Expectation

Suppose $\mathcal{E}$ is an experiment the set of possible outcomes of which is the sample space $S$. Let $E$ be an event, which is to say that $E \subseteq S$. Let us recall the relative frequency interpretation of $P(E)$, the probability of $E$. Let

$$s_1, s_2, \ldots, s_n, \ldots$$

be the outcomes of a never ending sequence of independent repetitions of $\mathcal{E}$. Let $\nu(E, n)$, $n = 1, 2 \ldots$ be the number of occurrences of $E$ in the first $n$ repetitions of $\mathcal{E}$; that is, $\nu(E, n)$ is the number of $i \in \{1, \ldots, n\}$ such that $s_i \in E$. Then

$$P(E) = \lim_{n \to \infty} \frac{\nu(E, n)}{n}.$$

Now suppose $X$ is a random variable on $S$ with finite range $x_1, \ldots, x_N$. We define the expectation $E(X)$ of $X$ to be

$$\lim_{n \to \infty} \frac{X(s_1) + X(s_2) + \cdots + X(s_n)}{n}$$

which is just the limit of the running average values of $X$ on the sequence of outcomes $s_1, s_2, \ldots, s_n, \ldots$. I claim that

$$E(X) = \sum_{i=1}^{N} x_i P(X = x_i).$$

Indeed,

$$\frac{X(s_1) + X(s_2) + \cdots + X(s_n)}{n} = \frac{\sum_{i=1}^{N} x_i \nu(X = x_i, n)}{n} = \frac{\sum_{i=1}^{N} x_i \nu(X = x_i, n)}{n} \to \sum_{i=1}^{N} x_i P(X = x_i)$$

$n \to \infty$.

Now suppose $X$ is a continuous random variable with range equal to $(a, b)$. For each $N = 1, 2, \ldots$ define the discrete random variable $X_N$ by requiring that

$$X_N = a + \frac{j}{N}(b - a)$$

if

$$a + \frac{j - 1}{N}(b - a) \leq X < a + \frac{j}{N}(b - a), \quad j = 1, \ldots, N.$$

Since $|X - X_N| \leq \frac{1}{N}$ for $N = 1, 2, \ldots$, $X_N$ is a better and better approximation to $X$ as $N \to \infty$. Now

$$E(X_N) = \sum_{j=1}^{N} (a + \frac{j}{N}(b - a)) P(\frac{j - 1}{N} \leq X < \frac{j}{N})$$

$$= \sum_{j=1}^{N} (a + \frac{j}{N}(b - a)) \int_{a + \frac{j - 1}{N}(b - a)}^{a + \frac{j}{N}(b - a)} f_X(x) \, dx$$

$$\to \int_{a}^{b} x f_X(x) \, dx$$

as $N \to \infty$. Thus it seems reasonable to define

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx$$
whenever $X$ is a continuous random variable and to expect that the expectation operator will have all of the properties it has in the discrete case.

Some of these properties are:

- $E(c) = c$;
- $E(cX) = cE(X)$;
- $E(X + Y) = E(X) + E(Y)$;
- $E(\phi(X_1, \ldots, X_n)) = \sum_{x_1, \ldots, x_n} \phi(x_1, \ldots, x_n) p_{X_1, \ldots, X_n}(x_1, \ldots, x_n)$ if $(X_1, \ldots, X_n)$ is discrete;
- $E(\phi(X_1, \ldots, X_n)) = \int \int \cdots \int \phi(x_1, \ldots, x_n) f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n$ if $(X_1, \ldots, X_n)$ is continuous.