

# *DATA ANALYSIS AND PROBABILITY FOR TEACHERS*

JIM ALBERT

## BRIEF TABLE OF CONTENTS

JUNE 2008

### **EXPLORING DATA**

Topic D1: Statistics, data, and variables

Topic D2: Graphing data

Topic D3: Summaries of data

Topic D4: Comparing batches and relative standing

Topic D5: Relationships between categorical variables

Topic D6: Relationships between quantitative variables

Topic D7: Relationships – summarizing by a line

### **COLLECTING DATA**

Topic C1: Obtaining Data by Sampling

Topic C2: Obtaining Data by Experiments

### **PROBABILITY**

Topic P1: Probability – a measure of uncertainty

Topic P2: Sample space and assigning probabilities

Topic P3: Let me count the ways

Topic P4: Computing probabilities by simulation

Topic P5: Conditional Probability

Topic P6: Probability distributions

Topic P7: Coin tossing distributions

Topic P8: Continuous distributions

Topic P9: The Normal distribution

### **INTRODUCTION TO STATISTICAL INFERENCE**

TOPIC I1: Introduction to Inference: Estimating a Proportion

TOPIC I2: Introduction to Hypothesis Testing

TOPIC I3: Learning about a Population Mean

# DATA ANALYSIS AND PROBABILITY FOR TEACHERS

JIM ALBERT

## EXTENDED TABLE OF CONTENTS

JULY 2008

### EXPLORING DATA

Topic D1: Statistics, data, and variables

SPOTLIGHT: The U.S. Census Bureau

WARM-UP ACTIVITY: Getting to know you

Statistics and data

Organizing and summarizing data: some initial thoughts

ACTIVITY: Meet the states data

Reading articles in the media

Graphs in the media

TECHNOLOGY ACTIVITY: Introduction to *Tinkerplots*

CLASSROOM CAPSULE: Children's Well-Being

EXERCISES

Topic D2: Graphing data

SPOTLIGHT: Who are the baseball players?

Graphing categorical data

Graphing quantitative data – distribution and shape

Stemplot

Histogram

Experiment with different graphs

TECHNOLOGY ACTIVITY (Fathom): Choosing the bins of a histogram

ACTIVITY: The shape of the data

Classes of data and shapes

ACTIVITY: Matching variables and shapes

CLASSROOM CAPSULE: Graphing health and liking school

EXERCISES

Topic D3: Summaries of data

SPOTLIGHT: Nutrition value of ice cream?

Summarizing categorical data

How many calories are there in an “average” scoop of ice cream? (Introducing the median)

Deviations and the mean

Geometrical interpretation of the mean

Comparing the median and the mean

Measures of spread: quartiles and the IQR

Measures of spread using deviations

Interpreting s: the 68/95/99.7 rule for bell-shaped data

ACTIVITY: Collecting some data on cities

ACTIVITY: V is for Variation

TECHNOLOGY ACTIVITY (Fathom): Deviations, the Mean, and Measures of Spread

ACTIVITY: Measurement bias

ACTIVITY: Matching statistics to histograms

CLASSROOM CAPSULE: Summarizing risky behavior

EXERCISES

Topic D4: Comparing batches and relative standing

SPOTLIGHT: Where's the best place to live?

Comparing batches of categorical data

Comparing batches of quantitative data

Relative standing

Flagging possible outliers

ACTIVITY: Comparing men and women in the class dataset

ACTIVITY: Matching statistics with boxplots

ACTIVITY: Counting pasta

CLASSROOM CAPSULE: Comparing achievement scores of rich and poor countries

EXERCISES

Topic D5: Relationships between categorical variables

SPOTLIGHT: The Titanic: women and children first?

