5

Characteristic Functions

5.1 INTRODUCTION

In Chapter 2 we examined the problem of finding probabilities of the form $P\{(R_1, \ldots, R_n) \in B\}$, where $R_1, \ldots, R_n$ were random variables on a given probability space. If $(R_1, \ldots, R_n)$ has density f, then

$$P\{(R_1, \ldots, R_n) \in B\} = \int \cdots \int_B f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n$$

In general, the evaluation of integrals of this type is quite difficult, if it is possible at all. In this chapter we describe an approach to a particular class of problems, those involving sums of independent random variables, which avoids integration in n dimensions. The approach is similar in spirit to the application of Fourier or Laplace transforms to a differential equation.

Let R be a random variable on a given probability space. We introduce the *characteristic function* of R, defined by

$$M_R(u) = E(e^{-iuR}), \qquad u \text{ real} \tag{5.1.1}$$

Here we meet complex-valued random variables for the first time. A complex-valued random variable on $(\Omega, \mathscr{F}, P)$ is a function T from Ω to the complex numbers C, such that the real part T_1 and the imaginary part T_2 of T are (real-valued) random variables. Thus $T(\omega) = T_1(\omega) + iT_2(\omega)$,

$\omega \in \Omega$. We define the *expectation* of T as the complex number $E(T) = E(T_1) + iE(T_2)$; $E(T)$ is defined only if $E(T_1)$ and $E(T_2)$ are both finite. In the present case we have $M_R(u) = E(\cos uR) - iE(\sin uR)$; since the cosine and the sine are ≤ 1 in absolute value, all expectations are finite. Thus M_R is a function from the reals to the complex numbers. If R has density f_R we obtain

$$M_R(u) = \int_{-\infty}^{\infty} e^{-iux} f_R(x)\, dx \tag{5.1.2}$$

which is the *Fourier transform* of f_R.

It will be convenient in many computations to use a Laplace rather than a Fourier transform. The *generalized characteristic function* of R is defined by

$$N_R(s) = E(e^{-sR}) \qquad s \text{ complex}\dagger \tag{5.1.3}$$

$N_R(s)$ is defined only for those s such that $E(e^{-sR})$ is finite. If s is imaginary, that is, if $s = iu$, u real, then $N_R(s) = M_R(u)$, so that $N_R(s)$ is defined at least for s on the imaginary axis. There will be situations in which $N_R(s)$ is not defined for any s off the imaginary axis, and other situations in which $N_R(s)$ is defined for all s.

If R has density f_R, we obtain

$$N_R(s) = \int_{-\infty}^{\infty} e^{-sx} f_R(x)\, dx \tag{5.1.4}$$

This is the (two-sided) *Laplace transform* of f_R.

The basic fact about characteristic functions is the following.

Theorem 1. Let $R_1, \ldots, R_n$ be independent random variables on a given probability space, and let $R_0 = R_1 + \cdots + R_n$. If $N_{R_i}(s)$ is finite for all $i = 1, 2, \ldots, n$, then $N_{R_0}(s)$ is finite, and

$$N_{R_0}(s) = N_{R_1}(s) N_{R_2}(s) \cdots N_{R_n}(s)$$

In particular, if we set $s = iu$, we obtain

$$M_{R_0}(u) = M_{R_1}(u) M_{R_2}(u) \cdots M_{R_n}(u)$$

Thus the characteristic function of a sum of independent random variables is the product of the characteristic functions.

† In doing most of the examples in this chapter, the student will not come to grief if he regards s as a real variable and replaces statements such as "$a < \text{Re}\, s < b$" by "$a < s < b$." Also, a comment about notation. We have taken $E(e^{-iuR})$ as the definition of the characteristic function rather than the more usual $E(e^{iuR})$ in order to preserve a notational symmetry between Fourier and Laplace transforms ($\int e^{-sx} f(x)\, dx$, not $\int e^{sx} f(x)\, dx$, is the standard notation for Laplace transform). Since u ranges over all real numbers, this change is of no essential significance.

PROOF.

$$E(e^{-sR_0}) = E(e^{-s(R_1+\cdots+R_n)}) = E\left[\prod_{k=1}^{n} e^{-sR_k}\right] = \prod_{k=1}^{n} E(e^{-sR_k})$$

by independence.

We have glossed over one point in this argument. If we take $n = 2$ for simplicity, we have complex-valued random variables $V = V_1 + iV_2$ and $W = W_1 + iW_2$ ($V = e^{-sR_1}$, $W = e^{-sR_2}$), where, by Theorem 2 of Section 2.7, V_j and W_k are independent ($j, k = 1, 2$), and all expectations are finite. We must show that $E(VW) = E(V)E(W)$, which we have proved only in the case when V and W are real-valued and independent. However, there is no difficulty.

$$\begin{aligned} E(VW) &= E[V_1W_1 - V_2W_2 + i(V_1W_2 + V_2W_1)] \\ &= E(V_1)E(W_1) - E(V_2)E(W_2) + i(E(V_1)E(W_2) + E(V_2)E(W_1)) \\ &= [E(V_1) + iE(W_1)][E(V_2) + iE(W_2)] = E(V)E(W) \end{aligned}$$

The proof for arbitrary n is more cumbersome, but the idea is exactly the same.

Thus we may find the characteristic function of a sum of independent random variables without any n-dimensional integration. However, this technique will not be of value unless it is possible to recover the distribution function from the characteristic function. In fact we have the following result, which we shall not prove.

Theorem 2 (Correspondence Theorem). *If $M_{R_1}(u) = M_{R_2}(u)$ for all u, then*

$$F_{R_1}(x) = F_{R_2}(x) \qquad \text{for all } x$$

For computational purposes we need some facts about the Laplace transform. Let f be a piecewise continuous function from E^1 to E^1 (not necessarily a density) and L_f its Laplace transform:

$$L_f(s) = \int_{-\infty}^{\infty} f(x)e^{-sx}\,dx$$

Laplace Transform Properties

1. If there are real numbers K_1 and K_2 and nonnegative real numbers A_1 and A_2 such that $|f(x)| \le A_1e^{K_1x}$ for $x \ge 0$, and $|f(x)| \le A_2e^{K_2x}$ for $x \le 0$, then $L_f(s)$ is finite for $K_1 < \operatorname{Re} s < K_2$. This follows, since

$$\int_0^{\infty} |f(x)e^{-sx}|\,dx \le \int_0^{\infty} A_1e^{(K_1-a)x}\,dx$$

and

$$\int_{-\infty}^{0} |f(x)e^{-sx}|\,dx \le \int_{-\infty}^{0} A_2e^{(K_2-a)x}\,dx$$

where $a = \operatorname{Re} s$. The integrals are finite if $K_1 < a < K_2$. Thus the class of functions whose Laplace transform can be taken is quite large.

