11.2.3. Suppose for the situation of Example 11.2.2, θ_1 has the prior distribution $N(75, 1/(5\theta_3))$ and θ_3 has the prior distribution $\Gamma(\alpha = 4, \beta = 0.5)$. Suppose the observed sample of size $n = 50$ resulted in $\overline{x} = 77.02$ and $s^2 = 8.2$.

(a) Find the Bayes point estimate of the mean θ_1.

(b) Determine an HDR interval estimate with $1 - \gamma = 0.90$.

11.2.4. Let $f(x|\theta)$, $\theta \in \Omega$, be a pdf with Fisher information, (6.2.4), $I(\theta)$. Consider the Bayes model

$$\begin{aligned} X|\theta &\sim f(x|\theta)\ ,\ \ \theta \in \Omega \\ \Theta &\sim h(\theta) \propto \sqrt{I(\theta)}. \end{aligned} \tag{11.2.2}$$

(a) Suppose we are interested in a parameter $\tau = u(\theta)$. Use the chain rule to prove that

$$\sqrt{I(\tau)} = \sqrt{I(\theta)}\left|\frac{\partial\theta}{\partial\tau}\right|. \tag{11.2.3}$$

(b) Show that for the Bayes model (11.2.2), the prior pdf for τ is proportional to $\sqrt{I(\tau)}$.

The class of priors given by expression (11.2.2) is often called the class of **Jeffreys' priors**; see Jeffreys (1961). This exercise shows that Jeffreys' priors exhibit an invariance in that the prior of a parameter τ, which is a function of θ, is also proportional to the square root of the information for τ.

11.2.5. Consider the Bayes model

$$\begin{aligned} X_i|\theta\ , i = 1, 2, \ldots, n &\sim \text{iid with distribution } \Gamma(1, \theta),\ \theta > 0 \\ \Theta &\sim h(\theta) \propto \frac{1}{\theta}. \end{aligned}$$

(a) Show that $h(\theta)$ is in the class of Jeffreys' priors.

(b) Show that the posterior pdf is

$$h(\theta|y) \propto \left(\frac{1}{\theta}\right)^{n+2-1} e^{-y/\theta},$$

where $y = \sum_{i=1}^n x_i$.

(c) Show that if $\tau = \theta^{-1}$, then the posterior $k(\tau|y)$ is the pdf of a $\Gamma(n, 1/y)$ distribution.

(d) Determine the posterior pdf of $2y\tau$. Use it to obtain a $(1-\alpha)100\%$ credible interval for θ.

(e) Use the posterior pdf in part (d) to determine a Bayesian test for the hypotheses $H_0 :\ \theta \geq \theta_0$ versus $H_1 :\ \theta < \theta_0$, where θ_0 is specified.

11.2.6. Consider the Bayes model

$$\begin{array}{rcl} X_i|\theta\ , i = 1, 2, \ldots, n & \sim & \text{iid with distribution Poisson } (\theta),\ \theta > 0 \\ \Theta & \sim & h(\theta) \propto \theta^{-1/2}. \end{array}$$

(a) Show that $h(\theta)$ is in the class of Jeffreys' priors.

(b) Show that the posterior pdf of $2n\theta$ is the pdf of a $\chi^2(2y+1)$ distribution, where $y = \sum_{i=1}^{n} x_i$.

(c) Use the posterior pdf of part (b) to obtain a $(1-\alpha)100\%$ credible interval for θ.

(d) Use the posterior pdf in part (d) to determine a Bayesian test for the hypotheses $H_0:\ \theta \geq \theta_0$ versus $H_1:\ \theta < \theta_0$, where θ_0 is specified.

11.2.7. Consider the Bayes model

$$X_i|\theta\ , i = 1, 2, \ldots, n \sim \text{iid with distribution } b(1, \theta),\ 0 < \theta < 1.$$

(a) Obtain the Jeffreys' prior for this model.

(b) Assume squared-error loss and obtain the Bayes estimate of θ.

11.2.8. Consider the Bayes model

$$\begin{array}{rcl} X_i|\theta\ , i = 1, 2, \ldots, n & \sim & \text{iid with distribution } b(1, \theta),\ 0 < \theta < 1 \\ \Theta & \sim & h(\theta) = 1. \end{array}$$

(a) Obtain the posterior pdf.

(b) Assume squared-error loss and obtain the Bayes estimate of θ.

11.2.9. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be a random sample from a multivariate normal normal distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)'$ and known positive definite covariance matrix $\boldsymbol{\Sigma}$. Let $\overline{\mathbf{X}}$ be the mean vector of the random sample. Suppose that $\boldsymbol{\mu}$ has a prior multivariate normal distribution with mean $\boldsymbol{\mu}_0$ and positive definite covariance matrix $\boldsymbol{\Sigma}_0$. Find the posterior distribution of μ, given $\overline{\mathbf{X}} = \overline{\mathbf{x}}$. Then find the Bayes estimate $E(\boldsymbol{\mu}\,|\,\overline{\mathbf{X}} = \overline{\mathbf{x}})$.

11.3 Gibbs Sampler

From the preceding sections, it is clear that integration techniques play a significant role in Bayesian inference. Hence, we now touch on some of the Monte Carlo techniques used for integration in Bayesian inference.

The Monte Carlo techniques discussed in Chapter 5 can often be used to obtain Bayesian estimates. For example, suppose a random sample is drawn from a

$N(\theta, \sigma^2)$, where σ^2 is known. Then $Y = \overline{X}$ is a sufficient statistic. Consider the Bayes model

$$\begin{array}{rcl} Y|\theta & \sim & N(\theta, \sigma^2/n) \\ \Theta & \sim & h(\theta) \propto b^{-1}\exp\{-(\theta-a)/b\}/(1+\exp\{-[(\theta-a)/b]\})^2,\ -\infty < \theta < \infty, \\ & & a \text{ and } b > 0 \text{ are known,} \end{array} \tag{11.3.1}$$

i.e., the prior is a logistic distribution. Thus the posterior pdf is

$$k(\theta|y) = \frac{\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left\{-\frac{1}{2}\frac{(y-\theta)^2}{\sigma^2/n}\right\}b^{-1}e^{-(\theta-a)/b}/(1+e^{-[(\theta-a)/b]})^2}{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left\{-\frac{1}{2}\frac{(y-\theta)^2}{\sigma^2/n}\right\}b^{-1}e^{-(\theta-a)/b}/(1+e^{-[(\theta-a)/b]})^2\,d\theta}.$$

Assuming squared-error loss, the Bayes estimate is the mean of this posterior distribution. Its computation involves two integrals, which cannot be obtained in closed form. We can, however, think of the integration in the following way. Consider the likelihood $f(y|\theta)$ as a function of θ; that is, consider the function

$$w(\theta) = f(y|\theta) = \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left\{-\frac{1}{2}\frac{(y-\theta)^2}{\sigma^2/n}\right\}.$$

We can then write the Bayes estimate as

$$\begin{aligned} \delta(y) &= \frac{\int_{-\infty}^{\infty}\theta w(\theta)b^{-1}e^{-(\theta-a)/b}/(1+e^{-[(\theta-a)/b]})^2\,d\theta}{\int_{-\infty}^{\infty}w(\theta)b^{-1}e^{-(\theta-a)/b}/(1+e^{-[(\theta-a)/b]})^2\,d\theta} \\ &= \frac{E[\Theta w(\Theta)]}{E[w(\Theta)]}, \end{aligned} \tag{11.3.2}$$

where the expectation is taken with Θ having the logistic prior distribution.

The estimation can be carried out by simple Monte Carlo. Independently, generate $\Theta_1, \Theta_2, \ldots, \Theta_m$ from the logistic distribution with pdf as in (11.3.1). This generation is easily computed because the inverse of the logistic cdf is given by $a + b\log\{u/(1-u)\}$, for $0 < u < 1$. Then form the random variable,

$$T_m = \frac{m^{-1}\sum_{i=1}^m \Theta_i w(\Theta_i)}{m^{-1}\sum_{i=1}^m w(\Theta_i)}. \tag{11.3.3}$$

By the Weak Law of Large Numbers (Theorem 5.1.1) and Slutsky's Theorem (Theorem 5.2.4), $T_m \to \delta(y)$, in probability. The value of m can be quite large. Thus simple Monte Carlo techniques enable us to compute this Bayes estimate. Note that we can bootstrap this sample to obtain a confidence interval for $E[\Theta w(\Theta)]/E[w(\Theta)]$; see Exercise 11.3.2.

Besides simple Monte Carlo methods, there are other more complicated Monte Carlo procedures that are useful in Bayesian inference. For motivation, consider the case in which we want to generate an observation that has pdf $f_X(x)$, but this generation is somewhat difficult. Suppose, however, that it is easy to generate both Y, with pdf $f_Y(y)$, and an observation from the conditional pdf $f_{X|Y}(x|y)$. As the following theorem shows, if we do these sequentially, then we can easily generate from $f_X(x)$.

Theorem 11.3.1. *Suppose we generate random variables by the following algorithm:*

$$\begin{array}{ll} \textit{1.} & \textit{Generate } Y \sim f_Y(y), \\ \textit{2.} & \textit{Generate } X \sim f_{X|Y}(x|Y). \end{array}$$

Then X has pdf $f_X(x)$.

Proof: To avoid confusion, let T be the random variable generated by the algorithm. We need to show that T has pdf $f_X(x)$. Probabilities of events concerning T are conditional on Y and are taken with respect to the cdf $F_{X|Y}$. Recall that probabilities can always be written as expectations of indicator functions and, hence, for events concerning T, are conditional expectations. In particular, for any $t \in R$,

$$\begin{aligned} P[T \leq t] &= E[F_{X|Y}(t)] \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{t} f_{X|Y}(x|y)\, dx \right] f_Y(y)\, dy \\ &= \int_{-\infty}^{t} \left[\int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y)\, dy \right] dx \\ &= \int_{-\infty}^{t} \left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \right] dx \\ &= \int_{-\infty}^{t} f_X(x)\, dx. \end{aligned}$$

Hence the random variable generated by the algorithm has pdf $f_X(x)$, as was to be shown. ■

In the situation of this theorem, suppose we want to determine $E[W(X)]$, for some function $W(x)$, where $E[W^2(X)] < \infty$. Using the algorithm of the theorem, generate independently the sequence $(Y_1, X_1), (Y_2, X_2), \ldots, (Y_m, X_m)$, for a specified value of m, where Y_i is drawn from the pdf $f_Y(y)$ and X_i is generated from the pdf $f_{X|Y}(x|Y)$. Then by the Weak Law of Large Numbers,

$$\overline{W} = \frac{1}{m} \sum_{i=1}^{m} W(X_i) \xrightarrow{P} \int_{-\infty}^{\infty} W(x) f_X(x)\, dx = E[W(X)].$$

Furthermore, by the Central Limit Theorem, $\sqrt{m}(\overline{W} - E[W(X)])$ converges in distribution to a $N(0, \sigma_W^2)$ distribution, where $\sigma_W^2 = \mathrm{Var}(W(X))$. If $w_1, w_2, \ldots, w_m$ is a realization of such a random sample, then an approximate $(1-\alpha)100\%$ (large sample) confidence interval for $E[W(X)]$ is

$$\overline{w} \pm z_{\alpha/2} \frac{s_W}{\sqrt{m}}, \tag{11.3.4}$$

where $s_W^2 = (m-1)^{-1} \sum_{i=1}^{m} (w_i - \overline{w})^2$.

To set ideas, we present the following simple example.

Example 11.3.1. Suppose the random variable X has pdf

$$f_X(x) = \begin{cases} 2e^{-x}(1-e^{-x}) & 0 < x < \infty \\ 0 & \text{elsewhere.} \end{cases} \tag{11.3.5}$$

Suppose Y and $X|Y$ have the respective pdfs

$$f_Y(y) = \begin{cases} 2e^{-2y} & 0 < x < \infty \\ 0 & \text{elsewhere} \end{cases} \tag{11.3.6}$$

$$f_{X|Y}(x|y) = \begin{cases} e^{-(x-y)} & y < x < \infty \\ 0 & \text{elsewhere.} \end{cases} \tag{11.3.7}$$

Suppose we generate random variables by the following algorithm:

1. Generate $Y \sim f_Y(y)$ as in expression (11.3.6).
2. Generate $X \sim f_{X|Y}(x|Y)$ as in expression (11.3.7).

Then, by Theorem 11.3.1, X has the pdf (11.3.5). Furthermore, it is easy to generate from the pdfs (11.3.6) and (11.3.7) because the inverses of the respective cdfs are given by $F_Y^{-1}(u) = -2^{-1}\log(1-u)$ and $F_{X|Y}^{-1}(u) = -\log(1-u) + Y$.

As a numerical illustration, the R function `condsim1` (found at the site listed in the Preface) uses this algorithm to generate observations from the pdf (11.3.5). Using this function, we performed $m = 10{,}000$ simulations of the algorithm. The sample mean and standard deviation were $\overline{x} = 1.495$ and $s = 1.112$. Hence a 95% confidence interval for $E(X)$ is $(1.473, 1.517)$, which traps the true value $E(X) = 1.5$; see Exercise 11.3.4. ■

For the last example, Exercise 11.3.3 establishes the joint distribution of (X, Y) and shows that the marginal pdf of X is given by (11.3.5). Furthermore, as shown in this exercise, it is easy to generate from the distribution of X directly. In Bayesian inference, though, we are often dealing with conditional pdfs, and theorems such as Theorem 11.3.1 are quite useful.

The main purpose of presenting this algorithm is to motivate another algorithm, called the **Gibbs Sampler**, which is useful in Bayes methodology. We describe it in terms of two random variables. Suppose (X, Y) has pdf $f(x, y)$. Our goal is to generate two streams of iid random variables, one on X and the other on Y. The Gibbs sampler algorithm is:

Algorithm 11.3.1 (Gibbs Sampler)**.** *Let m be a positive integer, and let X_0, an initial value, be given. Then for $i = 1, 2, 3, \ldots, m$,*

1. Generate $Y_i|X_{i-1} \sim f(y|x)$.

2. Generate $X_i|Y_i \sim f(x|y)$.

Note that before entering the ith step of the algorithm, we have generated X_{i-1}. Let x_{i-1} denote the observed value of X_{i-1}. Then, using this value, generate sequentially the new Y_i from the pdf $f(y|x_{i-1})$ and then draw (the new) X_i from the pdf $f(x|y_i)$, where y_i is the observed value of Y_i. In advanced texts, it is shown that

$$\begin{aligned} Y_i &\xrightarrow{D} Y \sim f_Y(y) \\ X_i &\xrightarrow{D} X \sim f_X(x), \end{aligned} \tag{11.3.8}$$

as $i \to \infty$, and

$$\frac{1}{m}\sum_{i=1}^{m} W(X_i) \xrightarrow{P} E[W(X)], \text{ as } m \to \infty. \tag{11.3.9}$$

Note that the Gibbs sampler is similar but not quite the same as the algorithm given by Theorem 11.3.1. Consider the sequence of generated pairs

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k), (X_{k+1}, Y_{k+1}).$$

Note that to compute (X_{k+1}, Y_{k+1}), we need only the pair (X_k, Y_k) and none of the previous pairs from 1 to $k-1$. That is, given the present state of the sequence, the future of the sequence is independent of the past. In stochastic processes such a sequence is called a **Markov chain**. Under general conditions, the distribution of Markov chains stabilizes (reaches an equilibrium or steady-state distribution) as the length of the chain increases. For the Gibbs sampler, the equilibrium distributions are the limiting distributions in the expression (11.3.8) as $i \to \infty$. How large should i be? In practice, usually the chain is allowed to run to some large value i before recording the observations. Furthermore, several recordings are run with this value of i and the resulting empirical distributions of the generated random observations are examined for their similarity. Also, the starting value for X_0 is needed; see Casella and George (1992) for a discussion. The theory behind the convergences given in the expression (11.3.8) is beyond the scope of this text. There are many excellent references on this theory. A discussion from an elementary level can be found in Casella and George (1992). An informative overview can be found in Chapter 7 of Robert and Casella (1999); see also Lehmann and Casella (1998). We next provide a simple example.

