Problems and Sample Exams for

STATISTICS

A CONCISE MATHEMATICAL INTRODUCTION
FOR STUDENTS, SCIENTISTS, AND ENGINEERS

DAVID W. SCOTT
Rice University
Houston, Texas

October 13, 2020

Contents

Chapter 1

Data Analysis and Understanding

1. In a typical introductory statistics course, the construction of a histogram from data boils down to the choice of the number of bins, k. Equivalently, you can focus on the bin width, h, which would be found by dividing the width of the sample range (a, b) by the number of bins: $h = (b - a)/k$. Often you will be taught that the choice of k should be any "convenient" value. In Section 9.1, we learn that the formula $h^* = 1.06\, s_x\, n^{-1/5}$ can be justified for normal data.

   (a) Generate a normal random sample in R by

   ```
   set.seed(246); x = rnorm(10000) .
   ```

   The standard deviation, s_x, is `sd(x) = 1.012`. Examine the default histogram in R by `hist(x,col=2)`. What value of k is used? (Hint: $k = 16$.) Show the corresponding $h = 0.474$. Look at two other histograms choosing bin widths $h/2$ and $2h$, using the optional second argument `breaks` via

   ```
   hist( x, breaks = seq(-4.5,4.5,.948), col=2 )
   ```

   (b) The formula given above results in $h^* = 0.170$. Examine the 3 histograms using $h^*/2$, h^*, and $2h^*$.

   (c) Discuss and evaluate your favorite histogram among these six.

   Hint: Use the on-line R help by invoking `help.start()`*.*

2. TBA

1.1 Exploring the Distribution of Data

1. TBA

1.1.1 Pearson's Father-Son Height Data

1. TBA

1.1.2 Lord Rayleigh's Data

1. TBA

1.1.3 Discussion

1. TBA

1.2 Exploring Prediction Using Data

1. TBA

1.2.1 Body and Brain Weights of Land Mammals

1. TBA

1.2.2 Space Shuttle Flight 25

1. TBA

1.2.3 Pearson's Father-Son Height Data Revisited

1. TBA

1.2.4 Discussion

1. TBA

1.3 Problems

1. A **frequency histogram** of continuous data is constructed by counting the number of data points that fall into equally-spaced bins of width h. h is called the **bin width**. Typically the bin edges are 0, $\pm h$, $\pm 2h$, $\pm 3h$, and so on. If the bin count in the k^{th} bin is denoted by ν_k, then the frequency histogram is defined as

$$\boxed{\hat{f}(x) = \nu_k\,, \qquad \text{for } x \text{ in the } k^{th} \text{ bin.}} \tag{1.1}$$

   (a) Show that the total area of the frequency histogram is nh, where $n = \sum_k \nu_k$. *Hint: The histogram is made up of rectangular blocks of width h and height v_i.*

   (b) A **probability histogram** is defined to have total area of one. Show that the following definition of a histogram has area equal to one:

$$\boxed{\hat{f}(x) = \frac{\nu_k}{nh}\,, \qquad \text{for } x \text{ in the } k^{th} \text{ bin.}} \tag{1.2}$$

2. One of the most famous epidemiological cases occurred in 1854 when Dr. John Snow successfully tracked down the source of an outbreak of **cholera** in the London suburb of SoHo. He mapped the households of some 500 victims over a 10-day period that lived within a quarter of mile of each other. However, many tens of thousands had died of cholera in England during the prior two decades. Dr. Snow believed contaminated water was a primary cause. Just as in the space shuttle example, there are choices of an appropriate time interval and the geographical extent that can influence your conclusions. Using the descriptions and maps conveniently assembled at

   `http://www.ph.ucla.edu/epi/snow/snowcricketarticle.html` ,

   discuss the evidence and choices that were and could have been made. *Hint: These data have been conveniently collected in CRAN Library* `HistData` *by Michael Friendly. Look at the help file for dataset* `snow` *and its example code.*

3. (a) The Tukey power transformation of a variable x is x^λ for any nonzero $\lambda \in \mathbb{R}^1$. To better understand why the $\log(x)$ is used in

place of x^0 when $\lambda = 0$, we consider the linear re-expression of the Tukey transformation given by the formula

$$\frac{x^\lambda - 1}{\lambda} \,. \tag{1.3}$$

Since (1.3) is 0/0 when $\lambda = 0$, use l'Hôpital's rule to find the limit transformation as $\lambda \to 0$. The scatter diagram using either formula for fixed nonzero λ will be visually identical. Formula (1.3) is referred to as the Box-Cox transformation; see Figure 1.1.

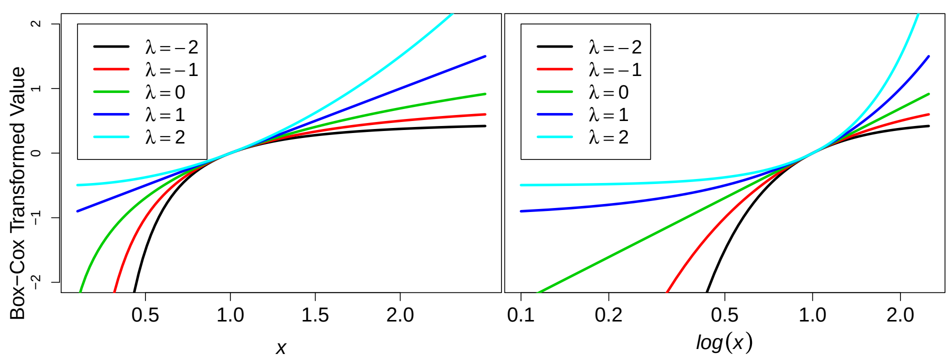

Figure 1.1: Box-Cox Transformation on natural and log scales

(b) Sometimes the transformation $\log(1+x)$ is in place of $\log(x)$ when $x \geq 0$ and x can take on the value 0. In this case, the original and transformed values of 0 are both 0. Try this form on the body-brain data and compare to Figure 1.4.

Chapter 2

Classical Probability

2.1 Experiments with Equally Likely Outcomes

1. TBA

2.1.1 Simple Outcomes

1. TBA

2.1.2 Compound Events and Set Operations

1. **Simple Set Operations:** Find the following simple sets:

   (a) If $A \supseteq B$, what is $A \cup B$?

   (b) If $A \supseteq B$, what is $A \cap B$?

   (c) $A \cap A^c$?

   (d) $A \cup A^c$?

2. **Counting Events:** List *all* possible events associated with a single toss of a regular die (6-sided) with faces showing the digits 1, 2, 3, 4, 5, or 6. **Hint:** Record the result of each roll of the die by the digit on the top.

3. **Counting Events (Part II):** Show

   (a) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

(b) $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$

(c) What is the application to the previous die rolling problem?