2. If $g(x) = f(x - a)$ and L_f is finite at s, then L_g is also finite at s and $L_g(s) = e^{-as} L_f(s)$. This follows, since

$$\int_{-\infty}^{\infty} f(x - a)e^{-sx}\, dx = e^{-as}\int_{-\infty}^{\infty} f(x - a)e^{-s(x-a)}\, d(x - a)$$

3. If $h(x) = f(-x)$ and L_f is finite at s, then L_h is finite at $-s$ and $L_h(-s) = L_f(s)$ [or $L_h(s) = L_f(-s)$ if L_f is finite at $-s$]. To verify this, write

$$\int_{-\infty}^{\infty} h(x)e^{sx}\, dx = (\text{with } y = -x)\int_{-\infty}^{\infty} f(y)e^{-sy}\, dy$$

4. If $g(x) = e^{-ax}f(x)$ and L_f is finite at s, then L_g is finite at $s - a$ and $L_g(s - a) = L_f(s)$ [or $L_g(s) = L_f(s + a)$ if L_f is finite at $s + a$]. For

$$\int_{-\infty}^{\infty} e^{-ax}e^{-(s-a)x}f(x)\, dx = \int_{-\infty}^{\infty} e^{-sx}f(x)\, dx$$

We now construct a very brief table of Laplace transforms for use in the examples. In Table 5.1.1, $u(x)$ is the *unit step function*, defined by $u(x) = 1$,

Table 5.1.1 Laplace Transforms

$f(x)$		$L_f(s)$	Region of Convergence
$u(x)$		$1/s$	$\operatorname{Re} s > 0$
$e^{-ax}u(x)$		$1/(s + a)$	$\operatorname{Re} s > -a$
$x^n e^{-ax}u(x)$,	$n = 0, 1, \ldots$	$n!/(s + a)^{n+1}$	$\operatorname{Re} s > -a$
$x^\alpha e^{-ax}u(x)$,	$\alpha > -1$	$\Gamma(\alpha + 1)/(s + a)^{\alpha+1}$	$\operatorname{Re} s > -a$

$x \geq 0$; $u(x) = 0$, $x < 0$. If we verify the last entry in the table the others will follow. Now

$$\int_0^{\infty} x^\alpha e^{-ax}e^{-sx}\, dx = [\text{with } y = (s + a)x]\int_0^{\infty} \frac{y^\alpha e^{-y}}{(s + a)^{\alpha+1}}\, dy$$

$$= \frac{\Gamma(\alpha + 1)}{(s + a)^{\alpha+1}}\dagger$$

† Strictly speaking, these manipulations are only valid for s *real* and $> -a$. However, one can show that under the hypothesis of property 1 L_f is analytic for $K_1 < \operatorname{Re} s < K_2$. In the present case $K_1 = -a$ and K_2 can be taken arbitrarily large, so that L_f is analytic for $\operatorname{Re} s > -a$. Now $L_f(s) = \Gamma(\alpha + 1)/(s + a)^{\alpha+1}$ for s real and $> -a$, and therefore, by the identity theorem for analytic functions, the formula holds for all s with $\operatorname{Re} s > -a$. This technique, which allows one to treat certain complex integrals as if the integrands were real-valued, will be used several times without further comment.

REMARK. $u(x)$ and $-u(-x)$ have the same Laplace transform $1/s$, but the regions of convergence are disjoint:

$$\int_{-\infty}^{\infty} u(x)e^{-sx}\,dx = \int_{0}^{\infty} e^{-sx}\,dx = \frac{1}{s}, \qquad \text{Re } s > 0$$

and

$$\int_{-\infty}^{\infty} -u(-x)e^{-sx}\,dx = \int_{-\infty}^{0} -e^{-sx}\,dx = \frac{1}{s}, \qquad \text{Re } s < 0$$

This indicates that any statement about Laplace transforms should be accompanied by some information about the region of convergence.

We need the following result in doing examples; the proof is measure-theoretic and will be omitted.

5. Let R be an absolutely continuous random variable. If h is a nonnegative (piecewise continuous) function and $L_h(s)$ is finite and coincides with the generalized characteristic function $N_R(s)$ for all s on the line Re $s = a$, then h is the density of R.

5.2 EXAMPLES

We are going to examine some typical problems involving sums of independent random variables. We shall use the result, to be justified in Example 6, that if *$R_1, R_2, \ldots, R_n$ are independent, each absolutely continuous, then $R_1 + \cdots + R_n$ is also absolutely continuous.*

In all examples $N_i(s)$ will denote the generalized characteristic function of the random variable R_i.

▶ **Example 1.** Let R_1 and R_2 be independent random variables, with R_1 uniformly distributed between -1 and $+1$, and R_2 having the exponential density $e^{-y}u(y)$. Find the density of $R_0 = R_1 + R_2$.

We have

$$N_1(s) = \int_{-1}^{1} \tfrac{1}{2}e^{-sx}\,dx = \frac{1}{2s}(e^{s} - e^{-s}), \qquad \text{all } s$$

$$N_2(s) = \int_{0}^{\infty} e^{-sy}e^{-y}\,dy = \frac{1}{s+1} \qquad \text{Re } s > -1$$

Thus, by Theorem 1 of Section 5.1,

$$N_0(s) = N_1(s)N_2(s) = \frac{1}{2s(s+1)}(e^{s} - e^{-s})$$

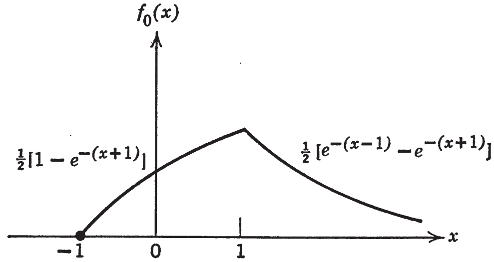

FIGURE 5.2.1

at least for Re $s > -1$. To find a function with this Laplace transform, we use *partial fraction expansion* of the rational function part of $N_0(s)$:

$$\frac{1}{2s(s+1)} = \frac{1}{2s} - \frac{1}{2(s+1)}$$

Now from Table 5.1.1, $u(x)$ has transform $1/s$ (Re $s > 0$) and $e^{-x}u(x)$ has transform $1/(s+1)$ (Re $s > -1$). Thus $(1/2)(1 - e^{-x})\ u(x)$ has transform $1/2s(s+1)$ (Re $s > 0$). By property 2 of Laplace transforms (Section 5.1), $(1/2)(1 - e^{-(x+1)})u(x+1)$ has transform $e^s/2s(s+1)$ and $(1/2)(1 - e^{-(x-1)})u(x-1)$ has transform $e^{-s}/2s(s+1)$ (Re $s > 0$). Thus a function h whose transform is $N_0(s)$ for Re $s > 0$ is

$$h(x) = \tfrac{1}{2}(1 - e^{-(x+1)})u(x+1) - \tfrac{1}{2}(1 - e^{-(x-1)})u(x-1)$$

By property 5 of Laplace transforms, h is the density of R_0; for a sketch, see Figure 5.2.1. ◀

▶ **Example 2.** Let $R_0 = R_1 + R_2 + R_3$, where R_1, R_2, and R_3 are independent with densities $f_1(x) = f_2(x) = e^x u(-x), f_3(x) = e^{-(x-1)}u(x-1)$. Find the density of R_0.

