Basic Statistics for Environmental Studies: A Distance-Learning Course

Richard Cellarius, Ph.D., Prescott College, Prescott, AZ
e-mail: rcellarius@prescott.edu
Web: http://myweb.cableone.net/rcellarius/Statistics/ or http://moodle.prescott.edu/course/edit.php?id=316 [key: mapstats]

Texts: Triola, Mario F. (2001). *Elementary Statistics Using Excel* . (1st Ed.) Boston: Addison-Wesley Longman.
ISBN 0-201-69942-7 [This edition is specifically required]
Moore, David S.(2001). *Statistics: Concepts and Controversies* (5th Ed.). New York: W. H. Freeman.
ISBN 0-7167-4008-7 [This edition is specifically required]
Tools: Hand calculator with statistical functions (graphing calculator not required); Microsoft Excel spreadsheet program
Credit: 2-3 semester credits, depending on number of topics completed, Minimum 2 credits for Topics 1-3; 1 credit additional for Topics 4-5.

This course comprises a series of five reading and problem assignments based on the above texts. Each assignment begins with a brief overview of the essential points to focus on. There is also a series of PowerPoint presentations and video files that assist in the elaboration of the concepts and problem-solving techniques. The problems from the texts are to be worked either on paper and hand calculator or using Excel. For the full 3-credit course, these are best completed as indicated on the attached schedule, usually about 3 weeks from when the assignment is started, based on a time commitment of about 9-10 hours per week. Some folks will find a slower pace will fit their other obligations better; however, completion of Topics 1-3 is essential for a basic understanding of the statistical approach to data analysis, and must be completed for the minimum 2-semester credits. Specific details for how to include the course in a study plan and review need to be worked out on an individual basis by the student and Graduate Advisor. Password protected answer files are available for self-directed study. Contact the Prescott College Master of Arts Program (MAP) office or Environmental Studies Faculty for information on setting up the course and access to the course materials.

IMPORTANT—PLEASE NOTE: I am no longer mentoring students in the course. The process I recommend is that when a student includes the course in the semester's study plan, it is up to the student and advisor to determine the process for evaluating it. I have created a set of pdf answer files, which are also available on this site in the section for the respective topic. They are password protected, and once you have finished one of the five topics, you can contact me or Frankie Cardamone (fcardamone@prescott.edu) for the password to the answer file(s) for that topic. I can also send the answer files to your graduate advisor, if he is willing to review your work in detail. My recommendation on the evaluation process is that when you have completed a topic, you evaluate your answers using the answer key(s) and report the results—number correct, number in the right direction but with things like numerical errors as minimal deviations from the correct answer, number totally missed—to your advisor. Your final self-evaluation for the course should summarize what you learned and how you did. Your advisor's evaluation could support that conclusion based on your reporting of your overall performance. One simple alternative for stronger evaluation is to have your GA review and correct your answers to the problems specifically identified as "self-evaluation" at the end of Topics 3 etc.

Topic 1 : Overview; Descriptive Statistics – Sampling, the Nature of Data, Displaying Data
This first installment introduces some of the initial concepts of **data acquisition** and **description** . This topic covers a great many topics that might be review of previous learning, at least in part. The fundamental issue here is to understand the different types of data and how data should properly be obtained and displayed. Data can often be displayed in a way that misleads or hides critical information, and it is important to be able to recognize this and evaluate those data more correctly.

Topic 2 : Probability and Probability Distributions; Sampling Distributions and the Central Limit Theorem
Probability is one of the most fundamental concepts underlying statistical analysis: given a number of events, what is the probability that a specific event will occur or, if there are repetitive events, what is the proportion of a specific kind of event that is likely to occur. For statistics, probability can be approached with a full-blown coverage of probability analysis or with a survey of the basic fundamentals sufficient to demonstrate that the statistical procedures encountered later have an adequate mathematical basis. I have chosen a route closer to the latter, which does not require learning elaborate notations or doing excessive calculations. The ultimate conclusions to focus on are the nature of **probability distributions** including the **Normal Distribution**, **sampling distributions** , and the **Central Limit Theorem** , which provides the fundamental rationale for relying on random samples of a population to make statistical conclusions about the entire population.

