CHAPTER 3

Moments and Generating Functions

3.1 INTRODUCTION

The study of probability distributions of a random variable is essentially the study of some numerical characteristics associated with them. These *parameters of the distribution* play a key role in mathematical statistics. In Section 3.2 we introduce some of these parameters, namely, moments and order parameters, and investigate their properties. In Section 3.3 the idea of generating functions is introduced. In particular, we study probability generating functions, moment generating functions, and characteristic functions. In Section 3.4 we deal with some moment inequalities.

3.2 MOMENTS OF A DISTRIBUTION FUNCTION

In this section we investigate some numerical characteristics, called *parameters*, associated with the distribution of an RV X. These parameters are *moments* and their functions and *order parameters*. We concentrate mainly on moments and their properties.

Let X be a random variable of the discrete type with probability mass function $p_k = P\{X = x_k\}, k = 1, 2, \ldots$. If

$$\sum_{k=1}^{\infty} |x_k| p_k < \infty, \tag{1}$$

we say that the *expected value* (or the *mean* or the *mathematical expectation*) of X exists and write

$$\mu = EX = \sum_{k=1}^{\infty} x_k p_k. \tag{2}$$

Note that the series $\sum_{k=1}^{\infty} x_k p_k$ may converge but the series $\sum_{k=1}^{\infty} |x_k| p_k$ may not. In that case we say that EX does not exist.

Example 1. Let X have the PMF given by

$$p_j = P\left\{X = (-1)^{j+1}\frac{3^j}{j}\right\} = \frac{2}{3^j}, \qquad j = 1, 2, \ldots.$$

Then

$$\sum_{j=1}^{\infty} |x_j| p_j = \sum_{j=1}^{\infty} \frac{2}{j} = \infty,$$

and EX does not exist, although the series

$$\sum_{j=1}^{\infty} x_j p_j = \sum_{j=1}^{\infty} (-1)^{j+1}\frac{2}{j}$$

is convergent.

If X is of the continuous type and has PDF f, we say that EX exists and equals $\int x f(x)\,dx$, provided that

$$\int |x| f(x)\,dx < \infty.$$

A similar definition is given for the mean of any Borel-measurable function $h(X)$ of X. Thus if X is of the continuous type and has PDF f, we say that $Eh(X)$ exists and equals $\int h(x) f(x)\,dx$, provided that

$$\int |h(x)| f(x)\,dx < \infty.$$

We emphasize that the condition $\int |x| f(x)\,dx < \infty$ must be checked before it can be concluded that EX exists and equals $\int x f(x)\,dx$. Moreover, it is worthwhile to recall at this point that the integral $\int_{-\infty}^{\infty} \varphi(x)\,dx$ exists, provided that the limit $\lim_{b\to\infty}^{a\to\infty} \int_{-b}^{a} \varphi(x)\,dx$ exists. It is quite possible for the limit $\lim_{a\to\infty} \int_{-a}^{a} \varphi(x)\,dx$ to exist without the existence of $\int_{-\infty}^{\infty} \varphi(x)\,dx$. As an example, consider the Cauchy PDF:

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

Clearly,

$$\lim_{a\to\infty} \int_{-a}^{a} \frac{x}{\pi}\frac{1}{1+x^2}\,dx = 0.$$

However, EX does not exist since the integral $(1/\pi)\int_{-\infty}^{\infty} |x|/(1+x^2)\,dx$ diverges.

Remark 1. Let $X(\omega) = I_A(\omega)$ for some $A \in \mathcal{S}$. Then $EX = P(A)$.

Remark 2. If we write $h(X) = |X|$, we see that EX exists if and only if $E|X|$ does.

Remark 3. We say that an RV X is *symmetric* about a point α if

$$P\{X \geq \alpha + x\} = P\{X \leq \alpha - x\} \qquad \text{for all } x.$$

In terms of DF F of X, this means that if

$$F(\alpha - x) = 1 - F(\alpha + x) + P\{X = \alpha + x\}$$

holds for all $x \in \mathcal{R}$, we say that the DF F (or the RV X) is symmetric with α as the *center* of *symmetry*. If $\alpha = 0$, then for every x,

$$F(-x) = 1 - F(x) + P\{X = x\}.$$

In particular, if X is an RV of the continuous type, X is symmetric with center α if and only if the PDF f of X satisfies

$$f(\alpha - x) = f(\alpha + x) \qquad \text{for all } x.$$

If $\alpha = 0$, we will say simply that X is symmetric (or that F is symmetric).

As an immediate consequence of this definition we see that if X is symmetric with α as the center of symmetry and $E|X| < \infty$, then $EX = \alpha$. A simple example of a symmetric distribution is the Cauchy PDF considered above (before Remark 1). We will encounter many such distributions later.

Remark 4. If a and b are constants and X is an RV with $E|X| < \infty$, then $E|aX + b| < \infty$ and $E\{aX + b\} = aEX + b$. In particular, $E\{X - \mu\} = 0$, a fact that should not come as a surprise.

Remark 5. If X is bounded, that is, if $P\{|X| < M\} = 1, 0 < M < \infty$, then EX exists.

Remark 6. If $\{X \geq 0\} = 1$ and EX exists, then $EX \geq 0$.

Theorem 1. Let X be an RV and g be a Borel-measurable function on $\mathcal{R}$. Let $Y = g(X)$. If X is of discrete type, then

$$EY = \sum_{j=1}^{\infty} g(x_j)P\{X = x_j\} \tag{3}$$

in the sense that if either side of (3) exists, so does the other, and then the two are equal. If X is of continuous type with PDF f, then $EY = \int g(x)f(x)\,dx$ in the sense that if either of the two integrals converges absolutely, so does the other, and the two are equal.

Remark 7. Let X be a discrete RV. Then Theorem 1 says that

$$\sum_{j=1}^{\infty} g(x_j) P\{X = x_j\} = \sum_{k=1}^{\infty} y_k P\{Y = y_k\}$$

in the sense that if either of the two series converges absolutely, so does the other, and the two sums are equal. If X is of the continuous type with PDF f, let $h(y)$ be the PDF of $Y = g(X)$. Then, according to Theorem 1,

$$\int g(x) f(x)\, dx = \int y\, h(y)\, dy,$$

provided that $E|g(X)| < \infty$.

Proof of Theorem 1. In the discrete case, suppose that $P\{X \in A\} = 1$. If $y = g(x)$ is a one-to-one mapping of A onto some set B, then

$$P\{Y = y\} = P\{X = g^{-1}(y)\}, \qquad y \in B.$$

We have

$$\sum_{x \in A} g(x) P\{X = x\} = \sum_{y \in B} y P\{Y = y\}.$$

In the continuous case, suppose that g satisfies the conditions of Theorem 2.5.3. Then

$$\int g(x) f(x)\, dx = \int_{\alpha}^{\beta} y f[g^{-1}(y)] \frac{d}{dy} g^{-1}(y)|dy$$

by changing the variable to $y = g(x)$. Thus

$$\int g(x) f(x)\, dx = \int_{\alpha}^{\beta} y\, h(y)\, dy.$$

The functions $h(x) = x^n$, where n is a positive integer, and $h(x) = |x|^{\alpha}$, where α is a positive real number, are of special importance. If EX^n exists for some positive integer n, we call EX^n the nth *moment* of (the distribution function of) X *about the origin.* If $E|X|^{\alpha} < \infty$ for some positive real number α, we call $E|X|^{\alpha}$ the αth absolute moment of X. We shall use the notation

(4) $$m_n = EX^n \quad \text{and} \quad \beta_{\alpha} = E|X|^{\alpha}$$

whenever the expectations exist.

Example 2. Let X have the *uniform* distribution on the first N natural numbers; that is, let

$$P\{X = k\} = \frac{1}{N}, \qquad k = 1, 2, \ldots, N.$$

Clearly, moments of all order exist:

$$EX = \sum_{k=1}^{N} k \cdot \frac{1}{N} = \frac{N+1}{2},$$

$$EX^2 = \sum_{k=1}^{N} k^2 \cdot \frac{1}{N} = \frac{(N+1)(2N+1)}{6}.$$

Example 3. Let X be an RV with PDF

$$f(x) = \begin{cases} \dfrac{2}{x^3}, & x \ge 1, \\ 0, & x < 1. \end{cases}$$

Then

$$EX = \int_1^{\infty} \frac{2}{x^2}\, dx = 2.$$

But

$$EX^2 = \int_1^{\infty} \frac{2}{x}\, dx$$

does not exist. Indeed, it is easily possible to construct examples of random variables for which all moments of a specified order exist but no higher-order moments do.

