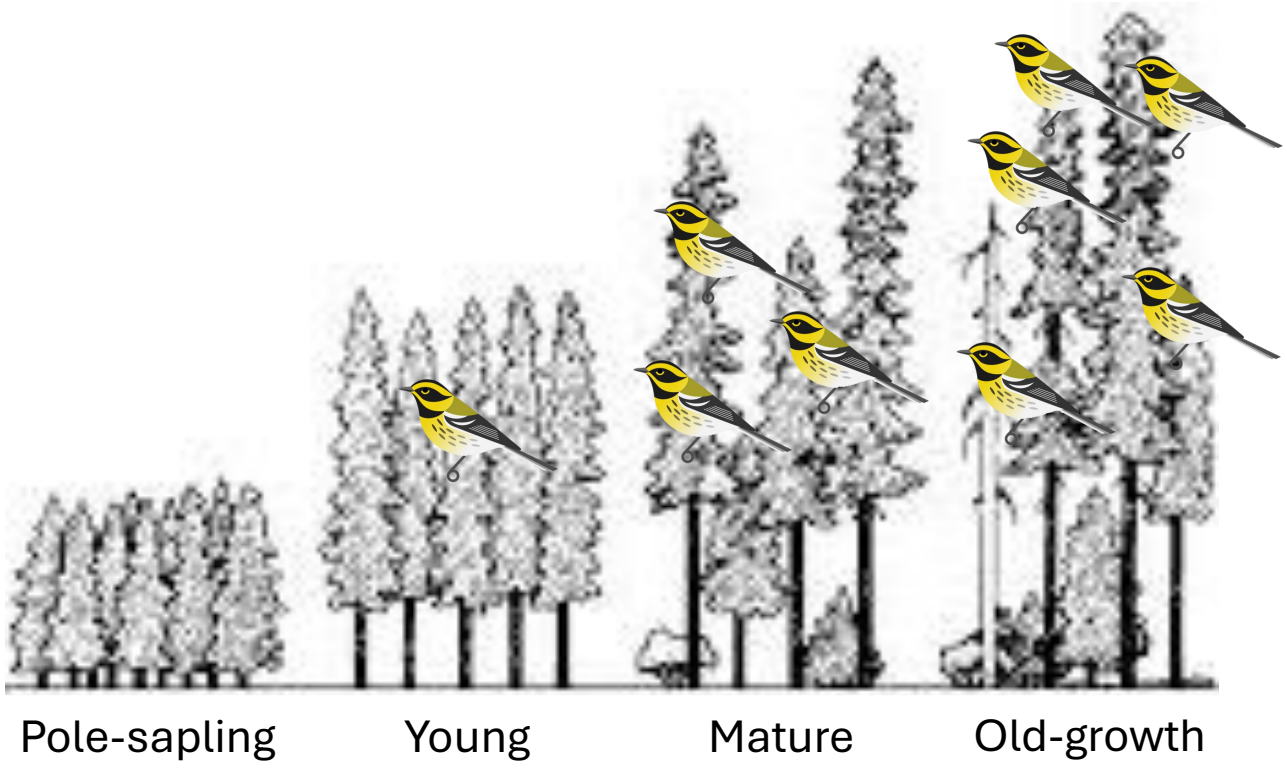
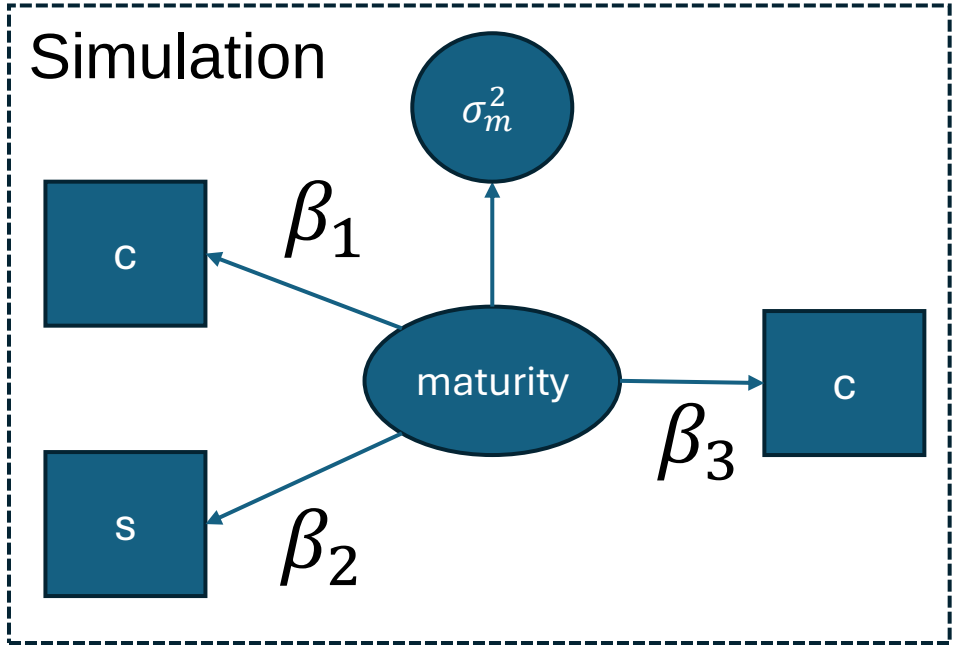
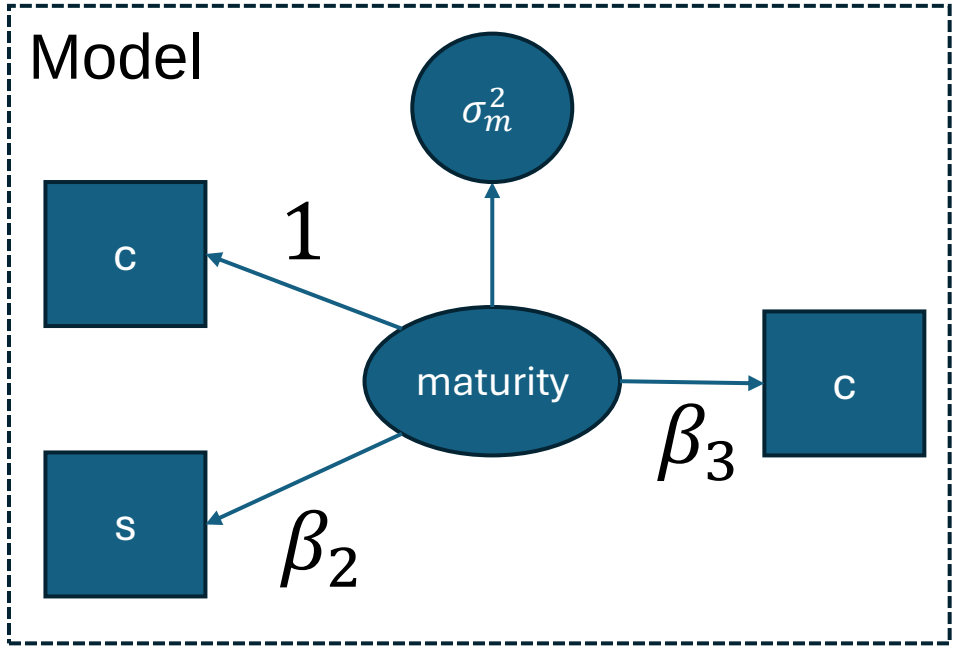
Latent variables: part II







What is a latent (or hidden) variable?

A random variable that is unmeasured but not necessarily unmeasureable.
-P Spirtes (2001)

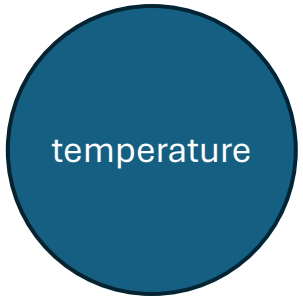
A variable that is not directly observable but is inferred from other variables that can be measured
-Generative AI (yesterday)

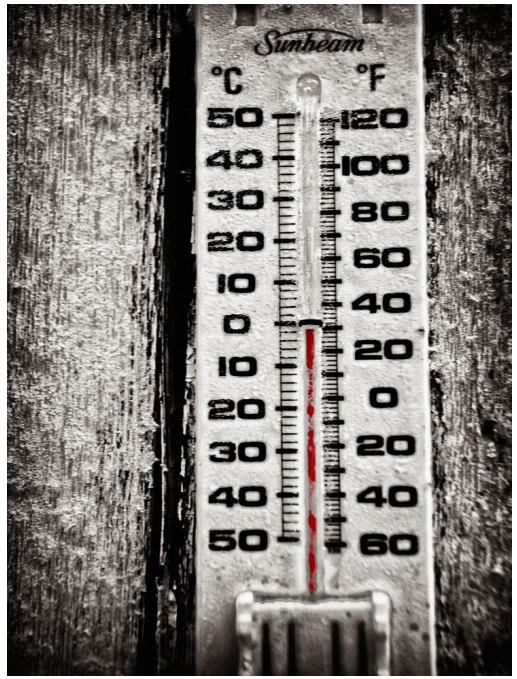
Variables that can only be inferred indirectly through a mathematical model from other observable variables
-Wikipedia (also yesterday)

What is a latent (or hidden) variable?

Everything is a latent variable **– LA Dyer**

Is temperature a latent variable?

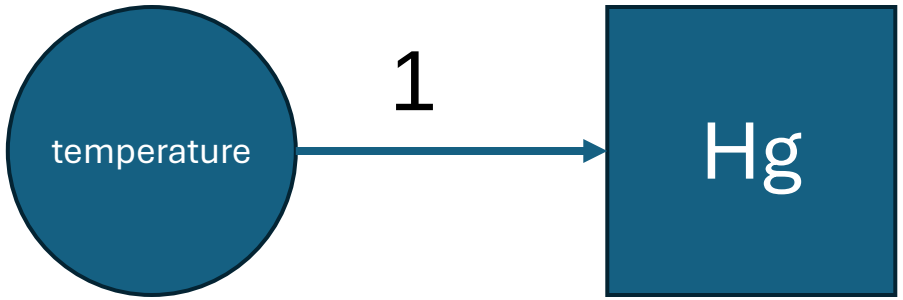




A random variable that is unmeasured but not necessarily unmeasureable.
-P Spirtes (2001)

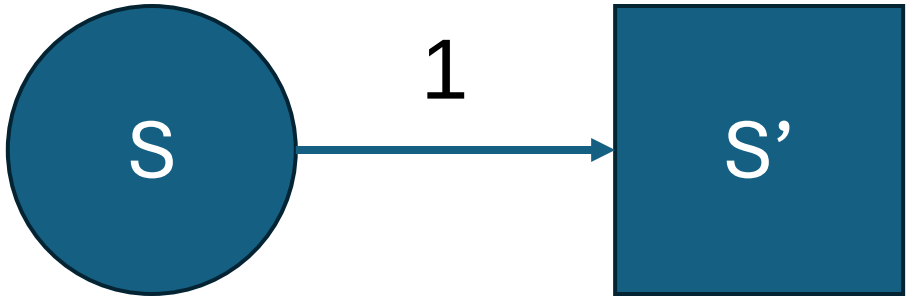
Temperature is the average kinetic energy of particles

Temperature is a latent variable



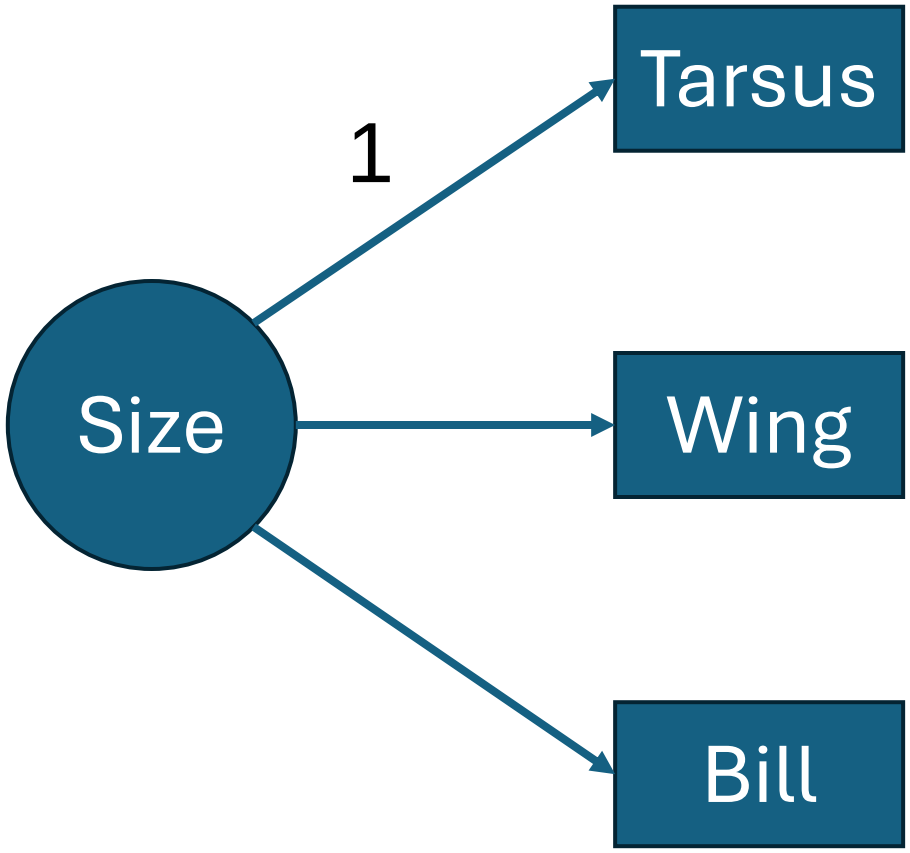
Temperature is the average kinetic energy of particles

Survival



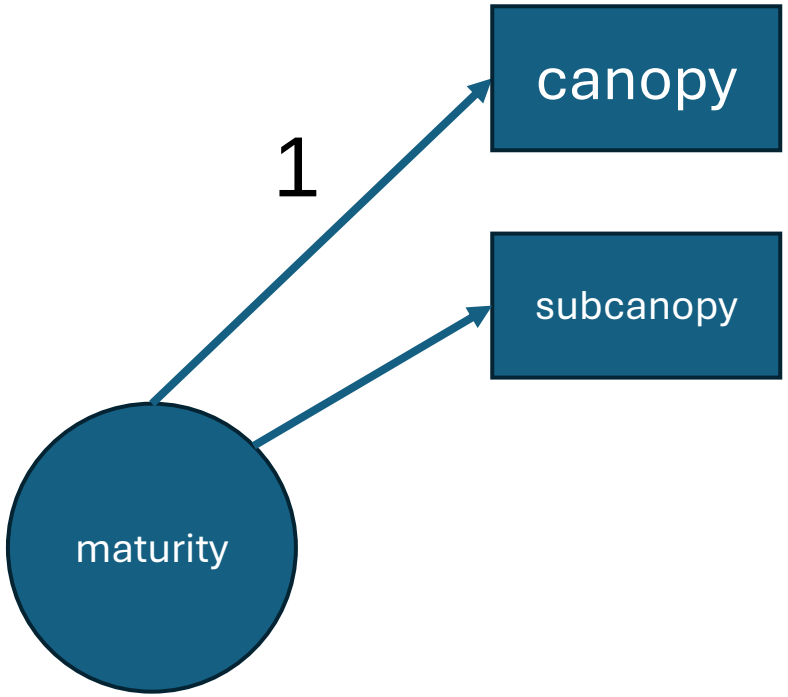
S: survival of a population, S’: survival of a marked sample