Titanic survival data

ACTIVITY: Predictable pairs

Simpson's Paradox [TO BE DONE]

CLASSROOM CAPSULE: Exploring two-way tables

EXERCISES

Topic D6: Relationships between quantitative variables

SPOTLIGHT: Measuring climate

Introduction: Looking at weather data

Relationships – scatterplots

A simple correlation formula – the QCR

A correlation coefficient

Interpreting the correlation coefficient

ACTIVITY: Matching correlations with scatterplots

TECHNOLOGY ACTIVITY (Fathom): Guessing correlations

EXERCISES

Topic D7: Relationships – summarizing by a line

SPOTLIGHT: Measuring a car

Relationships – summarizing by a least-squares line  
Making appropriate and inappropriate predictions  
ACTIVITY: Fitting a line by eye to Galton’s data  
The median-median line – a robust alternative method of fitting a line  
TECHNOLOGY ACTIVITY (Fathom): Fitting a “best” line  
Plotting residuals  
TECHNOLOGY ACTIVITY (Fathom): Exploring some Olympics data  
ACTIVITY: Regression to the mean  
Different ways of looking at relationships  
TECHNOLOGY ACTIVITY: Using *Tinkerplots* to study relationships  
CLASSROOM CAPSULE: Summarizing association between two measures of family health  
EXERCISES

## COLLECTING DATA

### Topic C1: Obtaining Data by Sampling

SPOTLIGHT [TO BE DONE]  
Population and sample  
Examples of popular sampling methods: Elvis Presley and Alf Landon  
Simple random sampling  
ACTIVITY: Random rectangles  
TECHNOLOGY ACTIVITY (Fathom): Biased sampling of rectangles  
Practical concerns in sampling  
EXERCISES

### Topic C2: Obtaining Data by Experiments

SPOTLIGHT [TO BE DONE]  
An apple a day: different ways of collecting data  
Music and math achievement  
An experiment to detect the Mozart effect  
Basic principles of experiments  
ACTIVITY: Jumping frogs  
EXERCISES

## PROBABILITY

Topic P1: Probability – a measure of uncertainty  
SPOTLIGHT: How Risky is ...?  
WARM-UP ACTIVITY: Some questions on probability  
The classical view of a probability

The frequency view of a probability

ACTIVITY: Tossing and spinning a poker chip

The subjective view of a probability

A calibration experiment

CLASSROOM CAPSULE: Thinking about probability

EXERCISES

Topic P2: Sample space and assigning probabilities

SPOTLIGHT: The casino game of Roulette

WARM-UP ACTIVITY: Writing down some sample spaces

Different representations of a sample space

Assigning probabilities

A more formal look at probability

The three probability axioms

The complement and addition rules

CLASSROOM CAPSULE: What can happen?

EXERCISES

Topic P3: Let me count the ways

SPOTLIGHT: Rolling dice and Yahtzee

Equally likely outcomes

The multiplication rule

Permutations

Combinations

Arrangements of non-distinct objects

Which rule?

Playing Yahtzee

ACTIVITY: Mothers and babies

ACTIVITY: Sampling from a bag

CLASSROOM CAPSULE: Playing scrabble

EXERCISES

Topic P4: Computing probabilities by simulation

SPOTLIGHT: Buffon's needle simulation

Simulating a lottery game

Basic components of a simulation experiment

The collector's problem

TECHNOLOGY ACTIVITY (Fathom): Mixed-up letters

ACTIVITY (TI -84 Plus Calculator, Fathom): The longest run

ACTIVITY (TI -84 Plus Calculator, Fathom): Sampling people from a room

ACTIVITY (TI -84 Plus Calculator, Fathom): : The birthday problem

CLASSROOM CAPSULE: Random ties

CLASSROOM CAPSULE: Waiting in line

## EXERCISES

### Topic P5: Conditional Probability

SPOTLIGHT: The three-card problem

New information, reduced sample space, and conditional probability

Conditional probability in everyday life

Conditional probability in a two-way table

Definition of conditional probability and multiplying probabilities

The multiplication rule under independence

Learning using Bayes' rule

TECHNOLOGY ACTIVITY (Fathom): Rolling two dice

TECHNOLOGY ACTIVITY (Fathom): How many defectives?

