Random variables

Remark on Notations

1. When X is a number chosen uniformly from a data set, What I call $\mathbb{P}(X = k)$ is called Freq$[k, X]$ in the courseware.

2. When X is a random variable, what I call $F(x)$ or $F_X(x)$ is called cumdist[x, X] in the courseware..

3. What I call $E[X]$ is called Expect$[X]$ in the courseware.

Introduction

Often when an experiment is performed, we are mainly interested in some function of the outcome as opposed to the actual outcome itself. For instance, in tossing two dice we are often interested in the sum of the two dice and are not really concerned with the separate values of each die. A real-valued function defined on a sample space is known as a *random variable*. Because the value of a random variable is determined by the outcome of the experiment, we can find the probability that the random variable takes on each value.

Example 1 *Suppose that you toss 3 fair coins. For each Heads, you win \$1, and for each Tails, you lose \$1. Let X be you net winning. Then X is a random variable that can take the values $-3, -1, 1, 3$. And we have*

$$\mathbb{P}(X = -3) = \mathbb{P}(X = 3) = \frac{1}{8}, \quad \mathbb{P}(X = -1) = \mathbb{P}(X = 1) = \frac{3}{8}.$$

Example 2 *Suppose that we are giving a coin having probability $p \in (0, 1)$ coming up Heads. This coin is flipped until a Heads occurs. Let X be the number of flips needed. Then X is a random variable that can takes on the values $1, 2, \ldots$. And we have*

$$\mathbb{P}(X = n) = (1 - p)^{n-1}p, \quad n = 1, 2, \ldots.$$

By the way, this is a Geometric(p) random variable.

For any random variable X, the function F defined by

$$F(x) = \mathbb{P}(X \leq x), \quad -\infty < x < \infty$$

is called the *cumulative distribution function*, or simply, the *distribution function* of X. Thus the distribution function specifies, for all real x, the probability that the random variable X is less than or equal to X. It is obvious that the distribution function F is an increasing function of x. We will discuss other properties of distribution functions later.

A random variable that can take on at most countably many values is called a *discrete random variable.* For a discrete random variable X, we define the *probability mass function p of X* by

$$p(x) = \mathbb{P}(X = x), \quad -\infty < x < \infty.$$

The probability mass function p is positive for at most a countable number of values of x. That is, if X must assume one of the values $x_1, x_2, \ldots$, then

$$p(x_i) \geq 0, \quad i = 1, 2, \ldots,$$
$$p(x) = 0, \quad \text{for all other values of } x.$$

Since X must assume one of the values x_i, we have

$$\sum_{i=1}^{\infty} p(x_i) = 1.$$

Given any nonnegative function $p(x)$ defined on $\mathbb{R}$ such that $p(x)$ is positive for at most countably many values of x and that

$$\sum_{x} p(x) = 1,$$

there is a random variable X with $p(x)$ as its probability mass function.

The (cumulative) distribution function F of a discrete random variable X can be expressed in terms of the mass function p of X:

$$F(x) = \sum_{t:t \leq x} p(t), \quad -\infty < x < \infty.$$

It is often instructive to present the probability mass function in a graphical format by plotting $p(x_i)$ on the y-axis against the x_i on the x-axis. You have seen plenty of examples of this sort in the courseware.

If X is a discrete random variable whose possible values are $x_1, x_2, \ldots$ where $x_1 < x_2 < \cdots$, then its distribution function F is a step function, that is the value of F is constant in the intervals $[x_{i-1}, x_i)$ and then takes a step (or jump) of size $p(x_i)$ at x_i. Thus in this case,

$$p(x_i) = F(x_i) - F(x_i-), \quad i = 1, 2, \ldots.$$

Example 3 *If X is a discrete random variable with probability mass function given by*

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8},$$

the its distribution function F is given by

$$F(x) = \begin{cases} 0 & x < 1, \\ \frac{1}{4} & 1 \leq x < 2 \\ \frac{3}{4} & 2 \leq x < 3 \\ \frac{7}{8} & 3 \leq x < 4 \\ 1 & 4 \leq x. \end{cases}$$

In the courseware there are examples of the graphs of distribution functions.

Expectations of discrete random variables

One of the most important concepts in probability theory is that of expectation of a random variable. If X is a discrete random variable with probability mass function p, the *expectation* or *expected value* of X, denoted by $E[X]$, is defined by

$$E[X] = \sum_{x:p(x)>0} p(x).$$

Example 4 *Let X be the random variable in Example 1. Then*

$$E[X] = (-3) \cdot \frac{1}{8} + (-1) \cdot \frac{3}{8} + 1 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = 0.$$

Example 5 *Let X be the random variable in Example 2. Then, by putting $q = 1 - p$,*

$$
\begin{aligned}
E[X] &= \sum_{n=1}^{\infty} nq^{n-1}p = p\sum_{n=0}^{\infty} \frac{d}{dq}(q^n) \\
&= p\frac{d}{dq}\left(\sum_{n=0}^{\infty} q^n\right) = p\frac{d}{dq}\left(\frac{1}{1-q}\right) \\
&= \frac{p}{(1-q)^2} = \frac{1}{p}.
\end{aligned}
$$

Expectation of a function of a discrete random variable

Suppose that we are given a discrete random variable along with its probability mass function, and that we want to compute the expectation of some function of X, say $g(X)$. How can we accomplis this? One way is as follows: Since $g(X)$ is itself a discrete random variable, it has a probability mass function which can be determined from the probability mass function of X. Once we get the probability mass function of $g(X)$, we can find its expectation easily. However, the following result gives a more convenient method.

