Chapter 6

Order Statistics and Quantiles

6.1 Extreme Order Statistics

Suppose we have a finite sample $X_1, \ldots, X_n$. Conditional on this sample, we define the values $X_{(1)}, \ldots, X_{(n)}$ to be a permutation of $X_1, \ldots, X_n$ such that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$. We call $X_{(i)}$ the ith order statistic of the sample. In much of this chapter, we will be concerned with the marginal distributions of these order statistics.

Even though the notation $X_{(i)}$ does not explicitly use the sample size n, the distribution of $X_{(i)}$ depends essentially on n. For this reason, some textbooks use slightly more complicated notation such as

$$X_{(1:n)}, X_{(2;n)}, \ldots, X_{(n;n)}$$

for the order statistics of a sample. We choose to use the simpler notation here, though we will always understand the sample size to be n.

The asymptotic distributions of order statistics at the extremes of a sample may be derived without any specialized knowledge other than the limit formula

$$\left(1 + \frac{c}{n}\right)^n \to e^c \quad \text{as } n \to \infty \tag{6.1}$$

and its generalization

$$\left(1 + \frac{c_n}{b_n}\right)^{b_n} \to e^c \quad \text{if } c_n \to c \text{ and } b_n \to \infty \tag{6.2}$$

[see Exercise 1.7(b)]. Recall that by "asymptotic distribution of $X_{(1)}$," we mean sequences k_n and a_n, along with a nondegenerate random variable X, such that $k_n(X_{(1)} - a_n) \xrightarrow{d} X$. This section consists mostly of a series of illustrative examples.

Example 6.1 Suppose $X_1, \ldots, X_n$ are independent and identically distributed uniform(0,1) random variables. What is the asymptotic distribution of $X_{(n)}$?

Since $X_{(n)} \le t$ if and only if $X_1 \le t$, $X_2 \le t$, $\ldots$, and $X_n \le t$, by independence we have

$$P(X_{(n)} \le t) = \begin{cases} 0 & \text{if } t \le 0 \\ t^n & \text{if } 0 < t < 1 \\ 1 & \text{if } t \ge 1. \end{cases} \tag{6.3}$$

From Equation (6.3), it is apparent that $X_{(n)} \xrightarrow{P} 1$, though this limit statement does not fully reveal the asymptotic distribution of $X_{(n)}$. We desire sequences k_n and a_n such that $k_n(X_{(n)} - a_n)$ has a nondegenerate limiting distribution. Evidently, we should expect $a_n = 1$, a fact we shall rederive below.

Computing the distribution function of $k_n(X_{(n)} - a_n)$ directly, we find

$$F(u) = P\{k_n(X_{(n)} - a_n) \le u\} = P\left\{X_{(n)} \le \frac{u}{k_n} + a_n\right\}$$

as long as $k_n > 0$. Therefore, we see that

$$F(u) = \left(\frac{u}{k_n} + a_n\right)^n \quad \text{for} \;\; 0 < \frac{u}{k_n} + a_n < 1. \tag{6.4}$$

We would like this expression to tend to a limit involving only u as $n \to \infty$. Keeping expression 6.2 in mind, we take $a_n = 1$ and $k_n = n$ so that $F(u) = (1 + u/n)^n$, which tends to e^u.

However, we are not quite finished, since we have not determined which values of u make the above limit valid. Equation 6.4 required that $0 < a_n + (u/k_n) < 1$, which in this case becomes $-1 < u/n < 0$. This means u may be any negative real number, since for any $u < 0$, $-1 < u/n < 0$ for all $n > |u|$. We conclude that if the random variable U has distribution function

$$F(u) = \begin{cases} \exp(u) & \text{if } u \le 0 \\ 1 & \text{if } u > 0, \end{cases}$$

then $n(X_{(n)} - 1) \xrightarrow{d} U$. Since $-U$ is simply a standard exponential random variable, we may also write

$$n(1 - X_{(n)}) \xrightarrow{d} \text{Exponential}(1).$$

Example 6.2 Suppose $X_1, X_2, \ldots$ are independent and identically distributed exponential random variables with mean 1. What is the asymptotic distribution of $X_{(n)}$?