Example 11.3.2. Suppose (X, Y) has the mixed discrete-continuous pdf given by

$$f(x, y) = \begin{cases} \frac{1}{\Gamma(\alpha)}\frac{1}{x!}y^{\alpha+x-1}e^{-2y} & y > 0;\ x = 0, 1, 2, \ldots \\ 0 & \text{elsewhere,} \end{cases} \tag{11.3.10}$$

for $\alpha > 0$. Exercise 11.3.5 shows that this is a pdf and obtains the marginal pdfs. The conditional pdfs, however, are given by

$$f(y|x) \propto y^{\alpha+x-1}e^{-2y} \tag{11.3.11}$$

and

$$f(x|y) \propto e^{-y}\frac{y^x}{x!}. \tag{11.3.12}$$

Hence the conditional densities are $\Gamma(\alpha+x, 1/2)$ and Poisson (y), respectively. Thus the Gibbs sampler algorithm is, for $i = 1, 2, \ldots, m$,

1. Generate $Y_i|X_{i-1} \sim \Gamma(\alpha + X_{i-1}, 1/2)$.
2. Generate $X_i|Y_i \sim \text{Poisson}(Y_i)$.

In particular, for large m and $n > m$,

$$\overline{Y} = (n-m)^{-1} \sum_{i=m+1}^{n} Y_i \xrightarrow{P} E(Y) \tag{11.3.13}$$

$$\overline{X} = (n-m)^{-1} \sum_{i=m+1}^{n} X_i \xrightarrow{P} E(X). \tag{11.3.14}$$

In this case, it can be shown (see Exercise 11.3.5) that both expectations are equal to α. The R function `gibbser2.s`, found at the site listed in the Preface, computes this Gibbs sampler. Using this routine, the authors obtained the following results upon setting $\alpha = 10$, $m = 3000$, and $n = 6000$:

Parameter	Estimate	Sample Estimate	Sample Variance	Approximate 95% Confidence Interval
$E(Y) = \alpha = 10$	$\overline{y}$	10.027	10.775	$(9.910, 10.145)$
$E(X) = \alpha = 10$	$\overline{x}$	10.061	21.191	$(9.896, 10.225)$

where the estimates $\overline{y}$ and $\overline{x}$ are the observed values of the estimators in expressions (11.3.13) and (11.3.14), respectively. The confidence intervals for α are the large sample confidence intervals for means discussed in Example 4.2.2, using the sample variances found in the fourth column of the above table. Note that both confidence intervals trapped $\alpha = 10$. ■

EXERCISES

11.3.1. Suppose Y has a $\Gamma(1,1)$ distribution while X given Y has the conditional pdf

$$f(x|y) = \begin{cases} e^{-(x-y)} & 0 < y < x < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Note that both the pdf of Y and the conditional pdf are easy to simulate.

(a) Set up the algorithm of Theorem 11.3.1 to generate a stream of iid observations with pdf $f_X(x)$.

(b) State how to estimate $E(X)$.

(c) Using your algorithm found in part (a), write an R function to estimate $E(X)$.

(d) Using your program, obtain a stream of 2000 simulations. Compute your estimate of $E(X)$ and find an approximate 95% confidence interval.

(e) Show that X has a $\Gamma(2,1)$ distribution. Did your confidence interval trap the true value 2?

11.3.2. Carefully write down the algorithm to obtain a bootstrap percentile confidence interval for $E[\Theta w(\Theta)]/E[w(\Theta)]$, using the sample $\Theta_1, \Theta_2, \ldots, \Theta_m$ and the estimator given in expression (11.3.3). Write R code for this bootstrap.

11.3.3. Consider Example 11.3.1.

(a) Show that $E(X) = 1.5$.

(b) Obtain the inverse of the cdf of X and use it to show how to generate X directly.

11.3.4. Obtain another 10,000 simulations similar to those discussed at the end of Example 11.3.1. Use your simulations to obtain a confidence interval for $E(X)$.

11.3.5. Consider Example 11.3.2.

(a) Show that the function given in expression (11.3.10) is a joint, mixed discrete-continuous pdf.

(b) Show that the random variable Y has a $\Gamma(\alpha, 1)$ distribution.

(c) Show that the random variable X has a negative binomial distribution with pmf

$$p(x) = \begin{cases} \frac{(\alpha+x-1)!}{x!(\alpha-1)!}2^{-(\alpha+x)} & x = 0, 1, 2, \ldots \\ 0 & \text{elsewhere.} \end{cases}$$

(d) Show that $E(X) = \alpha$.

11.3.6. Write an R function (or use `gibbser2.s`) for the Gibbs sampler discussed in Example 11.3.2. Run your function for $\alpha = 10$, $m = 3000$, and $n = 6000$. Compare your results with those of the authors tabled in the example.

11.3.7. Consider the following mixed discrete-continuous pdf for a random vector (X, Y) (discussed in Casella and George, 1992):

$$f(x, y) \propto \begin{cases} \binom{n}{x} y^{x+\alpha-1}(1-y)^{n-x+\beta-1} & x = 0, 1, \ldots, n,\ 0 < y < 1 \\ 0 & \text{elsewhere,} \end{cases}$$

for $\alpha > 0$ and $\beta > 0$.

(a) Show that this function is indeed a joint, mixed discrete-continuous pdf by finding the proper constant of proportionality.

(b) Determine the conditional pdfs $f(x|y)$ and $f(y|x)$.

(c) Write the Gibbs sampler algorithm to generate random samples on X and Y.

(d) Determine the marginal distributions of X and Y.

11.3.8. Write an R function for the Gibbs sampler of Exercise 11.3.7. Run your program for $\alpha = 10$, $\beta = 4$, $m = 3000$, and $n = 6000$. Obtain estimates (and confidence intervals) of $E(X)$ and $E(Y)$ and compare them with the true parameters.

11.4 Modern Bayesian Methods

The prior pdf has an important influence in Bayesian inference. We need only consider the different Bayes estimators for the normal model based on different priors, as shown in Examples 11.1.3 and 11.2.1. One way of having more control over the prior is to model the prior in terms of another random variable. This is called the **hierarchical Bayes** model, and it is of the form

$$\begin{aligned} X|\theta &\sim f(x|\theta) \\ \Theta|\gamma &\sim h(\theta|\gamma) \\ \Gamma &\sim \psi(\gamma). \end{aligned} \tag{11.4.1}$$

With this model we can exert control over the prior $h(\theta|\gamma)$ by modifying the pdf of the random variable Γ. A second methodology, **empirical Bayes**, obtains an estimate of γ and plugs it into the posterior pdf. We offer the reader a brief introduction of these procedures in this section. There are several good books on Bayesian methods. In particular, Chapter 4 of Lehmann and Casella (1998) discusses these procedures in some detail.

Consider first the hierarchical Bayes model given by (11.4.1). The parameter γ can be thought of a nuisance parameter. It is often called a **hyperparameter**. As with regular Bayes, the inference focuses on the parameter θ; hence, the posterior pdf of interest remains the conditional pdf $k(\theta|\mathbf{x})$.

These discussions often involve several pdfs; hence, we frequently use g as a generic pdf. It will always be clear from its arguments what distribution it represents. Keep in mind that the conditional pdf $f(\mathbf{x}|\theta)$ does not depend on γ; hence,

$$\begin{aligned} g(\theta,\gamma|\mathbf{x}) &= \frac{g(\mathbf{x},\theta,\gamma)}{g(\mathbf{x})} \\ &= \frac{g(\mathbf{x}|\theta,\gamma)g(\theta,\gamma)}{g(\mathbf{x})} \\ &= \frac{f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)}{g(\mathbf{x})}. \end{aligned}$$

Therefore, the posterior pdf is given by

$$k(\theta|\mathbf{x}) = \frac{\int_{-\infty}^{\infty} f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)\,d\gamma}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)\,d\gamma d\theta}. \tag{11.4.2}$$

Furthermore, assuming squared-error loss, the Bayes estimate of $W(\theta)$ is

$$\delta_W(\mathbf{x}) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(\theta)f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)\,d\gamma d\theta}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)\,d\gamma d\theta}. \tag{11.4.3}$$

Recall that we defined the Gibbs sampler in Section 11.3. Here we describe it to obtain the Bayes estimate of $W(\theta)$. For $i = 1, 2, \ldots, m$, where m is specified, the

ith step of the algorithm is

$$\begin{aligned}\Theta_i|\mathbf{x},\gamma_{i-1} &\sim g(\theta|\mathbf{x},\gamma_{i-1})\\ \Gamma_i|\mathbf{x},\theta_i &\sim g(\gamma|\mathbf{x},\theta_i).\end{aligned}$$

Recall from our discussion in Section 11.3 that

$$\begin{aligned}\Theta_i &\xrightarrow{D} k(\theta|\mathbf{x})\\ \Gamma_i &\xrightarrow{D} g(\gamma|\mathbf{x}),\end{aligned}$$

as $i \to \infty$. Furthermore, the arithmetic average

$$\frac{1}{m}\sum_{i=1}^{m} W(\Theta_i) \xrightarrow{P} E[W(\Theta)|\mathbf{x}] = \delta_W(\mathbf{x}) \text{ as } m \to \infty. \tag{11.4.4}$$

In practice, to obtain the Bayes estimate of $W(\theta)$ by the Gibbs sampler, we generate by Monte Carlo the stream of values $(\theta_1,\gamma_1),(\theta_2,\gamma_2)\ldots.$ Then choosing large values of m and $n^* > m$, our estimate of $W(\theta)$ is the average,

$$\frac{1}{n^*-m}\sum_{i=m+1}^{n^*} W(\theta_i). \tag{11.4.5}$$

Because of the Monte Carlo generation these procedures are often called **MCMC**, for **Markov Chain Monte Carlo** procedures. We next provide two examples.

Example 11.4.1. Reconsider the conjugate family of normal distributions discussed in Example 11.1.3, with $\theta_0 = 0$. Here we use the model

$$\begin{aligned}\overline{X}|\Theta &\sim N\left(\theta,\frac{\sigma^2}{n}\right),\ \sigma^2 \text{ is known}\\ \Theta|\tau^2 &\sim N(0,\tau^2)\\ \frac{1}{\tau^2} &\sim \Gamma(a,b),\ a \text{ and } b \text{ are known.}\end{aligned} \tag{11.4.6}$$

To set up the Gibbs sampler for this hierarchical Bayes model, we need the conditional pdfs $g(\theta|\overline{x},\tau^2)$ and $g(\tau^2|\overline{x},\theta)$. For the first, we have

$$g(\theta|\overline{x},\tau^2) \propto f(\overline{x}|\theta)h(\theta|\tau^2)\psi(\tau^{-2}).$$

As we have been doing, we can ignore standardizing constants; hence, we need only consider the product $f(\overline{x}|\theta)h(\theta|\tau^2)$. But this is a product of two normal pdfs which we obtained in Example 11.1.3. Based on those results, $g(\theta|\overline{x},\tau^2)$ is the pdf of a $N(\{\tau^2/[(\sigma^2/n)+\tau^2]\}\overline{x},(\tau^2\sigma^2)/[\sigma^2+n\tau^2])$. For the second pdf, by ignoring standardizing constants and simplifying, we obtain

$$\begin{aligned}g\left(\frac{1}{\tau^2}|\overline{x},\theta\right) &\propto f(\overline{x}|\theta)g(\theta|\tau^2)\psi(1/\tau^2)\\ &\propto \frac{1}{\tau}\exp\left\{-\frac{1}{2}\frac{\theta^2}{\tau^2}\right\}\left(\frac{1}{\tau^2}\right)^{a-1}\exp\left\{-\frac{1}{\tau^2}\frac{1}{b}\right\}\\ &\propto \left(\frac{1}{\tau^2}\right)^{a+(1/2)-1}\exp\left\{-\frac{1}{\tau^2}\left[\frac{\theta^2}{2}+\frac{1}{b}\right]\right\},\end{aligned} \tag{11.4.7}$$

which is the pdf of a $\Gamma\{a+(1/2), [(\theta^2/2)+(1/b)]^{-1}\}$ distribution. Thus the Gibbs sampler for this model is given by:

$$\begin{aligned}
\Theta_i|\overline{x}, \tau_{i-1}^2 &\sim N\left(\frac{\tau_{i-1}^2}{(\sigma^2/n)+\tau_{i-1}^2}\overline{x}, \frac{\tau_{i-1}^2\sigma^2}{\sigma^2+n\tau_{i-1}^2}\right) \\
\frac{1}{\tau_i^2}|\overline{x}, \Theta_i &\sim \Gamma\left(a+\frac{1}{2}, \left(\frac{\theta_i^2}{2}+\frac{1}{b}\right)^{-1}\right),
\end{aligned} \tag{11.4.8}$$

for $i = 1, 2, \ldots, m$. As discussed above, for a specified values of large m and $n^* > m$, we collect the chain's values $((\Theta_m, \tau_m), (\Theta_{m+1}, \tau_{m+1}), \ldots, (\Theta_{n^*}, \tau_{n^*}))$ and then obtain the Bayes estimate of θ (assuming squared-error loss):

$$\widehat{\theta} = \frac{1}{n^*-m}\sum_{i=m+1}^{n^*}\Theta_i. \tag{11.4.9}$$

The conditional distribution of Θ given $\overline{x}$ and τ_{i-1}, though, suggests the second estimate given by

$$\widehat{\theta}^* = \frac{1}{n^*-m}\sum_{i=m+1}^{n^*}\frac{\tau_i^2}{\tau_i^2+(\sigma^2/n)}\overline{x}. \quad \blacksquare \tag{11.4.10}$$

Example 11.4.2. Lehmann and Casella (1998, p. 257) presented the following hierarchical Bayes model:

$$\begin{aligned}
X|\lambda &\sim \text{Poisson}(\lambda) \\
\Lambda|b &\sim \Gamma(1, b) \\
B &\sim g(b) = \tau^{-1}b^{-2}\exp\{-1/b\tau\}, \quad b > 0, \tau > 0.
\end{aligned}$$

For the Gibbs sampler, we need the two conditional pdfs, $g(\lambda|x, b)$ and $g(b|x, \lambda)$. The joint pdf is

$$g(x, \lambda, b) = f(x|\lambda)h(\lambda|b)\psi(b). \tag{11.4.11}$$

Based on the pdfs of the model, (11.4.11), for the first conditional pdf we have

$$\begin{aligned}
g(\lambda|x, b) &\propto e^{-\lambda}\frac{\lambda^x}{x!}\frac{1}{b}e^{-\lambda/b} \\
&\propto \lambda^{x+1-1}e^{-\lambda[1+(1/b)]},
\end{aligned} \tag{11.4.12}$$

which is the pdf of a $\Gamma(x+1, b/[b+1])$ distribution.

For the second conditional pdf, we have

$$\begin{aligned}
g(b|x, \lambda) &\propto \frac{1}{b}e^{-\lambda/b}\tau^{-1}b^{-2}e^{-1/(b\tau)} \\
&\propto b^{-3}\exp\left\{-\frac{1}{b}\left[\frac{1}{\tau}+\lambda\right]\right\}.
\end{aligned}$$

In this last expression, making the change of variable $y = 1/b$ which has the Jacobian $db/dy = -y^{-2}$, we obtain

$$\begin{aligned} g(y|x,\lambda) &\propto y^3 \exp\left\{-y\left[\frac{1}{\tau}+\lambda\right]\right\} y^{-2} \\ &\propto y^{2-1}\exp\left\{-y\left[\frac{1+\lambda\tau}{\tau}\right]\right\}, \end{aligned}$$

which is easily seen to be the pdf of the $\Gamma(2, \tau/[\lambda\tau+1])$ distribution. Therefore, the Gibbs sampler is, for $i = 1, 2, \ldots, m$, where m is specified,

$$\begin{aligned} \Lambda_i|x, b_{i-1} &\sim \Gamma\left(x+1, b_{i-1}/[1+b_{i-1}]\right) \\ B_i = Y_i^{-1}, \text{ where } Y_i|x,\lambda_i &\sim \Gamma\left(2, \tau/[\lambda_i\tau+1]\right). \ \blacksquare \end{aligned}$$

As a numerical illustration of the last example, suppose we set $\tau = 0.05$ and observe $x = 6$. The R function[1] `hierarch1.s` computes the Gibbs sampler given in the example. It requires specification of the value of i at which the Gibbs sample commences and the length of the chain beyond this point. We set these values at $m = 1000$ and $n^* = 4000$, respectively, i.e., the length of the chain used in the estimate is 3000. To see the effect that varying τ has on the Bayes estimator, we performed five Gibbs samplers, with these results:

τ	0.040	0.045	0.050	0.055	0.060
$\widehat{\delta}$	6.600	6.490	6.530	6.500	6.440

There is some variation. As discussed in Lehmann and Casella (1998), in general, there is less effect on the Bayes estimator due to variability of the hyperparameter than in regular Bayes due to the variance of the prior.

11.4.1 Empirical Bayes

The empirical Bayes model consists of the first two lines of the hierarchical Bayes model; i.e.,

$$\begin{aligned} \mathbf{X}|\theta &\sim f(\mathbf{x}|\theta) \\ \Theta|\gamma &\sim h(\theta|\gamma). \end{aligned}$$

Instead of attempting to model the parameter γ with a pdf as in hierarchical Bayes, empirical Bayes methodology estimates γ based on the data as follows. Recall that

$$\begin{aligned} g(\mathbf{x},\theta|\gamma) &= \frac{g(\mathbf{x},\theta,\gamma)}{\psi(\gamma)} \\ &= \frac{f(\mathbf{x}|\theta)h(\theta|\gamma)\psi(\gamma)}{\psi(\gamma)} \\ &= f(\mathbf{x}|\theta)h(\theta|\gamma). \end{aligned}$$

[1]Downloadable at the site listed in the Preface

Consider, then, the likelihood function

$$m(\mathbf{x}|\gamma) = \int_{-\infty}^{\infty} f(\mathbf{x}|\theta)h(\theta|\gamma)\,d\theta. \tag{11.4.13}$$

Using the pdf $m(\mathbf{x}|\gamma)$, we obtain an estimate $\widehat{\gamma} = \widehat{\gamma}(\mathbf{x})$, usually by the method of maximum likelihood. For inference on the parameter θ, the empirical Bayes procedure uses the posterior pdf $k(\theta|\mathbf{x}, \widehat{\gamma})$.

We illustrate the empirical Bayes procedure with the following example.

