The science of uncertainty is called **probability theory**.

### 4.1 Probability Basics

**Experiment**

- Is a process that results in one of a number of possible outcomes.
- The outcome that occurs cannot be predicted with certainty.

**Example**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flip a coin</td>
<td></td>
</tr>
<tr>
<td>Roll a die</td>
<td></td>
</tr>
<tr>
<td>Dow Jones</td>
<td></td>
</tr>
</tbody>
</table>
Actual outcome can not be determined (predicted) in advance;

but

can

4.2 Events

Sample space

List of all possible outcomes for an experiment.

i.e. Coin: \( S = \)

Event

A collection of outcomes for the experiment

i.e. Dow Jones doesn't increase.
Definition - Probability

Number of ways event can occur

Probability of an event =

Total number of possible outcomes

Meaning of Probability

Probability is

- a generalization of the concept of percentage.
- a numerical measure of how likely it is that an event will occur.

0 0.5 1
Venn Diagrams

Useful for illustrating probability concepts.

- Sample space - depicted as a rectangle
- Events - disks inside the rectangle

Venn diagram for event A
Union

The **union** of any two events $A$ and $B$, denoted $A \cup B$ (A or B), occurs if

**Example** - Tossing a die.

Even numbers: $C =$

Less than 3: $D =$

$C \cup D =$
Intersection

The intersection of any two events A and B, denoted $A \cap B$ (A&B) occurs

Example

$C \cap D =$
Mutually Exclusive

Two events $A$ and $B$ are mutually exclusive if they do not have outcomes in common.

\[ A \cap B = \]

Example

Even numbers: $C = \quad \quad$ Odd numbers: $F = \quad \quad$

\[ C \cap F = \]

Hence, $C$ & $F$ are
Complement

The **complement** of any event $A$, denoted $\bar{A}$, is the set of all

![Diagram of complement](https://via.placeholder.com/150)

**Example** - $F = \{1, 3, 5\}$

$\bar{F} =$

For any event $A$,

1. $A \cap \bar{A} =$

2. $A \cup \bar{A} =$

   (Exhaustive).
4.3 Some Rules of Probability

1. Probabilities between 0 and 1

\[ 0 \leq P(A) \leq 1 \]

2. Probability of sample space

That is:

\[ P(S) = \quad \text{and} \quad P(\emptyset) = P(\{\}) = . \]

3. **Addition Rule**

(Mutually Exclusive Events):

When A and B are mutually exclusive

\[ P(A \cup B) = \]

**Ex.** Die: A={1} and B={2,4}

\[ P(A \cup B) = \]

\[ = \]

\[ = 3/6 = 0.5 \]

Note: can be extended to more than 2 events.
4. **Complement Rule**

\[ \bar{A} : \text{Event that } A \text{ does not occur, then} \]

Know that: \[ A \cup \bar{A} = \]

then \[ P(A \cup \bar{A}) = \]

and \[ \]

Hence: \[ P(A) = \]

or \[ P(\bar{A}) = \]

**Ex.** Die example: Assume that \( P(\bar{A}) = 2/6 \), what is the \( P(A) \)?

\[ P(A) = \]
5. **Addition Rule (General)** - events A and B do not have to be mutually exclusive events.

\[
P(A \cup B) =
\]

**Ex.**  
A: even number tossed  
B: number < 3 tossed

\[
P(A \cup B) = P(A) + P(B) - P(A \cap B)
\]
Ex. As reported in *Employment and Earnings*, the age dist’n of employed persons 16 years old and over is

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency (000's)</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-19</td>
<td>6,500</td>
<td>A</td>
</tr>
<tr>
<td>20-24</td>
<td>12,138</td>
<td>B</td>
</tr>
<tr>
<td>25-34</td>
<td>32,077</td>
<td>C</td>
</tr>
<tr>
<td>35-44</td>
<td>35,051</td>
<td>D</td>
</tr>
<tr>
<td>45-54</td>
<td>25,514</td>
<td>E</td>
</tr>
<tr>
<td>55-64</td>
<td>11,739</td>
<td>F</td>
</tr>
<tr>
<td>65 &amp; over</td>
<td>3,690</td>
<td>G</td>
</tr>
<tr>
<td>Total</td>
<td>126,709</td>
<td></td>
</tr>
</tbody>
</table>

An employed person is selected at random. Let the following events be defined

\[
W = \text{the person is between 20 and 64}
\]
\[
Y = \text{the person is under 65}
\]
\[
Z = \text{the person is 55 or over.}
\]

Describe each of the following events in words and determine their probabilities.
a. (not Y)

Is the event that the person selected is at least 65 years old.

\[ P \text{ (not } Y) = \]

\[ = \]

\[ = \]

\[ = = 0.0291 \]

\[ P \text{ (} Y) = ? \]

Could use the

\[ P \text{ (} Y) = \]

\[ = \]
b. (not W)

Is the event that the person selected is

\[ P(\bar{W}) = \]

\[ = \]

Since \[ A \cap G = \]

\[ = \]

\[ = 0.0804 \]
c. \( Y \cap Z \) (same as \( Y_1 Z \))

Is the event that the person selected is

\[ P (Y \cap Z) = \]

\[ = \]

\[ = 0.0926 \]
4.4 Joint and Marginal Probabilities

Data obtained by observing values of two variables -on same unit- of a population are called

