Chapter 10

Introducing Probability
Idea of Probability

◆ Probability is the science of chance behavior
◆ Chance behavior is unpredictable in the short run but has a regular and predictable pattern in the long run
  – this is why we can use probability to gain useful results from random samples and randomized comparative experiments
Randomness and Probability

- **Random**: individual outcomes are uncertain but there is a regular distribution of outcomes in a large number of repetitions

- *Relative frequency* (proportion of occurrences) of an outcome settles down to one value over the long run. That *one value* is then defined to be the *probability* of that outcome.
Relative-Frequency Probabilities

- Can be determined (or checked) by observing a long series of independent trials (empirical data)
  - experience with many samples
  - simulation (computers, random number tables)
Relative-Frequency Probabilities

Coin flipping:

![Coin flipping image]

![Graph showing proportion of heads over number of tosses]
Probability Models

- The **sample space** $S$ of a random phenomenon is the set of all possible outcomes.
- An **event** is an outcome or a set of outcomes (subset of the sample space).
- A **probability model** is a mathematical description of long-run regularity consisting of a sample space $S$ and a way of assigning probabilities to events.
Probability Model for Two Dice

Random phenomenon: roll pair of fair dice.

Sample space:

Probabilities: each individual outcome has probability 1/36 (.0278) of occurring.
Probability Rule 1

Any probability is a number between 0 and 1.

- A probability can be interpreted as the proportion of times that a certain event can be expected to occur.
- If the probability of an event is more than 1, then it will occur more than 100% of the time (Impossible!).
Probability Rule 2

All possible outcomes together must have probability 1.

- Because some outcome must occur on every trial, the sum of the probabilities for all possible outcomes must be exactly one.

- If the sum of all of the probabilities is less than one or greater than one, then the resulting probability model will be incoherent.
Probability Rule 3

If two events have no outcomes in common, they are said to be disjoint. The probability that one or the other of two disjoint events occurs is the sum of their individual probabilities.

- Age of woman at first child birth
  - under 20: 25%
  - 20-24: 33%
  - 25+: ? Rule 3 (or 2): 42%

\[
\{ \text{under 20: } 25\%, \text{ 20-24: } 33\% \} \quad \text{24 or younger: } 58\%
\]
Probability Rule 4

The probability that an event does not occur is 1 minus the probability that the event does occur.

- As a jury member, you assess the probability that the defendant is guilty to be 0.80. Thus you must also believe the probability the defendant is not guilty is 0.20 in order to be coherent (consistent with yourself).
- If the probability that a flight will be on time is 0.70, then the probability it will be late is 0.30.
PROBABILITY RULES

Rule 1. The probability $P(A)$ of any event $A$ satisfies $0 \leq P(A) \leq 1$.

Rule 2. If $S$ is the sample space in a probability model, then $P(S) = 1$.

Rule 3. Two events $A$ and $B$ are disjoint if they have no outcomes in common and so can never occur together. If $A$ and $B$ are disjoint,

$$P(A \text{ or } B) = P(A) + P(B)$$

This is the addition rule for disjoint events.

Rule 4. For any event $A$,

$$P(A \text{ does not occur}) = 1 - P(A)$$
Probability Rules: Mathematical Notation

Random phenomenon: roll pair of fair dice and count the number of pips on the up-faces.

Find the probability of rolling a 5.

\[
P(\text{roll a 5}) = P(\text{1-1}) + P(\text{1-2}) + P(\text{1-3}) + P(\text{2-1})
\]

\[
= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}
\]

\[
= \frac{4}{36}
\]

\[
= 0.111
\]
Discrete Probabilities

- Finite (countable) number of outcomes
  - assign a probability to each individual outcome, where the probabilities are numbers between 0 and 1 and sum to 1
  - the probability of any event is the sum of the probabilities of the outcomes making up the event
  - see previous slide for an example
Continuous Probabilities

- Intervals of outcomes
  - cannot assign a probability to each individual outcome (because there are an infinite number of outcomes)
  - probabilities are assigned to intervals of outcomes by using areas under density curves
  - a density curve has area exactly 1 underneath it, corresponding to total probability 1
Assigning Probabilities: Random Numbers Example

Random number generators give output (digits) spread uniformly across the interval from 0 to 1.

Find the probability of getting a random number that is less than or equal to 0.5 OR greater than 0.8.

\[ P(X \leq 0.5 \text{ or } X > 0.8) = P(X \leq 0.5) + P(X > 0.8) \]
\[ = 0.5 + 0.2 \]
\[ = 0.7 \]
Normal Probability Models

- Often the density curve used to assign probabilities to intervals of outcomes is the Normal curve
  - **Normal distributions are probability models**: probabilities can be assigned to intervals of outcomes using the Standard Normal probabilities in Table A of the text (pp. 690-691)
  - the technique for finding such probabilities is found in Chapter 3
Normal Probability Models

**Example:** convert observed values of the endpoints of the interval of interest to standardized scores (z scores), then find probabilities from Table A.
Random Variables

- A **random variable** is a variable whose value is a numerical outcome of a random phenomenon
  - often denoted with capital alphabetic symbols (X, Y, etc.)
  - a normal random variable may be denoted as $X \sim N(\mu, \sigma)$
- The **probability distribution** of a random variable $X$ tells us what values $X$ can take and how to assign probabilities to those values
Random Variables

- Random variables that have a finite (countable) list of possible outcomes, with probabilities assigned to each of these outcomes, are called **discrete**

- Random variables that can take on any value in an interval, with probabilities given as areas under a density curve, are called **continuous**
Random Variables

- **Discrete random variables**
  - number of pets owned (0, 1, 2, … )
  - numerical day of the month (1, 2, …, 31)
  - how many days of class missed

- **Continuous random variables**
  - weight
  - temperature
  - time it takes to travel to work
Personal Probabilities

- The degree to which a given individual believes the event in question will happen
- Personal belief or judgment
- Used to assign probabilities when it is not feasible to observe outcomes from a long series of trials
  - assigned probabilities must follow established rules of probabilities (between 0 and 1, etc.)
Personal Probabilities

Examples:
- probability that an experimental (never performed) surgery will be successful
- probability that the defendant is guilty in a court case
- probability that you will receive a ‘B’ in this course
- probability that your favorite baseball team will win the World Series in 2020