Example 4. Two players, A and B, play a coin-tossing game. A gives B one dollar if a head turns up; otherwise, B pays A one dollar. If the probability that the coin shows a head is p, find the expected gain of A.

Let X denote the gain of A. Then

$$P\{X = 1\} = P\{\text{tails}\} = 1 - p, \qquad P\{X = -1\} = p,$$

and

$$EX = 1 - p - p = 1 - 2p \begin{cases} > 0 & \text{if and only if } p < \frac{1}{2}, \\ = 0 & \text{if and only if } p = \frac{1}{2}. \end{cases}$$

Thus $EX = 0$ if and only if the coin is *fair*.

Theorem 2. If the moment of order t exists for an RV X, moments of order $0 < s < t$ exist.

Proof. Let X be of the continuous type with PDF f. We have

$$\begin{aligned} E|X|^s &= \int_{|x|^s \le 1} |x|^s f(x)\,dx + \int_{|x|^s > 1} |x|^s f(x)\,dx \\ &\le P\{|X|^s \le 1\} + E|X|^t < \infty. \end{aligned}$$

A similar proof can be given when X is a discrete RV.

Theorem 3. Let X be an RV on a probability space $(\Omega, \mathcal{S}, P)$. Let $E|X|^k < \infty$ for some $k > 0$. Then

$$n^k P\{|X| > n\} \to 0 \qquad \text{as } n \to \infty.$$

Proof. We provide the proof for the case in which X is of the continuous type with density f. We have

$$\infty > \int |x|^k f(x)\,dx = \lim_{n\to\infty} \int_{|x|\le n} |x|^k f(x)\,dx.$$

It follows that

$$\lim_{n\to\infty} \int_{|x|>n} |x|^k f(x)\,dx \to 0 \qquad \text{as } n \to \infty.$$

But

$$\int_{|x|>n} |x|^k f(x)\,dx \ge n^k P\{|X| > n\},$$

completing the proof.

Remark 8. Probabilities of the type $P\{|X| > n\}$ or either of its components, $P\{X > n\}$ or $P\{X < -n\}$, are called *tail probabilities*. The result of Theorem 3, therefore, gives the rate at which $P\{|X| > n\}$ converges to 0 as $n \to \infty$.

Remark 9. The converse of Theorem 3 does not hold in general; that is,

$$n^k P\{|X| > n\} \to 0 \qquad \text{as } n \to \infty \text{ for some } k$$

does not necessarily imply that $E|X|^k < \infty$, for consider the RV

$$P\{X = n\} = \frac{c}{n^2 \log n}, \qquad n = 2, 3, \ldots,$$

where c is a constant determined from

$$\sum_{n=2}^{\infty} \frac{c}{n^2 \log n} = 1.$$

We have

$$P\{X > n\} \approx c \int_n^{\infty} \frac{1}{x^2 \log x}\, dx \approx cn^{-1}(\log n)^{-1}$$

and $nP\{X > n\} \to 0$ as $n \to \infty$. (Here and subsequently, $\approx$ means that the ratio of two sides $\to 1$ as $n \to \infty$.) But

$$EX = \sum \frac{c}{n \log n} = \infty.$$

In fact, we need

$$n^{k+\delta} P\{|X| > n\} \to 0 \qquad \text{as } n \to 0$$

for some $\delta > 0$ to ensure that $E|X|^k < \infty$. A condition such as this is called a *moment condition.*

For the proof we need the following lemma.

Lemma 1. Let X be a nonnegative RV with distribution function F. Then

$$EX = \int_0^{\infty} [1 - F(x)]\, dx, \tag{5}$$

in the sense that if either side exists, so does the other and the two are equal.

Proof. If X is of the continuous type with density f and $EX < \infty$, then

$$EX = \int_0^{\infty} xf(x)\, dx = \lim_{n\to\infty} \int_0^n xf(x)\, dx.$$

On integration by parts, we obtain

$$\int_0^n xf(x)\, dx = nF(n) - \int_0^n F(x)\, dx = -n[1 - F(n)] + \int_0^n [1 - F(x)]\, dx.$$

But

$$n[1 - F(n)] = n \int_n^{\infty} f(x)\, dx < \int_n^{\infty} xf(x)\, dx,$$

and since $E|X| < \infty$, it follows that

$$n[1 - F(n)] \to 0 \qquad \text{as } n \to \infty.$$

We have

$$EX = \lim_{n\to\infty} \int_0^n xf(x)\,dx = \lim_{n\to\infty} \int_0^n [1 - F(x)]\,dx = \int_0^\infty [1 - F(x)]\,dx.$$

If $\int_0^\infty [1 - F(x)]\,dx < \infty$, then

$$\int_0^n xf(x)\,dx \le \int_0^n [1 - F(x)]\,dx \le \int_0^\infty [1 - F(x)]\,dx,$$

and it follows that $E|X| < \infty$.

We leave the reader to complete the proof in the discrete case.

Corollary 1. For any RV X, $E|X| < \infty$ if and only if the integrals $\int_{-\infty}^0 P\{X \le x\}\,dx$ and $\int_0^\infty P\{X > x\}\,dx$ both converge, and in that case

$$EX = \int_0^\infty P\{X > x\}\,dx - \int_{-\infty}^0 P\{X \le x\}\,dx.$$

Actually, we can get a little more out of Lemma 1 than the corollary above. In fact,

$$E|X|^\alpha = \int_0^\infty P\{|X|^a > x\}\,dx = \alpha \int_0^\infty x^{\alpha-1} P\{|X| > x\}\,dx,$$

and we see that *an RV X possesses* an *absolute moment of order* $\alpha > 0$ *if and only if* $|x|^{\alpha-1} P\{|X| > x\}$ is integrable over $(0, \infty)$.

A simple application of the integral test leads to the following *moments lemma.*

Lemma 2

$$E|X|^\alpha < \infty \Leftrightarrow \sum_{n=1}^{\infty} P\{|X| > n^{1/\alpha}\} < \infty. \tag{6}$$

Note that an immediate consequence of Lemma 2 is Theorem 3. We are now ready to prove the following result.

Theorem 4. Let X be an RV with a distribution satisfying $n^\alpha P\{|X| > n\} \to 0$ as $n \to \infty$ for some $\alpha > 0$. Then $E|X|^\beta < \infty$ for $0 < \beta < \alpha$.

Proof. Given $\varepsilon > 0$, we can choose an $N = N(\varepsilon)$ such that

$$P\{|X| > n\} < \frac{\varepsilon}{n^\alpha} \qquad \text{for all } n \ge N.$$

It follows that for $0 < \beta < \alpha$,

$$\begin{aligned} E|X|^\beta &= \beta \int_0^N x^{\beta-1} P\{|X| > x\}\,dx + \beta \int_N^\infty x^{\beta-1} P\{|X| > x\}\,dx \\ &\le N^\beta + \beta\varepsilon \int_N^\infty x^{\beta-\alpha-1}\,dx \\ &< \infty. \end{aligned}$$

Remark 10. Using Theorems 3 and 4, we demonstrate the existence of random variables for which moments of any order do not exist, that is, for which $E|X|^\alpha = \infty$ for every $\alpha > 0$. For such an RV $n^\alpha P\{|X| > n\} \nrightarrow 0$ as $n \to \infty$ for any $\alpha > 0$. Consider, for example, the RV X with PDF

$$f(x) = \begin{cases} \dfrac{1}{2|x|(\log|x|)^2} & \text{for } |x| > e \\ 0 & \text{otherwise.} \end{cases}$$

The DF of X is given by

$$F(x) = \begin{cases} \dfrac{1}{2\log|x|} & \text{if } x \le -e, \\ \dfrac{1}{2} & \text{if } -e < x < e, \\ 1 - \dfrac{1}{2\log x} & \text{if } x \ge e. \end{cases}$$

Then for $x > e$,

$$\begin{aligned} P\{|X| > x\} &= 1 - F(x) + F(-x) \\ &= \frac{1}{2\log x}, \end{aligned}$$

and $x^\alpha P\{|X| > x\} \to \infty$ as $x \to \infty$ for any $\alpha > 0$. It follows that $E|X|^\alpha = \infty$ for every $\alpha > 0$. In this example we see that $P\{|X| > cx\}/P\{|X| > x\} \to 1$ as $x \to \infty$ for every $c > 0$. A positive function $L(\cdot)$ defined on $(0, \infty)$ is said to be a function of *slow variation* if and only if $L(cx)/L(x) \to 1$ as $x \to \infty$ for every $c > 0$. For such a function $x^\alpha L(x) \to \infty$ for every $\alpha > 0$ (see Feller [23, pp. 275–279]). It follows that if $P\{|X| > x\}$ is slowly varying, $E|X|^\alpha = \infty$ for every $\alpha > 0$. Functions of slow variation play an important role in the theory of probability.