Example 11.4.3. Consider the same situation discussed in Example 11.4.2, except assume that we have a random sample on X; i.e., consider the model

$$\begin{aligned} X_i|\lambda,\ i = 1, 2, \ldots, n &\sim \text{iid Poisson}(\lambda) \\ \Lambda|b &\sim \Gamma(1, b). \end{aligned}$$

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$. Hence,

$$g(\mathbf{x}|\lambda) = \frac{\lambda^{n\overline{x}}}{x_1! \cdots x_n!} e^{-n\lambda},$$

where $\overline{x} = n^{-1} \sum_{i=1}^{n} x_i$. Thus, the pdf we need to maximize is

$$\begin{aligned} m(\mathbf{x}|b) &= \int_0^{\infty} g(\mathbf{x}|\lambda)h(\lambda|b)\,d\lambda \\ &= \int_0^{\infty} \frac{1}{x_1! \cdots x_n!} \lambda^{n\overline{x}+1-1} e^{-n\lambda} \frac{1}{b} e^{-\lambda/b}\,d\lambda \\ &= \frac{\Gamma(n\overline{x}+1)[b/(nb+1)]^{n\overline{x}+1}}{x_1! \cdots x_n! b}. \end{aligned}$$

Taking the partial derivative of $\log m(\mathbf{x}|b)$ with respect to b, we obtain

$$\frac{\partial \log m(\mathbf{x}|b)}{\partial b} = -\frac{1}{b} + (n\overline{x}+1)\frac{1}{b(bn+1)}.$$

Setting this equal to 0 and solving for b, we obtain the solution

$$\widehat{b} = \overline{x}. \tag{11.4.14}$$

To obtain the empirical Bayes estimate of λ, we need to compute the posterior pdf with $\widehat{b}$ substituted for b. The posterior pdf is

$$\begin{aligned} k(\lambda|\mathbf{x}, \widehat{b}) &\propto g(\mathbf{x}|\lambda)h(\lambda|\widehat{b}) \\ &\propto \lambda^{n\overline{x}+1-1} e^{-\lambda[n+(1/\widehat{b})]}, \end{aligned} \tag{11.4.15}$$

which is the pdf of a $\Gamma(n\overline{x}+1, \widehat{b}/[n\widehat{b}+1])$ distribution. Therefore, the empirical Bayes estimator under squared-error loss is the mean of this distribution; i.e.,

$$\widehat{\lambda} = [n\overline{x}+1]\frac{\widehat{b}}{n\widehat{b}+1} = \overline{x}, \tag{11.4.16}$$

since $\widehat{b} = \overline{x}$. Thus, for the above prior, the empirical Bayes estimate agrees with the mle. ■

We can use our solution of this last example to obtain the empirical Bayes estimate for Example 11.4.2 also, for in this earlier example, the sample size is 1. Thus, the empirical Bayes estimate for λ is x. In particular, for the numerical case given at the end of Example 11.4.2, the empirical Bayes estimate has the value 6.

EXERCISES

11.4.1. Consider the Bayes model

$$\begin{aligned} X_i|\theta &\sim \text{iid } \Gamma\left(1, \frac{1}{\theta}\right) \\ \Theta|\beta &\sim \Gamma(2, \beta). \end{aligned}$$

By performing the following steps, obtain the empirical Bayes estimate of θ.

(a) Obtain the likelihood function

$$m(\mathbf{x}|\beta) = \int_0^\infty f(\mathbf{x}|\theta) h(\theta|\beta)\, d\theta.$$

(b) Obtain the mle $\widehat{\beta}$ of β for the likelihood $m(\mathbf{x}|\beta)$.

(c) Show that the posterior distribution of Θ given $\mathbf{x}$ and $\widehat{\beta}$ is a gamma distribution.

(d) Assuming squared-error loss, obtain the empirical Bayes estimator.

11.4.2. Consider the hierarchical Bayes model

$$\begin{aligned} Y &\sim b(n, p), \quad 0 < p < 1 \\ p|\theta &\sim h(p|\theta) = \theta p^{\theta - 1}, \quad \theta > 0 \\ \theta &\sim \Gamma(1, a), \quad a > 0 \text{ is specified.} \end{aligned} \tag{11.4.17}$$

(a) Assuming squared-error loss, write the Bayes estimate of p as in expression (11.4.3). Integrate relative to θ first. Show that both the numerator and denominator are expectations of a beta distribution with parameters $y + 1$ and $n - y + 1$.

(b) Recall the discussion around expression (11.3.2). Write an explicit Monte Carlo algorithm to obtain the Bayes estimate in part (a).

11.4.3. Reconsider the hierarchical Bayes model (11.4.17) of Exercise 11.4.2.

(a) Show that the conditional pdf $g(p|y, \theta)$ is the pdf of a beta distribution with parameters $y + \theta$ and $n - y + 1$.

(b) Show that the conditional pdf $g(\theta|y, p)$ is the pdf of a gamma distribution with parameters 2 and $\left[\frac{1}{a} - \log p\right]^{-1}$.

(c) Using parts (a) and (b) and assuming squared-error loss, write the Gibbs sampler algorithm to obtain the Bayes estimator of p.

11.4.4. For the hierarchical Bayes model of Exercise 11.4.2, set $n = 50$ and $a = 2$. Now, draw a θ at random from a $\Gamma(1, 2)$ distribution and label it θ^*. Next, draw a p at random from the distribution with pdf $\theta^* p^{\theta^*-1}$ and label it p^*. Finally, draw a y at random from a $b(n, p^*)$ distribution.

(a) Setting m at 3000, obtain an estimate of θ^* using your Monte Carlo algorithm of Exercise 11.4.2.

(b) Setting m at 3000 and n^* at 6000, obtain an estimate of θ^* using your Gibbs sampler algorithm of Exercise 11.4.3. Let $p_{3001}, p_{3002}, \ldots, p_{6000}$ denote the stream of values drawn. Recall that these values are (asymptotically) simulated values from the posterior pdf $g(p|y)$. Use this stream of values to obtain a 95% credible interval.

11.4.5. Write the Bayes model of Exercise 11.4.2 as

$$\begin{aligned} Y &\sim b(n, p), \quad 0 < p < 1 \\ p|\theta &\sim h(p|\theta) = \theta p^{\theta-1}, \quad \theta > 0. \end{aligned}$$

Set up the estimating equations for the mle of $g(y|\theta)$, i.e., the first step to obtain the empirical Bayes estimator of p. Simplify as much as possible.

11.4.6. Example 11.4.1 dealt with a hierarchical Bayes model for a conjugate family of normal distributions. Express that model as

$$\begin{aligned} \overline{X}|\Theta &\sim N\left(\theta, \frac{\sigma^2}{n}\right), \ \sigma^2 \text{ is known} \\ \Theta|\tau^2 &\sim N(0, \tau^2). \end{aligned}$$

Obtain the empirical Bayes estimator of θ.

This page intentionally left blank

Appendix A

Mathematical Comments

A.1 Regularity Conditions

These are the regularity conditions referred to in Sections 6.4 and 6.5 of the text. A discussion of these conditions can be found in Chapter 6 of Lehmann and Casella (1998).

Let X have pdf $f(x;\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Omega \subset R^p$. For these assumptions, X can be either a scalar random variable or a random vector in R^k. As in Section 6.4, let $\mathbf{I}(\boldsymbol{\theta}) = [I_{jk}]$ denote the $p \times p$ information matrix given by expression (6.4.4). Also, we will denote the true parameter $\boldsymbol{\theta}$ by $\boldsymbol{\theta}_0$.

Assumptions A.1.1. *Additional regularity conditions for Sections 6.4 and 6.5.*

(R6): *There exists an open subset $\Omega_0 \subset \Omega$ such that $\boldsymbol{\theta}_0 \in \Omega_0$ and all third partial derivatives of $f(x;\boldsymbol{\theta})$ exist for all $\boldsymbol{\theta} \in \Omega_0$.*

(R7) *The following equations are true (essentially, we can interchange expectation and differentiation):*

$$\begin{aligned} E_{\boldsymbol{\theta}}\left[\frac{\partial}{\partial\theta_j}\log f(x\,;\boldsymbol{\theta})\right] &= 0, \quad \textit{for } j = 1,\ldots,p \\ I_{jk}(\boldsymbol{\theta}) &= E_{\boldsymbol{\theta}}\left[-\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log f(x\,;\boldsymbol{\theta})\right], \quad \textit{for } j,k = 1,\ldots,p. \end{aligned}$$

(R8) *For all $\boldsymbol{\theta} \in \Omega_0$, $\mathbf{I}(\boldsymbol{\theta})$ is positive definite.*

(R9) *There exist functions $M_{jkl}(x)$ such that*

$$\left|\frac{\partial^3}{\partial\theta_j\partial\theta_k\theta_l}\log f(x\,;\boldsymbol{\theta})\right| \le M_{jkl}(x), \quad \textit{for all } \boldsymbol{\theta} \in \Omega_0,$$

and

$$E_{\boldsymbol{\theta}_0}[M_{jkl}] < \infty, \quad \textit{for all } j,k,l \in 1,\ldots,p. \quad \blacksquare$$

A.2 Sequences

The following is a short review of sequences of real numbers. In particular the liminf and limsup of sequences are discussed. As a supplement to this text, the authors offer a mathematical primer which can be downloaded at the site listed in the Preface. In addition to the following review of sequences, it contains a brief review of infinite series, and differentiable and integrable calculus including double integration. Students that need a review of these concepts can freely download this supplement.

Let $\{a_n\}$ be a sequence of real numbers. Recall from calculus that $a_n \to a$ ($\lim_{n\to\infty} a_n = a$) if and only if

$$\text{for every } \epsilon > 0\text{, there exists an } N_0 \text{ such that } n \geq N_0 \implies |a_n - a| < \epsilon. \quad \text{(A.2.1)}$$

Let A be a set of real numbers that is bounded from above; that is, there exists an $M \in R$ such that $x \leq M$ for all $x \in A$. Recall that a is the **supremum** of A if a is the least of all upper bounds of A. From calculus, we know that the supremum of a set bounded from above exists. Furthermore, we know that a is the supremum of A if and only if, for all $\epsilon > 0$, there exists an $x \in A$ such that $a - \epsilon < x \leq a$. Similarly, we can define the **infimum** of A.

We need three additional facts from calculus. The first is the Sandwich Theorem.

Theorem A.2.1 (Sandwich Theorem)**.** *Suppose for sequences* $\{a_n\}$, $\{b_n\}$, *and* $\{c_n\}$ *that* $c_n \leq a_n \leq b_n$, *for all* n, *and that* $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = a$. *Then* $\lim_{n\to\infty} a_n = a$.

Proof: Let $\epsilon > 0$ be given. Because both $\{b_n\}$ and $\{c_n\}$ converge, we can choose N_0 so large that $|c_n - a| < \epsilon$ and $|b_n - a| < \epsilon$, for $n \geq N_0$. Because $c_n \leq a_n \leq b_n$, it is easy to see that

$$|a_n - a| \leq \max\{|c_n - a|, |b_n - a|\},$$

for all n. Hence, if $n \geq N_0$, then $|a_n - a| < \epsilon$. ■

The second fact concerns subsequences. Recall that $\{a_{n_k}\}$ is a subsequence of $\{a_n\}$ if the sequence $n_1 \leq n_2 \leq \cdots$ is an infinite subset of the positive integers. Note that $n_k \geq k$.

Theorem A.2.2. *The sequence* $\{a_n\}$ *converges to* a *if and only if every subsequence* $\{a_{n_k}\}$ *converges to* a.

Proof: Suppose the sequence $\{a_n\}$ converges to a. Let $\{a_{n_k}\}$ be any subsequence. Let $\epsilon > 0$ be given. Then there exists an N_0 such that $|a_n - a| < \epsilon$, for $n \geq N_0$. For the subsequence, take k' to be the first index of the subsequence beyond N_0. Because for all k, $n_k \geq k$, we have that $n_k \geq n_{k'} \geq k' \geq N_0$, which implies that $|a_{n_k} - a| < \epsilon$. Thus, $\{a_{n_k}\}$ converges to a. The converse is immediate because a sequence is also a subsequence of itself. ■

Finally, the third theorem concerns monotonic sequences.

Theorem A.2.3. *Let $\{a_n\}$ be a nondecreasing sequence of real numbers; i.e., for all n, $a_n \leq a_{n+1}$. Suppose $\{a_n\}$ is bounded from above; i.e., for some $M \in R$, $a_n \leq M$ for all n. Then the limit of a_n exists.*

Proof: Let a be the supremum of $\{a_n\}$. Let $\epsilon > 0$ be given. Then there exists an N_0 such that $a - \epsilon < a_{N_0} \leq a$. Because the sequence is nondecreasing, this implies that $a - \epsilon < a_n \leq a$, for all $n \geq N_0$. Hence, by definition, $a_n \to a$. ■

Let $\{a_n\}$ be a sequence of real numbers and define the two subsequences

$$b_n = \sup\{a_n, a_{n+1}, \ldots\}, \quad n = 1, 2, 3 \ldots \tag{A.2.2}$$

$$c_n = \inf\{a_n, a_{n+1}, \ldots\}, \quad n = 1, 2, 3 \ldots. \tag{A.2.3}$$

It is obvious that $\{b_n\}$ is a nonincreasing sequence. Hence, if $\{a_n\}$ is bounded from below, then the limit of b_n exists. In this case, we call the limit of $\{b_n\}$ the **limit supremum** (limsup) of the sequence $\{a_n\}$ and write it as

$$\overline{\lim_{n\to\infty}} a_n = \lim_{n\to\infty} b_n. \tag{A.2.4}$$

Note that if $\{a_n\}$ is not bounded from below, then $\overline{\lim}_{n\to\infty} a_n = -\infty$. Also, if $\{a_n\}$ is not bounded from above, we define $\overline{\lim}_{n\to\infty} a_n = \infty$. Hence, the $\overline{\lim}$ of any sequence always exists. Also, from the definition of the subsequence $\{b_n\}$, we have

$$a_n \leq b_n, \quad n = 1, 2, 3, \ldots. \tag{A.2.5}$$

On the other hand, $\{c_n\}$ is a nondecreasing sequence. Hence, if $\{a_n\}$ is bounded from above, then the limit of c_n exists. We call the limit of $\{c_n\}$ the **limit infimum** (liminf) of the sequence $\{a_n\}$ and write it as

$$\underline{\lim_{n\to\infty}} a_n = \lim_{n\to\infty} c_n. \tag{A.2.6}$$

Note that if $\{a_n\}$ is not bounded from above, then $\underline{\lim}_{n\to\infty} a_n = \infty$. Also, if $\{a_n\}$ is not bounded from below, $\underline{\lim}_{n\to\infty} a_n = -\infty$. Hence, the $\underline{\lim}$ of any sequence always exists. Also, from the definition of the subsequences $\{c_n\}$ and $\{b_n\}$, we have

$$c_n \leq a_n \leq b_n, \quad n = 1, 2, 3, \ldots. \tag{A.2.7}$$

Also, because $c_n \leq b_n$ for all n, we have

$$\underline{\lim_{n\to\infty}} a_n \leq \overline{\lim_{n\to\infty}} a_n. \ ■ \tag{A.2.8}$$

Example A.2.1. Here are two examples. More are given in the exercises.

1. Suppose $a_n = -n$ for all $n = 1, 2, \ldots$. Then $b_n = \sup\{-n, -n-1, \ldots\} = -n \to -\infty$ and $c_n = \inf\{-n, -n-1, \ldots\} = -\infty \to -\infty$. So, $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n = -\infty$.

2. Suppose $\{a_n\}$ is defined by

$$a_n = \begin{cases} 1 + \frac{1}{n} & \text{if } n \text{ is even} \\ 2 + \frac{1}{n} & \text{if } n \text{ is odd.} \end{cases}$$

Then $\{b_n\}$ is the sequence $\{3, 2+(1/3), 2+(1/3), 2+(1/5), 2+(1/5), \ldots\}$, which converges to 2, while $\{c_n\} \equiv 1$, which converges to 1. Thus, $\underline{\lim}_{n\to\infty} a_n = 1$ and $\overline{\lim}_{n\to\infty} a_n = 2$. ■

It is useful that the $\underline{\lim}_{n\to\infty}$ and $\overline{\lim}_{n\to\infty}$ of every sequence exists. Also, the sandwich effects of expressions (A.2.7) and (A.2.8) lead to the following theorem.

Theorem A.2.4. *Let $\{a_n\}$ be a sequence of real numbers. Then the limit of $\{a_n\}$ exists if and only if $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$, in which case, $\lim_{n\to\infty} a_n = \underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$.*

Proof: Suppose first that $\lim_{n\to\infty} a_n = a$. Because the sequences $\{c_n\}$ and $\{b_n\}$ are subsequences of $\{a_n\}$, Theorem A.2.2 implies that they converge to a also. Conversely, if $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$, then expression (A.2.7) and the Sandwich Theorem, A.2.1, imply the result. ■

Based on this last theorem, we have two interesting applications that are frequently used in statistics and probability. Let $\{p_n\}$ be a sequence of probabilities and let $b_n = \sup\{p_n, p_{n+1}, \ldots\}$ and $c_n = \inf\{p_n, p_{n+1}, \ldots\}$. For the first application, suppose we can show that $\overline{\lim}_{n\to\infty} p_n = 0$. Then, because $0 \leq p_n \leq b_n$, the Sandwich Theorem implies that $\lim_{n\to\infty} p_n = 0$. For the second application, suppose we can show that $\underline{\lim}_{n\to\infty} p_n = 1$. Then, because $c_n \leq p_n \leq 1$, the Sandwich Theorem implies that $\lim_{n\to\infty} p_n = 1$.

