HIERARCHICAL MODELING!! DAY 3

DAY 3 Board work

* Modeling the effect of light variation on sapling height

$$\mu_i = g(\alpha, r, c, x_i) = \frac{\alpha(x_i - c)}{\frac{\alpha}{r} + (x_i - c)}$$

α → asymptote; x_i → light level (lumens)

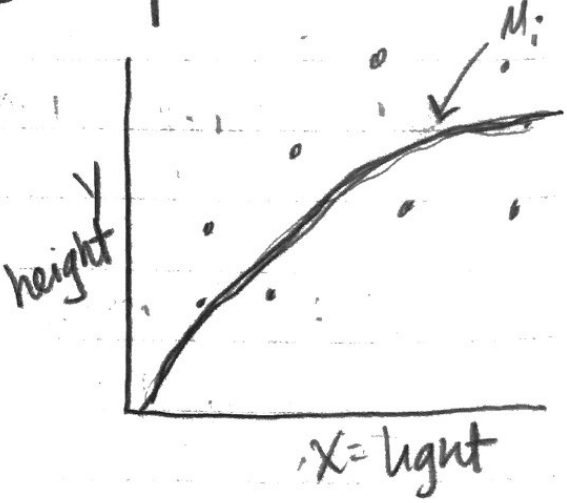
r = slope @ low light (competitive ability)

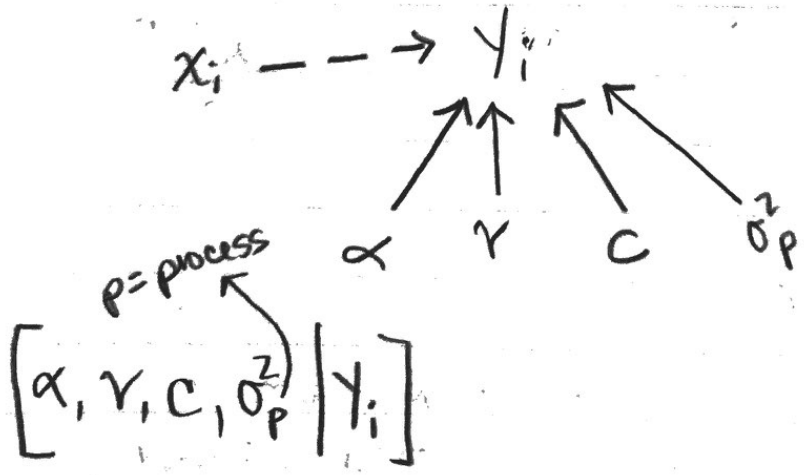
c = light level at which growth is zero (x-axis intercept)

y = Δ height = DVariable (measured perfectly)

For now x_i also measured perfectly

Priors on α mean = 35, sd = 4.25





σ_p^2 = The stuff that influences the response that is not included in the Model

Remember:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

$$\epsilon_i \sim \text{normal}(0, \sigma^2)$$

} Have to use normal

IS THE SAME AS

same

$$\mu_i = \beta_0 + \beta_1 X_i$$

$$Y_i \sim \text{normal}(\mu_i, \sigma^2)$$

} can use any dist for Y_i

Factor using the DATA as a guide

$$\left[\alpha, r, c, \sigma_p^2 \mid \underline{Y}\right] \propto \left[\alpha, r, c, \sigma_p^2, \underline{Y}\right] \quad \text{(joint)}$$

$$\propto \prod_{i=1}^{n} \left[Y_i \mid g(\alpha, r, c, x_i), \sigma_p^2\right] \quad \text{(likelihood)}$$

$$\times [\alpha][r][c][\sigma_p^2] \quad \text{(priors)}$$

μ_i

↓ only random variables can be in the posterior

★ Simple Bayesian Model because there are no random variables that are on both sides of the conditional system

★ You can ignore the Marginal (denom) because of

Choose Distributions

$$Y_i \sim \text{normal}\left(g(\alpha, r, c, x_i), \sigma_p^2\right)$$

$$\alpha \sim \text{gamma}\left(\frac{35^2}{4.5^2}, \frac{35}{4.5^2}\right)$$

$$r \sim \text{uniform}(0.1, 200)$$

$$c \sim \text{normal}(0, 10000)$$

↑ needs to be scaled appropriately relative to the mean

$$\sigma_p^2 \sim \text{inversegamma}(.001, .001)$$

↳ could put a uniform on σ because it's easier to think about sd, rather than var

$$\sigma \sim \text{uniform}(0, 100)$$

↳ coding use precision

$\tau \leftarrow 1/\sigma^2$

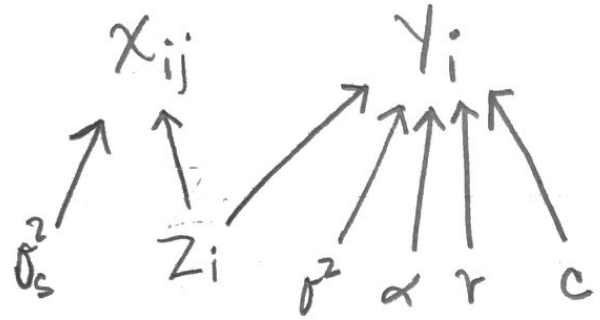
~~~

- x_i doesn't go in the posterior because it's known and putting a prior on it makes no sense

~~~

Errors in the variables

i = tree

j = reps (observations)

* What if x arises from a dist (because it is not "perfectly" known)



Z_i = mean light seen by the tree

you could estimate sigma at each tree ($\sigma^2_{s,i}$)

→ Do not have to assume homoskedasticity ☺

response depends on true light measured w/ uncertainty

$$\mu_i = g(\alpha, \gamma, c, Z_i)$$

matrix (row for each tree)

$$\left[\alpha, \gamma, c, \sigma^2_p, \underline{Z}, \sigma^2_s \middle| \underline{Y}\ \underline{X}\right] \propto \left[\alpha, \gamma, c, \sigma^2_p, \underline{Z}, \underline{Y}, \sigma^2_s, \underline{X}\right]$$

$$\propto \prod_{i=1}^{n} \left[Y_i \middle| g(\alpha, \gamma, c, Z_i), \sigma^2_p\right] \times \prod_{j=1}^{8} \left[X_{ij} \middle| Z_i, \sigma^2_s\right]$$

$$\times [\alpha][\gamma][c][\sigma^2_p][Z_i][\sigma^2_s]$$

* generally, you need more subscripts at top of your DAGs (X_{ij} vs. Z_i). If not the case, this is an indication something might be wrong.

need to add distributions on σ^2_s & Z_i