$$= E\left[\frac{d}{dt}(Xe^{tX})\right]$$
$$= E[X^2 e^{tX}]$$

and so

$$\phi''(0) = E[X^2]$$

In general, the nth derivative of $\phi(t)$ evaluated at $t = 0$ equals $E[X^n]$; that is,

$$\phi^n(0) = E[X^n], \quad n \geq 1$$

An important property of moment generating functions is that the *moment generating function of the sum of independent random variables is just the product of the individual moment generating functions*. To see this, suppose that X and Y are independent and have moment generating functions $\phi_X(t)$ and $\phi_Y(t)$, respectively. Then $\phi_{X+Y}(t)$, the moment generating function of $X + Y$, is given by

$$\begin{aligned}\phi_{X+Y}(t) &= E[e^{t(X+Y)}] \\ &= E[e^{tX}e^{tY}] \\ &= E[e^{tX}]E[e^{tY}] \\ &= \phi_X(t)\phi_Y(t)\end{aligned}$$

where the next to the last equality follows from Theorem 4.7.4 since X and Y, and thus e^{tX} and e^{tY}, are independent.

Another important result is that the *moment generating function uniquely determines the distribution*. That is, there exists a one-to-one correspondence between the moment generating function and the distribution function of a random variable.

4.9 CHEBYSHEV'S INEQUALITY AND THE WEAK LAW OF LARGE NUMBERS

We start this section by proving a result known as Markov's inequality.

PROPOSITION 4.9.1 MARKOV'S INEQUALITY

If X is a random variable that takes only nonnegative values, then for any value $a > 0$

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

Proof

We give a proof for the case where X is continuous with density f.

$$\begin{aligned} E[X] &= \int_0^\infty xf(x)\,dx \\ &= \int_0^a xf(x)\,dx + \int_a^\infty xf(x)\,dx \\ &\geq \int_a^\infty xf(x)\,dx \\ &\geq \int_a^\infty af(x)\,dx \\ &= a\int_a^\infty f(x)\,dx \\ &= aP\{X \geq a\} \end{aligned}$$

and the result is proved. □

As a corollary, we obtain Proposition 4.9.2.

PROPOSITION 4.9.2 CHEBYSHEV'S INEQUALITY
If X is a random variable with mean μ and variance σ^2, then for any value $k > 0$

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Proof

Since $(X - \mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a = k^2$) to obtain

$$P\{(X - \mu)^2 \geq k^2\} \leq \frac{E[(X - \mu)^2]}{k^2} \tag{4.9.1}$$

But since $(X - \mu) \geq k^2$ if and only if $|X - \mu| \geq k$, Equation 4.9.1 is equivalent to

$$P\{|X - \mu| \geq k\} \leq \frac{E[(X - \mu)^2]}{k^2} = \frac{\sigma^2}{k^2}$$

and the proof is complete. □

The importance of Markov's and Cheybyshev's inequalities is that they enable us to derive bounds on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known. Of course, if the actual distribution were known, then the desired probabilities could be exactly computed and we would not need to resort to bounds.

EXAMPLE 4.9a Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.

(a) What can be said about the probability that this week's production will exceed 75?

(b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60?

SOLUTION Let X be the number of items that will be produced in a week:

(a) By Markov's inequality

$$P\{X > 75\} \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}$$

(b) By Chebyshev's inequality

$$P\{|X - 50| \geq 10\} \leq \frac{\sigma^2}{10^2} = \frac{1}{4}$$

Hence

$$P\{|X - 50| < 10\} \geq 1 - \frac{1}{4} = \frac{3}{4}$$

and so the probability that this week's production will be between 40 and 60 is at least .75. ■

By replacing k by $k\sigma$ in Equation 4.9.1, we can write Chebyshev's inequality as

$$P\{|X - \mu| > k\sigma\} \leq 1/k^2$$

Thus it states that the probability a random variable differs from its mean by more than k standard deviations is bounded by $1/k^2$.

We will end this section by using Chebyshev's inequality to prove the weak law of large numbers, which states that the probability that the average of the first n terms in a sequence of independent and identically distributed random variables differs by its mean by more than ε goes to 0 as n goes to infinity.

Theorem 4.9.3 The Weak Law of Large Numbers

Let $X_1, X_2, \ldots,$ be a sequence of independent and identically distributed random variables, each having mean $E[X_i] = \mu$. Then, for any $\varepsilon > 0$,

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} \to 0 \quad \text{as } n \to \infty$$

Proof

We shall prove the result only under the additional assumption that the random variables have a finite variance σ^2. Now, as

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mu \quad \text{and} \quad \text{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

it follows from Chebyshev's inequality that

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right\} \leq \frac{\sigma^2}{n\epsilon^2}$$

and the result is proved. □

For an application of the above, suppose that a sequence of independent trials is performed. Let E be a fixed event and denote by $P(E)$ the probability that E occurs on a given trial. Letting

$$X_i = \begin{cases} 1 & \text{if } E \text{ occurs on trial } i \\ 0 & \text{if } E \text{ does not occur on trial } i \end{cases}$$

it follows that $X_1 + X_2 + \cdots + X_n$ represents the number of times that E occurs in the first n trials. Because $E[X_i] = P(E)$, it thus follows from the weak law of large numbers that for any positive number ε, no matter how small, the probability that the proportion of the first n trials in which E occurs differs from $P(E)$ by more than ε goes to 0 as n increases.

Problems

1. Five men and 5 women are ranked according to their scores on an examination. Assume that no two scores are alike and all 10! possible rankings are equally likely. Let X denote the highest ranking achieved by a woman (for instance, $X = 2$ if the top-ranked person was male and the next-ranked person was female). Find $P\{X = i\}, i = 1, 2, 3, \ldots, 8, 9, 10$.
2. Let X represent the difference between the number of heads and the number of tails obtained when a coin is tossed n times. What are the possible values of X?
3. In Problem 2, if the coin is assumed fair, for $n = 3$, what are the probabilities associated with the values that X can take on?

4. The distribution function of the random variable X is given

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x}{2} & 0 \le x < 1 \\ \dfrac{2}{3} & 1 \le x < 2 \\ \dfrac{11}{12} & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

- **(a)** Plot this distribution function.
- **(b)** What is $P\{X > \frac{1}{2}\}$?
- **(c)** What is $P\{2 < X \le 4\}$?
- **(d)** What is $P\{X < 3\}$?
- **(e)** What is $P\{X = 1\}$?

5. Suppose you are given the distribution function F of a random variable X. Explain how you could determine $P\{X = 1\}$. (*Hint*: You will need to use the concept of a limit.)

6. The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function between 50 and 150 hours before breaking down? What is the probability that it will function less than 100 hours?

7. The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \le 100 \\ \dfrac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events $E_i, i = 1, 2, 3, 4, 5$, that the ith such tube will have to be replaced within this time are independent.

8. If the density function of X equals

$$f(x) = \begin{cases} ce^{-2x} & 0 < x < \infty \\ 0 & x < 0 \end{cases}$$

find c. What is $P\{X > 2\}$?