We list some other properties in a theorem and ask the reader to provide the proofs in Exercise A.2.2:

Theorem A.2.5. *Let $\{a_n\}$ and $\{d_n\}$ be sequences of real numbers. Then*

$$\overline{\lim_{n\to\infty}}\,(a_n + d_n) \leq \overline{\lim_{n\to\infty}}\, a_n + \overline{\lim_{n\to\infty}}\, d_n \tag{A.2.9}$$

$$\underline{\lim_{n\to\infty}}\, a_n = -\overline{\lim_{n\to\infty}}\,(-a_n). \tag{A.2.10}$$

EXERCISES

A.2.1. Calculate the $\underline{\lim}$ and $\overline{\lim}$ of each of the following sequences:

(a) For $n = 1, 2, \ldots$, $a_n = (-1)^n \left(2 - \frac{4}{2^n}\right)$.

(b) For $n = 1, 2, \ldots$, $a_n = n^{\cos(\pi n/2)}$.

(c) For $n = 1, 2, \ldots$, $a_n = \frac{1}{n} + \cos \frac{\pi n}{2} + (-1)^n$.

A.2.2. Prove properties (A.2.9) and (A.2.10).

A.2.3. Let $\{a_n\}$ and $\{d_n\}$ be sequences of real numbers. Show that

$$\varliminf_{n\to\infty} (a_n + d_n) \geq \varliminf_{n\to\infty} a_n + \varliminf_{n\to\infty} d_n.$$

A.2.4. Let $\{a_n\}$ be a sequence of real numbers. Suppose $\{a_{n_k}\}$ is a subsequence of $\{a_n\}$. If $\{a_{n_k}\} \to a_0$ as $k \to \infty$, show that $\varliminf_{n\to\infty} a_n \leq a_0 \leq \varlimsup_{n\to\infty} a_n$.

This page intentionally left blank

Appendix B

R Primer

The package R can be downloaded at CRAN (https://cran.r-project.org/). It is freeware and there are versions for most platforms including Windows, Mac, and Linux. To install R simply follow the directions at CRAN. Installation should only take a few minutes. For more information on R, there are free downloadable manuals on its use at the CRAN website. There are many reference texts that the reader can consult, including the books by Venables and Ripley (2002), Verzani (2014), Crawley (2007), and Chapter 1 of Kloke and McKean (2014).

Once R is installed, in Windows, click on the R icon to begin an R session. The R prompt is a >. To exit R, type `q()`, which results in the query `Save workspace image? [y/n/c]:`. Upon typing `y`, the workspace will be saved for the next session. R has a built-in help (documentation) system. For example, to obtain help on the `mean` function, simply type `help(mean)`. To exit help, type `q`. We would recommend using R while working through the sections in this primer.

B.1 Basics

The commands of R work on numerical data, character strings, or logical types. To separate commands on the same line, use semicolons. Also, anything to the right of the symbol # is disregarded by R; i.e., to the right of # can be used for comments. Here are some arithmetic calculations:

```
> 8+6 - 7*2

[1] 0

> (150/3) + 7^2 -1 ; sqrt(50) - 50^(1/2)

[1] 98

[1] 0

> (4/3)*pi*5^3    # The volume of a sphere with radius 5
```

```
[1] 523.5988

> 2*pi*5          # The circumference of a sphere with radius 5

[1] 31.41593
```

Results can be saved for later calculation by either the **assignment** function `<-` or equivalently the equal symbol `=`. Names can be a mixture of letters, numbers, or symbols. For example:

```
> r <- 10 ; Vol <- (4/3)*pi*r^3 ; Vol

[1] 4188.79

> r = 100 ; circum = 2*pi*r ; circum

[1] 628.3185
```

Variables in R include scalars, vectors, or matrices. In the last example the variables `r` and `Vol` are scalars. Scalars can be combined into vectors with the `c` function. Further, arithmetic functions on vectors are performed componentwise. For instance, here are two ways to compute the volumes of spheres with radii $5, 6, \ldots, 9$.

```
> r <- c(5,6,7,8,9) ; Vol <- (4/3)*pi*r^3 ; Vol

[1]  523.5988  904.7787 1436.7550 2144.6606 3053.6281

> r <- 5:9 ; Vol <- (4/3)*pi*r^3 ; Vol

[1]  523.5988  904.7787 1436.7550 2144.6606 3053.6281
```

Components of a vector are referred to by using brackets. For example, the 5th component of the vector `vec` is `vec[5]`. Matrices can be formed from vectors using the commands `rbind` (combine rows) and `cbind` (combine columns) on vectors. To illustrate let $\mathbf{A}$ and $\mathbf{B}$ be the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}.$$

Then $\mathbf{AB}$, $\mathbf{A}^{-1}$, and $\mathbf{B}'\mathbf{A}$ are computed by

```
> c1 <- c(1,3) ; c2 <- c(4,2); a <- cbind(c1,c2)
> r1 <- c(1,3,5,7); r2 <- c(2,4,6,8); b <- rbind(r1,r2)
> a%*%b; solve(a) ; t(b)%*%a

     [,1] [,2] [,3] [,4]
[1,]    9   19   29   39
[2,]    7   17   27   37

   [,1] [,2]
c1 -0.2  0.4
c2  0.3 -0.1
```

```
     c1 c2
[1,]  7  8
[2,] 15 20
[3,] 23 32
[4,] 31 44
```

Brackets are also used to refer to elements of matrices. Let `amat` be a 4×4 matrix. Then the $(2, 3)$ element is `amat[2,3]` and the upper right corner 2×2 submatrix is `amat[1:2,3:4]`. This last item is an example of subsetting of a matrix. Subsetting is easy in R. For example, the following commands obtain the negative, positive, and elements of 0 for a vector `x`:

```
> x = c(-2,0,3,4,-7,-8,11,0);  xn = x[x<0]; xn

[1] -2 -7 -8

> xp = x[x>0]; xp

[1]  3  4 11

> x0 = x[x==0]; x0

[1] 0 0
```

For R vectors `x` and `y` of the same length, the plot of `y` versus `x` is obtained by the command `plot(y ~ x)`. The following segment of R code obtains plots found in Figure 2.1.1 of the volume and circumference of the sphere versus the radius for a sequence of radii from 0 to 8 in steps of 0.1. The first plot is a simple plot; the second plot adds some labeling and a title; the third plot draws a curve of the relationship; and the fourth plot shows the relationship between the circumference of the circle versus the radius.

```
par(mfrow=c(2,2))      # This sets up a 2 by 2 page of plots
r <- seq(0,8,.1);  Vol <- (4/3)*pi*r^3 ; plot(Vol ~ r)     # Plot 1
title("Simple Plot")
plot(Vol ~ r,xlab="Radius",ylab="Volume")                  # Plot 2
title("Volume vs Radius")
plot(Vol ~ r,pch=" ",xlab="Radius",ylab="Volume")
lines(Vol ~ r)                                             # Plot 3
title("Curve")
circum <- 2*pi*r
plot(circum ~ r,pch=" ",xlab="Radius",ylab="Circumference")
lines(circum ~ r); title("Circumference vs Radius")        # Plot 4
```

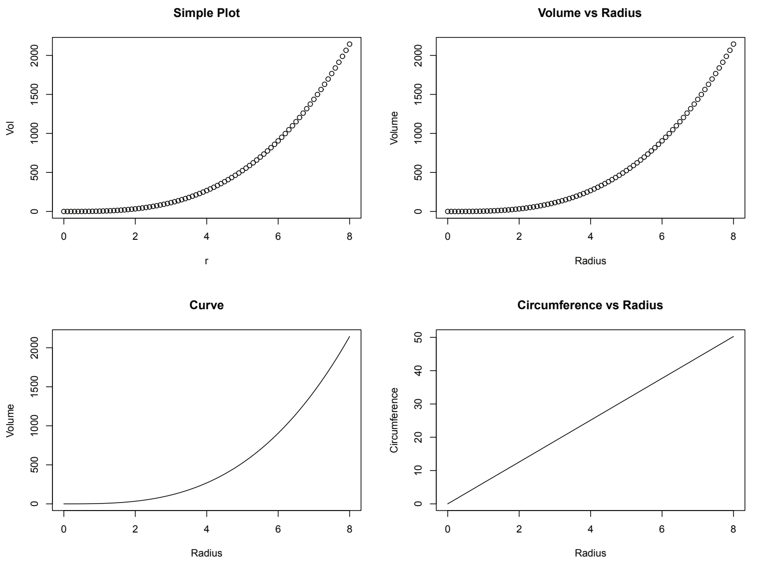

Figure 2.1.1: Spherical Plots discussed in Text.

B.2 Probability Distributions

For many distributions, R has functions that obtain probabilities, compute quantiles, and generate random variates. Here are two common examples. Let X be a random variable with a $N(\mu, \sigma^2)$ distribution. In R, let `mu` and `sig` denote the mean and standard deviation of X, respectively. Then the R commands and meanings are:

`pnorm(x,mu,sig)`	$P(X \leq x)$.
`qnorm(p,mu,sig)`	$P(X \leq q) = p$.
`dnorm(x,mu,sig)`	$f(x)$, where f is the pdf of X.
`rnorm(n,mu,sig)`	n variates generated from distribution of X.

As a numerical illustration, suppose the height of a male is normally distributed with mean 70 inches and standard deviation 4 inches.

```
> 1-pnorm(72,70,4)      # Prob. man exceeds 6 foot in ht.

[1] 0.3085375

> qnorm(.90,70,4)       # The upper 10th percentile in ht.

[1] 75.12621

> dnorm(72,70,4)        # value of density at 72
```

```
[1] 0.08801633

> rnorm(6,70,4)        # sample of size 6 on X

[1] 72.12486 75.25811 71.26661 63.36465 74.19436 69.71513
```

For the next figure, 2.2.2, we generate 100 variates, histogram the sample, and overlay the plot of the density of X on the histogram. Note the `pr=T` argument in the histogram. This scales the histogram to have area 1.

```
> x = rnorm(100,70,4);  x=sort(x)
> hist(x,pr=T,main="Histogram of Sample")
> y = dnorm(x,70,4)
> lines(y~x)
```

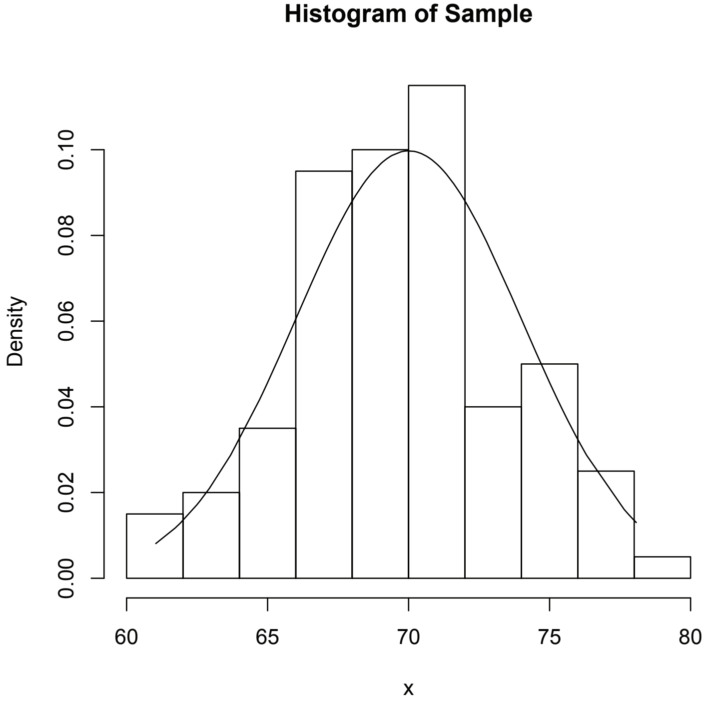

Figure 2.2.2: Histogram of a Random Sample from a $N(70, 4^2)$ distribution overlaid with the pdf of this normal.

For a discrete random variable the pdf is the probability mass function (pmf). Suppose X is binomial with 100 trials and 0.6 as the probability of success.

```
> pbinom(55,100,.6)      #  Probability of at most 55 successes

[1] 0.1789016

> dbinom(55,100,.6)      #  Probability of exactly 55 successes

[1] 0.04781118
```

Most other well known distributions are in core R. For example, here is the probability that a χ^2 random variable with 30 degrees of freedom exceeds 2 standard deviations form its mean, along with a Γ-distribution confirmation.

```
> mu=30; sig=sqrt(2*mu); 1-pchisq(mu+2*sig,30)

[1] 0.03471794

> 1-pgamma(mu+2*sig,15,1/2)

[1] 0.03471794
```

The `sample` command returns a random sample from a vector. It can either be sampling with replacement (`replace=T`) or sampling without replacement (`replace=F`). Here are samples of size 12 from the first 20 positive integers.

```
> vec = 1:20
> sample(vec,12,replace=T)

 [1] 14 20  7 17  6  6 11 11  9  1 10 14

> sample(vec,12,replace=F)

 [1] 12  1 14  5  4 11  3 17 16 19 20 15
```

B.3 R Functions

The syntax for R functions is the same as the syntax in R. This easily allows for the development of packages, a collection of R functions, for specific tasks. The schematic for an R function is

```
name-function <- function(arguments){
          ...  body of function  ...
}
```

Example B.3.1. Consider a process where a measurement is taken over time. At each time n, $n = 1, 2, \ldots$, the measurement x_n is observed but only the sample mean $\overline{x}_n = (1/n)\sum_{i=1}^{n} x_i$ of the measurements at time n is recorded and the point $(n, \overline{x}_n)$ is added to the running plot of sample means. How is this possible? There is a simple update formula for the sample mean that is easily derived. It is given by

$$\overline{x}_{n+1} = \frac{n}{n+1}\overline{x}_n + \frac{1}{n+1}x_{n+1}; \tag{B.3.1}$$

hence, the sample mean for the sequence $x_1, \ldots, x_{n+1}$ can be expressed as a linear combination of the sample mean at time n and the $(n+1)st$ measurement. The following R function codes this update formula:

```
mnupdate <- function(n,xbarn,xnp1){
#   Input: n is sample size; xbarn is mean of sample of size n;
#          xnp1 is (n+1) (new) observation
#   Output: mean of sample of size (n+1)
       mnupdate <- (n/(n+1))*xbarn + xnp1/(n+1)
       return(mnupdate)
}
```

To run this function we first source it in R. If the function is in the file `mnupdate.R` in the current directory then the source command is `source("mnupdate.R")`. It can also be copied and pasted into the current R session. Here is an execution of it:

```
> source("mnupdate.R")
> x = c(3,5,12,4); n=4; xbarn = mean(x);
> x; xbarn          #Old sample and its mean

[1]  3  5 12  4

[1] 6

> xp1 = 30        # New observation
> mnupdate(n,xbarn,xp1)    #  Mean of updated sample

[1] 10.8
```

■

B.4 Loops

Occasionally in the text, we use a loop in an R program to compute a result. Usually it is a simple `for` loop of the form

```
for(i in 1:n){
    ... R code often as a function of i ...
    #  For the n-iterations of the loop, i runs through
    #  the values i=1, i=2, ... , i=n.
}
```

For example, the following code segment produces a table of squares, cubes, square-roots, and cube-roots, for the integers from 1 to n.

```
#   set n at some value
tab <- c()            #  Initialize the table
for(i in 1:n){
     tab <- rbind(tab,c(i,i^2,i^3,i^(1/2),i^(1/3)))
}
tab
```

B.5 Input and Output

Many texts on R, including the references cited above, have information on input and output (I/O) in R. We only discuss several ways which are useful for the R discussion in our text. For output, we discuss two commands. The first writes an array (matrix) to a text file. Suppose `amat` is a matrix with p columns. Then the command `write(t(amat),ncol=p,file="amatrix.dat")` writes the matrix `amat` to the text file `amatrix.dat` in the current directory. Simply put the "Path" before the file as `file="Path/amatrix.dat"` to send it to another directory. The second way writes out variables to an R object file called an "rda" file. The variables can include scalars, vectors, matrices, and strings. For example the next line of code writes to an rda file the scalars `avar` and `bscale` and the matrix `amat` along with an information string.

```
info <- "This file contains the variable ....."
save(avar,bscale,amat,info,file="try.rda")
```

The command `load("try.rda")` will load these variables (names and values) into the current session. Most of the data sets discussed in the text are in rda files.

For input, we have already discussed the `c` and `load` functions. The `c` function is tedious, though, and a much easier way is to use the `scan` function. For example, the following lines of code assign the vector $(1, 2, 3)$ to `x`:

```
x <- scan()
1   2
  3

```

The separator between values is white space and the empty line after the data signals the end of `x`'s values. Note that this allows data to be copied and pasted into R. A matrix can also be scanned similarly by using the `read.table` function; for example, the following command inputs the above matrix **A** with column header "c1" and "c2":

```
a <- read.table(header = TRUE, text = "
       c1 c2
       1   4
       3   2
       ")
```

Notice that copy and paste is also easily used with this command. If the matrix **A** is in the file `amat.dat` with no header, it can be read in as

```
a <- matrix(scan("amat.dat"),ncol=2,byrow=T)
```

B.6 Packages

An R package is a collection of R functions designed for specified tasks. For example, in Chapter 10, the packages `Rfit` and `npsm` are discussed that compute rank-based robust and nonparametric procedures. There are thousands of free packages available to users at the site CRAN. The package `hmcpkg` contains all the R functions and R data sets discussed in this text. It can be downloaded at the site:

```
http://www.stat.wmich.edu/mckean/hmchomepage/Pkg/
```

Once it is installed on your computer use the library command as shown next to use the package in an R session. The next segment of code prints out the first 3 lines of the baseball data set discussed in Example 4.2.4. The attach command allows us to access the variables of the data set, as we show for the variable `height`.

```
library(hmcpkg)
head(bb,3)
hand height weight hitind hitpitind average
1    1      74     218      1           0   3.330
2    0      75     185      1           1   0.286
3    1      77     219      2           0   3.040
attach(bb);  head(height,4)    # accessing the variable height
[1] 74 75 77 73
```

In Example 1.3.3, the derivation of the probability that in a group of n people at least 2 have the same birthday is given. The R function `bday`, included in the package, computes this probability. The following segment of code computes it for a group of size 10.

```
library(hmcpkg)
bday(10)
[1] 0.1169482
```

This page intentionally left blank

Appendix C

Lists of Common Distributions

In this appendix, we provide a short list of common distributions. For each distribution, we note the expression where the pmf or pdf is defined in the text, the formula for the pmf or pdf, its mean and variance, and its mgf. The first list contains common discrete distributions, and the second list contains common continuous distributions.