As in Equation 6.3, if $u/k_n + a_n > 0$ then

$$P\left\{k_n(X_{(n)} - a_n) \le u\right\} = P\left(X_{(n)} \le \frac{u}{k_n} + a_n\right) = \left\{1 - \exp\left(-a_n - \frac{u}{k_n}\right)\right\}^n.$$

Taking $a_n = \log n$ and $k_n = 1$, the rightmost expression above simplifies to

$$\left\{1 - \frac{e^{-u}}{n}\right\}^n,$$

which has limit $\exp(-e^{-u})$. The condition $u/k_n + a_n > 0$ becomes $u + \log n > 0$, which is true for all $u \in \mathbb{R}$ as long as $n > \exp(-u)$. Therefore, we conclude that $X_{(n)} - \log n \xrightarrow{d} U$, where

$$P(U \le u) \stackrel{\text{def}}{=} \exp\{-\exp(-u)\} \text{ for all } u. \tag{6.5}$$

The distribution of U in Equation 6.5 is known as the extreme value distribution or the Gumbel distribution.

In Examples 6.1 and 6.2, we derived the asymptotic distribution of a maximum from a simple random sample. We did this using only the definition of convergence in distribution without relying on any results other than expression 6.2. In a similar way, we may derive the joint asymptotic distribution of multiple order statistics, as in the following example.

Example 6.3 *Range of uniform sample:* Let $X_1, \ldots, X_n$ be a simple random sample from Uniform(0, 1). Let $R_n = X_{(n)} - X_{(1)}$ denote the range of the sample. What is the asymptotic distribution of R_n?

To answer this question, we begin by finding the joint asymptotic distribution of $(X_{(n)}, X_{(1)})$, as follows. For sequences k_n and ℓ_n, as yet unspecified, consider

$$\begin{aligned} P(k_n X_{(1)} > x \text{ and } \ell_n(1 - X_{(n)}) > y) &= P(X_{(1)} > x/k_n \text{ and } X_{(n)} < 1 - y/\ell_n) \\ &= P(x/k_n < X_{(1)} < \cdots < X_{(n)} < 1 - y/\ell_n), \end{aligned}$$

where we have assumed that k_n and ℓ_n are positive. Since the probability above is simply the probability that the entire sample is to be found in the interval $(x/k_n, 1 - y/\ell_n)$, we conclude that as long as

$$0 < \frac{x}{k_n} < 1 - \frac{y}{\ell_n} < 1, \tag{6.6}$$

we have

$$P(k_n X_{(1)} > x \text{ and } \ell_n(1 - X_{(n)}) > y) = \left(1 - \frac{y}{\ell_n} - \frac{x}{k_n}\right)^n.$$

Expression (6.1) suggests that we set $k_n = \ell_n = n$, resulting in

$$P(nX_{(1)} > x \text{ and } n(1 - X_{(n)}) > y) \quad = \quad \left(1 - \frac{y}{n} - \frac{x}{n}\right)^n.$$

Expression 6.6 becomes

$$0 < \frac{x}{n} < 1 - \frac{y}{n} < 1,$$

which is satisfied for large enough n if and only if x and y are both positive. We conclude that for $x > 0$, $y > 0$,

$$P(nX_{(1)} > x \text{ and } n(1 - X_{(n)}) > y) \to e^{-x}e^{-y}.$$

Since this is the joint distribution of independent standard exponential random variables, say, Y_1 and Y_2, we conclude that

$$\begin{pmatrix} nX_{(1)} \\ n(1 - X_{(n)}) \end{pmatrix} \stackrel{d}{\to} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}.$$

Therefore, applying the continuous function $f(a, b) = a + b$ to both sides gives

$$n(1 - X_{(n)} + X_{(1)}) = n(1 - R_n) \stackrel{d}{\to} Y_1 + Y_2 \sim \text{Gamma}(2, 1).$$

Let us consider a different example in which the asymptotic joint distribution does not involve independent random variables.

Example 6.4 As in Example 6.3, let $X_1, \ldots, X_n$ be independent and identically distributed from uniform$(0, 1)$. What is the joint asymptotic distribution of $X_{(n-1)}$ and $X_{(n)}$?

