==(1)

4.5 Jensen Inequality

Definition 14 Convex function A continuous function $g(\cdot)$ with domain and counterdomain the real line is called *convex* if for every x_0 on the real line, there exists a line which goes through the point $(x_0, g(x_0))$ and lies on or under the graph of the function $g(\cdot)$. ////

Theorem 6 Jensen inequality Let X be a random variable with mean $\mathscr{E}[X]$, and let $g(\cdot)$ be a convex function; then $\mathscr{E}[g(X)] \geq g(\mathscr{E}[X])$.

PROOF Since $g(x)$ is continuous and convex, there exists a line, say $l(x) = a + bx$, satisfying $l(x) = a + bx \leq g(x)$ and $l(\mathscr{E}[X]) = g(\mathscr{E}[X])$. $l(x)$ is a line given by the definition of continuous and convex that goes through the point $(\mathscr{E}[X], g(\mathscr{E}[X]))$. Note that $\mathscr{E}[l(X)] = \mathscr{E}[(a + bX)] = a + b\mathscr{E}[X] = l(\mathscr{E}[X])$; hence $g(\mathscr{E}[X]) = l(\mathscr{E}[X]) = \mathscr{E}[l(X)] \leq \mathscr{E}[g(X)]$ [using property (iv) of expected values (see Theorem 3) for the last inequality]. ////

The Jensen inequality can be used to prove the Rao-Blackwell theorem to appear in Chap. VII. We point out that, in general, $\mathscr{E}[g(X)] \neq g(\mathscr{E}[X])$; for example, note that $g(x) = x^2$ is convex; hence $\mathscr{E}[X^2] \geq (\mathscr{E}[X])^2$, which says that the variance of X, which is $\mathscr{E}[X^2] - (\mathscr{E}[X])^2$, is nonnegative.

4.6 Moments and Moment Generating Functions

The *moments* (or *raw moments*) of a random variable or of a distribution are the expectations of the powers of the random variable which has the given distribution.

==(2)

Definition 15 Moments If X is a random variable, the *rth moment* of X, usually denoted by μ'_r, is defined as

$$\mu'_r = \mathscr{E}[X^r] \qquad (15)$$

if the expectation exists. ////

Note that $\mu'_1 = \mathscr{E}[X] = \mu_X$, the mean of X.

Definition 16 Central moments If X is a random variable, the *rth central moment* of X about a is defined as $\mathscr{E}[(X-a)^r]$. If $a = \mu_X$, we have the *rth central moment* of X about μ_X, denoted by μ_r, which is

$$\mu_r = \mathscr{E}[(X - \mu_X)^r]. \qquad (16)$$

////

Definition 17 Quantile The *qth quantile* of a random variable X or of its corresponding distribution is denoted by ξ_q and is defined as the smallest number ξ satisfying $F_X(\xi) \geq q$. ////

If X is a continuous random variable, then the *qth quantile* of X is given as the smallest number ξ satisfying $F_X(\xi) = q$. See Fig. 6.

Definition 18 Median The *median* of a random variable X. denoted by

med_X, med (X), or $\xi_{.50}$, is the .5th quantile. ////

Remark In some texts the median of X is alternatively defined as any number, say med (X), satisfying $P[X \leq \text{med}(X)] \geq \frac{1}{2}$ and $P[X \geq \text{med}(X)] \geq \frac{1}{2}$. ////

If X is a continuous random variable, then the median of X satisfies

$$\int_{-\infty}^{\text{med}(X)} f_X(x)\,dx = \frac{1}{2} = \int_{\text{med}(X)}^{\infty} f_X(x)\,dx;$$

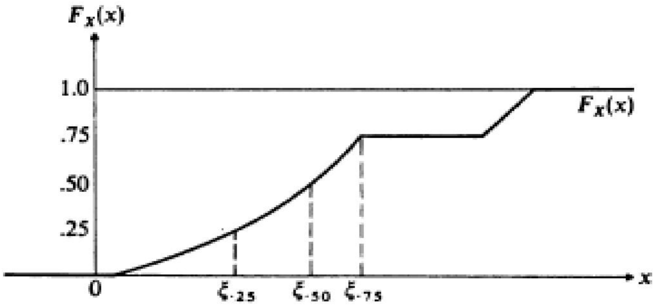

FIGURE 6

so the median of X is any number that has half the mass of X to its right and the other half to its left, which justifies use of the word "median."

