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

Variance of $\hat{\lambda}$:

- $\hat{\lambda}$ is a non-linear function of the vital rates, i.e. the elements of $\boldsymbol{A}$.
- In real world, a_{ij} are estimated.
- Variability in $\hat{a}_{ij}$ propagates into uncertainty in λ
- Delta method: approximation to the variance of a non-linear function
- One parameter: $\theta = f(\beta)$, e.g., $\theta = \exp \beta$

$$\text{Var } \theta \approx \left(\frac{d\theta}{d\beta}\right)^2 \text{Var } \beta$$

- If derivative is a function of β, it's evaluated at $\hat{\beta}$
- Multiple parameters, applied to λ

$$\text{Var } \lambda \approx \sum_{i,j} \left(\frac{\partial \lambda}{\partial a_{ij}}\right)^2 \text{Var } a_{ij} + \sum_{(i,j)\neq(k,l)} \left(\frac{\partial \lambda}{\partial a_{ij}}\right)\left(\frac{\partial \lambda}{\partial a_{kl}}\right) \text{Cov } a_{ij} a_{kl}$$

- Don't include $(k, l) = (i, j)$ in the second sum because those are the variance terms
- The partial derivatives are the sensitivity values!

$$\text{Var } \lambda \approx \sum_{i,j} S_{ij}^2 \text{Var } a_{ij} + \sum_{(i,j)\neq(k,l)} S_{ij} S_{kl} \text{Cov } a_{ij} a_{kl}$$

- Big simplification if you can assume all estimates are independent, so all Covariances = 0
  - Reality is that estimates are often correlated
  - Song sparrows, f_1 and f_2 are a single estimate, so correlation = 1

Song sparrow example:

- The uncertainty in each estimate

Parameter	s.e.	Variance	Sensitivity
f_1	0.52	0.27	0.091
f_2	0.52	0.27	0.057
ϕ_0	0.060	0.0036	1.96
ϕ_1	0.057	0.0032	0.26

- f_1 and f_2 are a single estimate, so correlation = 1 and covariance = 0.27
- ϕ_0 and ϕ_1 are independent and derived from different data than f_1 and f_2
  - So all correlations involving ϕ_0 or ϕ_1 are 0, so those covariances = 0
  - Reminder:

$$\text{Cor } X, Y = \frac{\text{Cov } X, Y}{\sqrt{\text{Var } X \times \text{Var } Y}}$$

$$\text{Cov } X, Y = (\text{Cor } X, Y)\sqrt{\text{Var } X \times \text{Var } Y}$$

- Putting the pieces together
  - Only have to consider the non-zero matrix elements
  - The zeros in $\boldsymbol{A}$ are fixed at zero, so their Var = 0

$$\begin{aligned} \text{Var } \hat{\lambda} &= 0.091^2 \times 0.27 \\ &+0.057^2 \times 0.27 \\ &+1.96^2 \times 0.0036 \\ &+0.26^2 \times 0.0032 \\ &+0.091 \times 0.057 \times 0.27 \\ &+0.091 \times 0.057 \times 0.27 \end{aligned}$$

  - Var $\hat{\lambda} = 0.020$, se $\hat{\lambda} = 0.14$
- Shortcut matrix computation of Var $\hat{\lambda}$:
  - Write $\boldsymbol{S}$ as a column vector of the non-zero sensitivities

$$\boldsymbol{S}' = [0.091, 0.057, 1.96, 0.26]$$

– and $\boldsymbol{V}$ as the variance-covariance for the matrix elements

$$\boldsymbol{V} = \begin{bmatrix} 0.27 & 0.27 & 0 & 0 \\ 0.27 & 0.27 & 0 & 0 \\ 0 & 0 & 0.0036 & 0 \\ 0 & 0 & 0 & 0.0032 \end{bmatrix}$$

– Var $\hat{\lambda} = \boldsymbol{S}\,\boldsymbol{V}\,\boldsymbol{S}$

Other choices of models:

- What if song sparrows live longer than 3 years?
- 5 years, $\lambda = 1.028$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 2.6 & 2.6 & 2.6 & 2.6 \\ 0.57 & 0 & 0 & 0 & 0 \\ 0 & 0.57 & 0 & 0 & 0 \\ 0 & 0 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0.57 & 0 \end{bmatrix}$$