List of Common Discrete Distributions	
Bernoulli $0 < p < 1$	(3.1.1) $p(x) = p^x(1-p)^{1-x}, \quad x = 0, 1$ $\mu = p, \quad \sigma^2 = p(1-p)$ $m(t) = [(1-p) + pe^t], \quad -\infty < t < \infty$
Binomial $0 < p < 1$ $n = 1, 2, \ldots$	(3.1.2) $p(x) = \binom{n}{x} p^x(1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$ $\mu = np, \quad \sigma^2 = np(1-p)$ $m(t) = [(1-p) + pe^t]^n, \quad -\infty < t < \infty$
Geometric $0 < p < 1$	(3.1.4) $p(x) = p(1-p)^x, \quad x = 0, 1, 2, \ldots$ $\mu = \frac{q}{p}, \quad \sigma^2 = \frac{1-p}{p^2}$ $m(t) = p[1 - (1-p)e^t]^{-1}, \quad t < -\log(1-p)$
Hypergeometric (N, D, n) $n = 1, 2, \ldots, \min\{N, D\}$	(3.1.7) $p(x) = \frac{\binom{N-D}{n-x}\binom{D}{x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \ldots, n$ $\mu = n\frac{D}{N}, \quad \sigma^2 = n\frac{D}{N}\frac{N-D}{N}\frac{N-n}{N-1}$ The above pmf is the probability of obtaining x Ds in a sample of size n, without replacement.
Negative Binomial $0 < p < 1$ $r = 1, 2, \ldots$	(3.1.3) $p(x) = \binom{x+r-1}{r-1} p^r(1-p)^x, \quad x = 0, 1, 2, \ldots$ $\mu = \frac{rq}{p}, \quad \sigma^2 = \frac{r(1-p)}{p^2}$ $m(t) = p^r[1 - (1-p)e^t]^{-r}, \quad t < -\log(1-p)$
Poisson $m > 0$	(3.2.1) $p(x) = e^{-m}\frac{m^x}{x!}, \quad x = 0, 1, 2, \ldots$ $\mu = m, \quad \sigma^2 = m$ $m(t) = \exp\{m(e^t - 1)\}, \quad -\infty < t < \infty$

List of Common Continuous Distributions

beta $\alpha > 0$ $\beta > 0$	(3.3.9) $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$ $\mu = \frac{\alpha}{\alpha+\beta}, \quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ $m(t) = 1 + \sum_{i=1}^{\infty} \left(\prod_{j=0}^{k-1} \frac{\alpha+j}{\alpha+\beta+j} \right) \frac{t^i}{i!}, \quad -\infty < t < \infty$
Cauchy	(1.9.2) $f(x) = \frac{1}{\pi} \frac{1}{x^2+1}, \quad -\infty < x < \infty$ Neither the mean nor the variance exists. The mgf does not exist.
Chi-squared, $\chi^2(r)$ $r > 0$	(3.3.7) $f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{(r/2)-1} e^{-x/2}, \quad x > 0$ $\mu = r, \quad \sigma^2 = 2r$ $m(t) = (1-2t)^{-r/2}, \quad t < \frac{1}{2}$ $\chi^2(r) \Leftrightarrow \Gamma(r/2, 2)$ r is called the degrees of freedom.
Exponential $\lambda > 0$	(3.3.6) $f(x) = \lambda e^{-\lambda x}, \quad x > 0$ $\mu = \frac{1}{\lambda}, \quad \sigma^2 = \frac{1}{\lambda^2}$ $m(t) = [1 - (t/\lambda)]^{-1}, \quad t < \lambda$ Exponential$(\lambda) \Leftrightarrow \Gamma(1, 1/\lambda)$
F, $F(r_1, r_2)$ $r_1 > 0$ $r_2 > 0 > 0$	(3.6.6) $f(x) = \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{(x)^{r_1/2-1}}{(1+r_1x/r_2)^{(r_1+r_2)/2}}, \quad x > 0$ If $r_2 > 2$, $\mu = \frac{r_2}{r_2-2}$. If $r > 4$, $\sigma^2 = 2\left(\frac{r_2}{r_2-2}\right)^2 \frac{r_1+r_2-2}{r_1(r_2-4)}$. The mgf does not exist. r_1 is called the numerator degrees of freedom. r_2 is called the denominator degrees of freedom.
Gamma, $\Gamma(\alpha, \beta)$ $\alpha > 0$ $\beta > 0$	(3.3.2) $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x > 0$ $\mu = \alpha\beta, \quad \sigma^2 = \alpha\beta^2$ $m(t) = (1-\beta t)^{-\alpha}, \quad t < \frac{1}{\beta}$

Continuous Distributions, Continued	
Laplace $-\infty < \theta < \infty$	(2.2.4) $f(x) = \frac{1}{2}\, e^{-\lvert x-\theta\rvert}, \quad -\infty < x < \infty$ $\mu = \theta, \quad \sigma^2 = 2$ $m(t) = e^{t\theta}\frac{1}{1-t^2}, \quad -1 < t < 1$
Logistic $-\infty < \theta < \infty$	(6.1.8) $f(x) = \frac{\exp\{-(x-\theta)\}}{(1+\exp\{-(x-\theta)\})^2}, \quad -\infty < x < \infty$ $\mu = \theta, \quad \sigma^2 = \frac{\pi^2}{3}$ $m(t) = e^{t\theta}\Gamma(1-t)\Gamma(1+t), \quad -1 < t < 1$
Normal, $N(\mu, \sigma^2)$ $-\infty < \mu < \infty$ $\sigma > 0$	(3.4.6) $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \quad -\infty < x < \infty$ $\mu = \mu, \quad \sigma^2 = \sigma^2$ $m(t) = \exp\{\mu t + (1/2)\sigma^2 t^2\}, \quad -\infty < t < \infty$
t, $t(r)$ $r > 0$	(3.6.2) $f(x) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+x^2/r)^{(r+1)/2}}, \quad -\infty < x < \infty$ If $r > 1$, $\mu = 0$. If $r > 2$, $\sigma^2 = \frac{r}{r-2}$. The mgf does not exist. The parameter r is called the degrees of freedom.
Uniform $-\infty < a < b < \infty$	(1.7.4) $f(x) = \frac{1}{b-a}, \quad a < x < b$ $\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$ $m(t) = \frac{e^{bt}-e^{at}}{(b-a)t}, \quad -\infty < t < \infty$

Appendix D

Tables of Distributions

Prior to the current age of computing, probability tables for certain distributions were part of many text books in probability and statistics. These are not needed any longer. Most statistical computing packages offer easy-to-use calls to determine these probabilities and quantiles. This is certainly true of the language R as we have discussed through out this text. Also, many hand calculators have such functions.

Tables for the following distributions are presented:

Table I Selected quantiles for chi-square distributions.

Table II Cumulative distribution function for the standard normal random variable.

Table III Selected quantiles for t-distributions.

Table IV Selected quantiles for F-distributions.

Table I
Chi-Square Distribution

The following table presents selected quantiles of chi-square distribution, i.e., the values x such that

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw,$$

for selected degrees of freedom r. The R function `chistable.s` generates this table.

	$P(X \leq x)$							
r	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892

Table II
Normal Distribution

The following table presents the standard normal distribution. The probabilities tabled are

$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw.$$

Note that only the probabilities for $z \ge 0$ are tabled. To obtain the probabilities for $z < 0$, use the identity $\Phi(-z) = 1 - \Phi(z)$. At the bottom of the table, some useful quantiles of the standard normal distribution are displayed. The R function `normaltable.s` generates this table.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.241	2.326	2.576	2.807	3.291

Table III
t-Distribution

The following table presents selected quantiles of the t-distribution, i.e., the values t such that

$$P(T \leq t) = \int_{-\infty}^{t} \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)(1+w^2/r)^{(r+1)/2}}\, dw,$$

for selected degrees of freedom r. The last row gives the standard normal quantiles.

	$P(T \leq t)$					
r	0.900	0.950	0.975	0.990	0.995	0.999
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
∞	1.282	1.645	1.960	2.326	2.576	3.090

Table IV
F-Distribution
Upper 0.05 Critical Points

The following table presents selected 0.95 and 0.99 quantiles of the F-distribution, i.e., for $\alpha = 0.05, 0.01$, the values $F_\alpha(r_1, r_2)$ such that

$$\alpha = P(X \geq F_\alpha(r_1, r_2)) = \int_{F_\alpha(r_1,r_2)}^{\infty} \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1+r_1 w/r_2)^{(r_1+r_2)/2}}\, dw,$$

where r_1 and r_2 are the numerator and denominator degrees of freedom, respectively. The R function `fp1.r` generates this table.

$F_{0.05}(r_1, r_2)$									
	r_1								
r_2	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Table IV
F-Distribution, Continued
Upper 0.05 Critical Points

Generated by the R function `fp2.r`.

$F_{0.05}(r_1, r_2)$									
	r_1								
r_2	10	15	20	25	30	40	60	120	∞
1	241.88	245.95	248.01	249.26	250.10	251.14	252.20	253.25	254.31
2	19.40	19.43	19.45	19.46	19.46	19.47	19.48	19.49	19.50
3	8.79	8.70	8.66	8.63	8.62	8.59	8.57	8.55	8.53
4	5.96	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.62	4.56	4.52	4.50	4.46	4.43	4.40	4.36
6	4.06	3.94	3.87	3.83	3.81	3.77	3.74	3.70	3.67
7	3.64	3.51	3.44	3.40	3.38	3.34	3.30	3.27	3.23
8	3.35	3.22	3.15	3.11	3.08	3.04	3.01	2.97	2.93
9	3.14	3.01	2.94	2.89	2.86	2.83	2.79	2.75	2.71
10	2.98	2.85	2.77	2.73	2.70	2.66	2.62	2.58	2.54
11	2.85	2.72	2.65	2.60	2.57	2.53	2.49	2.45	2.40
12	2.75	2.62	2.54	2.50	2.47	2.43	2.38	2.34	2.30
13	2.67	2.53	2.46	2.41	2.38	2.34	2.30	2.25	2.21
14	2.60	2.46	2.39	2.34	2.31	2.27	2.22	2.18	2.13
15	2.54	2.40	2.33	2.28	2.25	2.20	2.16	2.11	2.07
16	2.49	2.35	2.28	2.23	2.19	2.15	2.11	2.06	2.01
17	2.45	2.31	2.23	2.18	2.15	2.10	2.06	2.01	1.96
18	2.41	2.27	2.19	2.14	2.11	2.06	2.02	1.97	1.92
19	2.38	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.20	2.12	2.07	2.04	1.99	1.95	1.90	1.84
21	2.32	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.15	2.07	2.02	1.98	1.94	1.89	1.84	1.78
23	2.27	2.13	2.05	2.00	1.96	1.91	1.86	1.81	1.76
24	2.25	2.11	2.03	1.97	1.94	1.89	1.84	1.79	1.73
25	2.24	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.07	1.99	1.94	1.90	1.85	1.80	1.75	1.69
27	2.20	2.06	1.97	1.92	1.88	1.84	1.79	1.73	1.67
28	2.19	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.03	1.94	1.89	1.85	1.81	1.75	1.70	1.64
30	2.16	2.01	1.93	1.88	1.84	1.79	1.74	1.68	1.62
40	2.08	1.92	1.84	1.78	1.74	1.69	1.64	1.58	1.51
60	1.99	1.84	1.75	1.69	1.65	1.59	1.53	1.47	1.39
120	1.91	1.75	1.66	1.60	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.67	1.57	1.51	1.46	1.39	1.32	1.22	1.00

Table IV
F-Distribution, Continued
Upper 0.01 Critical Points

The R function `fp3.r` generates this table.

$F_{0.01}(r_1, r_2)$									
	r_1								
r_2	1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.1	6022.5
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Table IV
F-Distribution, Continued
Upper 0.01 Critical Points

The R function `fp4.r` generates this table.

$F_{0.01}(r_1, r_2)$									
	r_1								
r_2	10	15	20	25	30	40	60	120	∞
1	6055.9	6157.3	6208.7	6239.8	6260.7	6286.8	6313.0	6339.4	6365.9
2	99.40	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	26.87	26.69	26.58	26.50	26.41	26.32	26.22	26.13
4	14.55	14.20	14.02	13.91	13.84	13.75	13.65	13.56	13.46
5	10.05	9.72	9.55	9.45	9.38	9.29	9.20	9.11	9.02
6	7.87	7.56	7.40	7.30	7.23	7.14	7.06	6.97	6.88
7	6.62	6.31	6.16	6.06	5.99	5.91	5.82	5.74	5.65
8	5.81	5.52	5.36	5.26	5.20	5.12	5.03	4.95	4.86
9	5.26	4.96	4.81	4.71	4.65	4.57	4.48	4.40	4.31
10	4.85	4.56	4.41	4.31	4.25	4.17	4.08	4.00	3.91
11	4.54	4.25	4.10	4.01	3.94	3.86	3.78	3.69	3.60
12	4.30	4.01	3.86	3.76	3.70	3.62	3.54	3.45	3.36
13	4.10	3.82	3.66	3.57	3.51	3.43	3.34	3.25	3.17
14	3.94	3.66	3.51	3.41	3.35	3.27	3.18	3.09	3.00
15	3.80	3.52	3.37	3.28	3.21	3.13	3.05	2.96	2.87
16	3.69	3.41	3.26	3.16	3.10	3.02	2.93	2.84	2.75
17	3.59	3.31	3.16	3.07	3.00	2.92	2.83	2.75	2.65
18	3.51	3.23	3.08	2.98	2.92	2.84	2.75	2.66	2.57
19	3.43	3.15	3.00	2.91	2.84	2.76	2.67	2.58	2.49
20	3.37	3.09	2.94	2.84	2.78	2.69	2.61	2.52	2.42
21	3.31	3.03	2.88	2.79	2.72	2.64	2.55	2.46	2.36
22	3.26	2.98	2.83	2.73	2.67	2.58	2.50	2.40	2.31
23	3.21	2.93	2.78	2.69	2.62	2.54	2.45	2.35	2.26
24	3.17	2.89	2.74	2.64	2.58	2.49	2.40	2.31	2.21
25	3.13	2.85	2.70	2.60	2.54	2.45	2.36	2.27	2.17
26	3.09	2.81	2.66	2.57	2.50	2.42	2.33	2.23	2.13
27	3.06	2.78	2.63	2.54	2.47	2.38	2.29	2.20	2.10
28	3.03	2.75	2.60	2.51	2.44	2.35	2.26	2.17	2.06
29	3.00	2.73	2.57	2.48	2.41	2.33	2.23	2.14	2.03
30	2.98	2.70	2.55	2.45	2.39	2.30	2.21	2.11	2.01
40	2.80	2.52	2.37	2.27	2.20	2.11	2.02	1.92	1.80
60	2.63	2.35	2.20	2.10	2.03	1.94	1.84	1.73	1.60
120	2.47	2.19	2.03	1.93	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.04	1.88	1.77	1.70	1.59	1.47	1.32	1.00

Appendix E

References

Abebe, A., Crimin, K., McKean, J. W., Haas, J. V., and Vidmar, T. J. (2001), Rank-based procedures for linear models: applications to pharmaceutical science data, *Drug Information Journal*, **35**, 947–971.

Afifi, A. A. and Azen, S. P. (1972), *Statistical Analysis: A Computer Oriented Approach*, New York: Academic Press.

Arnold, S. F. (1981), *The Theory of Linear Models and Multivariate Analysis*, New York: John Wiley and Sons.

Azzalini, A. A. (1985), A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

Box, G. E. P. and Muller, M. (1958), A note on the generation of random normal variates, *Annals of Mathematical Statistics*, **29**, 610–611.

Breiman, L. (1968), *Probability*, Reading, MA: Addison-Wesley.

Buck, R. C. (1965), *Advanced Calculus*, New York: McGraw-Hill.

Canty, A. and Ripley, B. (2017), boot: Bootstrap R (S-Plus) Functions. R package version 1.3-19.

Carmer, S. G. and Swanson, M. R. (1973), An evaluation of ten multiple comparison procedures by Monte Carlo methods,o *Journal of the American Statistical Association* 68, 66–74.

Casella, G. and George, E. I. (1992), Explaining the Gibbs sampler, *American Statistician*, **46**, 167–174.

Chang, W. H., McKean, J. W., Naranjo, J. D., and Sheather, S. J. (1999), High breakdown rank-based regression, *Journal of the American Statistical Association*, **94**, 205–219.