4. TBA

2.2 Probability Laws

1. When rolling a pair of die, find

   (a) the probability getting at least 1 six?

   (b) their sum is less than 5?

2. TBA

2.2.1 Union and Intersection of Events A and B

1. TBA

2.2.1.1 Case (i):

1. TBA

2.2.1.2 Cases (ii) and (iii):

1. TBA

2.2.1.3 Case (iv):

1. TBA

2.2.2 Conditional Probability

1. TBA

2.2.2.1 Definition of Conditional Probability

1. (a) If $A \cap B = \emptyset$, find $P(A|B)$?

   (b) If $B \subset A$, find $P(A|B)$?

2. When rolling a pair of dice, define the events

$$\begin{aligned} A &= \{\text{at least one die shows a } 3\} \\ B &= \{\text{sum is at least } 5\} \end{aligned}$$

   Find

   (a) $P(B|A)$?

   (b) $P(A|B)$?

3. TBA

2.2.2.2 Conditional Probability With More Than Two Events

1. **Conditional probability with 3 events:** Suppose a shipment of 25 replacement iPhone screens contains 4 cracked displays. Three are chosen sequentially at random from the shipment box. What is the probability that the first two are OK but the third is broken?

2. **Alternative Birthday Problem:** How many individuals with random birthdays (of 365 possible) would make the probability of at least one duplicate birthday at least 90%?

3. **Scrabble Problem:** In the game of Scrabble, each player draws a tile to see who draws the lowest letter to see who goes first. (If players tie with the lowest letter, they draw again.) Here is a table of all 100 tiles, including two blank ('b') tiles. A blank tile beats the letter 'A.'

b	A	B	C	D	E	F	G	H	I	J	K	L	M
2	9	2	2	4	12	2	3	2	9	1	1	4	2
2	11	13	15	19	31	33	36	38	47	48	49	53	55
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	6	8	2	1	6	4	6	4	2	2	1	2	1
	61	69	71	72	78	82	88	92	94	96	97	99	100

(a) In a game with 4 players, what is the probability you would go first if you drew the letter 'J'? *Hint: This is a conditional probability given that you drew the letter 'J'.*

(b) For all 27 letters, what is the probability you not be immediately eliminated from going first? Graph these probabilities. *Hint: Unlike part (a) where you cannot be tied, here you can be tied but not eliminated.*

4. TBA

2.2.3 Independent Events

1. **Independent events identity:** Suppose $A_1, A_2, \ldots, A_n$ are independent events. Show

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n} P(A_i^c) \, .$$

2. TBA

2.2.4 Bayes Theorem

1. **Equivalence of Independent Events:** Prove that if $\{A, B\}$ are independent events, then so are $\{A, B^c\}$, $\{A^c, B\}$, and $\{A^c, B^c\}$.

2. TBA

2.2.5 Partitions and Total Probability

1. TBA

2.3 Counting Methods

1. A statistics professor is writing a quiz that covers 5 topics. She plans to reuse questions from previous semesters, one question per topic. If she has 3, 2, 4, 1, and 3 questions for each topic, respectively, how many different exams could she write?

2. TBA

2.3.1 With Replacement

1. TBA

2.3.2 Without Replacement (Permutations)

1. A research group meets once a month for a beer tasting excuse. One person volunteers to pick the order that the 6 beers will be offered, and to record the scores 1-6. Assuming the order in which the beers is presented matters, how many orderings are there?

2. In the previous problem, suppose 3 of the beers are pale ales, while 3 are its stronger couisin, IPA's. As a result, the pale ales will be offered before the IPA's. How many orderings are there?

3. TBA

2.3.3 Without Replacement Nor Order (Combinations)

1. In a department of statistics with 10 faculty members, the department chair has determined she has sufficient funds for 3 to attend the joint statistical meetings. Including herself, how many groups of 3 must she consider?

2. In the previous problem, suppose the department chair determines she must attend the meetings in order to represent the department. How many groups must she consider in this scenario?

3. TBA

2.3.4 Examples

1. **Some Poker Hands:** For 5-card poker hands, find

   - (a) ν(3 of a kind);
   - (b) ν(full house);
   - (c) ν(1 pair);

- (d) ν(Jack high).

Note: $\nu(\cdot)$ *denotes the "number of ways."*

2. TBA

2.3.5 Extended Combinations (Multinomial)

1. **Counting Committees:** From a Board of Directors with 4 men and 5 women, how many committees can be formed with

   - (a) 2 men and 3 women;
   - (b) 5 members, with at least 3 women?

2. TBA

2.4 Countable Sets: Implications As $n \to \infty$

1. TBA

2.4.1 Selecting Even or Odd Integers

1. TBA

2.4.2 Selecting Rational Versus Irrational Numbers

1. TBA

2.5 Kolmogorov's Axioms

1. **A Probability Inequality:** Prove $P(A \cup B) \leq P(A) + P(B)$.

2. **A Decomposition of the Union of 3 Sets:** What is $P(A \cup B \cup C)$ in terms of known probabilities of events A, B, C, $A \cap B$, $A \cap C$, $B \cap C$, $A \cap B \cap C$? **Hint:** You probably can guess what the formula is if you draw a general Venn diagram with 3 intersecting circles. But use the formula for $P(U \cup V)$; then let $U = A$ and $V = B \cup C$ to work it all out.

---

3. TBA

2.6 Reliability: Series versus Parallel Networks

1. TBA

2.6.1 Series Network

1. TBA

2.6.2 Parallel Network

1. TBA

2.7 Problems

1. Show that the formula (2.8) for case (iv) also covers cases (i)-(iii).

2. How does $P(T|C)$ change depending on increasing $P(C)$ versus the PSA test characteristics?

3. Show $P(A|B) + P(A^{\mathbf{c}}|B) = 1$, as does $P(A|B^{\mathbf{c}}) + P(A^{\mathbf{c}}|B^{\mathbf{c}}) = 1$.

4. Examine how $P(C|T)$ improves as a function of $P(C)$, $P(T|C)$, and $P(T^{\mathbf{c}}|C^{\mathbf{c}})$.

5. Prove these and/or draw Venn Diagrams; Equations (2.4) and (2.5).

6. Show that for case (ii), $P(B|A)$ in Equation (2.13) can be expressed in terms of $P(A)$ and $P(B)$ to reach the same conclusion.

7. Show formula (2.22) is correct.

8. Prove that a series network is no more reliable than its weakest component, while a parallel network is more reliable than its strongest component.

