Topics in Probability

MAT 152 – STATISTICAL METHODS I
LECTURE 2
INSTRUCTOR: DUSTIN ROTEN
FALL 2020

Tools for Tackling Probability

There are several graphical tools that can be used to better understand probabilities. These are:

1. Contingency Tables

   A way of portraying data that can facilitate calculating probabilities.

2. Tree Diagrams

   A special type of graph used to determine the outcomes of an experiment. It consists of "branches" that are labeled with either frequencies or probabilities. Trees are useful for conditional probabilities.

3. Venn Diagrams

   A picture that represents the outcomes of an experiment. It generally consists of a box that represents the sample space with circles or ovals.

Example (Contingency Tables)

Consider the contingency table below. It describes the distribution of a random sample of 100 individuals organized by gender and whether they are right- or left-handed.

M = the subject is male

F = the subject is female

R = the subject is right-handed

L = the subject is left-handed

	Right-Handed	Left-Handed	Total
Males	43	9	52
Females	44	4	48
Total	87	13	100

Determine *P(M)*.

P(M) is the probability of an individual being a male. This is done by taking the total number of males (left- and right-handed) and dividing by the number of individuals in the sample.

$$P(M) = \frac{52}{100} = 0.52$$

Determine *P(L)*.

P(L) is the probability of an individual being left-handed. This is done by taking the total number of left-handed individuals and dividing by the number of individuals in the sample.

$$P(L) = \frac{13}{100} = 0.13$$

	Right-Handed	Left-Handed	Total
Males	43	9	52
Females	44	4	48
Total	87	13	100

Example (cont.)

Determine $P(F \cap L)$.

Here, we need to find individuals that are both female AND left-handed.

$$P(F \cap L) = \frac{4}{100} = 0.04$$

Determine $P(F \cup L)$.

Here we must consider the union of all female individuals and left-handed individuals. This will be the 48 women (right- and left-handed) plus the 9 left-handed men.

$$P(F \cup L) = \frac{48 + 9}{100} = \frac{57}{100} = 0.57$$

	Right-Handed	Left-Handed	Total
Males	43	9	52
Females	44	4	48
Total	87	13	100

Example (Tree Diagram)

Consider a scenario where 10 cards were placed in a well-shuffled deck. 5 are red and 5 are blue. Two cards are drawn sequentially.

Scenario #1:
The first card is drawn and replaced before the second card is drawn. This is called sampling **with replacement.**

Scenario #2:
The first card is drawn. Then, the second card is drawn. This is called sampling **without replacement.**

How do the probabilities change in these scenarios?

Scenario #1
Sampling with Replacement

There are two options for the first draw: Red or Blue.

$P(R) = \frac{\text{Red Cards}}{\text{All Cards}} = \frac{5}{10} = 0.5$ and $P(B) = \frac{\text{Blue Cards}}{\text{All Cards}} = \frac{5}{10} = 0.5$

Remember, the first card is placed back into the deck so there are still 10 cards. There are four outcomes for the second draw: (Red, Red), (Red, Blue), (Blue, Red), (Blue, Blue).

Since we are replacing the first card, we are “resetting” the probabilities – these events are **independent**.

$P(RR) = P(R_1)P(R_2) = \left(\frac{5}{10}\right)\left(\frac{5}{10}\right) = 0.25$

$P(RB) = P(R_1)P(B_2) = \left(\frac{5}{10}\right)\left(\frac{5}{10}\right) = 0.25$

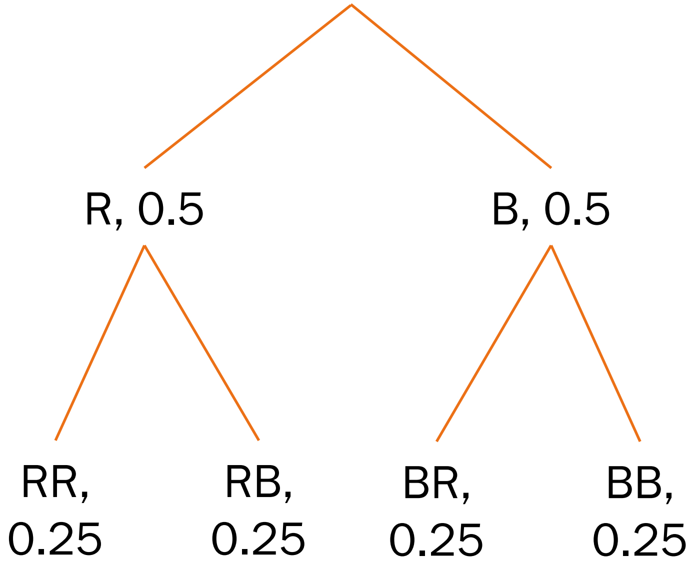
$P(BR) = P(B_1)P(R_2) = \left(\frac{5}{10}\right)\left(\frac{5}{10}\right) = 0.25$

$P(BB) = P(B_1)P(B_2) = \left(\frac{5}{10}\right)\left(\frac{5}{10}\right) = 0.25$

Notice that the probabilities add up to 1.

In this scenario, each outcome is equally likely.

Scenario #1
Sampling with Replacement



Draw #1 – Two options, equal probability

Draw #2 – four options, equal probability

For each step in the tree, the probability is “updated”

$$P(RR \text{ or } BR) = P(RR) + P(BR) = 0.25 + 0.25 = 0.5$$

Scenario #2
Sampling without Replacement

For the first draw, the initial probability is still the same:

$$P(R) = \frac{\text{Red Cards}}{\text{All Cards}} = \frac{5}{10} = 0.5 \text{ and } P(B) = \frac{\text{Blue Cards}}{\text{All Cards}} = \frac{5}{10} = 0.5$$

Now, the first card is NOT replaced, so there are 9 cards in the deck.
There are four still outcomes for the second draw: (Red, Red), (Red, Blue), (Blue, Red), (Blue, Blue).