Proceeding as in Example 6.3, we obtain

$$P\left[\begin{pmatrix} n(1 - X_{(n-1)}) \\ n(1 - X_{(n)}) \end{pmatrix} > \begin{pmatrix} x \\ y \end{pmatrix}\right] = P\left(X_{(n-1)} < 1 - \frac{x}{n} \text{ and } X_{(n)} < 1 - \frac{y}{n}\right). \quad (6.7)$$

We consider two separate cases: If $0 < x < y$, then the right hand side of (6.7) is simply $P(X_{(n)} < 1 - y/n)$, which converges to e^{-y} as in Example 6.1. On the other hand, if $0 < y < x$, then

$$\begin{aligned} P\left(X_{(n-1)} < 1 - \frac{x}{n} \text{ and } X_{(n)} < 1 - \frac{y}{n}\right) &= P\left(X_{(n)} < 1 - \frac{x}{n}\right) \\ &\quad + P\left(X_{(n-1)} < 1 - \frac{x}{n} < X_{(n)} < 1 - \frac{y}{n}\right) \\ &= \left(1 - \frac{x}{n}\right)^n + n\left(1 - \frac{x}{n}\right)^{n-1}\left(\frac{x}{n} - \frac{y}{n}\right) \\ &\to e^{-x}(1 + x - y). \end{aligned}$$

The second equality above arises because $X_{(n-1)} < a < X_{(n)} < b$ if and only if exactly $n-1$ of the X_i are less than a and exactly one is between a and b. We now know the joint asymptotic distribution of $n(1 - X_{(n-1)})$ and $n(1 - X_{(n-1)})$; but can we describe this joint distribution in a simple way? Suppose that Y_1 and Y_2 are independent standard exponential variables. Consider the joint distribution of Y_1 and $Y_1 + Y_2$: If $0 < x < y$, then

$$P(Y_1 + Y_2 > x \text{ and } Y_1 > y) = P(Y_1 > y) = e^{-y}.$$

On the other hand, if $0 < y < x$, then

$$\begin{aligned} P(Y_1 + Y_2 > x \text{ and } Y_1 > y) &= P(Y_1 > \max\{y, x - Y_2\}) = \mathrm{E}\, e^{-\max\{y, x-Y_2\}} \\ &= e^{-y} P(y > x - Y_2) + \int_0^{x-y} e^{t-x} e^{-t}\, dt \\ &= e^{-x}(1 + x - y). \end{aligned}$$

Therefore, we conclude that

$$\begin{pmatrix} n(1 - X_{(n-1)}) \\ n(1 - X_{(n)}) \end{pmatrix} \stackrel{d}{\to} \begin{pmatrix} Y_1 + Y_2 \\ Y_1 \end{pmatrix}.$$

Notice that marginally, we have shown that $n(1 - X_{(n-1)}) \stackrel{d}{\to}$ Gamma(2, 1).

Recall that if F is a continuous, invertible distribution function and U is a standard uniform random variable, then $F^{-1}(U) \sim F$. The proof is immediate, since $P\{F^{-1}(U) \le t\} = P\{U \le F(t)\} = F(t)$. We may use this fact in conjunction with the result of Example 6.4 as in the following example.

Example 6.5 Suppose $X_1, \ldots, X_n$ are independent standard exponential random variables. What is the joint asymptotic distribution of $(X_{(n-1)}, X_{(n)})$?

The distribution function of a standard exponential distribution is $F(t) = 1 - e^{-t}$, whose inverse is $F^{-1}(u) = -\log(1 - u)$. Therefore,

$$\begin{pmatrix} -\log(1 - U_{(n-1)}) \\ -\log(1 - U_{(n)}) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} X_{(n-1)} \\ X_{(n)} \end{pmatrix},$$

where $\stackrel{d}{=}$ means "has the same distribution". Thus,

$$\begin{pmatrix} -\log\left[n(1 - U_{(n-1)})\right] \\ -\log\left[n(1 - U_{(n)})\right] \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} X_{(n-1)} - \log n \\ X_{(n)} - \log n \end{pmatrix}.$$

We conclude by the result of Example 6.4 that

$$\begin{pmatrix} X_{(n-1)} - \log n \\ X_{(n)} - \log n \end{pmatrix} \stackrel{d}{\to} \begin{pmatrix} -\log(Y_1 + Y_2) \\ -\log Y_1 \end{pmatrix},$$

where Y_1 and Y_2 are independent standard exponential variables.

