Chapter 8: Probability: The Mathematics of Chance
Lesson Plan

- Probability Models and Rules
- Discrete Probability Models
- Equally Likely Outcomes
- Continuous Probability Models
- The Mean and Standard Deviation of a Probability Model
- The Central Limit Theorem
Chapter 8: Probability: The Mathematics of Chance
Probability Models and Rules

**Probability Theory**
- The mathematical description of randomness.
- Companies rely on profiting from known probabilities.
  - **Examples**: Casinos know every dollar bet will yield revenue; insurance companies base their premiums on known probabilities.

**Randomness –** A phenomenon is said to be random if individual outcomes are uncertain but the long-term pattern of many individual outcomes is predictable.

**Probability –** For a random phenomenon, the probability of any outcome is the proportion of times the outcome would occur in a very long series of repetitions.
Chapter 8: Probability: The Mathematics of Chance

Probability Models and Rules

- **Probability Model**
  - A mathematical description of a random phenomenon consisting of two parts: a sample space $S$ and a way of assigning probabilities to events.
  - **Sample Space** – The set of all possible outcomes.
  - **Event** – A subset of a sample space (can be an outcome or set of outcomes).

- **Probability Model Rolling Two Dice**
  - Rolling two dice and summing the spots on the up faces.

### Rolling Two Dice: Sample Space and Probabilities

<table>
<thead>
<tr>
<th>Outcome</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$\frac{1}{36}$</td>
<td>$\frac{2}{36}$</td>
<td>$\frac{3}{36}$</td>
<td>$\frac{4}{36}$</td>
<td>$\frac{5}{36}$</td>
<td>$\frac{6}{36}$</td>
<td>$\frac{5}{36}$</td>
<td>$\frac{4}{36}$</td>
<td>$\frac{3}{36}$</td>
<td>$\frac{2}{36}$</td>
<td>$\frac{1}{36}$</td>
</tr>
</tbody>
</table>

The probability of an $8$ is $\frac{5}{36} = 0.14$. 

[Probability histogram chart with bars for each outcome from 2 to 12, probabilities range from 0.00 to 0.25]
Probability Rules

1. The probability \( P(A) \) of any event \( A \) satisfies \( 0 \leq P(A) \leq 1 \).
   - Any probability is a number between 0 and 1.

2. If \( S \) is the sample space in a probability model, the \( P(S) = 1 \).
   - All possible outcomes together must have probability of 1.

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, \( P(A \text{ or } B) = P(A) + P(B) \) (addition rule for disjoint events).
   - If two events have no outcomes in common, the probability that one or the other occurs is the sum of their individual probabilities.

4. The complement of any event \( A \) is the event that \( A \) does not occur, written as \( A^c \). The complement rule: \( P(A^c) = 1 - P(A) \).
   - The probability that an event does not occur is 1 minus the probability that the event does occur.
Chapter 8: Probability: The Mathematics of Chance

Discrete Probability Models

- **Discrete Probability Model**
  - A probability model with a finite sample space is called discrete.
  - To assign probabilities in a discrete model, list the probability of all the individual outcomes.
  - These probabilities must be between 0 and 1, and the sum is 1.
  - The probability of any event is the sum of the probabilities of the outcomes making up the event.

- **Benford’s Law**
  - The first digit of numbers (not including zero, 0) in legitimate records (tax returns, invoices, etc.) often follow this probability model.
  - Investigators can detect fraud by comparing the first digits in business records (i.e., invoices) with these probabilities.

<table>
<thead>
<tr>
<th>First digit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.301</td>
<td>0.176</td>
<td>0.125</td>
<td>0.097</td>
<td>0.079</td>
<td>0.067</td>
<td>0.058</td>
<td>0.051</td>
<td>0.046</td>
</tr>
</tbody>
</table>

**Example:**
Event \( A = \{ \text{first digit is 1} \} \)
\[ P(A) = P(1) = 0.301 \]
Chapter 8: Probability: The Mathematics of Chance

Equally Likely Outcomes

- **Equally Likely Outcomes**
  - If a random phenomenon has \( k \) possible outcomes, all equally likely, then each individual outcome has probability of \( \frac{1}{k} \).
  - The probability of any event \( A \) is:

\[
P(A) = \frac{\text{count of outcomes in } A}{\text{count of outcomes in } S}
\]

\[
= \frac{\text{count of outcomes in } A}{k}
\]

**Example:**
Suppose you think the first digits are distributed “at random” among the digits 1 though 9; then the possible outcomes are equally likely.

<table>
<thead>
<tr>
<th>First digit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
</tr>
</tbody>
</table>

If business records are unlawfully produced by using (1 – 9) random digits, investigators can detect it.
Chapter 8: Probability: The Mathematics of Chance

Equally Likely Outcomes

- Comparing Random Digits (1 – 9) and Benford’s Law
  - Probability histograms of two models for first digits in numerical records (again, not including zero, 0, as a first digit).

  Figure (a) shows equally likely digits (1 – 9).
  Each digit has an equally likely probability to occur $P(1) = \frac{1}{9} = 0.111$.

  Figure (b) shows the digits following Benford’s law.
  In this model, the lower digits have a greater probability of occurring.

  The vertical lines mark the means of the two models.
Chapter 8: Probability: The Mathematics of Chance
Equally Likely Outcomes

- **Combinatorics**
  - The branch of mathematics that counts arrangement of objects when outcomes are equally likely.
  - **Fundamental Principle of Counting** (Multiplication Method of Counting)
    For both rules, we have a collection of $n$ distinct items, and we want to arrange $k$ of these items in order, such that:

**Rule A**
In the arrangement, the same item can appear several times.
The number of possible arrangements: $n \times n \times \ldots \times n = n^k$

**Rule B**
In the arrangement, any item can appear no more than once.
The number of possible arrangements: $n \times (n - 1) \times \ldots \times (n - k + 1)$
Chapter 8: Probability: The Mathematics of Chance

Equally Likely Outcomes

Two Examples of Fundamental Principle of Counting

Rule A  The number of possible arrangements:  \( n \times n \times \ldots \times n = n^k \)

*Same item can appear several times.*

Example: What is the probability a three-letter code has no X in it?

