Dr. Kim's Note (December 15th)

Bayes' Theorem

Let A_1, …, A_k be mutually exclusive and exhaustive events with *prior* probabilities $P(A_i)$ ($i=1, \ldots, k$). Then for any other event B for which $P(B)>0$, the *posterior* probability of A_j given that B has occurred is

$$P(A_j \mid B) = P(A_j \cap B)/P(B)$$

$$= P(B \mid A_j)P(A_j)/\sum_{i=1}^{k} P(B \mid A_i)P(A_i).$$

If A and B are any events whose probabilities are not 0 or 1,

$$P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B \mid A) \times P(A) + P(B \mid A^c) \times P(A^c)}.$$

Random variable

A **random variable** is a function that takes each possible outcome in the sample space and maps it into a numeric value. (Domain: the sample space *S*, Range: the set of real numbers.)

For example define the random variable *X* as the number of heads in 2 tosses of a fair, 50-50 coin. The sample space is $S = \{HT, HH, TH, TT\}$ the corresponding outcomes in this sample space get associated with values of the random variable X as $\{1,2,1,0\}$ because the outcomes have 1, 2, 1, and 0 heads respectively.

The Bernoulli Distribution

There is an experiment that can result in one of only two outcomes (success or failure). If we assume the constant success probability is *p*, then we call the experiment Bernoulli trial with *p*.

Example Let *X* be the number of heads in one toss of a coin. And say "success" event here is to obtain "Heads". Then, *X*~Bernoulli (1/2).

For Bernoulli trial, we define a random variable X as follow: If the experiment results in success, then X=1. Otherwise X=0.

$$p(0)=P(X=0)=1-p \leftarrow \text{failure probability}$$

$$p(1)=P(X=1)=p \leftarrow \text{success probability}$$

The random variable X is said to have the Bernoulli distribution with parameter p.

The notation is X~Bernoulli (p).

Discrete Random variable

A **discrete random variable** X has a finite number of possible values. The probability distribution of X lists the values and their probabilities:

Value of X	Probability
X_1	p_1
X_2	p_2
:	:
X_k	p_k

The probabilities p_i must satisfy two requirements:

1. Every probability p_i is a number between 0 and 1.
2. $p_1+p_2+\cdots+p_k = 1$

We usually summarize all the information about a random variable with a probability table like:

X	0	1	2
$P(x)$	1/4	1/2	1/4

This is the probability table representing the random variable X defined above for the 2 toss coin tossing experiment. There is one outcome with zero heads, 2 with one head, and one with 2 heads.

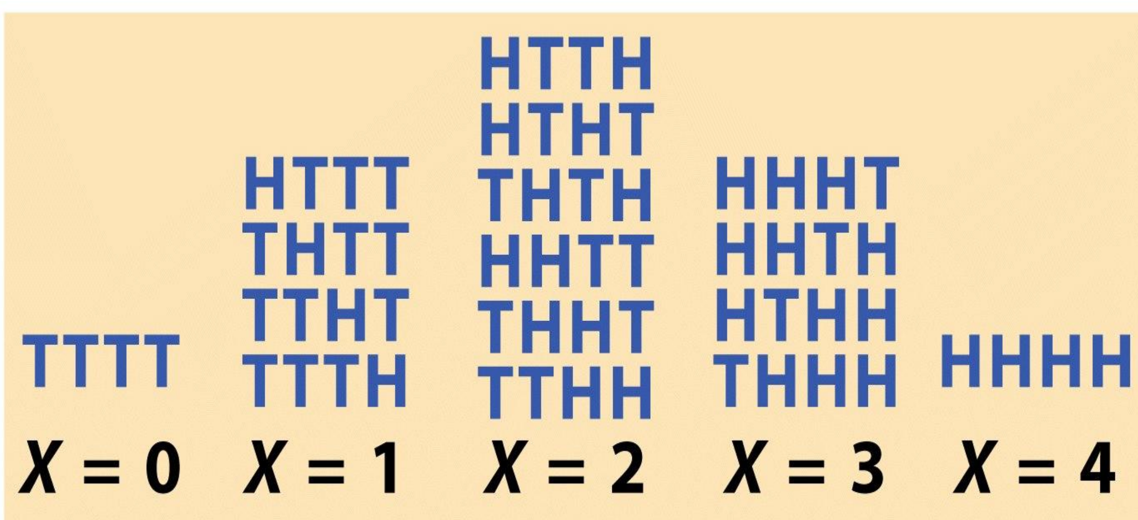

Figure 4-6
Introduction to the Practice of Statistics, Fifth Edition

All outcomes are equally likely, and this means the probabilities are defined as the number of outcomes in the event divided by the total number of outcomes.

We can find the probability of each value of X from Figure. Here is the result:

Value of X	Probability
0	0.0625
1	0.25
2	0.375
3	0.25
4	0.0625

- The probability mass function (**pmf**) of a discrete random variable X is the function $p(x)=P(X=x)$. The cumulative distribution function (**cdf**) of a random variable X is the function $F(x)=P(X\leq x)$.

