CHAPTER 5

Limiting Distributions

5.1 Convergence in Distribution

In some of the preceding chapters it has been demonstrated by example that the distribution of a random variable (perhaps a statistic) often depends upon a positive integer n. For example, if the random variable X is $b(n, p)$, the distribution of X depends upon n. If $\bar{X}$ is the mean of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$, then $\bar{X}$ is itself $N(\mu, \sigma^2/n)$ and the distribution of $\bar{X}$ depends upon n. If S^2 is the variance of this random sample from the normal distribution to which we have just referred, the random variable nS^2/σ^2 is $\chi^2(n-1)$, and so the distribution of this random variable depends upon n.

We know from experience that the determination of the probability density function of a random variable can, upon occasion, present rather formidable computational difficulties. For example, if $\bar{X}$ is the mean of a random sample $X_1, X_2, \ldots, X_n$ from a distribution that has the following p.d.f.

$$f(x) = 1, \qquad 0 < x < 1,$$
$$= 0 \qquad \text{elsewhere,}$$

then (Exercise 4.85) the m.g.f. of $\bar{X}$ is given by $[M(t/n)]^n$, where here

$$M(t) = \int_0^1 e^{tx}\, dx = \frac{e^t - 1}{t}, \qquad t \neq 0,$$
$$= 1, \qquad t = 0.$$

Hence

$$E(e^{t\bar{X}}) = \left(\frac{e^{t/n} - 1}{t/n}\right)^n, \qquad t \neq 0,$$
$$= 1, \qquad t = 0.$$

Since the m.g.f. of $\bar{X}$ depends upon n, the distribution of $\bar{X}$ depends upon n. It is true that various mathematical techniques can be used to determine the p.d.f. of $\bar{X}$ for a fixed, but arbitrarily fixed, positive integer n. But the p.d.f. is so complicated that few, if any, of us would be interested in using it to compute probabilities about $\bar{X}$. One of the purposes of this chapter is to provide ways of approximating, for large values of n, some of these complicated probability density functions.

Consider a distribution that depends upon the positive integer n. Clearly, the distribution function F of that distribution will also depend upon n. Throughout this chapter, we denote this fact by writing the distribution function as F_n and the corresponding p.d.f. as f_n. Moreover, to emphasize the fact that we are working with sequences of distribution functions and random variables, we place a subscript n on the random variables. For example, we shall write

$$F_n(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{1}{\sqrt{1/n}\sqrt{2\pi}} e^{-nw^2/2}\, dw$$

for the distribution function of the mean $\bar{X}_n$ of a random sample of size n from a normal distribution with mean zero and variance 1.

We now define convergence in distribution of a sequence of random variables.

Definition 1. Let the distribution function $F_n(y)$ of the random variable Y_n depend upon n, $n = 1, 2, 3, \ldots$. If $F(y)$ is a distribution function and if $\lim_{n\to\infty} F_n(y) = F(y)$ for every point y at which $F(y)$ is continuous, then the sequence of random variables, $Y_1, Y_2, \ldots$, *converges in distribution* to a random variable with distribution function $F(y)$.

The following examples are illustrative of this convergence in distribution.

Example 1. Let Y_n denote the nth order statistic of a random sample $X_1, X_2, \ldots, X_n$ from a distribution having p.d.f.

$$f(x) = \frac{1}{\theta}, \qquad 0 < x < \theta, \quad 0 < \theta < \infty,$$
$$= 0 \qquad \text{elsewhere.}$$

The p.d.f. of Y_n is

$$g_n(y) = \frac{ny^{n-1}}{\theta^n}, \qquad 0 < y < \theta,$$
$$= 0 \qquad \text{elsewhere,}$$

and the distribution function of Y_n is

$$F_n(y) = 0, \qquad y < 0,$$
$$= \int_0^y \frac{nz^{n-1}}{\theta^n}\,dz = \left(\frac{y}{\theta}\right)^n, \qquad 0 \le y < \theta,$$
$$= 1, \qquad \theta \le y < \infty.$$

Then

$$\lim_{n\to\infty} F_n(y) = 0, \qquad -\infty < y < \theta,$$
$$= 1, \qquad \theta \le y < \infty.$$

Now

$$F(y) = 0, \qquad -\infty < y < \theta,$$
$$= 1, \qquad \theta \le y < \infty,$$

is a distribution function. Moreover, $\lim_{n\to\infty} F_n(y) = F(y)$ at each point of continuity of $F(y)$. Recall that a distribution of the discrete type which has a probability of 1 at a single point has been called a *degenerate distribution.* Thus, in this example, the sequence of the nth order statistics, Y_n, $n = 1, 2, 3, \ldots$, converges in distribution to a random variable that has a degenerate distribution at the point $y = \theta$.

Example 2. Let $\bar{X}_n$ have the distribution function

$$F_n(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{1}{\sqrt{1/n}\sqrt{2\pi}} e^{-nw^2/2}\, dw.$$

If the change of variable $v = \sqrt{n}w$ is made, we have

$$F_n(\bar{x}) = \int_{-\infty}^{\sqrt{n}\bar{x}} \frac{1}{\sqrt{2\pi}} e^{-v^2/2}\, dv.$$

It is clear that

$$\begin{aligned} \lim_{n\to\infty} F_n(\bar{x}) &= 0, \qquad \bar{x} < 0, \\ &= \tfrac{1}{2}, \qquad \bar{x} = 0, \\ &= 1, \qquad \bar{x} > 0. \end{aligned}$$

Now the function

$$\begin{aligned} F(\bar{x}) &= 0, \qquad \bar{x} < 0, \\ &= 1, \qquad \bar{x} \geq 0, \end{aligned}$$

is a distribution function and $\lim_{n\to\infty} F_n(\bar{x}) = F(\bar{x})$ at every point of continuity of $F(\bar{x})$. To be sure, $\lim_{n\to\infty} F_n(0) \neq F(0)$, but $F(\bar{x})$ is not continuous at $\bar{x} = 0$. Accordingly, the sequence $\bar{X}_1, \bar{X}_2, \bar{X}_3, \ldots$ converges in distribution to a random variable that has a degenerate distribution at $\bar{x} = 0$.

Example 3. Even if a sequence $X_1, X_2, X_3, \ldots$ converges in distribution to a random variable X, we cannot in general determine the distribution of X by taking the limit of the p.d.f. of X_n. This is illustrated by letting X_n have the p.d.f.

$$\begin{aligned} f_n(x) &= 1, \qquad x = 2 + \frac{1}{n}, \\ &= 0 \qquad \text{elsewhere.} \end{aligned}$$

Clearly, $\lim_{n\to\infty} f_n(x) = 0$ for all values of x. This may suggest that X_n, $n = 1, 2, 3, \ldots$, does not converge in distribution. However, the distribution function of X_n is

$$\begin{aligned} F_n(x) &= 0, \qquad x < 2 + \frac{1}{n}, \\ &= 1, \qquad x \geq 2 + \frac{1}{n}, \end{aligned}$$

and

$$\lim_{n\to\infty} F_n(x) = 0, \qquad x \le 2,$$

$$= 1, \qquad x > 2.$$

Since

$$F(x) = 0, \qquad x < 2,$$

$$= 1, \qquad x \ge 2,$$

is a distribution function, and since $\lim_{n\to\infty} F_n(x) = F(x)$ at all points of continuity of $F(x)$, the sequence $X_1, X_2, X_3, \ldots$ converges in distribution to a random variable with distribution function $F(x)$.

