MthStat 465, Spring 2005, Lecture Number 11
Distribution Functions and Probability Mass Functions of Random Variables.

To each random variable R we may assign two functions whose domain is the real numbers and whose range is contained in $[0, 1]$. The first is the **probability mass function** of R, denoted p_R. For each real number t we define $p_R(t) = \Pr(\{s \in S : R(s) = t\})$. Note that if t is not in the range of R then $p_R(t) = 0$.

The second function is the **probability distribution function** of R, denoted by $F_R(t)$. For each real number t we define $F_R(t) = \Pr(\{s \in S : R(s) \leq t\})$. Note that $F_R(t) = 0$ if t is smaller than any range value of R and $F_R(t) = 1$ if t is greater than any range value of R.

For example, suppose that the values of a random variable R give real numbers drawn at random from the interval $[0, 1]$. By this we mean that for any $0 \leq a \leq b \leq 1$, $\Pr(a < R \leq b) = b - a$. In particular, since for any $0 < a < b \leq 1$ we have

$$0 \leq \Pr(R = b) \leq \Pr(a < R \leq b) = b - a,$$

we see by letting a approach b that $\Pr(R = b) = 0$. Therefore, the probability mass function of R is always equal to 0. As for the distribution function of R, if $t < 0$, the event $\{R \leq t\}$ is empty, so $F_R(t) = 0$. If $t = 0$, since we have $\{R \leq 0\} = \{R < t\} \cup \{R = 0\}$ and $\{R < t\} = \emptyset$, we have $F_R(0) = 0$ too. On the other hand, if $t \geq 1$, $\{R \leq 1\}$ is the whole sample space, so $F_R(t) = 1$. Finally, if $0 < t < 1$ then

$$\begin{aligned}
\{R \leq t\} &= \{R \leq 0\} \cup \{0 < R \leq t\} \\
F_R(t) &= F_R(0) + \Pr(\{0 < R \leq t\}) \\
&= 0 + (t - 0) \\
&= t.
\end{aligned}$$

If we make a new random variable S by the rule $S(s) = (R(s))^2$, S still has range $[0, 1]$ so $F_S(t) = 0$ if $t \leq 0$ and $F_S(t) = 1$ if $t \geq 1$. However, if $t \in (0, 1)$,

$$\begin{aligned}
F_S(t) &= \Pr(\{S \leq t\}) \\
&= \Pr(\{S \leq 0\} + \Pr(\{0 < S \leq t\}) \\
&= 0 + \Pr(\{0 < R^2 \leq t\}) \\
&= \Pr(\{0 < R \leq \sqrt{t}\}) \\
&\quad \text{because } 0 \leq R \leq 1 \\
&= \sqrt{t} - 0 \\
&= \sqrt{t}
\end{aligned}$$

It is clear that all distribution functions have the following properties:

1. $F_R(t) \to 1$ as $t \to +\infty$;
2. $F_R(t) \to 0$ as $t \to -\infty$;
3. If $x < y$ then $F_R(x) \leq F_R(y)$.

Although distribution functions are need not be continuous, they do also have the following properties:

1. $F_R(x) \to F(y)$ as $x \searrow y$;
2. $F_R(y) - \lim_{x \nearrow y} F(x) = \Pr(\{R = y\}) = p_R(y)$

The first of these properties says that distribution functions are continuous from the right. The second says that the size of the jump of the distribution function at y is the probablity that the random variable is equal to y, and provides the link between mass functions and distribution functions. In particular, it shows that the mass function can always be recovered from the distribution function. Our example with $S = R^2$ shows that the reverse is not always possible.