A Mathematical Demographic Analysis of Enrollment at a Regional University

John Lagergren

East Tennessee State University

Collaborators:

Diana Gonzalez

California State Polytechnic University

David Spencer

University of North Carolina

Advisor:

Dr. Ariel Cintron-Arias

East Tennessee State University

DEPARTMENT *of* MATHEMATICS & STATISTICS

College of Arts & Sciences

EAST TENNESSEE STATE UNIVERSITY

Grant # DUE -1356397

Dataset: ETSU Office of Planning and Analysis

Year	# Fresh.	# Soph.	# Jun.	# Sen.	Total	New Fr.
2000	2615	1763	1840	2612	8830	1467
2001	2646	1780	1789	2771	8986	1488
2002	2600	1761	1884	2753	8998	1476
2003	2622	1820	1900	2879	9221	1474
2004	2544	1862	1941	3008	9355	1569
2005	2659	1773	1978	3076	9486	1484
2006	2842	1997	1972	3084	9895	1569
2007	3020	1972	2158	3120	10270	1717
2008	3081	2082	2183	3308	10654	1924
2009	3197	2273	2293	3523	11286	1924
2010	3277	2281	2447	3804	11809	2033
2011	3333	2291	2565	3949	12138	2052
2012	3213	2090	2454	4071	11828	2105

Matrix Models: Terminology

Demography or Ecology	Student Enrollment
Age group (juvenile, adult, etc.)	Student classification (F, S, J, U)
Survival	Retention
Birth or Immigration	Recruitment
Living	Enrolled
Death	Departure

Matrix Model

- The state variable denotes the size of each student classification at a time point
- Denote the number of freshmen, sophomores, juniors, and seniors at year n:
  $\mathbf{x}_1(n), \mathbf{x}_2(n), \mathbf{x}_3(n), \mathbf{x}_4(n)$
- The general form of the matrix models is
  $\mathbf{x}(n+1) = \mathbf{A}\mathbf{x}(n) + \mathbf{b}(n)$

Models 1 and 2

- Model 1

$$\mathbf{x}(n+1) = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \end{bmatrix} \mathbf{x}(n)$$

- Model 2

$$\mathbf{x}(n+1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} u_1(n) + \alpha_1 + \beta_1 n \\ \alpha_2 + \beta_2 n \\ \alpha_3 + \beta_3 n \\ \alpha_4 + \beta_4 n \end{bmatrix}$$

Models 3 and 4

- Model 3

$$\mathbf{x}(n+1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} u_1(n) \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

- Model 4

$$\mathbf{x}(n+1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} u_1(n) \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Model 5

- Model 5

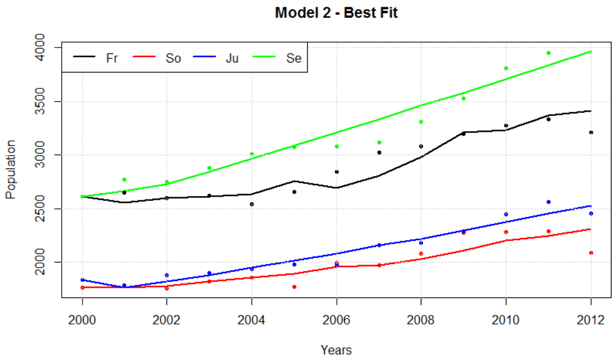
$$\mathbf{x}(n+1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} \alpha_1 + \beta_1 n \\ \alpha_2 + \beta_2 n \\ \alpha_3 + \beta_3 n \\ \alpha_4 + \beta_4 n \end{bmatrix}$$

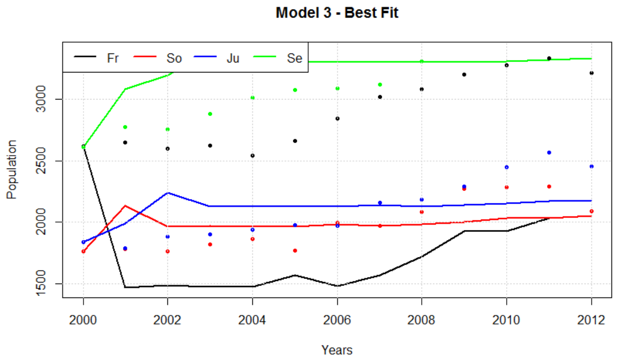
Vector Ordinary Least Squares

$$\min_{\theta}[\omega_1 \sum_{n=1}^{13}(x_1(n,\theta) - F_n)^2 + \omega_2 \sum_{n=1}^{13}(x_2(n,\theta) - S_n)^2 +$$
$$\omega_3 \sum_{n=1}^{13}(x_3(n,\theta) - J_n)^2 + \omega_4 \sum_{n=1}^{13}(x_4(n,\theta) - U_n)^2]$$