We have

$$N_1(s) = N_2(s) = \int_{-\infty}^{0} e^{x} e^{-sx}\, dx = \frac{1}{1-s}, \qquad \text{Re } s < 1$$

and

$$N_3(s) = \int_{1}^{\infty} e^{-(x-1)} e^{-sx}\, dx = \frac{e^{-s}}{s+1}, \qquad \text{Re } s > -1$$

Thus

$$N_0(s) = N_1(s)N_2(s)N_3(s) = \frac{e^{-s}}{(s-1)^2(s+1)}, \qquad -1 < \text{Re } s < 1$$

We expand the rational function in partial fractions.

$$\frac{1}{(s-1)^2(s+1)} = \frac{A}{(s-1)^2} + \frac{B}{s-1} + \frac{C}{s+1}$$

The coefficients may be found as follows.

$$A = [(s-1)^2G(s)]_{s=1} = \tfrac{1}{2}$$

$$B = \left[\frac{d}{ds}((s-1)^2G(s))\right]_{s=1} = -\tfrac{1}{4}$$

$$C = [(s+1)G(s)]_{s=-1} = \tfrac{1}{4}$$

From Table 5.1.1, the transform of $xe^{-x}u(x)$ is $1/(s+1)^2$, $\operatorname{Re} s > -1$. By Laplace transform property 3, the transform of $-xe^{x}u(-x)$ is $1/(1-s)^2$, $\operatorname{Re} s < 1$. The transform of $e^{-x}u(x)$ is $1/(s+1)$, $\operatorname{Re} s > -1$, so that, again by property 3, the transform of $e^{x}u(-x)$ is $1/(1-s)$, $\operatorname{Re} s < 1$.

Thus the transform of

$$-\tfrac{1}{2}xe^{x}u(-x) + \tfrac{1}{4}e^{x}u(-x) + \tfrac{1}{4}e^{-x}u(x)$$

is

$$G(s) = \frac{1/2}{(s-1)^2} - \frac{1/4}{s-1} + \frac{1/4}{s+1}, \qquad -1 < \operatorname{Re} s < 1$$

By property 2, the transform of

$$h(x) = [\tfrac{1}{4} - \tfrac{1}{2}(x-1)]e^{x-1}u(-(x-1)) + \tfrac{1}{4}e^{-(x-1)}u(x-1)$$

is

$$e^{-s}G(s) = N_0(s), \qquad -1 < \operatorname{Re} s < 1$$

By property 5, h is the density of R_0 (see Figure 5.2.2). ◄

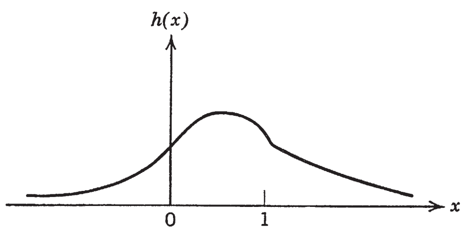

FIGURE 5.2.2

$$\begin{aligned} h(x) &= (\tfrac{1}{4} + \tfrac{1}{2}(1-x))e^{x-1}, \qquad x < 1 \\ &= \tfrac{1}{4}e^{-(x-1)}, \qquad x \geq 1 \end{aligned}$$

▶ **Example 3.** Let R have the *Cauchy density;* that is,

$$f_R(x) = \frac{1}{\pi(1 + x^2)}, \qquad -\infty < x < \infty$$

The characteristic function of R is

$$M_R(u) = \int_{-\infty}^{\infty} e^{-iux} f_R(x)\, dx$$

[In this case $N_R(s)$ is finite only for s on the imaginary axis.] $M_R(u)$ turns out to be $e^{-|u|}$. This may be verified by complex variable methods (see Problem 9), but instead we give a rough sketch of another attack. If the characteristic function of a random variable R is integrable, that is,

$$\int_{-\infty}^{\infty} |M_R(u)|\, du < \infty$$

it turns out that R has a density and in fact f_R is given by the *inverse Fourier transform.*

$$f_R(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M_R(u) e^{iux}\, du \tag{5.2.1}$$

In the present case

$$\int_{-\infty}^{\infty} e^{-|u|}\, du = \int_{-\infty}^{0} e^{u}\, du + \int_{0}^{\infty} e^{-u}\, du = 2 < \infty$$

and thus the density corresponding to $e^{-|u|}$ is

$$\begin{aligned}\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-|u|} e^{iux}\, du &= \frac{1}{2\pi} \int_{-\infty}^{0} e^{u(1+ix)}\, du + \frac{1}{2\pi} \int_{0}^{\infty} e^{-u(1-ix)}\, du \\ &= \frac{1}{2\pi}\left[\frac{1}{1 + ix} + \frac{1}{1 - ix}\right] = \frac{1}{\pi(1 + x^2)}\end{aligned}$$

Thus the Cauchy density in fact corresponds to the characteristic function $e^{-|u|}$.

This argument has a serious gap. We started with the assumption that $e^{-|u|}$ was the characteristic function of some random variable, and deduced from this that the random variable must have density $1/\pi(1 + x^2)$. We must establish that $e^{-|u|}$ is in fact a characteristic function (see Problem 8).

Now let $R_0 = R_1 + \cdots + R_n$, where the R_i are independent, each with the Cauchy density. Let us find the density of R_0. We have

$$M_0(u) = M_1(u) M_2(u) \cdots M_n(u) = (e^{-|u|})^n = e^{-n|u|}$$

If instead we consider R_0/n, we obtain

$$M_{R_0/n}(u) = E[e^{-iuR_0/n}] = M_0\left(\frac{u}{n}\right) = e^{-|u|}$$

Thus R_0/n has the Cauchy density. Now if $R_2 = nR_1$, then $f_2(y) = (1/n)f_1(y/n)$ (see Section 2.4), and so the density of R_0 is

$$f_0(y) = \frac{1}{n\pi(1 + y^2/n^2)} = \frac{n}{\pi(y^2 + n^2)}$$

REMARKS.

1. The arithmetic average R_0/n of a sequence of independent Cauchy distributed random variables has the same density as each of the components. There is no convergence of the arithmetic average to a constant, as we might expect physically. The trouble is that $E(R)$ does not exist.

2. If R has the Cauchy density and $R_1 = c_1R$, $R_2 = c_2R$, c_1, c_2 constant and > 0, then

$$M_1(u) = E(e^{-iuR_1}) = E(e^{-iuc_1R}) = M_R(c_1u) = e^{-c_1|u|}$$

and similarly

$$M_2(u) = e^{-c_2|u|}$$

Thus, if $R_0 = R_1 + R_2 = (c_1 + c_2)R$,

$$M_0(u) = e^{-(c_1+c_2)|u|}$$

which happens to be $M_1(u)M_2(u)$. This shows that if the characteristic function of the sum of two random variables is the product of the characteristic functions, the random variables need not be independent.