Chapter 3

Random Variables and Models Derived From Classical Probability and Postulates

3.1 Random Variables and Probability Distributions: Discrete Uniform Example

1. TBA

3.1.1 Toss of a Single Die

1. TBA

3.1.2 Toss of a Pair of Dice

1. TBA

3.2 Continuous Uniform Example: The Univariate Probability Density Function

1. TBA

3.2.1 Using the PDF to Compute Probabilities

1. TBA

3.2.2 Using the PDF To Compute Relative Odds

1. TBA

3.3 Summary Statistics: Central and Non-Central Moments

1. TBA

3.3.1 Expectation, Average, and Mean

1. Suppose the domain of a continuous random variable, X, is $(0, \infty)$, show
$$EX = \int_0^\infty \Pr(X > x)\, dx.$$
Hint: Use integration by parts.

2. TBA

3.3.2 Expectation as a Linear Operator

1. TBA

3.3.3 The Variance of a Random Variable

1. Suppose you are given the two expectations EX and $E[X(X-1)]$ for a random variable X. How can you use these to compute the variance σ_X^2?

2. TBA

3.3.4 Standardized Random Variables

1. TBA

3.3.5 Higher Order Moments

1. TBA

3.3.6 Moment Generating Function

1. TBA

3.3.7 Measurement Scales and Units of Measurement

1. TBA

3.3.7.1 The Four Measurement Scales

1. TBA

3.3.7.2 Units of Measurement

1. TBA

3.4 Binomial Experiments

1. TBA

3.5 Success Waiting Time: Geometric PDF

1. TBA

3.6 Waiting Time For r Successes: Negative Binomial

1. TBA

3.7 Poisson Process and Distribution

1. TBA

3.7.1 Moments of the Poisson PMF

1. TBA

3.7.2 Examples

1. An organic honey seller has found that she makes 8 sales on an average Saturday farmer's market over a 4-hour morning. On a slow Saturday, she sells less than 4. What is the probability of a slow Saturday if X is Poisson?

2. TBA

3.8 Waiting Time for Poisson Events: Negative Exponential PDF

1. An organic honey seller has found that she makes 8 sales on an average Saturday farmer's market over a 4-hour morning. She is thinking of going home early because she has had no sales since 8 a.m. and it now 11 a.m. Use the random variable T to measure the time to the first sale (of honey), assuming $\lambda = 8/4$ sales per hour. What is $P(T \geq 3.0\, hrs)$?

2. Suppose the lifetime of a light bulb follows a negative exponential distribution with a mean of 1000 hours. Find the probability the light bulb

   (a) lasts 1000 hours?

   (b) lasts 2000 hours?

   (c) burns out before 500 hours?

3. TBA

3.9 Normal or Gaussian Distribution

1. Suppose your laptop battery will run for X hours before recharging, where $X \sim N(7, 1/2^2)$. Find

(a) The probability it will run more than 9 hours?

(b) At least 8 hours?

(c) Between 6 and 7 hours?

2. TBA

3.9.1 Standard Normal Distribution

1. What is the probability a Gaussian random variable is *further than* 1, 2, 3, 4, 5 or 6 standard deviations of its mean? *(Hint: Use the R function* `pnorm(q,mu,sd)`*.)*

2. TBA

3.9.2 Sums of Independent Normal Random Variables

1. TBA

3.9.3 Normal Approximation to the Poisson Distribution

1. Plot the PMF of a Poisson PMF with $m = 36$, and overlay the approximating Gaussian PDF given by Eqn (3.44). Explain why these curve are close (without having to rescale the vertical scales) simply knowing that the former sums to 1 and latter integrates to 1. (Note the R function `dnorm(x,mu,sd)` computes the Gaussian/Normal PDF.)

2. TBA

3.10 Problems

1. Throw 3 dice and label the results X_1, X_2, and X_3. What is the PMF of the total number of pips, $X = X_1 + X_2 + X_3$?

2. Find the skewness and kurtosis coefficients in terms of μ and the noncentral moments μ_2', μ_3', and μ_4'.

3. If $X \sim NegBinom(r, p)$, find its MGF, and its mean and variance.

4. Use Mathematica to find the Taylor Series approximation to the difference of the exact MGF in Equation (3.52) and the MGF of a standard Normal. (We used a Taylor Series in the exponent when deriving Equation (3.53).

5. Use the MGF technique to prove the result in Equation (3.50). You should use the results from Section 5.4.1.

Chapter 4

Bivariate Random Variables, Transformations, and Simulations

4.1 Bivariate Continuous Random Variables

1. Let $f(x, y)$ be a uniform PDF over the region bounded by straight lines connecting the 5 points $(0, 0)$, $(4, 0)$, $(4, 1)$, $(2, 1)$, back to $(0, 0)$. *Remark: These problems cover all of this Section.*

   (a) What is the PDF $f(x, y)$?

   (b) Find the marginal pdf of X and its mean and variance?

   (c) Same for Y?

   (d) What is $m(x) = E[Y|X = x]$? Sketch it over the support of (X, Y).
   (Hint: This will be defined in a piecewise fashion.)

   (e) What is conditional variance of $m(x)$? Compare it to the unconditional variance of Y? Sketch $m(x)$ and add the 2 curves that are the conditional mean plus/minus one conditional standard deviation.

   (f) What is the covariance of X and Y? The correlation?

2. TBA

4.1.1 Joint CDF and PDF Functions

1. TBA

4.1.2 Marginal PDF

1. TBA

4.1.3 Conditional Probability Density Function

1. TBA

4.1.4 Independence of Two Random Variables

1. TBA

4.1.5 Expectation, Correlation, and Regression

1. TBA

4.1.5.1 Covariance and Correlation

1. Suppose X is a discrete random variable that takes on the values $-2, -1, 1, 2$ with equal probabilities $p_x = 1/4$. If $Y = X^2$, compute the correlation between X and Y. *Hint: Compute the means and variances of X and Y, and finally their covariance.*

2. Repeat the previous problem but with $Y = X^3$.

3. TBA

4.1.5.2 Regression Function

1. TBA

4.1.6 Independence of n Random Variables

1. TBA

4.1.7 Bivariate Normal PDF

1. TBA

4.1.8 Correlation, Independence, and Confounding Variables

1. TBA

4.2 Change of Variables

1. TBA

4.2.1 Examples: Two Uniform Transformations

1. TBA

4.2.2 One-Dimensional Transformations

1. Let X be a standard Gaussian r.v. and let $Y = exp(X)$ be the transformed r.v. The PDF of Y is called the **log-normal**. Using the change-of-variables technique, find the formula for $f_Y(y)$? Plot the PDF for $0 \leq y \leq 5$.

2. If X is distributed $Unif(0,1)$, what is the CDF of $Y = e^X$? Find the PDF and plot it.

3. If X is distributed as negative exponential with paramter $\lambda = 1$, what is the CDF of $Y = \log X$? Find the PDF and plot it.