- Goal: minimize sum of weighted sums of squares
- Genetic Algorithms
- Broyden-Fletcher-Goldfarb-Shanno (BFGS)

OLS Results – Best Fit Solutions





Akaike Information Criterion (AIC)

- Provides a means to quantify the quality of a statistical model

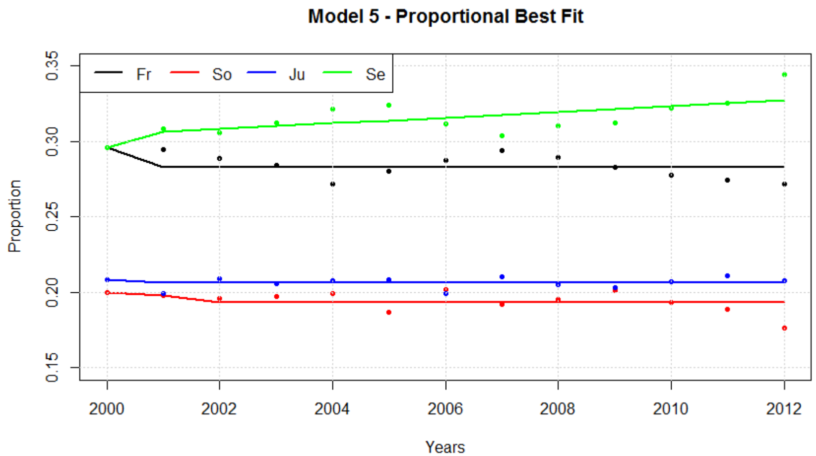
$$\text{AIC} = N\nu \ln\left(\frac{\sum_{j=1}^{N}(\mathbf{y}_j - \mathbf{f}(t_j;\hat{\mathbf{q}}_{\text{OLS}}))^T(\mathbf{y}_j - \mathbf{f}(t_j;\hat{\mathbf{q}}_{\text{OLS}}))}{N\nu}\right) + 2(\kappa_q + 1)$$

- We want to use the model with the lowest AIC score

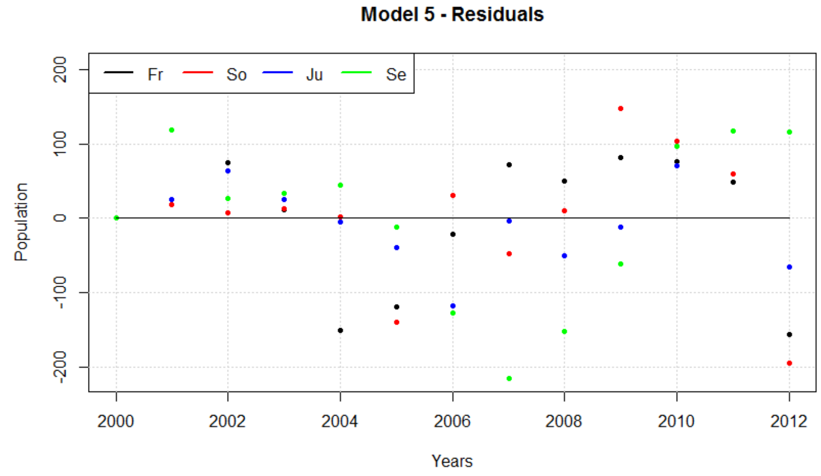
	Model 1	Model 2	Model 3	Model 4	Model 5
AIC Score	682.546	493.031	686.904	732.842	491.751

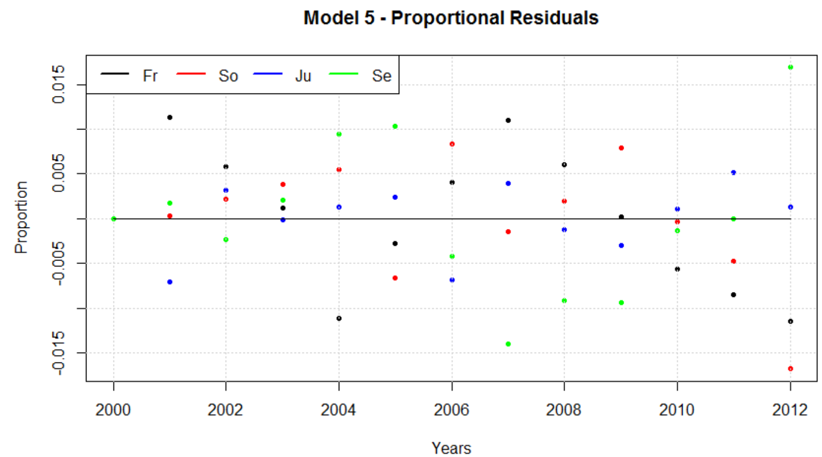
Residual Plots of Selected Model Under AIC

