1.6 Expected Value

A random variable $X$ is **continuous** if there is a function $f$, called its density function, so that $P(X \leq x) = \int_{-\infty}^{x} f(t)dt$ for all $x$. A random variable is **discrete** if it can only take a countable number of different values. In elementary textbooks you usually see two separate definitions for expected value:

$$E[X] = \begin{cases} \sum_{i} x_{i}P(X = x_{i}) & \text{if } X \text{ is discrete} \\ \int x f(x)dx & \text{if } X \text{ is continuous with density } f. \end{cases}$$

But it’s possible to have a random variable which is neither continuous nor discrete. For example, with $U \sim U(0,1)$, the variable $X = UI_{U>0.5}$ is neither continuous nor discrete. It’s also possible to have a sequence of continuous random variables which converges to a discrete random variable – or vice versa. For example, if $X_{n} = U/n$, then each $X_{n}$ is a continuous random variable but $\lim_{n \to \infty} X_{n}$ is a discrete random variable (which equals zero). This means it would be better to have a single more general definition which covers all types of random variables. We introduce this next.

A **simple** random variable is one which can take on only a finite number of different possible values, and its expected value is defined as above for discrete random variables. Using these, we next define the expected value of a more general non-negative random variable. We will later define it for general random variables $X$ by expressing it as the difference of two nonnegative random variables $X = X^+ - X^-$, where $x^+ = \max(0, x)$ and $x^- = \max(-x, 0)$. 

Definition 1.24 If $X \geq 0$, then we define

$$E[X] \equiv \sup_{\text{all simple variables } Y \leq X} E[Y].$$

We write $Y \leq X$ for random variables $X, Y$ to mean $P(Y \leq X) = 1$, and this is sometimes written as "$Y \leq X$ almost surely" and abbreviated "$Y \leq X$ a.s.". For example if $X$ is nonnegative and $a \geq 0$ then $Y = aI_{X \geq a}$ is a simple random variable such that $Y \leq X$. And by taking a supremum over "all simple variables" we of course mean the simple random variables must be measurable with respect to some given sigma field. Given a nonnegative random variable $X$, one concrete choice of simple variables is the sequence $Y_n = \min(\lfloor 2^n X \rfloor / 2^n, n)$, where $\lfloor x \rfloor$ denotes the integer portion of $x$. We ask you in exercise 17 at the end of the chapter to show that $Y_n \uparrow X$ and $E[X] = \lim_n E[Y_n]$.

Another consequence of the definition of expected value is that if $Y \leq X$, then $E[Y] \leq E[X]$.

Given any random variable $X \geq 0$ with $E[X] < \infty$, and any $\epsilon > 0$, we can find a simple random variable $Y$ with $E[X] - \epsilon \leq E[Y] \leq E[X]$. Our definition of the expected value also gives what is called the Lebesgue integral of $X$ with respect to the probability measure $P$, and is sometimes denoted $E[X] = \int X \, dP$.

So far we have only defined the expected value of a nonnegative random variable. For the general case we first define $X^+ = X I_{X \geq 0}$ and $X^- = -X I_{X < 0}$ so that we can define $E[X] = E[X^+] - E[X^-]$, with the convention that $E[X]$ is undefined if $E[X^+] = E[X^-] = \infty$. 
Remark 1.27 The definition of expected value covers random variables which are neither continuous nor discrete, but if $X$ is continuous with density function $f$ it is equivalent to the familiar definition $E[X] = \int xf(x)dx$. For example when $0 \leq X \leq 1$ the definition of the Riemann integral in terms of Riemann sums implies, with $[x]$ denoting the integer portion of $x$,

$$\int_0^1 xf(x)dx = \lim_{n \to \infty} \sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} xf(x)dx$$

$$\leq \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{i+1}{n} P(i/n \leq X \leq \frac{i+1}{n})$$

$$= \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{i}{n} P(i/n \leq X \leq \frac{i+1}{n})$$

$$= \lim_{n \to \infty} E([nX]/n)$$

$$\leq E[X],$$

where the last line follows because $[nX]/n \leq X$ is a simple random variable.

Using that the density function $g$ of $1-X$ is $g(x) = f(1-x)$, we obtain

$$1 - E[X] = E[1 - X]$$

$$\geq \int_0^1 xf(1-x)dx$$

$$= \int_0^1 (1-x)f(x)dx$$

$$= 1 - \int_0^1 xf(x)dx.$$