ENTM 798V: Introduction to *R* for computation and analysis in Ecology and Evolutionary Biology
(1 credit graduate seminar, Fall 2012)

Time: 2 hrs weekly TBA after polling students; **Location:** 4102 Plant Sciences Bldg

Instructors: Alexander Forde [4148 Plant Sciences Bldg., ajforde@umd.edu]
Office hours: 11-12 Tu & Th, 9–10 F or by appointment
Dan Gruner [4142 Plant Sciences Bldg., 5-3957, dsgruner@umd.edu]
Office hours: TBA

What is R?

R is an integrated suite of software facilities for data manipulation, calculation and graphical display (http://www.r-project.org/). This open source software is free, unlike most other off-the-shelf statistical or modeling packages such as SAS, SPSS, Systat and Matlab, and it runs on a variety of platforms including Windows, Linux and MacOS. Among other utilities it has:

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display, and
- a well developed, simple and effective programming language (based on `S') which includes conditionals, loops, user defined recursive functions and input and output facilities.

R is more than a statistical program, although most modern and classical statistical procedures are implemented in one or more of the many *packages* that can be loaded when needed. These tools are developed by statisticians and by users.

R is NOT:

- A commercial, drop-down menu "GUI" program for canned statistics.

R IS:

- Tremendously flexible, versatile, powerful, interactive, and increasingly popular;
- Absolutely, completely, unequivocally FREE forever;
- A challenge, and if you like this sort of thing, a lot of fun. New doors will be open to you.

Text: no required texts, but one book will be used frequently and is strongly recommended:
Zuur, A. F., E. N. Ieno, and E. Meesters. 2009. A Beginner's Guide to R. Springer, New York. (*Use R!* series). This book (along with several others as *.pdf) is posted on the blackboard system (www.elms.umd.edu/).

Beginning with the first meeting, **please bring notebook computers with *R* installed** to work through problems interactively in class; if you cannot bring a computer, please alert the instructor in advance to arrange a loaner to use during the class session. Basic instructions for ***R*** download and installation are posted on elms.

R neophytes and advanced users are equally welcome in this seminar. However, statistical neophytes, or first-year students without data, are asked to postpone enrollment until these limitations are cleared. As this is not a course to teach statistics, the equivalent of *BIOM601 (Biostatistics 1)* is strongly recommended. The seminar will also be more useful if you have preliminary data to play with, and some objectives for those data. The seminar will begin with ~4 instructor-led workshop sessions to acquaint students with basic concepts and operations in the "object-oriented" language of ***R***, frequently used commands, and available resources (books, websites, etc). Thereafter, as needed, instructors may offer irregular short modules (30 min.) with code, tips and specific applications and exercises. The focus of the seminar will shift to student-led workshop modules on selected applications of ***R***. Students may work in small teams but each student must take a leadership role on one topic during the semester. Modules or problem sets can be developed from outside resources (e.g. exercises from a text) or derived directly from a student's research. A list of possible topics will be provided. At all points we will rely on class participation for active input and discussion. Grade will be determined entirely on participation, effort and attendance; please notify the instructor in advance if you expect an absence.

Students are expected to follow the University of Maryland Code of Academic Integrity at all times.