4. TBA

4.2.2.1 Example 1: Exponential PDF

1. TBA

4.2.2.2 Example 2: Cauchy PDF

1. TBA

4.2.2.3 Example 3: Chi-Squared PDF With 1 Degree of Freedom

1. TBA

4.2.3 Two-Dimensional Transformations

1. TBA

4.3 Simulations

1. TBA

4.3.1 Generating Uniform Pseudo-Random Numbers

1. TBA

4.3.1.1 Reproducibility

1. TBA

4.3.1.2 RANDU

1. TBA

4.3.2 Probability Integral Transformation

1. We want to generate samples from the $Beta(1,3)$ PDF

$$f(x) = 3(1-x)^2 ,$$

where the support is $0 < x < 1$, using the Probability Integral Transformation.

(a) Show the CDF is $F(x) = 3x - 3x^2 + x^3$

(b) Find the closed form solution for x in $u = F(x)$, given a random number $u \sim U(0,1)$. (Note: Obtaining $x = F^{-1}(u)$.)

(c) Generate a sample of size $n = 10^4$ and plot a histogram and superimpose the PDF. How does it look to you? **Hints:**

```
> hist( 1-(1-runif(1e4))**(1/3), prob=50, T, col=2 )
> xs = seq(0,1,.01); lines( xs, 3*(1-xs)**2, col=4, lwd=3
)
```

2. TBA

4.3.3 Event-Driven Simulation

1. TBA

4.4 Problems

1. An alternative and informal derivation of Equation (4.5) may be illuminating. As we did in Figure (3.5), we write

$$\begin{aligned} P(X \approx x) &= P\Big(X \in \big(x - \tfrac{\Delta}{2}, x + \tfrac{\Delta}{2}\big)\Big) \\ &= P\Big(X \in \big(x - \tfrac{\Delta}{2}, x + \tfrac{\Delta}{2}\big), Y \in \big(-\infty, \infty\big)\Big) \\ &= \int_{y=\infty}^{\infty} \left[\int_{s=x-\Delta/2}^{x+\Delta/2} f_{X,Y}(s,y)\, ds \right] dy \\ &\approx \int_{y=\infty}^{\infty} \big[f_{X,Y}(x,y) \cdot \Delta \big]\, dy\,. \end{aligned}$$

Since $P(X \approx x) \approx \Delta \cdot f_X(x)$, verify this sequence of equations and show that Equation (4.5) follows. Draw the relevant bivariate event.

2. Show that the sequence of pseudo-random variables $\{(x_n, x_{n+1}, x_{n+2}),\ n = 1, 2, \ldots\}$ defined in Equation (4.32) satisfy the equation

$$x_{n+2} = 6\, x_{n+1} - 9\, x_n \qquad \bmod 2^{31}\,;$$

hence, the points (y_n, y_{n+1}, y_{n+2}), where $y_n = x_n/2^{31}$, fall on 15 planes in $\mathbb{R}^3$.

3. Use the bivariate change-of-variables approach to find the PDF of a bivariate normal, where the r.v.'s are **not** in standard form. Conclude the correlation coefficient is unchanged. *Hint:* You know $f_{X,Y}(x,y)$ where X and Y are in standard form. Define two new r.v.'s, $U = \mu_x + \sigma_x X$ and $V = \mu_y + \sigma_y Y$. Note that $U \sim N(\mu_X, \sigma_X^2)$ and $V \sim N(\mu_Y, \sigma_Y^2)$. Find $g_{U,V}(u,v)$, which is the general form of the bivariate Normal PDF with all 5 parameters.

Chapter 5

Approximations and Asymptotics

5.1 Why Do We Like Random Samples?

5.1.1 When $u(\boldsymbol{X})$ Takes a Product Form

5.1.2 When $u(\boldsymbol{X})$ Takes a Summation Form

5.2 Useful Inequalities

5.2.1 Markov's Inequality

5.2.2 Chebyshev's Inequality

1. Using Chebyschev's inequality to find n: Supposedly Asian Elephants are 200 pounds at birth. We think $\sigma = 5$ pounds. How large a sample n should we collect so that the probability $\bar{X}$ is within one (1) pound of μ is at least 99%?

2. TBA

5.2.3 Jensen's Inequality

5.2.4 Cauchy-Schwarz Inequality

5.3 Sequences of Random Variables

5.3.1 Weak Law of Large Numbers

5.3.2 Consistency of the Sample Variance

5.3.3 Relationships Among the Modes of Convergence

5.3.3.1 Proof of Result (5.21)

5.3.3.2 Proof of Result (5.22)

5.4 Central Limit Theorem

5.4.1 Moment Generating Function for Sums

1. The MGF technique and the CLT:

   (a) Show by the MGF technique that the random variable $Y = \sum_{i=1}^{n} X_i$ will be exactly $\chi^2(n)$ if the X_i are a random sample from the $\chi^2(1)$ PDF.

   (b) May we conclude the PDF $\chi^2(n)$ is approximately Gaussian?

2. **Analyzing the Average of a Normal Random Sample**

   - Use the moment generating function technique to find the exact PDF of the random variable $\bar{X}$ for a $N(\mu, \sigma^2)$ random sample.
   - First, use Mathematica to compute the MGF of a $N(\mu, \sigma^2)$ r.v.:

$$\begin{aligned} M_X(t) &= E[e^{tX}] \\ &= \exp\left(\mu\, t + \frac{1}{2}\sigma^2\, t^2\right) . \end{aligned}$$

- Next use the MGF approach to find the PDF of $\bar{X}$. *Hint:*

$$\begin{aligned}
M_{\bar{X}}(t) &= E[e^{t\bar{X}}] \\
&= E[e^{(t/n)(X_1+X_2+\cdots+X_n)}] \\
&= E\prod_{\ell=1}^{n} e^{(t/n)X_\ell} = \prod_{\ell=1}^{n} E[e^{(t/n)X_\ell}] = \prod_{\ell=1}^{n} M_{X_\ell}(t/n) \\
&= [M_X(t/n)]^n \ .
\end{aligned}$$

3. **Analyzing the Average of a Cauchy Random Sample**

- We wish to attempt to use the moment generating function technique to see if we can find the PDF of the random variable $\bar{X}$ for a Cauchy random sample. This cannot be done, as the Cauchy PDF, where

$$f_X(x) = \frac{1}{\pi(1+x^2)} \quad x \in \mathbb{R}$$

does not have a well-defined MGF. Instead, we need to use the Fourier (rather than LaPlace) transform, which is called the Characteristic Function (CF) in statistics.

