way. We have data $\mathbf{x}$ that we regard as being generated by a probability distribution $F(\mathbf{x}|\theta)$, which depends on a parameter θ. We wish to know $Eh(\mathbf{X}, \theta)$ for some function $h(\,)$. For example, if θ itself is estimated from the data as $\hat{\theta}(\mathbf{x})$ and $h(\mathbf{X}, \theta) = [\hat{\theta}(\mathbf{X}) - \theta]^2$, then $Eh(\mathbf{X}, \theta)$ is the mean square error of the estimate. As another example, if

$$h(\mathbf{X}, \theta) = \begin{cases} 1 \text{ if } |\hat{\theta}(\mathbf{X}) - \theta| > \Delta \\ 0 \text{ otherwise} \end{cases}$$

then $Eh(\mathbf{X}, \theta)$ is the probability that $|\hat{\theta}(\mathbf{X}) - \theta| > \Delta$. We realize that if θ were known, we could use the computer to generate independent random variables $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_B$ from $F(\mathbf{x}|\theta)$ and then appeal to the law of large numbers:

$$Eh(\mathbf{X}, \theta) \approx \frac{1}{B} \sum_{i=1}^{B} h(\mathbf{X}_i, \theta)$$

This approximation could be made arbitrarily precise by choosing B sufficiently large. The parametric bootstrap principle is to perform this Monte Carlo simulation using $\hat{\theta}$ in place of the unknown θ—that is, using $F(\mathbf{x}|\hat{\theta})$ to generate the $\mathbf{X}_i$. It is difficult to give a concise answer to the natural question: How much error is introduced by using $\hat{\theta}$ in place of θ? The answer depends on the continuity of $Eh(\mathbf{X}, \theta)$ as a function of θ—if small changes in θ can give rise to large changes in $Eh(\mathbf{X}, \theta)$, the parametric bootstrap will not work well.

8.10 Problems

1. The following table gives the observed counts in 1-second intervals for Berkson's data (Section 8.2). What are the expected counts from a Poisson distribution? Do they match the observed counts?

n	Observed
0	5267
1	4436
2	1800
3	534
4	111
5+	21

2. The Poisson distribution has been used by traffic engineers as a model for light traffic, based on the rationale that if the rate is approximately constant and the traffic is light (so the individual cars move independently of each other), the distribution of counts of cars in a given time interval or space area should be nearly Poisson (Gerlough and Schuhl 1955). The following table shows the number of right turns during 300 3-min intervals at a specific intersection. Fit a Poisson distribution. Comment on the fit by comparing observed and expected counts. It is useful to know that the 300 intervals were distributed over various hours of the day and various days of the week.

n	Frequency
0	14
1	30
2	36
3	68
4	43
5	43
6	30
7	14
8	10
9	6
10	4
11	1
12	1
13+	0

3. One of the earliest applications of the Poisson distribution was made by Student (1907) in studying errors made in counting yeast cells or blood corpuscles with a haemacytometer. In this study, yeast cells were killed and mixed with water and gelatin; the mixture was then spread on a glass and allowed to cool. Four different concentrations were used. Counts were made on 400 squares, and the data are summarized in the following table:

Number of Cells	Concentration 1	Concentration 2	Concentration 3	Concentration 4
0	213	103	75	0
1	128	143	103	20
2	37	98	121	43
3	18	42	54	53
4	3	8	30	86
5	1	4	13	70
6	0	2	2	54
7	0	0	1	37
8	0	0	0	18
9	0	0	1	10
10	0	0	0	5
11	0	0	0	2
12	0	0	0	2

a. Estimate the parameter λ for each of the four sets of data.
b. Find an approximate 95% confidence interval for each estimate.
c. Compare observed and expected counts.

4. Suppose that X is a discrete random variable with

$$P(X = 0) = \frac{2}{3}\theta$$

$$P(X = 1) = \frac{1}{3}\theta$$
$$P(X = 2) = \frac{2}{3}(1 - \theta)$$
$$P(X = 3) = \frac{1}{3}(1 - \theta)$$

where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations were taken from such a distribution: (3, 0, 2, 1, 3, 2, 1, 0, 2, 1).

a. Find the method of moments estimate of θ.
b. Find an approximate standard error for your estimate.
c. What is the maximum likelihood estimate of θ?
d. What is an approximate standard error of the maximum likelihood estimate?
e. If the prior distribution of Θ is uniform on [0, 1], what is the posterior density? Plot it. What is the mode of the posterior?

5. Suppose that X is a discrete random variable with $P(X = 1) = \theta$ and $P(X = 2) = 1 - \theta$. Three independent observations of X are made: $x_1 = 1, x_2 = 2, x_3 = 2$.

a. Find the method of moments estimate of θ.
b. What is the likelihood function?
c. What is the maximum likelihood estimate of θ?
d. If Θ has a prior distribution that is uniform on [0, 1], what is its posterior density?