3. If R has the Cauchy density and $R_1 = \theta R$, $\theta > 0$, then by the calculation performed before Remark 1, R_1 has density $f_1(y) = \theta/\pi(y^2 + \theta^2)$ and (as in Remark 2) characteristic function $M_1(u) = e^{-\theta|u|}$. A random variable with this density is said to be of the *Cauchy type with parameter* θ or to have the *Cauchy density with parameter* θ. The formula for $M_1(u)$ shows immediately that if $R_1, \dots, R_n$ are independent and R_i is of the Cauchy type with parameter θ_i, $i = 1, \dots, n$, then $R_1 + \cdots + R_n$ is of the Cauchy type with parameter $\theta_1 + \cdots + \theta_n$. ◄

► **Example 4.** If $R_1, R_2, \dots, R_n$ are independent and normally distributed, then $R_0 = R_1 + \cdots + R_n$ is also normally distributed.

We first show that if R is normally distributed with mean m and variance σ^2, then

$$N_R(s) = e^{-sm}e^{s^2\sigma^2/2} \qquad \text{(all } s\text{)} \tag{5.2.2}$$

Now

$$N_R(s) = \int_{-\infty}^{\infty} e^{-sx} f_R(x)\, dx = \int_{-\infty}^{\infty} e^{-sx} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-m)^2/2\sigma^2}\, dx$$

Let $y = (x - m)/\sqrt{2}\,\sigma$ and complete the square to obtain

$$N_R(s) = \frac{1}{\sqrt{\pi}} e^{-sm} \int_{-\infty}^{\infty} \exp\left[-\left(y^2 + s\sqrt{2}\,\sigma y + \frac{s^2\sigma^2}{2}\right)\right] e^{s^2\sigma^2/2}\, dy$$

$$= \frac{1}{\sqrt{\pi}} e^{-sm} e^{s^2\sigma^2/2} \int_{-\infty}^{\infty} e^{-t^2}\, dt = e^{-sm}e^{s^2\sigma^2/2} \qquad \text{by (2.8.2)}$$

(See the footnote on page 157.) Now if $E(R_i) = m_i$, Var $R_i = \sigma_i^2$, then

$$N_0(s) = N_1(s)N_2(s)\cdots N_n(s) = e^{-s(m_1+\cdots+m_n)}\, e^{s^2(\sigma_1^2+\cdots+\sigma_n^2)/2}$$

But this is the characteristic function of a normally distributed random variable, and the result follows. Note that $m_0 = m_1 + \cdots + m_n$, $\sigma_0^2 = \sigma_1^2 + \cdots + \sigma_n^2$, as we should expect from the results of Section 3.3. ◀

▶ **Example 5.** Let R have the Poisson distribution.

$$p_R(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, \ldots$$

We first show that the generalized characteristic function of R is

$$N_R(s) = \exp\left[\lambda(e^{-s} - 1)\right] \qquad \text{(all } s\text{)} \tag{5.2.3}$$

We have

$$N_R(s) = E(e^{-sR}) = \sum_{k=0}^{\infty} e^{-sk} p_R(k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} e^{-sk}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{-s})^k}{k!} = e^{-\lambda} \exp(\lambda e^{-s})$$

as asserted.

We now show that if $R_1, \ldots, R_n$ are independent random variables, each with the Poisson distribution, then $R_0 = R_1 + \cdots + R_n$ also has the poisson distribution.

If R_i has the Poisson distribution with parameter λ_i, then

$$N_0(s) = N_1(s)N_2(s)\cdots N_n(s) = \exp\left[(\lambda_1 + \cdots + \lambda_n)(e^{-s} - 1)\right]$$

This is the characteristic function of a Poisson random variable, and the result follows. Note that if R has the Poisson distribution with parameter

λ, then $E(R) = \operatorname{Var} R = \lambda$ (see Problem 8, Section 3.2). Thus the result that the parameter of R_0 is $\lambda_1 + \cdots + \lambda_n$ is consistent with the fact that $E(R_0) = E(R_1) + \cdots + E(R_n)$ and $\operatorname{Var} R_0 = \operatorname{Var} R_1 + \cdots + \operatorname{Var} R_n$. ◀

▶ **Example 6.** In certain situations (especially when the Laplace transforms cannot be expressed in closed form) it may be convenient to use a convolution procedure rather than the transform technique to find the density of a sum of independent random variables. The method is based on the following result.

Convolution Theorem. *Let R_1 and R_2 be independent random variables, having densities f_1 and f_2, respectively. Let $R_0 = R_1 + R_2$. Then R_0 has a density given by*

$$f_0(z) = \int_{-\infty}^{\infty} f_2(z - x) f_1(x)\, dx = \int_{-\infty}^{\infty} f_1(z - y) f_2(y)\, dy \tag{5.2.4}$$

(Intuitively, the probability that R_1 lies in $(x, x + dx]$ is $f_1(x)\, dx$; given that $R_1 = x$, the probability that R_0 lies in $(z, z + dz]$ is the probability that R_2 lies in $(z - x, z - x + dz]$, namely, $f_2(z - x)\, dz$. Integrating with respect to x, we obtain the result that the probability that R_0 lies in $(z, z + dz]$ is

$$dz \int_{-\infty}^{\infty} f_2(z - x) f_1(x)\, dx$$

Since the probability is $f_0(z)\, dz$, (5.2.4) follows.)

PROOF. To prove the convolution theorem, observe that

$$\begin{aligned} F_0(z) = P\{R_1 + R_2 \le z\} &= \iint\limits_{x+y \le z} f_1(x) f_2(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f_2(y)\, dy \right] f_1(x)\, dx \end{aligned}$$

Let $y = u - x$ to obtain

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{z} f_2(u - x)\, du \right] f_1(x)\, dx = \int_{-\infty}^{z} \left[\int_{-\infty}^{\infty} f_1(x) f_2(u - x)\, dx \right] du$$

This proves the first relation of (5.2.4); the other follows by a symmetrical argument.

We consider a numerical example. Let $f_1(x) = 1/x^2$, $x \ge 1$; $f_1(x) = 0$, $x < 1$. Let $f_2(y) = 1$, $0 \le y \le 1$; $f_2(y) = 0$ elsewhere. If $z < 1$, $f_0(z) = 0$; if $1 \le z \le 2$,

$$f_0(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z - x)\, dx = \int_1^z \frac{1}{x^2}\, dx = 1 - \frac{1}{z}$$

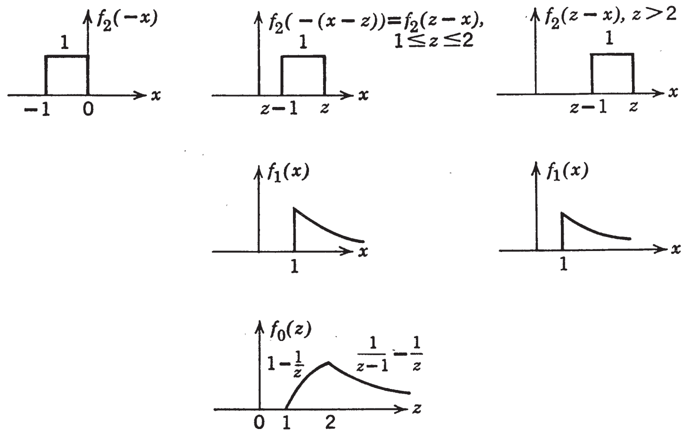

FIGURE 5.2.3 Application of the Convolution Theorem.