- First, use Mathematica to compute the CF of a Cauchy r.v.:

$$\begin{aligned}
CF_X(t) &= E[e^{itX}] \qquad \text{where } i = \sqrt{-1} \\
&= \int_{-\infty}^{\infty} e^{itx} \, \frac{1}{\pi(1+x^2)} \, dx \\
&= e^{-|t|} \ .
\end{aligned}$$

- Next use the CF approach to find the PDF of $\bar{X}$. *Hint:*

$$\begin{aligned}
CF_{\bar{X}}(t) &= E[e^{it\bar{X}}] = E[e^{i(t/n)(X_1+X_2+\cdots+X_n)}] \\
&= E\prod_{\ell=1}^{n} e^{i(t/n)X_\ell} = \prod_{\ell=1}^{n} E[e^{i(t/n)X_\ell}] = \prod_{\ell=1}^{n} CF_{X_\ell}(t/n) \\
&= [CF_X(t/n)]^n \ .
\end{aligned}$$

- **Bonus:** Use the CF approach to show that if $U \sim Unif(-\pi/2, \pi.2)$, then $X = \tan(U)$ is Cauchy.

5.4.2 Standardizing the Sum S_n

1. An old way of generating (approximately) $N(0,1)$ samples was to use the formula
$$Y = \sum_{i=1}^{12} U_i - 6\,.$$
   (a) Why should this work if $U_i \sim U(0,1)$?
   Hints: Compute the moments of Y and invoke the CLT.
   (b) Try it 1000 times and plot a histogram of the pseudo-random sample. **Hint:**
```
> y=NULL; for(i in 1:1000) {y=c(y,sum(runif(12))-6)}; hist(y,40,T,col
```

2. The Central Limit Theorem says that for any random sample $\{X_i\}$,
$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu_X}{\sigma_X} \right) \xrightarrow{\mathscr{D}} N(0,1)\,.$$
If, in fact, the sample is itself Normal, i.e. $X_i \sim N(\mu_X, \sigma_X^2)$, use the MGF technique to show that Y_n is exactly $N(0,1)$ for all sample sizes n.

3. **Where we derive a general formula for the first two moments of a linear combination of independent, but not identically distributed, random variables.** Given n independent r.v.'s X_i with means μ_i and variances σ_i^2, consider the random variable
$$Y = \sum_{i=1}^{n} \big(a_i\, X_i + b_i\big)\,,$$
where a_i and b_i are known constants.
   (a) Find the mean of Y.
   (b) Find the variance of Y.

   Hint: Define $Y_i = X_i - \mu_i$ and show $Y - \mu_Y = \sum_i a_i\, Y_i$. Then easy to compute the variance of Y, since the Y_i's are independent with $EY_i = 0$ and $EY_i^2 = \sigma_i^2$.

4. **This problem is a generalization of the previous problem. Here, we derive the formula for the first two moments of a linear combination of dependent, and non-identically distributed, random variables.** The n r.v.'s X_i have means μ_i and variances σ_i^2, and the correlation coefficient between X_i and X_j is ρ_{ij}. Note that $cov(X_i, X_j) = \sigma_i \sigma_j \rho_{ij}$. Consider the random variable

$$Y = \sum_{i=1}^{n} \left(a_i\, X_i + b_i\right), \quad \text{where } a_i \text{ and } b_i \text{ are known constants.}$$

   (a) Find the mean of Y.

   (b) Find the variance of Y.

   (c) What is the variance of Y when $n = 2$?

   Hint: Let $Y_i = X_i - \mu_i$; hence, $Y - \mu_Y = \sum_i a_i\, Y_i$. Then it is easier to compute the $var(Y)$, since the Y_i's satisfy $EY_i = 0$, $EY_i^2 = \sigma_i^2$, and $cov(Y_i, Y_j) = E[Y_i Y_j] = \sigma_i \sigma_j \rho_{ij}$ as well.

5. Consider a random sample of n Bernoulli events, that is, $X_i \sim B(1, p)$.

   (a) Show $X_i^2 = X_i$. Then show

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 = \frac{n}{n-1} \bar{X}\left(1 - \bar{X}\right).$$

   (b) By computing the first two non-central moments of $\bar{X}$, show that

$$E\, S^2 = \sigma_X^2 = p(1-p)\,, \qquad \text{i.e., } S^2 \text{ is unbiased for } \sigma_X^2.$$

   Hint: First show that $E\bar{X} = \mu_{\bar{X}} = \mu_X = p$ and $E\bar{X}^2 = Var\,\bar{X} + (E\bar{X})^2$.

6. TBA

5.4.3 Proof of Central Limit Theorem

5.5 Delta Method and Variance-Stabilizing Transformations

5.6 Problems

1. Prove Equation (5.8) assuming $f(\boldsymbol{x}) = \prod_{i=1}^{n} f_{X_i}(x_i)$, i.e. $\{X_1, X_2, \ldots, X_n\}$ is an i.i.d. sample. Note: This is a stronger assumption than required.

2. Show Chebyshev's inequality can also be expressed as

$$P\left(|X - \mu_X| > c\right) \leq \frac{\sigma_X^2}{c^2} \, .$$

3. Show the delta method given in Equation (5.37) for a Poisson r.v. with $g(x) = \sqrt{x}$ gives the same answer as in Equation (5.36).

Chapter 6

Parameter Estimation

6.1 Desirable Properties of an Estimator

1. TBA

6.2 Moments of the Sample Mean and Variance

1. TBA

6.2.1 Theoretical Mean and Variance of the Sample Mean

1. TBA

6.2.2 Theoretical Mean of the Sample Variance

1. TBA

6.2.3 Theoretical Variance of the Sample Variance

1. TBA

6.3 Method of Moments (MoM)

1. TBA

6.4 Sufficient Statistics and Data Compression

1. For a random sample of size n from the Geometric PMF, $Geom(p)$:

   (a) Find the sufficient statistic, $T(\boldsymbol{X})$, for the parameter $\theta = p$.

   (b) Find a Method of Moments estimator for p.

2. Consider the $Gamma(\alpha, \beta)$ PDF, which may also be found in the Appendix. What are the two sufficient statistics, $T_1(\boldsymbol{X})$ and $T_2(\boldsymbol{X})$, for the parameters α and β?

3. TBA

6.5 Bayesian Parameter Estimation

1. TBA

6.6 Maximum Likelihood Parameter Estimation

1. For a random sample of size n from the Geometric PMF, $Geom(p)$:

   (a) Find the MLE $\hat{p}_{MLE}$.

   (b) Show that it truly is a maximum.

2. Consider a random sample of size n from the $Unif(\theta, 1)$ PDF, where $\theta < 1$.

   (a) What is the likelihood function $L(\theta \mid \boldsymbol{X})$?

   (b) What is the MLE for θ?

