Stat 534: formulae referenced in lecture, week 10:
Hierarchical modeling

Example: Grizzly bear counts in Yellowstone Park

- Looks approx. linear on log scale after 1980
  - Exponential growth
  - Want to estimate population per capita growth rate
- Two ways to do this
  - Ignore random variation, initially
  - Model 1:

$$N_t = r^t \, N_0$$

$$\log N_t = t \, \log r + \log N_0$$

  - Model 2:

$$N_t = r \, N_{t-1}$$

$$\log r = \log N_t - \log N_{t-1}$$

  - Mathematically identical
- Now add random variation $\Rightarrow$ 2 estimators
- Model 1:
  - Linear regression

$$\log N_t = \log r \, t + \log N_0 + \varepsilon_t$$

  - $\widehat{\log r}$ = regression slope
- Model 2:
  - average change

$$\log r_t = \log N_t - \log N_{t-1} + \tau_t$$

  - $\widehat{\log r} = \overline{\log r_t}$
- Applied to grizzly bear data

Model	$\widehat{\log r}$	se
1	0.054	0.006
2	0.061	0.034

- Why?
  - ε_t and τ_t not the same "sort" of variability
  - Biologically and statistically different
  - ε_t: independent, $\log N_t$ independent
    * Each ε_t acts only on one $\log N_t$
  - τ_t: independent, $\log N_t$ correlated!
    * τ_t accumulate
    * $N_2 = \log r_t + \log N_1 + \tau_2$
      $= 2 \log r_t + \log N_0 + \tau_1 + \tau_2$
- Observation error (model 1)
  - population grows at r each year
  - errors are deviations from E $\log N_t$
- Process error (model 2)
  - population growth is random,
  - different r_t each year
  - τ_t is variability of growth rate

Estimating $\log r$ when both types of error

- Need to combine the process (model 2) and observation (model 1)
- Introduce an unmeasured "latent" variable, ν_t
- Process model:

$$\log \nu_t = \log r + \log \nu_{t-1} + \tau_t$$

- Observation model:

$$\log N_t = \log \nu_t + \varepsilon_t$$

- This is a hierarchical model
- Results for Yellowstone grizzlies

Parameter	estimate	se
r	0.060	0.025
σ^2_ε	0.083	
σ^2_τ	0.090	

- Process error and observation error approximately the same
- Could write the model without ν_t
  - Esp. when rv's have normal distributions
  - Latent variable simplifies writing the model
- Characteristics of hierarchical model
  - Multiple "levels" of variability
  - Almost always include a latent variable

Other examples of hierarchical models

- density dependent population growth
  - latent variable: # individuals that "drive" the density dependence
- classifying fish eggs
  - latent variable: species identity
- animal movement:
  - latent variable: "state", resting, searching, traveling