If $z > 2$,

$$f_0(z) = \int_{z-1}^{z} \frac{1}{x^2}\,dx = \frac{1}{z-1} - \frac{1}{z}$$

(see Figure 5.2.3).

REMARK. The successive application of the convolution theorem shows that if $R_1, \ldots, R_n$ are independent, each absolutely continuous, then $R_1 + \cdots + R_n$ is absolutely continuous. ◀

PROBLEMS

1. Let R_1, R_2, and R_3 be independent random variables, each uniformly distributed between -1 and $+1$. Find and sketch the density function of the random variable $R_0 = R_1 + R_2 + R_3$.
2. Two independent random variables R_1 and R_2 each have the density function $f(x) = 1/3$, $-1 \le x < 0$; $f(x) = 2/3$, $0 \le x \le 1$; $f(x) = 0$ elsewhere. Find and sketch the density function of $R_1 + R_2$.
3. Let $R = R_1^2 + \cdots + R_n^2$, where $R_1, \ldots, R_n$ are independent, and each R_i is normal with mean 0 and variance 1. Show that the density of R is

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)}\, x^{(n/2)-1} e^{-x/2}, \qquad x \ge 0$$

(R is said to have the "chi-square" distribution with n "degrees of freedom.")

4. A random variable R is said to have the "gamma distribution" if its density is, for some α, $\beta > 0$,

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, x \geq 0; f(x) = 0, x < 0$$

Show that if R_1 and R_2 are independent random variables, each having the gamma distribution with the same β, then $R_1 + R_2$ also has the gamma distribution.

5. If $R_1, \ldots, R_n$ are independent nonnegative random variables, each with density $\lambda e^{-\lambda x}u(x)$, find the density of $R_0 = R_1 + \cdots + R_n$.

6. Let θ be uniformly distributed between $-\pi/2$ and $\pi/2$. Show that $\tan \theta$ has the Cauchy density.

7. Let R have density $f(x) = 1 - |x|, |x| \leq 1$; $f(x) = 0$, $|x| > 1$. Show that $M_R(u) = 2(1 - \cos u)/u^2$.

***8.** (a) Suppose that f is the density of a random variable and the associated characteristic function M is real-valued, nonnegative, and integrable. Show that $kf(u)$, $-\infty < u < \infty$, is the characteristic function of a random variable with density $kM(x)/2\pi$, where k is chosen so that $kf(0) = 1$, that is,

$$\int_{-\infty}^{\infty} [kM(x)/2\pi]\, dx = 1$$

(b) Use part (a) to show that the following are characteristic functions of random variables: (i) $e^{-|u|}$, (ii) $M(u) = 1 - |u|$, $|u| \leq 1$; $M(u) = 0$, $|u| > 1$.

***9.** Use the calculus of residues to evaluate the characteristic function of the Cauchy density.

10. Calculate the characteristic function of the normal (0, 1) random variable as follows. Differentiate

$$M(u) = \int_{-\infty}^{\infty} (\cos ux) \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx$$

under the integral sign; then integrate by parts to obtain $M'(u) = -uM(u)$. Solve the resulting differential equation to obtain $M(u) = e^{-u^2/2}$. From this, find the characteristic function of a random variable that is normal with mean m and variance σ^2.

5.3 PROPERTIES OF CHARACTERISTIC FUNCTIONS

Let R be a random variable with characteristic function M and generalized characteristic function N. We shall establish several properties of M and N.

1. $M(0) = N(0) = 1$.
This follows, since $M(0) = N(0) = E(e^0)$.
2. $|M(u)| \leq 1$ for all u.
If R has a density f, we have

$$|M(u)| = \left|\int_{-\infty}^{\infty} e^{-iux} f(x)\, dx\right| \leq \int_{-\infty}^{\infty} |e^{-iux} f(x)|\, dx = \int_{-\infty}^{\infty} f(x)\, dx = 1$$

The general case can be handled by replacing $f(x)\, dx$ by $dF(x)$, where F is the distribution function of R. This involves Riemann-Stieltjes integration, which we shall not enter into here.

3. If R has a density f, and f is even, that is, $f(-x) = f(x)$ for all x, then $M(u)$ is real-valued for all u. For

$$M(u) = \int_{-\infty}^{\infty} f(x) \cos ux\, dx - i \int_{-\infty}^{\infty} f(x) \sin ux\, dx$$

Since $f(x)$ is an even function of x and $\sin ux$ is an odd function of x, $f(x) \sin ux$ is odd; hence the second integral is 0.

It turns out that the assertion that $M(u)$ is real for all u is equivalent to the statement that R has a symmetric distribution, that is, $P\{R \in B\} = P\{R \in -B\}$ for every Borel set B. ($-B = \{-x : x \in B\}$.)

4. If R is a discrete random variable taking on only integer values, then $M(u + 2\pi) = M(u)$ for all u.

To see this, write

$$M(u) = E(e^{-iuR}) = \sum_{n=-\infty}^{\infty} p_n e^{-iun} \tag{5.3.1}$$

where $p_n = P\{R = n\}$. Since $e^{-iun} = e^{-i(u+2\pi)n}$, the result follows.

Note that the p_n are the coefficients of the *Fourier series* of M on the interval $[0, 2\pi]$. If we multiply (5.3.1) by e^{iuk} and integrate, we obtain

$$p_k = \frac{1}{2\pi} \int_0^{2\pi} M(u) e^{iuk}\, du \tag{5.3.2}$$

We come now to the important *moment-generating property*. Suppose that $N(s)$ can be expanded in a power series about $s = 0$.

$$N(s) = \sum_{k=0}^{\infty} a_k s^k$$

where the series converges in some neighborhood of the origin. This is just the Taylor expansion of N; hence the coefficients must be given by

$$a_k = \frac{1}{k!} \left.\frac{d^k N(s)}{ds^k}\right]_{s=0}$$

But if R has density f and we can differentiate $N(s) = \int_{-\infty}^{\infty} e^{-sx} f(x)\, dx$ under the integral sign, we obtain $N'(s) = \int_{-\infty}^{\infty} -xe^{-sx} f(x)\, dx$; if we can differentiate k times, we find that

$$N^{(k)}(s) = \int_{-\infty}^{\infty} (-1)^k x^k e^{-sx} f(x)\, dx$$

Thus

$$N^{(k)}(0) = (-1)^k E(R^k) \tag{5.3.3}$$

and hence

$$a_k = \frac{(-1)^k}{k!} E(R^k)$$

The precise statement is as follows.

