## 5.9 General approximation methods

Theoretical statistical computations frequently involve values of functions such as tail areas from common probability distributions such as the normal, chi-squared, and Student distributions, and their inverses (*percent points* or *quantiles*). With few exceptions, these functions do not have closed-form representations, so that one must make do with approximations that can be simply expressed. For the most common distributions there are many standard approximations now in use.

Closely related to the problem of approximating probabilities and percent points is the more general problem of approximating the distribution of a function of a random variable $X$ whose distribution is known. In this section, we discuss in barest outline a few general methods that can be used to obtain approximations that can be converted to useful computing formulæ. Our focus will be on general aspects of computing tail areas and their inverses. Section 5.10 will then deal with specific algorithms for a few of the most common distributions. For a more complete discussion, including a useful collection of algorithms, one should consult Kennedy and Gentle (1980).

In the remainder of this section, we shall employ the following notation. The random variable $X$ will be assumed to have a cumulative distribution function (*cdf*) given by $F_X(t) = \Pr\{X \leq t\}$, with a density function $f_X(t)$ with respect to Lebesgue measure. Of interest will be approximations to cdf $F_X(t)$, and its inverse, $F_X^{-1}(p)$, for $0 < p < 1$. When we wish to obtain either $F_X(t) = p$ for small values of $t$ (so that $p$ is close to zero), or values of the complementary cdf $G_X(t) \equiv 1 - F_X(t) = \int_t^\infty f_X(x)\,dx$ for large $t$, we speak of evaluating the *tail area* of $F_X$, and we treat this problem separately.

Kennedy and Gentle (1980) quite correctly point out that cdf's, tail areas, and percent points are often used as intermediate quantities in computations, and thus may require high accuracy. A very simple example comes from the literature of ranking and selection. Suppose that $X_1$ and $X_2$ are independent, with $X_1 \sim \mathcal{N}(\mu, 1)$ and $X_2 \sim \mathcal{N}(0, 1)$. Let $Y = \max(X_1, X_2)$. The density of $Y$ is given by $p(y) = \phi(y-\mu)\Phi(y) + \phi(y)\Phi(y-\mu)$, where $\phi(t)$ denotes the standard normal density and $\Phi(t)$ the corresponding distribution function. Computing $E(Y) = \int yp(y)\,dy$ to six decimal places requires evaluation of the integrand to *at least* six digits. Thus, for theoretical computations it is important to have highly accurate approximations for such fundamental building blocks as cdf's and tail areas, even though for such tasks as computing p-values in applied work much cruder approximations will suffice.

### 5.9.1 Cumulative distribution functions

The distribution function $\boxed{F_X(t) = \Pr\{X \leq t\}}$ can be written as the integral

$$p_b(t) \qquad F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx \qquad (5.9.1)$$

which suggests that quadrature methods can be used to approximate the integral. Indeed, numerical quadrature can often be a good choice, although for specific common distribution functions alternative approximations can be derived which either require less computational effort than quadrature of similar accuracy, or which have guaranteed (small) error bounds, which are generally not obtainable from the general quadrature theory.

Some general methods for obtaining approximations to integrals involve series expansions, continued fractions, and rational approximations. These methods are discussed in Section 5.9.4.

### 5.9.2 Tail areas

At first glance it would seem that computing the tail area $G(t)$ for large $t$ would be easy; given an approximation for $F(t)$ one could simply evaluate $1 - F(t)$. Excellent approximations exist, for instance, for the cdf of a normal random variable, so that normal tail areas would seem to be trivial to obtain. Unfortunately, life is not so simple. Consider single-precision floating-point approximation of $1 - \Phi(4) \approx 3.16713 \times 10^{-5}$, in a format that supplies five decimal digits. Then an accurate approximation to $\Phi(4)$ is 0.99997, the complement of which, $3 \times 10^{-5}$, is only good to a single decimal place. Arguments much larger than four produce answers with no significant digits!

$t = p_b^{-1}(s)$ on $[0,1]$.

### 5.9.3 Percent points

Let $p$ be a probability in $(0, 1)$, and let $F(x)$ denote the cumulative distribution function of a random variable $X$. A point $x_p$ for which $F(x_p) = p$ is called a $p$th quantile or fractile of $F$. If $F$ is a continuous monotone increasing function, then $x_p = F^{-1}(p)$.

A common approach for evaluating $x_p$ is to solve for $x$ in the nonlinear system $F(x) = p$ using any of the methods of Chapter 4. For many standard distributions the density $f(x) = F'(x)$ is available in closed form, and adequate approximations to $F$ itself are also at hand. In this case, Newton's method can be used effectively, provided only that a starting value sufficiently close to the answer can be obtained.

COMMENT. For small values of $p$, computing $F^{-1}$ is an ill-conditioned problem. This makes it very difficult to compute $F^{-1}$ to a fixed absolute accuracy. On the other hand, it is less difficult to compute percent points to a given *percentage* accuracy.