# SDS 321

## Lectures 2 & 3

# Conditional Probability
# and
# Statistical Independence

# Recall

Last week, we introduced the building blocks of probability:

- Experiment: The underlying process whose outcome we are interested in.
- Sample space: The set of all possible outcomes of our experiment. Outcomes must be distinct, mutually exclusive and collectively exhaustive.
- Event: A subset of our sample space.
- Probability of an event: A number assigned to that event that quantifies how likely it is.
- Probability law: A function that assigns probabilities to events.

# Recall

We looked at Axioms of Probability:

- Axioms of probability:
  - **Nonnegativity**: $P(A) \geq 0$, for every event $A$.
  - **Additivity**: If $A$ and $B$ are two disjoint events, then the probability of their union satisfies $P(A \cup B) = P(A) + P(B)$.
    This extends to the union of infinitely many disjoint events:

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$$

  - **Normalization**: The probability of the entire sample space $\Omega$ is equal to 1, i.e. $P(\Omega) = 1$

# Partial (prior) Information

- So far, we have assumed we know nothing about the outcome of our experiment, except for the information encoded in the probability law.

- Sometimes, however, we have partial information that may affect the likelihood of a given event.

  - ➢ The experiment involves rolling a die. You are told that the number is odd.
  - ➢ The experiment involves the weather tomorrow. You know that the weather today is rainy.
  - ➢ The experiment involves the presence or absence of a disease. A blood test comes back positive.

- In each case, knowing about some event B (e.g. “it is raining today”) changes our beliefs about event A (“Will it rain tomorrow?”).

- We want to update our probability law to incorporate this new knowledge.

# Conditional Probability

Original problem:

•What is the probability of some event A?

- Ex: Let A = the event that we roll a number less than 4. What is P(A)?

New problem:

•Assuming info about event B (i.e., “given” event B), what is the probability of event A?

- Ex. Let B = the event that the number is even.
- We want to know what is P(A | B)?

P(A | B) is read as P(A “given” B)

- This is still the probability of event A, except, we now know something about event B.

# Conditional Probability

Example 1:

•Experiment: Let's toss two coins simultaneously.

- What is the Sample Space?

  {HH, HT, TH, TT}

- Let Event A = We get at least one Tail

  3/4

- What is P(A)?

  3/4

- Let Event B = We get the same outcome on the two tosses

  2/4

- What is P(A | B)?

  {TT, HH} = 1/2

- What is P(B | A)?

  {TH, HT, TT} = 1/3

# Conditional Probability

Example 2:

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

Let's first define our events:

A: event that the tosses yield alternating Tails and Heads
B: event that the first toss is a Head

We want P(A|B)

# Conditional Probability

Example 2:

- Sample space for three coin tosses is {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- B: event that the first toss is a Head
- We are conditioning on event B, so we know event B has occurred.
- Our new sample space conditional on B having occurred is: {HHT , HTH , HTT , HHH }
- A : event that the tosses yield alternating Tails and Heads
- Now probability of A given B has occurred, is P(A|B):
- P(A|B) = ¼

So $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

# Conditional Probability Axioms

We can show that the conditional probability $P(A \mid B)$ forms a legitimate probability law that satisfies the three axioms of probability, for a fixed event B.

(1)Non-negativity: $P(A \mid B) \geq 0$ for every A.

(2) Normalization:

Since we are conditioning on B, we can think of the sample space as being confined to the new universe defined by B. So $P(B \mid B) = 1$.

(or) $P(\Omega \mid B) = P(\Omega \cap B) / P(B) = P(B)/P(B) = 1.$

(3) Additivity:

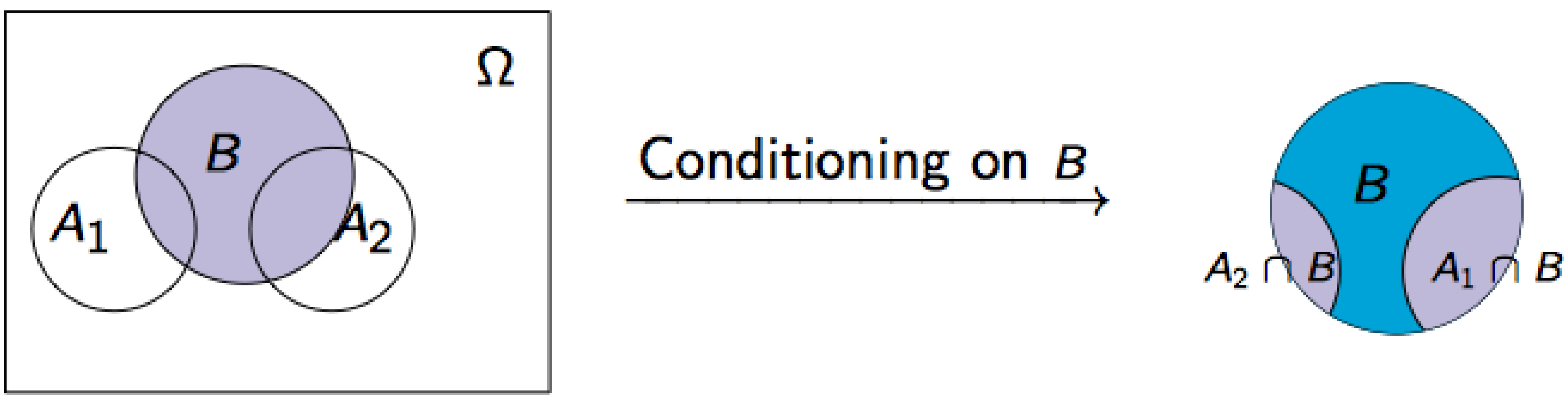
$P(A1 \cup A2 \mid B) = P(A1|B) + P(A2|B)$ for two disjoint sets, A1 and A2.