Random variables for which $P\{|X| > x\}$ is slowly varying are clearly excluded from the domain of the following result.

Theorem 5. Let X be an RV satisfying

$$\frac{P\{|X| > cx\}}{P\{|X| > x\}} \to 0 \qquad \text{as } x \to \infty \quad \text{for all } c > 1; \tag{7}$$

then X possesses moments of all orders. [Note that if $c = 1$, the limit in (7) is 1, whereas if $c < 1$, the limit will not go to 0 since $P\{|X| > cx\} \geq P\{|X| > x\}$.]

Proof. Let $\varepsilon > 0$ (we will choose ε later), choose x_0 so large that

$$\frac{P\{|X| > cx\}}{P\{|X| > x\}} < \varepsilon \qquad \text{for all } x \geq x_0, \tag{8}$$

and choose x_1 so large that

$$P\{|X| > x\} < \varepsilon \qquad \text{for all } x \geq x_1. \tag{9}$$

Let $N = \max(x_0, x_1)$. We have for a fixed positive integer r,

$$\frac{P\{|X| > c^r x\}}{P\{|X| > x\}} = \prod_{p=1}^{r} \frac{P\{|X| > c^p x\}}{P\{|X| > c^{p-1} x\}} \leq \varepsilon^r \tag{10}$$

for $x \geq N$. Thus for $x \geq N$ we have, in view of (9),

$$P\{|X| > c^r x\} \leq \varepsilon^{r+1}. \tag{11}$$

Next note that for any fixed positive integer n,

$$\begin{aligned} E|X|^n &= n\int_0^\infty x^{n-1} P\{|X| > x\}\,dx \\ &= n\int_0^N x^{n-1} P\{|X| > x\}\,dx + n\int_N^\infty x^{n-1} P\{|X| > x\}\,dx. \end{aligned} \tag{12}$$

Since the first integral in (12) is finite, we need only show that the second integral is also finite. We have

$$\begin{aligned} \int_N^\infty x^{n-1} P\{|X| > x\}\,dx &= \sum_{r=1}^\infty \int_{c^{r-1}N}^{c^r N} x^{n-1} P\{|X| > x\}\,dx \\ &\leq \sum_{r=1}^\infty (c^r N)^{n-1} \varepsilon^r \cdot 2c^r N \\ &= 2N^n \sum_{r=1}^\infty (\varepsilon c^n)^r \\ &= 2N^n \frac{\varepsilon c^n}{1 - \varepsilon c^n} < \infty, \end{aligned}$$

provided that we choose ε such that $\varepsilon c^n < 1$. It follows that $E|X|^n < \infty$ for $n = 1, 2, \ldots$. Actually, we have shown that (7) implies that $E|X|^\delta < \infty$ for all $\delta > 0$.

Theorem 6. If $h_1, h_2, \ldots, h_n$ are Borel-measurable functions of an RV X and $Eh_i(X)$ exists for $i = 1, 2, \ldots, n$, then $E\left[\sum_{i=1}^n h_i(X)\right]$ exists and equals $\sum_{i=1}^n Eh_i(X)$.

Definition 1. Let k be a positive integer and c be a constant. If $E(X-c)^k$ exists, we call it the *moment of order* k about the point c. If we take $c = EX = \mu$, which exists since $E|X| < \infty$, we call $E(X-\mu)^k$ the *central moment of order* k or the moment of order k about the mean. We shall write

$$\mu_k = E(X-\mu)^k.$$

If we know $m_1, m_2, \ldots, m_k$, we can compute $\mu_1, \mu_2, \ldots, \mu_k$, and conversely. We have

$$\mu_k = E(X-\mu)^k = m_k - \binom{k}{1}\mu m_{k-1} + \binom{k}{2}\mu^2 m_{k-2} - \cdots + (-1)^k \mu^k \tag{13}$$

and

$$m_k = E(X-\mu+\mu)^k = \mu_k + \binom{k}{1}\mu\mu_{k-1} + \binom{k}{2}\mu^2\mu_{k-2} + \cdots + \mu^k. \tag{14}$$

The case $k = 2$ is of special importance.

Definition 2. If EX^2 exists, we call $E(X-\mu)^2$ the *variance* of X, and we write $\sigma^2 = \text{var}(X) = E(X-\mu)^2$. The quantity σ is called the *standard deviation* (SD) of X.

From Theorem 6 we see that

$$\sigma^2 = \mu_2 = EX^2 - (EX)^2. \tag{15}$$

Variance has some important properties.

Theorem 7. $\text{Var}(X) = 0$ if and only if X is degenerate.

Theorem 8. $\text{Var}(X) < E(X-c)^2$ for any $c \neq EX$.

Proof. We have

$$\text{var}(X) = E(X-\mu)^2 = E(X-c)^2 + (c-\mu)^2.$$

Note that

$$\text{var}(aX+b) = a^2\,\text{var}(X).$$

Let $E|X|^2 < \infty$. Then we define

(16) $$Z = \frac{X - EX}{\sqrt{\text{var}(X)}} = \frac{X - \mu}{\sigma}$$

and see that $EZ = 0$ and $\text{var}(Z) = 1$. We call Z a *standardized* RV.

Example 5. Let X be an RV with *binomial* PMF

$$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, 2, \ldots, n; \quad 0 < p < 1.$$

Then

$$\begin{aligned}
EX &= \sum_{k=0}^{\infty} k \binom{n}{k} p^k (1-p)^{n-k} \\
&= np \sum \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\
&= np; \\
EX^2 &= E[X(X-1) + X] \\
&= \sum k(k-1) \binom{n}{k} p^k (1-p)^{n-k} + np \\
&= n(n-1)p^2 + np; \\
\text{var}(X) &= n(n-1)p^2 + np - n^2 p^2 \\
&= np(1-p); \\
EX^3 &= E[X(X-1)(X-2) + 3X(X-1) + X] \\
&= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np;
\end{aligned}$$

and

$$\begin{aligned}
\mu_3 &= m_3 - 3\mu m_2 + 2\mu^3 \\
&= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np - 3np[n(n-1)p^2 + np] + 2n^3 p^3 \\
&= np(1-p)(1-2p).
\end{aligned}$$

In the example above we computed *factorial moments* $EX(X-1)(X-2)\cdots(X-k+1)$ for various values of k. For some discrete integer-valued RVs whose PMF contains factorials or binomial coefficients, it may be more convenient to compute factorial moments.

We have seen that for some distributions, even the mean does not exist. We next consider some parameters, called *order parameters*, which always exist.

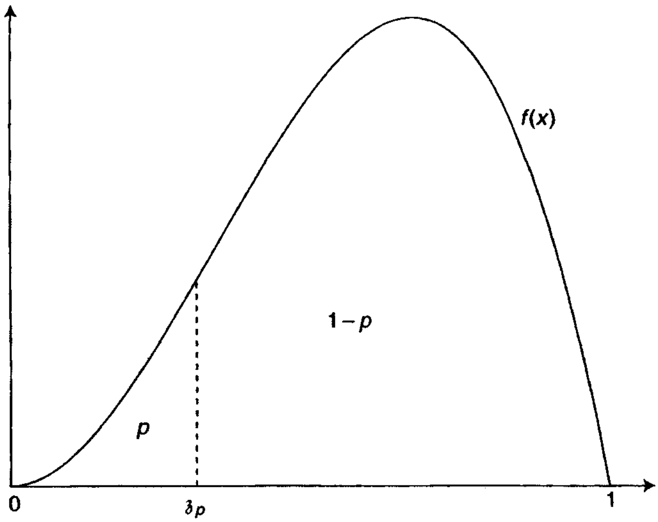

Fig. 1. Quantile of order p.

Definition 3. A number x (Fig. 1) satisfying

$$P\{X \le x\} \ge p, \qquad P\{X \ge x\} \ge 1 - p, \qquad 0 < p < 1, \tag{17}$$

is called a *quantile of order* p [or $(100p)$th *percentile*] for the RV X (or for the DF F of X). We write $\mathfrak{z}_p(X)$ for a quantile of order p for the RV X.

If x is a quantile of order p for an RV X with DF F, then

$$p \le F(x) \le p + P\{X = x\}. \tag{18}$$

If $P\{X = x\} = 0$, as is the case—in particular, if X is of the continuous type—a quantile of order p is a solution of the equation

$$F(x) = p. \tag{19}$$

If F is strictly increasing, (19) has a unique solution. Otherwise (Fig. 2), there may be many (even uncountably many) solutions of (19), each of which is then called a quantile of order p. Quantiles are of great deal of interest in testing hypotheses.