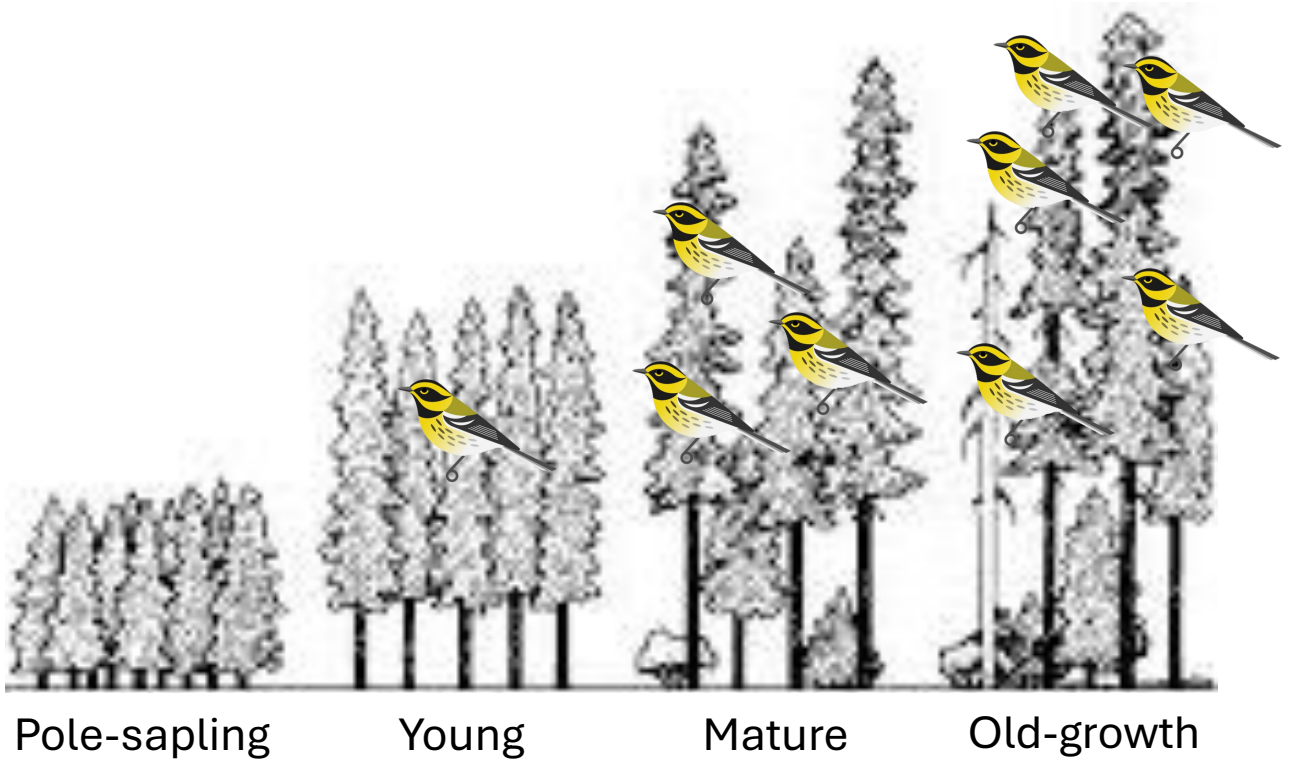
Size



‘Size’ is a human construct

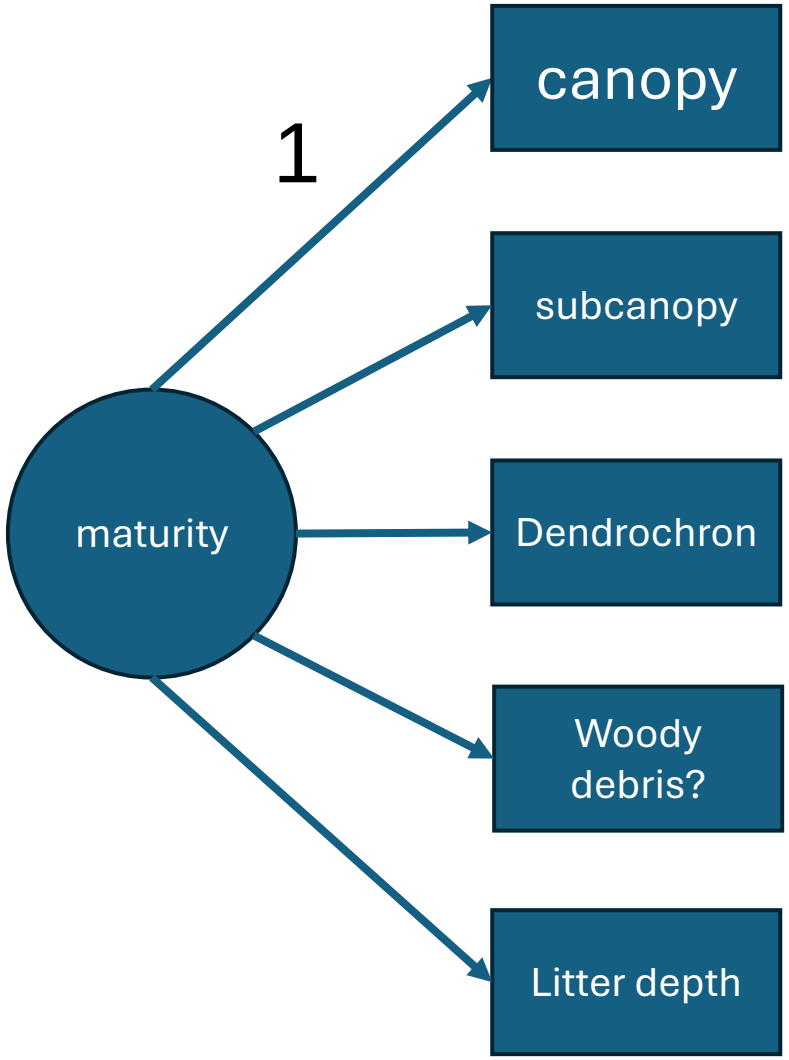
Forest maturity

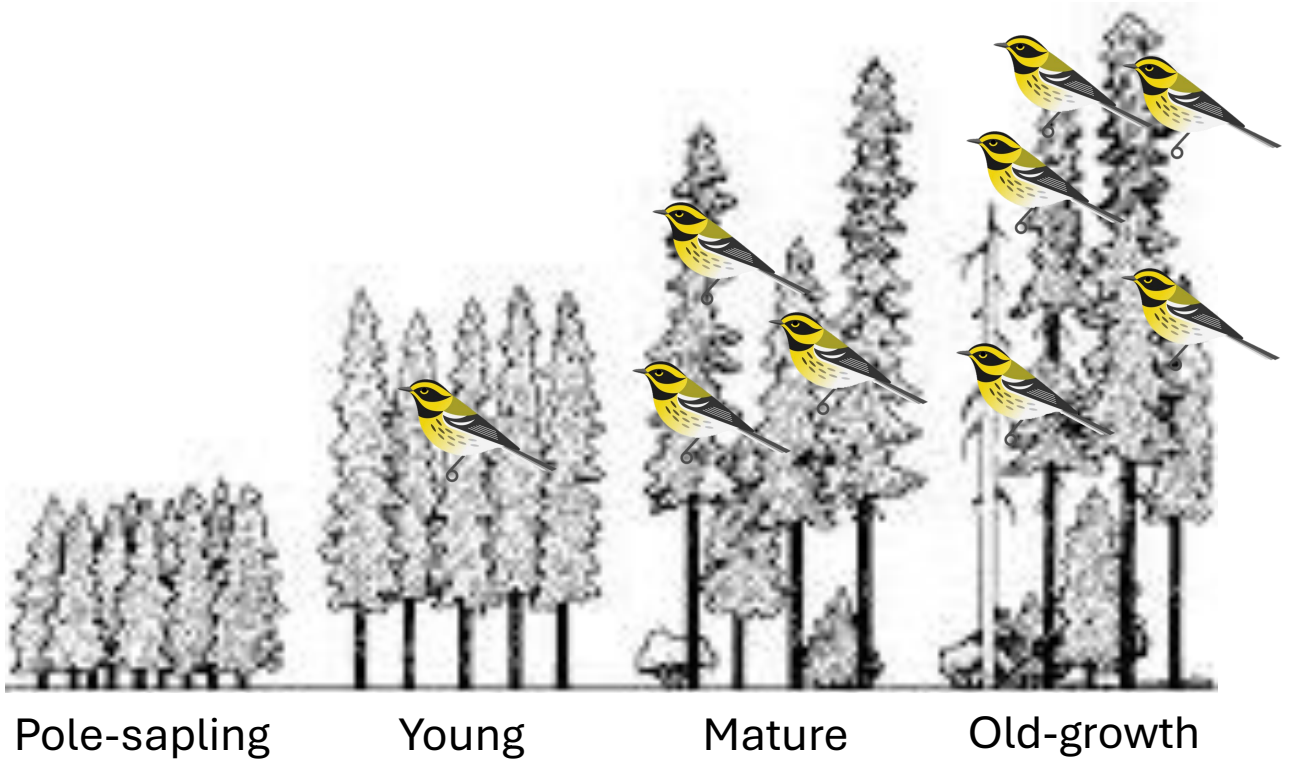




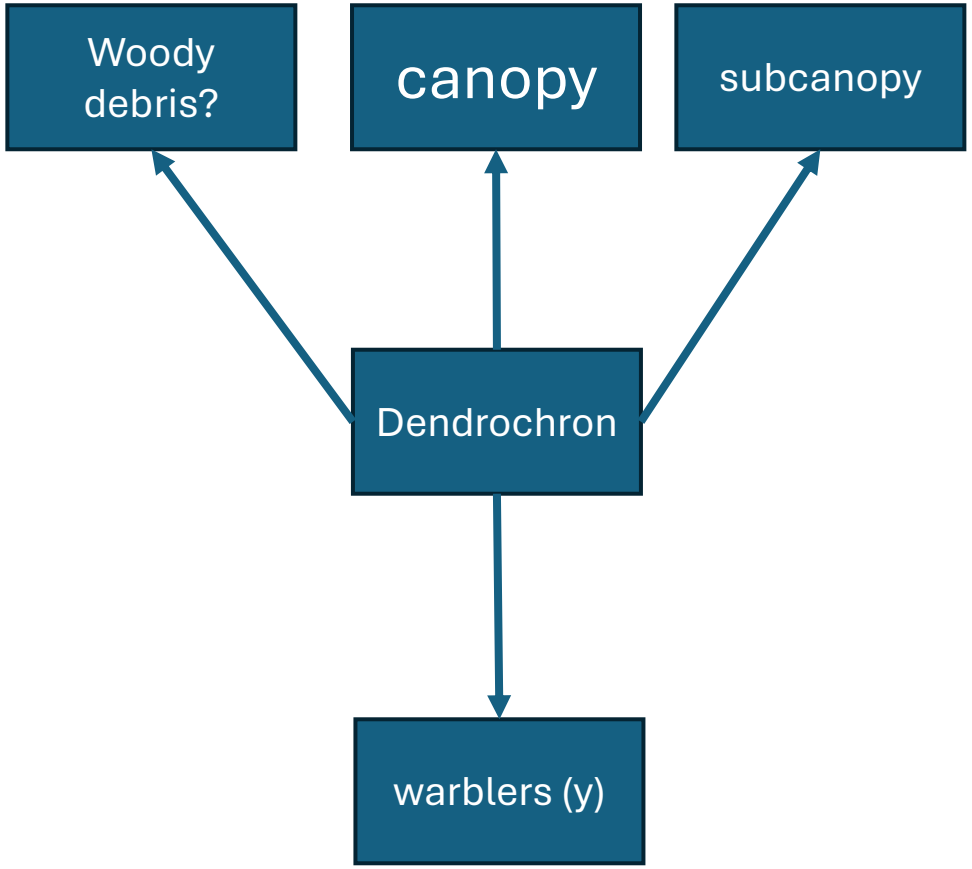
These ‘seral stages’ are human constructs

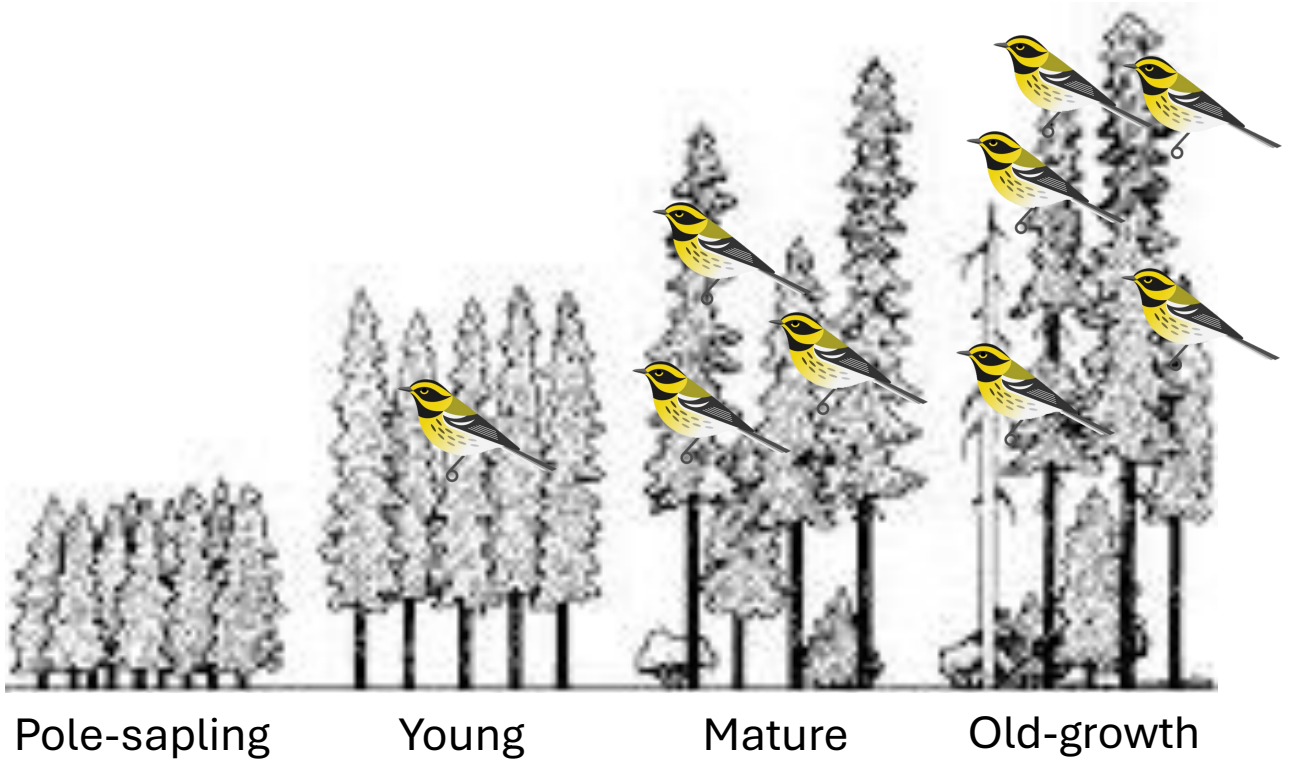
Forest maturity [expanded]





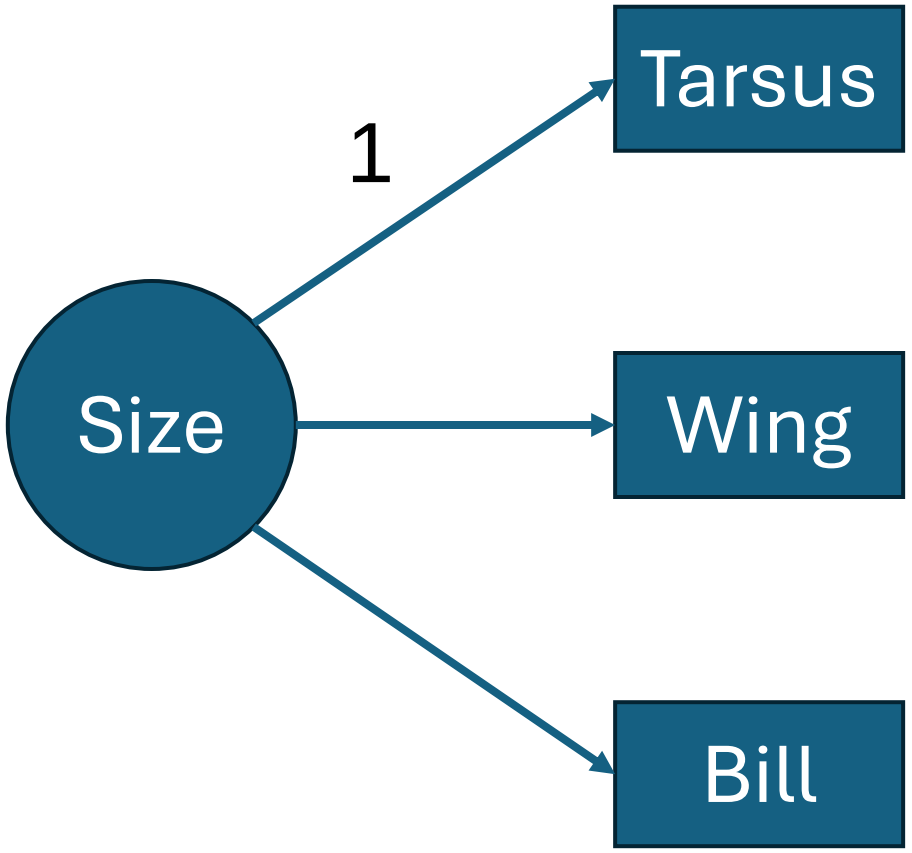
We could structure this differently





We don’t have to use latent variables!

