In Example 10.2, we derived the asymptotic distribution of the maximum from a random sample from a uniform distribution. We did this using only the definition of convergence in distribution without relying on any results other than the fact that

\[ \left(1 + \frac{c_n}{b_n}\right)^{b_n} \to e^c \]  

(32)

if \( c_n \to c \) and \( b_n \to \infty \). In a similar way, we may derive the joint asymptotic distribution of several order statistics, as seen in the following example.

**Example 17.1** Range of uniform sample: Let \( X_1, \ldots, X_n \) be iid from Uniform(0,1). Let \( R_n = X_{(n)} - X_{(1)} \) denote the range of the sample. What is the asymptotic distribution of \( R_n \)?

We begin to answer this question by finding the joint asymptotic distribution of \((X_{(n)}, X_{(1)})\), as follows. For certain sequences \( k_n \) and \( \ell_n \), as yet unspecified, consider

\[ P(k_n X_{(1)} > x \text{ and } \ell_n (1 - X_{(n)}) > y) = P(X_{(1)} > x/k_n \text{ and } X_{(n)} < 1 - y/\ell_n) \]

\[ = P(x/k_n < X_{(1)} < \cdots < X_{(n)} < 1 - y/\ell_n), \]

where we have assumed that \( k_n \) and \( \ell_n \) 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

\[ 0 < x \frac{k_n}{k_n} < 1 - y \frac{\ell_n}{\ell_n} < 1, \]

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 (32) now makes it clear that \( k_n = n \) and \( \ell_n = n \) are sensible choices for \( k_n \) and \( \ell_n \), and the result is that

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

as long as

\[ 0 < x \frac{n}{n} < 1 - \frac{y}{n} < 1. \]  

(33)

Notice that condition (33) will be satisfied for large \( n \) 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 iid standard exponential random variables, say, \( Y_1 \) and \( Y_2 \), we conclude that

\[ \left( \frac{nX_{(1)}}{n(1 - X_{(n)})} = \frac{Y_1}{Y_2} \right). \]

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

\[ n(1 - X_{(n)} + X_{(1)}) = n(1 - R_n) \overset{d}{=} Y_1 + Y_2, \]

and the sum of independent gamma(1, 1) variables is gamma(2, 1).
Let’s consider a similar example in which the asymptotic joint distribution does not involve independent random variables.

**Example 17.2** As in Example 17.1, let \( X_1, \ldots, X_n \) be iid from uniform(0,1). We now consider the joint asymptotic distribution of \( X_{(n-1)} \) and \( X_{(n)} \). Omitting the step in which \( k_n \) and \( \ell_n \) are to be found, since they would again both be set to \( n \), we proceed as follows:

\[
P\{n(1 - X_{(n-1)}) > x \text{ and } n(1 - X_{(n)}) > y\} = P\left(X_{(n-1)} < 1 - \frac{x}{n} \text{ and } X_{(n)} < 1 - \frac{y}{n}\right). \tag{34}
\]

We consider two separate cases: If \( 0 < x < y \), then the right hand side of (34) is simply \( P(X_{(n)} < 1 - y/n) \), which converges to \( e^{-y} \). On the other hand, if \( 0 < y < x \), then

\[
P\left(X_{(n-1)} < 1 - \frac{x}{n} \text{ and } X_{(n)} < 1 - \frac{y}{n}\right) = P\left(X_{(n)} < 1 - \frac{x}{n}\right) + P\left(X_{(n-1)} < 1 - \frac{x}{n} \text{ and } X_{(n)} < 1 - \frac{y}{n}\right)
\]

\[
= \left(1 - \frac{x}{n}\right)^n + n \left(1 - \frac{x}{n}\right)^{n-1} \left(\frac{x}{n} - \frac{y}{n}\right)
\]

\[
\rightarrow e^{-x}(1 + x - y).
\]

What is this joint asymptotic distribution? Suppose that \( Y_1 \) and \( Y_2 \) are iid standard exponential variables as in the previous example. 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

\[
P(Y_1 + Y_2 > x \text{ and } Y_1 > y) = P(Y_1 > \max\{y, x - Y_2\}) = E e^{-\max\{y, x - Y_2\}}
\]

\[
= e^{-y}P(y > x - Y_2) + \int_0^{x-y} e^{t-x}e^{-t} dt = e^{-y}(1 + x - y).
\]

Therefore, we conclude that

\[
\left(\frac{n(1 - X_{(n-1)})}{n(1 - X_{(n)})}\right) \overset{D}{\sim} \left(\frac{Y_1 + Y_2}{Y_1}\right).
\]

Notice that this means that the asymptotic marginal distribution of \( n(1-X_{(n-1)}) \) is gamma(2,1).