5. If $N_R(s)$ is analytic at $s = 0$ (i.e., expandable in a power series in a neighborhood of $s = 0$), then all moments of R are finite, and

$$N_R(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} E(R^k) s^k \tag{5.3.4}$$

within the radius of convergence of the series. In particular, (5.3.3) holds for all k.

We shall not give a proof of (5.3.4). The above remarks make it at least plausible; further evidence is presented by the following argument. If R has density f, then

$$\begin{aligned} N(s) &= \int_{-\infty}^{\infty} e^{-sx} f(x)\, dx \\ &= \int_{-\infty}^{\infty} \left(1 - sx + \frac{s^2x^2}{2!} - \frac{s^3x^3}{3!} + \cdots + \frac{(-1)^k s^k x^k}{k!} + \cdots\right) f(x)\, dx \end{aligned}$$

If we are allowed to integrate term by term, we obtain (5.3.4).

Let us verify (5.3.4) for a numerical example. Let $f(x) = e^{-x}u(x)$, so that $N(s) = 1/(s + 1)$, $\operatorname{Re} s > -1$. We have a power series expansion for $N(s)$ about $s = 0$.

$$\frac{1}{1 + s} = 1 - s + s^2 - s^3 + \cdots + (-1)^k s^k + \cdots \qquad |s| < 1$$

Equation 5.3.4 indicates that we should have $(-1)^k E(R^k)/k! = (-1)^k$, or $E(R^k) = k!$ To check this, notice that

$$E(R^k) = \int_0^{\infty} x^k e^{-x}\, dx = \Gamma(k + 1) = k!$$

Remark. Let R be a discrete random variable taking on only nonnegative integer values. In the generalized characteristic function

$$N(s) = \sum_{k=0}^{\infty} p_k e^{-sk}, \qquad p_k = P\{R = k\}$$

make the substitution $z = e^{-s}$. We obtain

$$A(z) = N(s)]_{z=e^{-s}} = E(z^R) = \sum_{k=0}^{\infty} p_k z^k$$

A is called the *generating function* of R; it is finite at least for $|z| \le 1$, since $\sum_{k=0}^{\infty} p_k = 1$.

We consider generating functions in detail in connection with the random walk problem in Chapter 6.

PROBLEMS

1. Could $[2/(s + 1)] - (1/s)(1 - e^{-s})(\operatorname{Re} s > 0)$ be the generalized characteristic function of an (absolutely continuous) random variable? Explain.
2. If the density of a random variable R is zero for $x \notin$ the finite interval $[a, b]$, show that $N_R(s)$ is finite for all s.
3. We have stated that if $M_R(u)$ is integrable, R has a density [see (5.2.1)]. Is the converse true?
4. Let R have a *lattice distribution*; that is, R is discrete and takes on the values $a + nd$, where a and d are fixed real numbers and n ranges over the integers. What can be said about the characteristic function of R?
5. If R has the Poisson distribution with parameter λ, calculate the mean and variance of R by differentiating $N_R(s)$.

5.4 THE CENTRAL LIMIT THEOREM

The weak law of large numbers states that if, for each n, $R_1, R_2, \dots, R_n$ are independent random variables with finite expectations and uniformly bounded variances, then, for every $\varepsilon > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}(R_i - ER_i)\right| \ge \varepsilon\right\} \to 0 \qquad \text{as } n \to \infty$$

In particular, if the R_i are independent observations of a random variable R (with finite mean m and finite variance σ^2), then

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} R_i - m\right| \geq \varepsilon\right\} \to 0 \qquad \text{as } n \to \infty$$

The central limit theorem gives further information; it says roughly that for large n, the sum $R_1 + \cdots + R_n$ of n independent random variables is approximately normally distributed, under wide conditions on the individual R_i.

To make the idea of "approximately normal" more precise, we need the notion of *convergence in distribution.* Let $R_1, R_2, \ldots$ be random variables with distribution functions $F_1, F_2, \ldots$, and let R be a random variable with distribution function F. We say that the sequence $R_1, R_2, \ldots$ converges in distribution to R (notation: $R_n \xrightarrow{d} R$) iff $F_n(x) \to F(x)$ at all points x at which F is continuous.

To see the reason for the restriction to continuity points of F, consider the following example.

▶ **Example 1.** Let R_n be uniformly distributed between 0 and $1/n$ (see Figure 5.4.1). Intuitively, as $n \to \infty$, R_n approximates more and more closely a random variable R that is identically 0. But $F_n(x) \to F(x)$ when $x \neq 0$, but not at $x = 0$, since $F_n(0) = 0$ for all n, and $F(0) = 1$. Since $x = 0$ is not a continuity point of F, we have $R_n \xrightarrow{d} R$. ◀

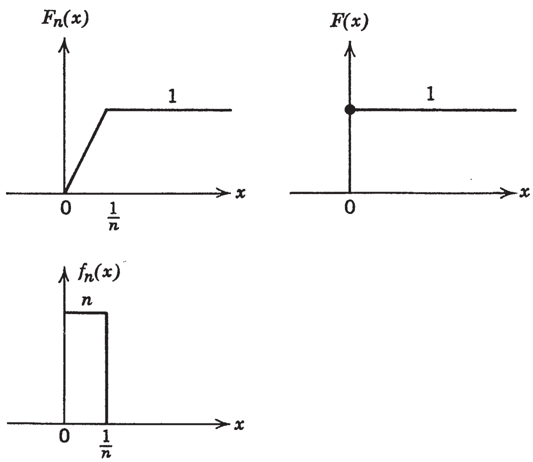

FIGURE 5.4.1 Convergence in Distribution.

REMARK. The type of convergence involved in the weak law of large numbers is called *convergence in probability*. The sequence R_1, $R_2, \ldots$ is said to converge in probability to R (notation: $R_n \xrightarrow{P} R$) iff for every $\varepsilon > 0$, $P\{|R_n - R| \geq \varepsilon\} \to 0$ as $n \to \infty$. Intuitively, for large n, R_n is very likely to be very close to R. Thus the weak law of large numbers states that $(1/n) \sum_{i=1}^{n} (R_i - E(R_i)) \xrightarrow{P} 0$; in the case in which $E(R_i) = m$ for all i, we have

$$\frac{1}{n} \sum_{i=1}^{n} R_i \xrightarrow{P} m$$

The relation between convergence in probability and convergence in distribution is outlined in Problem 1.

The basic result about convergence in distribution is the following.

Theorem 1. *The sequence $R_1, R_2, \ldots$ converges in distribution to R if and only if $M_n(u) \to M(u)$ for all u, where M_n is the characteristic function of R_n, and M is the characteristic function of R.*

The proof is measure-theoretic, and will be omitted.