The outcomes are now **conditional** because the results of the first draw will affect the results of the second draw!

$$P(RR) = P(R_1)P(R_2) = \left(\frac{5}{10}\right)\left(\frac{4}{9}\right) \approx 0.22$$

$$P(RB) = P(R_1)P(B_2) = \left(\frac{5}{10}\right)\left(\frac{5}{9}\right) \approx 0.28$$

$$P(BR) = P(B_1)P(R_2) = \left(\frac{5}{10}\right)\left(\frac{5}{9}\right) \approx 0.28$$

$$P(BB) = P(B_1)P(B_2) = \left(\frac{5}{10}\right)\left(\frac{4}{9}\right) \approx 0.22$$

On the second draw:

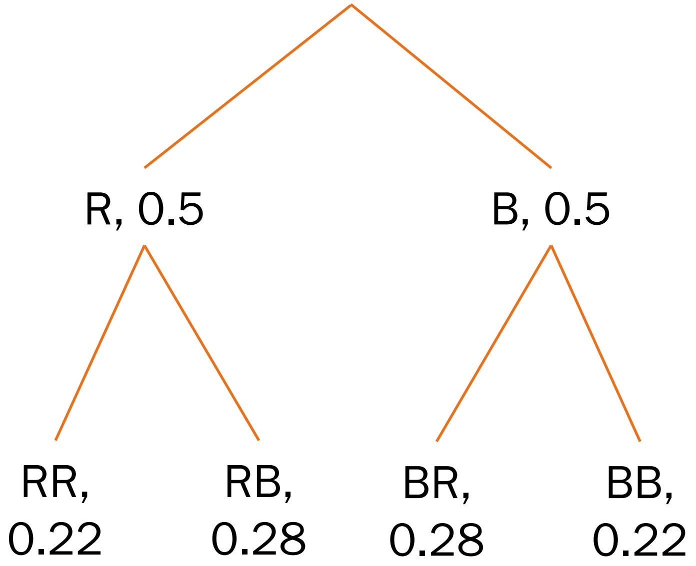
4 red cards remain with 9 cards in the deck

5 blue cards remain with 9 cards in the deck

5 red cards remain with 9 cards in the deck

4 blue cards remain with 9 cards in the deck

Scenario #2
Sampling with Replacement



Draw #1 – Two options, equal probability

For each step in the tree, the probability is "updated"

Draw #2 – four options, unequal probability

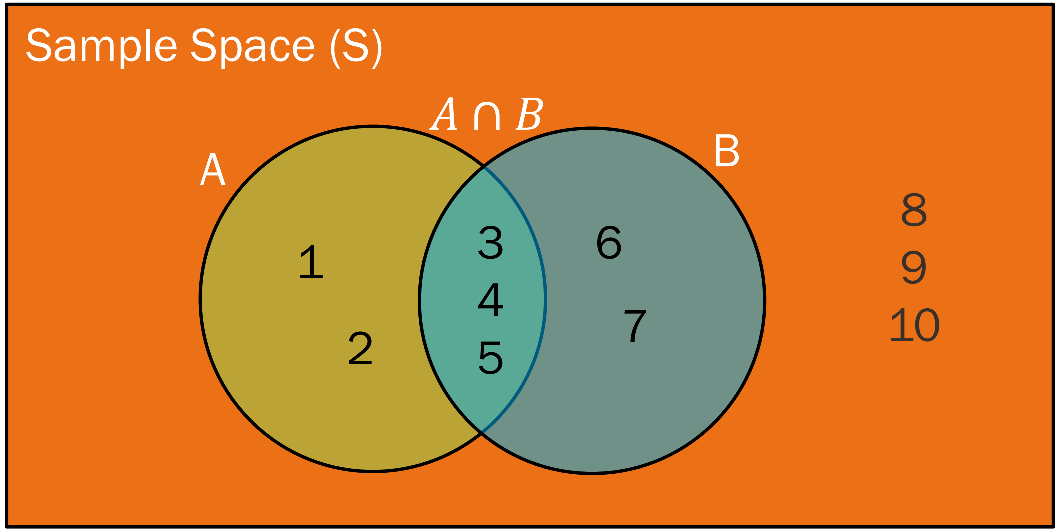
$$P(RR \text{ or } BB) = P(RR) + P(BB) = 0.22 + 0.22 = 0.44$$

Example (Venn Diagram)

Consider the following sample space:
$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

$A = \{1, 2, 3, 4, 5\}$ and $B = \{3, 4, 5, 6, 7\}$

These can be represented with a Venn Diagram



$$P(A) = \frac{\{1, 2, 3, 4, 5\}}{\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}} = \frac{5}{10} = 0.5$$

(Everything in circle "A")

$$P(B) = \frac{\{3, 4, 5, 6, 7\}}{\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}} = \frac{5}{10} = 0.5$$

(Everything in circle "B")

$$P(A \cap B) = \frac{\{3, 4, 5\}}{\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}} = \frac{3}{10} = 0.3$$

(Only the values in both circles)

$$P(A \cup B) = \frac{\{1, 2, 3, 4, 5, 6, 7\}}{\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}} = \frac{7}{10} = 0.7$$

(The values in both circles)