6. Suppose that $X \sim \text{bin}(n, p)$.

a. Show that the mle of p is $\hat{p} = X/n$.
b. Show that mle of part (a) attains the Cramér-Rao lower bound.
c. If $n = 10$ and $X = 5$, plot the log likelihood function.

7. Suppose that X follows a geometric distribution,

$$P(X = k) = p(1 - p)^{k-1}$$

and assume an i.i.d. sample of size n.

a. Find the method of moments estimate of p.
b. Find the mle of p.
c. Find the asymptotic variance of the mle.
d. Let p have a uniform prior distribution on [0, 1]. What is the posterior distribution of p? What is the posterior mean?

8. In an ecological study of the feeding behavior of birds, the number of hops between flights was counted for several birds. For the following data, (a) fit a geometric distribution, (b) find an approximate 95% confidence interval for p, (c)

examine goodness of fit. (d) If a uniform prior is used for p, what is the posterior distribution and what are the posterior mean and standard deviation?

Number of Hops	Frequency
1	48
2	31
3	20
4	9
5	6
6	5
7	4
8	2
9	1
10	1
11	2
12	1

9. How would you respond to the following argument? This talk of sampling distributions is ridiculous! Consider Example A of Section 8.4. The experimenter found the mean number of fibers to be 24.9. How can this be a "random variable" with an associated "probability distribution" when it's just a number? The author of this book is guilty of deliberate mystification!

10. Use the normal approximation of the Poisson distribution to sketch the approximate sampling distribution of $\hat{\lambda}$ of Example A of Section 8.4. According to this approximation, what is $P(|\lambda_0 - \hat{\lambda}| > \delta)$ for $\delta = .5, 1, 1.5, 2$, and 2.5, where λ_0 denotes the true value of λ?

11. In Example A of Section 8.4, we used knowledge of the exact form of the sampling distribution of $\hat{\lambda}$ to estimate its standard error by

$$s_{\hat{\lambda}} = \sqrt{\frac{\hat{\lambda}}{n}}$$

This was arrived at by realizing that $\sum X_i$ follows a Poisson distribution with parameter $n\lambda_0$. Now suppose we hadn't realized this but had used the bootstrap, letting the computer do our work for us by generating B samples of size $n = 23$ of Poisson random variables with parameter $\lambda = 24.9$, forming the mle of λ from each sample, and then finally computing the standard deviation of the resulting collection of estimates and taking this as an estimate of the standard error of $\hat{\lambda}$. Argue that as $B \to \infty$, the standard error estimated in this way will tend to $s_{\hat{\lambda}}$.

12. Suppose that you had to choose either the method of moments estimates or the maximum likelihood estimates in Example C of Section 8.4 and C of Section 8.5. Which would you choose and why?

13. In Example D of Section 8.4, the method of moments estimate was found to be $\hat{\alpha} = 3\overline{X}$. In this problem, you will consider the sampling distribution of $\hat{\alpha}$.

a. Show that $E(\hat{\alpha}) = \alpha$—that is, that the estimate is unbiased.

b. Show that $\text{Var}(\hat{\alpha}) = (3 - \alpha^2)/n$. [*Hint:* What is $\text{Var}(\overline{X})$?]
c. Use the central limit theorem to deduce a normal approximation to the sampling distribution of $\hat{\alpha}$. According to this approximation, if $n = 25$ and $\alpha = 0$, what is the $P(|\hat{\alpha}| > .5)$?

14. In Example C of Section 8.5, how could you use the bootstrap to estimate the following measures of the accuracy of $\hat{\alpha}$: (a) $P(|\hat{\alpha} - \alpha_0| > .05)$, (b) $E(|\hat{\alpha} - \alpha_0|)$, (c) that number Δ such that $P(|\hat{\alpha} - \alpha_0| > \Delta) = .5$.

15. The upper quartile of a distribution with cumulative distribution F is that point $q_{.25}$ such that $F(q_{.25}) = .75$. For a gamma distribution, the upper quartile depends on α and λ, so denote it as $q(\alpha, \lambda)$. If a gamma distribution is fit to data as in Example C of Section 8.5 and the parameters α and λ are estimated by $\hat{\alpha}$ and $\hat{\lambda}$, the upper quartile could then be estimated by $\hat{q} = q(\hat{\alpha}, \hat{\lambda})$. Explain how to use the bootstrap to estimate the standard error of $\hat{q}$.

16. Consider an i.i.d. sample of random variables with density function

$$f(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right)$$

a. Find the method of moments estimate of σ.
b. Find the maximum likelihood estimate of σ.
c. Find the asymptotic variance of the mle.
d. Find a sufficient statistic for σ.

17. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables on the interval $[0, 1]$ with the density function

$$f(x|\alpha) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2}[x(1-x)]^{\alpha-1}$$

where $\alpha > 0$ is a parameter to be estimated from the sample. It can be shown that

$$E(X) = \frac{1}{2}$$
$$\text{Var}(X) = \frac{1}{4(2\alpha + 1)}$$

a. How does the shape of the density depend on α?
b. How can the method of moments be used to estimate α?
c. What equation does the mle of α satisfy?
d. What is the asymptotic variance of the mle?
e. Find a sufficient statistic for α.

18. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables on the interval $[0, 1]$ with the density function

$$f(x|\alpha) = \frac{\Gamma(3\alpha)}{\Gamma(\alpha)\Gamma(2\alpha)}x^{\alpha-1}(1-x)^{2\alpha-1}$$

where $\alpha > 0$ is a parameter to be estimated from the sample. It can be shown

that

$$E(X) = \frac{1}{3}$$
$$\text{Var}(X) = \frac{2}{9(3\alpha + 1)}$$

a. How could the method of moments be used to estimate α?
b. What equation does the mle of α satisfy?
c. What is the asymptotic variance of the mle?
d. Find a sufficient statistic for α.

19. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$.

a. If μ is known, what is the mle of σ?
b. If σ is known, what is the mle of μ?
c. In the case above (σ known), does any other unbiased estimate of μ have smaller variance?

20. Suppose that $X_1, X_2, \ldots, X_{25}$ are i.i.d. $N(\mu, \sigma^2)$, where $\mu = 0$ and $\sigma = 10$. Plot the sampling distributions of $\overline{X}$ and $\hat{\sigma}^2$.

21. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. with density function

$$f(x|\theta) = e^{-(x-\theta)}, \qquad x \geq \theta$$

and $f(x|\theta) = 0$ otherwise.

a. Find the method of moments estimate of θ.
b. Find the mle of θ. (*Hint:* Be careful, and don't differentiate before thinking. For what values of θ is the likelihood positive?)
c. Find a sufficient statistic for θ.

22. The Weibull distribution was defined in Problem 67 of Chapter 2. This distribution is sometimes fit to lifetimes. Describe how to fit this distribution to data and how to find approximate standard errors of the parameter estimates.

23. A company has manufactured certain objects and has printed a serial number on each manufactured object. The serial numbers start at 1 and end at N, where N is the number of objects that have been manufactured. One of these objects is selected at random, and the serial number of that object is 888. What is the method of moments estimate of N? What is the mle of N?

24. Find a very new shiny penny. Hold it on its edge and spin it. Do this 20 times and count the number of times it comes to rest heads up. Letting π denote the probability of a head, graph the log likelihood of π. Next, repeat the experiment in a slightly different way: This time spin the coin until 10 heads come up. Again, graph the log likelihood of π.

25. If a thumbtack is tossed in the air, it can come to rest on the ground with either the point up or the point touching the ground. Find a thumbtack. Before doing any experiment, what do you think π, the probability of it landing point up, is? Next, toss the thumbtack 20 times and graph the log likelihood of π. Then do

another experiment: Toss the thumbtack until it lands point up 5 times, and graph the log likelihood of π based on this experiment.

Find and graph the posterior distribution arising from a uniform prior on π. Find the posterior mean and standard deviation and compare the posterior with a normal distribution with that mean and standard deviation. Finally, toss the thumbtack 20 more times and compare the posterior distribution based on all 40 tosses to that based on the first 20.

26. In an effort to determine the size of an animal population, 100 animals are captured and tagged. Some time later, another 50 animals are captured, and it is found that 20 of them are tagged. How would you estimate the population size? What assumptions about the capture/recapture process do you need to make? (See Example I of Section 1.4.2.)

27. Suppose that certain electronic components have lifetimes that are exponentially distributed: $f(t|\tau) = (1/\tau)\exp(-t/\tau)$, $t \geq 0$. Five new components are put on test, the first one fails at 100 days, and no further observations are recorded.

a. What is the likelihood function of τ?
b. What is the mle of τ?
c. What is the sampling distribution of the mle?
d. What is the standard error of the mle?

(*Hint:* See Example A of Section 3.7.)

28. Why do the intervals in the left panel of Figure 8.8 have different centers? Why do they have different lengths?

29. Are the estimates of σ^2 at the centers of the confidence intervals shown in the right panel of Figure 8.8? Why are some intervals so short and others so long? For which of the samples that produced these confidence intervals was $\hat{\sigma}^2$ smallest?