It is interesting to note that although we refer to a sequence of random variables, $X_1, X_2, X_3, \ldots,$ converging in distribution to a random variable X having some distribution function $F(x)$, it is actually the distribution functions $F_1, F_2, F_3, \ldots$ that converge. That is,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

at all points x for which $F(x)$ is continuous. For that reason we often find it convenient to refer to $F(x)$ as the *limiting distribution*. Moreover, it is then a little easier to say that X_n, representing the sequence $X_1, X_2, X_3, \ldots,$ has a limiting distribution with distribution function $F(x)$. Henceforth, we use this terminology.

Example 4. Let Y_n denote the nth order statistic of a random sample from the uniform distribution of Example 1. Let $Z_n = n(\theta - Y_n)$. The p.d.f. of Z_n is

$$h_n(z) = \frac{(\theta - z/n)^{n-1}}{\theta^n}, \qquad 0 < z < n\theta,$$

$$= 0 \qquad \text{elsewhere},$$

and the distribution function of Z_n is

$$G_n(z) = 0, \qquad z < 0,$$

$$= \int_0^z \frac{(\theta - w/n)^{n-1}}{\theta^n}\, dw = 1 - \left(1 - \frac{z}{n\theta}\right)^n, \qquad 0 \le z < n\theta,$$

$$= 1, \qquad n\theta \le z.$$

Hence

$$\lim_{n\to\infty} G_n(z) = 0, \qquad z \le 0,$$
$$= 1 - e^{-z/\theta}, \qquad 0 < z < \infty.$$

Now

$$G(z) = 0, \qquad z < 0,$$
$$= 1 - e^{-z/\theta}, \qquad 0 \le z,$$

is a distribution function that is everywhere continuous and $\lim_{n\to\infty} G_n(z) = G(z)$ at all points. Thus Z_n has a limiting distribution with distribution function $G(z)$. This affords us an example of a limiting distribution that is not degenerate.

Example 5. Let T_n have a t-distribution with n degrees of freedom, $n = 1, 2, 3, \ldots$. Thus its distribution function is

$$F_n(t) = \int_{-\infty}^{t} \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\,\Gamma(n/2)} \frac{1}{(1 + y^2/n)^{(n+1)/2}}\, dy,$$

where the integrand is the p.d.f. $f_n(y)$ of T_n. Accordingly,

$$\lim_{n\to\infty} F_n(t) = \lim_{n\to\infty} \int_{-\infty}^{t} f_n(y)\, dy$$
$$= \int_{-\infty}^{t} \lim_{n\to\infty} f_n(y)\, dy.$$

The change of the order of the limit and integration is justified because $|f_n(y)|$ is dominated by a function, like $10f_1(y)$, with a finite integral. That is,

$$|f_n(y)| \le 10 f_1(y)$$

and

$$\int_{-\infty}^{t} 10 f_1(y)\, dy = \frac{10}{\pi} \arctan t < \infty,$$

for all real t. Hence, here we can find the limiting distribution by finding the limit of the p.d.f. of T_n. It is

$$\lim_{n\to\infty} f_n(y) = \lim_{n\to\infty} \left[\frac{\Gamma[(n+1)/2]}{\sqrt{n/2}\,\Gamma(n/2)}\right]$$
$$\times \lim_{n\to\infty} \frac{1}{(1 + y^2/n)^{1/2}} \lim_{n\to\infty} \left\{\frac{1}{\sqrt{2\pi}} \left[\left(1 + \frac{y^2}{n}\right)\right]^{-n/2}\right\}.$$

Using the fact from elementary calculus that

$$\lim_{n\to\infty}\left(1+\frac{y^2}{n}\right)^n = e^{y^2},$$

the limit associated with the third factor is clearly the p.d.f. of the standard normal distribution. The second limit obviously equals 1. If we knew more about the gamma function, it is easy to show that the first limit also equals 1. Thus we have

$$\lim_{n\to\infty} F_n(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy,$$

and hence T_n has a limiting standard normal distribution.

EXERCISES

5.1. Let $\bar{X}_n$ denote the mean of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. Find the limiting distribution of $\bar{X}_n$.

5.2. Let Y_1 denote the first order statistic of a random sample of size n from a distribution that has the p.d.f. $f(x) = e^{-(x-\theta)}$, $\theta < x < \infty$, zero elsewhere. Let $Z_n = n(Y_1 - \theta)$. Investigate the limiting distribution of Z_n.

5.3. Let Y_n denote the nth order statistic of a random sample from a distribution of the continuous type that has distribution function $F(x)$ and p.d.f. $f(x) = F'(x)$. Find the limiting distribution of $Z_n = n[1 - F(Y_n)]$.

5.4. Let Y_2 denote the second order statistic of a random sample of size n from a distribution of the continuous type that has distribution function $F(x)$ and p.d.f. $f(x) = F'(x)$. Find the limiting distribution of $W_n = nF(Y_2)$.

5.5. Let the p.d.f. of Y_n be $f_n(y) = 1$, $y = n$, zero elsewhere. Show that Y_n does not have a limiting distribution. (In this case, the probability has "escaped" to infinity.)

5.6. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution that is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$. Show that the sum $Z_n = \sum_1^n X_i$ does not have a limiting distribution.

5.2 Convergence in Probability

In the discussion concerning convergence in distribution, it was noted that it was really the sequence of distribution functions that converges to what we call a limiting distribution function.

Convergence in probability is quite different, although we demonstrate that in a special case there is a relationship between the two concepts.

Definition 2. A sequence of random variables $X_1, X_2, X_3, \ldots$ *converges in probability* to a random variable X if, for every $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr(|X_n - X| < \epsilon) = 1,$$

or equivalently,

$$\lim_{n\to\infty} \Pr(|X_n - X| \ge \epsilon) = 0.$$

Statisticians are usually interested in this convergence when the random variable X is a constant, that is, when the random variable X has a degenerate distribution at that constant. Hence we concentrate on that situation.

Example 1. Let $\bar{X}_n$ denote the mean of a random sample of size n from a distribution that has mean μ and positive variance σ^2. Then the mean and variance of $\bar{X}_n$ are μ and σ^2/n. Consider, for every fixed $\epsilon > 0$, the probability

$$\Pr(|\bar{X}_n - \mu| \ge \epsilon) = \Pr\left(|\bar{X}_n - \mu| \ge \frac{k\sigma}{\sqrt{n}}\right),$$

where $k = \epsilon\sqrt{n}/\sigma$. In accordance with the inequality of Chebyshev, this probability is less than or equal to $1/k^2 = \sigma^2/n\epsilon^2$. So, for every fixed $\epsilon > 0$, we have

$$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mu| \ge \epsilon) \le \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0.$$

Hence $\bar{X}_n$, $n = 1, 2, 3, \ldots$, converges in probability to μ if σ^2 is finite. (In a more advanced course, the student will learn that μ finite is sufficient to ensure this convergence in probability.) This result is called the *weak law of large numbers*.