Exercises for Section 6.1

Exercise 6.1 For a given n, let $X_1, \ldots, X_n$ be independent and identically distributed with distribution function

$$P(X_i \le t) = \frac{t^3 + \theta^3}{2\theta^3} \qquad \text{for } t \in [-\theta, \theta].$$

Let $X_{(1)}$ denote the first order statistic from the sample of size n; that is, $X_{(1)}$ is the smallest of the X_i.

(a) Prove that $-X_{(1)}$ is consistent for θ.

(b) Prove that

$$n(\theta + X_{(1)}) \stackrel{d}{\to} Y,$$

where Y is a random variable with an exponential distribution. Find $\mathrm{E}\,(Y)$ in terms of θ.

(c) For a fixed α, define

$$\delta_{\alpha,n} = -\left(1 + \frac{\alpha}{n}\right) X_{(1)}.$$

Find, with proof, α^* such that

$$n\,(\theta - \delta_{\alpha^*,n}) \stackrel{d}{\to} Y - \mathrm{E}\,(Y),$$

where Y is the same random variable as in part (b).

(d) Compare the two consistent θ-estimators $\delta_{\alpha^*,n}$ and $-X_{(1)}$ empirically as follows. For $n \in \{10^2, 10^3, 10^4\}$, take $\theta = 1$ and simulate 1000 samples of size n from the distribution of X_i. From these 1000 samples, estimate the bias and mean squared error of each estimator. Which of the two appears better? Do your empirical results agree with the theoretical results in parts (c) and (d)?

Exercise 6.2 Let $X_1, X_2, \ldots$ be independent uniform $(0, \theta)$ random variables. Let $X_{(n)} = \max\{X_1, \ldots, X_n\}$ and consider the three estimators

$$\delta_n^0 = X_{(n)} \qquad \delta_n^1 = \frac{n}{n-1} X_{(n)} \qquad \delta_n^2 = \left(\frac{n}{n-1}\right)^2 X_{(n)}.$$

(a) Prove that each estimator is consistent for θ.

(b) Perform an empirical comparison of these three estimators for $n = 10^2, 10^3, 10^4$. Use $\theta = 1$ and simulate 1000 samples of size n from uniform $(0, 1)$. From these 1000 samples, estimate the bias and mean squared error of each estimator. Which one of the three appears to be best?

(c) Find the asymptotic distribution of $n(\theta - \delta_n^i)$ for $i = 0, 1, 2$. Based on your results, which of the three appears to be the best estimator and why? (For the latter question, don't attempt to make a rigorous mathematical argument; simply give an educated guess.)

Exercise 6.3 Find, with proof, the asymptotic distribution of $X_{(n)}$ if $X_1, \ldots, X_n$ are independent and identically distributed with each of the following distributions. (That is, find k_n, a_n, and a nondegenerate random variable X such that $k_n(X_{(n)} - a_n) \stackrel{d}{\to} X$.)

(a) Beta(3, 1) with distribution function $F(x) = x^2$ for $x \in (0, 1)$.

(b) Standard logistic with distribution function $F(x) = e^x/(1 + e^x)$.

Exercise 6.4 Let $X_1, \ldots, X_n$ be independent uniform$(0, 1)$ random variables. Find the joint asymptotic distribution of $\left[nX_{(2)}, n(1 - X_{(n-1)})\right]$.

Hint: To find a probability such as $P(a < X_{(2)} < X_{(n)} < b)$, consider the trinomial distribution with parameters $[n; (a, b - a, 1 - b)]$ and note that the probability in question is the same as the probability that the numbers in the first and third categories are each ≤ 1.

Exercise 6.5 Let $X_1, \ldots, X_n$ be a simple random sample from the distribution function $F(x) = [1 - (1/x)]I\{x > 1\}$.

(a) Find the joint asymptotic distribution of $(X_{(n-1)}/n, X_{(n)}/n)$.

