

Moments and Generating Functions

3.1 INTRODUCTION

The study of the probability distributions of a random variable is essentially the study of some numerical characteristics associated with them. These so-called *parameters of the distribution* play a key role in mathematical statistics. In Section 2 we introduce some of these parameters, namely, moments and order parameters, and investigate their properties. In Section 3 the idea of generating functions is introduced. In particular, we study probability generating functions and moment generating functions. Section 4 deals 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 (a) *moments* and their functions and (b) *order parameters*. We will concentrate mainly on moments and their properties.

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

$$(1) \quad \sum_{k=1}^{\infty} |x_k| p_k < \infty,$$

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

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

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

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

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

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

Similar definition is given for the mean of any Borel-measurable function $h(X)$ of X .

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 $\lim_{b \rightarrow \infty} \int_{-b}^b \varphi(x) dx$ exists. It is quite possible for the limit $\lim_{a \rightarrow \infty} \int_{-a}^a \varphi(x) dx$ to exist without the existence of $\int_{-\infty}^{\infty} \varphi(x) dx$. As an example consider the Cauchy pdf:

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

Clearly

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

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

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

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

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

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 $EX = \alpha$. Examples of symmetric df's are easy to construct, and we will encounter many such distributions in this book.

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, $P\{|X| < M\} = 1$, $0 < M < \infty$, then EX exists.

Remark 6. If $P\{X \geq 0\} = 1$, 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 $Y = g(X)$. Then

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

in the sense that, if either side of (3) exists, so does the other, and then 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. Let X be discrete, and suppose that $P\{X \in A\} = 1$. If $y = g(x)$ is a one-to-one mapping of A onto some set B , then

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

We have

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

If X is of the continuous type with pdf f , and g satisfies the conditions of Theorem 2.5.3, then

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

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

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

In the general case, including the case where g may not be one-to-one, we refer the reader to Loève [71], page 166.

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 *n th 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 *α th absolute moment* of X . We shall use the following notation:

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

whenever the expectations exist.

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

$$P\{X = k\} = \frac{1}{N}, \quad k = 1, 2, \dots, N.$$

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

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

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 gain of A .

Let X denote the expected gain of A . Then

$$P\{X = 1\} = P\{\text{Tails}\} = 1 - p, \quad P\{X = -1\} = p$$

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

$$E|X|^s = \int_{|x|^s \leq 1} |x|^s f(x) dx + \int_{|x|^s > 1} |x|^s f(x) dx$$

$$\leq P\{|X|^t \leq 1\} + E|X|^t < \infty.$$

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

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

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

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

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

It follows that

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

But

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

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 $n \rightarrow \infty$.

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

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

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

$$P\{X = n\} = \frac{c}{n^2 \log n}, \quad n = 2, 3, \dots,$$

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

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

In fact, we need

$$n^{k+\delta} P\{|X| > n\} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

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

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

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

$$EX = \int_0^\infty x f(x) dx = \lim_{n \rightarrow \infty} \int_0^n x f(x) dx.$$

On integration by parts we obtain

$$\begin{aligned} \int_0^n x f(x) dx &= n F(n) - \int_0^n F(x) dx \\ &= -n[1 - F(n)] + \int_0^n [1 - F(x)] dx. \end{aligned}$$

But

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

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

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

We have

$$\begin{aligned} EX &= \lim_{n \rightarrow \infty} \int_0^n x f(x) dx = \lim_{n \rightarrow \infty} \int_0^n [1 - F(x)] dx \\ &= \int_0^\infty [1 - F(x)] dx. \end{aligned}$$

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

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

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

If, on the other hand, X is a discrete rv, let us write $P\{X = x_j\} = p_j$. Then

$$EX = \sum_{j=1}^\infty x_j p_j.$$

Let $I = \int_0^\infty [1 - F(x)] dx$. Then

$$I = \sum_{k=1}^\infty \int_{(k-1)/n}^{k/n} P\{X > x\} dx,$$

and since $P\{X > x\}$ is nonincreasing, we have

$$\frac{1}{n} \sum_{k=1}^\infty P\left\{X > \frac{k}{n}\right\} \leq I \leq \frac{1}{n} \sum_{k=1}^\infty P\left\{X > \frac{k-1}{n}\right\}$$

for every positive integer n . Let

$$U_n = \frac{1}{n} \sum_{k=1}^\infty P\left\{X > \frac{k-1}{n}\right\}, \quad \text{and} \quad L_n = \frac{1}{n} \sum_{k=1}^\infty P\left\{X > \frac{k}{n}\right\}.$$