Proof?

# Conditional Probability Axioms

(3) Additivity:

$P(A1 \cup A2 \mid B) = P(A1|B) + P(A2|B)$ for two disjoint sets, A1 and A2.




Using additivity, $P((A_1 \cup A_2) \cap B) = P(A_1 \cap B) + P(A_2 \cap B)$, so

$$P(A_1 \cup A_2|B) = \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} = P(A_1|B) + P(A_2|B)$$

# Properties of Conditional Probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then $P(\cup_{i=1}^{n} A_i | B) = \sum_{i=1}^{n} P(A_i | B)$.
- If $A_1 \subseteq A_2$, then $P(A_1|B) \leq P(A_2|B)$.
- $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2 | B)$
- Union bound: $P(A_1 \cup A_2 | B) \leq P(A_1|B) + P(A_2|B)$
  - $P(\cup_{i=1}^{n} A_i | B) \leq \sum_{i=1}^{n} P(A_i | B)$.

# Example: Conditional Probability & Probability Trees

Example:

Anita loves to run, and she especially loves to go running in the rain.

- If it is raining, the probability that she goes for a run is 0.8
- If it isn’t raining, the probability that she goes for a run is 0.25
- The probability that it rains tomorrow is 0.1

What is the probability that it doesn’t rain tomorrow, and she goes for a run?

Let’s define the events first:

R = event that it rains tomorrow
$R^c$ = the event that it doesn’t rain tomorrow
W = event that Anita goes for a run
$W^c$ = event that Anita doesn’t go for a run

# Example: Conditional Probability

We defined these events:

R = event that it rains tomorrow
$R^c$ = the event that it doesn't rain tomorrow
W = event that Anita goes for a run
$W^c$ = event that Anita doesn't go for a run

What do we know?

- If it is raining, the probability that she goes for a run is 0.8.

  $P(W \mid R) = 0.8$ $\qquad$ $P(W^c \mid R) = ?$

- If it isn't raining, the probability that she goes for a run is 0.25.

  $P(W \mid R^c) = 0.25$ $\qquad$ $P(W^c \mid R^c) = ?$

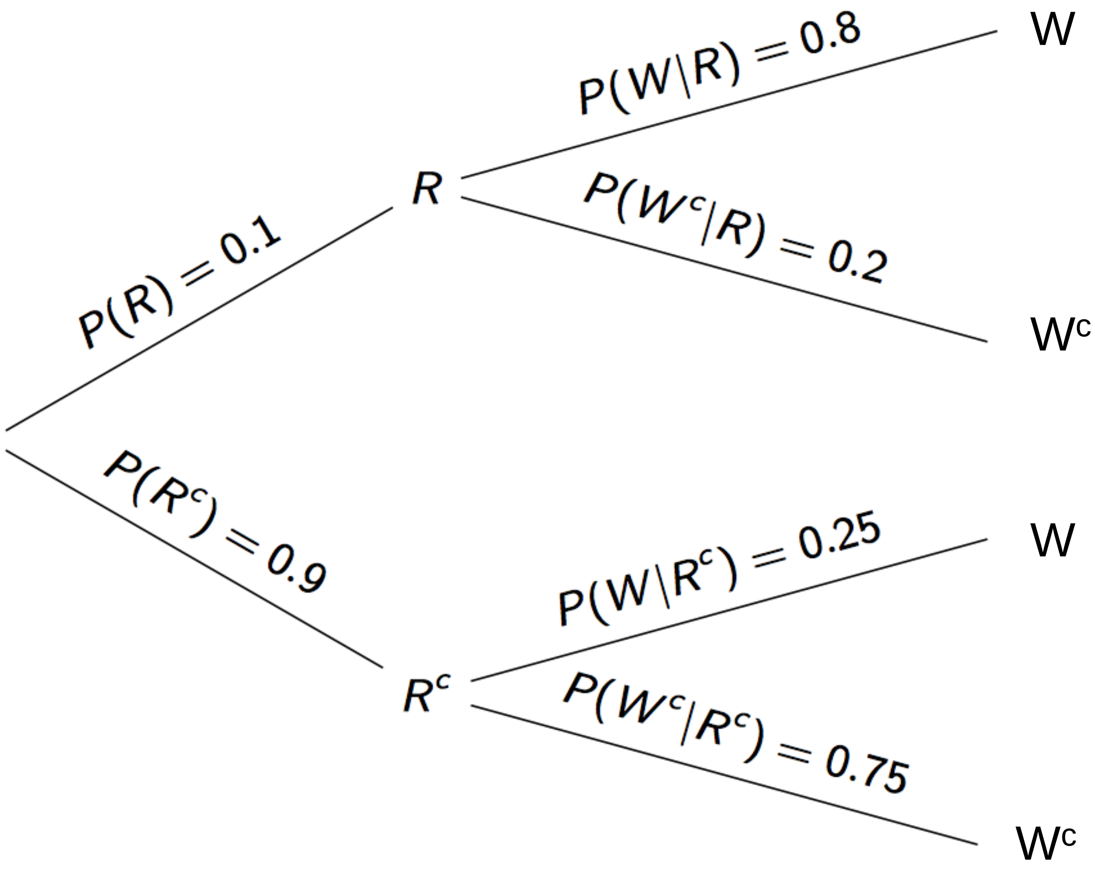
- The probability that it rains tomorrow is 0.1.

