Random Variables and Expectation

A *random variable* arises when we assign a numeric value to each elementary event. For example, if each elementary event is the result of a series of three tosses of a fair coin, then X = "the number of Heads" is a random variable. Associated with any random variable is its *probability distribution* (sometimes called its *density function*), which indicates the likelihood that each possible value is assumed. For example, Pr(X=0) = 1/8, Pr(X=1) = 3/8, Pr(X=2) = 3/8, and Pr(X=3) = 1/8.

The *cumulative distribution function* indicates the likelihood that the random variable is less-than-or-equal-to any particular value. For example, $\Pr(X \le x)$ is 0 for $x < 0$, 1/8 for $0 \le x < 1$, 1/2 for $1 \le x < 2$, 7/8 for $2 \le x < 3$, and 1 for all $x \ge 3$.

Two random variables X and Y are *independent* if all events of the form "$X \le x$" and "$Y \le y$" are independent events.

The *expected value* of X is the average value of X, weighted by the likelihood of its various possible values. Symbolically,

$$E[X] = \sum_x x \cdot \Pr(X = x)$$

where the sum is over all values taken by X with positive probability. Multiplying a random variable by any constant simply multiplies the expectation by the same constant, and adding a constant just shifts the expectation:

$$E[kX+c] = k \cdot E[X]+c \ .$$

For any event A, the *conditional expectation* of X given A is defined as

$$E[X|A] = \Sigma_x \, x \cdot \Pr(X=x \mid A) \ .$$

A useful way to break down some calculations (when your natural response to "What's E[X]?" is, "Well, it depends on whether A occurs or not") is $E[X] = E[X|A] \cdot \Pr(A) + E[X|A^c] \cdot \Pr(A^c)$.

The expected value of the sum of several random variables is equal to the sum of their expectations, e.g.,

$$E[X+Y] = E[X] + E[Y] \ .$$

On the other hand, the expected value of the product of two random variables is not necessarily the product of the expected values. For example, if they tend to be "large" at the same time, and "small" at the same time, $E[XY] > E[X] \cdot E[Y]$, while if one tends to be large when the other is small, $E[XY] < E[X] \cdot E[Y]$. However, in the special case in which X and Y are independent, equality *does* hold: $E[XY] = E[X] \cdot E[Y]$.

Why is expected value important? As you'll discuss in your core Finance course, a well-diversified investor will want the managers in all the companies in which he or she has invested to consistently make decisions which maximize expected (after-tax) return on investment.

Let the random variable X represent uncertainty you're facing (for example, the price of a barrel of oil six months from now), and let p represent a policy you might adopt (such as the quantity of oil on which you might purchase a six-month call option today).

Your actual profit six months from now, looked at from today's perspective, is Profit(X, p) (that is, it depends on both the value ultimately taken by X, and on your choice of p). It takes different values as X varies, and therefore your profit is itself a random variable.

In other words, when you choose a policy, you're actually choosing between random variables. How should you make that choice?

If you work for a publicly-traded firm, the shareholders will want you to choose the policy which maximizes E[Profit(X, p)].

If you're making a personal decision, your choice will depend on your attitude towards risk, which can be represented (if you satisfy certain rationality criteria) by a utility function which associates changes in net wealth with changes in net "happiness," and you will **want** to choose the policy that maximizes your expected utility.

[For example, if my utility from a positive financial reward of $w is the square-root of w (one particular utility function over positive gains), then I'm indifferent between receiving $250,000 for sure, and having a 50% chance of receiving $1,000,000 (and otherwise nothing), since both offer me the same expected utility.]

Your personal preferences might not always coincide with those of the shareholders: That's why executive compensation packages must be set carefully. But ultimately, you always will want to make decisions which maximize *some* expected payoff.