We have

$$\begin{aligned} L_n &= \frac{1}{n} \sum_{k=1}^\infty \sum_{j=k}^\infty P\left\{\frac{j}{n} < X \leq \frac{j+1}{n}\right\} \\ &= \frac{1}{n} \sum_{k=2}^\infty (k-1) P\left\{\frac{k-1}{n} < X \leq \frac{k}{n}\right\} \end{aligned}$$

on rearranging the series. Thus

$$\begin{aligned} L_n &= \sum_{k=2}^\infty \frac{k}{n} P\left\{\frac{k-1}{n} < X \leq \frac{k}{n}\right\} - \sum_{k=2}^\infty \frac{1}{n} P\left\{\frac{k-1}{n} < X \leq \frac{k}{n}\right\} \\ &= \sum_{k=2}^\infty \frac{k}{n} \sum_{(k-1)/n < x_j \leq k/n} p_j - \frac{1}{n} P\left\{X > \frac{1}{n}\right\} \\ &\geq \sum_{k=1}^\infty \sum_{(k-1)/n < x_j \leq k/n} x_j p_j - \frac{1}{n} \\ &= EX - \frac{1}{n}. \end{aligned}$$

Similarly we can show that for every positive integer n

$$U_n \leq EX + \frac{1}{n}.$$

Thus we have

$$EX - \frac{1}{n} \leq \int_0^\infty [1 - F(x)] dx \leq EX + \frac{1}{n}$$

for every positive integer n . Taking the limit as $n \rightarrow \infty$, we see that

$$EX = \int_0^\infty [1 - F(x)] dx.$$

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

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

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

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

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 (see Apostol [3], 361) leads to the following *moments lemma*.

Lemma 2.

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

In Section 6.4 we will construct another proof of (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 $n^\alpha P\{|X| > n\} \rightarrow 0$ as $n \rightarrow \infty$ 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

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

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

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

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 $n^\alpha P\{|X| > n\} \rightarrow 0$ as $n \rightarrow \infty$ for any $\alpha > 0$. Consider, for example, the rv X with pdf

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

The df of X is given by

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

Then for $x > e$

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

and $x^\alpha P\{|X| > x\} \rightarrow \infty$ as $x \rightarrow \infty$ for any $\alpha > 0$. It follows that $E|X|^\alpha = \infty$ for every $\alpha > 0$. In this example we see that $P\{|X| > cx\}/P\{|X| > x\} \rightarrow 1$ as $x \rightarrow \infty$ 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 $L(cx)/L(x) \rightarrow 1$ as $x \rightarrow \infty$ for every $c > 0$. For such a function $x^\alpha L(x) \rightarrow \infty$ for every $\alpha > 0$ (see Feller [29], 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. We again refer the reader to Feller [29].

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

$$(7) \quad \frac{P\{|X| > \alpha k\}}{P\{|X| > k\}} \rightarrow 0 \quad \text{as } k \rightarrow \infty \quad \text{for all } \alpha > 1;$$

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

Proof. Let $\varepsilon > 0$ (we will choose ε later), choose K_0 so large that

$$(8) \quad \frac{P\{|X| > \alpha k\}}{P\{|X| > k\}} < \varepsilon \quad \text{for all } k \geq K_0,$$

and choose K_1 so large that

$$(9) \quad P\{|X| > k\} < \varepsilon \quad \text{for all } k \geq K_1.$$

Let $N = \max(K_0, K_1)$. We have, for a fixed positive integer r ,

$$(10) \quad \frac{P\{|X| > \alpha^r k\}}{P\{|X| > k\}} = \prod_{p=1}^r \frac{P\{|X| > \alpha^p k\}}{P\{|X| > \alpha^{p-1} k\}} \leq \varepsilon^r$$

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

$$(11) \quad P\{|X| > \alpha^r k\} \leq \varepsilon^{r+1}.$$

Next note that, for any fixed positive integer n ,

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

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

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

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

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

Proof. The proof is simple.