  $P(R) = 0.1$ $\qquad$ $P(R^c) = ?$

# Probability Tree

We can organize all this information using a tree structure:

- The first set of branches are simple probabilities
- The second set of branches are conditional probabilities



# Points of Caution

- Here are some common mistakes when working with conditional probabilities. Use them carefully!

- Do not confuse between $P(A|B)$, $P(B|A)$, and $P(A \cap B)$.
  They are all different!

- If you know $P(A|B)$, you automatically know $P(A^c|B)$.
  $P(A^c|B) = 1 - P(A|B)$

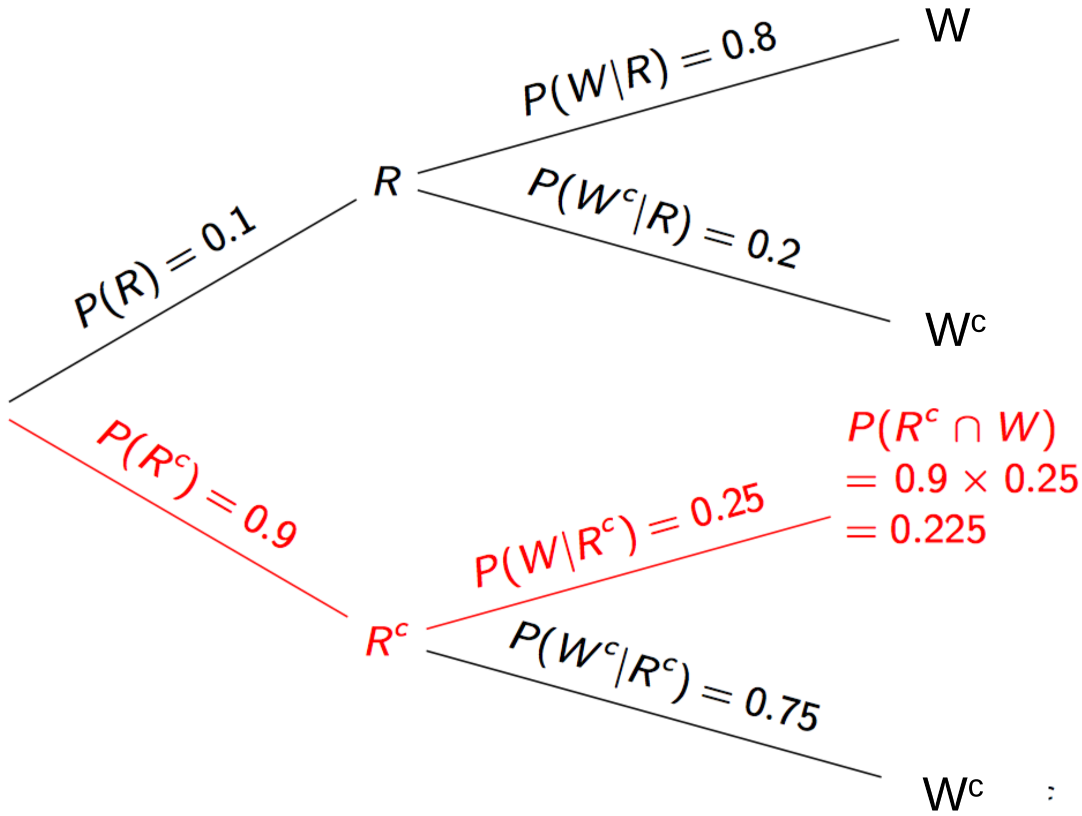
- If you know $P(A|B)$, it doesn't imply you know $P(A|B^c)$.
  Careful! : $P(A|B^c) \neq 1 - P(A|B)$.