They’re just very useful…



We can make some assumptions about our latent variable(s)

We generally assume they're normally distributed

$$\boldsymbol{m} \sim \text{normal}(\mu, \sigma_m^2)$$

We assume that they are zero-centered b/c they're human constructs

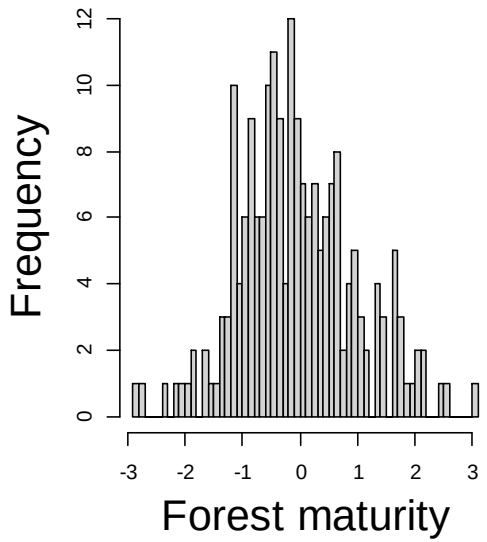
i.e., what should the scale of forest maturity be?

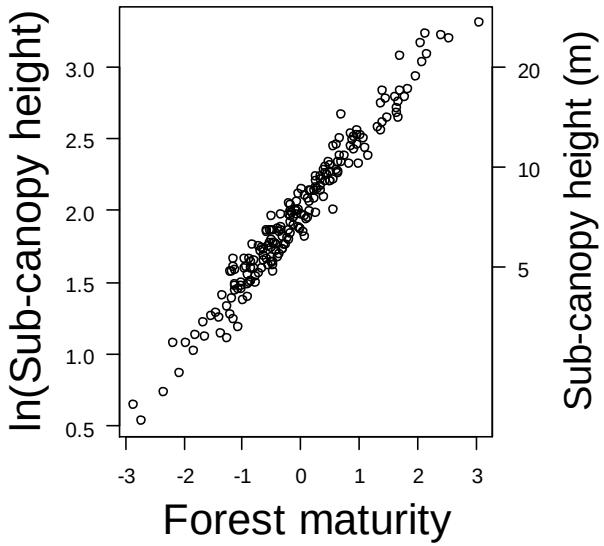
$$\boldsymbol{m} \sim \text{normal}(0, \sigma_m^2)$$

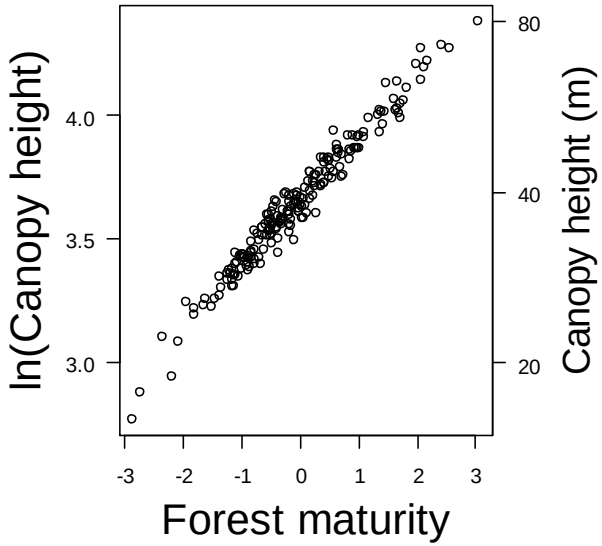
Assigning an intercept would be entirely subjective, plus the math is easier if μ = 0

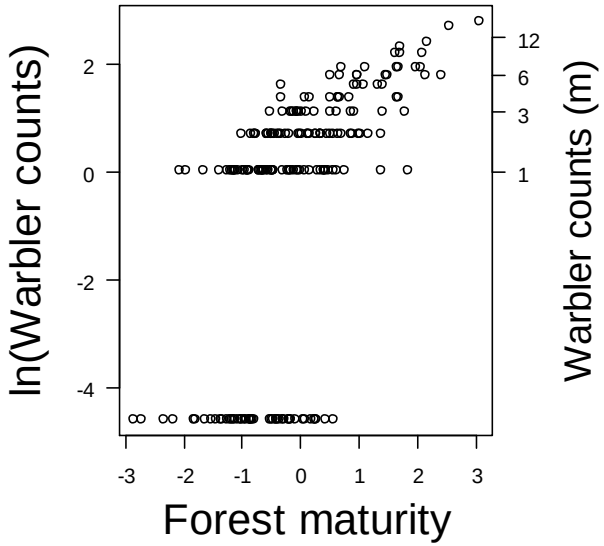
So let’s talk about this ‘fixing a loading to 1’ thing