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 $E|X|^k < \infty$, 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, \dots, m_k , we can compute $\mu_1, \mu_2, \dots, \mu_k$, and conversely. We have

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

and

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

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

$$(15) \quad \sigma^2 = \mu_2 = EX^2 - (EX)^2.$$

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

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

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*

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

Then

$$\begin{aligned} EX &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\ &= np; \end{aligned}$$

$$\begin{aligned} EX^2 &= E\{X(X-1) + X\} \\ &= \sum k(k-1) \binom{n}{k} p^k (1-p)^{n-k} + np \\ &= n(n-1)p^2 + np; \end{aligned}$$

$$\begin{aligned} \text{var}(X) &= n(n-1)p^2 + np - n^2p^2 \\ &= np(1-p); \end{aligned}$$

$$\begin{aligned} EX^3 &= E\{X(X-1)(X-2) + 3X(X-1) + X\} \\ &= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np; \end{aligned}$$

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

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

Definition 3. A number x satisfying

$$(17) \quad P\{X \leq x\} \geq p, \quad P\{X \geq x\} \geq 1-p, \quad 0 < p < 1,$$

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

$$(18) \quad p \leq F(x) \leq p + P\{X = x\}.$$

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

$$(19) \quad F(x) = p.$$

If F is strictly increasing, (19) has a unique solution. Otherwise there may be many (even uncountably many) solutions of (19), each one of which is then called a quantile of order p .

Definition 4. Let X be an rv with df F . A number x satisfying

$$(20) \quad \frac{1}{2} \leq F(x) \leq \frac{1}{2} + P\{X = x\}$$

or, equivalently,

$$(21) \quad P\{X \leq x\} \geq \frac{1}{2} \quad \text{and} \quad P\{X \geq x\} \geq \frac{1}{2}$$

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*

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

Then $E|X|$ is not finite. The median of the rv X is clearly $x = 0$.

Example 7. Let X be an rv with pmf

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

Then

$$P\{X \leq 0\} = \frac{1}{2} \quad \text{and} \quad P\{X \geq 0\} = \frac{3}{4} > \frac{1}{2}.$$

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

$$P\{X \leq x\} = P\{X = -2\} + P\{X = 0\} = \frac{1}{2}$$

and

$$P\{X \geq x\} = P\{X = 1\} + P\{X = 2\} = \frac{1}{2},$$

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

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

$$P\{X \leq -2\} = \frac{1}{4} > p \quad \text{and} \quad P\{X \geq -2\} = 1 > 1 - p.$$

PROBLEMS 3.2

- Find the expected number of throws of a fair die until a 6 is obtained.
- 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 .
- Let X be an rv with pdf

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

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

- Let X be an rv with pdf

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

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

- 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 .

- Pareto's distribution* with parameters α and β (both α and β positive) is defined by the pdf

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

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.

- 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 .

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

- For the Poisson rv X with pmf

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

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 $\mu_2 = \mu_3 = \lambda$, $\mu_4 = \lambda + 3\lambda^2$.

- For any rv X with $E|X|^4 < \infty$ define

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

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 pmf's of Problems 8 and 9.

- 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 pmf's of Example 5 and Problem 9.

- Prove Theorem 6.

- Prove Theorem 7.

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 rv's. Let X be an rv, and let

$$p_k = P\{X=k\}, \quad k = 0, 1, 2, \dots$$

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

Definition 1. The function defined by

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

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

Example 1. Consider the Poisson rv

$$P\{X = k\} = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

We have

$$P(s) = \sum_{k=0}^{\infty} (s\lambda)^k \frac{e^{-\lambda}}{k!} = e^{-\lambda} e^{s\lambda} = e^{-\lambda(1-s)}, \quad |s| \leq 1.$$

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

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

Then

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

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. The moments of the rv X , if they exist, can be determined by the derivative at the point $s = 1$ of the function $P(s)$. Thus

$$P'(s) = \sum_{k=1}^{\infty} kp_k s^{k-1}, \quad \text{so that } P'(1) = EX \quad \text{if } EX < \infty.$$

$$P''(s) = \sum_{k=2}^{\infty} k(k-1)p_k s^{k-2}, \quad \text{so that } P''(1) = E\{X(X-1)\} \quad \text{if } EX^2 < \infty,$$

and so on.

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

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

Next we consider the important concept of a moment generating function.

Definition 2. Let X be an rv defined on (Ω, \mathcal{F}, P) . The function

$$(2) \quad M(s) = E e^{sX}$$

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 4. Let X have the pmf

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

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 5. Let X have the pdf

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

Then

$$M(s) = \frac{1}{2} \int_0^{\infty} e^{(s-1/2)x} dx = \frac{1}{1-2s}, \quad s < \frac{1}{2}.$$

Example 6. Let X have pmf

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

Then

$$M(s) = E e^{sX} = e^{-\lambda} \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k}{k!} = e^{-\lambda(1-e^s)} \quad \text{for all } s.$$

The following result will be quite useful in what follows.