Remark. A stronger type of convergence is given by $\Pr(\lim_{n\to\infty} Y_n = c) = 1$; in this case we say that Y_n, $n = 1, 2, 3, \ldots$, converges to c *with probability 1*. Although we do not consider this type of convergence, it is known that the mean $\bar{X}_n$, $n = 1, 2, 3, \ldots$, of a random sample converges with probability 1 to the mean μ of the distribution, provided that the latter exists. This is one form of the *strong law of large numbers*.

We prove a theorem that relates a certain limiting distribution to convergence in probability to a constant.

Theorem 1. *Let $F_n(y)$ denote the distribution function of a random variable Y_n whose distribution depends upon the positive integer n. Let c denote a constant which does not depend upon n. The sequence Y_n, $n = 1, 2, 3, \ldots$, converges in probability to the constant c if and only if the limiting distribution of Y_n is degenerate at $y = c$.*

Proof. First, assume that the $\lim_{n\to\infty} \Pr(|Y_n - c| < \epsilon) = 1$ for every $\epsilon > 0$. We are to prove that the random variable Y_n is such that

$$\begin{aligned} \lim_{n\to\infty} F_n(y) &= 0, \qquad y < c, \\ &= 1, \qquad y > c. \end{aligned}$$

Note that we do not need to know anything about the $\lim_{n\to\infty} F_n(c)$. For if the limit of $F_n(y)$ is as indicated, then Y_n has a limiting distribution with distribution function

$$\begin{aligned} F(y) &= 0, \qquad y < c, \\ &= 1, \qquad y \ge c. \end{aligned}$$

Now

$$\Pr(|Y_n - c| < \epsilon) = F_n[(c + \epsilon) -] - F_n(c - \epsilon),$$

where $F_n[(c + \epsilon) -]$ is the left-hand limit of $F_n(y)$ at $y = c + \epsilon$. Thus we have

$$1 = \lim_{n\to\infty} \Pr(|Y_n - c| < \epsilon) = \lim_{n\to\infty} F_n[(c + \epsilon) -] - \lim_{n\to\infty} F_n(c - \epsilon).$$

Because $0 \le F_n(y) \le 1$ for all values of y and for every positive integer n, it must be that

$$\lim_{n\to\infty} F_n(c - \epsilon) = 0, \qquad \lim_{n\to\infty} F_n[(c + \epsilon) -] = 1.$$

Since this is true for every $\epsilon > 0$, we have

$$\begin{aligned} \lim_{n\to\infty} F_n(y) &= 0, \qquad y < c, \\ &= 1, \qquad y > c, \end{aligned}$$

as we were required to show.

To complete the proof of Theorem 1, we assume that

$$\lim_{n\to\infty} F_n(y) = 0, \qquad y < c,$$
$$= 1, \qquad y > c.$$

We are to prove that $\lim_{n\to\infty} \Pr(|Y_n - c| < \epsilon) = 1$ for every $\epsilon > 0$. Because

$$\Pr(|Y_n - c| < \epsilon) = F_n[(c + \epsilon) -] - F_n(c - \epsilon),$$

and because it is given that

$$\lim_{n\to\infty} F_n[(c + \epsilon) -] = 1,$$
$$\lim_{n\to\infty} F_n(c - \epsilon) = 0,$$

for every $\epsilon > 0$, we have the desired result. This completes the proof of the theorem.

For convenience, in the notation of Theorem 1, we sometimes say that Y_n, rather than the sequence $Y_1, Y_2, Y_3, \ldots$, converges in probability to the constant c.

EXERCISES

5.7. Let the random variable Y_n have a distribution that is $b(n, p)$.
(a) Prove that Y_n/n converges in probability to p. This result is one form of the weak law of large numbers.
(b) Prove that $1 - Y_n/n$ converges in probability to $1 - p$.

5.8. Let S_n^2 denote the variance of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. Prove that $nS_n^2/(n-1)$ converges in probability to σ^2.

5.9. Let W_n denote a random variable with mean μ and variance b/n^p, where $p > 0$, μ, and b are constants (not functions of n). Prove that W_n converges in probability to μ.
Hint: Use Chebyshev's inequality.

5.10. Let Y_n denote the nth order statistic of a random sample of size n from a uniform distribution on the interval $(0, \theta)$, as in Example 1 of Section 5.1. Prove that $Z_n = \sqrt{Y_n}$ converges in probability to $\sqrt{\theta}$.

5.3 Limiting Moment-Generating Functions

To find the limiting distribution function of a random variable Y_n by use of the definition of limiting distribution function obviously requires that we know $F_n(y)$ for each positive integer n. But, as indicated in the introductory remarks of Section 5.1, this is precisely the problem we should like to avoid. If it exists, the moment-generating function that corresponds to the distribution function $F_n(y)$ often provides a convenient method of determining the limiting distribution function. To emphasize that the distribution of a random variable Y_n depends upon the positive integer n, in this chapter we shall write the moment-generating function of Y_n in the form $M(t; n)$.

The following theorem, which is essentially Curtiss' modification of a theorem of Lévy and Cramér, explains how the moment-generating function may be used in problems of limiting distributions. A proof of the theorem requires a knowledge of that same facet of analysis that permitted us to assert that a moment-generating function, when it exists, uniquely determines a distribution. Accordingly, no proof of the theorem will be given.

Theorem 2. *Let the random variable Y_n have the distribution function $F_n(y)$ and the moment-generating function $M(t; n)$ that exists for $-h < t < h$ for all n. If there exists a distribution function $F(y)$, with corresponding moment-generating function $M(t)$, defined for $|t| \le h_1 < h$, such that $\lim_{n\to\infty} M(t; n) = M(t)$, then Y_n has a limiting distribution with distribution function $F(y)$.*

In this and the subsequent sections are several illustrations of the use of Theorem 2. In some of these examples it is convenient to use a certain limit that is established in some courses in advanced calculus. We refer to a limit of the form

$$\lim_{n\to\infty}\left[1 + \frac{b}{n} + \frac{\psi(n)}{n}\right]^{cn},$$

where b and c do not depend upon n and where $\lim_{n\to\infty} \psi(n) = 0$. Then

$$\lim_{n\to\infty}\left[1 + \frac{b}{n} + \frac{\psi(n)}{n}\right]^{cn} = \lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{cn} = e^{bc}.$$

For example,

$$\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^3}{n^{3/2}}\right)^{-n/2}=\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^3/\sqrt{n}}{n}\right)^{-n/2}.$$

Here $b = -t^2$, $c = -\frac{1}{2}$, and $\psi(n) = t^3/\sqrt{n}$. Accordingly, for every fixed value of t, the limit is $e^{t^2/2}$.

Example 1. Let Y_n have a distribution that is $b(n, p)$. Suppose that the mean $\mu = np$ is the same for every n; that is, $p = \mu/n$, where μ is a constant. We shall find the limiting distribution of the binomial distribution, when $p = \mu/n$, by finding the limit of $M(t; n)$. Now

$$M(t; n) = E(e^{tY_n}) = [(1-p)+pe^t]^n = \left[1+\frac{\mu(e^t-1)}{n}\right]^n$$

for all real values of t. Hence we have

$$\lim_{n\to\infty} M(t; n) = e^{\mu(e^t-1)}$$

for all real values of t. Since there exists a distribution, namely the Poisson distribution with mean μ, that has this m.g.f. $e^{\mu(e^t-1)}$, then, in accordance with the theorem and under the conditions stated, it is seen that Y_n has a limiting Poisson distribution with mean μ.