We have already mentioned that $\mathscr{E}[X]$, the first moment, locates the "center" of the density of X. The median of X is also used to indicate a central location of the density of X. A third measure of location of the density of X, though not necessarily a measure of central location, is the *mode* of X, which is defined as that point (if such a point exists) at which $f_X(\cdot)$ attains its maximum. Other measures of location [for example, $\frac{1}{2}(\xi_{.25} + \xi_{.75})$] could be devised, but three, mean, median, and mode, are the ones commonly used.

Definition 19 Factorial moment If X is a random variable, the *rth factorial moment of* X is defined as (r is a positive integer):

$$\mathscr{E}[X(X-1)\cdots(X-r+1)]. \tag{18}$$

////

Definition 20 **Moment generating function** Let X be a random variable with density $f_X(\cdot)$. The expected value of e^{tX} is defined to be the *moment generating function* of X if the expected value exists for every value of t in some interval $-h < t < h$; $h > 0$. The moment generating function, denoted by $m_X(t)$ or $m(t)$, is

$$m(t) = \mathscr{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx \tag{19}$$

if the random variable X is continuous and is

$$m(t) = \mathscr{E}[e^{tX}] = \sum_x e^{tx} f_X(x)$$

if the random variable is discrete. ////

One might note that a moment generating function is defined in terms of a density function, and since density functions were defined without reference to random variables (see Definitions 6 and 9), a moment generating function can be discussed without reference to random variables.

If a moment generating function exists, then $m(t)$ is continuously differentiable in some neighborhood of the origin. If we differentiate the moment generating function r times with respect to t, we have

$$\frac{d^r}{dt^r} m(t) = \int_{-\infty}^{\infty} x^r e^{xt} f_X(x)\,dx, \tag{20}$$

and letting $t \to 0$, we find

$$\frac{d^r}{dt^r} m(0) = \mathscr{E}[X^r] = \mu_r', \tag{21}$$

where the symbol on the left is to be interpreted to mean the rth derivative of $m(t)$ evaluated as $t \to 0$. Thus the moments of a distribution may be obtained from the moment generating function by differentiation, hence its name.

$$\begin{aligned} m(t) &= \mathscr{E}\left[1 + Xt + \frac{1}{2!}(Xt)^2 + \frac{1}{3!}(Xt)^3 + \cdots\right] \\ &= 1 + \mu_1' t + \frac{1}{2!}\mu_2' t^2 + \cdots \\ &= \sum_{i=0}^{\infty} \frac{1}{i!}\mu_i' t^i, \end{aligned} \tag{22}$$

from which it is again evident that μ_r' may be obtained from $m(t)$; μ_r' is the coefficient of $t^r/r!$.

EXAMPLE 17 Let X be a random variable with probability density function given by $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$.

$$m_X(t) = \mathscr{E}[e^{tX}] = \int_0^\infty e^{tx} \lambda e^{-\lambda x}\, dx = \frac{\lambda}{\lambda - t} \qquad \text{for } t < \lambda.$$

$$m'(t) = \frac{dm(t)}{dt} = \frac{\lambda}{(\lambda - t)^2} \qquad \text{hence} \quad m'(0) = \mathscr{E}[X] = \frac{1}{\lambda}.$$

And
$$m''(t) = \frac{2\lambda}{(\lambda - t)^3}, \qquad \text{so } m''(0) = \mathscr{E}[X^2] = \frac{2}{\lambda^2}. \qquad ////$$