30. The exponential distribution is $f(x;\lambda) = \lambda e^{-\lambda x}$ and $E(X) = \lambda^{-1}$. The cumulative distribution function is $F(x) = P(X \leq x) = 1 - e^{-\lambda x}$. Three observations are made by an instrument that reports $x_1 = 5$ and $x_2 = 3$, but x_3 is too large for the instrument to measure and it reports only that $x_3 > 10$. (The largest value the instrument can measure is 10.0.)

a. What is the likelihood function?
b. What is the mle of λ?

31. George spins a coin three times and observes no heads. He then gives the coin to Hilary. She spins it until the first head occurs, and ends up spinning it four times total. Let θ denote the probability the coin comes up heads.

a. What is the likelihood of θ?
b. What is the MLE of θ?

32. The following 16 numbers came from normal random number generator on a computer:

5.3299	4.2537	3.1502	3.7032	1.6070	6.3923	3.1181
6.5941	3.5281	4.7433	0.1077	1.5977	5.4920	1.7220
4.1547	2.2799					

a. What would you guess the mean and variance (μ and σ^2) of the generating normal distribution were?
b. Give 90%, 95%, and 99% confidence intervals for μ and σ^2.
c. Give 90%, 95%, and 99% confidence intervals for σ.
d. How much larger a sample do you think you would need to halve the length of the confidence interval for μ?

33. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$, where μ and σ are unknown. How should the constant c be chosen so that the interval $(-\infty, \overline{X} + c)$ is a 95% confidence interval for μ; that is, c should be chosen so that $P(-\infty < \mu \leq \overline{X} + c) = .95$.

34. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\mu_0, \sigma_0^2)$ and μ and σ^2 are estimated by the method of maximum likelihood, with resulting estimates $\hat{\mu}$ and $\hat{\sigma}^2$. Suppose the bootstrap is used to estimate the sampling distribution of $\hat{\mu}$.

a. Explain why the bootstrap estimate of the distribution of $\hat{\mu}$ is $N(\hat{\mu}, \frac{\hat{\sigma}^2}{n})$.
b. Explain why the bootstrap estimate of the distribution of $\hat{\mu} - \mu_0$ is $N(0, \frac{\hat{\sigma}^2}{n})$.
c. According to the result of the previous part, what is the form of the bootstrap confidence interval for μ, and how does it compare to the exact confidence interval based on the t distribution?

35. (Bootstrap in Example A of Section 8.5.1) Let $U_1, U_2, \ldots, U_{1029}$ be independent uniformly distributed random variables. Let X_1 equal the number of U_i less than .331, X_2 equal the number between .331 and .820, and X_3 equal the number greater than .820. Why is the joint distribution of X_1, X_2, and X_3 multinomial with probabilities .331, .489, and .180 and $n = 1029$?

36. How do the approximate 90% confidence intervals in Example E of Section 8.5.3 compare to those that would be obtained approximating the sampling distributions of $\hat{\alpha}$ and $\hat{\lambda}$ by normal distributions with standard deviations given by $s_{\hat{\alpha}}$ and $s_{\hat{\lambda}}$ as in Example C of Section 8.5?

37. Using the notation of Section 8.5.3, suppose that $\underline{\theta}$ and $\overline{\theta}$ are lower and upper quantiles of the distribution of θ^*. Show that the bootstrap confidence interval for θ can be written as $(2\hat{\theta} - \overline{\theta}, 2\hat{\theta} - \underline{\theta})$.

38. Continuing Problem 37, show that if the sampling distribution of θ^* is symmetric about $\hat{\theta}$, then the bootstrap confidence interval is $(\underline{\theta}, \overline{\theta})$.

39. In Section 8.5.3, the bootstrap confidence interval was derived from consideration of the sampling distribution of $\hat{\theta} - \theta_0$. Suppose that we had started with considering the distribution of $\hat{\theta}/\theta$. How would the argument have proceeded, and would the bootstrap interval that was finally arrived at have been different?

40. In Example A of Section 8.5.1, how could you use the bootstrap to estimate the following measures of the accuracy of $\hat{\theta}$: (a) $P(|\hat{\theta} - \theta_0| > .01)$, (b) $E(|\hat{\theta} - \theta_0|)$, (c) that number Δ such that $P(|\hat{\theta} - \theta_0| > \Delta) = .5$?

41. What are the relative efficiencies of the method of moments and maximum likelihood estimates of α and λ in Example C of Section 8.4 and Example C of Section 8.5?

42. The file `gamma-ray` contains a small quantity of data collected from the Compton Gamma Ray Observatory, a satellite launched by NASA in 1991 (`http://cossc.gsfc.nasa.gov/`). For each of 100 sequential time intervals of variable lengths (given in seconds), the number of gamma rays originating in a particular area of the sky was recorded. Assuming a model that the arrival times are a Poisson process with constant emission rate (λ = events per second), estimate λ. What is the estimated standard error? How might you informally check the assumption that the emission rate is constant? What is the posterior distribution of Λ if an improper gamma prior is used?