Whenever a random variable has a limiting distribution, we may, if we wish, use the limiting distribution as an approximation to the exact distribution function. The result of this example enables us to use the Poisson distribution as an approximation to the binomial distribution when n is large and p is small. This is clearly an advantage, for it is easy to provide tables for the one-parameter Poisson distribution. On the other hand, the binomial distribution has two parameters, and tables for this distribution are very ungainly. To illustrate the use of the approximation, let Y have a binomial distribution with $n = 50$ and $p = \frac{1}{25}$. Then

$$\Pr(Y \le 1) = (\tfrac{24}{25})^{50} + 50(\tfrac{1}{25})(\tfrac{24}{25})^{49} = 0.400,$$

approximately. Since $\mu = np = 2$, the Poisson approximation to this probability is

$$e^{-2} + 2e^{-2} = 0.406.$$

Example 2. Let Z_n be $\chi^2(n)$. Then the m.g.f. of Z_n is $(1-2t)^{-n/2}$, $t < \frac{1}{2}$. The mean and the variance of Z_n are, respectively, n and $2n$. The limiting distribution of the random variable $Y_n = (Z_n - n)/\sqrt{2n}$ will be investigated. Now the m.g.f. of Y_n is

$$M(t; n) = E\left\{\exp\left[t\left(\frac{Z_n - n}{\sqrt{2n}}\right)\right]\right\}$$

$$= e^{-tn/\sqrt{2n}} E\left(e^{tZ_n/\sqrt{2n}}\right)$$

$$= \exp\left[-\left(t\sqrt{\frac{2}{n}}\right)\left(\frac{n}{2}\right)\right]\left(1 - 2\frac{t}{\sqrt{2n}}\right)^{-n/2}, \qquad t < \frac{\sqrt{2n}}{2}.$$

This may be written in the form

$$M(t; n) = \left(e^{t\sqrt{2/n}} - t\sqrt{\frac{2}{n}}\, e^{t\sqrt{2/n}}\right)^{-n/2}, \qquad t < \sqrt{\frac{n}{2}}.$$

In accordance with Taylor's formula, there exists a number $\xi(n)$, between 0 and $t\sqrt{2/n}$, such that

$$e^{t\sqrt{2/n}} = 1 + t\sqrt{\frac{2}{n}} + \frac{1}{2}\left(t\sqrt{\frac{2}{n}}\right)^2 + \frac{e^{\xi(n)}}{6}\left(t\sqrt{\frac{2}{n}}\right)^3.$$

If this sum is substituted for $e^{t\sqrt{2/n}}$ in the last expression for $M(t; n)$, it is seen that

$$M(t; n) = \left(1 - \frac{t^2}{n} + \frac{\psi(n)}{n}\right)^{-n/2},$$

where

$$\psi(n) = \frac{\sqrt{2}t^3e^{\xi(n)}}{3\sqrt{n}} - \frac{\sqrt{2}t^3}{\sqrt{n}} - \frac{2t^4e^{\xi(n)}}{3n}.$$

Since $\xi(n) \to 0$ as $n \to \infty$, then $\lim \psi(n) = 0$ for every fixed value of t. In accordance with the limit proposition cited earlier in this section, we have

$$\lim_{n \to \infty} M(t; n) = e^{t^2/2}$$

for all real values of t. That is, the random variable $Y_n = (Z_n - n)/\sqrt{2n}$ has a limiting standard normal distribution.

EXERCISES

5.11. Let X_n have a gamma distribution with parameter $\alpha = n$ and β, where β is not a function of n. Let $Y_n = X_n/n$. Find the limiting distribution of Y_n.

5.12. Let Z_n be $\chi^2(n)$ and let $W_n = Z_n/n^2$. Find the limiting distribution of W_n.

5.13. Let X be $\chi^2(50)$. Approximate Pr $(40 < X < 60)$.

5.14. Let $p = 0.95$ be the probability that a man, in a certain age group, lives at least 5 years.

(a) If we are to observe 60 such men and if we assume independence, find the probability that at least 56 of them live 5 or more years.

(b) Find an approximation to the result of part (a) by using the Poisson distribution.
Hint: Redefine p to be 0.05 and $1 - p = 0.95$.

5.15. Let the random variable Z_n have a Poisson distribution with parameter $\mu = n$. Show that the limiting distribution of the random variable $Y_n = (Z_n - n)/\sqrt{n}$ is normal with mean zero and variance 1.

5.16. Let S_n^2 denote the variance of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. It has been proved that $nS_n^2/(n-1)$ converges in probability to σ^2. Prove that S_n^2 converges in probability to σ^2.

5.17. Let X_n and Y_n have a bivariate normal distribution with parameters μ_1, μ_2, σ_1^2, σ_2^2 (free of n) but $\rho = 1 - 1/n$. Consider the conditional distribution of Y_n, given $X_n = x$. Investigate the limit of this conditional distribution as $n \to \infty$. What is the limiting distribution if $\rho = -1 + 1/n$? Reference to these facts was made in the Remark in Section 2.3.

5.18. Let $\bar{X}_n$ denote the mean of a random sample of size n from a Poisson distribution with parameter $\mu = 1$.

(a) Show that the m.g.f. of $Y_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma = \sqrt{n}(\bar{X}_n - 1)$ is given by $\exp[-t\sqrt{n} + n(e^{t/\sqrt{n}} - 1)]$.

(b) Investigate the limiting distribution of Y_n as $n \to \infty$.
Hint: Replace, by its MacLaurin's series, the expression $e^{t/\sqrt{n}}$, which is in the exponent of the moment-generating function of Y_n.

5.19. Let $\bar{X}_n$ denote the mean of a random sample of size n from a distribution that has p.d.f. $f(x) = e^{-x}$, $0 < x < \infty$, zero elsewhere.

(a) Show that the m.g.f. $M(t; n)$ of $Y_n = \sqrt{n}(\bar{X}_n - 1)$ is equal to $[e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}]^{-n}$, $t < \sqrt{n}$.

(b) Find the limiting distribution of Y_n as $n \to \infty$.

This exercise and the immediately preceding one are special instances of an important theorem that will be proved in the next section.

5.4 The Central Limit Theorem

It was seen (Section 4.8) that, if $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution with mean μ and variance σ^2, the random variable

$$\frac{\sum_1^n X_i - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

is, for every positive integer n, normally distributed with zero mean and unit variance. In probability theory there is a very elegant theorem called the *central limit theorem*. A special case of this theorem asserts the remarkable and important fact that if $X_1, X_2, \ldots, X_n$ denote the observations of a random sample of size n from any distribution having positive variance σ^2 (and hence finite mean μ), then the random variable $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting standard normal distribution. If this fact can be established, it will imply, whenever the conditions of the theorem are satisfied, that (for large n) the random variable $\sqrt{n}(\bar{X} - \mu)/\sigma$ has an approximate normal distribution with mean zero and variance 1. It will then be possible to use this approximate normal distribution to compute approximate probabilities concerning $\bar{X}$.

The more general form of the theorem is stated, but it is proved only in the modified case. However, this is exactly the proof of the theorem that would be given if we could use the characteristic function in place of the m.g.f.