3. TBA

6.6.1 Relationship to Bayesian Parameter Estimation

1. TBA

6.6.2 Poisson MLE Example

1. TBA

6.6.3 Normal MLE Example

1. TBA

6.6.4 Uniform MLE Example

1. TBA

6.7 Information Inequalities and the Cramér-Rao Lower Bound

1. TBA

6.7.1 Score Function

1. TBA

6.7.2 Asymptotics of MLE

1. TBA

6.7.3 Minimum Variance of Unbiased Estimators

1. TBA

6.7.4 Examples

1. TBA

6.8 Problems

1. Given n independent r.v.'s X_i with means μ_i and variances σ_i^2, consider the random variable

$$Y = \sum_{i=1}^{n} \left(a_i \, X_i + b_i \right) ,$$

where a_i and b_i are known constants. Show

$$EY = \sum_{i=1}^{n} (a_i \, \mu_i + b_i)$$
$$\operatorname{var} Y = \sum_{i=1}^{n} a_i^2 \sigma_i^2 \, .$$

Hint: Define $Y_i = X_i - \mu_i$ and show $Y - \mu_Y = \sum a_i Y_i$.

2. For a Normal random sample, show that $\operatorname{var} S^2 = 2\sigma_X^4/(n-1)$ using the fact that $(n-1)S^2/\sigma_X^2 \sim \chi^2(n-1)$; see Equation (6.9) and following.

3. In the MoM example 3 in Section (6.3), show that $F(x) = 1 - \theta/x$, so that you can generate random samples as $x = \theta/(1-u)$ by the probability integral transformation method, where u is a pseudo-independent *Unif*$(0,1)$ sample. Choose $\theta = 3$. For a sample of size 10^3, plot a histogram of your sample. Does the default **R** histogram show much structure? Does the alternative MoM estimator of θ seem to work? Verify that $\hat{\theta}$ is a MoM estimator.

4. Show that if the sampling density $f(x|\boldsymbol{\theta})$ has m sufficient statistics, then the MLE $\hat{\boldsymbol{\theta}}$ must be a function of the data through the sufficient statistics alone.

Chapter 7

Hypothesis Testing

7.1 Setting up a Hypothesis Test

1. TBA

7.1.1 Example of a Critical Region

1. TBA

7.1.2 Accuracy and Errors in Hypothesis Testing

1. TBA

7.2 Best Critical Region for Simple Hypotheses

1. TBA

7.2.1 Simple Example Continued

1. TBA

7.2.2 Gaussian Shift Model with Common Variance

1. TBA

7.3 Best Critical Region for a Composite Alternative Hypothesis

1. Suppose we have a random sample of size n from a Poisson $Pois(m)$ PMF. We wish to test

$$\begin{aligned} H_0: &\quad m = m_0 \qquad\qquad \text{versus} \\ H_1: &\quad m = m_1 > m_0\,. \end{aligned}$$

   (a) Following Section 7.2, find the form of the best critical region.

   (b) The number of murders in Houston for the 4 years from 2014 to 2017 was 242, 297, 303, and 269. What is the best critical region for testing

$$\begin{aligned} H_0: &\quad m = 250 \qquad\qquad \text{versus} \\ H_1: &\quad m = 275 \end{aligned}$$

   at the 5% significance level?

   (c) What is the power of your test?

   (d) What is your decision for the data given?

2. Consider a hypothesis test of the mean of normal data with known variance of 1:

$$\begin{aligned} H_0: &\quad \mu = 0 \qquad\qquad \text{versus} \\ H_1: &\quad \mu \neq 0. \end{aligned}$$

   (a) If $n = 16$, what is the best critical region?

   (b) Perform the hypothesis test of these simulated samples. Generate 16 normal samples in R via

$$set.seed(123); \quad x = rnorm(16, 1, 1)\,.$$

   What is your decision?

3. TBA

7.3.1 Negative Exponential Composite Hypothesis Test

1. TBA

7.3.1.1 Example

1. TBA

7.3.1.2 Alternative Critical Regions

1. TBA

7.3.1.3 Mount St. Helens Example

1. TBA

7.3.2 Gaussian Shift Model with Common But Unknown Variance: The t-test

1. Consider a hypothesis test of the mean of normal data with unknown variance:

$$\begin{aligned} H_0 : \quad & \mu = 0 \qquad\qquad \text{versus} \\ H_1 : \quad & \mu \neq 0. \end{aligned}$$

   (a) If $n = 9$, what is the best critical region?

   (b) Perform the hypothesis test of these simulated samples. Generate 9 normal samples in R via

$$set.seed(456); \quad x = rnorm(9, 1, 2)\,.$$

   What is your decision?

2. TBA

7.3.3 The Random Variable T_{n-1}

1. TBA

7.3.3.1 Where We Show $\bar{X}$ and S^2 are Independent

1. TBA

7.3.3.2 Where We Show That S^2 Scaled is $\chi^2(n-1)$

1. TBA

7.3.3.3 Where We Finally Derive the T PDF

1. TBA

7.3.4 The One-Sample t-test

1. TBA

7.3.5 Example

1. TBA

7.3.6 Other t-tests

1. TBA

7.3.6.1 Paired t-test

1. TBA

7.3.6.2 Two-Sample t-test

1. TBA

7.3.6.3 Example Two-Sample t-test: Lord Rayleigh's Data

1. TBA

7.4 Reporting Results: p-Values and Power

1. TBA

7.4.1 Example When the Null Hypothesis Is Rejected

1. TBA

7.4.2 When the Null Hypothesis Is Not Rejected

1. TBA

7.4.3 The Power Function

1. TBA

7.5 Multiple Testing and the Bonferroni Correction

1. TBA

7.6 Problems

1. Show that the sum of n i.i.d. negative exponential r.v.'s with parameter λ has a $Gamma(n, \lambda)$ exactly. Hint: Use the MGF technique of Section 5.4.1.

2. Show that among the contiguous intervals (a, b) containing 95% of the probability for our example in Section 7.3.1.1, a necessary condition to minimize the width of the interval, $b - a$, is that the sampling density must be equal at a and b. This result holds generally if the sampling density is unimodal and monotone on either side of the mode.

3. Suppose the distribution of the r.v. used to find the various intervals (a, b) in Section 7.3.1.2 was not only unimodal and monotone on either side of the mode, but was also symmetric. Show that the equal-tail-area and narrowest-width intervals are identical.

4. Using the MGF technique, show that

$$Y = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)\,,$$

where Z_i is a random sample from the $N(0,1)$ PDF.

5. Using the MGF technique, show that if $U = S + T$, where (a) $U \sim \chi^2(n)$; (b) $T \sim \chi^2(1)$; and (c) S and T are independent, then $S \sim \chi^2(n-1)$.