Definition 4. Let X be an RV with DF F. A number x satisfying

$$\tfrac{1}{2} \le F(x) \le \tfrac{1}{2} + P\{X = x\} \tag{20}$$

or, equivalently,

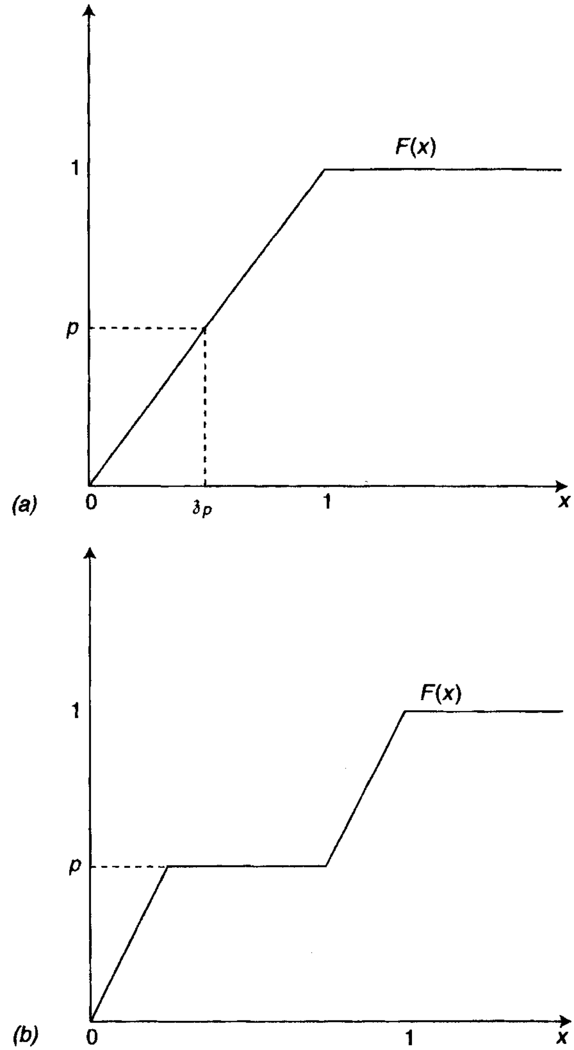

Fig. 2. (*a*) Unique quantile; (*b*) infinitely many solutions of $F(x) = p$.

(21) $$P\{X \le x\} \ge \tfrac{1}{2} \quad \text{and} \quad P\{X \ge x\} \ge \tfrac{1}{2}$$

is called a *median* of X (or F).

Again we note that there may be many values that satisfy (20) or (21). Thus a median is not necessarily unique.

If F is a symmetric DF, the center of symmetry is clearly the median of the DF F. The median is an important centering constant, especially in cases where the mean of the distribution does not exist.

Example 6. Let X be an RV with *Cauchy* PDF

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

Then $E|X|$ is not finite, but $E|X|^\delta < \infty$ for $0 < \delta < 1$. The median of the RV X is clearly $x = 0$.

Example 7. Let X be an RV with PMF

$$P\{X = -2\} = P\{X = 0\} = \tfrac{1}{4}, \quad P\{X = 1\} = \tfrac{1}{3}, \quad P\{X = 2\} = \tfrac{1}{6}.$$

Then

$$P\{X \le 0\} = \tfrac{1}{2} \quad \text{and} \quad P\{X \ge 0\} = \tfrac{3}{4} > \tfrac{1}{2}.$$

In fact, if x is any number such that $0 < x < 1$, then

$$P\{X \le x\} = P\{X = -2\} + P\{X = 0\} = \tfrac{1}{2}$$

and

$$P\{X \ge x\} = P\{X = 1\} + P\{X = 2\} = \tfrac{1}{2},$$

and it follows that every x, $0 \le x < 1$, is a median of the RV X.

If $p = 0.2$, the quantile of order p is $x = -2$, since

$$P\{X \le -2\} = \tfrac{1}{4} > p \quad \text{and} \quad P\{X \ge -2\} = 1 > 1 - p.$$

PROBLEMS 3.2

1. Find the expected number of throws of a fair die until a 6 is obtained.
2. From a box containing N identical tickets numbered 1 through N, n tickets are drawn with replacement. Let X be the largest number drawn. Find EX.
3. Let X be an RV with PDF

$$f(x) = \frac{c}{(1+x^2)^m}, \qquad -\infty < x < \infty, \quad m \ge 1,$$

where $c = \Gamma(m)/[\Gamma(\frac{1}{2})\Gamma(m - \frac{1}{2})]$. Show that EX^{2r} exists if and only if $2r < 2m - 1$. What is EX^{2r} if $2r < 2m - 1$?

4. Let X be an RV with PDF

$$f(x) = \begin{cases} \dfrac{ka^k}{(x+a)^{k+1}} & \text{if } x \geq 0, \\ 0 & \text{otherwise } (a > 0). \end{cases}$$

Show that $E|X|^\alpha < \infty$ for $\alpha < k$. Find the quantile of order p for the RV X.

5. Let X be an RV such that $E|X| < \infty$. Show that $E|X - c|$ is minimized if we choose c equal to the median of the distribution of X.

6. *Pareto's distribution* with parameters α and β (both α and β positive) is defined by the PDF

$$f(x) = \begin{cases} \dfrac{\beta\alpha^\beta}{x^{\beta+1}} & \text{if } x \geq \alpha, \\ 0 & \text{if } x < \alpha. \end{cases}$$

Show that the moment of order n exists if and only if $n < \beta$. Let $\beta > 2$. Find the mean and the variance of the distribution.

7. For an RV X with PDF

$$f(x) = \begin{cases} \frac{1}{2}x & \text{if } 0 \leq x < 1, \\ \frac{1}{2} & \text{if } 1 < x \leq 2, \\ \frac{1}{2}(3-x) & \text{if } 2 < x \leq 3, \end{cases}$$

show that moments of all order exist. Find the mean and the variance of X.

8. For the PMF of Example 5, show that

$$EX^4 = np + 7n(n-1)p^2 + 6n(n-1)(n-2)p^3 + n(n-1)(n-2)(n-3)p^4$$

and

$$\mu_4 = 3(npq)^2 + npq(1 - 6pq),$$

where $0 \leq p \leq 1, q = 1 - p$.

9. For the Poisson RV X with PMF

$$P\{X = x\} = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots,$$

show that $EX = \lambda$, $EX^2 = \lambda + \lambda^2$, $EX^3 = \lambda + 3\lambda^2 + \lambda^3$, $EX^4 = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4$, and $\mu_2 = \mu_3 = \lambda$, $\mu_4 = \lambda + 3\lambda^2$.

10. For any RV X with $E|X|^4 < \infty$, define

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}}, \qquad \alpha_4 = \frac{\mu_4}{\mu_2^2}.$$

Here α_3 is known as the *coefficient of skewness* and is sometimes used as a measure of asymmetry, and α_4 is known as *kurtosis* and is used to measure the peakedness ("flatness of the top") of a distribution. Compute α_3 and α_4 for the PMFs of Problems 8 and 9.

11. For a positive RV X define the negative moment of order n by EX^{-n}, where $n > 0$ is an integer. Find $E[1/(X+1)]$ for the PMFs of Example 5 and Problem 9.

12. Prove Theorem 6.

13. Prove Theorem 7.

14. In each of the following cases, compute EX, $\text{var}(X)$, and EX^n (for $n \geq 0$, an integer) whenever they exist:

(a) $f(x) = 1, -\frac{1}{2} \leq x \leq \frac{1}{2}$, and zero elsewhere.
(b) $f(x) = e^{-x}, x \geq 0$, and zero elsewhere.
(c) $f(x) = (k-1)/x^k, x \geq 1$, and zero elsewhere; $k > 1$ is a constant.
(d) $f(x) = 1/[\pi(1+x^2)], -\infty < x < \infty$.
(e) $f(x) = 6x(1-x), 0 < x < 1$, and zero elsewhere.
(f) $f(x) = xe^{-x}, x \geq 0$, and zero elsewhere.
(g) $P(X = x) = p(1-p)^{x-1}, x = 1, 2, \ldots$, and zero elsewhere: $0 < p < 1$.