Best Fit Plot

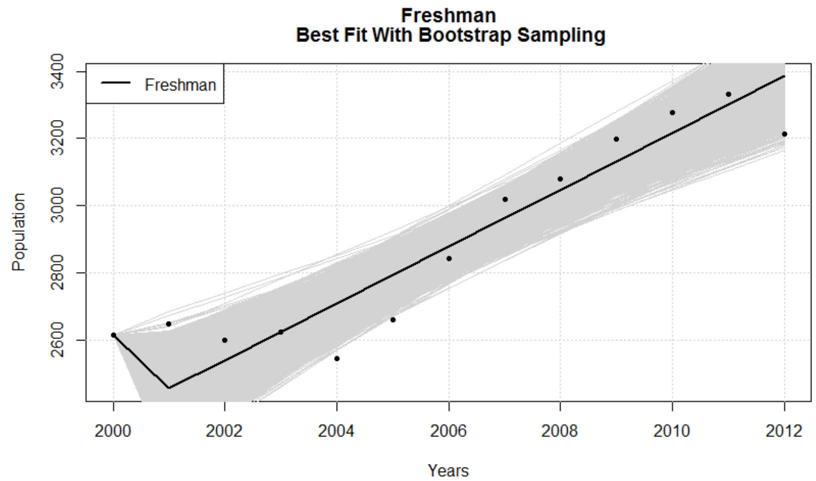


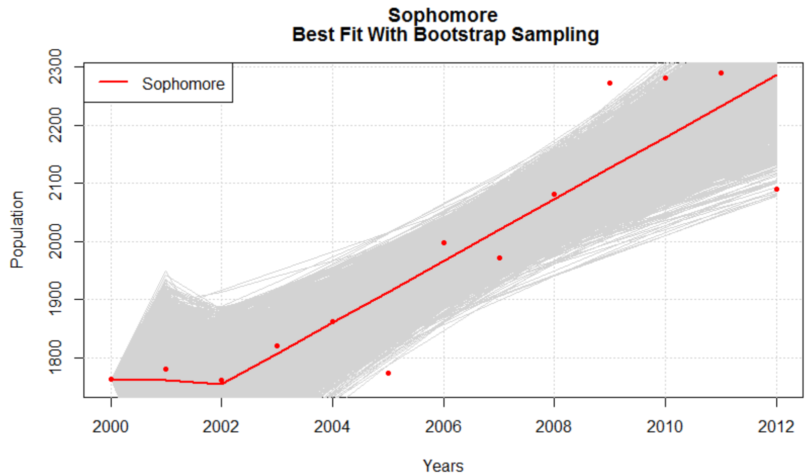
Residual Plot

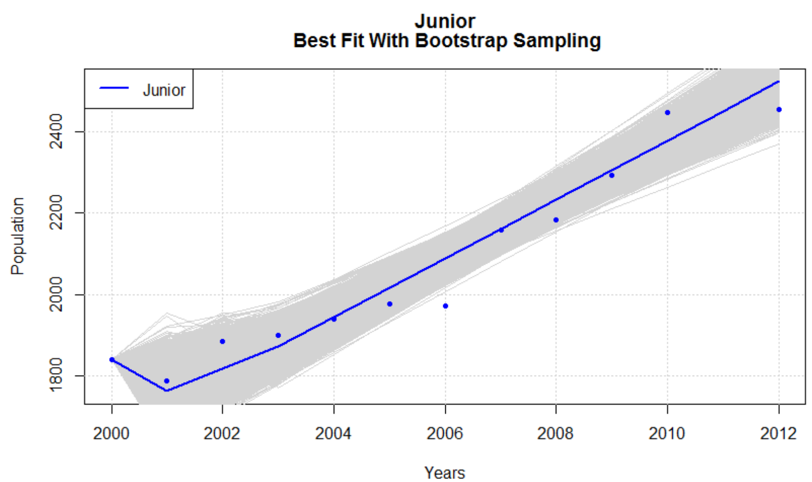


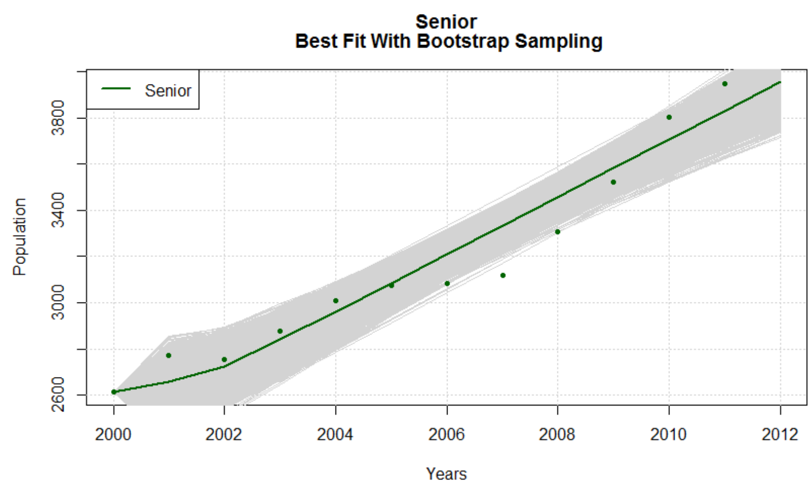


Bootstrap Sampling









Estimates from Bootstrap Sampling

- Parameter Estimates and Standard Error

Parameter	Estimate	Standard Error
F to **S** probability	0.251	0.002
S to **J** probability	0.291	0.002
J to **U** probability	0.368	0.003
F flat recruitment	2369.702	1.026
F yearly increase	84.673	0.141
S flat recruitment	1072.319	4.462
S yearly increase	31.967	0.199
J flat recruitment	1192.494	2.905
J yearly increase	56.648	0.141
U flat recruitment	1880.667	5.462
U yearly increase	97.719	0.275

Calculating the Student Life Table

- Retention Probability: $p_x = \hat{p}_i$
- Departure Probability: $q_x = 1 - \hat{p}_i$
- Number Enrolled: $l_x = 1000 \prod_{1}^{i-1} \hat{p}_i$
- Departure distribution: $d_x = l_x q_x$

$$L_x = l_x - (1 - a_x) d_x, \quad T_x = \sum_{i=x}^{4} L_i, \quad e_x = \frac{T_x}{l_x}$$

Life Table Estimates

	Freshman	Sophomore	Junior	Senior
Retention Probability	0.251	0.291	0.368	0.000
Departure Probability	0.749	0.709	0.632	1.000
# Enrolled in Student Class	1000.000	251.431	73.226	26.953
Departure Distribution	748.569	178.205	46.273	26.953
Person-years Enrolled	513.430	162.329	50.090	0.083
Person-years Remaining	725.932	212.501	50.173	0.083
Remaining Time Enrolled	0.726	0.845	0.685	0.003
Average Years in Class	0.350	0.500	0.500	0.003