43. The file `gamma-arrivals` contains another set of gamma-ray data, this one consisting of the times between arrivals (interarrival times) of 3,935 photons (units are seconds).

- **a.** Make a histogram of the interarrival times. Does it appear that a gamma distribution would be a plausible model?
- **b.** Fit the parameters by the method of moments and by maximum likelihood. How do the estimates compare?
- **c.** Plot the two fitted gamma densities on top of the histogram. Do the fits look reasonable?
- **d.** For both maximum likelihood and the method of moments, use the bootstrap to estimate the standard errors of the parameter estimates. How do the estimated standard errors of the two methods compare?
- **e.** For both maximum likelihood and the method of moments, use the bootstrap to form approximate confidence intervals for the parameters. How do the confidence intervals for the two methods compare?
- **f.** Is the interarrival time distribution consistent with a Poisson process model for the arrival times?

44. The file `bodytemp` contains normal body temperature readings (degrees Fahrenheit) and heart rates (beats per minute) of 65 males (coded by 1) and 65 females (coded by 2) from Shoemaker (1996). Assuming that the population distributions are normal (an assumption that will be investigated in a later chapter), estimate the means and standard deviations of the males and females. Form 95% confidence intervals for the means. Standard folklore is that the average body temperature is 98.6 degrees Fahrenheit. Does this appear to be the case?

45. A Random Walk Model for Chromatin. A human chromosome is a very large molecule, about 2 or 3 centimeters long, containing 100 million base pairs (Mbp). The cell nucleus, where the chromosome is contained, is in contrast only about a thousandth of a centimeter in diameter. The chromosome is packed in a series of coils, called *chromatin*, in association with special proteins (histones), forming a string of microscopic beads. It is a mixture of DNA and protein. In the G0/G1 phase of the cell cycle, between mitosis and the onset of DNA replication, the mitotic chromosomes diffuse into the interphase nucleus. At this stage, a number of important processes related to chromosome function take place. For example, DNA is made accessible for transcription and is duplicated, and repairs are made of DNA strand breaks. By the time of the next mitosis, the chromosomes have been duplicated. The complexity of these and other processes raises many

questions about the large-scale spatial organization of chromosomes and how this organization relates to cell function. Fundamentally, it is puzzling how these processes can unfold in such a spatially restricted environment.

At a scale of about 10^{-3} Mbp, the DNA forms a chromatin fiber about 30 nm in diameter; at a scale of about 10^{-1} Mbp the chromatin may form loops. Very little is known about the spatial organization beyond this scale. Various models have been proposed, ranging from highly random to highly organized, including irregularly folded fibers, giant loops, radial loop structures, systematic organization to make the chromatin readily accessible to transcription and replication machinery, and stochastic configurations based on random walk models for polymers.

A series of experiments (Sachs et al., 1995; Yokota et al., 1995) were conducted to learn more about spatial organization on larger scales. Pairs of small DNA sequences (size about 40 kbp) at specified locations on human chromosome 4 were flourescently labeled in a large number of cells. The distances between the members of these pairs were then determined by flourescence microscopy. (The distances measured were actually two-dimensional distances between the projections of the paired locations onto a plane.) The empirical distribution of these distances provides information about the nature of large-scale organization.

There has long been a tradition in chemistry of modeling the configurations of polymers by the theory of random walks. As a consequence of such a model, the two-dimensional distance should follow a Rayleigh distribution

$$f(r|\theta) = \frac{r}{\theta^2} \exp\left(\frac{-r^2}{2\theta^2}\right)$$

Basically, the reason for this is as follows: The random walk model implies that the joint distribution of the locations of the pair in R^3 is multivariate Gaussian; by properties of the multivariate Gaussian, it can be shown the joint distribution of the locations of the projections onto a plane is bivariate Gaussian. As in Example A of Section 3.6.2 of the text, it can be shown that the distance between the points follows a Rayleigh distribution.

In this exercise, you will fit the Rayleigh distribution to some of the experimental results and examine the goodness of fit. The entire data set comprises 36 experiments in which the separation between the pairs of flourescently tagged locations ranged from 10 Mbp to 192 Mbp. In each such experimental condition, about 100–200 measurements of two-dimensional distances were determined. This exercise will be concerned just with the data from three experiments (short, medium, and long separation). The measurements from these experiments is contained in the files `Chromatin/short`, `Chromatin/medium`, `Chromatin/long`.

a. What is the maximum likelihood estimate of θ for a sample from a Rayleigh distribution?

b. What is the method of moments estimate?

c. What are the approximate variances of the mle and the method of moments estimate?

d. For each of the three experiments, plot the likelihood functions and find the mle's and their approximate variances.

e. Find the method of moments estimates and the approximate variances.

f. For each experiment, make a histogram (with unit area) of the measurements and plot the fitted densities on top. Do the fits look reasonable? Is there any appreciable difference between the maximum likelihood fits and the method of moments fits?

g. Does there appear to be any relationship between your estimates and the genomic separation of the points?

h. For one of the experiments, compare the asymptotic variances to the results obtained from a parametric bootstrap. In order to do this, you will have to generate random variables from a Rayleigh distribution with parameter θ.