15. Find the quantile of order $p(0 < p < 1)$ for the following distributions.

(a) $f(x) = 1/x^2, x \geq 1$, and zero elsewhere.
(b) $f(x) = 2x\exp(-x^2), x \geq 0$, and zero otherwise.
(c) $f(x) = 1/\theta, 0 \leq x \leq \theta$, and zero elsewhere.
(d) $P(X = x) = \theta(1-\theta)^{x-1}, x = 1, 2, \ldots$, and zero otherwise; $0 < \theta < 1$.
(e) $f(x) = (1/\beta^2)x \exp(-x/\beta), x > 0$, and zero otherwise; $\beta > 0$.
(f) $f(x) = (3/b^3)(b-x)^2, 0 < x < b$, and zero elsewhere.

3.3 GENERATING FUNCTIONS

In this section we consider some functions that generate probabilities or moments of an RV. The simplest type of generating function in probability theory is the one associated with integer-valued RVs. Let X be an RV, and let

$$p_k = P\{X = k\}, \qquad k = 0, 1, 2, \ldots$$

with $\sum_{k=0}^{\infty} p_k = 1$.

Definition 1. The function defined by

$$P(s) = \sum_{k=0}^{\infty} p_k s^k, \tag{1}$$

which surely converges for $|s| \leq 1$, is called the *probability generating function* (PGF) of X.

Example 1. Consider the Poisson RV with PMF

$$P\{X = k\} = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots.$$

We have

$$P(s) = \sum_{k=0}^{\infty}(s\lambda)^k \frac{e^{-\lambda}}{k!} = e^{-\lambda}e^{s\lambda} = e^{-\lambda(1-s)} \qquad \text{for all } s.$$

Example 2. Let X be an RV with *geometric* distribution, that is, let

$$P\{X = k\} = pq^k, \qquad k = 0, 1, 2, \ldots; \quad 0 < p < 1, \quad q = 1 - p.$$

Then

$$P(s) = \sum_{k=0}^{\infty} s^k pq^k = p\frac{1}{1 - sq}, \qquad |s| \leq 1.$$

Remark 1. Since $P(1) = 1$, series (1) is uniformly and absolutely convergent in $|s| \leq 1$ and the PGF P is a continuous function of s. It determines the PGF uniquely, since $P(s)$ can be represented in a unique manner as a power series.

Remark 2. Since a power series with radius of convergence r can be differentiated termwise any number of times in $(-r, r)$, it follows that

$$P^{(k)}(s) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)P(X = n)s^{n-k},$$

where $P^{(k)}$ is the kth derivative of P. The series converges at least for $-1 < s < 1$. For $s = 1$ the right side reduces formally to $E[X(X-1)\ldots(X-k+1)]$, which is the kth factorial moment of X whenever it exists. In particular, if $EX < \infty$, then $P'(1) = EX$, and if $EX^2 < \infty$, then $P''(1) = EX(X-1)$ and $\text{var}(X) = EX^2 - (EX)^2 = P''(1) - [P'(1)]^2 + P'(1)$.

Example 3. In Example 1 we found that $P(s) = e^{-\lambda(1-s)}$, $|s| \leq 1$, for a Poisson RV. Thus

$$P'(s) = \lambda e^{-\lambda(1-s)},$$
$$P''(s) = \lambda^2 e^{-\lambda(1-s)}.$$

Also, $EX = \lambda$, $E(X^2 - X) = \lambda^2$, so that $\text{var}(X) = EX^2 - (EX)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

In Example 2 we computed $P(s) = p/(1 - sq)$, so that

$$P'(s) = \frac{pq}{(1 - sq)^2} \quad \text{and} \quad P''(s) = \frac{2pq^2}{(1 - sq)^3}.$$

Thus

$$EX = \frac{q}{p}, \qquad EX^2 = \frac{q}{p} + \frac{2pq^2}{p^3}, \qquad \text{and} \qquad \text{var}(X) = \frac{q^2}{p^2} + \frac{q}{p} = \frac{q}{p^2}.$$

Example 4. Consider the PGF

$$P(s) = \left(\frac{1 + s}{2}\right)^n, \qquad -\infty < s < \infty.$$

Expanding the right side into a power series, we get

$$P(s) = \sum_{k=0}^{n} \frac{1}{2^n} \binom{n}{k} s^{n-k} = \sum_{k=0}^{n} p_k s^k,$$

and it follows that

$$P(X = k) = p_k = \binom{n}{k} \bigg/ 2^n, \qquad k = 0, 1, \dots, n.$$

We note that the PGF, being defined only for discrete integer-valued RVs, has limited utility. We next consider a generating function that is quite useful in probability and statistics.

Definition 2. Let X be an RV defined on $(\Omega, \mathcal{S}, P)$. The function

$$M(s) = Ee^{sX} \tag{2}$$

is known as the *moment generating function* (MGF) of the RV X if the expectation on the right side of (2) exists in some neighborhood of the origin.

Example 5. Let X have the PMF

$$f(x) = \begin{cases} \dfrac{6}{\pi^2} \cdot \dfrac{1}{k^2}, & k = 1, 2, \dots, \\ 0, & \text{otherwise.} \end{cases}$$

Then $(1/\pi^2) \sum_{k=1}^{\infty} e^{sk}/k^2$, is infinite for every $s > 0$. We see that the MGF of X does not exist. In fact, $EX = \infty$.

Example 6. Let X have the PDF

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$
\begin{aligned}
M(s) &= \frac{1}{2}\int_0^\infty e^{(s-1/2)x}\,dx \\
&= \frac{1}{1-2s}, \qquad s < \frac{1}{2}.
\end{aligned}
$$

Example 7. Let X have the PMF

$$
P\{X=k\} = \begin{cases} e^{-\lambda}\dfrac{\lambda^k}{k!}, & k = 0, 1, 2, \dots, \\ 0, & \text{otherwise.} \end{cases}
$$

Then

$$
\begin{aligned}
M(s) = Ee^{sX} &= e^{-\lambda}\sum_{k=0}^{\infty} e^{sk}\frac{\lambda^k}{k!} \\
&= e^{-\lambda(1-e^s)} \qquad \text{for all } s.
\end{aligned}
$$

The following result will be quite useful subsequently.

Theorem 1. The MGF uniquely determines a DF and, conversely, if the MGF exists, it is unique.

For the proof we refer the reader to Widder [116, p. 460], or Curtiss [18]. Theorem 2 explains why we call $M(s)$ an MGF.

Theorem 2. If the MGF $M(s)$ of an RV X exists for s in $(-s_0, s_0)$, say, $s_0 > 0$, the derivatives of all order exist at $s = 0$ and can be evaluated under the integral sign, that is,

(3) $$M^{(k)}(s)\big|_{s=0} = EX^k \qquad \text{for positive integral } k.$$

For the proof of Theorem 2, we refer to Widder [116, pp. 446–447]. See also Problem 9.

Remark 3. Alternatively, if the MGF $M(s)$ exists for s in $(-s_0, s_0)$, say, $s_0 > 0$, one can express $M(s)$ (uniquely) in a Maclaurin series expansion:

(4) $$M(s) = M(0) + \frac{M'(0)}{1!}s + \frac{M''(0)}{2!}s^2 + \cdots,$$

so that EX^k is the coefficient of $s^k/k!$ in expansion (4).

Example 8. Let X be an RV with PDF $f(x) = \frac{1}{2}e^{-x/2}$, $x > 0$. From Example 6, $M(s) = 1/(1-2s)$ for $s < \frac{1}{2}$. Thus

$$M'(s) = \frac{2}{(1-2s)^2} \quad \text{and} \quad M''(s) = \frac{4 \cdot 2}{(1-2s)^3}, \qquad s < \frac{1}{2}.$$

It follows that

$$EX = 2, \quad EX^2 = 8, \quad \text{and} \quad \text{var}(X) = 4.$$

Example 9. Let X be an RV with PDF $f(x) = 1, 0 \le x \le 1$, and $= 0$ otherwise. Then

$$M(s) = \int_0^1 e^{sx}\,dx = \frac{e^s - 1}{s}, \qquad \text{all } s,$$

$$M'(s) = \frac{e^s \cdot s - (e^s - 1) \cdot 1}{s^2},$$

and

$$EX = M'(0) = \lim_{s \to 0} \frac{se^s - e^s + 1}{s^2} = \frac{1}{2}.$$

We emphasize that the expectation Ee^{sX} does not exist unless s is carefully restricted. In fact, the requirement that $M(s)$ exists in a neighborhood of zero is a very strong requirement that is not satisfied by some common distributions. We next consider a generating function that exists for all distributions.

Definition 3. Let X be an RV. The complex-valued function ϕ defined on $\mathcal{R}$ by

$$\phi(t) = E(e^{itX}) = E(\cos tX) + iE(\sin tX), \qquad t \in \mathcal{R}$$

where $i = \sqrt{-1}$ is the imaginary unit, is called the *characteristic function* (CF) of RV X.