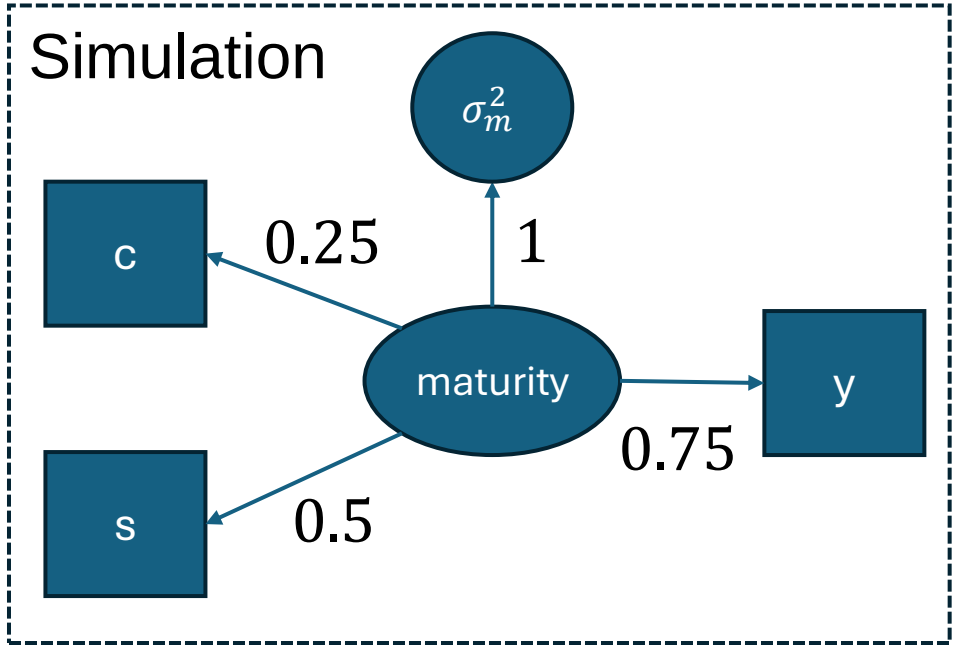
Let's simulate some data





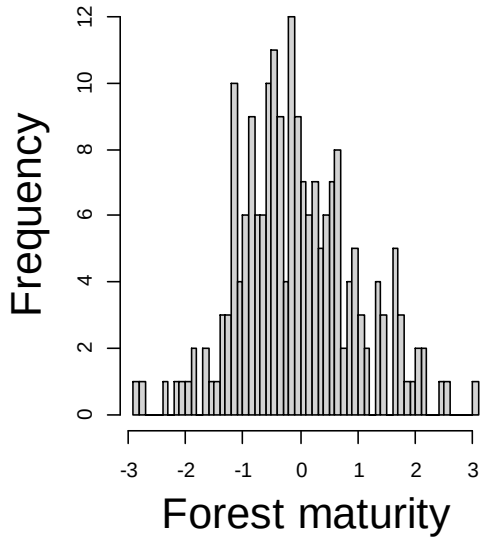


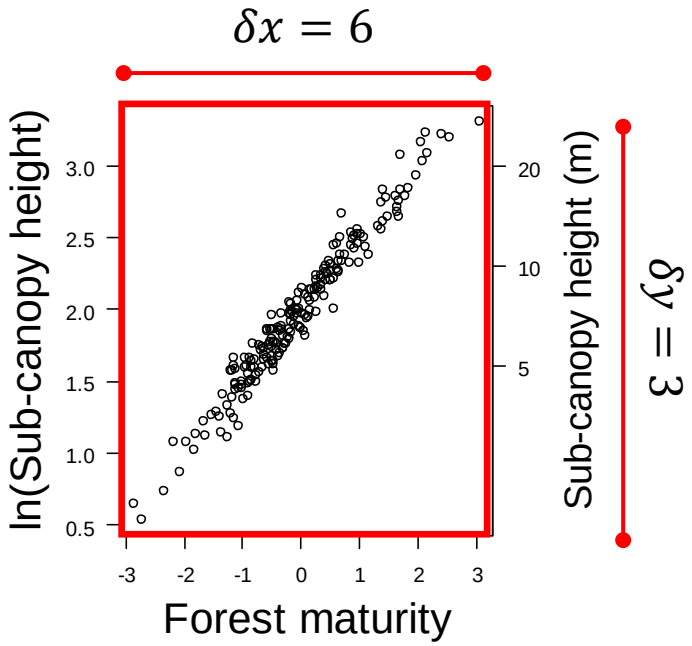


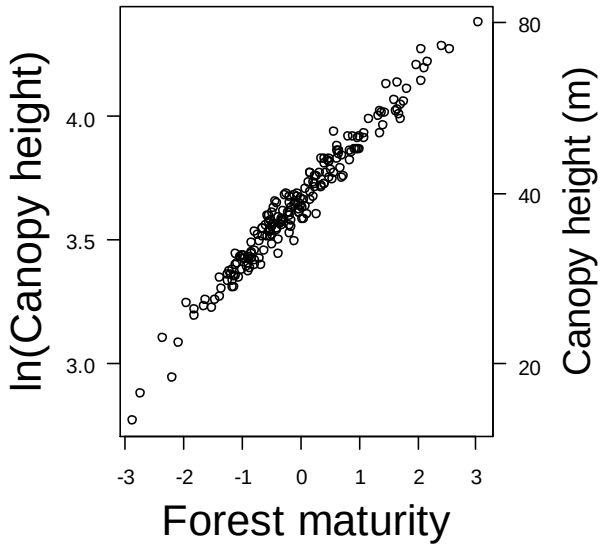


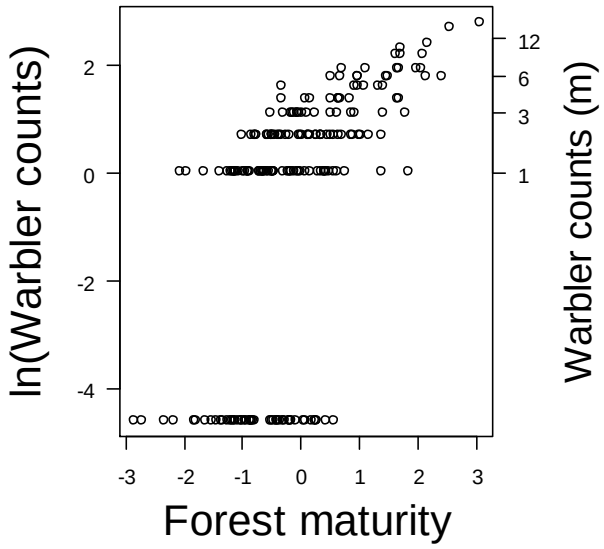
$$\delta y = \delta x \beta$$

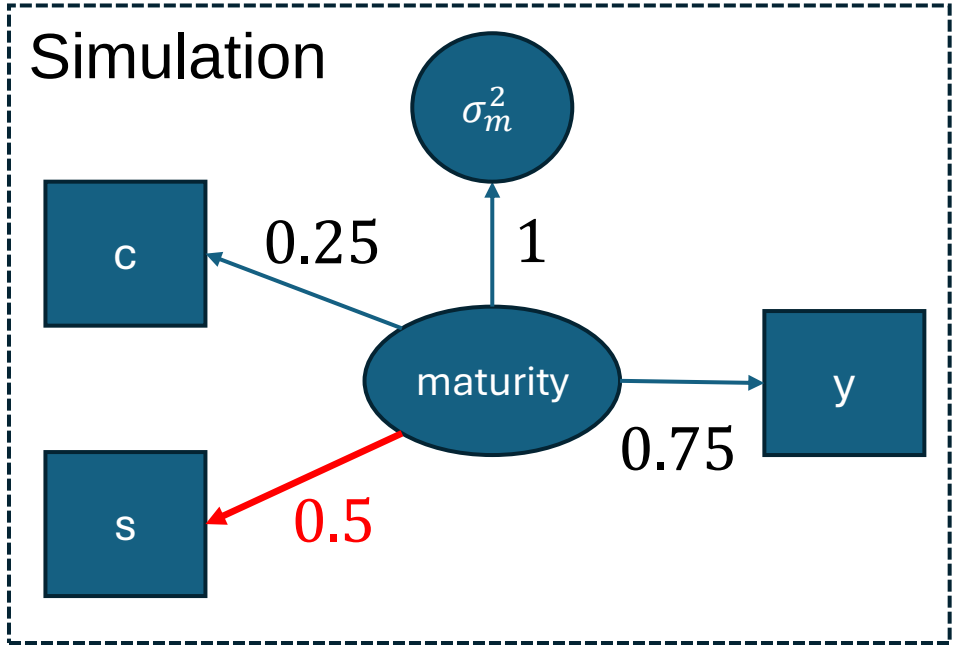
Clarifying a loading…









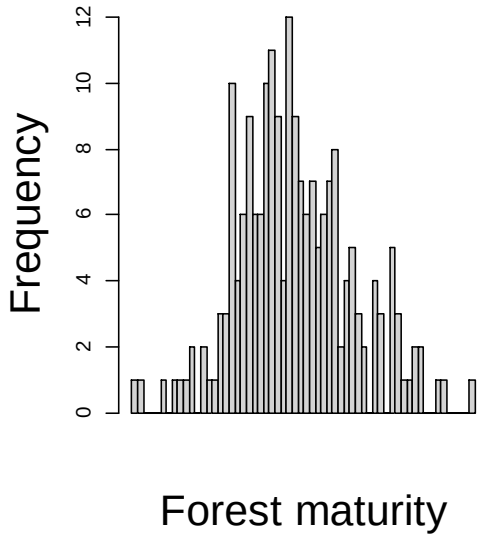


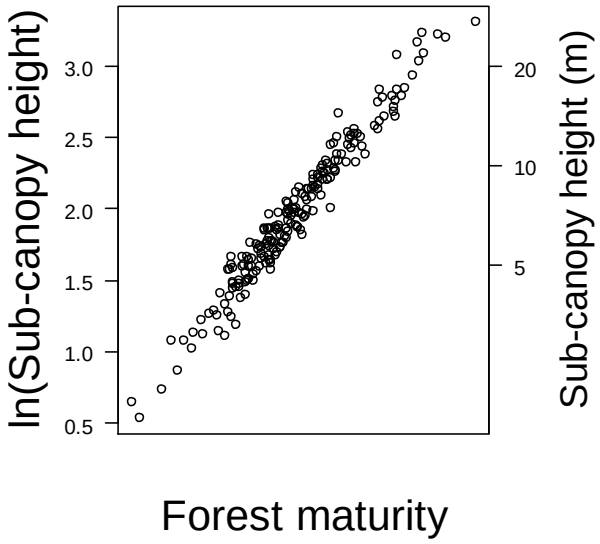
$$\delta y = \delta x \beta$$

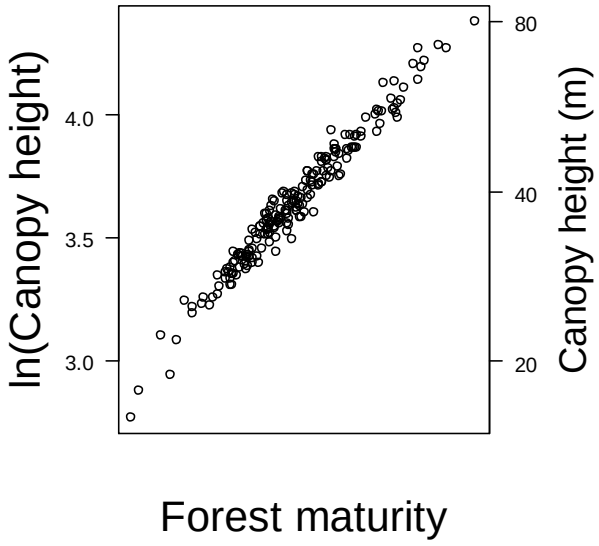
$$3 = 6\beta$$

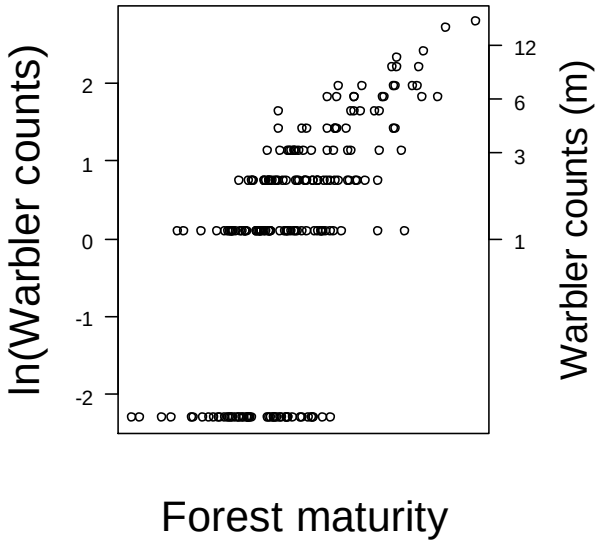
Now, let’s build a model

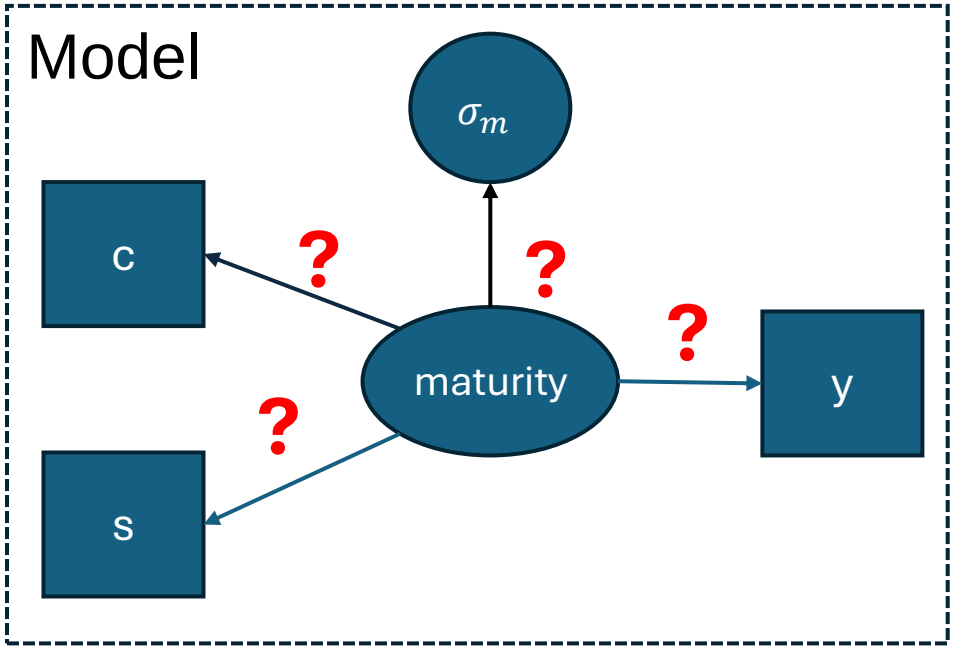
Shoot…







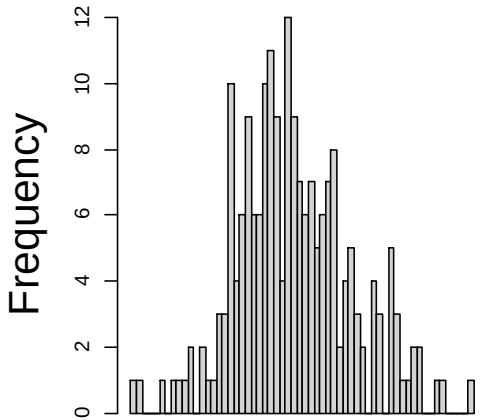




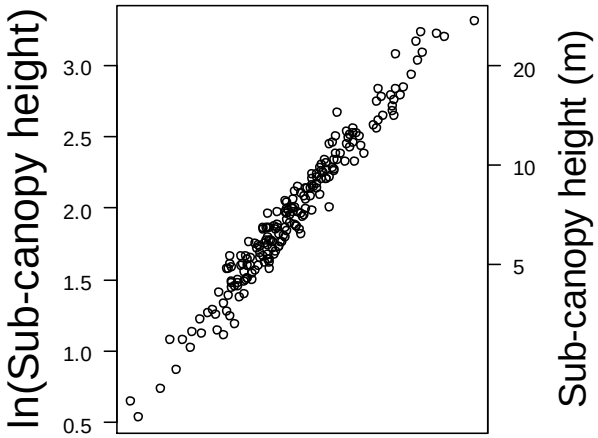
There's a big problem:
We don't know the range of maturity

$$\boldsymbol{m} \sim \text{normal}(0, \sigma_m^2)$$

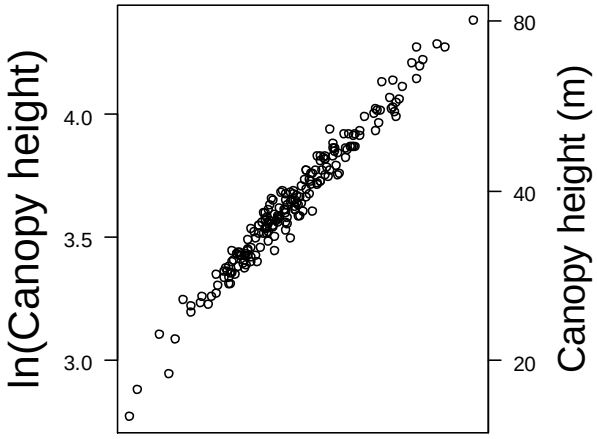
Small groups!



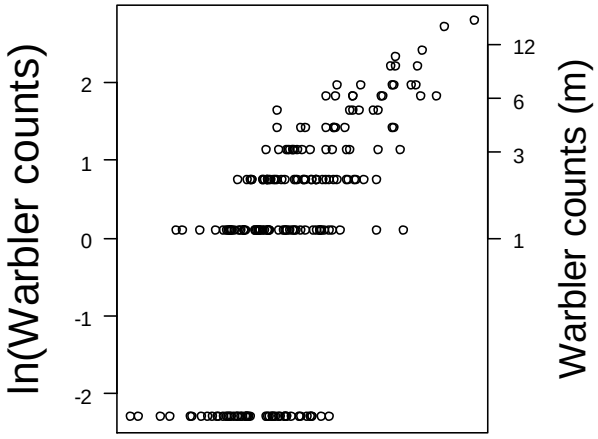
Forest maturity



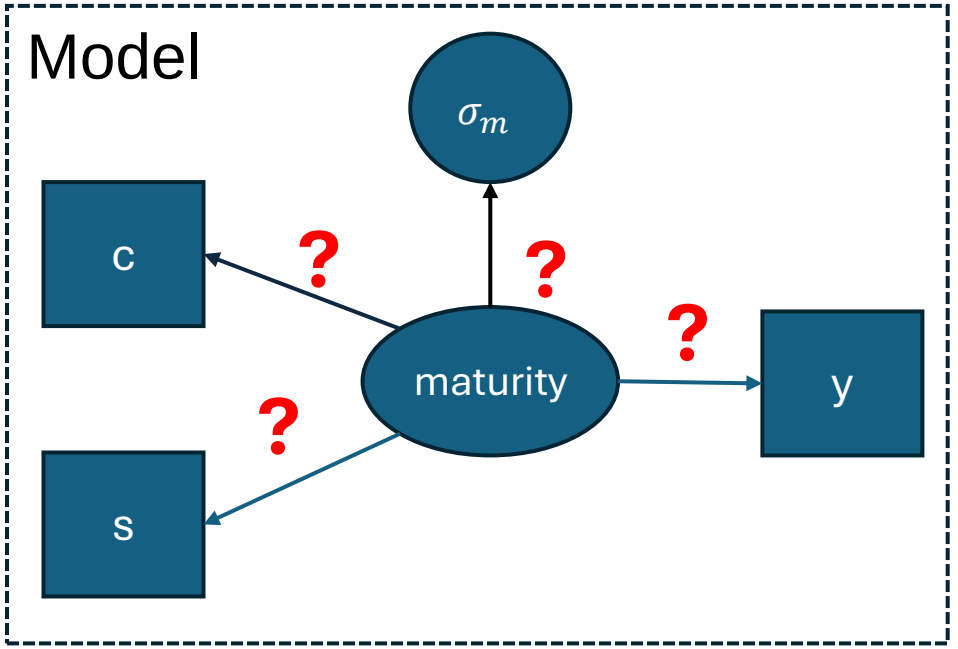
Forest maturity



Forest maturity



Forest maturity

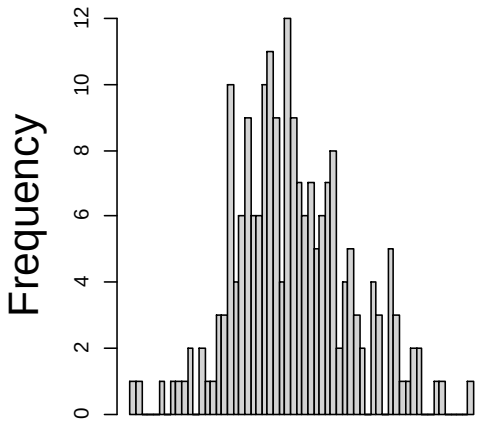


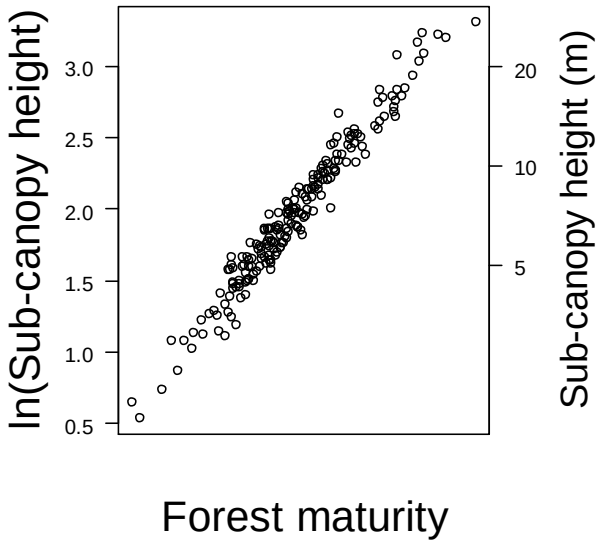
How could we scale maturity, i.e., what should the minimum and maximum values of maturity be?

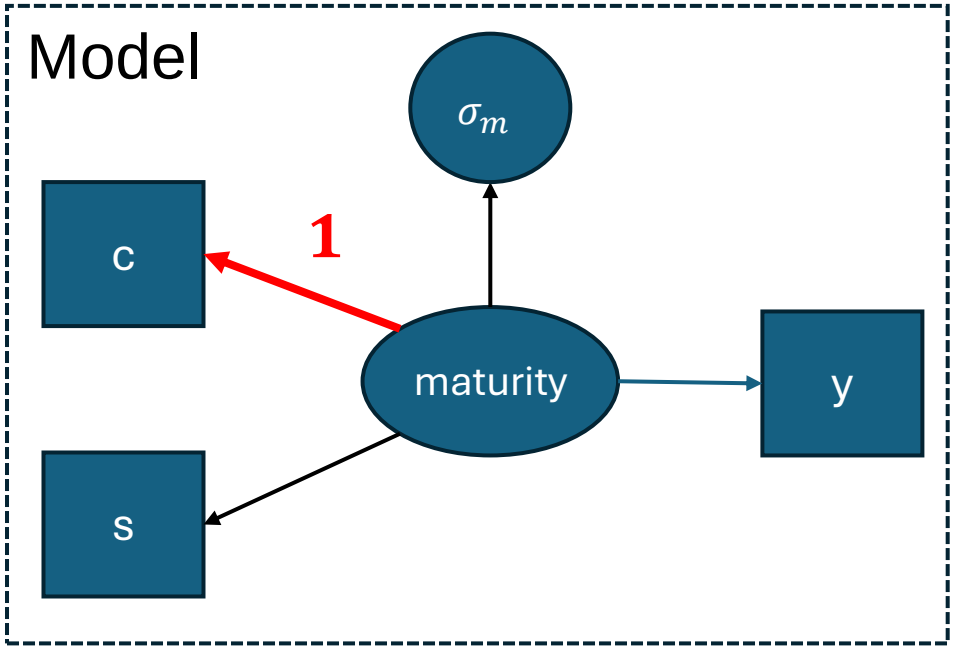
$$\delta y = \delta x \beta$$

$$\boldsymbol{m} \sim \text{normal}(0, \sigma_m^2)$$

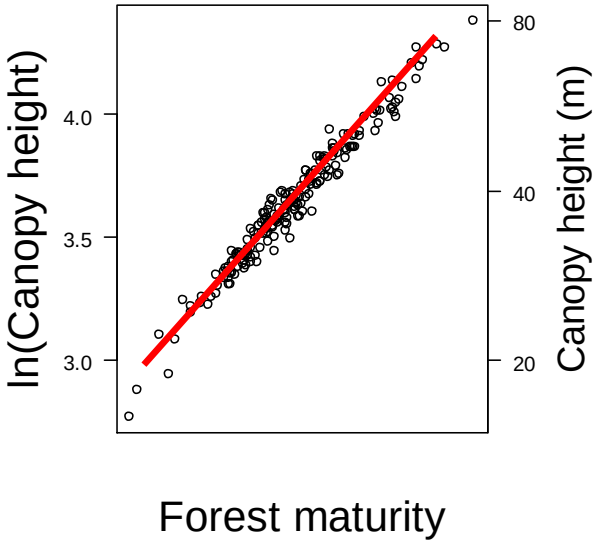
What if we fix a beta?