Hint: Proceed as in Example 6.5.

(b) Find the asymptotic distribution of $X_{(n-1)}/X_{(n)}$.

Exercise 6.6 If $X_1, \ldots, X_n$ are independent and identically distributed uniform$(0, 1)$ variables, prove that $X_{(1)}/X_{(2)} \stackrel{d}{\to}$ uniform$(0, 1)$.

Exercise 6.7 Let $X_1, \ldots, X_n$ be independent uniform$(0, 2\theta)$ random variables.

(a) Let $M = (X_{(1)} + X_{(n)})/2$. Find the asymptotic distribution of $n(M - \theta)$.

(b) Compare the asymptotic performance of the three estimators M, $\overline{X}_n$, and the sample median $\tilde{X}_n$ by considering their relative efficiencies.

(c) For $n \in \{101, 1001, 10001\}$, generate 500 samples of size n, taking $\theta = 1$. Keep track of M, $\overline{X}_n$, and $\tilde{X}_n$ for each sample. Construct a 3×3 table in which you report the sample variance of each estimator for each value of n. Do your simulation results agree with your theoretical results in part (b)?

Exercise 6.8 Let $X_1, \ldots, X_n$ be a simple random sample from a logistic distribution with distribution function $F(t) = e^{t/\theta}/(1 + e^{t/\theta})$ for all t.

(a) Find the asymptotic distribution of $X_{(n)} - X_{(n-1)}$.

Hint: Use the fact that $\log U_{(n)}$ and $\log U_{(n-1)}$ both converge in probability to zero.

(b) Based on part (a), construct an approximate 95% confidence interval for θ. Use the fact that the .025 and .975 quantiles of the standard exponential distribution are 0.0253 and 3.6889, respectively.

(c) Simulate 1000 samples of size $n = 40$ with $\theta = 2$. How many confidence intervals contain θ?

6.2 Sample Quantiles

To derive the distribution of sample quantiles, we begin by obtaining the exact distribution of the order statistics of a random sample from a uniform distribution. To facilitate this derivation, we begin with a quick review of changing variables. Suppose $\mathbf{X}$ has density $f_{\mathbf{X}}(\mathbf{x})$ and $\mathbf{Y} = \mathbf{g}(\mathbf{X})$, where $\mathbf{g} : \mathbb{R}^k \to \mathbb{R}^k$ is differentiable and has a well-defined inverse, which we denote by $\mathbf{h} : \mathbb{R}^k \to \mathbb{R}^k$. (In particular, we have $\mathbf{X} = \mathbf{h}[\mathbf{Y}]$.) The density for $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = |\text{Det}[\nabla \mathbf{h}(\mathbf{y})]| \, f_{\mathbf{X}}[\mathbf{h}(\mathbf{y})], \tag{6.8}$$

where $|\text{Det}[\nabla \mathbf{h}(\mathbf{y})]|$ is the absolute value of the determinant of the $k \times k$ matrix $\nabla \mathbf{h}(\mathbf{y})$.

6.2.1 Uniform Order Statistics

We now show that the order statistics of a uniform distribution may be obtained using ratios of gamma random variables. Suppose $X_1, \ldots, X_{n+1}$ are independent standard exponential, or Gamma(1, 1), random variables. For $j = 1, \ldots, n$, define

$$Y_j = \frac{\sum_{i=1}^{j} X_i}{\sum_{i=1}^{n+1} X_i}. \tag{6.9}$$

We will show that the joint distribution of $(Y_1, \ldots, Y_n)$ is the same as the joint distribution of the order statistics $(U_{(1)}, \ldots, U_{(n)})$ of a simple random sample from uniform$(0, 1)$ by demonstrating that their joint density function is the same as that of the uniform order statistics, namely $n! I\{0 < u_{(1)} < \ldots < u_{(n)} < 1\}$.