Chung, K. L. (1974), *A Course in Probability Theory*, New York: Academic Press.

Conover, W. J. and Iman, R. L. (1981), Rank transform as a bridge between parametric and nonparametric statistics, *American Statistician*, **35**, 124–133.

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton: Princeton Univerity Press.

Crawley, M. J. (2007), *The R Book*, Chichester, West Sussex, John Wiley & Sons, Ltd.

Curtiss, J. H. (1942), A note on the theory of moment generating functions, *Annals of Mathematical Statistics*, **13**, 430.

D'Agostino, R. B. and Stephens, M. A. (1986), *Goodness-of-Fit Techniques*, New York: Marcel Dekker.

Davison, A. C. and Hinkley, D. V. (1997), *Bootstrap Methods and Their Applications*, Cambridge, UK: Cambridge University Press.

Devore, J. L. (2012), *Probability & Statistics*, 8*th* Ed., Boston: Brooks/Cole.

Draper, N. R. and Smith, H. (1966), *Applied Regression Analysis*, New York: John Wiley & Sons.

DuBois, C., Ed. (1960), *Lowie's Selected Papers in Anthropology*, Berkeley: University of California Press.

Dunnett, C. W. (1980), *Journal of the American Statistical Association*, 50, 1096–1121.

Efron, B. (1979), Bootstrap methods: Another look at the jackknife, *Annals of Statistics*, **7**, 1–26.

Efron B. and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Graybill, F. A. (1969), *Introduction to Matrices with Applications in Statistics*, Belmont, CA: Wadsworth.

Graybill, F. A. (1976), *Theory and Application of the Linear Model*, North Scituate, MA: Duxbury.

Hald, A. (1952), *Statistical Theory with Engineering Applications*, New York: John Wiley & Sons.

Haldane, J. B. S. (1948), The precision of observed values of small frequencies, *Biometrika*, **35**, 297–303.

Hampel, F. R. (1974), The influence curve and its role in robust estimation, *Journal of the American Statistical Association*, **69**, 383–393.

Hardy, G. H. (1992), *A Course in Pure Mathematics*, Cambridge, UK: Cambridge University Press.

Hettmansperger, T. P. (1984), *Statistical Inference Based on Ranks*, New York: John Wiley & Sons.

Hettmansperger, T. P. and McKean, J. W. (2011), *Robust Nonparametric Statistical Methods*, 2nd Ed., Boca Raton, FL: CRC Press.

Hewitt, E. and Stromberg, K. (1965), *Real and Abstract Analysis*, New York: Springer-Verlag.

Hodges, J. L., Jr., and Lehmann, E. L. (1961), Comparison of the normal scores and Wilcoxon tests, In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 307–317, Berkeley: University of California Press.

Hodges, J. L., Jr., and Lehmann, E. L. (1963), Estimates of location based on rank tests, *Annals of Mathematical Statistics*, **34**, 598–611.

Hogg, R. V. and Craig, A. T. (1958), On the decomposition of certain chi-square variables, *Annals of Mathematical Statistics*, **29**, 608.

Hogg, R. V., Fisher, D. M., and Randles, R. H. (1975), A two-sample adaptive distribution-free test, *Journal of the American Statistical Association*, **70**, 656–661.

Hollander, M. and Wolfe, D. A. (1999), *Nonparametric Statistical Methods*, 2nd Ed., New York: John Wiley & Sons.

Hsu, J. C. (1996), *Multiple Comparisons*, London: Chapman Hall.

Huber, P. J. (1981), *Robust Statistics*, New York: John Wiley & Sons.

Ihaka, R. and Gentleman, R. (1996), R: A language for data analysis and graphics, *Journal of Computational and Graphical Statistics*, **5**, 229–314.

Jeffreys, H. (1961), *The Theory of Probability*, Oxford: Oxford University Press.

Johnson, R. A. and Wichern, D. W. (2008), *Applied Multivariate Statistical Analysis*, 6th Ed., Boston: Pearson.

Kendall, M. and Stuart, A. (1979), *The Advanced Theory of Statistics*, Vol. 2, New York: Macmillan.

Kendall, M. G. (1962), *Rank Correlation Methods*, 3rd Ed., London: Griffin.

Kennedy, W. J. and Gentle, J. E. (1980), *Statistical Computing*, New York: Marcel Dekker.

Kitchens, L. J. (1997), *Exploring Statistics: A Modern Introduction to Data Analysis and Inference*, *2nd*Ed., Wadsworth.

Kloke, J. D. and McKean, J. W. (2011), Rfit: R algorithms for rank-based fitting, Submitted.

Lehmann, E. L. (1983), *Theory of Point Estimation*, New York: John Wiley & Sons.

Lehmann, E. L. (1986), *Testing Statistical Hypotheses*, 2nd Ed., London: Chapman & Hall.

Lehmann, E. L. (1999), *Elements of Large Sample Theory*, New York: Springer-Verlag.

Lehmann, E. L. and Casella, G. (1998), *Theory of Point Estimation*, 2nd Ed., New York: Springer-Verlag.

Lehmann, E. L. and Scheffé, H. (1950), Completeness, similar regions, and unbiased estimation, *Sankhya*, **10**, 305–340.

Marsaglia, G. and Bray, T. A. (1964), A convenient method for generating normal variables, *SIAM Review*, **6**, 260–264.

McKean, J. W. (2004), Robust analyses of linear models, *Statistical Science*, **19**, 562–570.

McKean, J. W. and Vidmar, T. J. (1994), A comparison of two rank-based methods for the analysis of linear models, *American Statistician*, **48**, 220–229.

McKean, J. W., Vidmar, T. J. and Sievers, G. (1989), A Robust Two-Stage Multiple Comparison Procedure with Application to a Random Drug Screen, *Biometrics 45*, 1281-1297.

McLachlan, G. J. and Krishnan, T. (1997), *The EM Algorithm and Extensions*, New York: John Wiley & Sons.

Minitab (1991), MINITAB Reference Manual, Valley Forge, PA: Minitab, Inc.

Mosteller, F. and Tukey, J. W. (1977), *Data Reduction and Regression*, Reading, MA: Addison-Wesley.

Naranjo, J. D. and McKean, J. W. (1997), Rank regression with estimated scores, *Statistics and Probability Letters*, **33**, 209–216.

Nelson, W. (1982), *Applied Lifetime Data Analysis*, New York: John Wiley & Sons.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996), *Applied Linear Statistical Models*, 4th Ed., Chicago: Irwin.

Parzen, E. (1962), *Stochastic Processes*, San Francisco: Holden-Day.

Randles, R. H. and Wolfe, D. A. (1979), *Introduction to the Theory of Nonparametric Statistics*, New York: John Wiley and Sons.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, 2nd Ed., New York: John Wiley & Sons.

Rasmussen, S. (1992), *An Introduction to Statistics with Data Analysis*, Belmont, CA: Brroks/Cole.

Robert, C. P. and Casella, G. (1999), *Monte Carlo Statistical Methods*, New York: Springer-Verlag.

Rousseeuw, P. J. and Leroy, A. M. (1987), *Robust Regression and Outlier Detection*, New York: John Wiley & Sons.

Scheffé, H. (1959), *The Analysis of Variance*, New York: John Wiley & Sons.

Seber, G. A. F. (1984), *Multivariate Observations*, New York: John Wiley & Sons.

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons.

Sheather, S. J. and Jones M. C. (1991), A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society-Series B*, **53**, 683–690.

Silverman, B. W. (1986), *Density Estimation*, London: Chapman and Hall.

Shirley, E. A. C. (1981), A distribution-free method for analysis of covariance based on rank data, *Applied Statistics*, **30**, 158–162.

S-PLUS (2000), *S-PLUS 6.0 Guide to Statistics*, Vol. 2, Seattle: Data Analysis Division, MathSoft.

Stapleton, J. H. (2009), *Linear Statistical Models, 2nd ed.*, New York: John Wiley & Sons.

Stigler, S.M. (1977), Do robust estimators work with real data? *Annals of Statistics*, **5**, 1055–1078.

Terpstra, J. T. and McKean, J. W. (2005), Rank-based analyses of linear models using R, *Journal of Statistical Software*, **14**, http://www.jstatsoft.org/.

Tucker, H. G. (1967), *A Graduate Course in Probability*, New York: Academic Press.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley.

Venables, W. N. and Ripley, B. D. (2002), *Modern Applied Statistics with S*, 4th Ed., New York: Springer-Verlag.

Verzani, J. (2014), *Usng R for Introductory Statistics, 2nd Ed.*, Bocs Raton, FL: Chapman-Hall.

Willerman, L., Schultz, R., Rutledge, J. N., and Bigler, E. (1991), In vivo brain size and intelligence," *Intelligence*, **15**, 223–228.

This page intentionally left blank

Appendix F

Answers to Selected Exercises

Chapter 1

1.2.1 (a) $\{0,1,2,3,4\}$, $\{2\}$; (b) $(0,3)$, $\{x : 1 \leq x < 2\}$;
(c) $\{(x,y) : 1 < x < 2, 1 < y < 2\}$.

1.2.2 (a) $\{x : 0 < x \leq 5/8\}$.

1.2.3 $C_1 \cap C_2 = \{mary, mray\}$.

1.2.4 (c) $(\cup A_n)^c = \cap A_n^C$; $(\cap A_n)^c = \cup A_n^C$.

1.2.6 (a) $\{x : 0 < x < 3\}$,
(b) $\{(x,y) : 0 < x^2 + y^2 < 4\}$.

1.2.7 (a) $\{x : x = 2\}$, (b) ϕ,
(c) $\{(x,y) : x = 0, y = 0\}$.

1.2.8 (a) $\frac{80}{81}$, (b) 1.

1.2.9 $\frac{11}{16}, 0, 1$.

1.2.10 $\frac{8}{3}, 0, \frac{\pi}{2}$.

1.2.11 (a) $\frac{1}{2}$, (b) 0, (c) $\frac{2}{9}$.

1.2.12 (a) $\frac{1}{6}$, (b) 0.

1.2.14 10.

1.3.2 $\frac{1}{4}, \frac{1}{13}, \frac{1}{52}, \frac{4}{13}$.

1.3.3 $\frac{31}{32}, \frac{3}{64}, \frac{1}{32}, \frac{63}{64}$.

1.3.4 0.3.

1.3.5 $e^{-4}, 1 - e^{-4}, 1$.

1.3.6 $\frac{1}{2}$.

1.3.10 (a)$\binom{6}{4}/\binom{16}{4}$, (b) $\binom{10}{4}/\binom{16}{4}$.

1.3.11 $1 - \binom{990}{5}/\binom{1000}{5}$.

1.3.13 (b) $1 - \binom{10}{3}/\binom{20}{3}$.

1.3.15 (a) $1 - \binom{48}{5}/\binom{50}{5}$.

1.3.16 $n = 23$.

1.3.19 $13 \cdot 12\binom{4}{3}\binom{4}{2}/\binom{52}{5}$.

1.3.22 (a) $0 \leq \sum_{i=1}^{3} p_i \leq 1$, (b) no.

1.4.3 $\frac{9}{47}$.

1.4.4 $2\frac{13}{52}\frac{12}{51}\frac{26}{50}\frac{25}{49}$.

1.4.6 $\frac{111}{143}$.

1.4.8 (a) 0.022, (b) $\frac{5}{11}$.

1.4.9 $\frac{5}{14}$.

1.4.10 $\frac{3}{7}, \frac{4}{7}$.

1.4.12 (c) 0.88.

1.4.14 (a) 0.1764.

1.4.15 $4(0.7)^3(0.3)$.

1.4.16 0.75.

1.4.18 (a) $\frac{6}{11}$.

1.4.20 $\frac{1}{7}$.

1.4.21 (a) $1-\left(\frac{5}{6}\right)^6$, (b) $1-e^{-1}$.

1.4.23 $\frac{3}{4}$.

1.4.25 $\frac{43}{64}$.

1.4.27 (a) $\sum_{x=1}^{20} 4/[20(25-(x-1))]$
(b) `x=1:20;sum(4/((25-x+1)*20))`
(c) Download `ex1427.R`

1.4.28 $\frac{5\cdot4\cdot5\cdot4\cdot3}{10\cdot9\cdot8\cdot7\cdot6}$.

1.4.29 $\frac{13}{4}$.

1.4.30 $\frac{2}{3}$.

1.4.31 $0.518, 0.491$.

1.4.32 No.

1.5.1 $\frac{9}{13}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}$.

1.5.2 (a) $\frac{1}{2}$, (b) $\frac{1}{21}$.

1.5.3 $\frac{1}{5}, \frac{1}{5}, \frac{1}{5}$.

1.5.5 (a) $\frac{\binom{13}{x}\binom{39}{5-x}}{\binom{52}{5}}, x=0,1,2,3,4,5,$
(b) $\left[\binom{39}{5}+\binom{13}{1}\binom{39}{4}\right]/\binom{52}{5}$.

1.5.7 $\frac{3}{4}$.

1.5.8 For the plot download `ex158.R`
(a) $\frac{1}{4}$, (b) 0, (c) $\frac{1}{4}$, (d) 0.

1.6.2 (a) $p_X(x) = \frac{1}{10}, x = 1,2,\ldots,10,$
(b) $\frac{4}{10}$.

1.6.3 (a) $\left(\frac{5}{6}\right)^{x-1}\frac{1}{6}\ x=1,2,3,\ldots,$
(c) $\frac{6}{11}$.

1.6.4 $\frac{6}{36}, x=0; \frac{12-2x}{36}, x=1,2,3,4,5$.

1.6.5 (a) Download `dex165.R`.

1.6.7 $\frac{1}{3}, y=3,5,7$.

1.6.8 $\left(\frac{1}{2}\right)^{\sqrt[3]{y}}, y=1,8,27,\ldots$.

1.7.1 $F(x)=\frac{\sqrt{x}}{10}, 0\le x<100;$
$f(x)=\frac{1}{20\sqrt{x}}, 0<x<100$.

1.7.3 $\frac{5}{8}; \frac{7}{8}; \frac{3}{8}$.

1.7.5 $e^{-2}-e^{-3}$.

1.7.6 (a) $\frac{1}{27}, 1$; (b) $\frac{2}{9}, \frac{25}{36}$.

1.7.8 (a)1; (b) $\frac{2}{3}$; (c) 2.

1.7.9 (b) $\sqrt[3]{1/2}$; (c) 0.

1.7.10 $\sqrt[4]{0.2}$.

1.7.12 (a) $1-(1-x)^3, 0\le x<1$;
(b) $1-\frac{1}{x}, 1\le x<\infty$.

1.7.13 $xe^{-x}, 0<x<\infty$; mode is 1.

1.7.14 $\frac{7}{12}$.

1.7.17 $\frac{1}{2}$.

1.7.19 $-\sqrt{2}$.

1.7.20 (b) $f_y(y)=1/(5+y)^{1.2}$.
(c) `dlife <- function(y){1/(5+y)^(1.2)}`.

1.7.21 (a) $f(x)=(5/3)e^{-x}/[1+(2/3)e^{-x}]^{(7/2)}$.
(b) `f=function(x) {(1+(2/3)exp(-x))^(-5/2)}`

1.7.22 $\frac{1}{27}, 0<y<27$.

1.7.24 $\frac{1}{\pi(1+y^2)}, -\infty<y<\infty$.

1.7.25 cdf $1-e^{-y}, 0\le y<\infty$.

1.7.26 pdf $\frac{1}{3\sqrt{y}}, 0<y<1,$
$\frac{1}{6\sqrt{y}}, 1<y<4$.

1.8.3 $2, 86.4, -160.8$.

1.8.4 $3, 11, 27$.

1.8.5 $\frac{\log 100.5-\log 50.5}{50}$.

1.8.6 (a) $\frac{3}{4}$;(b) $\frac{1}{4}, \frac{1}{2}$.

1.8.7 $\frac{3}{20}$.

1.8.8 \$7.80.

1.8.9 (a) 2; (b) pdf is $\frac{2}{y^3}\,, 1 < y < \infty$; (c) 2.

1.8.10 $\frac{7}{3}$.

1.8.12 (a) $\frac{1}{2}$; (c) $\frac{1}{2}$.

1.8.13 $P[G = -p_0] = \frac{1}{3}$, $P[G = 1 - p_0] = \frac{2}{3}\frac{1}{2}$, $\ldots, P[G = 50 - p_0] = \frac{2}{3}\frac{1}{2}(0.0045)$.

1.8.14 Range of G: $\{2 - p_0, 5 - p_0, 8 - p_0\}$, Probs: $\frac{3}{10}, \frac{6}{10}, \frac{1}{10}$.

1.9.1 (a) $1.5\,, 0.75$; (b) $0.5, 0.05$; (c) 2, does not exist.

1.9.2 $\frac{e^t}{2-e^t}, t < \log 2; 2; 2$.

1.9.12 $10; 0; 2; -30$.

1.9.14 (a) $-\frac{2\sqrt{2}}{5}$; (b) 0; (c) $\frac{2\sqrt{2}}{5}$.

1.9.16 $\frac{1}{2p}; \frac{3}{2}; \frac{5}{2}; 5; 50$.

1.9.18 $\frac{31}{12}; \frac{167}{144}$.

1.9.19 $E(X^r) = \frac{(r+2)!}{2}$.

1.9.20 odd moments are 0, $E(X^{2n}) = (2n)!$.