Theorem 15 in Ferguson’s book is a generalization of Example 17.2.

Recall that if \( F \) is a continuous, invertible cdf and \( U \) is a standard uniform random variable, then \( F^{-1}(U) \sim F \). This is easy to prove, since \( P\{F^{-1}(U) \leq t\} = P(U \leq F(t)) = F(t) \). We may use this fact in conjunction with the result of Example 17.2 as in the following example.

**Example 17.3** Suppose \( X_1, \ldots, X_n \) are iid standard exponential random variables. What is the joint asymptotic distribution of \( (X_{(n-1)}, X_{(n)}) \)? Since the cdf of a standard exponential distribution is \( F(t) = 1 - e^{-t} \), whose inverse is \( F^{-1}(u) = -\log(1 - u) \), clearly

\[
\{-\log(1 - U_{(n-1)}), -\log(1 - U_{(n)})\} \overset{D}{=} \{X_{(n-1)}, X_{(n)}\},
\]

where \( \overset{D}{=\text{ means “has the same distribution”}} \). Therefore,

\[
-\log\{n(1 - U_{(n-1)}), n(1 - U_{(n)})\} = \{-\log(1 - U_{(n-1)}) - \log n, -\log(1 - U_{(n)}) - \log n\}
\]

\[
\overset{D}{=} \{X_{(n-1)} - \log n, X_{(n)} - \log n\}.
\]

We conclude by the result of Example 17.2 that

\[
\left(\frac{X_{(n-1)} - \log n}{X_{(n)} - \log n}\right) \overset{D}{\sim} \left(\frac{-\log(Y_1 + Y_2)}{-\log Y_1}\right),
\]

where \( Y_1 \) and \( Y_2 \) are iid standard exponential variables.

47
Problems

Problem 17.1 If \( X_1, \ldots, X_n \) are iid standard uniform variables, find the joint asymptotic distribution of \( \{nX(2), n(1 - X(n-1))\} \).

*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 \( \leq 1 \).

Problem 17.2 Let \( X_1, \ldots, X_n \) be a random sample from the distribution with cdf \( F(x) = [1 - (1/x)]I\{x > 1\} \).

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

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

*Hint:* In part (a), proceed as in Example 17.3.

Problem 17.3 If \( X_1, \ldots, X_n \) are iid \( \text{uniform}(0,1) \) variables, prove that \( X_{(1)}/X_{(2)} \overset{d}{\rightarrow} \text{uniform}(0,1) \).

Problem 17.4 Let \( X_1, \ldots, X_n \) be iid from uniform \((0, 2\theta)\).

(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 \( 3 \times 3 \) 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)?

Problem 17.5 Let \( X_1, \ldots, X_n \) be an iid sample from a logistic distribution with cdf \( F(t) = e^{t/\theta}/(1 + e^{t/\theta}) \) for all \( t \).

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

(b) Based on part (a), construct an approximate 95% confidence interval for \( \theta \). 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 \( \theta \)?

*Hint:* In part (a), use the fact that \( \log U_{(n)} \) and \( \log U_{(n-1)} \) both converge in probability to zero.
We begin with a density for a multivariate normal distribution on $\mathbb{R}^k$. However, note that not all multivariate normal distributions have densities—consider the univariate example of $N(0, 0)$.

Given a mean vector $\xi \in \mathbb{R}^k$ and a positive definite $k \times k$ covariance matrix $\Sigma$, $X$ has a multivariate normal distribution with mean $\xi$ and covariance $\Sigma$, written $X \sim N_k(\xi, \Sigma)$, if its density on $\mathbb{R}^k$ is

$$f(x) = C \exp \left\{ -\frac{1}{2} (x - \xi)^t \Sigma^{-1} (x - \xi) \right\}.$$  

(35)

In expression (35), the constant is $C = (2^k \pi^k |\Sigma|)^{-1/2}$, where $|\Sigma|$ denotes the determinant of $\Sigma$. Because of the assumption that $\Sigma$ is positive definite, an assumption that we will later relax, $|\Sigma|$ is guaranteed to be positive.