We derive the joint density of $(Y_1, \ldots, Y_n)$ as follows. As an intermediate step, define $Z_j = \sum_{i=1}^{j} X_i$ for $j = 1, \ldots, n+1$. Then

$$X_i = \begin{cases} Z_i & \text{if } i = 1 \\ Z_i - Z_{i-1} & \text{if } i > 1, \end{cases}$$

which means that the gradient of the transformation from $\mathbf{Z}$ to $\mathbf{X}$ is upper triangular with ones on the diagonal, a matrix whose determinant is one. This implies that the density for $\mathbf{Z}$ is

$$f_{\mathbf{Z}}(\mathbf{z}) = \exp\{-z_{n+1}\} I\{0 < z_1 < z_2 < \cdots < z_{n+1}\}.$$

Next, if we define $Y_{n+1} = Z_{n+1}$, then we may express $\mathbf{Z}$ in terms of $\mathbf{Y}$ as

$$Z_i = \begin{cases} Y_{n+1} Y_i & \text{if } i < n+1 \\ Y_{n+1} & \text{if } i = n+1. \end{cases} \tag{6.10}$$

The gradient of the transformation in Equation (6.10) is lower triangular, with y_{n+1} along the diagonal except for a 1 in the lower right corner. The determinant of this matrix is y_{n+1}^n, so the density of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = y_{n+1}^n \exp\{-y_{n+1}\} I\{y_{n+1} > 0\} I\{0 < y_1 < \cdots < y_n < 1\}. \tag{6.11}$$

Equation (6.11) reveals several things: First, $(Y_1, \ldots, Y_n)$ is independent of Y_{n+1} and the marginal distribution of Y_{n+1} is Gamma$(n+1, 1)$. More important for our purposes, the marginal joint density of $(Y_1, \ldots, Y_n)$ is proportional to $I\{0 < y_1 < \cdots < y_n < 1\}$, which is exactly what we needed to prove. We conclude that the vector $\mathbf{Y}$ defined in Equation (6.9) has the same distribution as the vector of order statistics of a simple random sample from uniform$(0, 1)$.

6.2.2 Uniform Sample Quantiles

Using the result of Section 6.2.1, we may derive the joint asymptotic distribution of a set of sample quantiles for a uniform simple random sample.

Suppose we are interested in the p_1 and p_2 quantiles, where $0 < p_1 < p_2 < 1$. The following argument may be generalized to obtain the joint asymptotic distribution of any finite number of quantiles. If $U_1, \ldots, U_n$ are independent uniform$(0,1)$ random variables, then the p_1 and p_2 sample quantiles may be taken to be the a_nth and b_nth order statistics, respectively, where $a_n \stackrel{\text{def}}{=} \lfloor .5+np_1 \rfloor$ and $b_n \stackrel{\text{def}}{=} \lfloor .5+np_2 \rfloor$ ($\lfloor .5+x \rfloor$ is simply x rounded to the nearest integer).

Next, let

$$A_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{a_n} X_i, \quad B_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=a_n+1}^{b_n} X_i, \quad \text{and} \quad C_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=b_n+1}^{n+1} X_i.$$

We proved in Section 6.2.1 that $(U_{(a_n)}, U_{(b_n)})$ has the same distribution as

$$\mathbf{g}(A_n, B_n, C_n) \stackrel{\text{def}}{=} \left(\frac{A_n}{A_n+B_n+C_n}, \frac{A_n+B_n}{A_n+B_n+C_n} \right). \tag{6.12}$$

The asymptotic distribution of $\mathbf{g}(A_n, B_n, C_n)$ may be determined using the delta method if we can determine the joint asymptotic distribution of (A_n, B_n, C_n).

A bit of algebra reveals that

$$\sqrt{n}\,(A_n - p_1) \;=\; \sqrt{\frac{a_n}{n}}\sqrt{a_n}\left(\frac{nA_n}{a_n} - \frac{np_1}{a_n}\right) \;=\; \sqrt{\frac{a_n}{n}}\sqrt{a_n}\left(\frac{nA_n}{a_n} - 1\right) + \frac{a_n - np_1}{\sqrt{n}}.$$