Thus, in order to show that a sequence converges in distribution to a normal random variable, it suffices to show that the corresponding sequence of characteristic functions converges to a normal characteristic function. This is the technique that will be used to prove the main theorem, which we now state.

Theorem 2. (Central Limit Theorem). *For each n, let $R_1, R_2, \ldots, R_n$ be independent random variables on a given probability space. Assume that the R_i all have the same density function f (and characteristic function M) with finite mean m and finite variance $\sigma^2 > 0$, and finite third moment as well. Let*

$$T_n = \frac{\sum_{j=1}^{n} R_j - nm}{\sqrt{n}\,\sigma}$$

($= [S_n - E(S_n)]/\sigma(S_n)$, where $S_n = R_1 + \cdots + R_n$ and $\sigma(S_n)$ is the standard deviation of S_n) so that T_n has mean 0 and variance 1. Then T_1, $T_2, \ldots$ converge in distribution to a random variable that is normal with mean 0 and variance 1.

*Before giving the proof, we need some preliminaries.

Theorem 3. *Let f be a complex-valued function on E^1 with n continuous derivatives on the interval $V = (-b, b)$. Then, on V,*

$$f(u) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)u^k}{k!} + u^n \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!} f^{(n)}(tu)\, dt$$

Thus, if $|f^{(n)}| \le M$ on V,

$$f(u) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)u^k}{k!} + \frac{\theta\,|u|^n}{n!} \qquad \text{where } |\theta| \le M \ (\theta \text{ depends on } u)$$

PROOF. Using integration by parts, we obtain

$$\begin{aligned}
\int_0^u f^{(n)}(t) \frac{(u-t)^{n-1}}{(n-1)!}\, dt &= \left[\frac{(u-t)^{n-1}}{(n-1)!} f^{(n-1)}(t)\right]_0^u + \int_0^u f^{(n-1)}(t) \frac{(u-t)^{n-2}}{(n-2)!}\, dt \\
&= -\frac{f^{(n-1)}(0)u^{n-1}}{(n-1)!} + \int_0^u f^{(n-1)}(t) \frac{(u-t)^{n-2}}{(n-2)!}\, dt \\
&= -\frac{f^{(n-1)}(0)u^{n-1}}{(n-1)!} - \frac{f^{(n-2)}(0)u^{n-2}}{(n-2)!} \\
&\qquad + \int_0^u f^{(n-2)}(t) \frac{(u-t)^{n-3}}{(n-3)!}\, dt \\
&= -\sum_{k=1}^{n-1} \frac{f^{(k)}(0)u^k}{k!} + \int_0^u f'(t)\, dt \qquad \text{by iteration}
\end{aligned}$$

Thus

$$f(u) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)u^k}{k!} + \int_0^u f^{(n)}(t) \frac{(u-t)^{n-1}}{(n-1)!}\, dt$$

The change of variables $t = ut'$ in the above integral yields the desired expression for $f(u)$.

Now if

$$I = u^n \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!} f^{(n)}(tu)\, dt$$

then

$$|I| \le M\,|u|^n \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!}\, dt = \frac{M\,|u|^n}{n!}$$

Let $\theta = In!/|u|^n$; then $|\theta| \le M$ and the result follows.

Theorem 4.

$$\text{(a)} \quad e^{iy} = 1 + iy + \frac{\theta y^2}{2}, \qquad |\theta| \le 1$$

$$e^{iy} = 1 + iy - \frac{y^2}{2} + \frac{\theta_1 |y|^3}{6}, \qquad |\theta_1| \le 1$$

where y is an arbitrary real number, θ, θ_1 depending on y.

PROOF. This is immediate from Theorem 3.

(b) *If z is a complex number and $|z| \le 1/2$, then $\ln(1 + z) = z + \theta |z|^2$, where $|\theta| \le 1$, θ depending on z. (Take $\ln 1 = 0$ to determine the branch of the logarithm.)*

PROOF.

$$\ln(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots$$

$$= z + z^2\left(-\frac{1}{2} + \frac{z}{3} - \frac{z^2}{4} + \cdots\right)$$

Now

$$\left|-\tfrac{1}{2} + \tfrac{1}{3}z - \tfrac{1}{4}z^2 + \cdots\right| \le \tfrac{1}{2} + \tfrac{1}{3}(\tfrac{1}{2}) + \tfrac{1}{4}(\tfrac{1}{2})^2 + \cdots \le \sum_{k=1}^{\infty} (\tfrac{1}{2})^k = 1$$

Since $z^2 = |z|^2 e^{i\,arg(z^2)}$ and $|e^{i\,arg(z^2)}| = 1$, the result follows.

PROOF OF THEOREM 2. By Theorem 1, we must show that $M_{T_n}(u) \to e^{-u^2/2}$, the characteristic function of a normal (0, 1) random variable. Now we may assume without loss of generality that $m = 0$. For if we have proved the theorem under this assumption, then write

$$T_n = \frac{\sum_{j=1}^{n} (R_j - m)}{\sqrt{n}\,\sigma}$$

The random variables $R_j - m$ have mean 0 and variance σ^2, and the result will follow.

The characteristic function of $\sum_{j=1}^{n} R_j$ is $(M(u))^n$; hence the characteristic function of T_n is

$$E(e^{-iuT_n}) = E\left[\exp\left(-i\frac{u}{\sqrt{n}\,\sigma}\sum_{j=1}^{n} R_j\right)\right] = \left[M\left(\frac{u}{\sqrt{n}\,\sigma}\right)\right]^n$$

Now

$$M\left(\frac{u}{\sqrt{n}\,\sigma}\right) = \int_{-\infty}^{\infty} e^{-iux/\sqrt{n}\sigma} f(x)\,dx$$

$$= \int_{-\infty}^{\infty} \left(1 - \frac{iux}{\sqrt{n}\,\sigma} - \frac{u^2x^2}{2n\sigma^2} + \frac{\theta_1\,|u|^3\,|x|^3}{6n^{3/2}\sigma^3}\right) f(x)\,dx \quad \text{by Theorem 4a}$$

But

$$\int_{-\infty}^{\infty} xf(x)\,dx = m = 0, \qquad \int_{-\infty}^{\infty} x^2 f(x)\,dx = \sigma^2$$

Thus

$$\left[M\left(\frac{u}{\sqrt{n}\,\sigma}\right)\right]^n = \left(1 - \frac{u^2}{2n} + \frac{c}{n^{3/2}}\right)^n$$

where c depends on u. Take logarithms to obtain, by Theorem 4b,

$$n \ln\left(1 - \frac{u^2}{2n} + \frac{c}{n^{3/2}}\right) = n\left(-\frac{u^2}{2n} + \frac{c}{n^{3/2}} + \theta\left|-\frac{u^2}{2n} + \frac{c}{n^{3/2}}\right|^2\right) \to -\frac{u^2}{2}$$

as $n \to \infty$

Thus

$$\left[M\left(\frac{u}{\sqrt{n}\,\sigma}\right)\right]^n \to e^{-u^2/2}$$

which is the desired result.*

REMARK. Convergence in distribution of $T_n = [S_n - E(S_n)]/\sigma(S_n)$, $S_n = \sum_{j=1}^n R_j$, $\sigma(S_n)$ = standard deviation of S_n, can be established under conditions much more general than those given in Theorem 2. For example, the finiteness of the third moment is not necessary in Theorem 2; neither is the assumption that the R_i have a density. We give two other sufficient conditions for normal convergence.