- 7 years, $\lambda = 1.052$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 2.6 & 2.6 & 2.6 & 2.6 & 2.6 & 2.6 \\ 0.57 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.57 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.57 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.57 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.57 & 0 \end{bmatrix}$$

- No fixed lifespan, $\lambda = 1.06$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 2.6 \\ 0.2 & 0.57 \end{bmatrix}$$

  – Stage structured with Juvenile and Adult

  – No longer a Leslie matrix

Stage structured population models

- Very common when size (continuous) determines the demography

- Northern Monkshood: size = stem basal diameter
- Seedlings always died, so demography driven by survival and clonal reproduction
  - Plant overwinters as rootstock,
  - can produce two or more stems the next spring
  - eventually develops two separate root systems
- Simplified version of Monkshood stem demography
  - Classify plants into 3 size categories: no stem, stem < 2mm, stem ≥ 2mm
- At a low elevation site: $\hat{\lambda} = 0.939$

$$\boldsymbol{A}_l = \begin{bmatrix} 0.72 & 0.13 & 0.71 \\ 0.10 & 0.70 & 0.10 \\ 0 & 0.08 & 0.75 \end{bmatrix}$$

- At a high elevation site: $\hat{\lambda} = 1.035$

$$\boldsymbol{A}_h = \begin{bmatrix} 0.60 & 0.42 & 1.11 \\ 0.15 & 0.73 & 0.09 \\ 0 & 0.09 & 0.77 \end{bmatrix}$$

Growth and survival models:

- Very common stage-structured model
- 4 stages:
  non-reproductive / small / medium / large

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & f_2 & f_3 & f_4 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 \\ 0 & 0 & a_{34} & a_{44} \end{bmatrix}$$

- Interpretations of a's:

a_{11}	survive and don't grow
a_{21}	survive and grow
a_{22}, a_{32}	similar for small
a_{33}, a_{43}	similar for medium
a_{44}	survival for large

- More useful to reparameterize

ϕ_1	P[non repro survives]
g_1	P [non repro grows | survived]
ϕ_2	P[small survives]
g_2	P [small grows | survived]
ϕ_3	P[medium survives]
g_3	P [medium grows | survived]
ϕ_4	P[large survives]

- The relationships:

$a_{11} = \phi_1(1 - g_1)$	survive and don't grow
$a_{21} = \phi_1 g_1$	survive and grow
$a_{22} = \phi_2(1 - g_2)$	similar for small
$a_{32} = \phi_2 g_2$	
$a_{33} = \phi_3(1 - g_3)$	similar for medium
$a_{43} = \phi_3 g_3$	
$a_{44} = \phi_4$	survival for large

- Know how to get sensitivities for matrix elements, a_{ij}
- Really want sensitivity to ϕ_i or g_i, e.g.

$$\frac{\partial \lambda}{\partial g_1} = \sum_{ij} \left(\frac{\partial \lambda}{\partial a_{ij}} \right) \left(\frac{\partial a_{ij}}{\partial g_1} \right)$$

- For this G-S model:

$$\frac{\partial \lambda}{\partial g_1} = \phi_1 S_{21} - \phi_1 S_{11}$$

Life Table Response Experiment (LTRE):

- "Retrospective" analysis of demography
- Treat $\boldsymbol{A}$ as the response in an observational or experimental study
- Observe differences in $\hat{\lambda}$
  - Ask, which demographic rates most responsible for the difference in $\hat{\lambda}$?
  - Rates most different between two conditions is not sufficient
  - also need large sensitivity for that rate

- First order Taylor expansion of $\hat{\lambda}(\boldsymbol{A}_h) - \hat{\lambda}(\boldsymbol{A}_l)$

$$\hat{\lambda}(\boldsymbol{A}_h) - \hat{\lambda}(\boldsymbol{A}_l) \approx \sum_{ij} S_{ij} \left(a_{ij}^{(h)} - a_{ij}^{(l)} \right)$$

- Which sensitivity matrix to use?
  - Caswell suggests the average transition matrix: $\boldsymbol{A}^* = (\boldsymbol{A}_l + \boldsymbol{A}_h)/2$

