Random Variables

Random Variables – A random variable is a real valued function defined on the sample space of an experiment. Associated with each random variable is a probability density function (pdf) for the random variable. The sample space is also called the support of a random variable.

Random variables can be classified into two categories based on their support; discrete or continuous. A discrete random variable is a random variable for which the support is a discrete set, otherwise the random variable is continuous.

Discrete Random Variable – For a discrete random variable, it is useful to think of the random variable and its pdf together in a probability distribution table.

Example: A fair coin is tossed three times. Let X = the random variable representing the total number of heads that turn up. Then we have *supp(X)* = {0,1,2,3}, a discrete set. The probability distribution table for X is

X	$f_X(x) = \Pr(X=x)=p(x)$
0	1/8
1	3/8
2	3/8
3	1/8

The pdf for X is the second column of the table. Note that for a discrete random variable, the pdf evaluated at a specific value of the random variable X equals the probability that the random variable X equals the specific value.

Continuous Random Variable – Problems involving continuous random variables will often state the pdf explicitly and ask for probabilities such as $\Pr(a < X < b)$. In fact, for a continuous random variable, $\Pr(X=x) = 0$ always. We still call $f_X(x)$ the density at x, but it is not equal to $\Pr(X=x)$. Notice that

$$\Pr(a < X < b) = \Pr(a \le X < b) = \Pr(a < X \le b) = \Pr(a \le X \le b).$$

Each of these probabilities is calculated as the value of $\int_a^b f(x)dx$.

Mixed Distributions – These are random variables that have certain points that have non-zero probability (a point mass) and also certain intervals with continuous pdf. We will see an example of a mixed distribution soon.

Other Random Variable Concepts and Relationships Among Them
The cumulative distribution function (aka distribution function) for the random variable X is defined by $F(x) = \Pr(X \le x)$. If the random variable X happens to continuous, then the relationship between the cdf and pdf is $F(x) = \int_{-\infty}^{x} f(t)dt,$ and so by the FTC $F'(x) = f(x).$

The survival function for the random variable X is defined by $S(x) = \Pr(X > x) = 1 - F(x).$ Note that $S'(x) = -F'(x) = -f(x)$ for continuous random variables.

For a continuous random variable X, the hazard rate (or failure rate) is defined by $h(x) = \frac{f(x)}{S(x)} = -\frac{d}{dx}\left[\ln(S(x))\right].$ Later, when studying for Exam MLC, we will call this the *force of mortality*.