Darryl Holman
117 Denny Hall
206-543-7586
djholman@u.washington.edu
http://faculty.washington.edu/djholman

Quantitative Methods and Modeling for Biocultural Anthropology. BIO A 526

Winter 2005

Scope: This course will introduce you to the concepts and methods of taking complex real-world systems and creating quantitative models of the system. By the end of the quarter, you will have gained experience in modeling the behavior of a system, developing testable hypotheses, and using observations (taken from fieldwork or data sets) to statistically evaluate hypotheses arising from the model. We will survey some of the concepts, tools, and methods for developing models based on underlying biocultural processes, as well as the methods of testing models from observations collected in anthropological field studies. We will focus on methods for longitudinal research of fertility, mortality, disease dynamics, population genetics, and other biocultural processes, but the concepts and methods are applicable to many other types of anthropological and biocultural research covering the life span.

After a brief introduction to modeling, we will focus on three modeling techniques. (1) We will briefly examine dynamic simulation models. (2) We will spend much time on developing likelihood models. These are models that can be developed from theoretical foundations and can be fit to real data. (3) We will examine simulation models and techniques.

Classes: Monday, Wednesday and Friday, 02:30-03:50 p.m. in 322 Parrington.

Office hours: After class. Other times can be arranged. You can use email anytime (djholman@u.washington.edu).

Textbook and readings: There are no required textbooks for this course. I will, however, recommend some of the more important books for this course. Most (if not all) readings will be provided on-line.

Grading: Your course grade will be based on six problem sets (8% each) and a final project (52%).

Problem sets: The six problem sets will consist of several modeling exercises. Frequently, the problems will require the use of computer software. I recommend that you get an account on the CSDE Windows network, as all the required software will be available on those systems. Data sets and other helpful material will be available on the web site. You are free to use books, readings, notes, and web pages to help you work on the problems. You can work in groups, but I recommend you tackle the problems yourselves. **Grades for late problem sets will depreciate by 10% per day**, including any fraction of a day late. Problem sets are due by the beginning of the class period, one calendar week after being handed out.

There are two software programs we will be using. The first is *STELLA*, a dynamic modeling and simulation program. A limited version of *STELLA* is available with one of the optional textbooks or can be downloaded from http://www.hps-inc.com. If you think *STELLA* will be useful for your research, a full-feature version is available at a substantial student discount from the publisher. One CSDE terminal server has *STELLA* installed.

The second software package is called *mle*, and is written by your instructor. You can use this program for maximum likelihood estimation and simulation programming. The software is free from http://faculty.washington.edu/~djholman/mle. Extensive documentation is available online. You can download a pdf version of the documents for browsing or printing. The *mle* program will also be installed on the CSDE terminal servers. Most of the exercises that use *mle* can be done in other statistical programming languages (*splus, R, Matlab, Gauss, Octave*). You are free to use any of these for your work under the idea that learning one such language will help you understand any other. For the statistical programming exercises, *mle* will be easiest program to use, but other languages (*Gauss, R, Matlab*) are suitable.

Projects: 52% of your course grade will be based on a project. This project can take one of several forms: (1) a new research proposal in preliminary form incorporating some of modeling ideas covered in the course; (2) a completed existing research or funding proposal (a proposal started in BIO A 525, for example) that is revised to include one or more of the methods covered in this course; (3) a new manuscript in which you have applied methods covered in this course; (4) a term paper in which a model is developed and explored.

Quantitative Methods and Modeling for Biocultural Anthropology. BIO A 526, Winter 2005

Topics and schedule

Week 1

Readings: Levins (1966), Hilborn and Mangel (1997) pp. 12-38.

Jan 3 Course introduction

Jan 5 Models and the scientific method

Jan 7 Overview of some modeling techniques

Week 2

Readings: Skellam (1955), Hannon and Ruth (1997) pp. 3-27.

Optional Reading: Review Thompson and Gardner (1998) as needed.

Jan 10 Some useful mathematical tricks (PS 1 assigned)

Jan 12 How to build a model

Jan 14 Dynamic systems models I

Week 3

Readings: Sattenspiel (1990), Wood JW (1998).

Jan 17 No class – Holiday

Jan 19 Dynamic systems models II (PS 1 due; PS 2 assigned)

Jan 21 Dynamic systems models III

Week 4

Readings: Hilborn and Mangel (1997) pp. 39-93, Holman (handouts 1 & 2).