Count the number of three-letter code with no X: \( 25 \times 25 \times 25 = 15,625 \).

Count all possible three-letter codes: \( 26 \times 26 \times 26 = 17,576 \).

\[
P(\text{no X}) = \frac{\text{Number of codes with no X}}{\text{Number of all possible codes}} = \frac{25 \times 25 \times 25}{26 \times 26 \times 26} = \frac{15,625}{17,576} = 0.889
\]

Rule B  The number of possible arrangements:  \( n \times (n - 1) \times \ldots \times (n - k + 1) \)

*Any item can appear no more than once.*

Example: What is the probability a three-letter code has no X and no repeats?

\[
P(\text{no X, no repeats}) = \frac{\text{Number of codes with no X, no repeats}}{\text{Number of all possible codes, no repeats}} = \frac{25 \times 24 \times 23}{26 \times 25 \times 24} = \frac{13,800}{15,600} = 0.885
\]
Chapter 8: Probability: The Mathematics of Chance

Continuous Probability Model

- **Density Curve**
  - A curve that is always on or above the horizontal axis.
  - The curve always has an area of exactly 1 underneath it.

- **Continuous Probability Model**
  - Assigns probabilities as areas under a density curve.
  - The area under the curve and above any range of values is the probability of an outcome in that range.

**Example: Normal Distributions**

- Total area under the curve is 1.
- Using the 68-95-99.7 rule, probabilities (or percents) can be determined.
- Probability of 0.95 that proportion $\hat{p}$ from a single SRS is between 0.58 and 0.62 (adults frustrated with shopping).
Mean of a Discrete Probability Model

Suppose that the possible outcomes $x_1, x_2, \ldots, x_k$ in a sample space $S$ are numbers and that $p_j$ is the probability of outcome $x_j$. The mean $\mu$ of this probability model is:

$$\mu = x_1 p_1 + x_2 p_2 + \ldots + x_k p_k$$

Mean of Random Digits Probability Model


$$= 45 (1/9) = 5$$

Mean of Benford’s Probability Model

$$\mu = (1)(0.301) + (2)(0.176) + (3)(0.125) + (4)(0.097) + (5)(0.079) + (6)(0.067) + (7)(0.058) + (8)(0.051) + (9)(0.046)$$

$$= 3.441$$
Mean of a Continuous Probability Model

- Suppose the area under a density curve was cut out of solid material. The mean is the point at which the shape would balance.

Law of Large Numbers

- As a random phenomenon is repeated a large number of times:
  - The proportion of trials on which each outcome occurs gets closer and closer to the probability of that outcome, and
  - The mean $\bar{x}$ of the observed values gets closer and closer to $\mu$.

(This is true for trials with numerical outcomes and a finite mean $\mu$.)

Chapter 8: Probability: The Mathematics of Chance
The Mean and Standard Deviation of a Probability Model
Chapter 8: Probability: The Mathematics of Chance

The Mean and Standard Deviation of a Probability Model

Standard Deviation of a Discrete Probability Model

- Suppose that the possible outcomes \( x_1, x_2, \ldots, x_k \) in a sample space \( S \) are numbers and that \( p_i \) is the probability of outcome \( x_i \).
- The variance \( \sigma^2 \) of this probability model is:
  \[
  \sigma^2 = (x_1 - \mu)^2 p_1 + (x_2 - \mu)^2 p_2 + \ldots + (x_k - \mu)^2 p_k
  \]
- The standard deviation \( \sigma \) is the square root of the variance.

Example: Find the standard deviation for the data that shows the probability model for Benford’s law.

<table>
<thead>
<tr>
<th>First Digit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.301</td>
<td>0.176</td>
<td>0.125</td>
<td>0.097</td>
<td>0.079</td>
<td>0.067</td>
<td>0.058</td>
<td>0.051</td>
<td>0.046</td>
</tr>
</tbody>
</table>

**Variance** \( \sigma^2 = (x_1 - \mu)^2 p_1 + (x_2 - \mu)^2 p_2 + \ldots + (x_k - \mu)^2 p_k \)

\[
= (1 - 3.441)^2 0.301 + (2 - 3.441)^2 0.176 + (3 - 3.441)^2 0.125 + (4 - 3.441)^2 0.097 + (5 - 3.441)^2 0.079 + (6 - 3.441)^2 0.067 + (7 - 3.441)^2 0.058 + (8 - 3.441)^2 0.051 + (9 - 3.441)^2 0.046 = 6.06
\]

\[
\sigma = \sqrt{\sigma^2} = \sqrt{6.06} = 2.46
\]
One of the most important results of probability theory is the central limit theorem, which says:

- The distribution of any random phenomenon tends to be Normal if we average it over a large number of independent repetitions.
- This theorem allows us to analyze and predict the results of chance phenomena when we average over many observations.

The Central Limit Theorem

- Draw a simple random sample (SRS) of size $n$ from any large population with mean $\mu$ and a finite standard deviation $\sigma$.

Then,

- The mean of the sampling distribution of $\bar{x}$ is $\mu$.
- The standard deviation of the sampling distribution of $\bar{x}$ is $\sigma/\sqrt{n}$.
- The central limit theorem says that the sampling distribution of $\bar{x}$ is approximately normal when the sample size $n$ is large.