Theorem 3. *Let $X_1, X_2, \ldots, X_n$ denote the observations of a random sample from a distribution that has mean μ and positive variance σ^2. Then the random variable $Y_n = \left(\sum_1^n X_i - n\mu\right)\Big/\sqrt{n}\sigma = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting distribution that is normal with mean zero and variance* 1.

Proof. In the modification of the proof, we assume the existence of the m.g.f. $M(t) = E(e^{tX})$, $-h < t < h$, of the distribution. However, this proof is essentially the same one that would be given for this theorem in a more advanced course by replacing the m.g.f. by the characteristic function $\varphi(t) = E(e^{itX})$.

The function

$$m(t) = E[e^{t(X-\mu)}] = e^{-\mu t}M(t)$$

also exists for $-h < t < h$. Since $m(t)$ is the m.g.f. for $X - \mu$, it must follow that $m(0) = 1$, $m'(0) = E(X - \mu) = 0$, and $m''(0) = E[(X - \mu)^2] = \sigma^2$. By Taylor's formula there exists a number ξ between 0 and t such that

$$m(t) = m(0) + m'(0)t + \frac{m''(\xi)t^2}{2}$$

$$= 1 + \frac{m''(\xi)t^2}{2}.$$

If $\sigma^2 t^2/2$ is added and subtracted, then

$$m(t) = 1 + \frac{\sigma^2 t^2}{2} + \frac{[m''(\xi) - \sigma^2]t^2}{2}. \tag{1}$$

Next consider $M(t; n)$, where

$$\begin{aligned}
M(t; n) &= E\left[\exp\left(t\frac{\sum X_i - n\mu}{\sigma\sqrt{n}}\right)\right] \\
&= E\left[\exp\left(t\frac{X_1 - \mu}{\sigma\sqrt{n}}\right)\exp\left(t\frac{X_2 - \mu}{\sigma\sqrt{n}}\right)\cdots\exp\left(t\frac{X_n - \mu}{\sigma\sqrt{n}}\right)\right] \\
&= E\left[\exp\left(t\frac{X_1 - \mu}{\sigma\sqrt{n}}\right)\right]\cdots E\left[\exp\left(t\frac{X_n - \mu}{\sigma\sqrt{n}}\right)\right] \\
&= \left\{E\left[\exp\left(t\frac{X - \mu}{\sigma\sqrt{n}}\right)\right]\right\}^n \\
&= \left[m\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n, \quad -h < \frac{t}{\sigma\sqrt{n}} < h.
\end{aligned}$$

In Equation (1) replace t by $t/\sigma\sqrt{n}$ to obtain

$$m\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + \frac{[m''(\xi) - \sigma^2]t^2}{2n\sigma^2},$$

where now ξ is between 0 and $t/\sigma\sqrt{n}$ with $-h\sigma\sqrt{n} < t < h\sigma\sqrt{n}$. Accordingly,

$$M(t; n) = \left\{1 + \frac{t^2}{2n} + \frac{[m''(\xi) - \sigma^2]t^2}{2n\sigma^2}\right\}^n.$$

Since $m''(t)$ is continuous at $t = 0$ and since $\xi \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} [m''(\xi) - \sigma^2] = 0.$$

The limit proposition cited in Section 5.3 shows that

$$\lim_{n\to\infty} M(t; n) = e^{t^2/2}$$

for all real values of t. This proves that the random variable $Y_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting standard normal distribution.

We interpret this theorem as saying that, when n is a large, fixed positive integer, the random variable $\bar{X}$ has an approximate normal distribution with mean μ and variance σ^2/n; and in applications we use the approximate normal p.d.f. as though it were the exact p.d.f. of $\bar{X}$.

Some illustrative examples, here and later, will help show the importance of this version of the central limit theorem.

Example 1. Let $\bar{X}$ denote the mean of a random sample of size 75 from the distribution that has the p.d.f.

$$\begin{aligned} f(x) &= 1, \qquad 0 < x < 1, \\ &= 0 \qquad \text{elsewhere.} \end{aligned}$$

It was stated in Section 5.1 that the exact p.d.f. of $\bar{X}$, say $g(\bar{x})$, is rather complicated. It can be shown that $g(\bar{x})$ has a graph when $0 < \bar{x} < 1$ that is composed of arcs of 75 different polynomials of degree 74. The computation of such a probability as $\Pr(0.45 < \bar{X} < 0.55)$ would be extremely laborious. The conditions of the theorem are satisfied, since $M(t)$ exists for all real values of t. Moreover, $\mu = \frac{1}{2}$ and $\sigma^2 = \frac{1}{12}$, so that we have approximately

$$\begin{aligned} \Pr(0.45 < \bar{X} < 0.55) &= \Pr\left[\frac{\sqrt{n}(0.45 - \mu)}{\sigma} < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < \frac{\sqrt{n}(0.55 - \mu)}{\sigma}\right] \\ &= \Pr[-1.5 < 30(\bar{X} - 0.5) < 1.5] \\ &= 0.866, \end{aligned}$$

from Table III in Appendix B.

Example 2. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution that is $b(1, p)$. Here $\mu = p$, $\sigma^2 = p(1 - p)$, and $M(t)$ exists for all real values of t. If $Y_n = X_1 + \cdots + X_n$, it is known that Y_n is $b(n, p)$. Calculation of probabilities concerning Y_n, when we do not use the Poisson approximation, can be greatly simplified by making use of the fact that $(Y_n - np)/\sqrt{np(1 - p)} = \sqrt{n}(\bar{X}_n - p)/\sqrt{p(1 - p)} = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting distribution that is normal with mean zero and variance 1. Frequently, statisticians say that Y_n, or more simply Y, has an approximate normal distribution with mean np and variance $np(1 - p)$. Even with n as small as 10, with $p = \frac{1}{2}$ so that the binomial distribution is symmetric about $np = 5$, we note in Figure 5.1 how well the normal distribution, $N(5, \frac{5}{2})$, fits the binomial distribution, $b(10, \frac{1}{2})$, where the heights of the rectangles represent the probabilities of the respective integers $0, 1, 2, \ldots, 10$. Note that the area of the rectangle whose base is $(k - 0.5, k + 0.5)$ and the area under the normal p.d.f. between $k - 0.5$ and $k + 0.5$ are *approximately* equal for each $k = 0, 1, 2, \ldots, 10$, even with $n = 10$. This example should help the reader understand Example 3.

