We have begun to study some basic ideas of probability:

- The **probability** of any outcome in a random process is the proportion of times that the outcome would occur in a very large number of repetitions (**relative frequency**).
- Probabilities can be approximated through **simulation**, artificially re-creating the random process a large number of times and determining the proportion of times that the outcome occurs.
- In the case of **equally likely** outcomes, theoretical probabilities can be calculated by enumerating the **sample space**, counting how many outcomes comprise the event of interest, and dividing by the total number of outcomes in the sample space.

Here are some basic probability rules:

- For any event E, \(0 < P(E) < 1\)
- \(P(S) = 1\), where S denotes the sample space
- For two events E and F that are disjoint (i.e., have no outcomes in common), \(P(E \text{ or } F) = P(E) + P(F)\) **addition rule for disjoint (mutually exclusive) events**
- For any event E, \(P(\text{not } E) = 1 - P(E)\) **complement rule**

**Example 9-1: Top 100 Films**

In 1998 the American Film Institute created a list of the top 100 American films ever made (www.afi.com/tvevents/100years/movies.aspx). The following 2×2 table classifies each movie according to whether it was seen by Allan and whether it was seen by Beth.

<table>
<thead>
<tr>
<th>Allan yes</th>
<th>Beth yes</th>
<th>Beth no</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allan yes</td>
<td>42</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Allan no</td>
<td>17</td>
<td>35</td>
<td>100</td>
</tr>
</tbody>
</table>

Suppose that one of these films is chosen “at random,” which means that each of the 100 films is equally likely to be chosen.

a) Determine the probability that Allan has seen the film. *(Hint: Count how many films he has seen and divide by 100.)*

b) Determine the probability that Beth has seen the film.

c) Determine the probability that Allan has not seen the film. Describe two different ways to find this probability.
d) Determine the probability that Beth has not seen the film. Describe two different ways to find this probability.

e) Determine the probability that both Allan and Beth have seen the film.

f) Determine the probability that neither Allan nor Beth has seen the film.

g) Add the counts in the appropriate cells of the table to determine the probability that either Allan or Beth (or both) have seen the film.

h) If you had not been given the table but instead had merely been told that P(A) = .48 and P(B) = .59, would you have been able to calculate P(A or B)? Explain.

i) One might naively think that P(A or B) = P(A) + P(B). Calculate this sum, and indicate whether it is larger or smaller than P(A or B) and by how much. Explain why this makes sense, and indicate how to adjust the right side of this expression to make the equality valid.

- **General addition rule**: For any two events E and F, P(E or F) = P(E) + P(F) – P(E and F).

j) Use this addition rule as a second way to calculate the probability that Allan or Beth has seen the movie, verifying your answer to (g).

k) As a third way to calculate this probability, first identify the complement (opposite) of the event \{Allan or Beth has seen the movie\}. Then find the probability of this complement from the table. Then use the complement rule to determine P(A or B). Are your answers to (g) and (j) confirmed?
Example 9-2: Graduate School Applications
Suppose that you have applied to two graduate schools and believe that you have a .6 probability of being accepted by school C, a .7 probability of being accepted by school D, and a .5 probability of being accepted by both.

a) Organize this information into a probability table. Also fill in the rest of the table.

b) Determine the probability of being accepted by at least one of these two schools, first using the table and then using the general addition rule and then using the complement rule.

Table:

General addition rule:

Complement rule:

c) Determine the probability of being rejected by both schools.

d) Determine the probability of acceptance by one school but not both schools.

e) Represent these probabilities in a Venn diagram.

For questions f) and g), suppose you know only that Pr(C) = .6 and Pr(D) = .7.

f) What is the smallest possible value for Pr(C and D)? What is the largest possible value for Pr(C and D)? Under what circumstances do these values occur?

Smallest: ___________________  Largest: ___________________

g) What is the smallest possible value for Pr(C or D)? What is the largest possible value for Pr(C or D)? Under what circumstances do these values occur?

Smallest: ___________________  Largest: ___________________