Optional Reading: Edwards (1992), Gage (1989), Konigsberg & Frankenberg (1992).

Jan 24 Probability models

Jan 26 Likelihood models I (PS 2 due)

Jan 28 Likelihood models II (PS 3 assigned)

Week 5

Readings: Raftery (1995), Holman (handout 2).

Optional Reading: Burnham and Anderson (1998).

Jan 31 Inference and model selection I

Feb 2 Inference and model selection II

Feb 4 Messy data I (PS 3 due)

Week 6

Readings: Holman (handout 3), Holman and Yamaguchi (*in press*), Wood et al. (1993).

Feb 7 Messy data II (PS 4 assigned)

Feb 9 Heterogeneity I – Measured covariates

Feb 11 Heterogeneity II – Hazard covariate models

Week 7

Readings: Holman (2003), Holman et al. (*in press*), Holman & Grimes (2003), Weinberg & Gladen (1986).

Optional Reading: Holman and Jones (1998).

Feb 14 Heterogeneity III – unobserved characteristics (PS 4 due)

Feb 16 Heterogeneity IV – Mixture models

Feb 18 Heterogeneity V – "Sterility" models (PS 5 assigned)

Week 8

Readings: Coale and McNeil (1972), Holman et al. (n.d.), Sarton-Miller et al. (2004).

Feb 21 No class – Holiday

Feb 23 Convolution models

Feb 25 Clustered observations (PS 5 due)

Week 9

Readings: Efron & Tibshirani (1991), Holman (handout 4).

Feb 28 Heterogeneity VI – time varying covariates

Mar 2 Bootstrapped estimates and other computer intensive methods

Mar 4 Simulation models I (PS 6 assigned)

Week 10

Mar 7 Simulation models II

Mar 9 Simulation models III

Mar 11 Odds and ends (PS 6 due)

Exam week

Mar 18 Projects due by 5 pm

Quantitative Methods and Modeling for Biocultural Anthropology. BIO A 526, Winter 2005

Bibliography

Math books

Bender EA (1978) *An Introduction to Mathematical Modeling*. New York: Wiley (Reprinted by Dover in 2000).

Cullen MR (1983) *Mathematics for the Biosciences.* New York: PWS-Kent Publishing (Reprinted by Ceramic Book and Literature Service, www.cbls.com). [Math course for biologists]

Simon W (1972) *Mathematical Techniques for Biology and Medicine.* NY: Academic Press (reprinted by Dover in 1986). [Mathematics review for biologists—somewhat advanced]

Thompson Sp and Gardner M (1998) *Calculus Made Easy*. New York: St. Martin's Press. [Recommended: This is a fantastically easy way to learn the most useful stuff in calculus].

Books on simulation

Bratley P, Fox BL, Schrage LE (1987) *A Guide to Simulation*. New York: Springer-Verlag.

Ferber J (1999) *Multi-agent Systems: An Introduction to Distributed Artificial Intelligence*. London: Addison-Wesley.

Hannon B, Ruth M (1997) *Modeling Dynamic Biological Systems.* New York: Springer-Verlag. [Dynamic models only]

Huckfeldt RR, Kohfeld CW, Likens TW (1982) *Dynamic Modeling: An Introduction*. Thousand Oaks, California: SAGE Publications, Inc.

Mooney CZ (1997) *Monte Carlo Simulation*. Thousand Oaks, California: SAGE Publications, Inc.

Books on probability, likelihood, stochastic modeling, model selection

Burnham KP, Anderson DR (1998) *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag.

Cullen AC, Frey HC (1999) *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs.* New York: Plenum Press.

Edwards AWF (1992) *Likelihood.* London: Johns Hopkins University Press (1972 edition by Cambridge University Press). [Strongly recommended].

Eliason SR (1993) *Maximum Likelihood Estimation: Logic and Practice.* Thousand Oaks, California: SAGE Publications, Inc.

Gentle JE (2002) *Elements of Computational Statistics*. New York: Springer-Verlag. [Advanced material].

Gilchrist WG (2000) *Statistical Modelling with Quantile Functions*. Boca Raton: Chapman & Hall/CRC. [interesting alternative modeling ideas].