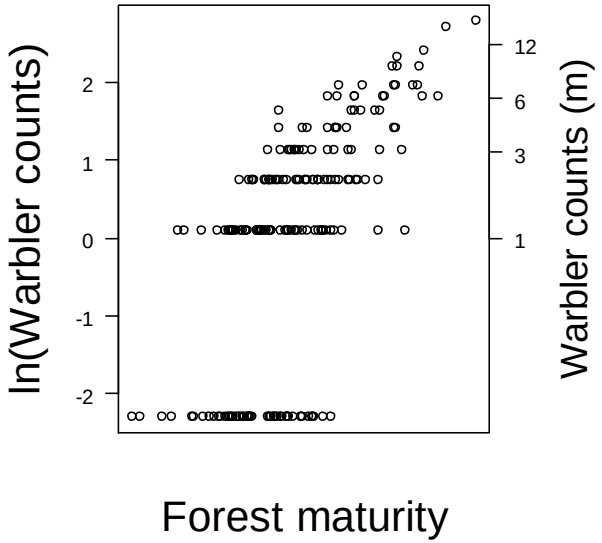






$\delta y = 1.5$



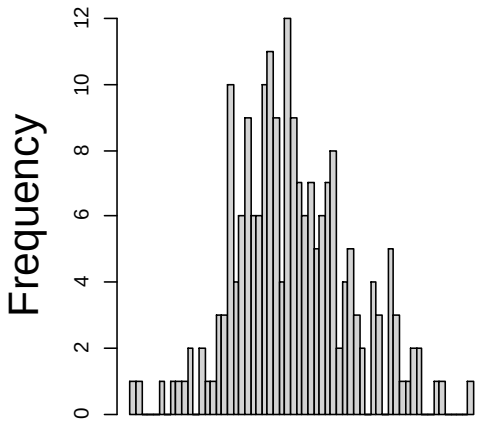


$$\delta y = \delta x \beta$$

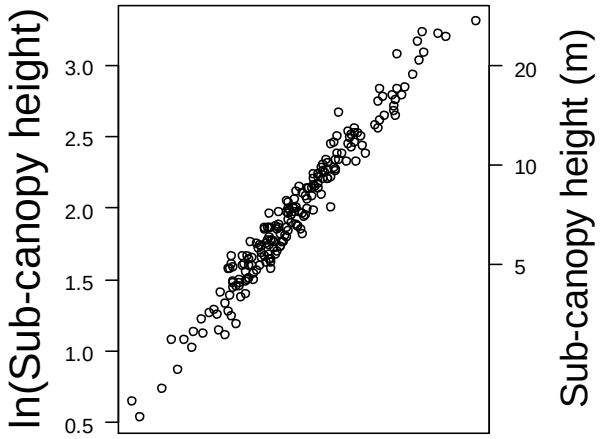
$$1.5 = ? \times 1$$

$$\boldsymbol{m} \sim \text{normal}(0, \sigma_m^2)$$

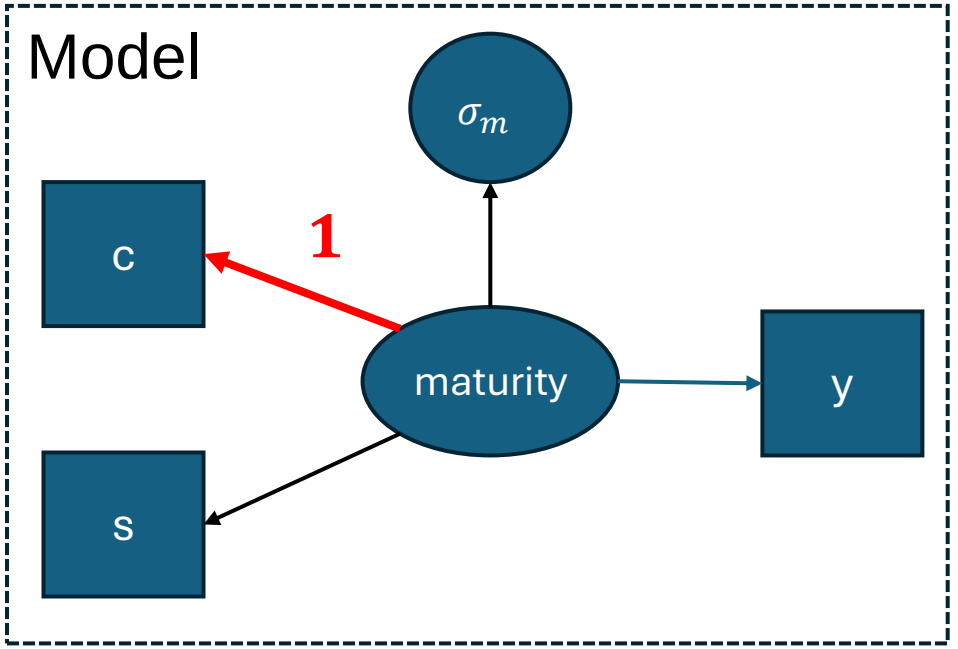
Now we can calculate σ_m



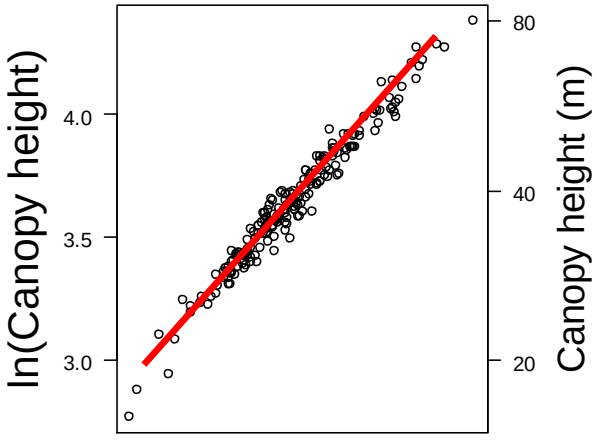
Forest maturity



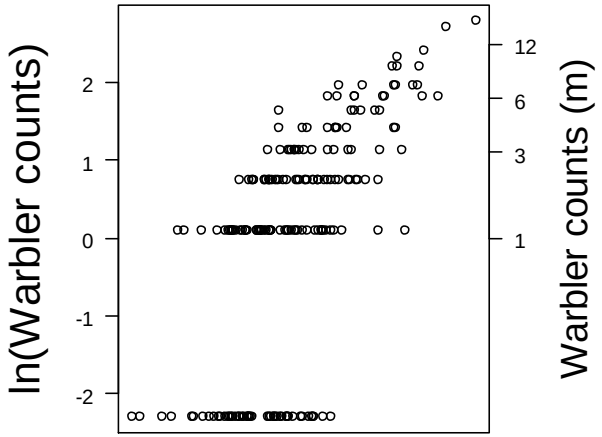
Forest maturity



$\delta y = 1.5$



Forest maturity



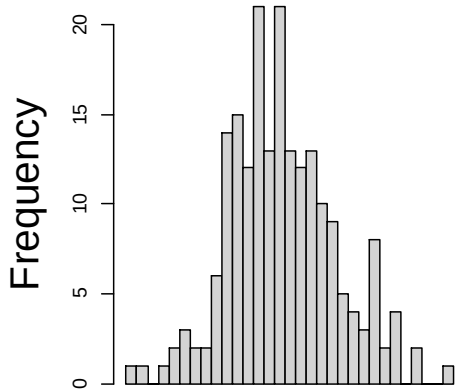
Forest maturity

$$\delta y = \delta x \beta$$

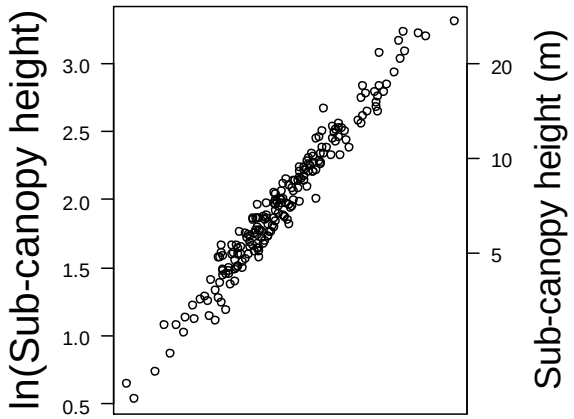
$$1.5 = ? \times 1$$

Our x-axis must have a range of 1.5!

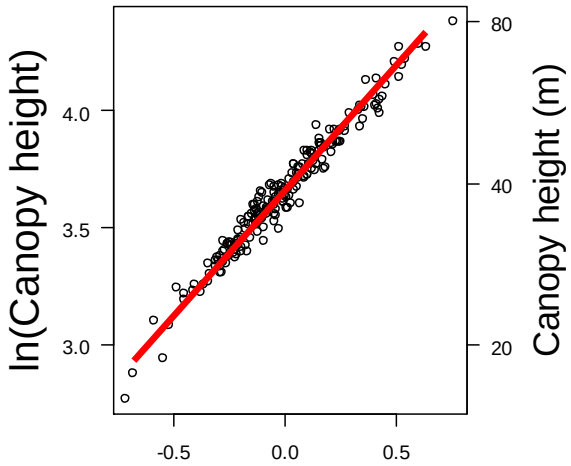
Now we have a scale!!



Forest maturity



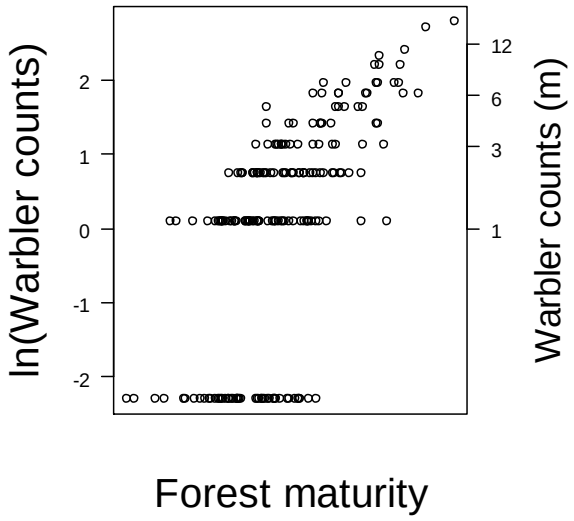
Forest maturity



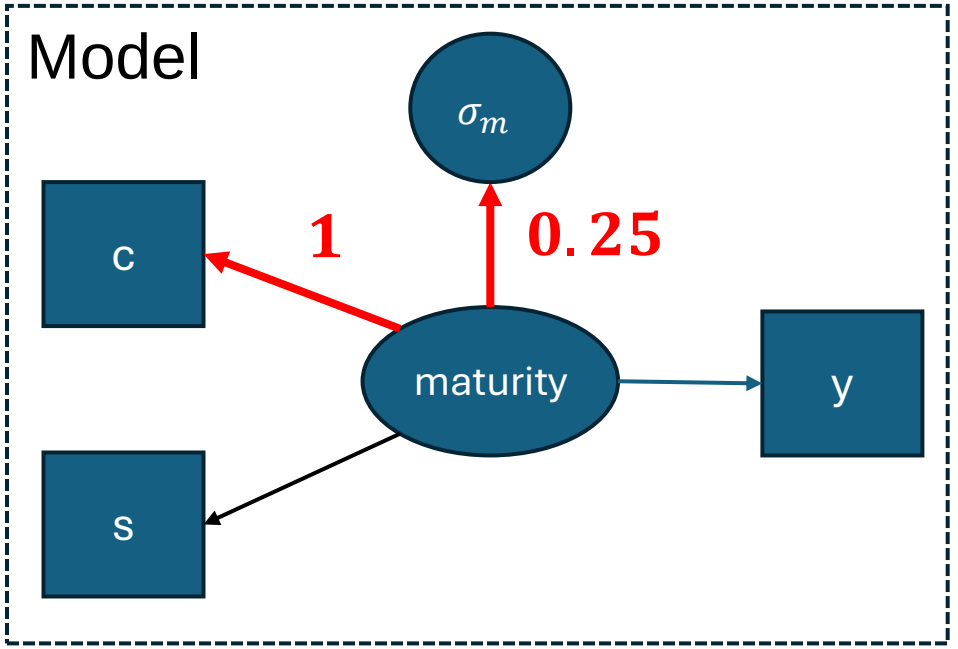
Forest maturity

$\delta y = 1.5$

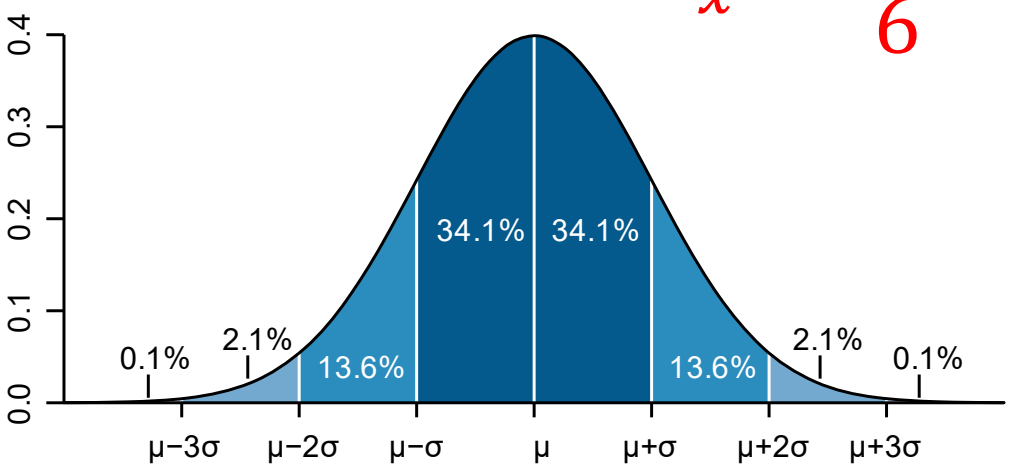
$\delta x = 1.5$



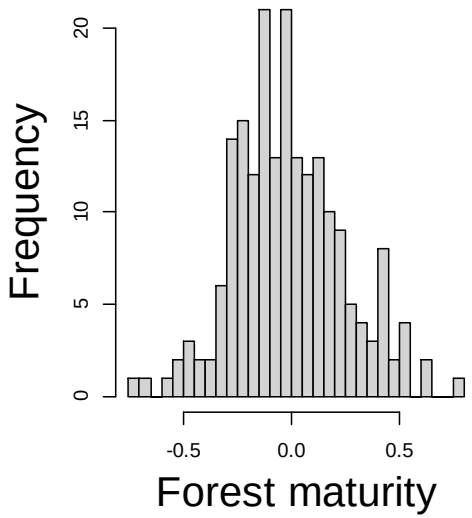
Forest maturity

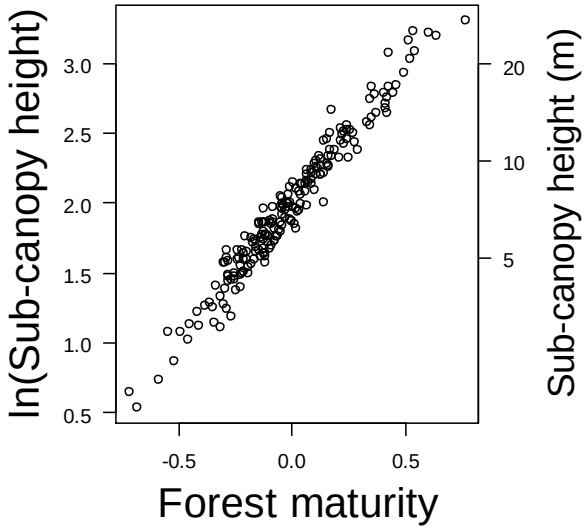


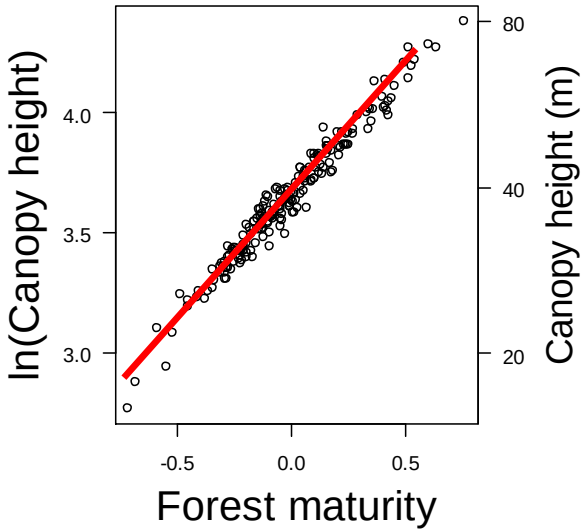
$$\sigma_x \cong \frac{\delta x}{6}$$



Now we know the x-axis for everything!