1. The R_i are uniformly bounded; that is, there is a constant k such that $|R_i| \le k$ for all i, and also Var $(S_n) \to \infty$.

   The requirement that Var $S_n \to \infty$ is necessary, for otherwise we could take R_1 to have an arbitrary distribution function, and $R_n = 0$ for $n \ge 2$. Then $S_n = R_1$,

$$T_n = \frac{S_n - ES_n}{\sigma(S_n)} = \frac{R_1 - ER_1}{\sigma(R_1)}$$

   Thus the functions F_{T_n} are the same for all n and hence in general cannot approach

$$F^*(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$$

which is the normal distribution function with mean 0 and variance 1.

2. (The Liapounov condition)

$$\frac{\sum_{k=1}^{n} E\,|R_k - ER_k|^{2+\delta}}{|\sigma(S_n)|^{2+\delta}} \to 0 \qquad \text{for some } \delta > 0$$

REMARK. For each n, let $R_1, R_2, \ldots, R_n$ be independent random variables, where all R_i have the same distribution function, with finite mean m, finite variance σ^2, and also finite third moment. It can be shown that there is a positive number k such that $|F_{T_n}(x) - F^*(x)| \le k/\sqrt{n}$ for all x and all n. It follows that, for large n, S_n is approximately normal with mean nm and variance $n\sigma^2$, in the sense that for all x,

$$\left|F_{S_n}(x) - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi n}\,\sigma}\, e^{-(t-nm)^2/2n\sigma^2}\, dt\right| \le \frac{k}{\sqrt{n}}$$

For

$$F_{S_n}(x) = P\{S_n \le x\} = P\left\{\frac{S_n - nm}{\sqrt{n}\,\sigma} \le \frac{x - nm}{\sqrt{n}\,\sigma}\right\}$$

and this differs from $F^*((x - nm)/\sqrt{n}\,\sigma)$ by $\le k/\sqrt{n}$. But

$$\begin{aligned} F^*\left(\frac{x - nm}{\sqrt{n}\,\sigma}\right) &= \int_{-\infty}^{(x-nm)/\sqrt{n}\,\sigma} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt \\ &= \left(\text{set } t = \frac{y - nm}{\sqrt{n}\,\sigma}\right) \frac{1}{\sqrt{2\pi n}\,\sigma} \int_{-\infty}^{x} e^{-(y-nm)^2/2n\sigma^2}\, dy \end{aligned}$$

and the result follows.

In particular, let R be the number of successes in n Bernoulli trials; then (Example 1, Section 3.5) $R = R_1 + \cdots + R_n$, where the R_i are independent, and $P\{R_i = 1\} = p$, $P\{R_i = 0\} = 1 - p$. Thus, for large n, *R is approximately normal* with mean np and variance $np(1 - p)$, in the sense described above.

PROBLEMS

1. Show that $R_n \xrightarrow{P} R$ implies $R_n \xrightarrow{d} R$, as follows. Let F_n be the distribution function of R_n, and F the distribution function of R.

(a) If $\epsilon > 0$, show that

$$P\{R_n \le x\} \le P\{|R_n - R| \ge \epsilon\} + P\{R \le x + \epsilon\}$$

and

$$P\{R \le x - \epsilon\} \le P\{|R_n - R| \ge \epsilon\} + P\{R_n \le x\}$$

Conclude that

$$F(x - \epsilon) - P\{|R_n - R| \geq \epsilon\} \leq F_n(x) \leq P\{|R_n - R| \geq \epsilon\} + F(x + \epsilon)$$

(b) If $R_n \xrightarrow{P} R$ and F is continuous at x, show that $F_n(x) \to F(x)$.

2. Give an example of a sequence $R_1, R_2, \ldots$ that converges in distribution to a random variable R, but does not converge in probability to R.

3. If R_n converges in distribution to a *constant* c, that is, $\lim_{n\to\infty} F_n(x) = 1$ for $x > c$, and $= 0$ for $x < c$, show that $R_n \xrightarrow{P} c$.

4. Let $R_1, R_2, \ldots$ be random variables such that $P\{R_n = e^n\} = 1/n$, $P\{R_n = 0\} = 1 - 1/n$. Show that $R_n \xrightarrow{P} 0$, but $E(R_n{}^k) \to \infty$ as $n \to \infty$, for any fixed $k > 0$.

5. Two candidates, A and B, are running for President and it is desired to predict the outcome of the election. Assume that n people are selected independently and at random and asked their preference. Suppose that the probability of selecting a voter who favors A in any particular observation is p. (p is fixed but unknown.) Let Q_n be the relative frequency of "A" voters in the sample; that is,

$$Q_n = \frac{\text{number of "}A\text{" voters in sample}}{\text{size of sample}}$$

(a) We wish to choose n large enough so that $P\{|Q_n - p| \leq .001\} \geq .99$ for all possible values of p. In other words, we wish to predict A's percentage of the vote to within .1%, with 99% confidence. Estimate the minimum value of n.

(b) Estimate the minimum value of n if we wish to predict A's percentage to within 1%, with 95% confidence. (Use the central limit theorem.)

6. (a) Show that the normal density function (with mean 0, variance 1) satisfies the inequality

$$\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \leq \frac{1}{\sqrt{2\pi}\, x} e^{-x^2/2}, \qquad x > 0$$

HINT: show that

$$\frac{1}{\sqrt{2\pi} x} e^{-x^2/2} = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\left(1 + \frac{1}{t^2}\right) dt$$

by differentiating both sides.

(b) Show that

$$\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \sim \frac{1}{\sqrt{2\pi} x} e^{-x^2/2}$$

in the sense that the ratio of the two sides approaches 1 as $x \to \infty$.

7. Consider a container holding $n = 10^6$ molecules. In the steady state it is reasonable that there be roughly as many molecules on the left side as on the right. Assume that the molecules are dropped independently and at random into the

container and that each molecule may fall with equal probability on the left or right side.

If R is the number of molecules on the right side of the container, we may invoke the central limit theorem to justify the physical assumption that for the purpose of calculating $P\{a \leq R \leq b\}$ we may regard R as normally distributed with mean $np = n/2$ and variance $np(1 - p) = n/4$.

Use Problem 6 to bound $P\{|R - n/2| > .005n\}$, the probability of a fluctuation about the mean of more than $\pm.5\%$ of the total number of molecules.

8. Let R be the number of successes in 10,000 Bernoulli trials, with probability of success .8 on a given trial. Use the central limit theorem to estimate $P\{7940 \leq R \leq 8080\}$.