Show that if X follows a Rayleigh distribution with $\theta = 1$, then $Y = \theta X$ follows a Rayleigh distribution with parameter θ. Thus it is sufficient to figure out how to generate random variables that are Rayleigh, $\theta = 1$. Show how Proposition D of Section 2.3 of the text can be applied to accomplish this.

$B = 100$ bootstrap samples should suffice for this problem. Make a histogram of the values of the θ^*. Does the distribution appear roughly normal? Do you think that the large sample theory can be reasonably applied here? Compare the standard deviation calculated from the bootstrap to the standard errors you found previously.

i. For one of the experiments, use the bootstrap to construct an approximate 95% confidence interval for θ using $B = 1000$ bootstrap samples. Compare this interval to that obtained using large sample theory.

46. The data of this exercise were gathered as part of a study to estimate the population size of the bowhead whale (Raftery and Zeh 1993). The statistical procedures for estimating the population size along with an assessment of the variability of the estimate were quite involved, and this problem deals with only one aspect of the problem—a study of the distribution of whale swimming speeds. Pairs of sightings and corresponding locations that could be reliably attributed to the same whale were collected, thus providing an estimate of velocity for each whale. The velocities, $v_1, v_2, \ldots, v_{210}$ (km/h), were converted into times $t_1, t_2, \ldots, t_{210}$ to swim 1 km—$t_i = 1/v_i$. The distribution of the t_i was then fit by a gamma distribution. The times are contained in the file `whales`.

a. Make a histogram of the 210 values of t_i. Does it appear that a gamma distribution would be a plausible model to fit?

b. Fit the parameters of the gamma distribution by the method of moments.

c. Fit the parameters of the gamma distribution by maximum likelihood. How do these values compare to those found before?

d. Plot the two gamma densities on top of the histogram. Do the fits look reasonable?

e. Estimate the sampling distributions and the standard errors of the parameters fit by the method of moments by using the bootstrap.

f. Estimate the sampling distributions and the standard errors of the parameters fit by maximum likelihood by using the bootstrap. How do they compare to the results found previously?

g. Find approximate confidence intervals for the parameters estimated by maximum likelihood.

47. The Pareto distribution has been used in economics as a model for a density function with a slowly decaying tail:

$$f(x|x_0, \theta) = \theta x_0^{\theta} x^{-\theta-1}, \qquad x \geq x_0, \quad \theta > 1$$

Assume that $x_0 > 0$ is given and that $X_1, X_2, \ldots, X_n$ is an i.i.d. sample.

a. Find the method of moments estimate of θ.
b. Find the mle of θ.
c. Find the asymptotic variance of the mle.
d. Find a sufficient statistic for θ.

48. Consider the following method of estimating λ for a Poisson distribution. Observe that

$$p_0 = P(X = 0) = e^{-\lambda}$$

Letting Y denote the number of zeros from an i.i.d. sample of size n, λ might be estimated by

$$\tilde{\lambda} = -\log\left(\frac{Y}{n}\right)$$

Use the method of propagation of error to obtain approximate expressions for the variance and the bias of this estimate. Compare the variance of this estimate to the variance of the mle, computing relative efficiencies for various values of λ. Note that $Y \sim \text{bin}(n, p_0)$.

49. For the example on muon decay in Section 8.4, suppose that instead of recording $x = \cos\theta$, only whether the electron goes backward ($x < 0$) or forward ($x > 0$) is recorded.

a. How could α be estimated from n independent observations of this type? (*Hint:* Use the binomial distribution.)
b. What is the variance of this estimate and its efficiency relative to the method of moments estimate and the mle for $\alpha = 0, .1, .2, .3, .4, .5, .6, .7, .8, .9$?

50. Let $X_1, \ldots, X_n$ be an i.i.d. sample from a Rayleigh distribution with parameter $\theta > 0$:

$$f(x|\theta) = \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}, \qquad x \geq 0$$

(This is an alternative parametrization of that of Example A in Section 3.6.2.)

a. Find the method of moments estimate of θ.
b. Find the mle of θ.
c. Find the asymptotic variance of the mle.