1.9.24 $\frac{5}{8}; \frac{37}{192}$.

1.9.27 $(1 - \beta t)^{-1}, \beta, \beta^2$.

1.10.3 0.84.

1.10.4 $P(|X| \geq 5) = 0.0067$.

Chapter 2

2.1.1 $\frac{15}{64}; 0; \frac{1}{2}; \frac{1}{2}$.

2.1.2 $\frac{1}{4}$.

2.1.7 $ze^{-z}, 0 < z < \infty$.

2.1.8 $-\log z,\ 0 < z < 1$.

2.1.9 $\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}/\binom{52}{13}$, x and y nonnegative integers such that $x + y \leq 13$.

2.1.11 $\frac{15}{2}x_1^2(1 - x_1^2), 0 < x_1 < 1$; $5x_2^4, 0 < x_2 < 1$.

2.1.14 $\frac{2}{3}; \frac{1}{2}; \frac{2}{3}; \frac{1}{2}; \frac{4}{9}$; yes; $\frac{11}{3}$.

2.1.15 $\frac{e^{t_1+t_2}}{(2-e^{t_1})(2-e^{t_2})}, t_i < \log 2$.

2.1.16 $(1 - t_2)^{-1}(1 - t_1 - t_2)^{-2}, t_2 < 1$, $t_1 + t_2 < 1$; no.

2.2.2

1	2	3	4	6	9
$\frac{1}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{12}{36}$	$\frac{9}{36}$.

2.2.3 $e^{-y_1-y_2}, 0 < y_i < \infty$.

2.2.4 $8y_1y_2^3, 0 < y_i < 1$.

2.2.6 (a) $y_1e^{-y_1}, 0 < y_1 < \infty$; (b) $(1 - t_1)^{-2}, t_1 < 1$.

2.3.1 $\frac{3x_1+2}{6x_1+3}; \frac{6x_1^2+6x_1+1}{2(6x_1+3)^2}$.

2.3.2 (a) $2, 5$; (b) $10x_1x_2^2, 0 < x_1 < x_2 < 1$; (c) $\frac{12}{25}$; (d) $\frac{449}{1536}$.

2.3.3 (a) $\frac{3x_2}{4}; \frac{3x_2^2}{80}$; (b) pdf is $7(4/3)^7y^6, 0 < y < \frac{3}{4}$; (c) $E(X) = E(Y) = \frac{21}{32}$; $\text{Var}(X_1) = \frac{553}{15360} > \text{Var}(Y) = \frac{7}{1024}$.

2.3.8 $x + 1, 0 < x < \infty$.

2.3.9 (a) $\binom{13}{x_1}\binom{13}{x_2}\binom{26}{5-x_1-x_2}/\binom{52}{5}$, x_1, x_2 nonnegative integers, $x_1 + x_2 \leq 5$; (c) $\binom{13}{x_2}\binom{26}{5-x_1-x_2}/\binom{39}{5-x_1}$, $x_2 \leq 5 - x_1$.

2.3.11 (a) $\frac{1}{x_1}, 0 < x_2 < x_1 < 1$; (b) $1 - \log 2$.

2.3.12 (b) e^{-1}.

2.5.1 (a) 1; (b) -1; (c) 0.

2.5.2 (a) $\frac{7}{\sqrt{804}}$.

2.5.8 $1, 2, 1, 2, 1$.

2.5.9 $\frac{1}{2}$.

2.4.4 $\frac{5}{81}$.

2.4.5 $\frac{7}{8}$.

2.4.6 $2; 2$.

2.4.8 $\frac{2(1-y^3)}{3(1-y^2)}, 0 < y < 1$.

2.4.9 $\frac{1}{2}$.

2.4.12 $\frac{4}{9}$.

2.4.13 $4; 4$.

2.6.1 (g) $\frac{2+3y+3z}{3+6y+6z}$.

2.6.2 (a) $\frac{1}{6}; 0$;
(b) $(1-t_1)^{-1}(1-t_2)^{-1}(1-t_3)^{-1}$; yes.

2.6.3 pdf is $12(1-y)^{11}, 0 < y < 1$.

2.6.4 pmf is $\frac{y^3-(y-1)^3}{6^3}$.

2.6.6 $\sigma_1(\rho_{12} - \rho_{13}\rho_{23})/\sigma_2(1-\rho_{23}^2)$;
$\sigma_1(\rho_{13} - \rho_{12}\rho_{23})/\sigma_3(1-\rho_{23}^2)$.

2.6.9 (a) $\frac{3}{4}$.

2.7.1 joint pdf $y_2y_3^2e^{-y_3}, 0 < y_1 < 1$,
$0 < y_2 < 1, 0 < y_3 < \infty$.

2.7.2 $\frac{1}{2\sqrt{y}}, 0 < y < 1$.

2.7.3 $\frac{1}{4\sqrt{y}}, 0 < y < 1; \frac{1}{8\sqrt{y}}, 1 \le y < 9$.

2.7.7 $24y_2y_3^2y_4^3, 0 < y_i < 1$.

2.7.8 (a) $\frac{9}{16}; \frac{6}{16}; \frac{1}{16}$; (b) $\left(\frac{3}{4} + \frac{1}{4}e^t\right)^6$.

2.8.2 $\frac{8}{3}; \frac{2}{9}$.

2.8.3 7.

2.8.5 2.5; 0.25.

2.8.7 -5; 30.6.

2.8.8 $\frac{\sigma_1}{\sqrt{\sigma_1^2+\sigma_2^2}}$.

2.8.10 0.265.

2.8.12 22.5; 65.25.

2.8.13 $\frac{\mu_2\sigma_1}{\sqrt{\sigma_1^2\sigma_2^2+\mu_1^2\sigma_2^2+\mu_2^2\sigma_1^2}}$.

2.8.15 0.801.

Chapter 3

3.1.1 $\frac{40}{81}$.

3.1.4 `1-pbinom(34,40,7/8)=0.6162`.

3.1.5 $P(X \ge 20) = 0.0009$.

3.1.6 5.

3.1.11 $\frac{3}{16}$.

3.1.13 $\frac{65}{81}$.

3.1.15 $\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^{x-3}, x = 3, 4, 5, \ldots$.

3.1.16 $\frac{5}{72}$.

3.1.18 (a) Negative binomial, parameters r and T/N.

3.1.19 (b) Code:
```
ps=c(.3,.2,.2,.2,.1)
coll=c()
for(i in 1:10000)
{coll<-c(coll,multitrial(ps))}
table(coll)/10000
```

3.1.20 (a) -$2.40

3.1.22 $\frac{1}{6}$.

3.1.23 $\frac{24}{625}$.

3.1.25 (a) $\frac{11}{6}$; (b) $\frac{x_1}{2}$; (c) $\frac{11}{6}$.

3.1.26 $\frac{25}{4}$.

3.1.30 (a) 0.0853; (b) 0.2637; (c) 0.0861, 0.2639.

3.2.1 0.09.

3.2.4 $4^xe^{-4}/x!, x = 0, 1, 2 \ldots$.

3.2.5 0.84., 0.9858.

3.2.11 About 6.7.

3.2.13 8.

3.2.14 2.

3.2.16 (a) $e^{-2}\exp\{(1+e^{t_1})e^{t_2}\}$.

3.3.1 (a) 0.05.; (b) 0.9592

3.3.2 $0.831; 12.8$.

3.3.3 (b) 0.1355

3.3.4 $\chi^2(4)$.

3.3.6 pdf is $3e^{-3y}, 0 < y < \infty$.

3.3.7 $2; 0.95$.

3.3.14 (a) 0.0839; (b) 0.2424

3.3.15 $\frac{11}{16}$.

3.3.16 $\chi^2(2)$.

3.3.18 $\frac{\alpha}{\alpha+\beta}; \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$.

3.3.19 (a) 20; (b) 1260; (c) 495.

3.3.20 $\frac{10}{243}$.

3.3.24 (a) $(1-6t)^{-8}, t < \frac{1}{6}$; (b) $\Gamma(\alpha = 8, \beta = 6)$.

3.4.2 $0.067; 0.685$.

3.4.3 1.645.

3.4.4 $71.4; 189.4$.

3.4.8 0.598.

3.4.10 0.774.

3.4.11 (a) $\sqrt{\frac{2}{\pi}}; \frac{\pi-2}{\pi}$.

3.4.12 0.90.

3.4.13 0.477.

3.4.14 0.461.

3.4.15 $N(0, 1)$.

3.4.16 0.433.

3.4.17 $0; 3$.

3.4.22 $N(0, 2)$.

3.4.25 (a) 0.04550; (b) 0.1649

3.4.28 Mean is $\sqrt{2/\pi}(\alpha/\sqrt{1+\alpha^2})$.

3.4.29 0.24.

3.4.30 0.159.

3.4.31 0.159.

3.4.33 $\chi^2(2)$.

3.5.1 (a) 0.574; (b) 0.735.

3.5.2 (a) 0.264; (b) 0.440; (c) 0.433; (d) 0.643.

3.5.7 $\frac{4}{5}$.

3.5.8 $(38.2, 43.4)$.

3.5.17 0.05.

3.6.1 0.05.

3.6.2 1.761.

3.6.5 (d) 0.0734; (e) 0.0546

3.6.6 1.732; 0.1817

3.6.10 $\frac{1}{4.74}; 3.33$.

3.6.13 (a) $f(y) = e^y[1 + (1/s)e^y]^{-(s+1)}$.

3.7.1 $E(X) = (1-\beta)^{-\alpha}$, if $\beta < 1$.

3.7.2 Download `dloggamma.R`.

Chapter 4

4.1.1 (b) 101.15; (c) 55.5; $\theta \log 2$ (d) 70.11.

4.1.2 (b) 201, 293.9, 17.14, 11.72; (c) 0.269; (d) 0.207

4.1.3 9.5.

4.1.10 (e) 0.65; 0.95.

4.1.11 (e) 0.92; 0.97

4.2.1 $(79.21, 83.19), 90\%$.

4.2.2 $(51.82, 150.48)$

4.2.4 $(6.46, 24.69)$.

4.2.5 $(0.143, 0.365)$.

4.2.6 24 or 25.

4.2.7 $(3.7, 5.7)$.

4.2.8 160.

4.2.9 (a) 1.31σ; (b) 1.49σ.

4.2.10 $c = \sqrt{\frac{n}{n+1}}; k = 1.50.$

4.2.13
```
ind=rep(0,numb);
for(i in 1:numb){if
(ci[i,1]*c[i,2]<0){ind[i]=1}}
```

4.2.14 $\left(\frac{5\overline{x}}{24}, \frac{5\overline{x}}{16}\right).$

4.2.16 6765.

4.2.17 $(3.19, 3.61).$

4.2.18 (b) $(3.625, 29.101).$

4.2.21 $(-3.32, 1.72).$

4.2.26 135 or 136.

4.3.1 (c) $(0.1637, 0.3642).$

4.3.3 $(0.4972, 0.6967).$

4.3.4 (c) $(0.197, 1.05).$

4.4.2 (a) 0.00697; (b) 0.0244; (c) 0.0625

4.4.5 (a) $4, 23, 67, 99, 301.$

4.4.5 $1 - (1 - e^{-3})^4.$

4.4.6 (a) $\frac{1}{8}$.

4.4.10 Weibull.

4.4.11 $\frac{5}{16}$.

4.4.12 pdf: $(2z_1)(4z_2^3)(6z_3^5)$, $0 < z_i < 1$.

4.4.13 $\frac{7}{12}$.

4.4.17 (a) $48y_3^5y_4, 0 < y_3 < y_4 < 1$; (b) $\frac{6y_3^5}{y_4^6}, 0 < y_3 < y_4$; (c) $\frac{6}{7}y_4$.

4.4.18 $\frac{1}{4}$.

4.4.19 $6uv(u + v), 0 < u < v < 1.$

4.4.24 14.

4.4.25 (a) $\frac{15}{16}$; (b) $\frac{675}{1024}$; (c) $(0.8)^4$.

4.4.26 0.824.

4.4.27 8.

4.4.28 (a) 1.13σ; (b) 0.92σ.

4.4.30 $(40, 124), 88\%.$

4.4.32 $(180, 190)$ and $(195, 210)$.

4.5.3 $1 - \left(\frac{3}{4}\right)^\theta + \theta\left(\frac{3}{4}\right)^\theta \log\left(\frac{3}{4}\right), \theta = 1, 2.$

4.5.4 $0.17; 0.78.$

4.5.8 $n = 19$ or 20.

4.5.9 $\gamma\left(\frac{1}{2}\right) = 0.062; \gamma\left(\frac{1}{12}\right) = 0.920.$

4.5.10 $n \approx 73; c \approx 42.$

4.5.12 (a) 0.051; (b) $0.256; 0.547; 0.780.$

4.5.13 (a) 0.154; (b) 0.154.

4.5.14 (1) 0.11514; (2) 0.0633.

4.6.5 (b) $t = -3.0442, p$-value $= 0.0033.$

4.6.6 (b) $t = 2.034$, p-value $= 0.06065.$

4.7.1 p-value $= 0.0184.$

4.7.2 $8.37 > 7.81;$ reject.

4.7.4 $b \leq 8$ or $b \geq 32$.

4.7.5 $2.44 < 11.3$; do not reject H_0.

4.7.6 $6.40 < 9.49$; do not reject H_0.

4.7.7 $\chi^2 = 49.731, p$-value $= 1.573e - 09.$

4.7.8 $k = 3.$

4.8.5 $F^{-1}(u) = \log[u/(1 - u)].$

4.8.7 For $0 < u < (1/2)$:
$F^{-1}(u) = \log[2u].$
For $(1/2) < u < 1$:
$F^{-1}(u) = \log[2(1 - u)].$

4.8.8 $F^{-1}(u) = \log[-\log(1 - u)].$

4.8.18 (a) $F^{-1}(u) = u^{1/\beta}$;
(b) e.g., dominated by a uniform pdf.

4.9.4 (a) $\beta \log 2.$

4.9.8 Use $s_x = 20.41; s_y = 18.59.$

4.9.10 (a) $\overline{y} - \overline{x} = 9.67$;
20 possible permutations;
(c) P_n^n/n^n.

4.9.11 $\mu_0; n^{-1}\sum_{i=1}^n (x_i - \overline{x})^2$.

4.10.1 8.

4.10.4 (a) Beta$(n-j+1, j)$;
(b) Beta$(n-j+i-1, j-i+2)$.

4.10.5 $\frac{10!}{1!3!4!} v_1 v_2^3 (1-v_1-v_2)^4$,
$0 < v_2, v_1 + v_2 < 1$.

Chapter 5

5.1.9 No; $Y_n - \frac{1}{n}$.

5.2.1 Degenerate at μ.

5.2.2 Gamma$(\alpha = 1, \beta = 1)$.

5.2.3 Gamma$(\alpha = 1, \beta = 1)$.

5.2.4 Gamma$(\alpha = 2, \beta = 1)$.

5.2.7 Degenerate at β.

5.2.9 0.682.
```
pchisq(60,50)
 - pchisq(40,50)=.686
```

5.2.10 Download function `cdistplt4`.

5.2.11 (a) `1-pbinom(55,60,.95)=0.820`
(b) 0.815.

5.2.14 Degenerate at $\mu_2 + \frac{\sigma_2}{\sigma_1}(x - \mu_1)$.

5.2.15 (b) $N(0,1)$.

5.2.17 (b) $N(0,1)$.

5.2.20 $\frac{1}{5}$.

5.3.2 0.954.

5.3.3 0.604.

5.3.4 0.840.

5.3.5 0.728.

5.3.7 0.08.

5.3.9 0.267.

Chapter 6

6.1.1 (a) $\hat{\theta} = \overline{X}/4$. (c) 5.03

6.1.2 (a) $-n/\log(\prod_{i=1}^n X_i)$.
(b)$Y_1 = \min\{X_1, \ldots, X_n\}$.

6.1.4 (a) $Y_n = \max\{X_1, \ldots, X_n\}$.
(b) $(2n+1)/(2n)$.
(c) $\sqrt{1/2}Y_n$.

6.1.5 (a) $X = \theta U^{1/2}$, U is unf$(0,1)$.
(b) 7.7, 5.4.

6.1.6 $1 - \exp\{-2/\overline{X}\}$.

6.1.7 $\widehat{p} = \frac{53}{125}$,
$\sum_{x=3}^5 \binom{5}{x}\widehat{p}^x(1-\widehat{p})^{5-x}$., 0.3597.

6.1.8 (b) -0.534.

6.1.9 $\overline{x}^2 e^{-\overline{x}}/2.$, 0.2699.

6.1.10 $\max\left\{\frac{1}{2}, \overline{X}\right\}$.

6.2.7 (a) $\frac{4}{\theta^2}$.
(c) $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{\mathcal{D}} N(0, \theta^2/4)$.
(d) 5.03 ± 0.99.

6.2.8 (a) $\frac{1}{2\theta^2}$.

6.2.13 (b) $\hat{\theta} = 3.547$.
(c) $(2.39, 4.92)$, Yes.

6.2.14 (a) $F(x) = 1 - [\theta^3/(x+\theta)^3]$.
(b)
```
g=function(n,t){u=runif(n)
t*((1-u)^(-1/3)-1)}
```

6.3.1 (b) Test-Stat = 17.28, Reject

6.3.2 $\gamma(\theta) = P[\chi^2(2n) < (\theta_0/\theta)c_1]$
$+P[\chi^2(2n) > (\theta_0/\theta)c_2]$.