EXAMPLE 18 Consider the random variable X having probability density function $f_X(x) = x^{-2} I_{[1,\infty)}(x)$. (See Example 13.) If the moment generating function of X exists, then it is given by $\int_1^\infty x^{-2} e^{tx}\, dx$. It can be shown, however, that the integral does not exist for any $t > 0$, and hence the moment generating function does not exist for this random variable X. ////

Definition 21 Factorial moment generating function Let X be a random variable. The *factorial moment generating function* is defined as $\mathscr{E}[t^X]$ if this expectation exists. ////

The factorial moment generating function is used to generate factorial moments in the same way as the raw moments are obtained from $\mathscr{E}[e^{tX}]$ except that t approaches 1 instead of 0. It sometimes simplifies finding moments of discrete distributions.

EXAMPLE 19 Suppose X has a discrete density function given by

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \text{for } x = 0, 1, 2, \ldots.$$

Then

$$\mathscr{E}[t^X] = \sum_{x=0}^\infty \frac{t^x e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}.$$

$$\frac{d}{dt}\mathscr{E}[t^X] = \frac{d}{dt} e^{\lambda(t-1)} = \lambda e^{\lambda(t-1)}; \qquad \text{hence} \quad \frac{d}{dt}\mathscr{E}[t^X]\bigg|_{t=1} = \lambda. \qquad ////$$

In addition to raw moments, central moments, and factorial moments, there are other kinds of moments, called *cumulants*, or *semi-invariants*. Cumulants will be defined in terms of the *cumulant generating function*. We will not make use of cumulants in this book.

Definition 22 **Cumulant and cumulant generating function** The logarithm of the moment generating function of X is defined to be the *cumulant generating function* of X. The *rth cumulant* of X, denoted by $\kappa_r(X)$ or κ_r, is the coefficient of $t^r/r!$ in the Taylor series expansion of the cumulant generating function. ////

PROBLEMS

1 (*a*) Show that the following are probability density functions (p.d.f.'s):

$$f_1(x) = e^{-x} I_{(0,\infty)}(x)$$
$$f_2(x) = 2e^{-2x} I_{(0,\infty)}(x)$$
$$f(x) = (\theta + 1) f_1(x) - \theta f_2(x) \qquad 0 < \theta < 1.$$

(*b*) Prove or disprove: If $f_1(x)$ and $f_2(x)$ are p.d.f.'s and if $\theta_1 + \theta_2 = 1$, then $\theta_1 f_1(x) + \theta_2 f_2(x)$ is a p.d.f.

9 Let $f_X(x) = k(1/\beta)\{1 - [(x-\alpha)/\beta]^2\} I_{(\alpha-\beta,\, \alpha+\beta)}(x)$, where $-\infty < \alpha < \infty$ and $\beta > 0$.

(*a*) Find k so that $f_X(\cdot)$ is a p.d.f., and sketch the p.d.f.

(*b*) Find the mean, median, and variance of X.

(*c*) Find $\mathscr{E}[|X - \alpha|]$.

(*d*) Find the qth quantile of X.

10 Let $f_X(x) = \frac{1}{2}\{\theta I_{(0,1)}(x) + I_{[1,2]}(x) + (1-\theta) I_{(2,3)}(x)\}$, where θ is a fixed constant satisfying $0 \le \theta \le 1$.

(*a*) Find the c.d.f. of X.

(*b*) Find the mean, median, and variance of X.

11 Let $f(x;\, \theta) = \theta f(x;\, 1) + (1-\theta) f(x;\, 0)$, where θ is a fixed constant satisfying $0 \le \theta \le 1$. Assume that $f(\cdot\,;0)$ and $f(\cdot\,;1)$ are both p.d.f.'s.

(*a*) Show that $f(\cdot\,;\theta)$ is also a p.d.f.

(*b*) Find the mean and variance of $f(\cdot\,;\theta)$ in terms of the mean and variance of $f(\cdot\,;0)$ and $f(\cdot\,;1)$, respectively.

(*c*) Find the m.g.f. of $f(\cdot\,;\theta)$ in terms of the m.g.f.'s of $f(\cdot\,;0)$ and $f(\cdot\,;1)$.