Topic 3 : Introduction to Inferential Statistics – Confidence Intervals; Hypothesis Testing and Tests of Significance
The two parts of this Topic introduce the essential concepts and tools of **Inferential Statistics** , the science and art of making conclusions about ***populations*** from ***statistical samples*** of those populations. In the first part, the task is to determine the probable range within which the "true" population value of the quantity in question lies based on analysis of the sample data. In the second part, the process is extended to make comparative statements or conclusions based on **hypotheses** about the population, again on the basis of the sample data from that population. (Supplementary file Temps-Table 6-1.xls)

Topic 4 : More Tests of Hypotheses: Inferences from Two Samples; Multinomial Experiments and Contingency Tables
The two tools covered in this Topic are among the most widely used statistical procedures. The first, "Inferences from Two Samples," generally follows quite directly and quite simply from the hypothesis testing of a single sample that was the subject of the previous topic; here one of two questions is asked: (1) "are the means of the populations represented by two different samples the same or different?" or (2) "has the population changed as the result of the treatment?" The second tool involves making inferences from data gathered when there are a number of categories, for example races of people or varieties of trees. Again we will consider two different

types of questions, (1) “how does the distribution in the categories compare with an assumed distribution?” which applies when you’re dealing with a sample from a single population – **a multinomial experiment** – often referred to as a **Chi-Square Test** by biologists, or (2) “are the distributions among the various categories the same or different in the different samples?” which involves analysis of **contingency tables**.

Topic 5 : Two Important Analysis Tools: Analysis of Variance (ANOVA) and Correlation & Regression
The two tools described here are again two of the most important tools used in statistical analysis, and for the purposes of this course, they complete our study of basic statistical analysis techniques. W cover them in the reverse order that they are discussed in the text, primarily because **Analysis of Variance** (**ANOVA**) is a logical extension to multiple samples of the hypothesis testing of sample data encountered in one and two sample tests (Topics 3 and 4). In contrast **Correlation** and **Regression** deal with the analysis of a different type of data: pairs of sample data, such as weight and height of individuals, weight and girth of bears, or time and distance for a mode of travel or a race.

Prescott College -- Standard MAP Term Schedule					
Week	**Day**	**Date**	**Mail Study Packets**	**Start Statistics Topic (Recommended dates in order to complete all five topics)**	**Date to Complete Statistics Topic**
	Fri		**Start of First Colloquium of the Term**		
1	Mon			**Order texts**	
	Fri				
2	Mon		Study plans due	**Topic 1: Overview, Sampling, the Nature of Data, Displaying Data**	
	Fri				
3	Mon				
	Fri				
4	Mon		1st		
	Fri				Topic 1
5	Mon			**Topic 2: Probability and Probability Distributions; Sampling Distributions and the Central Limit Theorem**	
	Fri				
6	Mon				
	Fri				
7	Mon				
	Fri				Topic 2
8	Mon		2nd	**Topic 3: Introduction to Inferential Statistics – Confidence Intervals; Hypothesis Testing and Tests of Significance**	
	Fri				
9	Mon				
	Fri				
10	Mon		3rd		
	Fri				Topic 3
11	Mon			**Topic 4: More Tests of Hypotheses: Inferences from Two Samples; Multinomial Experiments and Contingency Tables**	
	Fri				
12	Mon				
	Fri				Topic 4
13	Mon		4th	**Topic 5: Two Important Analysis Tools: Analysis of Variance (ANOVA) and Correlation & Regression**	
	Fri		**Second Colloquium**		
14	Mon				
	Fri				
15	Mon				Topic 5
	Fri				
16	Mon		5th & EoT summary		
	Fri				
EoT	Sat		**End of Term; EoT materials due in MAP office from advisor**		