Clearly,

$$\phi(t) = \sum_k (\cos tx_k + i \sin tx_k) P(X = x_k)$$

in the discrete case, and

$$\phi(t) = \int_{-\infty}^{\infty} \cos tx f(x)\,dx + i \int_{-\infty}^{\infty} \sin tx\ f(x)\,dx$$

in the continuous case.

Example 10. Let X be a normal RV with PDF

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \qquad x \in \mathcal{R}.$$

Then

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos tx\, e^{-x^2/2}\, dx + \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin tx\, e^{-x^2/2}\, dx.$$

Note that $\sin tx$ is an odd function and so also is $\sin tx\, e^{-x^2/2}$. Thus the second integral on the right side vanishes and we have

$$\begin{aligned}\phi(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos tx\, e^{-x^2/2}\, dx \\ &= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos tx\, e^{-x^2/2}\, dx = e^{-t^2/2}, \qquad t \in \mathcal{R}.\end{aligned}$$

Remark 4. Unlike an MGF that may not exist for some distributions, a CF always exists, which makes it a much more convenient tool. In fact, it is easy to see that ϕ is continuous on $\mathcal{R}$, $|\phi(t)| \leq 1$ for all t, and $\phi(-t) = \overline{\phi}(t)$ where $\overline{\phi}$ is the complex conjugate of ϕ. Thus $\overline{\phi}$ is the CF of $-X$. Moreover, ϕ uniquely determines the DF of RV X. For these and many other properties of characteristic functions, we need a comprehensive knowledge of complex variable theory, well beyond the scope of this book. We refer the reader to Lukacs [68].

Finally, we consider the problem of characterizing a distribution from its moments. Given a set of constants $\{\mu_0 = 1, \mu_1, \mu_2, \dots\}$, the problem of moments asks if they can be moments of a distribution function F. At this point it will be worthwhile to take note of some facts.

First, we have seen that if the $M(s) = Ee^{sX}$ exists for some X for s in some neighborhood of zero, then $E|X|^n < \infty$ for all $n \geq 1$. Suppose, however, that $E|X|^n < \infty$ for all $n \geq 1$. It does not follow that the MGF of X exists.

Example 11. Let X be an RV with PDF

$$f(x) = ce^{-|x|^\alpha}, \qquad 0 < \alpha < 1, \quad -\infty < x < \infty,$$

where c is a constant determined from

$$c \int_{-\infty}^{\infty} e^{-|x|^\alpha}\, dx = 1.$$

Let $s > 0$. Then

$$\int_0^\infty e^{sx}e^{-x^\alpha}\,dx = \int_0^\infty e^{x(s-x^{\alpha-1})}\,dx$$

and since $\alpha - 1 < 0$, $\int_0^\infty s^{sx}e^{-x^\alpha}\,dx$ is not finite for any $s > 0$. Hence the MGF does not exist. But

$$E|X|^n = c\int_{-\infty}^\infty |x|^n e^{-|x|^\alpha}\,dx = 2c\int_0^\infty x^n e^{-x^\alpha}\,dx < \infty \qquad \text{for each } n,$$

as is easily checked by substituting $y = x^\alpha$.

Second, two (or more) RVs may have the same set of moments.

Example 12. Let X have *lognormal* PDF

$$f(x) = (x\sqrt{2\pi})^{-1}e^{-(\log x)^2/2}, \qquad x > 0,$$

and $f(x) = 0$ for $x \le 0$. Let X_ε, $|\varepsilon| \le 1$, have PDF

$$f_\varepsilon(x) = f(x)[1 + \varepsilon\sin(2\pi\log x)], \qquad x \in \mathcal{R}.$$

[Note that $f_\varepsilon \ge 0$ for all ε, $|\varepsilon| \le 1$, and $\int_{-\infty}^\infty f_\varepsilon(x)\,dx = 1$, so f_ε is a PDF.] Since, however,

$$\begin{aligned}\int_0^\infty x^k f(x)\sin(2\pi\log x)\,dx &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-(t^2/2)+kt}\sin(2\pi t)\,dt\\ &= \frac{1}{\sqrt{2\pi}}e^{k^2/2}\int_{-\infty}^\infty e^{-y^2/2}\sin(2\pi y)\,dy\\ &= 0,\end{aligned}$$

we see that

$$\int_0^\infty x^k f(x)\,dx = \int_0^\infty x^k f_\varepsilon(x)\,dx$$

for all ε, $|\varepsilon| \le 1$, and $k = 0, 1, 2, \ldots$. But $f(x) \ne f_\varepsilon(x)$.

Third, moments of any RV X necessarily satisfy certain conditions. For example, if $\beta_\nu = E|X|^\nu$, we will see (Theorem 3.4.3) that $(\beta_\nu)^{1/\nu}$ is an increasing function of ν. Similarly, the quadratic form

$$E\left(\sum_{i=1}^n X^{\alpha_i}t_i\right)^2 \ge 0$$

yields a relation between moments of various orders of X.

The following result, which we do not prove here, gives a sufficient condition for unique determination of F from its moments.

Theorem 3. Let $\{m_k\}$ be the moment sequence of an RV X. If the series

$$\sum_{k=1}^{\infty} \frac{m_k}{k!} s^k \tag{5}$$

converges absolutely for some $s > 0$, then $\{m_k\}$ uniquely determines the DF F of X.

Example 13. Suppose that X has PDF

$$f(x) = e^{-x} \qquad \text{for } x \geq 0 \quad \text{and} \qquad = 0 \text{ for } x < 0.$$

Then $EX^k = \int_0^\infty x^k e^{-x}\, dx = k!$, and from Theorem 3,

$$\sum_{k=1}^{\infty} \frac{m_k}{k!} s^k = \sum_{k=1}^{\infty} s^k = \frac{s}{1-s}$$

for $0 < s < 1$, so that $\{m_k\}$ determine F uniquely. In fact, from Remark 3,

$$M(s) = \sum_{k=0}^{\infty} m_k \frac{s^k}{k!} = \sum_{k=0}^{\infty} s^k = \frac{1}{1-s},$$

$0 < s < 1$, which is the MGF of X.

In particular, if for some constant c,

$$|m_k| \leq c^k, \qquad k = 1, 2, \ldots,$$

then

$$\sum_{k=1}^{\infty} \frac{|m_k|}{k!} s^k \leq \sum_{1}^{\infty} \frac{(cs)^k}{k!} < e^{cs} \qquad \text{for } s > 0,$$

and the DF of X is determined uniquely. Thus if $P\{|X| \leq c\} = 1$ for some $c > 0$, then all moments of X exist, satisfying $|m_k| \leq c^k$, $k \geq 1$, and the DF of X is determined uniquely from its moments.

Finally, we mention some sufficient conditions for a moment sequence to determine a unique DF.

(i) The range of the RV is finite.
(ii) (Carleman) $\sum_{k=1}^{\infty} (m_{2k})^{-1/2k} = \infty$ when the range of the RV is $(-\infty, \infty)$. If the range is $(0, \infty)$, a sufficient condition is $\sum_{k=1}^{\infty} (m_k)^{-1/2k} = \infty$,
(iii) $\lim_{n\to\infty}[(m_{2n})^{1/2n}/2n]$ is finite.

PROBLEMS 3.3

1. Find the PGF of the RVs with the following PMFs:

(a) $P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, 2, \ldots, 0 \le p \le 1.$

(b) $P\{X = k\} = [e^{-\lambda}/(1-e^{-\lambda})](\lambda^k/k!), k = 1, 2, \ldots; \lambda > 0.$

(c) $P\{X = k\} = pq^k(1-q^{N+1})^{-1}, k = 0, 1, 2, \ldots, N; 0 < p < 1, q = 1-p.$

2. Let X be an integer-valued RV with PGF $P(s)$. Let α and β be nonnegative integers, and write $Y = \alpha X + b$. Find the PGF of Y.

3. Let X be an integer-valued RV with PGF $P(s)$, and suppose that the MGF $M(s)$ exists for $s \in (-s_0, s_0)$, $s_0 > 0$. How are $M(s)$ and $P(s)$ related? Using $M^{(k)}(s)|_{s=0} = EX^k$ for positive integral k, find EX^k in terms of the derivatives of $P(s)$ for values of $k = 1, 2, 3, 4$.

4. For the Cauchy PDF

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

does the MGF exist?