51. The double exponential distribution is

$$f(x|\theta) = \frac{1}{2} e^{-|x-\theta|}, \qquad -\infty < x < \infty$$

For an i.i.d. sample of size $n = 2m + 1$, show that the mle of θ is the median of the sample. (The observation such that half of the rest of the observations are

smaller and half are larger.) [*Hint:* The function $g(x) = |x|$ is not differentiable. Draw a picture for a small value of n to try to understand what is going on.]

52. Let $X_1, \ldots, X_n$ be i.i.d. random variables with the density function

$$f(x|\theta) = (\theta + 1)x^{\theta}, \qquad 0 \le x \le 1$$

a. Find the method of moments estimate of θ.
b. Find the mle of θ.
c. Find the asymptotic variance of the mle.
d. Find a sufficient statistic for θ.

53. Let $X_1, \ldots, X_n$ be i.i.d. uniform on $[0, \theta]$.

a. Find the method of moments estimate of θ and its mean and variance.
b. Find the mle of θ.
c. Find the probability density of the mle, and calculate its mean and variance. Compare the variance, the bias, and the mean squared error to those of the method of moments estimate.
d. Find a modification of the mle that renders it unbiased.

54. Suppose that an i.i.d. sample of size 15 from a normal distribution gives $\overline{X} = 10$ and $s^2 = 25$. Find 90% confidence intervals for μ and σ^2.

55. For two factors—starchy or sugary, and green base leaf or white base leaf—the following counts for the progeny of self-fertilized heterozygotes were observed (Fisher 1958):

Type	Count
Starchy green	1997
Starchy white	906
Sugary green	904
Sugary white	32

According to genetic theory, the cell probabilities are $.25(2 + \theta)$, $.25(1 - \theta)$, $.25(1 - \theta)$, and $.25\theta$, where $\theta (0 < \theta < 1)$ is a parameter related to the linkage of the factors.

a. Find the mle of θ and its asymptotic variance.
b. Form an approximate 95% confidence interval for θ based on part (a).
c. Use the bootstrap to find the approximate standard deviation of the mle and compare to the result of part (a).
d. Use the bootstrap to find an approximate 95% confidence interval and compare to part (b).

56. Referring to Problem 55, consider two other estimates of θ. (1) The expected number of counts in the first cell is $n(2 + \theta)/4$; if this expected number is equated to the count X_1, the following estimate is obtained:

$$\tilde{\theta}_1 = \frac{4X_1}{n} - 2$$

(2) The same procedure done for the last cell yields

$$\tilde{\theta}_2 = \frac{4X_4}{n}$$

Compute these estimates. Using that X_1 and X_4 are binomial random variables, show that these estimates are unbiased, and obtain expressions for their variances. Evaluate the estimated standard errors and compare them to the estimated standard error of the mle.

57. This problem is concerned with the estimation of the variance of a normal distribution with unknown mean from a sample $X_1, \ldots, X_n$ of i.i.d. normal random variables. In answering the following questions, use the fact that (from Theorem B of Section 6.3)

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

and that the mean and variance of a chi-square random variable with r df are r and $2r$, respectively.

a. Which of the following estimates is unbiased?

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

b. Which of the estimates given in part (a) has the smaller MSE?

c. For what value of ρ does $\rho\sum_{i=1}^{n}(X_i - \overline{X})^2$ have the minimal MSE?

58. If gene frequencies are in equilibrium, the genotypes AA, Aa, and aa occur with probabilities $(1-\theta)^2$, $2\theta(1-\theta)$, and θ^2, respectively. Plato et al. (1964) published the following data on haptoglobin type in a sample of 190 people:

Haptoglobin Type		
Hp1-1	Hp1-2	Hp2-2
10	68	112

a. Find the mle of θ.

b. Find the asymptotic variance of the mle.

c. Find an approximate 99% confidence interval for θ.

d. Use the bootstrap to find the approximate standard deviation of the mle and compare to the result of part (b).

e. Use the bootstrap to find an approximate 99% confidence interval and compare to part (c).

59. Suppose that in the population of twins, males (M) and females (F) are equally likely to occur and that the probability that twins are identical is α. If twins are not identical, their genes are independent.

a. Show that

$$P(MM) = P(FF) = \frac{1+\alpha}{4} \qquad P(MF) = \frac{1-\alpha}{2}$$

b. Suppose that n twins are sampled. It is found that n_1 are MM, n_2 are FF, and n_3 are MF, but it is not known which twins are identical. Find the mle of α and its variance.

