Fundamentals of Statistics - ORF 245/ EGR 245 (Spring 2015)
Syllabus

1 Schedule


Precepts (coordinated every week to cover similar material):
- Mon., 3:30 pm to 4:20 pm, Sherrerd Hall 101. By: Huanran Lu.
- Mon., 7:30 pm to 8:20 pm, Sherrerd Hall 101. By: Yuyan Wang.
- Tues., 3:30 pm to 4:20 pm, Friend Center 004. By: Byron Vickers.
- Tues., 7:30 pm to 4:20 pm, Sherrerd Hall 101. By: Kobby Aboagye.

Office hours:
- Wed. 1:30pm to 3pm: Huaran Lu (Sherrerd Hall, Library).
- Wed. 3pm to 4:30 pm: Byron Vickers (Sherrerd Hall, Library).
- Wed. 6pm to 7:30 pm: Samory Kpotufe (Sherrerd Hall, 327).
- Thur. 11am to 12:30 pm: Yuyan Wang (Sherrerd Hall, Library).
- Thur. 1:30pm to 3pm: Kobby Aboagye (Sherrerd Hall, Library).

2 Prerequisites

A good understanding of calculus, and some multivariate calculus. Basic counting (e.g., remember $n!$, \( C(n, k) \) a.k.a \( \binom{n}{k} \), \( P(n, k) \), and such?) will be assumed. However, initial precepts will review counting.

3 Course Evaluation

Exams

There will be a midterm and final exam each worth 20% of your grade. Any material covered in lectures or in problem sets and programming assignments can make it into the exams.

Homework (problem sets)

Weekly homeworks will be posted (usually on Friday) and due on Fridays at 5pm. Late homeworks will not be accepted, except under extreme circumstances.

A homework will consist of a number of problems\(^1\), where each problem is worth 10 points. Together, homeworks will be worth 40% of your grade.

Collaboration: Homeworks can be discussed but must be written up individually. The goal of homeworks is to help you exercise those concepts seen in class.

Programming assignments

Frequent programming assignments will be given, together worth 20% of a grade. Programming will be in Matlab. Programming assignments might be late by 48hrs, but will not be accepted beyond that.

Collaboration: Programming can be discussed but must be written up individually. You must write your own code, and turn in a printout of the code with your report. The aims of the assignments are (1) that you become familiar with important tools such as Matlab, (2) and that you get concrete evidence of abstract concepts seen in lectures.

---

\(^1\)Don’t tell anyone, but some problems will have solutions in the book. This is to encourage you to read the book.
4 Reading

Introductory books on Statistics sort of follow the same outline with rather standard topics. The decision to use Jay Devore, Probability and Statistics, 9th Edition as the main text came down to the clarity of notation and examples, the outlining of important concepts, and my liking books with colors in them.

The second book Sheldon M. Ross 5th Ed., Introduction to Probability and Statistics for Engineers & Scientists is a nice addition with different working examples.

5 Overview of topics

We will loosely follow the order of topics below, keeping in mind that some concepts will often appear throughout the course and some are best learned together.

5.1 Introductory ideas [1 week]
- Sample (observations) vs. Population (underlying truth).
- Notions of typicality (simple ways to summarize samples or population): mean, median, mode.
- Proportion/relative-frequency, histograms (a more sophisticated summary).

Note: We'll also introduce concepts such as random variables, estimation, variance.

5.2 Probability (essentially, proportions in the population) [1 week]
- Events, and probabilities of events, union and intersection of events.
- Conditioning and Bayes rule. Independence of events. Independent trials.

5.3 Random variable (r.v.) and important distributions [3 weeks]
- Discrete or continuous r.v. – call it $X$ – and its distribution (law governing the probability of events associated with $X$). Expectation and Variance of $X$.
- Sampling from a distribution (precept and programming).

Note: We'll also introduce concepts such as estimation error (mostly through programming).

5.4 Estimation (of distributions) and confidence [2 weeks]
- Maximum Likelihood principle: Bernouilli, Normal. (Bayesian MAP in precept and assignments).
- Bias, Variance principle: e.g. yields a different estimate of variance.
- Confidence intervals: Gaussian, Bernouilli.

5.5 Hypothesis testing [2 weeks]
- $p$-values, type-I and type-II errors, power of a test.
- Tests on a Normal population.
- Tests on a Bernouilli population.

5.6 Regression [2 weeks]
- Linear regression (least-squares).
- NonLinear regression (polynomial, maybe kernel regression).

5.7 Classification (if time permits) [1 week]
Simple classification rules such as nearest neighbors, perceptron (a simple neural network).

... so let’s have fun! 🎈