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 [137], page 460, or Curtiss [20]. See also P.2.14. 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,

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

For the proof of Theorem 2 we refer to Widder [137], pages 446-447. See also P.2.14 and 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:

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

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

Example 7. Let X be an rv with pdf $f(x) = (1/2)e^{-x/2}$, $x > 0$. From Example 5, $M(s) = 1/(1 - 2s)$ for $s < 1/2$. Thus

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

It follows that

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

Example 8. Let X be an rv with pdf $f(x) = 1$, $0 \leq x \leq 1$, and $= 0$ otherwise. Then

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

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

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

Remark 4. Since there exist rv's for which the mgf may not exist, its utility is somewhat limited. It is much more convenient to work with the *characteristic function* of an rv X , which is defined as $E(e^{itX})$, where $i = \sqrt{-1}$, the imaginary unit, and t is any real number. $E(e^{itX})$ exists for every distribution. Moreover, it uniquely determines the distribution of rv X . Since we do not assume a knowledge of complex variables in this book, we will deal with only mgf's whenever they exist.

We next consider the problem of characterizing a distribution from its moments. Let X be an rv with mgf $M(s)$. Since $n!e^{|sx|} \geq |sx|^n$ for $n > 0$, n integral, we see that $E|X|^n < \infty$ for any n . Given the mgf $M(s)$, we can determine EX^n for any n (positive integer) with the help of Theorem 2. Suppose now that moments of all orders exist for an rv X . It does not follow that the mgf exists.

Example 9. Let X be an rv with pdf

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

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

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

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

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

$$(5) \quad \sum_{k=1}^{\infty} \frac{m_k}{k!} s^k$$

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

The proof of this result is much too complicated to be included here, and we refer the reader to original papers by Hamburger [46]. It should be noted that condition (5) is not necessary (see Dharmadhikari [25]).

In particular if for some constant c

$$|m_k| \leq c^k, \quad k = 1, 2, \dots,$$

then

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

and the df of X is uniquely determined.

If not all the moments of an rv X exist, there is no chance of determining the df of X . The df of X is surely not determined uniquely by the moments that do exist.

Example 10. Let X be an rv with pmf

$$P\left\{X = \frac{3^k}{k^2}\right\} = \frac{2}{3^k}, \quad k = 1, 2, \dots$$

Then

$$EX = \sum_{k=1}^{\infty} \frac{2}{k^2} < \infty, \quad \text{and} \quad EX^2 = \sum_{k=1}^{\infty} \frac{2 \cdot 3^k}{k^4} = \infty.$$

Let Y be an rv with pmf

$$P\{Y = 0\} = \frac{1}{3}, \quad \text{and} \quad P\left\{Y = \frac{3^{k+1}}{k^2}\right\} = \frac{2}{3^{k+1}}, \quad k = 1, 2, \dots$$

Then

$$EY = \sum_{k=1}^{\infty} \frac{2}{k^2} < \infty, \quad \text{and} \quad EX = EY.$$

But

$$EY^2 = 2 \sum_{k=1}^{\infty} \frac{3^{k+1}}{k^4} = \infty,$$

and X and Y do not have the same distribution.

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) $\lim_{n \rightarrow \infty} \{(m_{2n})^{1/2n}/2n\}$ is finite.

PROBLEMS 3.3

1. Find the pgf of the rv's with the following pmf's:

- (a) $P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, 2, \dots, n$; $0 \leq p \leq 1$.
- (b) $P\{X = k\} = [e^{-\lambda}/(1 - e^{-\lambda})] (\lambda^k/k!)$, $k = 1, 2, \dots$; $\lambda > 0$.
- (c) $P\{X = k\} = pq^k(1 - q^{N+1})^{-1}$, $k = 0, 1, 2, \dots, N$; $0 < p < 1$, $q = 1 - p$.

2. Let X be an integer-valued rv with pgf $P(s)$. Let a and b be nonnegative integers, and write $Y = aX + b$. 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

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

does the mgf exist?

5. Let X be an rv with pmf

$$P\{X = j\} = p_j, \quad j = 0, 1, 2, \dots$$

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

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

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

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

where $a_j \geq 0$ 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

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

show that the mgf exists and equals

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

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 [106])

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 $\varepsilon > 0$,

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

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

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

where

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

Then

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

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

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

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

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

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

For rv's with finite second-order moments one cannot do better than the inequality in (3).

Example 1.

$$\begin{aligned} P\{X = 0\} &= 1 - \frac{1}{K^2}, & K > 1, \text{ constant,} \\ P\{X = \mp 1\} &= \frac{1}{2K^2}, \\ EX &= 0, & EX^2 &= \frac{1}{K^2}, & \sigma &= \frac{1}{K}, \\ P\{|X| \geq K\sigma\} &= P\{|X| \geq 1\} = \frac{1}{K^2}, \end{aligned}$$

so that equality is achieved.