6. Verify the critical regions and sample sizes for the example at the end of Section 7.4.3.

7. The 1969 military draft lottery numbers are shown in Figure 7.9. Individuals with numbers over 195 were not drafted. There appears to be a nonrandom downward trend. For example, only 5 birthdays in December had a lottery number above average. Run a two-sample T-test in **R** with the command `t.test(x[1:183],x[184:366])` and show the p-value is 5×10^{-5}. What do we conclude? Perform a simulation using data from `x = sample(366)` to verify the T-distribution is the appropriate test statistic.

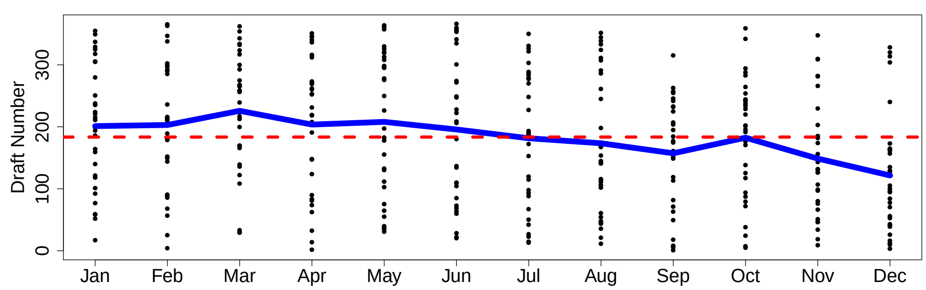

Figure 7.1: Draft lottery numbers 1–366 by month. The monthly average is the blue line. The overall average of 183.5 is the red dotted line.

Chapter 8

Confidence Intervals and Other Hypothesis Tests

8.1 Confidence Intervals

1. TBA

8.1.1 Confidence Interval for μ: Normal Data, σ^2 Known

1. TBA

8.1.2 Confidence Interval for μ: σ^2 Unknown

1. Suppose we have a random sample of size 50 from a $N(\mu, \sigma^2)$ PDF. We wish to test

$$\begin{aligned} H_0 : &\quad \mu = 10 \qquad\qquad \text{versus} \\ H_1 : &\quad \mu \neq 10\,. \end{aligned}$$

The sample moments are $\bar{x} = 13.4508$ and $s^2 = 65.8016$.

   (a) Find the critical region C and test the null hypothesis at the 5% level. What is your decision?

   (b) What is the p-value for your decision?

   (c) What is a 95% confidence interval for μ?

2. Use the same data as in the previous Problem.

   (a) Test the null hypothesis that $\sigma^2 = 64$ versus a two-sided alternative. First find the critical region and then give your decision.

   (b) Find a 95% confidence interval for σ^2?

   (c) If you are worried about performing 2 statistical tests on the same data and that the overall type I error might not be 5%, how might you modify your approach?

3. TBA

8.1.3 Confidence Intervals and p-Values

1. TBA

8.2 Hypotheses About the Variance and the F-Distribution

1. TBA

8.2.1 The F-Distribution

1. TBA

8.2.2 Hypotheses About the Value of σ^2

1. TBA

8.2.3 Confidence Interval for σ^2

1. TBA

8.2.4 Two-Sided Alternative for Testing $\sigma^2 = \sigma_0^2$

1. TBA

8.3 Pearson's Chi-Squared Tests

1. TBA

8.3.1 The Multinomial PMF

1. TBA

8.3.2 Goodness-of-Fit (GoF) Tests

1. TBA

8.3.3 Two-Category Binomial Case

1. TBA

8.3.4 m-Category Multinomial Case

1. TBA

8.3.5 Goodness-of-Fit Test for a Parametric Model

1. One hundred points were simulated using the R function `rexp(100,1)` for the negative exponential PDF with $\lambda = 1$. We are interested in testing a goodness-of-fit of the random numbers to the true model. A histogram of these data gave bin counts over the 9 bins

   $$\{(0, 0.5), (0.5, 1), \ldots, (4, 4.5)\} \qquad \text{of} \qquad (42,\ 16,\ 15,\ 11,\ 3,\ 5,\ 3,\ 1,\ 4)\,.$$

   Also, $\boxed{\bar{x} = 1.0865}$.

   (a) Perform a Pearson goodness-of-fit test at the 5% level. *Hint: Be careful that the expected counts all are greater than 5 and that they add up to $n = 100$.*

   (b) What is the p-value? How do your interpret this number?

   (c) If our null hypothesis was that the data came from a negative exponential PDF, but we did not know that the mean was 1, how would you perform the goodness-of-fit test? (Do it.)

2. TBA

8.3.6 Tests for Independence in Tables

1. **Pearson Test for Independence in Tables.** One hundred middle school students studying Spanish were randomly assigned to 2 classrooms. One classroom used Rosetta Stone and the other used Babbel software in the language lab and on their personal computers. At the end of one semester, the final grades turned out to be:

Grade	A	B	C	D	E	Totals
Classroom I	8	13	16	10	3	50
Classroom II	4	9	14	16	7	50

Table 8.1: Hypothetical Data Comparing Foreign Language Software

   (a) Test the null hypothesis that the software options are equally effective.

   (b) What is the p-value?

2. TBA

8.4 Correlation Coefficient Tests and C.I.'s

1. The correlation coefficient measures the degree to which the random variable Y is related to X in a linear fashion. If the relationship is in fact nonlinear, the intuition may be faculty. In, 1973, Professor Francis Anscombe devised a clever set of 4 datasets (each with $n = 11$) that illustrate this fact. These data are available in R as the object `anscombe`. Compute the correlation coefficient by the R expression `cor(x,y)` for each set of data and plot. Add the least squares regression line to each figure by using the R expression `abline( lsfit(x,y) )` . What do you conclude?