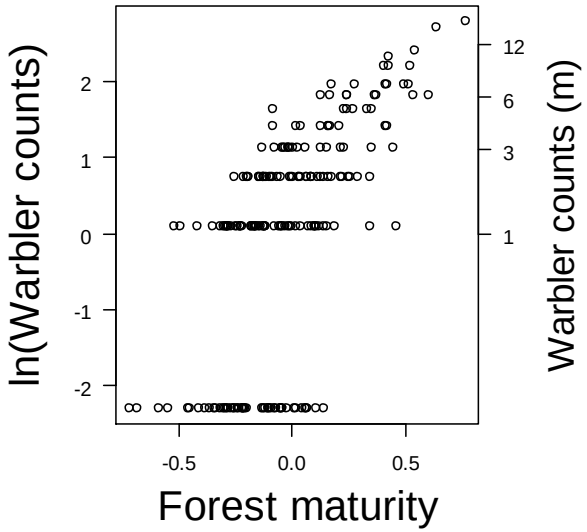


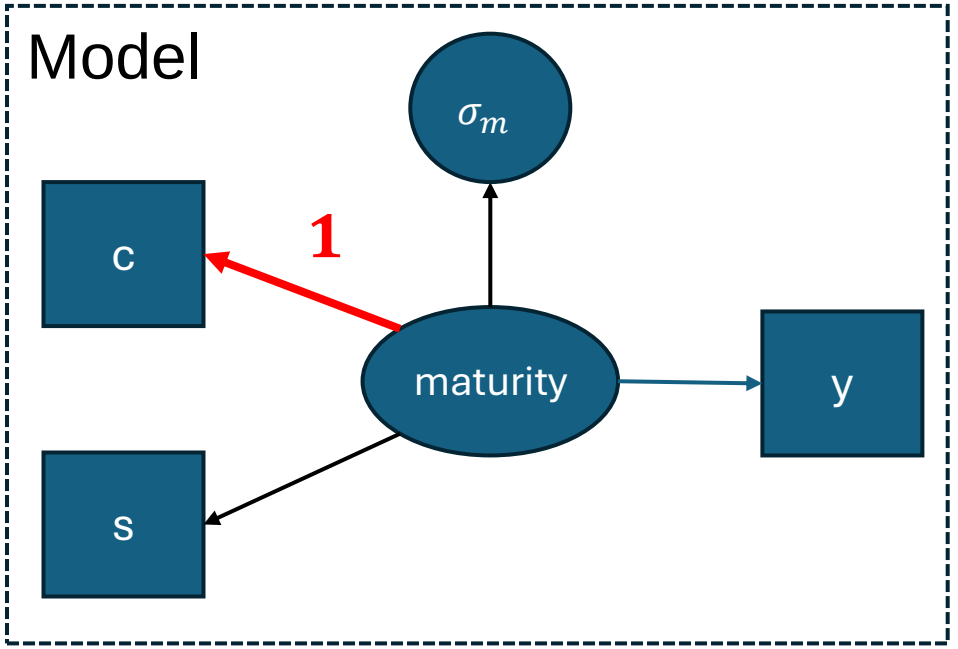




$\delta y = 1.5$

$\delta x = 1.5$

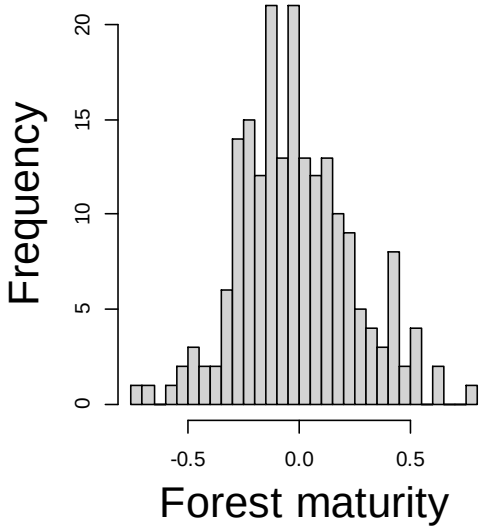


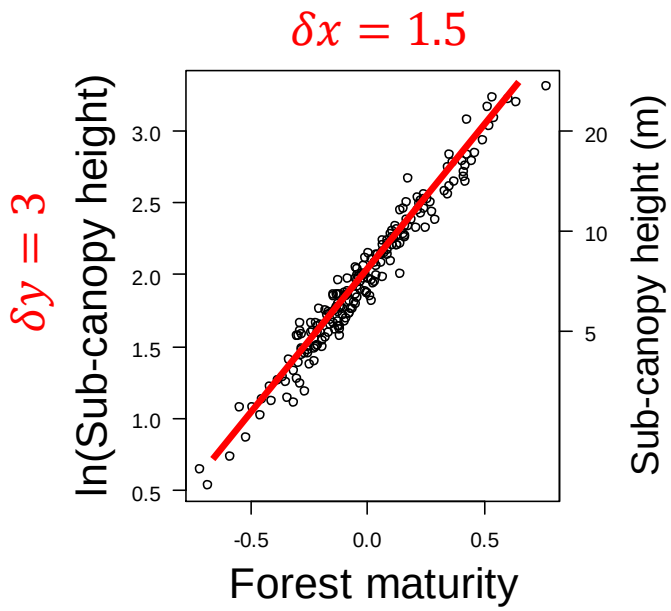


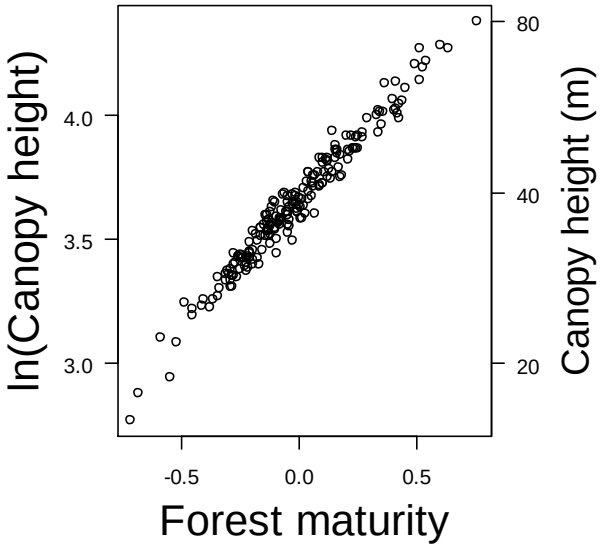
$$\delta y = \delta x \beta$$

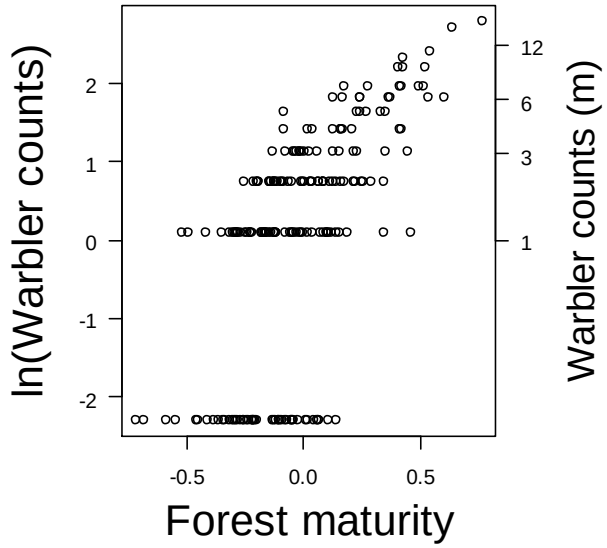
$$1.5 = 1.5 \times 1$$

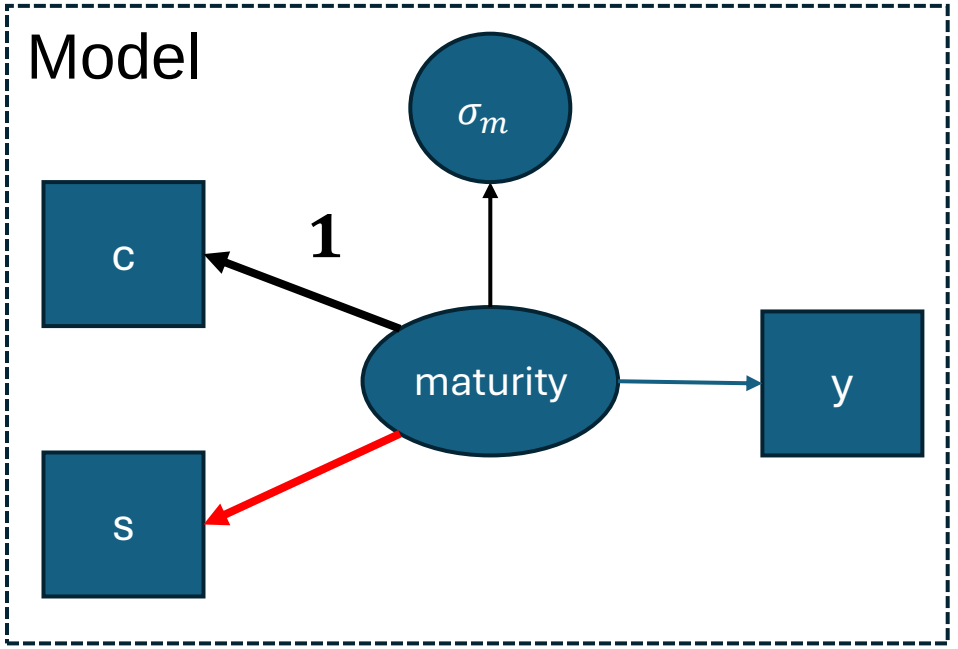
So what's the effect of *m* on *s*?







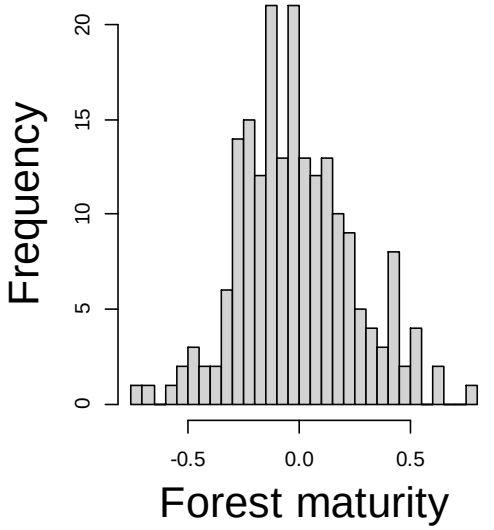


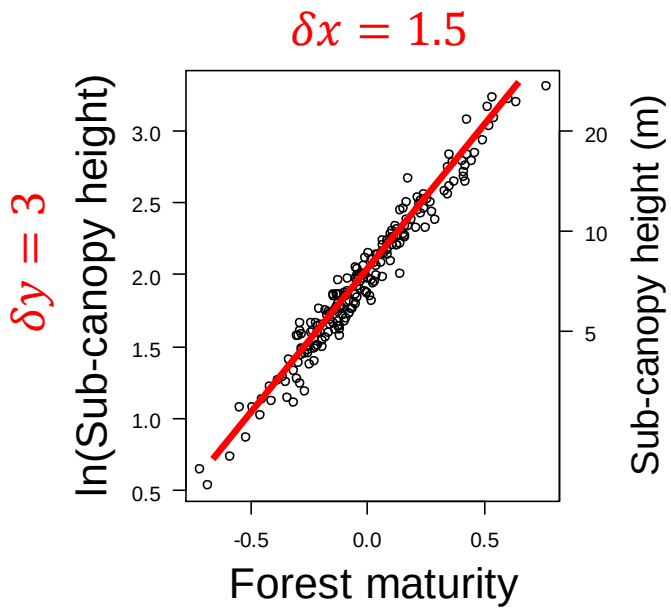


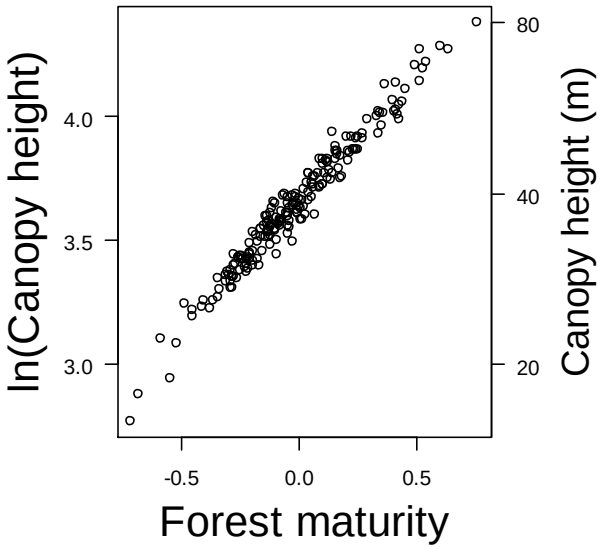
$$\delta y = \delta x \beta$$

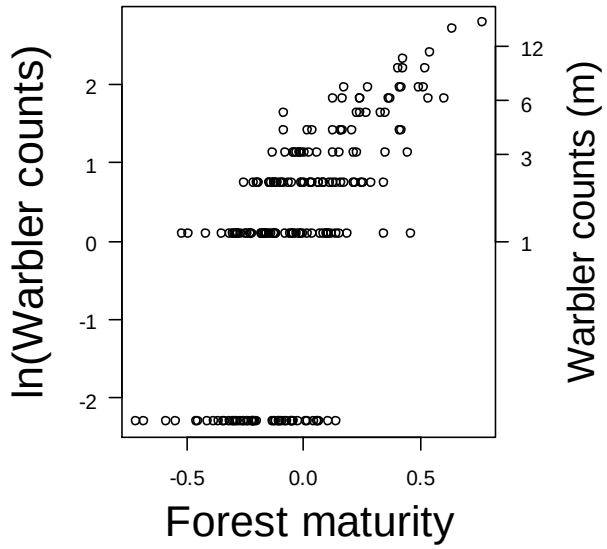
$$3 = 1.5 \times ?$$

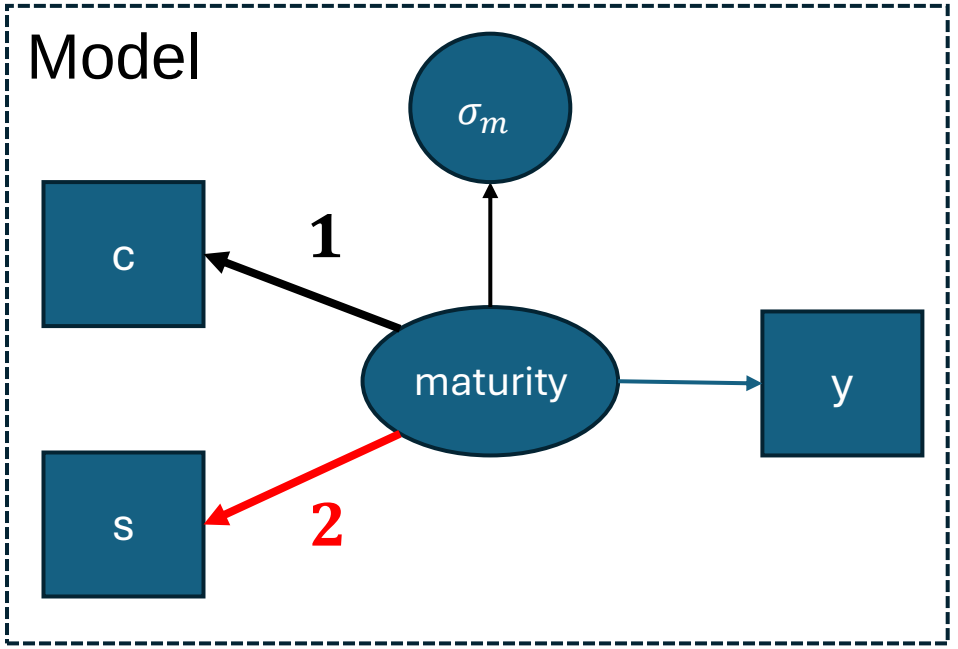
So what's the effect of *m* on *s*







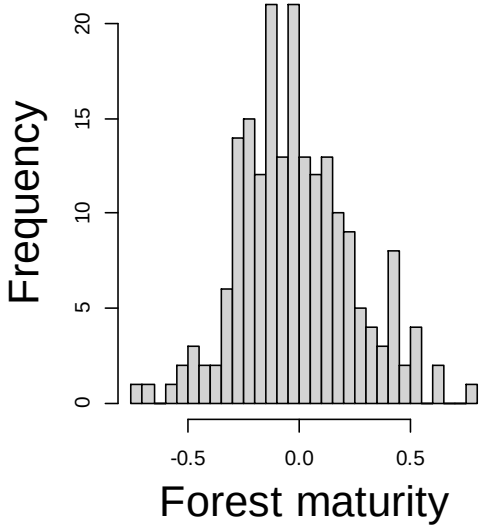


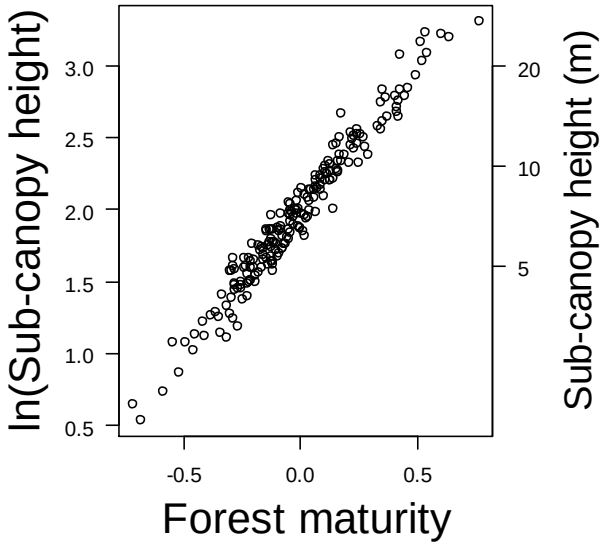


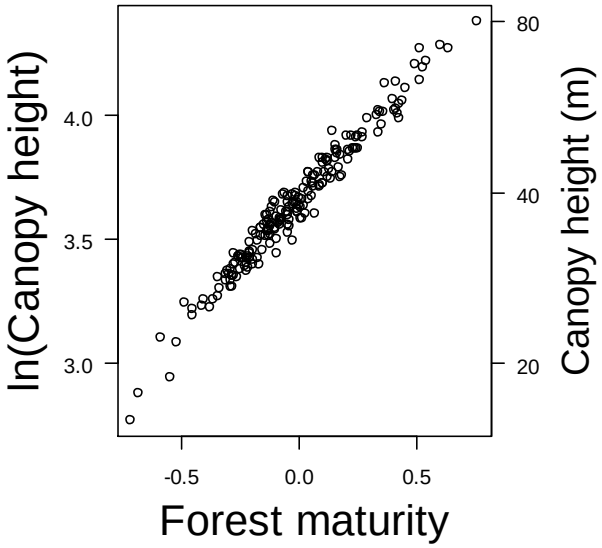
$$\delta y = \delta x \beta$$

$$3 = 1.5 \times 2$$

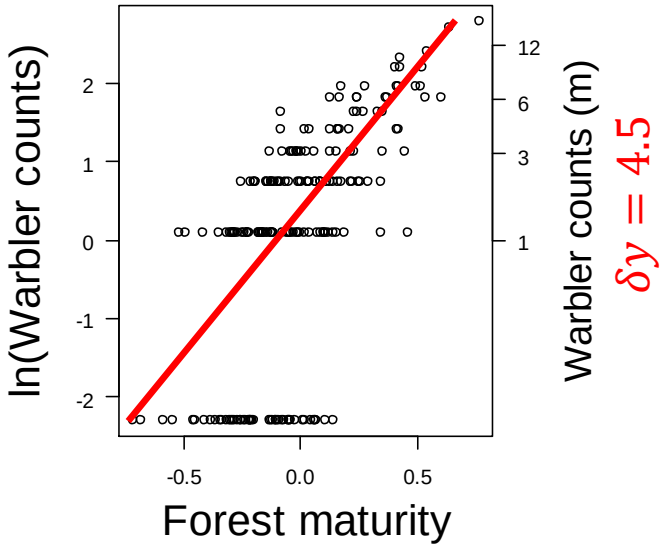
What about effect of *m* on *y*?