5. Let X be an RV with PMF

$$P\{X = j\} = p_j, \qquad j = 0, 1, 2, \ldots.$$

Set $P\{X > j\} = q_j$, $j = 0, 1, 2, \ldots$. Clearly, $q_j = p_{j+1} + p_{j+2} + \cdots$, $j \ge 0$. Write $Q(s) = \sum_{j=0}^{\infty} q_j s^j$. Then the series for $Q(s)$ converges in $|s| < 1$. Show that

$$Q(s) = \frac{1 - P(s)}{1 - s} \qquad \text{for } |s| < 1,$$

where $P(s)$ is the PGF of X. Find the mean and the variance of X (when they exist) in terms of Q and its derivatives.

6. For the PMF

$$P\{X = j\} = \frac{a_j \theta^j}{f(\theta)}, \qquad j = 0, 1, 2, \ldots, \quad \theta > 0,$$

where $a_j \ge 0$ and $f(\theta) = \sum_{j=0}^{\infty} a_j \theta^j$, find the PGF and the MGF in terms of f.

7. For the *Laplace* PDF

$$f(x) = \frac{1}{2\lambda} e^{-|x-\mu|/\lambda}, \qquad -\infty < x < \infty; \quad \lambda > 0, \quad -\infty < \mu < \infty,$$

show that the MGF exists and equals

$$M(t) = (1 - \lambda^2 t^2)^{-1} e^{\mu t}, \qquad |t| < \frac{1}{\lambda}.$$

8. For any integer-valued RV X, show that

$$\sum_{n=0}^{\infty} s^n P\{X \leq n\} = (1-s)^{-1} P(s),$$

where P is the PGF of X.

9. Let X be an RV with MGF $M(t)$, which exists for $t \in (-t_0, t_0)$, $t_0 > 0$. Show that

$$E|X|^n < n!\, s^{-n}[M(s) + M(-s)]$$

for any fixed s, $0 < s < t_0$, and for each integer $n \geq 1$. Expanding e^{tx} in a power series, show that for $t \in (-s, s)$, $0 < s < t_0$,

$$M(t) = \sum_{n=0}^{\infty} t^n \frac{EX^n}{n!}.$$

[Since a power series can be differentiated term by term within the interval of convergence, it follows that for $|t| < s$,

$$M^{(k)}(t)|_{t=0} = EX^k$$

for each integer $k \geq 1$.] (Roy, LePage, and Moore [93]]

10. Let X be an integer-valued random variable with

$$E[X(X-1)\cdots(X-k+1)] = \begin{cases} k!\binom{n}{k} & \text{if } k = 0, 1, 2, \ldots, n \\ 0 & \text{if } k > n. \end{cases}$$

Show that X must be degenerate at n. [*Hint:* Prove and use the fact that if $EX^k < \infty$ for all k, then

$$P(s) = \sum_{k=0}^{\infty} \frac{(s-1)^k}{k!} E[X(X-1)\cdots(X-k+1)].$$

Write $P(s)$ as

$$P(s) = \sum_{k=0}^{\infty} P(X=k)s^k = \sum_{k=0}^{\infty} P(X=k) \sum_{i=0}^{k} (s-1)^i$$

$$= \sum_{i=0}^{\infty}(s-1)^i \sum_{k=i}^{\infty}\binom{k}{i}P(X=k).]$$

11. Let $p(n,k) = f(n,k)/n!$ where $f(n,k)$ is given by

$$f(n+1,k) = f(n,k) + f(n,k-1) + \cdots + f(n,k-n)$$

for $k = 0, 1, \ldots, \binom{n}{2}$ and

$$f(n,k) = 0 \qquad \text{for } k < 0, \qquad f(1,0) = 1, \qquad f(1,k) = 0 \text{ otherwise.}$$

Let

$$P_n(s) = \frac{1}{n!}\sum_{k=0}^{\infty} s^k f(n,k)$$

be the probability generating function of $p(n,k)$. Show that

$$P_n(s) = (n!)^{-1}\prod_{k=2}^{n}\frac{1-s^k}{1-s}, \qquad |s| < 1.$$

(P_n is the generating function of Kendall's τ-statistic.)

12. For $k = 0, 1, \ldots, \binom{n}{2}$, let $u_n(k)$ be defined recursively by

$$u_n(k) = u_{n-1}(k-n) + u_{n-1}(k)$$

with $u_0(0) = 1$, $u_0(k) = 0$ otherwise and $u_n(k) = 0$ for $k < 0$. Let $P_n(s) = \sum_{k=0}^{\infty} s^k u_n(k)$ be the generating function of $\{u_n\}$. Show that

$$P_n(s) = \prod_{j=1}^{n}(1+s^j) \qquad \text{for } |s| < 1.$$

If $p_n(k) = u_n(k)/2^n$, find $\{p_n(k)\}$ for $n = 2, 3, 4$. (P_n is the generating function of the one-sample Wilcoxon test statistic.)

3.4 SOME MOMENT INEQUALITIES

In this section we derive some inequalities for moments of an RV. The main result of this section is Theorem 1 (and its corollary), which gives a bound for tail probability in terms of some moment of the random variable.

Theorem 1. Let $h(X)$ be a nonnegative Borel-measurable function of an RV X. If $Eh(X)$ exists, then for every $\varepsilon > 0$,

(1) $$P\{h(X) \geq \varepsilon\} \leq \frac{Eh(X)}{\varepsilon}.$$

Proof. We prove the result when X is discrete. Let $P\{X = x_k\} = p_k$, $k = 1, 2, \ldots$. Then

$$Eh(X) = \sum_k h(x_k)p_k$$
$$= \left(\sum_A + \sum_{A^c}\right) h(x_k)p_k,$$

where

$$A = \{k \colon h(x_k) \geq \varepsilon\}.$$

Then

$$Eh(X) \geq \sum_A h(x_k)p_k \geq \varepsilon \sum_A p_k$$
$$= \varepsilon P\{h(X) \geq \varepsilon\}.$$

Corollary. Let $h(X) = |X|^r$ and $\varepsilon = K^r$, where $r > 0$ and $K > 0$. Then

(2) $$P\{|X| \geq K\} \leq \frac{E|X|^r}{K^r},$$

which is *Markov's inequality*. In particular, if we take $h(X) = (X - \mu)^2$, $\varepsilon = K^2\sigma^2$, we get *Chebychev–Bienayme inequality*:

(3) $$P\{|X - \mu| \geq K\sigma\} \leq \frac{1}{K^2},$$

where $EX = \mu$, $\text{var}(X) = \sigma^2$.

Remark 1. The inequality (3) is generally attributed to Chebychev, although recent research has shown that credit should also go to I. J. Bienayme.

Remark 2. If we wish to be consistent with our definition of a DF as $F_X(x) = P(X \leq x)$, then we may want to reformulate (1) in the following form:

$$P\{h(X) > \varepsilon\} < \frac{Eh(X)}{\varepsilon}.$$

For RVs with finite second-order moments, one cannot do better than the inequality in (3).

Example 1

$$P\{X=0\}=1-\frac{1}{K^2}$$
$$K>1, \text{ constant,}$$
$$P\{X=\mp 1\}=\frac{1}{2K^2}$$

$$EX=0, \qquad EX^2=\frac{1}{K^2}, \qquad \sigma=\frac{1}{K},$$

and

$$P\{|X|\geq K\sigma\}=P\{|X|\geq 1\}=\frac{1}{K^2},$$

so that equality is achieved.

Example 2. Let X be distributed with PDF $f(x)=1$ if $0<x<1$, and $=0$ otherwise. Then

$$EX=\frac{1}{2}, \qquad EX^2=\frac{1}{3}, \qquad \text{var}(X)=\frac{1}{3}-\frac{1}{4}=\frac{1}{12},$$

and

$$P\left\{\left|X-\frac{1}{2}\right|<2\sqrt{\frac{1}{12}}\right\}=P\left\{\frac{1}{2}-\frac{1}{\sqrt{3}}<X<\frac{1}{2}+\frac{1}{\sqrt{3}}\right\}=1.$$

From Chebychev's inequality

$$P\left\{\left|X-\frac{1}{2}\right|<2\sqrt{\frac{1}{12}}\right\}\geq 1-\frac{1}{4}=0.75.$$

In Fig. 1 we compare the upper bound for $P\{|X-\frac{1}{2}|\geq k/\sqrt{12}\}$ with the exact probability.

It is possible to improve upon Chebychev's inequality, at least in some cases, if we assume the existence of higher-order moments. We need the following lemma.