60. Let $X_1, \ldots, X_n$ be an i.i.d. sample from an exponential distribution with the density function

$$f(x|\tau) = \frac{1}{\tau} e^{-x/\tau}, \qquad 0 \le x < \infty$$

a. Find the mle of τ.
b. What is the exact sampling distribution of the mle?
c. Use the central limit theorem to find a normal approximation to the sampling distribution.
d. Show that the mle is unbiased, and find its exact variance. (*Hint:* The sum of the X_i follows a gamma distribution.)
e. Is there any other unbiased estimate with smaller variance?
f. Find the form of an approximate confidence interval for τ.
g. Find the form of an exact confidence interval for τ.

61. Laplace's rule of succession. Laplace claimed that when an event happens n times in a row and never fails to happen, the probability that the event will occur the next time is $(n + 1)/(n + 2)$. Can you suggest a rationale for this claim?

62. Show that the gamma distribution is a conjugate prior for the exponential distribution. Suppose that the waiting time in a queue is modeled as an exponential random variable with unknown parameter λ, and that the average time to serve a random sample of 20 customers is 5.1 minutes. A gamma distribution is used as a prior. Consider two cases: (1) the mean of the gamma is 0.5 and the standard deviation is 1, and (2) the mean is 10 and the standard deviation is 20. Plot the two posterior distributions and compare them. Find the two posterior means and compare them. Explain the differences.

63. Suppose that 100 items are sampled from a manufacturing process and 3 are found to be defective. A beta prior is used for the unknown proportion θ of defective items. Consider two cases: (1) $a = b = 1$, and (2) $a = 0.5$ and $b = 5$. Plot the two posterior distributions and compare them. Find the two posterior means and compare them. Explain the differences.

64. This is a continuation of the previous problem. Let $X = 0$ or 1 according to whether an item is defective. For each choice of the prior, what is the marginal distribution of X before the sample is taken? What are the marginal distributions after the sample is taken? (*Hint*: for the second question, use the posterior distribution of θ.)

65. Suppose that a random sample of size 20 is taken from a normal distribution with unknown mean and known variance equal to 1, and the mean is found to be $\bar{x} = 10$. A normal distribution was used as the prior for the mean, and it was found that the posterior mean was 15 and the posterior standard deviation was 0.1. What were the mean and standard deviation of the prior?

66. Let the unknown probability that a basketball player makes a shot successfully be θ. Suppose your prior on θ is uniform on [0, 1] and that she then makes two shots in a row. Assume that the outcomes of the two shots are independent.

a. What is the posterior density of θ?

b. What would you estimate the probability that she makes a third shot to be?

67. Evans (1953) considered fitting the negative binomial distribution and other distributions to a number of data sets that arose in ecological studies. Two of these sets will be used in this problem. The first data set gives frequency counts of *Glaux maritima* made in 500 contiguous 20-cm^2 quadrants. For the second data set, a plot of potato plants 48 rows wide and 96 ft long was examined. The area was split into 2304 sampling units consisting of 2-ft lengths of row and in each unit the number of potato beetles was counted. Fit Poisson and negative binomial distributions, and comment on the goodness of fit. For these data, the method of moments should be fairly efficient.

Count	*Glaux maritima*	Potato Beetles
0	1	190
1	15	264
2	27	304
3	42	260
4	77	294
5	77	219
6	89	183
7	57	150
8	48	104
9	24	90
10	14	60
11	16	46
12	9	29
13	3	36
14	1	19
15		12
16		11
17		6
18		10
19		2
20		4
21		1
22		3
23		4
24		1
25		1
26		0
27		0
28		1

68. Let $X_1, \ldots, X_n$ be an i.i.d. sample from a Poisson distribution with mean λ, and let $T = \sum_{i=1}^{n} X_i$.

a. Show that the distribution of $X_1, \ldots, X_n$ given T is independent of λ, and conclude that T is sufficient for λ.

b. Show that X_1 is not sufficient.

c. Use Theorem A of Section 8.8.1 to show that T is sufficient. Identify the functions g and h of that theorem.

69. Use the factorization theorem (Theorem A in Section 8.8.1) to conclude that $T = \sum_{i=1}^{n} X_i$ is a sufficient statistic when the X_i are an i.i.d. sample from a geometric distribution.

70. Use the factorization theorem to find a sufficient statistic for the exponential distribution.

71. Let $X_1, \ldots, X_n$ be an i.i.d. sample from a distribution with the density function

$$f(x|\theta) = \frac{\theta}{(1+x)^{\theta+1}}, \qquad 0 < \theta < \infty \text{ and } 0 \le x < \infty$$

Find a sufficient statistic for θ.

72. Show that $\prod_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} X_i$ are sufficient statistics for the gamma distribution.

73. Find a sufficient statistic for the Rayleigh density,

$$f(x|\theta) = \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}, \qquad x \ge 0$$

74. Show that the binomial distribution belongs to the exponential family.

75. Show that the gamma distribution belongs to the exponential family.