CLASSROOM CAPSULE: Spinning a random spinner

## EXERCISES

### Topic P6: Probability distributions

SPOTLIGHT: The hat check problem

A random variable

Summarizing a probability distribution

Standard deviation of a probability distribution

Interpreting the standard deviation for a bell-shaped distribution

TECHNOLOGY ACTIVITY (Fathom): Constructing a probability distribution by simulation

TECHNOLOGY ACTIVITY (Fathom): Playing roulette

ACTIVITY: How many keys?

ACTIVITY: Investing money: comparing safe and risky investments

CLASSROOM CAPSULE: Playing the lottery

## EXERCISES

### Topic P7: Coin tossing distributions

SPOTLIGHT: Galton's board

Probabilities of a coin-tossing experiment

Binomial experiments

ACTIVITY: Coin flipping: Is it Real or Fake?

TECHNOLOGY ACTIVITY (Fathom): Simulated coin flipping

ACTIVITY: Is a professional athlete streaky?

Binomial computations

Mean and standard deviation of a binomial

Negative binomial experiments

ACTIVITY: Graphing binomial and negative binomial experiments

CLASSROOM CAPSULE: Probabilities of the Galton board

## EXERCISES

### Topic P8: Continuous distributions

SPOTLIGHT: A spinner baseball game

The uniform distribution

A probability density/waiting for a bus

The probability function (the cdf  $F(x)$ )

Finding probabilities using  $F$

Summarizing a continuous random variable

Percentiles

TECHNOLOGY ACTIVITY (Fathom): Spinning away

TECHNOLOGY ACTIVITY (Fathom): Waiting for the shuttle

TECHNOLOGY ACTIVITY(TI-84 Plus or Fathom): A test with a bimodal distribution

TECHNOLOGY ACTIVITY: Estimating areas by simulation [TO BE DONE]

CLASSROOM CAPSULE: Census ages

EXERCISES

Topic P9: The Normal distribution

SPOTLIGHT: Early use of the normal curve

Modeling data by a normal curve

Computing normal probabilities

Computing normal percentiles

Binomial probabilities and the normal curve

Sampling distribution of the mean: the central limit theorem

The central limit theorem works for any population

TECHNOLOGY ACTIVITY (Fathom): Sampling heights

TECHNOLOGY ACTIVITY (Fathom): Rolling biased dice

CLASSROOM CAPSULE: Sampling from a U-shaped population

EXERCISES

## **INTRODUCTION TO STATISTICAL INFERENCE**

TOPIC I1: Introduction to Inference: Estimating a Proportion

SPOTLIGHT [TO BE DONE]

A classroom survey

Population, parameter, sample, and statistic

Sample estimates: bias and variance

TECHNOLOGY ACTIVITY (Fathom): The taxi problem

Construction of a confidence interval

A large sample confidence interval for a proportion

Understanding a confidence interval

Choosing a sample size

Some cautions

TECHNOLOGY ACTIVITY (Fathom): Penny ages

EXERCISES

TOPIC I2: Introduction to Hypothesis Testing

SPOTLIGHT [TO BE DONE]

Introduction: A taste test

Stating the hypotheses

A statistical test

A large-sample test for a proportion

Statistical significance

Two-sided tests and confidence intervals

Decisions, two-errors, and confidence

TECHNOLOGY ACTIVITY (Fathom): Is the machine working?

EXERCISES

TOPIC I3: Learning about a Population Mean

SPOTLIGHT [TO BE DONE]

How long is a cell phone call?

Review: the pattern of sample means

A confidence interval for a mean

Understanding a confidence interval for a mean

Testing about a mean

More about a test for a mean

ACTIVITY: Estimating family size

ACTIVITY: Estimating the total of a restaurant bill

TECHNOLOGY ACTIVITY (Fathom): Is the neighborhood expanding?

EXERCISES