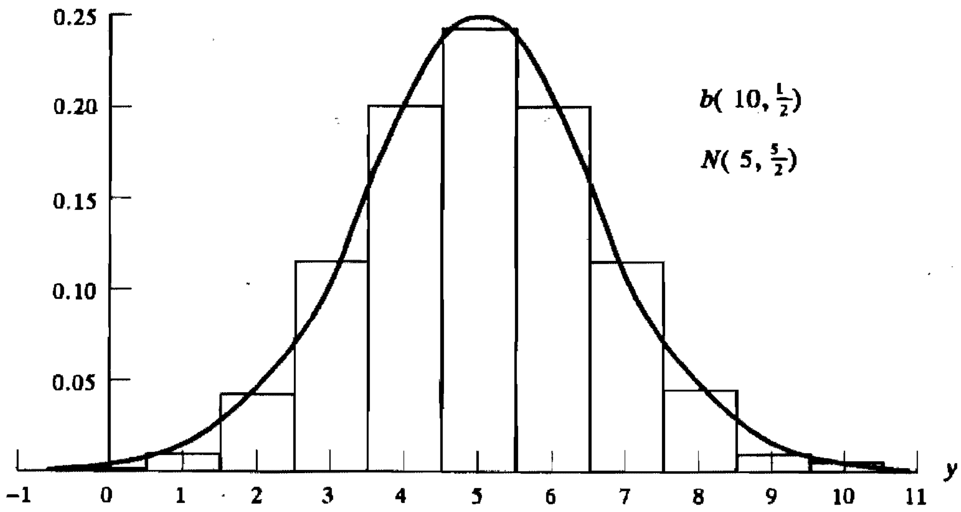

FIGURE 5.1

Example 3. With the background of Example 2, let $n = 100$ and $p = \frac{1}{2}$, and suppose that we wish to compute $\Pr(Y = 48, 49, 50, 51, 52)$. Since Y is a random variable of the discrete type, the events $Y = 48, 49, 50, 51, 52$ and $47.5 < Y < 52.5$ are equivalent. That is, $\Pr(Y = 48, 49, 50, 51, 52) = \Pr(47.5 < Y < 52.5)$. Since $np = 50$ and $np(1 - p) = 25$, the latter probability may be written

$$\Pr(47.5 < Y < 52.5) = \Pr\left(\frac{47.5 - 50}{5} < \frac{Y - 50}{5} < \frac{52.5 - 50}{5}\right)$$

$$= \Pr\left(-0.5 < \frac{Y - 50}{5} < 0.5\right).$$

Since $(Y - 50)/5$ has an approximate normal distribution with mean zero and variance 1, Table III shows this probability to be approximately 0.382.

The convention of selecting the event $47.5 < Y < 52.5$, instead of, say, $47.8 < Y < 52.3$, as the event equivalent to the event $Y = 48, 49, 50, 51, 52$ seems to have originated in the following manner: The probability, $\Pr(Y = 48, 49, 50, 51, 52)$, can be interpreted as the sum of five rectangular areas where the rectangles have bases 1 but the heights are, respectively, $\Pr(Y = 48), \ldots, \Pr(Y = 52)$. If these rectangles are so located that the midpoints of their bases are, respectively, at the points $48, 49, \ldots, 52$ on a horizontal axis, then in approximating the sum of these areas by an area bounded by the horizontal axis, the graph of a normal p.d.f., and two ordinates, it seems reasonable to take the two ordinates at the points 47.5 and 52.5.

We know that $\bar{X}$ and $\sum_{i=1}^{n} X_i$ have approximate normal distributions, provided that n is large enough. Later, we find that other statistics also have approximate normal distributions, and this is the reason that the normal distribution is so important to statisticians. That is, while not many underlying distributions are normal, the distributions of statistics calculated from random samples arising from these distributions are often very close to being normal.

Frequently, we are interested in functions of statistics that have approximate normal distributions. For illustration, Y_n of Example 2 has an approximate $N[np, np(1-p)]$. So $np(1-p)$ is an important function of p as it is the variance of Y_n. Thus, if p is unknown, we might want to estimate the variance of Y_n. Since $E(Y_n/n) = p$, we might use $n(Y_n/n)(1 - Y_n/n)$ as such an estimator and would want to know something about the latter's distribution. In particular, does it also have an approximate normal distribution? If so, what are its mean and variance? To answer questions like these, we use a procedure that is commonly called the *delta method*, which will be explained using the sample mean $\bar{X}_n$ as the statistic.

We know that $\bar{X}_n$ converges in probability to μ and $\bar{X}_n$ is approximately $N(\mu, \sigma^2/n)$. Suppose that we are interested in a function of $\bar{X}_n$, say $u(\bar{X}_n)$. Since, for large n, $\bar{X}_n$ is close to μ, we can approximate $u(\bar{X}_n)$ by the first two terms of Taylor's expansion about μ, namely

$$u(\bar{X}_n) \approx v(\bar{X}_n) = u(\mu) + (\bar{X}_n - \mu)u'(\mu),$$

where $u'(\mu)$ exists and is not zero. Since $v(\bar{X}_n)$ is a linear function of $\bar{X}_n$, it has an approximate normal distribution with mean

$$E[v(\bar{X}_n)] = u(\mu) + E[(\bar{X}_n - \mu)]u'(\mu) = u(\mu)$$

and variance

$$\text{var}\,[v(\bar{X}_n)] = [u'(\mu)]^2 \,\text{var}\,(\bar{X}_n - \mu) = [u'(\mu)]^2 \frac{\sigma^2}{n}.$$

Now, for large n, $u(\bar{X}_n)$ is approximately equal to $v(\bar{X}_n)$; so it has the same approximating distribution. That is, $u(\bar{X}_n)$ is approximately $N\{u(\mu), [u'(\mu)]^2\sigma^2/n\}$. More formally, we could say that

$$\frac{u(\bar{X}_n) - u(\mu)}{\sqrt{[u'(\mu)]^2\sigma^2/n}}$$

has a limiting standard normal distribution.

Example 4. Let Y_n (or Y for simplicity) be $b(n, p)$. Thus Y/n is approximately $N[p, p(1-p)/n)]$. Statisticians often look for functions of statistics whose variances do not depend upon the parameter. Here the variance of Y/n depends upon p. Can we find a function, say $u(Y/n)$, whose variance is essentially free of p? Since Y/n converges in probability to p, we can approximate $u(Y/n)$ by the first two terms of its Taylor's expansion about p, namely by

$$v\left(\frac{Y}{n}\right) = u(p) + \left(\frac{Y}{n} - p\right)u'(p).$$

Of course, $v(Y/n)$ is a linear function of Y/n and thus also has an approximate normal distribution; clearly, it has mean $u(p)$ and variance

$$[u'(p)]^2 \frac{p(1-p)}{n}.$$

But it is the latter that we want to be essentially free of p; thus we set it equal to a constant, obtaining the differential equation

$$u'(p) = \frac{c}{\sqrt{p(1-p)}}.$$

A solution of this is

$$u(p) = (2c)\arcsin \sqrt{p}.$$

If we take $c = \frac{1}{2}$, we have, since $u(Y/n)$ is approximately equal to $v(Y/n)$, that

$$u\left(\frac{Y}{n}\right) = \arcsin \sqrt{\frac{Y}{n}}.$$

This has an approximate normal distribution with mean $\arcsin \sqrt{p}$ and variance $1/4n$, which is free of p.

EXERCISES

5.20. Let $\bar{X}$ denote the mean of a random sample of size 100 from a distribution that is $\chi^2(50)$. Compute an approximate value of $\Pr(49 < \bar{X} < 51)$.

5.21. Let $\bar{X}$ denote the mean of a random sample of size 128 from a gamma distribution with $\alpha = 2$ and $\beta = 4$. Approximate $\Pr(7 < \bar{X} < 9)$.

5.22. Let Y be $b(72, \frac{1}{3})$. Approximate $\Pr(22 \le Y \le 28)$.

5.23. Compute an approximate probability that the mean of a random sample of size 15 from a distribution having p.d.f. $f(x) = 3x^2$, $0 < x < 1$, zero elsewhere, is between $\frac{3}{5}$ and $\frac{4}{5}$.

