# Probability And Statistics for Elementary Teachers at Western Oregon University 

Scott Beaver - Western Oregon University

MAA MathFest

August 3, 2007

## Course Structure

- Course meets for two hours, twice each week - once in a standard classroom and once in a computer lab
- The two subjects are developed in parallel wherever possible
- Statistics text is Workshop Statistics: Discovery with Data and Fathom ${ }^{\text {TM }}$
- No probability text
- About one-quarter of classtime is group work


## Two Surprises

Discuss, play the games and take (lots of) data

- The Monty Hall Problem
- Non-Transitive Dice


## Preliminary Definitions

- Probability: experiment, outcome, sample space, event, random variables, probability of an event
- Statistics: datum, data, sample, population, parameter, statistic

Note that a set of data can have repeated values unlike a mathematical set

Discuss raw data and derived data like sample proportions

## Fair and Unfair Games

Analysis of simple games is used as a tool for practicing sample space enumeration and probability computation

Typically, a game is presented, students construct a probability-based analysis, play it and take data, and then consider the statistics

## Display

- Begin with dotplots; discuss various features
- Dotplots lead to histograms or bar graphs
- Granularity of a dotplot leads to a discussion of the similarities between discrete quantitative and categorical RV's
- We also note that the real-life constraint of finite precision reduces all continuous RV's to discrete RV's
- Relabeling a histogram/bar graph with proportions instead of raw numbers leads to the idea of a probability distribution


## Measures of Center and of Spread

- Define the parameters and statistics and give symbols used
- Note potential misrepresentation of average by the mean, and resistance of median to single datum change
- Brief mention of why sample standard deviation has division by $n-1$ rather than $n$ (unbiased estimator of population standard deviation)
- Students generate small data sets having particular properties


## Flipping 5 coins, or one coin 5 times

An extremely rich topic:

- Enables a discussion of the occasional non-uniqueness of the experiment-RV pair
- Leads to combinations and discussion of when order matters
- Under assumption of fairness, construct probability distribution, get in groups of three and perform the 5-coin experiment 50 times each
- Compare individual and aggregate proportion distributions with theoretical - leads to discussion of sample size
- Discuss what we'd be able to conclude in the absence of fairness assumption
- Discuss what would happen if a different number coins were flipped - leads toward binomial distribution


## Normal Distribution

Derived as limiting case of binomial distribution

Brief mention of the calculus involved, but we just use the table to compute the four standard kinds of probabilities associated with z-scores

Largely in anticipation of central limit theorem and confidence intervals

## Sampling

- Essentially performing a single-stage experiment repeatedly
- Most interesting experiments are those involving people and their opinions or activities - very little theoretical understanding, so we take statistics
- Discuss bias; sampling and non-sampling sources
- Discuss Literary Digest debacle, and pitfalls of mailed questionnaire surveys


## Sample Means, the Central Limit Theorem, and Confidence Intervals

If the population is normally distributed with respect to some RV with population mean $\mu$, then the sample means will be distributed normally about $\mu$ as well

This leads to conclusions about the chances of getting a sample mean within some range

Dually, if we assume that the population RV is normally distributed with unknown $\mu$, we can draw conclusions about the chances of $\mu$ being in some range

Question: which of the above two results is more useful?

## Association of two RV's

Qualitative treatment of scatterplots

Correlation does not imply causation

Small correlation coefficient does not imply lack of causation

## Misleading Statistics

We survey various techniques often employed to mislead consumers or voters

- Two- or three-dimensional figures displaying one-dimensional data
- Y-axis expansion
- Excision of part of the $y$-axis
- Omitted measure of spread
- Use of raw data rather than proportions
- And many more!


## Counting

The final 2-3 weeks of the course is a systematic treatment of more difficult probability calculations involving counting arguments

- Multiplication and addition rules
- Combinations and permutations
- Complements, replacement, independence, and mutually exclusive events
- Conditional probability

We always do the Birthday Problem, but with 17-20 students in the class, we've never had a duplicated birthday...

## Assessment

- Weekly homework
- Two midterms and a final exam
- Three-class lesson plan project


## Contact Information

Scott Beaver - Western Oregon University

beavers@wou.edu

www.wou.edu/~beavers