Guttorp P (1995) *Stochastic Modeling of Scientific Data*. London: Chapman & Hall. [Advanced modeling]

Hilborn R, Mangel M (1997) *The Ecological Detective: Confronting Models with Data.* Princeton, New Jersey:Princeton University Press. [Strongly recommended]

Mooney CZ, Duval RD (1993) *Bootstrapping: A Nonparametric Approach to Statistical Inference.* Thousand Oaks, California: SAGE Publications, Inc.

Morgan BJT (2000) *Applied Stochastic Modelling.* London: Arnold. (Recommended).

Ross GJS (1990) *Nonlinear Estimation.* New York: Springer-Verlag.

Royall R (1997) *Statistical Evidence: A likelihood paradigm*. Boca Raton: Chapman & Hall/CRC.

Readings

Coale AJ and McNeil DR (1972). The distribution of age at first marriage in a female cohort. *J Am Stat Assoc* **67**:743-9.

Efron B, Tibshirani R (1991) Statistical data analysis in the computer age. *Science* **253**:390-395.

Gage TB (1989) Bio-mathematical approaches to the study of human variation in mortality. *Yearbook of Physical Anthropology* **32**:185-214.

Holman DJ (2003) Unobserved heterogeneity. In: Lewis-Beck MS, Bryman A, Liao TF, (eds.) *Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage Publications.

Holman DJ, Grimes MA, Achterberg JT, Brindle E, O'Connor KA. (*in press*) The distribution of postpartum amenorrhea in rural Bangladeshi women. *The American Journal of Physical Anthropology*. (Also: Working Paper 04-08 http://csde.washington.edu/downloads/downloader/dl.pl?id=04-08.pdf, Center for Studies in Demography & Ecology, University of Washington).

Holman DJ, Yamaguchi K (*in press*) Longitudinal analysis of deciduous tooth emergence: IV. Covariate effects in Japanese children. *American Journal of Physical Anthropology.*

Konigsberg LW, and Frankenberg SR (1992) Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology* **89**: 235-256.

Holman DJ, Grimes MA. (2003) Patterns for the initiation of breastfeeding in humans. *The American Journal of Human Biology* **15**:765-780.

Holman DJ, Jones RE (1998) Longitudinal analysis of deciduous tooth emergence II: Parametric survival analysis in Bangladeshi, Guatemalan, Japanese and Javanese children. *American Journal of Physical Anthropology* **105**(2):209-30.

Holman DJ, Brunson E, Newell LL, Jones RE, Streissguth A, (n.d.) Measuring developmental noise from bilateral traits. Manuscript.

Levins R (1966) The strategy of model building in Population Biology. *American Scientist* **54**(4):421-31.

Raftery AE (1995) Bayesian model selection in social research. *Sociological Methodology* **25**:111-195.

Sarton-Miller I, Holman DJ, Spielvogel H (2003) Regression-based prediction of net energy expenditure in children performing activities at high altitude. *American Journal of Human Biology* **15**:554-565.

Sattenspiel L (1990) Modeling the spread of infectious disease in human populations. *Yearbook of Physical Anthropology* **33**:245-276.

Skellam JG (1955) The mathematical approach to population dynamics. In: Cragg JB and Pirie NW. *The Numbers of Man and Animals*.

Weinberg CR, Gladen BC (1986) The beta-geometric distribution applied to comparative fecundability studies. *Biometrics* **42**:547-60.

Wood JW (1998) A theory of preindustrial population dynamics. *Current Anthropology*. **39**(1):99-135.

Wood JW, Holman DJ, Weiss KM, Buchanan AV, LeFor B (1992) Hazards models for human biology. *Yearbook of Physical Anthropology* **35**:43-87.

Software

mle: http://faculty.washington.edu/~djholman/mle. [free]
R: http://cran.r-project.org/. [free SPlus-like language]
Octave: http://www.octave.org/. [free Matlab-like language]
Matlab: http://www.mathworks.com/products/ [commercial]
Gauss: http://www.aptech.com/ [commercial]
Stella: www.hps-inc.com [commercial, for dynamic models]
Madonna: http://www.berkeleymadonna.com/ [commercial, for dynamic models]
Simulink: http://www.mathworks.com/products/ [commercial, for dynamic models]
Mathematica: http://www.wolfram.com/ [commercial, symbolic programming language]
Maple: http://www.maplesoft.com/ [commercial, symbolic programming language]