MATH STAT 465
EXPECTED VALUE OF A RANDOM VARIABLE
February 27, 2004

ERIC S. KEY

Abstract. We define the expected value for all random variables, and give two important theorems on the calculation of expected value.

1. Introduction

Expected value is most easily defined for simple random variables. Using this definition, expected value is then defined for positive random variables and then any random variable. We conclude with some theorems that allow us to compute expected value in special cases.

2. Simple Random Variables

Suppose that $\{A_1, A_2, \ldots, A_N\}$ is a partition of the sample space Ω, that I_j is the indicator of the event A_j, and that $(a_1, a_2, \ldots a_N)$ is a sequence of real numbers. Let

$$X = a_1 I_1 + a_2 I_2 + \cdots + a_N I_N.$$

As we know, X is called a **simple random variable** and has range $\{a_1, a_2, \ldots, a_N\}$. We define the **expected value of** X, denoted by E[X], by

$$\mathrm{E[X] = a_1\,Pr(A_1) + a_2\,Pr(A_2) + \cdots + a_N\,Pr(A_N)}.$$

Now for something a little subtle. There is no requirement that the values a_1, $a_2 \ldots$, a_N are different. Let $\{r_1, r_2, \ldots, r_M\}$ be a listing of the *distinct* elements of the range of X, that is $r_j \neq r_k$ if $j \neq k$. Then it is easy to check that

$$\mathrm{E[X] = r_1\,Pr(\{\omega : X(\omega) = r_1\}) + \cdots + r_M\,Pr(\{\omega : X(\omega) = r_M\})}.$$

This shows that the expected value of X does not depend on which partition of Ω is chosen in the formula that gives X. This gives the important result:

Theorem 1. *If X and Y are simple random variables, and a is a constant, then*

$$\mathrm{E[aX + Y] = aE[X] + E[Y]}.$$

3. Positive random variables

Suppose that X is any random variable whose range is contained in the non-negative real numbers. We now will define its expected value. To do so, recall that a set of real numbers, U, is said to be **bounded** if there is some real number M so that if $u \in U$ then $|u| \leq M$. For example, the set of integers is not bounded, but the set of rational numbers whose squares are less than 2 is bounded. If U is a bounded set, then the **least upper bound of** U is the smallest number M that is an upper bound for U. It is a fact about the real numbers that every bounded set has exactly one least upper bound.

Associate to X the set

$$U = \{\mathrm{E[Y]} : 0 \leq \mathrm{Y} \leq \mathrm{X} \text{ and Y is a simple random variable}\}.$$

If U is a bounded set, we define E[X] to be the least upper bound of U. If U is not a bounded set, then the expected value of X cannot be defined.

This construction is a generalization of the idea of integration that you studied in calculus. Observe that if X is a simple random variable itself, then we get the same value for E[X] as in the previous section.

4. General random variables

If u is a real number, define the **positive part of** u, denote u^+ by $u^+ = (|u| + u)/2$, and the **negative part of** u, denoted u^- by $u^- = (|u| - u)/2$. Observe that $u^+ \geq 0$, $u^- \geq 0$, and $u = u^+ - u^-$. If X is any random variable and $\mathrm{E[X^+]}$ and $\mathrm{E[X^-]}$ are defined, then we define $\mathrm{E[X]} = \mathrm{E[X^+]} - \mathrm{E[X^-]}$.

5. Three important theorems

Theorem 2. *If* E[X] *and* E[Y] *are defined, and a is a real number, then* $\mathrm{E[aX+Y]} = \mathrm{aE[X]} + \mathrm{E[Y]}$.

As application of this theorem, notice that since $(X - \mathrm{E[X]})^2 = \mathrm{X}^2 - 2\mathrm{E[X]} \cdot \mathrm{X} + (\mathrm{E[X]})^2$,

$$\mathrm{Var[X]} := \mathrm{E[(X - E[X])^2]} = \mathrm{E[X^2 - 2E[X] \cdot X + (E[X])^2]} = \mathrm{E[X^2]} - (\mathrm{E[X]})^2.$$

Theorem 3. *If X is a discrete random variable with range R, and g is a function defined on the range of X such that $g(X)$ is a random variable and*

$$\sum_{r \in R} |g(r)| \Pr(\{\omega : X(\omega) = r\}) < \infty,$$

then

$$\mathrm{E[g(X)]} = \sum_{\mathrm{r} \in \mathrm{R}} \mathrm{g(r)} \Pr(\{\omega : \mathrm{X}(\omega) = \mathrm{r}\}).$$

Theorem 4. *If X is an absolutely continuous random variable with range R and density f, and g is a function defined on the range of X such that $g(X)$ is a random variable and*

$$\int_{-\infty}^{\infty} |g(r)| \cdot f(r)\, dr < \infty,$$

then

$$\mathrm{E[g(X)]} = \int_{-\infty}^{\infty} \mathrm{g(r) \cdot f(r)\, dr}.$$

Note that these theorems can be applied with $g(r) = r$ to yield formulae for E[X] itself.