5.24. Let Y denote the sum of the observations of a random sample of size 12 from a distribution having p.d.f. $f(x) = \frac{1}{6}$, $x = 1, 2, 3, 4, 5, 6$, zero elsewhere. Compute an approximate value of $\Pr(36 \le Y \le 48)$.

Hint: Since the event of interest is $Y = 36, 37, \ldots, 48$, rewrite the probability as $\Pr(35.5 < Y < 48.5)$.

5.25. Let Y be $b(400, \frac{1}{5})$. Compute an approximate value of $\Pr(0.25 < Y/n)$.

5.26. If Y is $b(100, \frac{1}{2})$, approximate the value of $\Pr(Y = 50)$.

5.27. Let Y be $b(n, 0.55)$. Find the smallest value of n so that (approximately) $\Pr(Y/n > \frac{1}{2}) \geq 0.95$.

5.28. Let $f(x) = 1/x^2$, $1 < x < \infty$, zero elsewhere, be the p.d.f. of a random variable X. Consider a random sample of size 72 from the distribution having this p.d.f. Compute approximately the probability that more than 50 of the observations of the random sample are less than 3.

5.29. Forty-eight measurements are recorded to several decimal places. Each of these 48 numbers is rounded off to the nearest integer. The sum of the original 48 numbers is approximated by the sum of these integers. If we assume that the errors made by rounding off are i.i.d. and have uniform distributions over the interval $(-\frac{1}{2}, \frac{1}{2})$, compute approximately the probability that the sum of the integers is within 2 units of the true sum.

5.30. We know that $\bar{X}$ is approximately $N(\mu, \sigma^2/n)$ for large n. Find the approximate distribution of $u(\bar{X}) = \bar{X}^3$.

5.31. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean μ. Thus $Y = \sum_{i=1}^{n} X_i$ has a Poisson distribution with mean $n\mu$. Moreover, $\bar{X} = Y/n$ is approximately $N(\mu, \mu/n)$ for large n. Show that $u(Y/n) = \sqrt{Y/n}$ is a function of Y/n whose variance is essentially free of μ.

5.5 Some Theorems on Limiting Distributions

In this section, we shall present some theorems that can often be used to simplify the study of certain limiting distributions.

Theorem 4. *Let $F_n(u)$ denote the distribution function of a random variable U_n whose distribution depends upon the positive integer n. Let U_n converge in probability to the constant $c \neq 0$. The random variable U_n/c converges in probability to* 1.

The proof of this theorem is very easy and is left as an exercise.

Theorem 5. *Let $F_n(u)$ denote the distribution function of a random variable U_n whose distribution depends upon the positive integer n. Further, let U_n converge in probability to the positive constant c and let*

$\Pr(U_n < 0) = 0$ *for every n. The random variable* $\sqrt{U_n}$ *converges in probability to* $\sqrt{c}$.

Proof. We are given that the $\lim_{n\to\infty} \Pr(|U_n - c| \ge \epsilon) = 0$ for every $\epsilon > 0$.

We are to prove that the $\lim_{n\to\infty} \Pr(|\sqrt{U_n} - \sqrt{c}| \ge \epsilon') = 0$ for every $\epsilon' > 0$. Now the probability

$$\begin{aligned}\Pr(|U_n - c| \ge \epsilon) &= \Pr[|(\sqrt{U_n} - \sqrt{c})(\sqrt{U_n} + \sqrt{c})| \ge \epsilon] \\ &= \Pr\left(|\sqrt{U_n} - \sqrt{c}| \ge \frac{\epsilon}{\sqrt{U_n} + \sqrt{c}}\right) \\ &\ge \Pr\left(|\sqrt{U_n} - \sqrt{c}| \ge \frac{\epsilon}{\sqrt{c}}\right) \ge 0.\end{aligned}$$

If we let $\epsilon' = \epsilon/\sqrt{c}$, and if we take the limit, as n becomes infinite, we have

$$0 = \lim_{n\to\infty} \Pr(|U_n - c| \ge \epsilon) \ge \lim_{n\to\infty} \Pr(|\sqrt{U_n} - \sqrt{c}| \ge \epsilon') = 0$$

for every $\epsilon' > 0$. This completes the proof.

The conclusions of Theorems 4 and 5 are very natural ones and they certainly appeal to our intuition. There are many other theorems of this flavor in probability theory. As exercises, it is to be shown that if the random variables U_n and V_n converge in probability to the respective constants c and d, then $U_n V_n$ converges in probability to the constant cd, and U_n/V_n converges in probability to the constant c/d, provided that $d \ne 0$. However, we shall accept, without proof, the following theorem, which is a modification of Slutsky's theorem.

Theorem 6. *Let* $F_n(u)$ *denote the distribution function of a random variable* U_n *whose distribution depends upon the positive integer n. Let* U_n *have a limiting distribution with distribution function* $F(u)$. *Let a random variable* V_n *converge in probability to* 1. *The limiting distribution of the random variable* $W_n = U_n/V_n$ *is the same as that of* U_n; *that is,* W_n *has a limiting distribution with distribution function* $F(w)$.

Example 1. Let Y_n denote a random variable that is $b(n, p)$, $0 < p < 1$. We know that

$$U_n = \frac{Y_n - np}{\sqrt{np(1-p)}}$$

has a limiting distribution that is $N(0, 1)$. Moreover, it has been proved that Y_n/n and $1 - Y_n/n$ converge in probability to p and $1 - p$, respectively; thus $(Y_n/n)(1 - Y_n/n)$ converges in probability to $p(1 - p)$. Then, by Theorem 4, $(Y_n/n)(1 - Y_n/n)/[p(1 - p)]$ converges in probability to 1, and Theorem 5 asserts that the following does also:

$$V_n = \left[\frac{(Y_n/n)(1 - Y_n/n)}{p(1 - p)}\right]^{1/2}.$$

Thus, in accordance with Theorem 6, the ratio $W_n = U_n/V_n$, namely

$$\frac{Y_n - np}{\sqrt{n(Y_n/n)(1 - Y_n/n)}},$$

has a limiting distribution that is $N(0, 1)$. This fact enables us to write (with n a large, fixed positive integer)

$$\Pr\left[-2 < \frac{Y - np}{\sqrt{n(Y/n)(1 - Y/n)}} < 2\right] = 0.954,$$

approximately.

Example 2. Let $\bar{X}_n$ and S_n^2 denote, respectively, the mean and the variance of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$, $\sigma^2 > 0$. It has been proved that $\bar{X}_n$ converges in probability to μ and that S_n^2 converges in probability to σ^2. Theorem 5 asserts that S_n converges in probability to σ and Theorem 4 tells us that S_n/σ converges in probability to 1. In accordance with Theorem 6, the random variable $W_n = \sigma\bar{X}_n/S_n$ has the same limiting distribution as does $\bar{X}_n$. That is, $\sigma\bar{X}_n/S_n$ converges in probability to μ.

EXERCISES

5.32. Prove Theorem 4.
Hint: Note that $\Pr(|U_n/c - 1| < \epsilon) = \Pr(|U_n - c| < \epsilon|c|)$, for every $\epsilon > 0$. Then take $\epsilon' = \epsilon|c|$.

5.33. Let $\bar{X}_n$ denote the mean of a random sample of size n from a gamma distribution with parameters $\alpha = \mu > 0$ and $\beta = 1$. Show that the limiting distribution of $\sqrt{n}(\bar{X}_n - \mu)/\sqrt{\bar{X}_n}$ is $N(0, 1)$.