Continuous Random variable

A **continuous random variable** takes all values in an interval of numbers. The probability distribution of *X* is described by a density curve. The probability of any event is the area under density curve and above the values of *X* that make up the event.

Suppose that we want to choose a number at random between 0 and 1, allowing any number between 0 and 1 as the outcome. Software random number generators will do this. You can visualize such a random number by thinking of a spinner. The sample space *S* is now an entire interval of numbers:

$$S=\{\text{all number } x \text{ such } 0 \le x \le 1\}$$

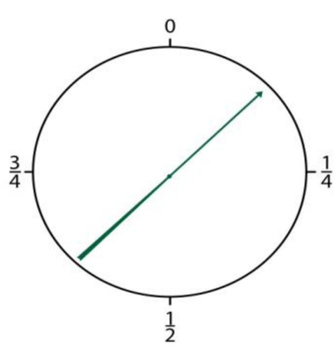

Figure A spinner that generates a random number between 0 and 1.

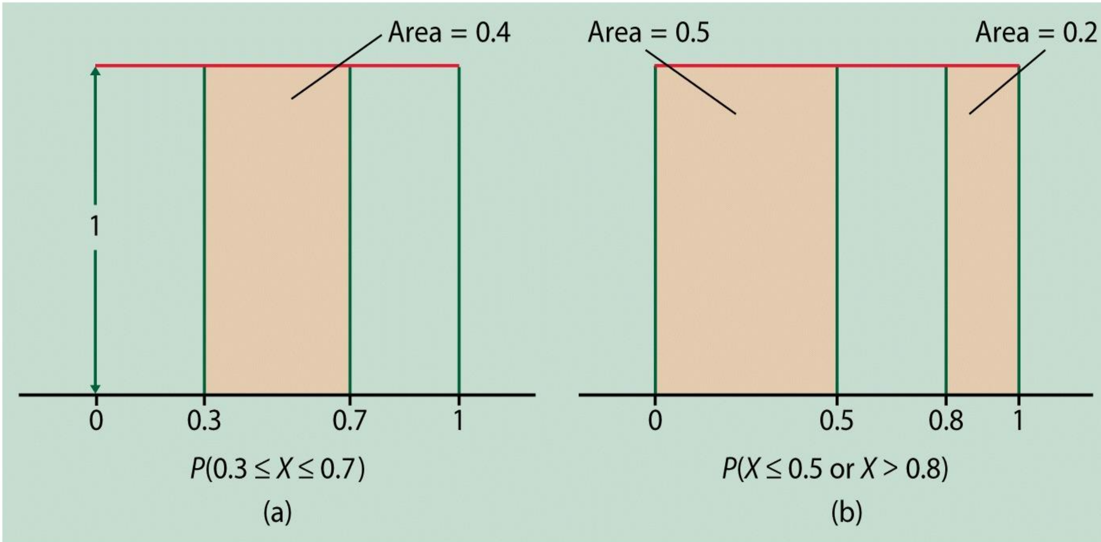

Figure Assigning probabilities for generating a random number between 0 and 1. The probability of any interval of numbers is the area above the interval and under curve.

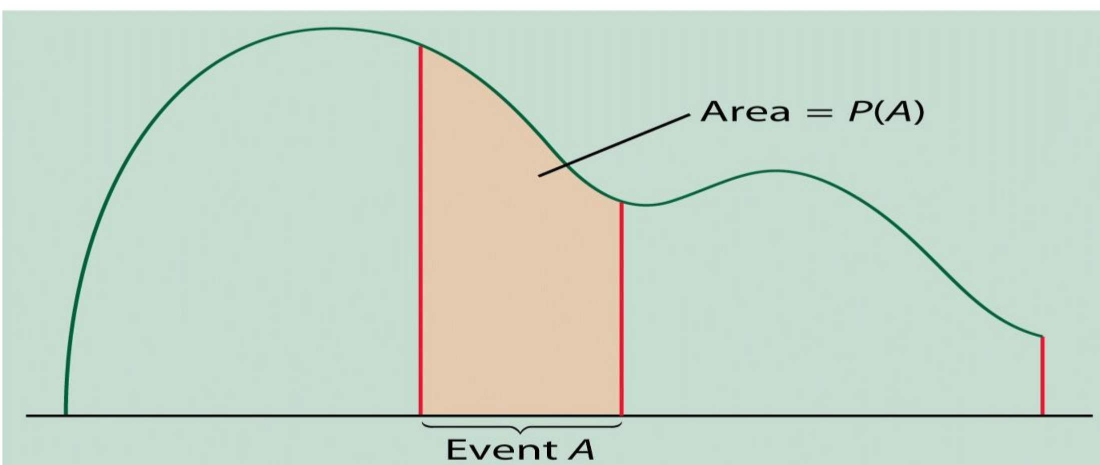

Figure The probability distribution of a continuous random variable assigns probabilities as area under a density curve.