MthStat 465, Spring 2005, Lecture Number 6
Probability Models

A probability model has three components.

Sample Space:: The sample space in a probability model is an non-empty set. Each element of the sample space should correspond to a different possible outcome of the experiment. For example, if a coin is to be tossed twice, and it is supposed that each time it lands it either shows heads or tails, then one possible sample space is $\{(H,H),(H,T),(T,H),(T,T)\}$ where for each ordered pair (x,y), $x = H$ means a head was obtained on the first toss, and so on.

In what follows, let S denote a sample space.

Complete set of events:: An event is a subset of the sample space, S. Events are the sets of outcomes corresponding to the answers to yes/no questions. For example, in the coin tossing experiment, if the question is "Did we get at least one head?", the answer "Yes" corresponds to $\{(H,H),(H,T),(T,H)\}$ and the answer "No" to $\{(T,T)\}$. For a set of events to be complete it must have three properties:

1. It must contain the sample space S itself.
2. It must be closed under complements: If F is an event, so must be F^c. (Just rephrase your question to reverse the meaning of "Yes" and "No".
3. It must be closed under countable unions: If $F_1, F_2, \ldots$ is an infinite sequence of events in S, then $F_1 \cup F_2 \cup \cdots$ must be an event as well. This allows for asking an infinite sequence of questions, and getting at least one answer of "Yes".

A complete set of events is called a **sigma algebra**. If S is our example from coin tossing, then

$$\{\{(H,H),(T,T)\},\{(H,T),(T,H)\},\emptyset,S\}$$

is a sigma algebra, as is the set of all subsets of S.

Note that these properties imply that the empty set is an event, as is any finite union of events.

Probability measure:: If we let S be our sample space, and Σ be a sigma algebra of events, we need to assign a probability to each event. That means we must have a function whose domain is the set of events, Σ, and whose range is contained in $[0,1]$. This function is called a **probability measure**, and its values are called **probabilities**. Let us denote this function by Pr. We have $\Pr : \Sigma \to [0,1]$. Pr must have two properties:

1. $\Pr(S) = 1$.
2. If $F_1, F_2, \cdots$ is a sequence of events and for $i \neq j$ we always have $F_i \cap F_j = \emptyset$, then

$$\Pr(F_1 \cup F_2 \cup \cdots) = \sum_{k=1}^{\infty} \Pr(F_k).$$

It follows from these assumptions that $\Pr(\emptyset) = 0$ and for any pair of events A and B with $A \cap B = \emptyset$, $\Pr(A \cup B) = \Pr(A) + \Pr(B)$. In particular, $1 = \Pr(A) + \Pr(A^c)$.

The heart of probability modeling usually lies in choosing Pr. If S has only a finite number of elements, the model is called **classical** if the probability of an event is equal to the number of its elements divided by the number of elements of the sample space.

A major difference between probability and statistics is that in probability you assume you know Pr and try to predict the data before you perform the experiment, while in statistics, you have the data and you try to decide what Pr is.