5.34. Let $T_n = (\bar{X}_n - \mu)/\sqrt{S_n^2/(n-1)}$, where $\bar{X}_n$ and S_n^2 represent, respectively, the mean and the variance of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. Prove that the limiting distribution of T_n is $N(0, 1)$.

5.35. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be the observations of two independent random samples, each of size n, from the distributions that have the respective means μ_1 and μ_2 and the common variance σ^2. Find the limiting distribution of

$$\frac{(\bar{X}_n - \bar{Y}_n) - (\mu_1 - \mu_2)}{\sigma\sqrt{2/n}},$$

where $\bar{X}_n$ and $\bar{Y}_n$ are the respective means of the samples.

Hint: Let $\bar{Z}_n = \sum_1^n Z_i/n$, where $Z_i = X_i - Y_i$.

5.36. Let U_n and V_n converge in probability to c and d, respectively. Prove the following.

(a) The sum $U_n + V_n$ converges in probability to $c + d$.
Hint: Show that $\Pr(|U_n + V_n - c - d| \geq \epsilon) \leq \Pr(|U_n - c| + |V_n - d| \geq \epsilon) \leq \Pr(|U_n - c| \geq \epsilon/2 \text{ or } |V_n - d| \geq \epsilon/2) \leq \Pr(|U_n - c| \geq \epsilon/2) + \Pr(|V_n - d| \geq \epsilon/2)$.

(b) The product $U_n V_n$ converges in probability to cd.

(c) If $d \neq 0$, the ratio U_n/V_n converges in probability to c/d.

5.37. Let U_n converge in probability to c. If $h(u)$ is a continuous function at $u = c$, prove that $h(U_n)$ converges in probability to $h(c)$.

Hint: For each $\epsilon > 0$, there exists a $\delta > 0$ such that $\Pr[|h(U_n) - h(c)| < \epsilon] \geq \Pr[|U_n - c| < \delta]$. Why?

ADDITIONAL EXERCISES

5.38. A nail manufacturer guarantees that not more than one nail in a box of 100 nails is defective. If, in fact, the probability of each individual nail being defective is $p = 0.005$, compute the probability that:

(a) The next box of nails violates the guarantee. Use the Poisson approximation, after assuming independence.

(b) The guarantee is violated at least once in the next 25 boxes.

5.39. Let $\bar{X}_n$ and $\bar{Y}_n$ be the means of two independent random samples of size n from a distribution having variance σ^2. Determine n so that $\Pr(|\bar{X}_n - \bar{Y}_n| \leq \sigma/2) = 0.98$, approximately.

5.40. Let $X_1, X_2, \ldots, X_{25}$ be a random sample from a distribution with p.d.f. $f(x) = 6x(1 - x)$, $0 < x < 1$, zero elsewhere. Find $\Pr[0.48 < \bar{X}_n < 0.52]$ approximately.

5.41. A rolls an unbiased die 100 independent times and B rolls an unbiased die 100 independent times. What is the approximate probability that A will total at least 25 points more than B?

5.42. Compute, approximately, the probability that the sum of the observations of a random sample of size 24 from a chi-square distribution with 3 degrees of freedom is between 70 and 80.

5.43. Let X be the number of times that no heads appear on two coins when these two coins are tossed together n times. Find the smallest value of n so that $\Pr(0.24 \le X/n \le 0.26) \ge 0.954$, approximately.

5.44. Two persons have 16 and 32 dollars, respectively. They bet one dollar on each of 900 independent tosses of an unbiased coin. What is an approximation to the probability that neither person is in debt at the end of the 900 trials?

5.45. A die is rolled 720 independent times. Compute, approximately, the probability that the number of fives that appear will be between 110 and 125 inclusive.

5.46. A part is produced with a mean of 6.2 ounces and a standard deviation of 0.2 ounce. What is the probability that the weight of 100 such items is between 616 and 624 ounces?

5.47. Let $X_1, \ldots, X_{25}$ be a random sample of size 25 from a distribution having p.d.f. $f(x) = x/6$, $x = 1, 2, 3$, zero elsewhere. Approximate

$$\Pr\left(\sum_{i=1}^{25} X_i = 50, 51, \ldots, \text{or } 60\right).$$

5.48. Say that a lot of 1000 items contains 20 defective items. A sample of size 50 is taken at random and without replacement from the lot. If 3 or fewer defective items are found in this sample, the lot is accepted. Approximate the probability of accepting this lot.

5.49. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having finite $E(X^m)$, $m > 0$. Show that $\sum_{i=1}^{n} X_i^m/n$ converges in probability to $E(X^m)$. Was an additional assumption needed?

5.50. It can be proved that the mean $\bar{X}_n$ of a random sample of size n from a Cauchy distribution has that same Cauchy distribution for every n. Thus $\bar{X}_n$ does not converge in probability to zero. How can this be, as earlier, under certain conditions, we proved that $\bar{X}_n$ converges in probability to the mean of the distribution?

5.51. Let Y be $\chi^2(n)$. What is the limiting distribution of $Z = \sqrt{Y} - \sqrt{n}$?

5.52. Let $\bar{X}$ be the mean of a random sample of size n from a Poisson distribution with parameter μ. Find the function $Y = u(\bar{X})$ so that Y has an approximate normal distribution with mean $u(\mu)$ and variance that is free of μ.

5.53. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution with distribution function $F(x)$ and p.d.f. $f(x) = F'(x)$. Say $F(\xi_p) = p$ and $f(\xi_p) > 0$. Consider the order statistic $Y_{[np]}$, where $[np]$ is the greatest integer in np.

(a) Note that the event $\sqrt{n}(Y_{[np]} - \xi_p) \le u$ is equivalent to $Z \ge [np]$, where Z is the number of X-values less than or equal to $\xi_p + u/\sqrt{n}$.

(b) Write $Z \ge np$, an approximation to $Z \ge [np]$, as

$$\frac{Z - nF(\xi_p + u/\sqrt{n})}{\sqrt{np(1-p)}} \ge \frac{-f(\xi_p)u}{\sqrt{p(1-p)}}, \quad \text{approximately,}$$

using $F(\xi_p + u/\sqrt{n}) \approx p + f(\xi_p)u/\sqrt{n}$.

(c) Since the left-hand member of the inequality in part (b) is approximately $N(0, 1)$, argue that $Y_{[np]}$ has an approximate normal distribution with mean ξ_p and variance $p(1-p)/n[f(\xi_p)]^2$.

CHAPTER 6

Introduction to Statistical Inference

6.1 Point Estimation

The first five chapters of this book deal with certain concepts and problems of probability theory. Throughout we have carefully distinguished between a sample space $\mathscr{C}$ of outcomes and the space $\mathscr{A}$ of one or more random variables defined on $\mathscr{C}$. With this chapter we begin a study of some problems in statistics and here we are more interested in the number (or numbers) by which an outcome is represented than we are in the outcome itself. Accordingly, we shall adopt a frequently used convention. We shall refer to a random variable X as the outcome of a random experiment and we shall refer to the space of X as the sample space. Were it not so awkward, we would call X the numerical outcome. Once the experiment has been performed and it is found that $X = x$, we shall call x the experimental value of X for that performance of the experiment.