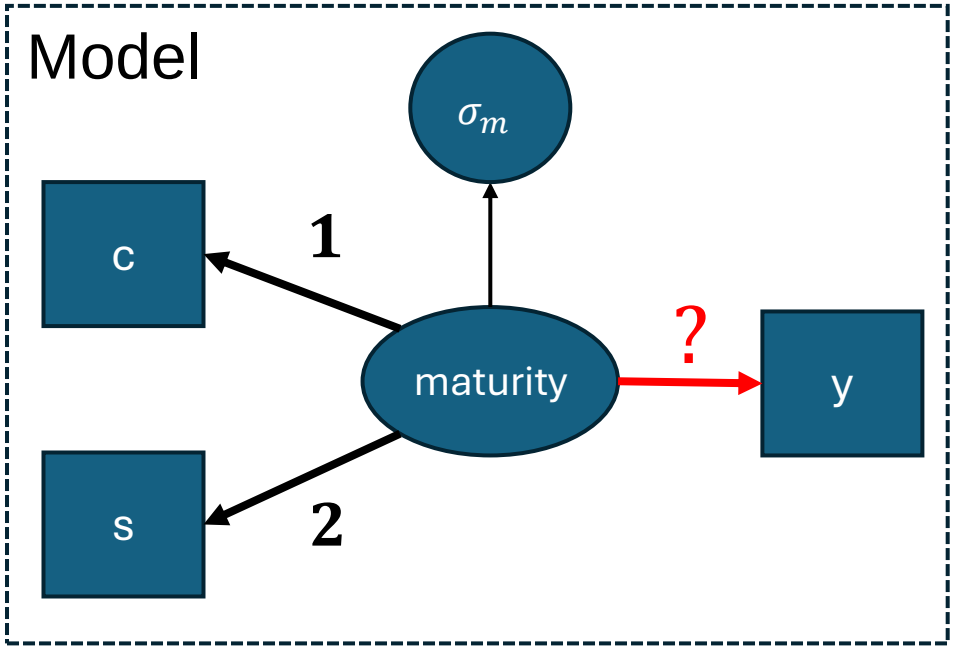




$\delta x = 1.5$



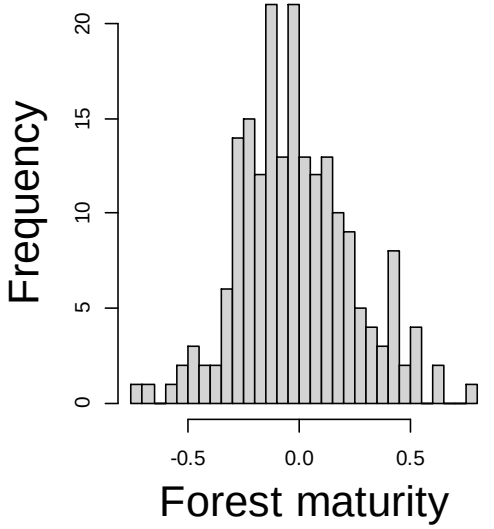
$\delta y = 4.5$

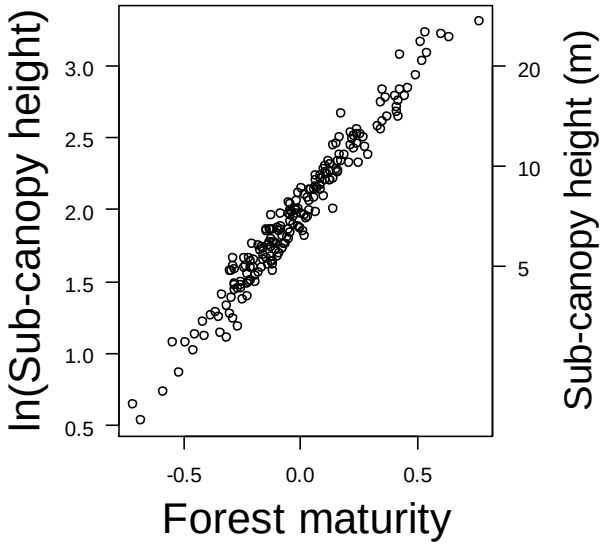


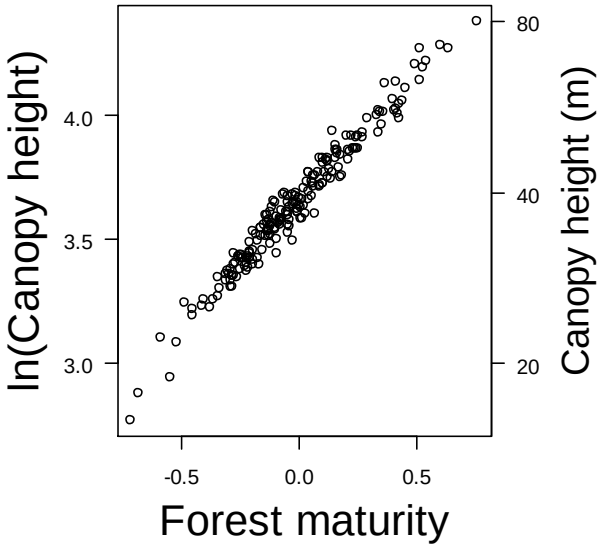
$$\delta y = \delta x \beta$$

$$4.5 = 1.5 \times ?$$

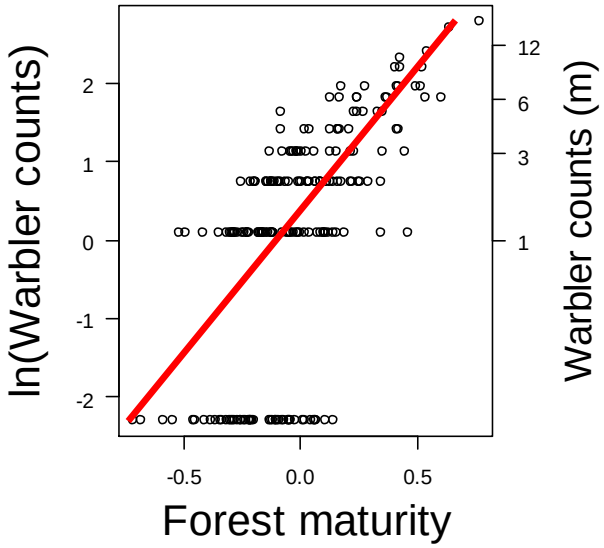
What about effect of *m* on *y*?



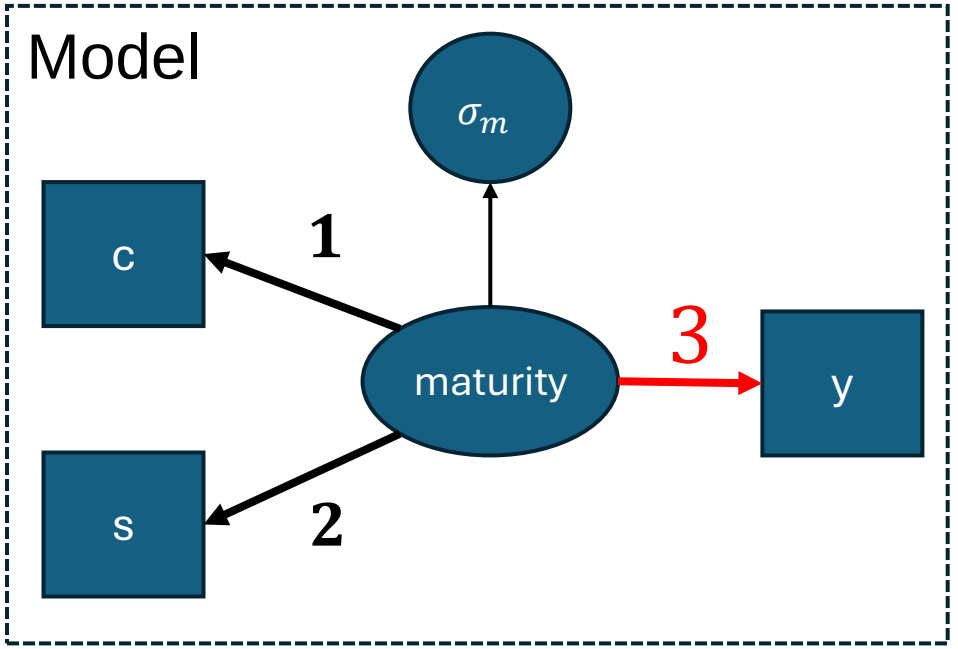




$\delta x = 1.5$



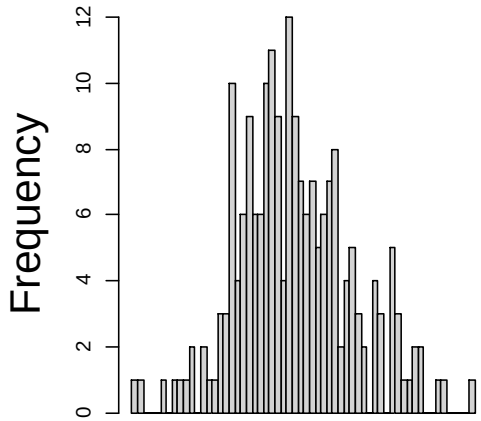
$\delta y = 4.5$

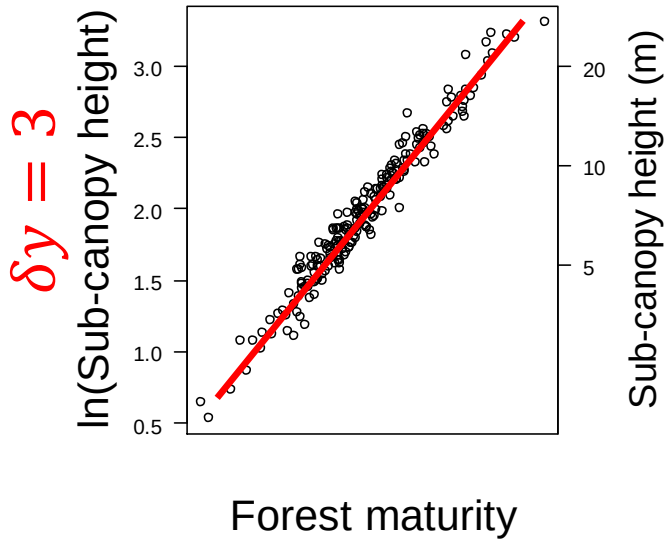


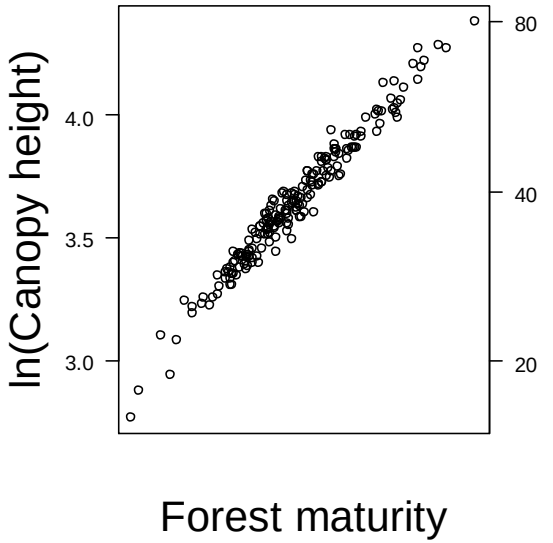
$$\delta y = \delta x \beta$$

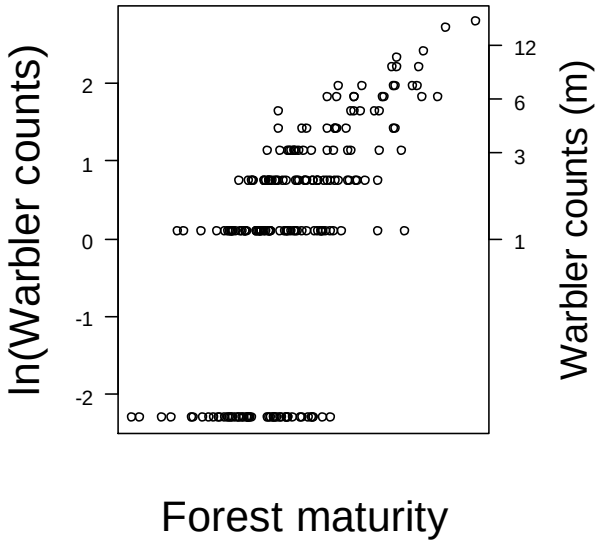
$$4.5 = 1.5 \times 3$$

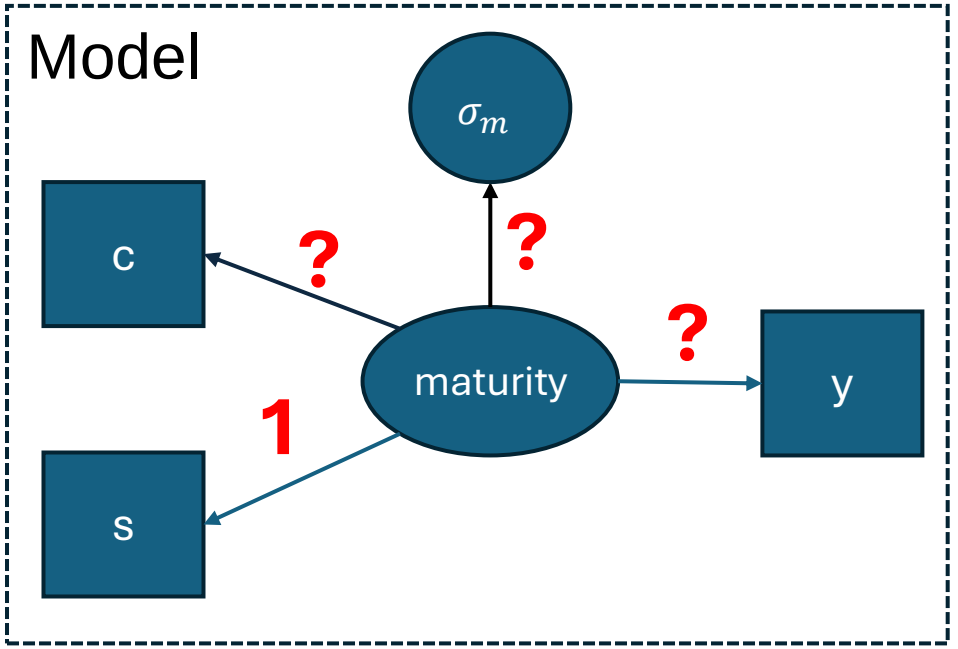
Let's do it again!!