Discussion

- Extremely high departure rates indicate very few students graduate after four years spending **one** year in **each** student class
- Is the use of fixed demographic parameters valid for college enrollment?
- Are the assumptions too strong to find valuable qualities in the dynamics of student enrollment?

Future Work

- Sensitivity analysis on the Model and the Life Table?
- Deeper analysis on Life Table values?
- Allow more time for GA convergence?
- More models?
- Look for more data?
- Other school systems?

References

- ETSU Fact Book. *http://www.etsu.edu/opa/factbook.aspx*
- Modeling and Inverse Problems in the Presence of Uncertainty. *H.T. Banks, S. Hu, and W.C. Thompson*. 2013.
- Methods Protocol for the Human Mortality Database. *J.R. Wilmoth, K. Andreev, D. Jdanov, and D.A. Glei*. 2007.
- Introduction to the Mathematics of Population. *N. Keyfitz*. 1968.
- Mathematical Models for the Growth of Human Populations. *J.H. Pollard*. 1973.
- National Center for Education Statistics. https://nces.ed.gov/ipeds/

Questions?

- Email me at lagergren@goldmail.etsu.edu
- Special thanks:
  - To the National Institute for Mathematical and Biological Synthesis for funding and hosting this opportunity
  - To NSF for funding this research
  - My advisor, Dr. Ariel Cintron-Arias
  - My research peers, D. Gonzalez and D. Spencer

DEPARTMENT *of* MATHEMATICS & STATISTICS
College of Arts & Sciences
EAST TENNESSEE STATE UNIVERSITY

Grant # DUE -1356397