6.3.8 Reject if $2\sum_{i=1}^n Y_i < \chi^2_{1-\alpha/2}(2n)$ or
$2\sum_{i=1}^n Y_i > \chi^2_{\alpha/2}(2n)$.

6.3.16 (a) $\left(\frac{1}{3\overline{x}}\right)^{n\overline{x}} \left(\frac{2}{3(1-\overline{x})}\right)^{n-n\overline{x}}$.

6.3.17 (a) $\chi^2_W = \{\sqrt{nI(\overline{X})}(\overline{X} - \theta_0)\}^2$.
(b) Download `waldpois.R`.
(c) $\chi^2_W = 6.90$, $p-$value $= 0.0172$.

6.3.18 $\left(\frac{\overline{x}/\alpha}{\beta_0}\right)^{n\alpha}$
$\times \exp\left\{-\sum_{i=1}^n x_i \left(\frac{1}{\beta_0} - \frac{\alpha}{\overline{x}}\right)\right\}$.

6.4.1 (a) $0.300, 0.225, 0.350, 0.125$.
(b) CI for p_2: $(0.167, 0.283)$

6.4.2 (a) $\overline{x}, \overline{y}$,
$\frac{1}{n+m}\left[\sum_{i=1}^{n}(x_i-\overline{x})^2+\sum_{i=1}^{m}(y_i-\overline{y})^2\right]$.
(b) $\frac{n\overline{x}+m\overline{y}}{n+m}$,
$\left[\sum_{i=1}^{n}(x_i-\widehat{\theta}_1)^2+\sum_{i=1}^{m}(y_i-\widehat{\theta}_1)^2\right]$
$(n+m)^{-1}$.

6.4.3 $\widehat{\theta}_1=\min\{X_i\}, \frac{1}{n}\sum_{i=1}^{n}(X_i-\widehat{\theta}_1)$.

6.4.4 $\widehat{\theta}_1=\min\{X_i\}$,
$n/\log\left[\prod_{i=1}^{n}X_i/\widehat{\theta}_1^n\right]$.

6.4.5 $(Y_1+Y_n)/2, (Y_n-Y_1)/2$; no.

6.4.6 (a) $\overline{X}+1.282\sqrt{\frac{n-1}{n}}S$;
(b) $\Phi\left(\frac{c-\overline{X}}{\sqrt{(n-1)/nS}}\right)$.

6.4.7 (a) mle is 0.7263, $\hat{p}=0.76$
A run of BS: $(0.629, 0.828)$.
Via $\hat{p}$: $(0.642, 0.878)$.

6.4.8 (a) mle is 64.83, $x_{(45)}=64.6$

6.4.9 If $\frac{y_1}{n_1}\le\frac{y_2}{n_2}$, then $\widehat{p}_1=\frac{y_1}{n_1}$ and
$\widehat{p}_2=\frac{y_2}{n_2}$; else, $\widehat{p}_1=\widehat{p}_2=\frac{y_1+y_2}{n_1+n_2}$.

6.5.1 $t=-8.64$, $p-$ value $=0.0001$.

6.5.2 $(81.30004, 81.30156)$.

6.5.3 (b) $(-0.0249, 0.1749)$.

6.5.6 (b) $c\frac{\sum_{i=1}^{n}X_i^2}{\sum_{i=1}^{m}Y_i^2}$.

6.5.7 $F=\frac{\overline{X}}{\overline{Y}}$.

6.5.8 (b) $F=\overline{x}/\overline{y}=0.3389$, Reject.

6.5.9 $c\frac{[\max\{-X_1,X_{n_1}\}]^{n_1}[\max\{-Y_1,Y_{n_2}\}]^{n_2}}{[\max\{-X_1,-Y_1,X_{n_1},Y_{n_2}\}]^{n_1+n_2}}$,
$\chi^2(2)$.

6.6.8 The R function `mixnormal`, at site listed in the Preface produced these results:
(first row are initial estimates, second row are the estimates after 500 iterations):

μ_1	μ_2	σ_1	σ_2	π
105.00	130.00	15.00	25.00	0.600
98.76	133.96	9.88	21.50	0.704

Chapter 7

7.1.4 $\frac{1}{3}, \frac{2}{3}$.

7.1.5 $\delta_1(y)$.

7.1.6 $b=0$, does not exist.

7.1.7 does not exist.

7.2.8 $\prod_{i=1}^{n}[X_i(1-X_i)]$.

7.2.9 (a) $\frac{n!\theta^{-r}}{(n-r)!}e^{-\frac{1}{\theta}\left[\sum_{i=1}^{r}y_i+(n-r)y_r\right]}$.
(b) $r^{-1}[\sum_{i=1}^{r}y_i+(n-r)y_r]$.

7.3.2 $60y_3^2(y_5-y_3)/\theta^5$;
$0<y_3<y_5<\theta$;
$6y_5/5; \theta^2/7; \theta^2/35$.

7.3.3 $\frac{1}{\theta^2}e^{-y_1/\theta}, 0<y_2<y_1<\infty$;
$y_1/2; \theta^2/2$.

7.3.5 $n^{-1}\sum_{i=1}^{n}X_i^2; n^{-1}\sum_{i=1}^{n}X_i$;
$(n+1)Y_n/n$.

7.3.6 $6\overline{X}$.

7.4.2 (a) X ; (b) X

7.4.3 Y/n.

7.4.5 $Y_1-\frac{1}{n}$.

7.4.7 (a) Yes; (b) yes.

7.4.8 (a) $E(X)=0$.

7.4.9 (a) $\max\{-Y_1, 0.5Y_n\}$; (b) yes;
(c) yes.

7.5.1 $Y_1=\sum_{i=1}^{n}X_i; Y_1/4n$; yes.

7.5.4 $\overline{x}/\alpha$.

7.5.9 $\overline{x}$.

7.5.11 (b) Y_1/n; (c) θ; (d) Y_1/n.

7.6.1 $\overline{X}^2-\frac{1}{n}$.

7.6.2 $Y^2/(n^2+2n)$.

7.6.3 (a) 0.8413; (b) 0.7702 (c) Our run 0.0584.

7.6.4 (a) 49.4; (b) Our run: 4.405

7.6.6 (a) $\left(\frac{n-1}{n}\right)^Y \left(1+\frac{Y}{n-1}\right)$;
(b)$\left(\frac{n-1}{n}\right)^{n\overline{X}} \left(1+\frac{n\overline{X}}{n-1}\right)$;
(c) $N\left(\theta, \frac{\theta}{n}\right)$.

7.6.9 $1-e^{-2/\overline{X}}; 1-\left(1-\frac{2/\overline{X}}{n}\right)^{n-1}$.

7.6.10 (b) $\overline{X}$; (c) $\overline{X}$; (d) $1/\overline{X}$.

7.7.3 Yes.

7.7.5 (a) $\frac{\Gamma[(n-1)/2]}{\Gamma[n/2]}\sqrt{\frac{n-1}{2}}S$.
(b) Download `bootse6.R`
10.1837; Our run: 1.156828

7.7.6 (b) $\frac{Y_1+Y_n}{2}; \frac{(n+1)(Y_n-Y_1)}{2(n-1)}$.

7.7.7 (a) $K = (\Gamma((n-1)/2)/\Gamma(n/2)) \times\sqrt{((n-1)/2)}$
mvue $= \Phi^{-1}(p)KS+\overline{x}$
(c) 59.727; Our run 3.291479.

7.7.9 (a) $\frac{1}{n-1}\sum_{h=1}^n(X_{ih}-\overline{X}_i) \times (X_{jh}-\overline{X}_j)$;
(b) $\sum_{i=1}^n a_i\overline{X}_i$.

7.7.10 $\left(\sum_{i=1}^n x_i, \sum_{i=1}^n \frac{1}{x_i}\right)$.

7.8.3 $Y_1, ; \sum_{i=1}^n (Y_i - Y_1)/n$.

7.9.13 (a) $\Gamma(3n, 1/\theta)$, no;
(c) $(3n-1)/Y$;
(e) Beta$(3, 3n-3)$.

Chapter 8

8.1.4 $\sum_{i=1}^{10} x_i^2 \geq 18.3$; yes; yes.

8.1.5 $\prod_{i=1}^n x_i \geq c$.

8.1.6 $3\sum_{i=1}^{10} x_i^2 + 2\sum_{i=1}^{10} x_i \geq c$.

8.1.7 About 96; 76.7.

8.1.8 $\prod_{i=1}^n [x_i(1-x_i)] \geq c$.

8.1.9 About 39; 15.

8.1.10 0.08; 0.875.

8.2.1 $(1-\theta)^9(1+9\theta)$.

8.2.2 $1-\frac{15}{16\theta^4}, 1<\theta$.

8.2.3 $1-\Phi\left(\frac{3-5\theta}{2}\right)$.

8.2.4 About 54; 5.6.

8.2.7 Reject H_0 if $\overline{x} \geq 77.564$.

8.2.8 About 27; reject H_0 if $\overline{x} \leq 24$.

8.2.10 $\Gamma(n, \theta)$;
Reject H_0 if $\sum_{i=1}^n x_i \geq c$.

8.2.12 (b)$\frac{6}{32}$; (c) $\frac{1}{32}$.
(d) reject if $y=0$;
if $y=1$, reject with probability $\frac{1}{5}$.

8.3.1 (b) $t=-2.2854, p=0.02393$;
(c) $(-0.5396 - 0.0388)$.

8.3.5 (d) $n=90$.

8.3.6 78; 0.7608.

8.3.10 Under H_1, $(\theta_4/\theta_3)F$ has an $F(n-1, m-1)$ distribution.

8.3.12 Reject H_0 if $|y_3-\theta_0| \geq c$.

8.3.14 (a) $\prod_{i=1}^n (1-x_i) \geq c$.

8.3.17 (b) $F=1.34$; $p=0.088$.

8.4.1 $5.84n - 32.42; 5.84n+41.62$.

8.4.2 $0.04n - 1.66; 0.04n + 1.20$.

8.4.4 $0.025, 29.7, -29.7$.

8.5.5 $(9y-20x)/30 \leq c \Rightarrow (x,y) \in$ 2nd.

8.5.7 $2w_1^2 + 8w_2^2 \geq c \Rightarrow (w_1, w_2) \in$ II.

Chapter 9

9.2.3 6.39.

9.2.6 (b) $F=1.1433$, $p=0.3451$.

9.2.7 $7.875 > 4.26$; reject H_0.

9.2.8 $10.224 > 4.26$; reject H_0.

9.3.2 $2r + 4\theta$.

9.3.3 (a) $5m/3$; (b) 0.6174; 0.9421; (c) 7

9.4.1 None. For B − C: $(-0.199, 10.252)$.

9.4.2 No significant differences.

9.4.3 (a) CI's of form: (4.2.14) using α/k.

9.4.4 (a)$(-0.103, 0.0214)$ (b) $\chi^2 = 24.4309$, $p = 0.00367$, $(-0.103, 0.021)$.

9.5.6 $7.00; 9.98$.

9.5.8 $4.79; 22.82; 30.73$.

9.5.10 (a) $7.624 > 4.46$, reject H_A; (b) $15.538 > 3.84$, reject H_B.

9.5.11 $8; 0; 0; 0; 0; -3; 1; 2; -2; 2; -2; 2; 2; -2; 2; -2; 0; 0; 0; 0$.

9.6.1 $N(\alpha^*, \sigma^2(n^{-1} + \overline{x}^2/\sum(x_i - \overline{x})^2))$.

9.6.2 (a) $6.478{+}4.483x$; (d) $(-0.026, 8.992)$.

9.6.3 (a) $-983.8868 + 0.5041x$.

9.6.8 PI: $(3.27, 3.70)$

9.6.10 $\hat{\beta} = n^{-1}\sum_i Y_i/x_i$; $\hat{\gamma} = n^{-1}\sum_i[(Y_i/x_i) - n^{-1}\sum_j(Y_j/x_j)]^2$.

9.6.14 $\widehat{a} = \frac{5}{3}$.

9.7.2 Reject H_0.

9.7.6 Lower Bound: $\tanh\left[\frac{1}{2}\log\frac{1+r}{1-r} - \frac{z_{\alpha/2}}{\sqrt{n-3}}\right]$.

9.7.7 (a) $0.710, (0.555, 0.818)$; (b) Pitchers: $0.536, (0.187, 0.764)$.

9.8.2 $2; \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}; \mu_1 = \mu_2 = 0$.

9.8.3 (b) $\mathbf{A}^2 = \mathbf{A}; \mathrm{tr}(\mathbf{A}) = 2$; $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/8 = 6$.

9.8.4 (a) $\sum \sigma_i^2/n^2$.

9.8.5 (a) $[1 + (n-1)\rho](\sigma^2/n)$.

9.9.1 Dependent.

9.9.3 $0, 0, 0, 0$.

9.9.4 $\sum_{i=1}^n a_{ij} = 0$.

Chapter 10

10.2.3 (a) 0.1148; (b) 0.7836.

10.2.4 (a) 425; $(380, 500)$; (b) 591.18; $(508.96, 673.41)$.

10.2.9 (a) $P(Z > z_\alpha - (\sigma/\sqrt{n})\theta)$, where $E(Z) = 0$ and $\mathrm{Var}(Z) = 1$; (c) Use the Central Limit Theorem; (d) $\left[\frac{(z_\alpha - z_{\gamma^*})\sigma}{\theta^*}\right]^2$.

10.4.2 $1 - \Phi[z_\alpha - \sqrt{\lambda_1\lambda_2}(\delta/\sigma)]$.

10.4.3 Conf.Int for MWW: $(0.0483, 00571)$.

10.4.4 Our run: $n_1 = n_2 = 39$ yielded 0.8025 power.

10.3.4 (a) $T^+ = 174$, p-value $= 0.0083$. (b) $t = 3.0442$, p-value $= 0.0067$.

10.5.3 $\frac{n(n-1)}{n+1}$.

10.7.1 (b) $(0.156, 0.695)$.

10.5.14 (a) $W_S^* = 9; W_{XS}^* = 6$; (b) 1.2; (c) 9.5.

10.8.3 $\widehat{y}_{\mathrm{LS}} = 205.9 + 0.015x$; $\widehat{y}_W = 211.0 + 0.010x$.

10.8.4 (a) $\widehat{y}_{\mathrm{LS}} = 265.7 - 0.765(x - 1900)$; $\widehat{y}_W = 246.9 - 0.436(x - 1900)$; (b) $\widehat{y}_{\mathrm{LS}} = 3501.0 - 38.35(x - 1900)$; $\widehat{y}_W = 3297.0 - 35.52(x - 1900)$.

10.8.9 $r_{qc} = 16/17 = 0.941$ (zeroes were excluded).

10.8.10 $r_N = 0.835; z = 3.734$.

10.9.4 Cases: $t < y$ and $t > y$.

10.9.5 (c) $y^2 - \sigma^2$.

10.9.7 (a)$n^{-1}\sum_{i=1}^n(Y_i - \overline{Y})^2$; (c) $y^2 - \sigma^2$.

10.9.9 $0; [4f^2(\theta)]^{-1}$.

10.9.14 $\widehat{y}_{\mathrm{LS}} = 3.14 + .028x$;
$\widehat{y}_W = 0.214 + .020x$.

Chapter 11

11.1.1 $0.45; 0.55$.

11.1.3 $[y\tau^2 + \mu\sigma^2/n]/(\tau^2 + \sigma^2/n)$.

11.1.4 $\beta(y + \alpha)/(n\beta + 1)$.

11.1.6 $\frac{y_1+\alpha_1}{n+\alpha_1+\alpha_2+\alpha_3}; \frac{y_2+\alpha_2}{n+\alpha_1+\alpha_2+\alpha_3}$.

11.1.8 (a) $\left(\theta - \frac{10+30\theta}{45}\right)^2 + \left(\frac{1}{45}\right)^2 30\theta(1-\theta)$.

11.1.9 $\sqrt[6]{2}, y_4 < 1; \sqrt[6]{2}y_4, 1 \le y_4$.

11.2.1 (a) $\frac{\theta_2^2}{[\theta_2^2+(x_1-\theta_1)^2][\theta_2^2+(x_2-\theta_1)^2]}$.

11.2.3 (a) 76.84; (b) $(76.25, 77.43)$.

11.2.5 (a) $I(\theta) = \theta^{-2}$; (d) $\chi^2(2n)$.

11.2.8 (a) $\mathrm{beta}(n\overline{x} + 1, n + 1 - n\overline{x})$.

11.3.1 (a) Let U_1 and U_2 be iid uniform(0,1):

1. Draw $Y = -\log(1 - U_1)$
2. Draw $X = Y - \log(1 - U_2)$.

11.3.3 (b) $F_X^{-1}(u) = -\log(1 - \sqrt{u})$, $0 < u < 1$.

11.3.7 (b) $f(x|y)$ is a $b(n, y)$ pmf; $f(y|x)$ is a $\mathrm{beta}(x + \alpha, n - x + \beta)$ pdf.

11.4.1 (b) $\widehat{\beta} = \frac{1}{2\overline{x}}$; (d) $\widehat{\theta} = \frac{1}{\overline{x}}$.

11.4.2 (a) $\delta(y) = \frac{\int_0^1 \left[\frac{a}{1-a\log p}\right]^2 p^y(1-p)^{n-y}\,dp}{\int_0^1 \left[\frac{a}{1-a\log p}\right]^2 p^{y-1}(1-p)^{n-y}\,dp}$.

This page intentionally left blank

Index