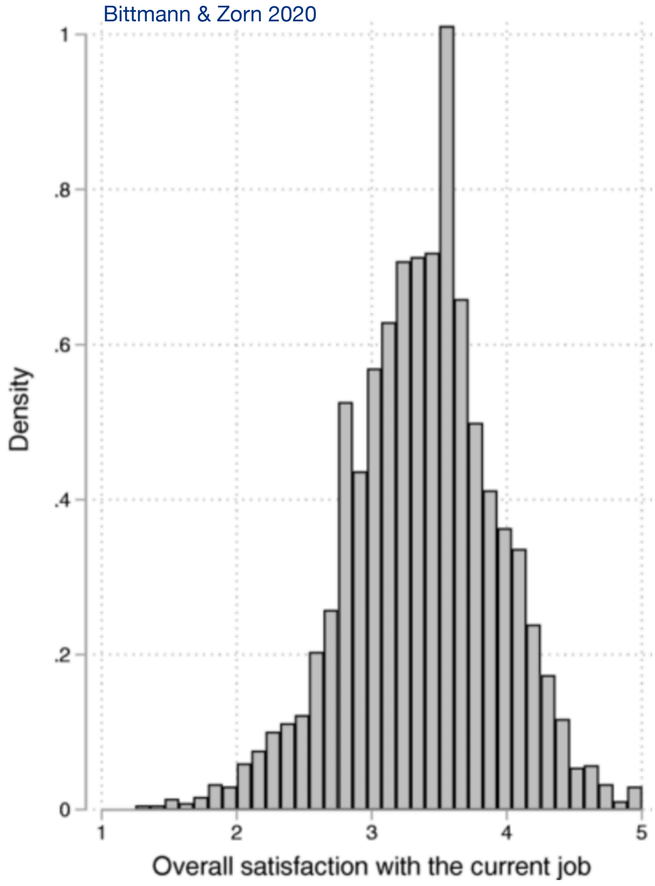
Why are quantitative traits often normally distributed?

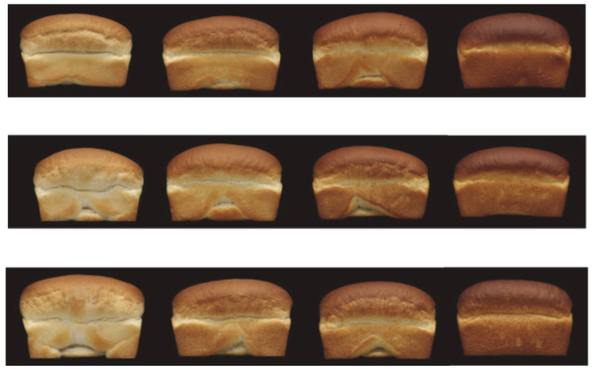




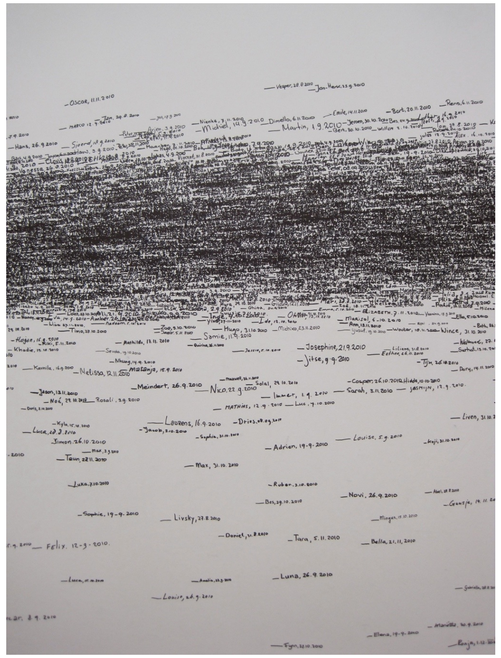
Non-biological variables are, too!

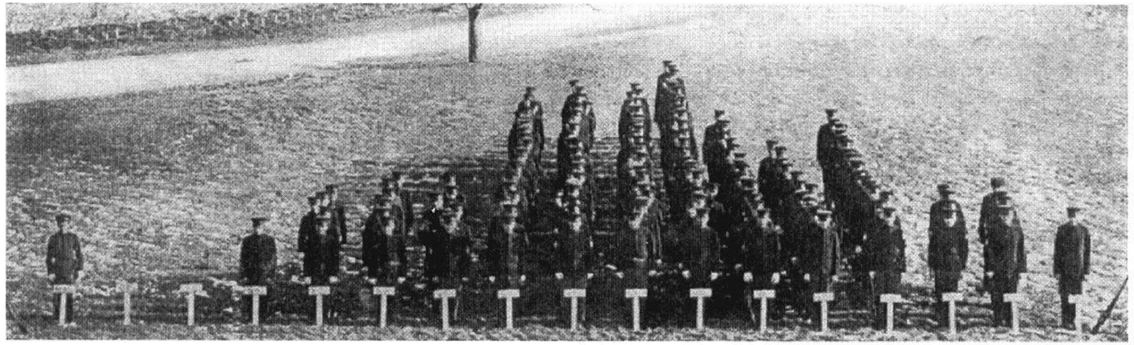
Bittmann & Zorn 2020





Why are quantitative traits often normally distributed?





"fuzzy analogue" of the central limit theorem:

Variables which are, themselves, influenced by the sum of many, independent underlying variables tend to be normally distributed.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

μ ("mu") = mean

σ ("sigma") = standard deviation

$$\boldsymbol{V_P} = V_G + V_E + V_{GxE}$$

$$\overbrace{V_A + V_D + V_I}$$

Summary and questions for further consideration

Quantitative Traits | Inheritance | Simulation of CLT

- What assumptions does our simulation make?
- How might we introduce environmental contributions?
- How might we include dominance in our simulation?