By the central limit theorem, $\sqrt{a_n}(nA_n/a_n - 1) \stackrel{d}{\to} N(0,1)$ because the X_i have mean 1 and variance 1. Furthermore, $a_n/n \to p_1$ and the rightmost term above goes to 0, so Slutsky's theorem gives

$$\sqrt{n}\,(A_n - p_1) \stackrel{d}{\to} N(0, p_1).$$

Similar arguments apply to B_n and to C_n. Because A_n and B_n and C_n are independent of one another, we may stack them as in Exercise 2.17 to obtain

$$\sqrt{n}\left\{ \begin{pmatrix} A_n \\ B_n \\ C_n \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 - p_1 \\ 1 - p_2 \end{pmatrix} \right\} \stackrel{d}{\to} N_3 \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 - p_1 & 0 \\ 0 & 0 & 1 - p_2 \end{pmatrix} \right\}$$

by Slutsky's theorem. For $\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^2$ defined in Equation (6.12), we obtain

$$\nabla \mathbf{g}(a,b,c) = \frac{1}{(a+b+c)^2} \begin{pmatrix} b+c & c \\ -a & c \\ -a & -a-b \end{pmatrix}.$$

Therefore,

$$[\nabla g(p_1, p_2 - p_1, 1 - p_2)]^T = \begin{pmatrix} 1-p_1 & -p_1 & -p_1 \\ 1-p_2 & 1-p_2 & -p_2 \end{pmatrix},$$

so the delta method gives

$$\sqrt{n}\left\{\begin{pmatrix} U_{(a_n)} \\ U_{(b_n)} \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}\right\} \stackrel{d}{\to} N_2\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} p_1(1-p_1) & p_1(1-p_2) \\ p_1(1-p_2) & p_2(1-p_2) \end{pmatrix}\right\}. \tag{6.13}$$

The method used above to derive the joint distribution (6.13) of two sample quantiles may be extended to any number of quantiles; doing so yields the following theorem:

Theorem 6.6 Suppose that for given constants $p_1, \ldots, p_k$ with $0 < p_1 < \cdots < p_k < 1$, there exist sequences $\{a_{1n}\}, \ldots, \{a_{kn}\}$ such that for all $1 \le i \le k$,

$$\sqrt{n}\left(\frac{a_{in}}{n} - p_i\right) \to 0.$$

Then if $U_1, \ldots, U_n$ is a sample from Uniform(0,1),

$$\sqrt{n}\left\{\begin{pmatrix} U_{(a_{1n})} \\ \vdots \\ U_{(a_{kn})} \end{pmatrix} - \begin{pmatrix} p_1 \\ \vdots \\ p_k \end{pmatrix}\right\} \stackrel{d}{\to} N_k\left\{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} p_1(1-p_1) & \cdots & p_1(1-p_k) \\ \vdots & & \vdots \\ p_1(1-p_k) & \cdots & p_k(1-p_k) \end{pmatrix}\right\}.$$

Note that for $i < j$, both the (i, j) and (j, i) entries in the covariance matrix above equal $p_i(1-p_j)$ and $p_j(1-p_i)$ never occurs in the matrix.

6.2.3 General sample quantiles

Let $F(x)$ be the distribution function for a random variable X. The quantile function $F^-(u)$ of Definition 3.13 is nondecreasing on $(0, 1)$ and it has the property that $F^-(U)$ has the same distribution as X for $U \sim$ uniform$(0, 1)$. This property follows from Lemma 3.14, since

$$P[F^-(U) \le x] = P[U \le F(x)] = F(x)$$

for all x. Since $F^-(u)$ is nondecreasing, it preserves ordering; thus, if $X_1, \ldots, X_n$ is a random sample from $F(x)$, then

$$(X_{(1)}, \ldots, X_{(n)}) \stackrel{d}{=} \left[F^-(U_{(1)}), \ldots, F^-(U_{(n)})\right].$$

(The symbol $\stackrel{d}{=}$ means "has the same distribution as".)