As a special case, consider the bivariate normal distribution, where for some $\sigma^2 > 0$, $\tau^2 > 0$, and $-1 < \rho < 1$ we have

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho \sigma \tau \\ \rho \sigma \tau & \tau^2 \end{pmatrix}$$

and thus

$$|\Sigma| = \sigma^2 \tau^2 (1 - \rho^2) \quad \text{and} \quad \Sigma^{-1} = \frac{1}{\sigma^2 \tau^2 (1 - \rho^2)} \begin{pmatrix} \tau^2 & -\rho \sigma \tau \\ -\rho \sigma \tau & \sigma^2 \end{pmatrix}.$$

In this case, of course, $X_1$ and $X_2$ have correlation $\rho$ and marginal variances $\sigma^2$ and $\tau^2$, respectively. To see the bivariate normal density really written out, see Lehmann page 287. However, here we will use matrix notation whenever possible because of its elegance compared to the componentwise expansions.

To define the multivariate normal distribution in full generality, we first consider the case in which $\Sigma$ is diagonal, say $\Sigma = D = \text{diag}(d_1, \ldots, d_k)$. Of course, if $D$ is a legitimate covariance matrix it must be nonnegative definite (or positive semidefinite), which means that all of its eigenvalues are nonnegative. Since the eigenvalues of a diagonal matrix are simply its diagonal elements, we see immediately that $d_i \geq 0$ for all $i$. We may now define a multivariate normal distribution with covariance matrix $D$.

**Definition 18.1** Suppose that $D = \text{diag}(d_1, \ldots, d_k)$ for nonnegative real numbers $d_1, \ldots, d_k$. Then for $\xi \in \mathbb{R}^k$, the multivariate normal distribution with mean $\xi$ and covariance $D$, denoted $N_k(\xi, D)$, is the joint distribution of the independent random variables $X_1, \ldots, X_k$, where $X_i \sim N(\xi_i, d_i)$.

To define a multivariate normal distribution for a general covariance matrix $\Sigma$, we make use of the fact that any symmetric matrix may be diagonalized by an orthogonal matrix. We first define orthogonal, then state the diagonalizability result as a lemma.

**Definition 18.2** A square matrix $Q$ is orthogonal if $Q^{-1}$ exists and is equal to $Q^t$.

**Lemma 18.1** If $A$ is a symmetric $k \times k$ matrix, then there exists an orthogonal matrix $Q$ such that $QAQ^t$ is diagonal.

Note that the diagonal elements of the matrix $QAQ^t$ in the matrix above must be the eigenvalues of $A$. This is easy to prove, since if $\lambda$ is a diagonal element of $QAQ^t$ then it is an eigenvalue of $QAQ^t$ and hence there exists a vector $x$ such that $QAQ^t x = \lambda x$, which implies that $A(Q^tx) = \lambda(Q^tx)$ and so $\lambda$ is an eigenvalue of $A$.
**Definition 18.3** Suppose $\Sigma$ is an arbitrary symmetric $k \times k$ matrix with nonnegative eigenvalues.

Let $Q$ be an orthogonal matrix such that $Q\Sigma Q^T$ is diagonal. Then for $\xi \in \mathbb{R}^k$, the multivariate normal distribution with mean $\xi$ and covariance $\Sigma$, denoted $N_k(\xi, \Sigma)$, is the distribution of $\xi + Q^T Y$, where $Y \sim N_k(0, Q\Sigma Q^T)$.

It is not immediately clear that $N_k(\xi, \Sigma)$ is well-defined by Definition 18.3, since it is not clear that $\xi + Q^T Y$ and $\xi + Q_2^T Y$ must have the same distribution whenever $Q_1 \Sigma Q_1^T$ and $Q_2 \Sigma Q_2^T$ are both diagonal for orthogonal $Q_1$ and $Q_2$. However, this is indeed true (though we will not prove it here), so Definition 18.3 is a valid definition.