Proposition 6 *If X is a discrete random variable with mass function p, then for any real-valued function g,*

$$E[g(X)] = \sum_{x} g(x)p(x).$$

As a consequence of this theorem, we immediately get the following:

Corollary 7 *If X is a a discrete random variable, and a, b are constants, then*

$$E[aX + b] = a\,E[X] + b.$$

This appears in the courseware, in a slightly different form.

Example 8 *Let X be a random variable that takes on the values $-1, 0, 1$ with respective probabilities*

$$\mathbb{P}(X = -1) = .2, \quad \mathbb{P}(X = 0) = .5, \quad \mathbb{P}(X = 1) = .3.$$

3

Find $E[X^2]$.

Solution 1 *Letting $Y = X^2$, then the probability mass function of Y is given by*

$$\mathbb{P}(Y = 1) = .5, \quad \mathbb{P}(Y = 0) = .5.$$

Hence $E[X^2] = E(Y) = 1(.5)+)(.5) = .5$.

Solution 2 *Applying the above proposition,*

$$E[X^2] = (-1)^2(.2) + 0^2(.5) + 1^2(.3) = .5.$$

Variances of discrete random variables

Given a random variable X, it would be extremely useful if we are able to summarize the essential statistical properties of X by certain suitable defined measure. One such measure is $E[X]$, the expectation (mean) of X. However, $E[X]$ does not tell us anything about how spreadout the random variable is from its mean. For example, the following three random variables Y and Z with probability mass functions given by

$$\mathbb{P}(Y = 1) = \mathbb{P}(Y = -1) = \frac{1}{2}, \quad \mathbb{P}(Z = 100) = \mathbb{P}(Z = -100) = \frac{1}{2},$$

both have mean 0, but the second variable spread out much more than the first variable.

We measure the spreadout of a random variable from its mean by the following quantity.

If X is a random variable with mean μ, then the *variance* of X, denoted by $\text{Var}(X)$, is defined by

$$\text{Var}(X) = E[(X - \mu)^2].$$

An alternative formula for $\text{Var}(X)$ is derived a s follows (in the discrete case):

$$
\begin{aligned}
\text{Var}(X) &= E[(X - \mu)^2] = \sum_x (x - \mu)^2 p(x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2.
\end{aligned}
$$

That is

$$\text{Var}(X) = E[X^2] - (E[X])^2.$$

The following result follows immediately from Proposition **??**.

Proposition 9 *If X is a discrete random variable, a, b are constants, then*

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

Example 10 *Let X be the outcome when a fair die is rolled. Find $\text{Var}(X)$. We can easily find $E[X] = \frac{7}{2}$. Also*

$$E[X^2] = 1^2(\frac{1}{6}) + 2^2(\frac{1}{6}) + 3^2(\frac{1}{6}) + 4^2(\frac{1}{6}) + 5^2(\frac{1}{6}) + 6^2(\frac{1}{6}) = \frac{91}{6}.$$

Hence $\text{Var}(X) = \frac{91}{6} = (\frac{7}{2})^2 = \frac{35}{12}$.

Properties of distributions functions

If F is the (cumulative) distribution function of a random variable X, then it follows from the definition of the distribution function and properties of probability measures that F satisfies the following properties:

1. F is a nondecreasing function; that is, if $x < y$, then $F(x) \leq F(y)$.

2. $\lim_{x \to \infty} F(x) = 1$.

3. $\lim_{x \to -\infty} F(x) = 0$.

4. F is right continuous; that is, for any x, we have $F(x+) = F(x)$, where $F(x+) = \lim_{y \downarrow x} F(y)$ is the right limit at x.

All probability questions about X can be answered in terms of F. For instance, for any real numbers a, b,

$$
\begin{aligned}
\mathbb{P}(X < a) &= F(a-), \\
\mathbb{P}(X = a) &= F(a) - F(a-), \\
\mathbb{P}(X > a) &= 1 - F(a), \\
\mathbb{P}(a < X \leq b) &= F(b) - F(a), \\
\mathbb{P}(a \leq X \leq b) &= F(b) - F(a-), \\
\mathbb{P}(a < X < b) &= F(b-) - F(a), \\
\mathbb{P}(a \leq X < b) &= F(b-) - F(a-),
\end{aligned}
$$

If F is any function on $\mathbb{R}$ satisfying condition 1 through 4 above, then it is the distribution function of some random variable.

Here is an example of finding probabilities using the distribution function.

Example 11 *The distribution function of a random variable X is given by*

$$
F(x) = \begin{cases}
0 & x < 0, \\
\frac{x}{2} & 0 \leq x < 1 \\
\frac{2}{3} & 1 \leq x < 2 \\
\frac{11}{12} & 2 \leq x < 3 \\
1 & 3 \leq x.
\end{cases}
$$

Find (a) $\mathbb{P}(X < 3)$, (b) $\mathbb{P}(X = 1)$, (c) $\mathbb{P}(X > \frac{1}{2})$ and (d) $\mathbb{P}(2 < X \leq 4)$.