Now, suppose that at some point ξ, the derivative $F'(\xi)$ exists and is positive. Then $F(x)$ must be continuous and strictly increasing in a neighborhood of ξ. This implies that in this neighborhood, $F(x)$ has a well-defined inverse, which must be differentiable at the point $p \stackrel{\text{def}}{=} F(\xi)$. If $F^{-1}(u)$ denotes the inverse that exists in a neighborhood of p, then

$$\frac{dF^{-1}(p)}{dp} = \frac{1}{F'(\xi)}. \tag{6.14}$$

Equation (6.14) may be derived by differentiating the equation $F[F^{-1}(p)] = p$. Note also that whenever the inverse $F^{-1}(u)$ exists, it must coincide with the quantile function $F^{-}(u)$. Thus, the condition that $F'(\xi)$ exists and is positive is a sufficient condition to imply that $F^{-}(u)$ is differentiable at p. This differentiability is important: If we wish to transform the uniform order statistics $U_{(a_{in})}$ of Theorem 6.6 into order statistics $X_{(a_{in})}$ using the quantile function $F^{-}(u)$, the delta method requires the differentiability of $F^{-}(u)$ at each of the points $p_1, \ldots, p_k$.

The delta method, along with Equation (6.14), yields the following corollary of Theorem 6.6:

Theorem 6.7 Let $X_1, \ldots, X_n$ be a simple random sample from a distribution function $F(x)$ such that $F(x)$ is differentiable at each of the points $\xi_1 < \cdots < \xi_k$ and $F'(\xi_i) > 0$ for all i. Denote $F(\xi_i)$ by p_i. Then under the assumptions of Theorem 6.6,

$$\sqrt{n}\left\{\begin{pmatrix} X_{(a_{1n})} \\ \vdots \\ X_{(a_{kn})} \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}\right\} \stackrel{d}{\to} N_k\left\{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{p_1(1-p_1)}{F'(\xi_1)^2} & \cdots & \frac{p_1(1-p_k)}{F'(\xi_1)F'(\xi_k)} \\ \vdots & & \vdots \\ \frac{p_1(1-p_k)}{F'(\xi_1)F'(\xi_k)} & \cdots & \frac{p_k(1-p_k)}{F'(\xi_k)^2} \end{pmatrix}\right\}.$$

Exercises for Section 6.2

Exercise 6.9 Let X_1 be Uniform$(0, 2\pi)$ and let X_2 be standard exponential, independent of X_1. Find the joint distribution of $(Y_1, Y_2) = (\sqrt{2X_2}\cos X_1, \sqrt{2X_2}\sin X_1)$.

Note: Since $-\log U$ has a standard exponential distribution if $U \sim$ uniform$(0, 1)$, this problem may be used to simulate normal random variables using simulated uniform random variables.

Exercise 6.10 Suppose $X_1, \ldots, X_n$ is a simple random sample from a distribution that is symmetric about θ, which is to say that $P(X_i \le x) = F(x - \theta)$, where $F(x)$ is the distribution function for a distribution that is symmetric about zero. We wish to estimate θ by $(Q_p + Q_{1-p})/2$, where Q_p and Q_{1-p} are the p and $1-p$ sample quantiles, respectively. Find the smallest possible asymptotic variance for the estimator and the p for which it is achieved for each of the following forms of $F(x)$:

(a) Standard Cauchy

(b) Standard normal

(c) Standard double exponential

Hint: For at least one of the three parts of this question, you will have to solve for a minimizer numerically.

Exercise 6.11 When we use a boxplot to assess the symmetry of a distribution, one of the main things we do is visually compare the lengths of $Q_3 - Q_2$ and $Q_2 - Q_1$, where Q_i denotes the ith sample quartile.

(a) Given a random sample of size n from $N(0, 1)$, find the asymptotic distribution of $(Q_3 - Q_2) - (Q_2 - Q_1)$.

(b) Repeat part (a) if the sample comes from a standard logistic distribution.

(c) Using 1000 simulations from each distribution, use graphs to assess the accuracy of each of the asymptotic approximations above for $n = 5$ and $n = 13$. (For a sample of size $4k + 1$, define Q_i to be the $(ik + 1)$th order statistic.) For each value of n and each distribution, plot the empirical distribution function against the theoretical limiting distribution function.

Exercise 6.12 Let $X_1, \ldots, X_n$ be a random sample from Uniform$(0, 2\theta)$. Find the asymptotic distributions of the median, the midquartile range, and $\frac{2}{3}Q_3$, where Q_3 denotes the third quartile and the midquartile range is the mean of the 1st and 3rd quartiles. Compare these three estimates of θ based on their asymptotic variances.