Frequency distributions of bivariate data is called a

Ex. Following is a contingency table giving the number of institutions of higher education in the US by region and type.
<table>
<thead>
<tr>
<th>REGION</th>
<th>TYPE</th>
<th>Public $T_1$</th>
<th>Private $T_2$</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>$R_1$</td>
<td>266</td>
<td>555</td>
<td>821</td>
</tr>
<tr>
<td>Midwest</td>
<td>$R_2$</td>
<td>359</td>
<td>504</td>
<td>863</td>
</tr>
<tr>
<td>South</td>
<td>$R_3$</td>
<td>533</td>
<td>502</td>
<td>1035</td>
</tr>
<tr>
<td>West</td>
<td>$R_4$</td>
<td>313</td>
<td>242</td>
<td>555</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>1471</td>
<td>1803</td>
<td>3274</td>
</tr>
</tbody>
</table>
a. How many cells does this contingency table have?

b. What is the total number of institutions of higher education in the US?

c. How many institutions are in the Midwest?

d. How many are public?

e. How many are private schools in the South?
$\mathbf{R}_i$ : Events associated with Region
$i = 1, 2, 3 \& 4$

e.i. $R_2$ : event the higher education institution is in the Midwest

\[ P(R_2) = ? \]
\[ = 0.2636 \]

$\mathbf{T}_i$ : Events associated with Type
$i = 1 \& 2$

e.i. $T_1$ : event the higher education institution is public

\[ P(T_1) = ? \]
\[ = 0.4493 \]
Could also consider events jointly.

e.i. Event the higher education institution is public in the Midwest.

\[
P(\quad) = 0.1097
\]

Note:

\[P(R_2)\text{ and } P(T_1)\] are called

\[P(T_1 \cap R_2)\] is called a
4.5 Conditional Probability

The conditional probability $P(A \mid B)$ means the probability that $A$ will occur given that $B$ has occurred.

**Ex.** Select randomly from standard pack of 52 cards (no replacement). Let

- $A_1$ : 1st card is an ace
- $A_2$ : 2nd card is an ace

$P(A_1) = \, ?$

$P(A_2 \mid A_1) = \, ?$

Not always easy to calculate them and need to use **General Formula:**

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$
Ex. A joint frequency dist’n for the number of injuries in the US by circumstance and sex is as shown in the following contingency table. Frequencies are in millions.

<table>
<thead>
<tr>
<th>Circumst.</th>
<th>Work $C_1$</th>
<th>Home $C_2$</th>
<th>Other $C_3$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male $S_1$</td>
<td>8.0</td>
<td></td>
<td>17.8</td>
<td>35.6</td>
</tr>
<tr>
<td>Female $S_2$</td>
<td></td>
<td>11.6</td>
<td>12.9</td>
<td>25.8</td>
</tr>
<tr>
<td>TOTAL</td>
<td>9.3</td>
<td>21.4</td>
<td>30.7</td>
<td>61.4</td>
</tr>
</tbody>
</table>
a. Fill in the two empty cells

Missing frequencies in cells

b. How many cells does the contingency table have?

c. Find the probability that an injured person was hurt at work.

\[ P(\quad) = \quad \]

\[ P(\quad) = 0.1515 \]
d. Find the probability that the injured person is female.

\[ P(\text{female}) = ? \]

\[ P(\text{female}) = 0.4202 \]

e. Find the probability that the injured person is female and was hurt at work.

\[ P(\text{female and hurt at work}) = ? \]

\[ P(\text{female and hurt at work}) = 0.0212 \]
f. Given that an individual was hurt at work, what is the probability that it is a female. Obtain this probability directly from the table.

\[ P(\quad \quad ) = ? \]

\[ P(\quad \quad ) = \quad = 0.1398 \]

g. Obtain \( P(S_2|C_1) \) using the conditional probability rule and your answers from part (c) and (e).

\[ P(S_2|C_1) = \]

\[ = \frac{1.3}{9.3} = 0.1398 \]
4.6 The Multiplication Rule; Independence

To find the probability of a joint event i.e. \( P(A \cap B) \), for any two events \( A \) and \( B \), rearrange the conditional probability rule:

\[
P(A \mid B) = \frac{P(A \cap B)}{P(B)}.
\]

Multiplying both sides by \( P(B) \) gives

\[
P(A \cap B) = P(A \mid B) \cdot P(B).
\]

Alternatively:

\[
P(B \mid A) = \frac{P(A \cap B)}{P(A)}.
\]

\[\rightarrow P(A \cap B) = \]
Independent Events

Two events A and B are said to be independent if

\[ P(A \mid B) = P(A) \quad \text{or} \quad P(B \mid A) = P(B). \]

Otherwise the events are dependent.

Multiplication Rule for Independent Events

With independent events \( P(A \mid B) = P(A) \)

hence, the multiplication rule then reduces to

\[ P(A \cap B) = \quad \text{.} \]
h. Are $S_2$ and $C_1$ independent? Explain.

\[ P(S_2 | C_1) = \]

\[ = \]

\[ 0.1398 \times 0.4202 \]

Hence, $C_1$ and $S_2$ are

i. Obtain $P(S_2 \cap C_1)$ using the multiplication rule and your answers from parts (c) and (f).

\[ P(S_2 \cap C_1) = \]

\[ = x = 0.0212 \]