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

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

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

From Chebychev's inequality

$$P\left\{|X - \frac{1}{2}| \leq 2\sqrt{\frac{1}{12}}\right\} \geq 1 - \frac{1}{4} = .75.$$

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

$$(4) \quad P\{X > x\} \leq \frac{\sigma^2}{\sigma^2 + x^2} \quad \text{if } x > 0,$$

$$(5) \quad P\{X > x\} \geq \frac{x^2}{\sigma^2 + x^2} \quad \text{if } x < 0.$$

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

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

It follows that

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

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

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

Similar proof holds for (5).

Remark 1. 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

$$(7) \quad P\{|X| \geq K\sigma\} \geq \frac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 K^4 - 2K^2\sigma^4} \quad \text{for } K > 1,$$

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

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

as asserted.

Remark 2. 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

$$EX = \frac{1}{2}, \quad \text{var}(X) = \frac{1}{12}, \quad \mu_4 = E\{X - \frac{1}{2}\}^4 = \frac{1}{80},$$

and

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

that is

$$P\{|X - \frac{1}{2}| \leq 2\sqrt{\frac{1}{12}}\} \geq \frac{45}{49} \approx .92,$$

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

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

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

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

Clearly $Q \geq 0$ for all u, v real. It follows (see P. 2.4) that

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

implying that

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

Thus

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

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

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

The equality holds if and only if

$$\beta_k^{1/k} = \beta_{k+1}^{1/(k+1)} \quad \text{for } k = 1, 2, \dots,$$

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

PROBLEMS 3.4

1. For the rv with pdf

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

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 $g(|X|) < \infty$. Show that

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

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

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

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. (a) Let X be an rv with df F , and let g be a strictly convex function on the range of F . Let φ be a Borel-measurable function. Suppose that EX , $E\varphi(X)$, and $Eg(X)$ all exist, and write $EX = \mu$. Let $t(x) = g(\mu) + K(x - \mu)$ be a line of support for g at $x = \mu$. Also, let $h(x) = g(x) - t(x)$. Then for every $\varepsilon > 0$ we have

$$E|\varphi(X)| \leq \sup_{|x-\mu|<\varepsilon} |\varphi(x)| + \left[\sup_{|x-\mu|\geq\varepsilon} \frac{|\varphi(x)|}{h(x)} \right] Eh(X).$$

(This inequality is due to M. Riesz; see, for example, Lukacs [76]. See P.2.3 for definitions of strictly convex functions and line of support.)

- (b) Derive Markov's inequality from Riesz's inequality. (c) Show that for $a > 0$ and $\varepsilon > 0$

$$P\{|X - \mu| > \varepsilon\} \leq Me^{-a\varepsilon} [Ee^{aX} - e^{a\mu}],$$

where $M = \{e^{-a\varepsilon} - 1 + a\varepsilon\}^{-1}$.

(Hint: Take $g(x) = e^{ax}$, $\varphi(x) = 0$ if $|x - \mu| < \varepsilon$, and $= 1$ if $|x - \mu| \geq \varepsilon$;
 $h(x) = e^{a\mu} [e^{a(x-\mu)} - 1 - a(x - \mu)].$)

7. For any rv X , show that

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

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

CHAPTER 4

Random Vectors

4.1 INTRODUCTION

In many experiments an observation is expressible, not as a single numerical quantity, but as a family of several separate numerical quantities. Thus, for example, if a pair of distinguishable dice is tossed, the outcome is a pair (x, y) , where x denotes the face value on the first die, and y , the face value on the second die. Similarly, to record the height and weight of every person in a certain community we need a pair (x, y) , where the components represent, respectively, the height and the weight of a particular individual. To be able to describe such experiments mathematically we must study the *multi-dimensional random variables* or *random vectors*.

In Section 2 we introduce the basic notions involved and study joint, marginal, and conditional distributions. In Section 3 we examine independent random variables and investigate some consequences of independence. Sections 4 and 5 deal with functions of random vectors and their induced distributions. Sections 6 and 7 consider moments and their generating functions, and in Section 8 we study the functional relationship between two dependent random variables.

4.2 RANDOM VECTORS

In this section we study multidimensional rv's. Let (Ω, \mathcal{F}, P) be a fixed but otherwise arbitrary probability space.

Definition 1. The collection $\mathbf{X} = (X_1, X_2, \dots, X_n)$ defined on (Ω, \mathcal{F}, P) into \mathcal{R}_n by

$$\mathbf{X}(\omega) = (X_1(\omega), X_2(\omega), \dots, X_n(\omega)), \quad \omega \in \Omega,$$