Now that $N_k(\xi, \Sigma)$ is defined, we may state the multivariate version of the central limit theorem:

**Theorem 18.1** If $X^{(1)}, X^{(2)}, \ldots$ are iid with mean vector $\mu \in \mathbb{R}^k$ and covariance $\Sigma$ (where $\Sigma$ has finite entries), then

$$\sqrt{n} (\bar{X}_n - \mu) \xrightarrow{d} N_k(0, \Sigma).$$

**Example 18.1** Let $X_1, X_2, \ldots$ be iid with $E X_i = \xi$, $Var X_i = \sigma^2$, $E (X_i - \xi)^3 = \gamma$, and $Var (X_i - \xi)^2 = \tau^2 < \infty$. Let

$$S_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$  \hspace{1cm} (36)

We have shown earlier that $\sqrt{n} (S_n^2 - \sigma^2) \xrightarrow{d} N(0, \tau^2)$. The same fact may be proven using the multivariate results of Chapter 5 as follows.

First, let $Y_i = X_i - \xi$ and $Z_i = Y_i^2$. We may use the multivariate central limit theorem to find the joint asymptotic distribution of $\bar{Y}_n$ and $\bar{Z}_n$, namely

$$\sqrt{n} \left\{ \begin{pmatrix} \bar{Y}_n \\ \bar{Z}_n \end{pmatrix} - \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix} \right\} \xrightarrow{d} N_2 \left\{ \begin{pmatrix} 0 \\ (\sigma^2, \gamma, \tau^2) \end{pmatrix} \right\}.$$  \hspace{1cm} (36)

Note that the above result uses the fact that $Cov(Y_1, Z_1) = \gamma$, which is easy to check.

We may write $S_n^2 = Z_n - (\bar{Y}_n)^2$. Therefore, define the function $g(a, b) = b - a^2$ and note that this gives $\hat{g}(a, b) = (-2a, 1)$. To use the delta method, we should evaluate

$$\hat{g}(0, \sigma^2) \left( \begin{array}{c} \sigma^2 \\ \gamma \\ \tau^2 \end{array} \right) \hat{g}(0, \sigma^2)^T = (0, 1) \left( \begin{array}{c} \sigma^2 \\ \gamma \\ \tau^2 \end{array} \right) \left( \begin{array}{c} 0 \\ 1 \end{array} \right) = \hat{g}(0, \sigma^2)^T = \tau^2$$

We conclude that

$$\sqrt{n} \left\{ g(\bar{Y}_n) - g(\bar{Z}_n) \right\} = \sqrt{n} (S_n^2 - \sigma^2) \xrightarrow{d} N(0, \tau^2)$$

as found earlier.

---

**Problems**

**Problem 18.1** Suppose $X \sim N_k(\xi, \Sigma)$, where $\Sigma$ is invertible. Prove that

$$(X - \xi)^T \Sigma^{-1} (X - \xi) \sim \chi^2_k.$$

**Hint:** If $Q$ diagonalizes $\Sigma$, say $Q \Sigma Q^T = \Lambda$, let $\Lambda^{1/2}$ have the obvious definition and consider $\Lambda^{1/2} Y$, where $Y = (\Lambda^{1/2})^{-1} Q(X - \xi)$.  

50
**Problem 18.2**  Let $X_1, X_2, \ldots$ be iid from $N(\xi, \sigma^2)$ where $\xi \neq 0$. Let $S_n^2$ be defined as in equation (36). Find the asymptotic distribution of the coefficient of variation $S_n/X_n$.

**Problem 18.3**  If $X$ and $Y$ are standard normal distributions, then any mixture of these distributions is trivially standard normal since $\alpha f_X(x) + (1 - \alpha)f_Y(x)$ is a standard normal density. Is the same true of bivariate normal distributions? In other words, if $X$ and $Y$ are bivariate normal, where each of the marginal distributions is standard normal (that is, $X_1, X_2, Y_1,$ and $Y_2$ are all standard normal), is $\alpha f_X(x) + (1 - \alpha)f_Y(x)$ a bivariate normal density? If yes, prove it; if no, provide a counterexample.