Monkshood LTRE

- $$\boldsymbol{A}^* = \begin{bmatrix} 0.66 & 0.275 & 0.91 \\ 0.125 & 0.715 & 0.095 \\ 0 & 0.085 & 0.76 \end{bmatrix}$$

- $$\boldsymbol{S}_{A^*} = \begin{bmatrix} 0.294 & 0.156 & 0.058 \\ 0.768 & 0.407 & 0.153 \\ 1.50 & 0.796 & 0.299 \end{bmatrix}$$

- The components

Element	Interp.	$a_{ij}^u - a_{ij}^l$	$\Delta\lambda_{ij}$
a_{11}	survival	-0.12	-0.035
a_{12}	fecundity	0.29	0.045
a_{13}	fecundity	0.40	0.023
a_{21}	growth	0.05	0.038
a_{22}	survival	0.03	0.012
a_{23}	clonal repro	-0.01	-0.002
a_{31}		0	0
a_{32}	growth	0.01	0.008
a_{33}	survival	0.02	0.060

- How good is the approximation?
  - Calculate $\sum_{ij} \Delta\lambda_{ij} = 0.095$
  - Observed $a_{ij}^u - a_{ij}^l = 0.096$
  - Pretty good approximation!
  - Expect less good if a_{ij} more different

Summary of what we've seen so far:

- Short-term dynamics: use $\boldsymbol{A}$ and $\boldsymbol{N}_0$ to project $\boldsymbol{N}_t$
- Long-term dynamics:
  - asympototic growth rate, λ
  - stationary age/stage distribution, $\boldsymbol{U}$
  - reproductive values, $\boldsymbol{V}$
  - how quickly transient dies out, $\mid \lambda^1 \mid / \mid \lambda^2 \mid$
- Prospective analysis
  - If you could change elements of $\boldsymbol{A}$, what should you focus on?
  - What gives biggest impact on long-term growth?
  - Answered by $\boldsymbol{S}$ or $\boldsymbol{E}$
- Retrospective analysis
  - 2 or more treatments / conditions
  - different transition matrices, diff. λ
  - which elements contributed most to difference

Sources of estimates and se's

- Mark-recapture $\rightarrow \hat{\phi}$, se
- Observing nests, dens
  - empirical means and se's
- Tagging plants, watching transitions
  - Binomial distributions, se $= \sqrt{p(1-p)/N}$
- Counting seeds, relating to seedling numbers
  - Poisson distributions for seedling numbers
- Many other possibilities

Bias-variance tradeoff:

- Structured populations with continuous "stage"
- Matrix models construct bins, estimate transition prob. for each combination of bins
- Big Q: what's the best set of bins?
- Bias vs. variance tradeoff
  - Frequent statistical issue
  - Many narrow bins: low bias, high variance
  - Few wide bins: high bias, low variance
- Illustration: 100 plants
  - If initial size = 0.4, what is distribution of next year's size?
  - Wide bins: 0-0.5, 0.5-1, $\cdots$
P[0-0.5 | 0.4] = 0.91,
P[0.5-1 | 0.4] = 0.09
  - Narrow bins: 0-0.1, 0.1-0.3, 0.3-0.5, $\cdots$
P[0.3-0.5 | 0.4] = 0.5,
P[0.5-0.75 | 0.4] = 0.35,
P[0.75-1 | 0.4] = 0.06
- What are the correct probabilities?
  - Wide bins: 0-0.5, 0.5-1, $\cdots$
P[0-0.5 | 0.4] = 0.73,
P[0.5-1 | 0.4] = 0.27
  - Narrow bins: 0-0.1, 0.1-0.3, 0.3-0.5, $\cdots$
P[0.1-0.3 | 0.4] = 0.026,
P[0.3-0.5 | 0.4] = 0.70,
P[0.5-0.75 | 0.4] = 0.26,
P[0.75-1 | 0.4] = 0.004
- s.e. of "same size bin":
  - Wide bins: 0.032, $N = 79$
  - Narrow bins: 0.12, $N = 17$
- Problem is that probabilities for S=0.4 estimated only from part of the data