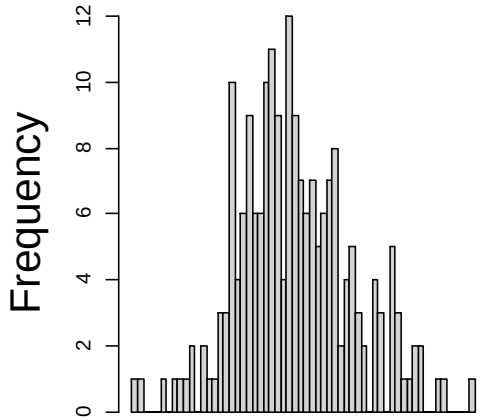


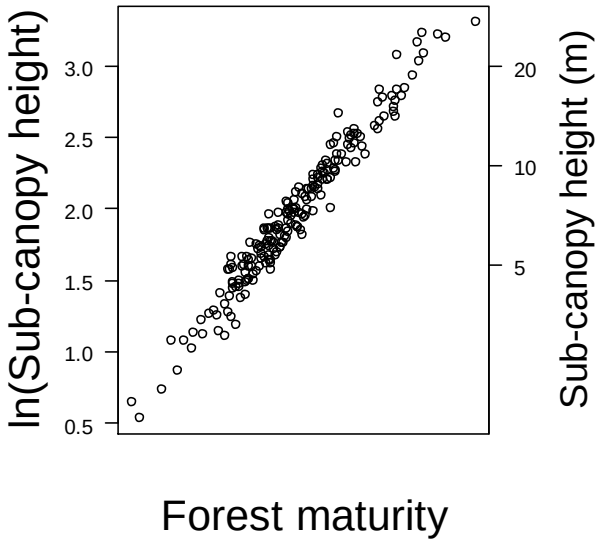


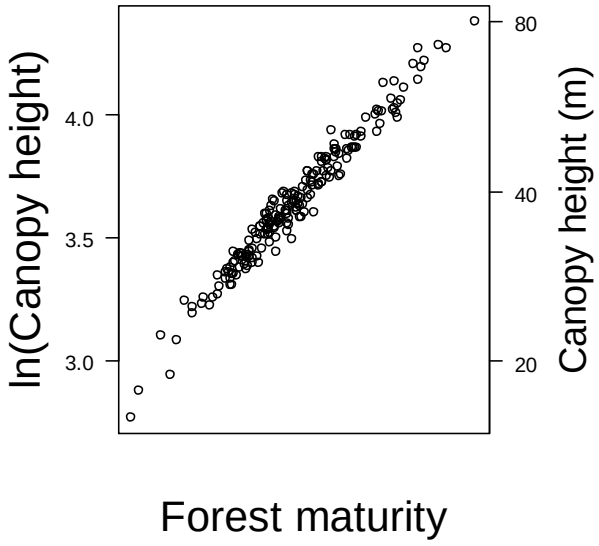


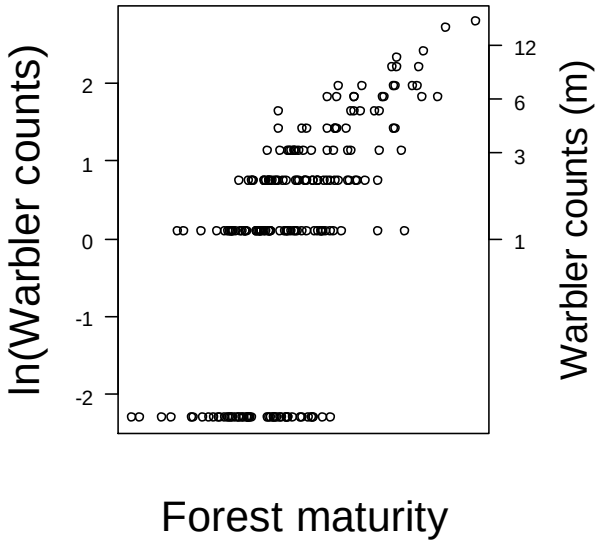
$$\delta y = \delta x \beta$$

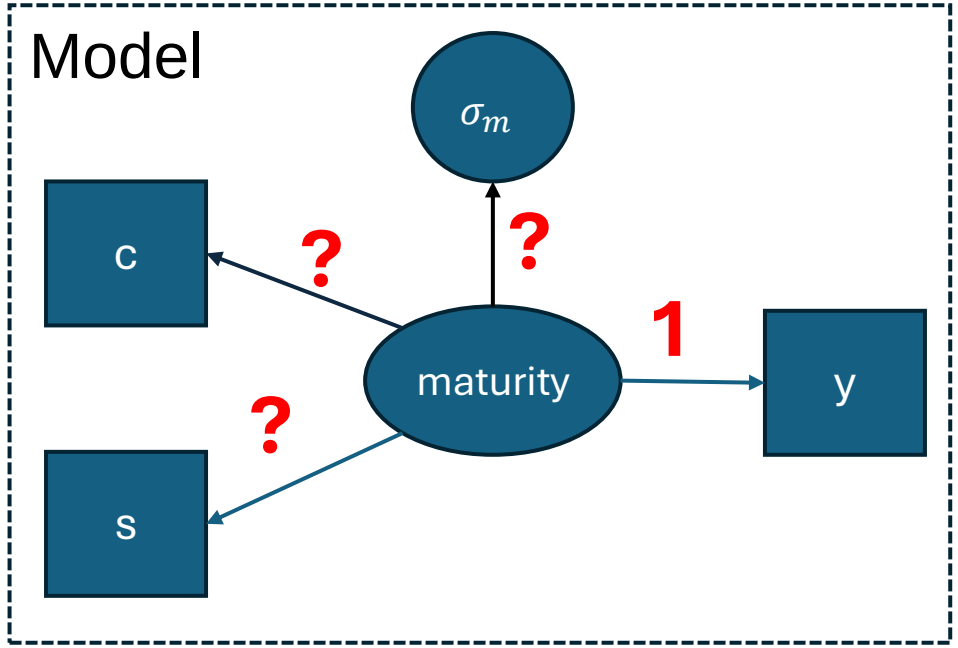
Play with the script some later!





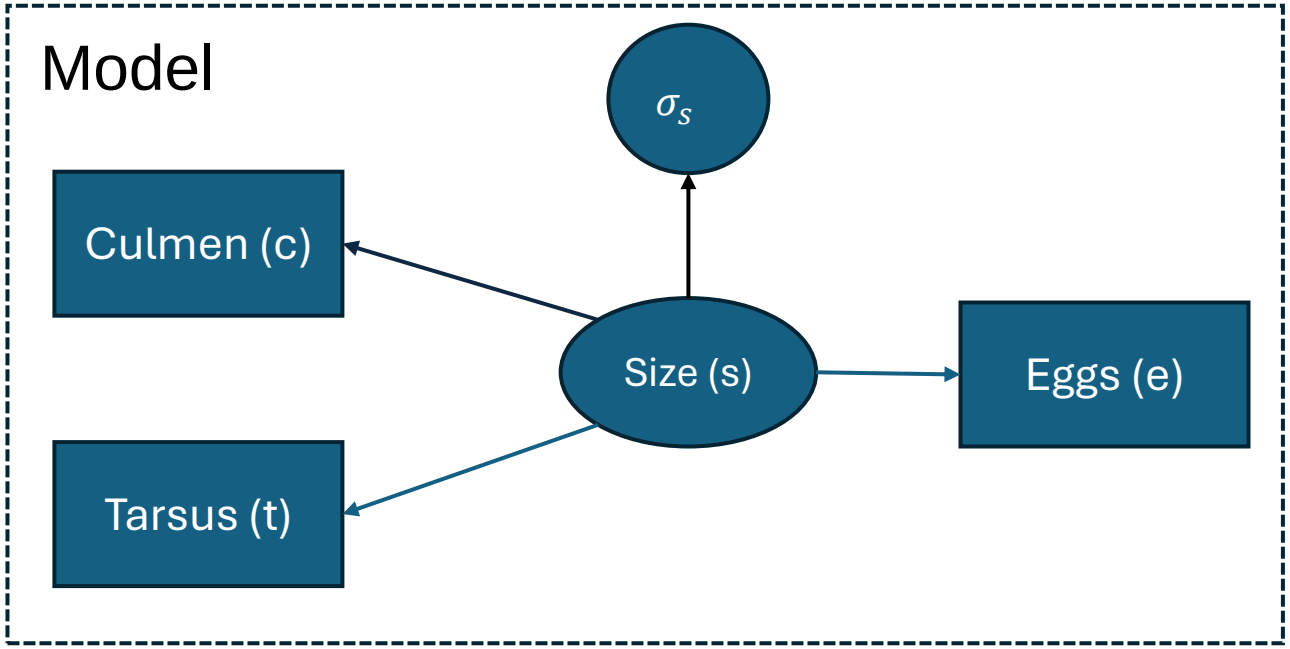




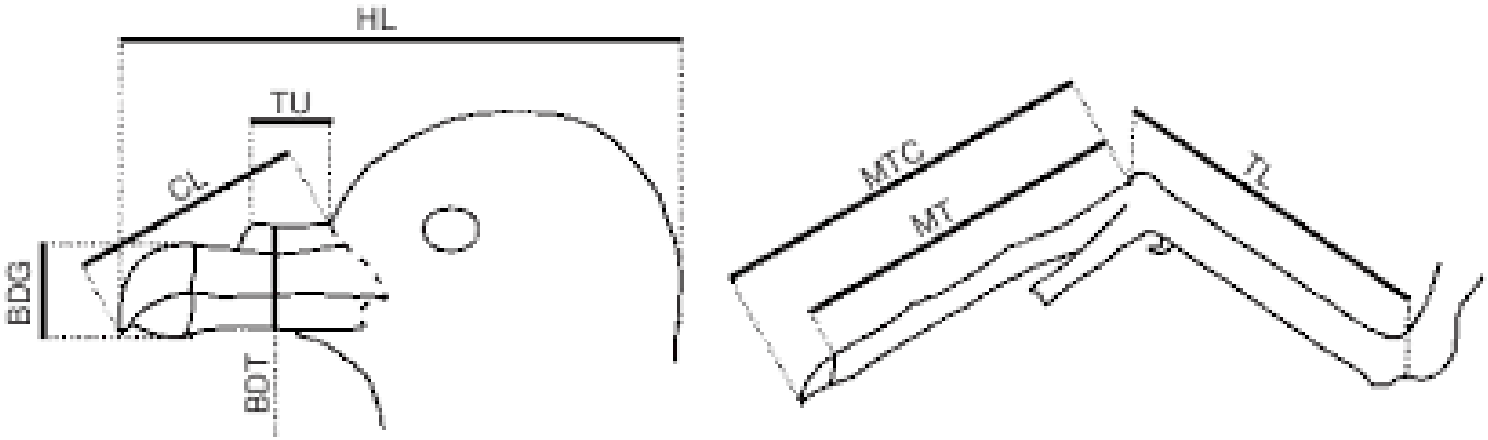


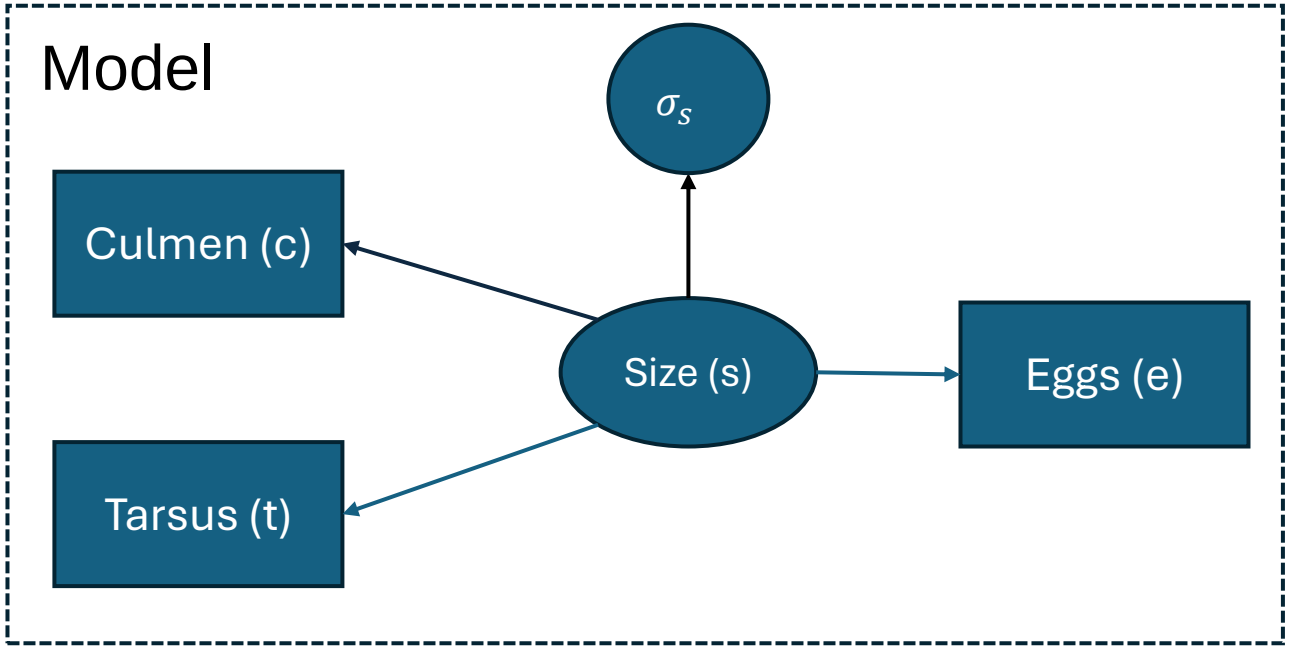
Today's coding exercise… build a SEM to estimate 'size'

- Adult body size varies by ~25%
- Egg size can range from 50 cm^3 to >100 cm^3

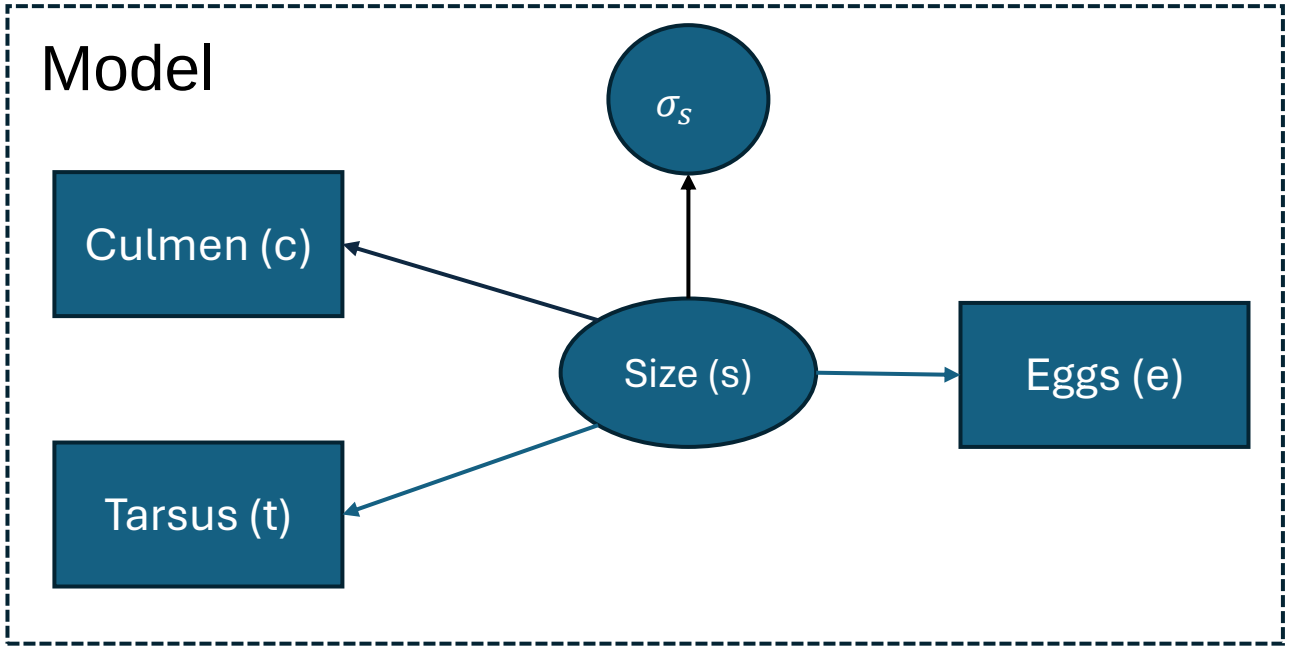


We take multiple morphometric measurements





And we'll attempt to estimate the effect of size on clutch size



I've left a couple of things 'blank' in the model for you to supply