2. Consider the full 5-parameter bivariate normal PDF. Given a random sample $\{(x_i, y_i),\ i = 1, \ldots, n\}$, assume you know the MLE's for μ_x, μ_y,

σ_x, and σ_y are

$$\hat{\mu}_x = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \hat{\sigma}_x = s_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\hat{\mu}_y = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad \hat{\sigma}_y = s_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}\,.$$

Verify the formula for the MLE of the correlation coefficient ρ in Equation (8.23)

$$\hat{\rho} = r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad \text{or equivalently}$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{s_x\, s_y}$$

Hints: Now the pdf is

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\ \exp\left[\frac{-Q(x,y)}{2(1-\rho^2)}\right] \quad \text{where}$$

$$Q(x,y) = \left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\,.$$

Thus the average (i.e. divided by n) log-likelihood for ρ is given by

$$\ell(\rho|\boldsymbol{x},\boldsymbol{y}) = -\log 2\pi - \log\sigma_x - \log\sigma_y - \frac{1}{2}\log(1-\rho^2) - \frac{\bar{Q}(\boldsymbol{x},\boldsymbol{y})}{2(1-\rho^2)}$$

where

$$\bar{Q}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{n}\sum_{i=1}^{n} Q(x_i, y_i)\,.$$

Plugging in the 4 MLE's we know, we find

$$
\begin{aligned}
Q(x,y) &= \frac{(x-\bar{x})^2}{s_x^2} - 2\rho\frac{(x-\bar{x})(y-\bar{y})}{s_x\, s_y} + \frac{(y-\bar{y})^2}{s_y^2}\,; \quad \text{hence,} \\
\bar{Q}(\boldsymbol{x},\boldsymbol{y}) &= \frac{1}{n}\sum_{i=1}^{n} Q(x_i,y_i) \\
&= \frac{\frac{1}{n}\sum_i (x_i-\bar{x})^2}{s_x^2} - 2\rho\frac{\frac{1}{n}\sum_i (x_i-\bar{x})(y_i-\bar{y})}{s_x\, s_y} + \frac{\frac{1}{n}\sum_i (y_i-\bar{y})^2}{s_y^2} \\
&= \frac{s_x^2}{s_x^2} - 2\rho\frac{r s_x s_y}{s_x s_y} + \frac{s_y^2}{s_y^2} \\
&= 2(1-\rho\, r)\,.
\end{aligned}
$$

Thus the relevant terms in the average log-likelihood to be maximized are

$$\tilde{\ell}(\rho|\boldsymbol{x},\boldsymbol{y}) = -\frac{1}{2}\log(1-\rho^2) - \frac{2(1-\rho\, r)}{2(1-\rho^2)}$$

Use Mathematica to show

$$\frac{\partial\tilde{\ell}(\rho|\boldsymbol{x},\boldsymbol{y})}{\partial\rho} = \frac{r-\rho+r\rho^2-\rho^3}{(1-\rho^2)^2}$$

which equals 0 when $\hat{\rho} = r$ (also $\pm i$).

3. TBA

8.4.1 How to Test if the Correlation $\rho = 0$

1. TBA

8.4.2 Confidence Intervals and Tests for a General Correlation Coefficient

1. **Correlation Coefficient Test and Confidence Interval.** This case study examined the relationship between an isometric strength test, X, and job performance, Y, for 50 warehouse workers. You may assume

these data follow a bivariate normal PDF. The following statistics were calculated from these hypothetical data:

$$\sum_{i=1}^{50} x_i = 250.0109$$
$$\sum_{i=1}^{50} x_i^2 = 1335.2427$$
$$\sum_{i=1}^{50} y_i = 494.6882$$
$$\sum_{i=1}^{50} y_i^2 = 4934.5076$$
$$\sum_{i=1}^{50} x_i y_i = 2515.2075\,.$$

(a) What is the sample correlation coefficient?

(b) Test the null hypothesis that $H_0 : \rho = 0$ versus a two-side alternative. First find the critical region and then give your decision.

(c) Find a 95% confidence interval for ρ?

2. TBA

8.5 Linear Regression

1. TBA

8.5.1 Least Squares Regression

1. TBA

8.5.2 Distribution of the Least-Squares Parameters

1. TBA

8.5.3 A Confidence Interval for the Slope

1. TBA

8.5.4 A Two-Side Hypothesis Test for the Slope

1. TBA

8.5.5 Predictions at a New Value

1. TBA

8.5.6 Population Interval at a New Value

1. **Linear Regression Confidence Intervals.** A chemical engineering graduate student ran 11 experiments as follows. Given 500 grams of input chemicals by weight, a process was run at 11 temperatures $(50, 60, 70, \ldots, 140, 150)$. The output material was separated into product (y_i) and waste. The weight of the product was also measured in grams; see the Figure on the next page, on which both the data and the least-squares line are displayed.

   You may assume the ϵ_i are normally distributed. The following statistics were calculated from these hypothetical data:

$$\sum_{i=1}^{11} x_i = 1100$$
$$\sum_{i=1}^{11} x_i^2 = 121{,}000$$
$$\sum_{i=1}^{11} y_i = 1{,}844.812$$
$$\sum_{i=1}^{11} y_i^2 = 369{,}372.483$$
$$\sum_{i=1}^{11} x_i y_i = 205{,}743.937\,.$$

(a) Consider the pointwise confidence interval for $Y|x$ at the 95% level. Sketch the (upper and lower) curves onto the graph provided over the range of the data. (Compute these CI's at enough points to sketch the continuous lines. Never mind that the overall type I error may be wrong.)

(b) Next, consider the pointwise confidence interval for the predictions, also at the 95% level for each x. Sketch the (upper and lower) curves onto the graph provided over the range of the data. (Again, ignore the fact that the overall type I error will be wrong.)

2. TBA

8.6 Analysis of Variance

1. TBA

8.7 Problems

1. Find confidence intervals for paired and two-sample t-tests.

2. Show that the inequality in Equation (8.6) may be written in the form

$$Prob\left(\frac{\sqrt{n}\,\bar{X}}{\sqrt{\frac{(n-1)\,S^2}{n-1}}} < t_{0.975,n-1}\right) = 20\%\,,$$

where $\sqrt{n}\bar{X} \sim N(\sqrt{n}\mu_1, 1)$. The quantity follows the non-central T distribution, where the numerator is $N(\mu, 1)$ rather than $N(0, 1)$. For the parameters in the example,

```
> pt( qt(.975,15), 15, 2.997865 ) = 0.20000;
```

where the noncentrality parameter $\mu = 2.997865$ was found by interpolation. Hence, $4\mu_1 = 2.997865$ or $\mu_1 = 0.7495$.

Chapter 9

Topics in Statistics

9.1 MSE and Histogram Bin Width Selection

9.1.1 MSE Criterion for Biased Estimators

9.1.2 Case Study: Optimal Histogram Bin Widths

9.1.3 Examples with Gaussian Data

9.1.4 Normal Reference Rules for the Histogram Bin Width

9.1.4.1 Scott's Rule

9.1.4.2 Friedman-Diaconis Rule

9.1.4.3 Sturges' Rule

9.1.4.4 Comparison of the Three Rules

9.2 Optimal Stopping Time Problem

9.3 Compound Random Variables

9.3.1 Computing Expectations with Conditioning

9.3.2 Compound Random Variables

9.4 Simulation and the Bootstrap

9.5 Multiple Linear Regression

9.6 Experimental Design

9.7 Logistic Regression, Poisson Regression, and the Generalized Linear Model

9.8 Robustness

9.9 Conclusions

Subject Index