Integral Projection Models: General concept

- Vital rates are continuous functions of size, not discrete bins
- Use a model to estimate size-specific probabilities
- All observations used to estimate all probabilities
- Easterling, Ellner and Dixon (2000) Ecology
- e.g. for growth (change in size):

$$P[S_{t+1} = x] = g(S_t, \theta)$$

- Need to determine the form of $g()$:
  - How do mean(S_{t+1}) and variance(S_{t+1}) depend on S_t?
  - What's the appropriate distribution?
  - and all parameters, θ
- Projecting forward 1 year
  - Start # individuals in each size j, how many individuals with size i?
  - Matrix model:

$$N_{t+1}^{(i)} = \sum_j a_{ij} N_t^{(j)}$$

  - IPM:

$$n_{t+1}(s) = \int_j g(x, s, \theta)\, n_t(x)\, dx$$

Integral Projection Models: more details

- Notation:
  - x: size of an individual at time t
  - y: size of an individual at time $t+1$
  - $n(x,t)$: # individuals in size $(y, y+\delta y)$ at time t
- Need three functions, for time $t \rightarrow t+1$:

- $s(x)$: P[size x survives]
- $g(x, y)$: P[surviving individual of size x grows to size $(y, y + \delta y)$]
- $f(x, y)$: E # newborns of size $(y, y + \delta y)$ per size x individual
- clonal reproduction, if any, goes into $f()$

- Combine into the "kernel"

$$k(x, y) = f(x, y) + s(X)g(x, y)$$

  - $k(x, y)$ is a 2D surface
  - E # individuals next year in size $(y, y + \delta)$ per individual in size x

- With all the pieces:

$$n(y, t + 1) = \int_x k(x, y)\, n(x, t)\, dx$$

- This has a steady state when

$$n(y, t + 1) = \int_x k(x, y)\, n(x, t)\, dx = \lambda n(x, t)$$

  - Need to find λ and $n(y)$ numerically

"Solving" the IPM:

- Can evaluate any integral numerically
  - Approximate the integral by a sum

$$\int_{x=l}^{u} f(x)\, dx \approx \sum_{x=l}^{u} f(x) \Delta x$$

    * l and u are the smallest and largest feasible sizes
  - Choose m, the number of steps in the sum
  - $\Delta x = (U - L)/m$
  - For each step, evaluate $f(x)$ at the midpoint of the step

- Applied to a population model:

- Evaluate the kernel, $k(x, y)$ on a grid of mid-points (for x and y) $\Rightarrow \boldsymbol{A}$
- The integral model is now approximated by

$$\boldsymbol{N}_{t+1} = \boldsymbol{A}\boldsymbol{N}_t$$

- use the eigenvalues and eigenvectors of $\boldsymbol{A}$!
- Want m large (100, more?), so $\boldsymbol{A}$ is large
  - $\Rightarrow$ low bias
  - $\Rightarrow$ but also low variance
  - because all data used to estimate each element of $\boldsymbol{A}$
- Does assume (critically) the correct model for the kernel components
  - But, we have data analysis tools for that

The really neat part of an IPM:

- I thought of the IPM as a way around the bias-variance tradeoff
  - Used 2D smoothing to estimate $s(x), g(x, y), f(x, y)$
- Others, Steve Ellner?, realized what you could do with parametric models for $s(x)$, $g(x, y)$, $f(x, y)$
- Back to the monkshood LTRE data:
  - I have data from multiple sites
  - different annual temp and annual rainfall
  - What if you want to extend LTRE model to continuous covariates?
  - Requires models for each matrix element

IPMs with covariates:

- Consider data from all sites together
- Start with a parametric model, e.g.:
  - survival: logit $s(x) = \beta_0 + \beta_1 x$

- add environmental covariates

$$\text{logit } s(x) = \beta_0 + \beta_1 x + \beta_2 Temp + \beta_3 x * Temp$$

- Now have kernel functions for any covariate value(s)
- And hence can answer questions like:
  - what does the demography of a rare plant look like if the temperature increases by 2°C and precip increases by 5cm?
- Practical issues:
  - need to estimate 3 components of the kernel
  - growth function prob. requires modeling both mean and variance
  - Lots of bookkeeping
  - Much more work than just counting #'s of individuals in different groups
- R library IPMpack
  - Was on CRAN, seems no longer maintained
  - removed early 2020
  - can get from the archive or sourceforge