Lemma 1. Let X be an RV with $EX=0$ and $\text{var}(X)=\sigma^2$. Then

$$P\{X\geq x\}\leq\frac{\sigma^2}{\sigma^2+x^2} \qquad \text{if } x>0, \tag{4}$$

and

$$P\{X\geq x\}\geq\frac{x^2}{\sigma^2+x^2} \qquad \text{if } x<0. \tag{5}$$

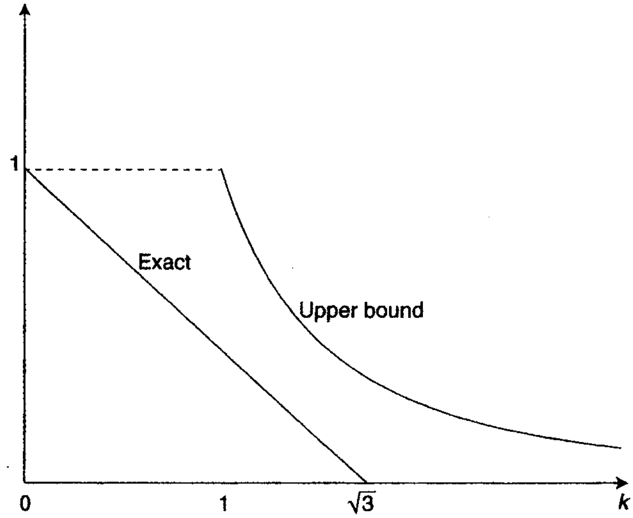

Fig. 1. Chebychev upper bound versus exact probability.

Proof. Let $h(t) = (t + c)^2$, $c > 0$. Then $h(t) \geq 0$ for all t and

$$h(t) \geq (x + c)^2 \qquad \text{for } t \geq x > 0.$$

It follows that

$$P\{X \geq x\} \leq P\{h(X) \geq (x + c)^2\} \tag{6}$$
$$\leq \frac{E(X + c)^2}{(x + c)^2} \qquad \text{for all } c > 0, \quad x > 0.$$

Since $EX = 0$, $EX^2 = \sigma^2$, and the right side of (6) is minimum when $c = \sigma^2/x$. We have

$$P\{X \geq x\} \leq \frac{\sigma^2}{\sigma^2 + x^2}, \qquad x > 0.$$

A similar proof holds for (5).

Remark 3. Inequalities (4) and (5) cannot be improved (Problem 3).

Theorem 2. Let $E|X|^4 < \infty$, and let $EX = 0$, $EX^2 = \sigma^2$. Then

$$P\{|X| \geq K\sigma\} \leq \frac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 K^4 - 2K^2\sigma^4} \qquad \text{for } K > 1, \tag{7}$$

where $\mu_4 = EX^4$.

Proof. For the proof, let us substitute $(X^2 - \sigma^2)/(K^2\sigma^2 - \sigma^2)$ for X and take $x = 1$ in (4). Then

$$
\begin{aligned}
P\{X^2 - \sigma^2 \geq K^2\sigma^2 - \sigma^2\} &\leq \frac{\operatorname{var}[(X^2 - \sigma^2)/(K^2\sigma^2 - \sigma^2)]}{1 + \operatorname{var}[(X^2 - \sigma^2)/(K^2\sigma^2 - \sigma^2)]} \\
&= \frac{\mu_4 - \sigma^4}{\sigma^4(K^2 - 1)^2 + \mu_4 - \sigma^4} \\
&= \frac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 K^4 - 2K^2\sigma^4}, \qquad K > 1,
\end{aligned}
$$

as asserted.

Remark 4. Bound (7) is better than bound (3) if $K^2 \geq \mu_4/\sigma^4$ and worse if $1 \leq K^2 < \mu_4/\sigma^4$ (Problem 5).

Example 3. Let X have the uniform density

$$
f(x) = \begin{cases} 1 & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}
$$

Then

$$
EX = \frac{1}{2}, \qquad \operatorname{var}(X) = \frac{1}{12}, \qquad \mu_4 = E\left(X - \frac{1}{2}\right)^4 = \frac{1}{80},
$$

and

$$
P\left\{\left|X - \frac{1}{2}\right| \geq 2\sqrt{\frac{1}{12}}\right\} \leq \frac{\frac{1}{80} - \frac{1}{144}}{\frac{1}{80} + \frac{1}{144} \cdot 16 - 8\frac{1}{144}} = \frac{4}{49},
$$

that is,

$$
P\left\{\left|X - \frac{1}{2}\right| < 2\sqrt{\frac{1}{12}}\right\} \geq \frac{45}{49} \approx 0.92,
$$

which is much better than the bound given by Chebychev's inequality (Example 2).

Theorem 3 (Lyapunov Inequality). Let $\beta_n = E|X|^n < \infty$. Then for arbitrary k, $2 \leq k \leq n$, we have

$$
\beta_{k-1}^{1/(k-1)} \leq \beta_k^{1/k}. \tag{8}
$$

Proof. Consider the quadratic form:

$$Q(u, v) = \int_{-\infty}^{\infty} (u|x|^{(k-1)/2} + v|x|^{(k+1)/2})^2 f(x)\,dx,$$

where we have assumed that X is continuous with PDF f. We have

$$Q(u, v) = u^2\beta_{k-1} + 2uv\beta_k + \beta_{k+1}v^2.$$

Clearly, $Q \geq 0$ for all u, v real. It follows that

$$\begin{vmatrix} \beta_{k-1} & \beta_k \\ \beta_k & \beta_{k+1} \end{vmatrix} \geq 0,$$

implying that

$$\beta_k^{2k} \leq \beta_{k-1}^k \beta_{k+1}^k.$$

Thus

$$\beta_1^2 \leq \beta_0^1\beta_2^1, \qquad \beta_2^4 \leq \beta_1^2\beta_3^2, \quad \ldots, \quad \beta_{n-1}^{2(n-1)} \leq \beta_{n-2}^{n-1}\beta_n^{n-1},$$

where $\beta_0 = 1$. Multiplying successive $k-1$ of these, we have

$$\beta_{k-1}^k \leq \beta_k^{k-1} \quad \text{or} \quad \beta_{k-1}^{1/(k-1)} \leq \beta_k^{1/k}.$$

It follows that

$$\beta_1 \leq \beta_2^{1/2} \leq \beta_3^{1/3} \leq \cdots \leq \beta_n^{1/n}.$$

The equality holds if and only if

$$\beta_k^{1/k} = \beta_{k+1}^{1/(k+1)} \qquad \text{for } k = 1, 2, \ldots;$$

that is, $\{\beta_k^{1/k}\}$ is a constant sequence of numbers, which happens if and only if $|X|$ is degenerate; that is, for some c, $P\{|X| = c\} = 1$.

PROBLEMS 3.4

1. For the RV with PDF

$$f(x; \lambda) = \frac{e^{-x}x^\lambda}{\lambda!}, \qquad x > 0,$$

where $\lambda \geq 0$ is an integer, show that

$$P\{0 < X < 2(\lambda + 1)\} > \frac{\lambda}{\lambda + 1}.$$

2. Let X be any RV, and suppose that the MGF of X, $M(t) = Ee^{tx}$, exists for every $t > 0$. Then for any $t > 0$,

$$P\{tX > s^2 + \log M(t)\} < e^{-s^2}.$$

3. Construct an example to show that inequalities (4) and (5) cannot be improved.

4. Let $g(\cdot)$ be a function satisfying $g(x) > 0$ for $x > 0$, $g(x)$ increasing for $x > 0$, and $E|g(X)| < \infty$. Show that

$$P\{|X| > \varepsilon\} < \frac{Eg(|X|)}{g(\varepsilon)} \qquad \text{for every } \varepsilon > 0.$$

5. Let X be an RV with $EX = 0$, $\text{var}(X) = \sigma^2$, and $EX^4 = \mu_4$. Let K be any positive real number. Show that

$$P\{|X| \geq K\sigma\} \leq \begin{cases} 1 & \text{if } K^2 < 1, \\ \dfrac{1}{K^2} & \text{if } 1 \leq K^2 < \dfrac{\mu_4}{\sigma^4}, \\ \dfrac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 K^4 - 2K^2\sigma^4} & \text{if } K^2 \geq \dfrac{\mu_4}{\sigma^4}. \end{cases}$$

In other words, show that bound (7) is better than bound (3) if $K^2 \geq \mu_4/\sigma^4$ and worse if $1 \leq K^2 < \mu_4/\sigma^4$. Construct an example to show that the last inequalities cannot be improved.

6. Use Chebychev's inequality to show that for any $k > 1$, $e^{k+1} \geq k^2$.

7. For any RV X, show that

$$P\{X \geq 0\} \leq \inf[\varphi(t) : t \geq 0] \leq 1,$$

where $\varphi(t) = Ee^{tX}$, $0 < \varphi(t) \leq \infty$.

8. Let X be an RV such that $P(a \leq X \leq b) = 1$ where $-\infty < a < b < \infty$. Show that $\text{var}(X) \leq (b - a)^2/4$.