- Use these correctly:
  $P(A^c|B) = 1 - P(A|B)$
  $P(A^c|B^c) = 1 - P(A|B^c)$

# Probability Tree

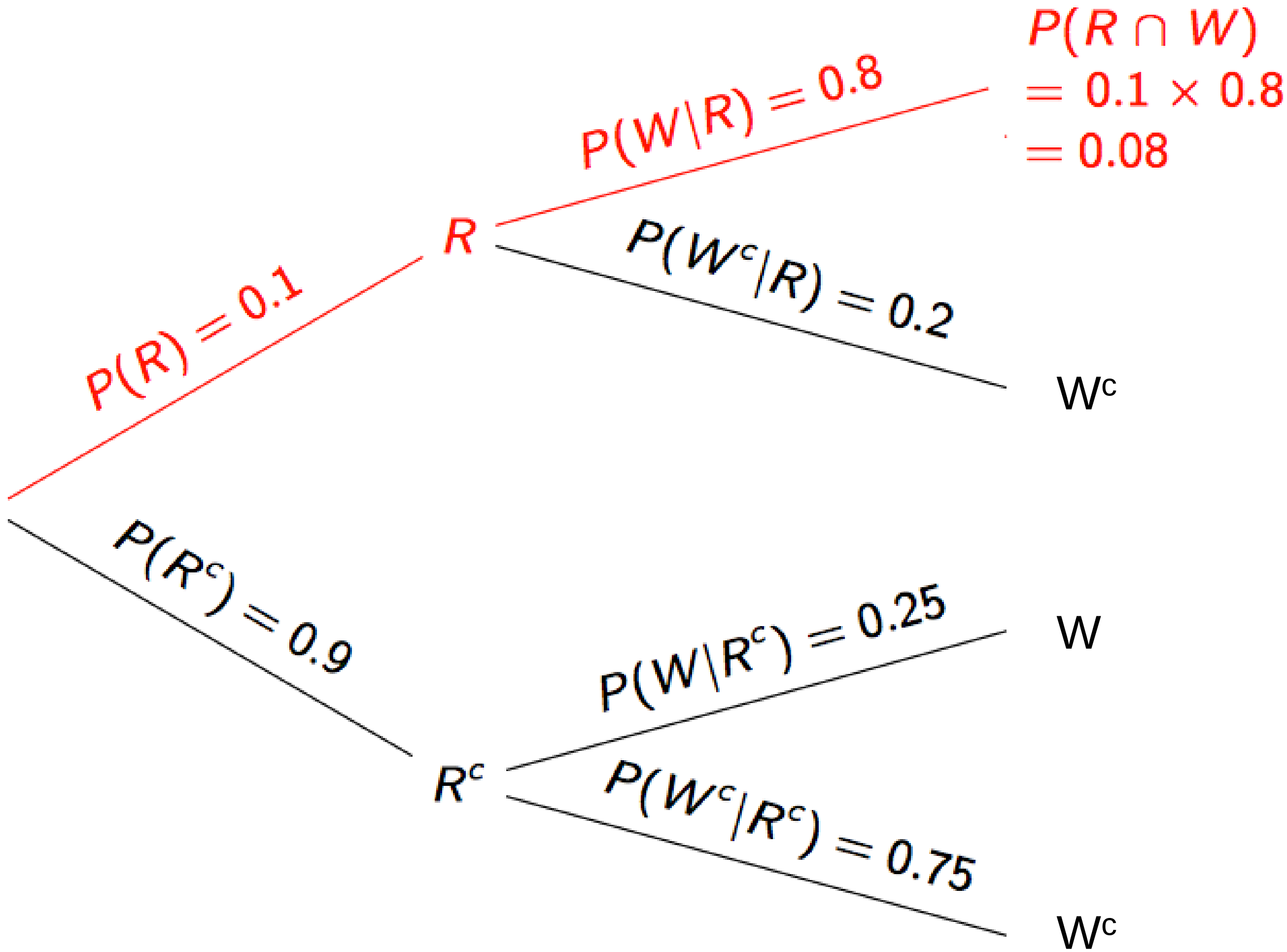
Joint probabilities are obtained by multiplying down the tree.

We want to find:
The probability that it doesn't rain tomorrow and Anita goes for a run =
$P(R^c \cap W) = 0.9 \times 0.25 = 0.225$



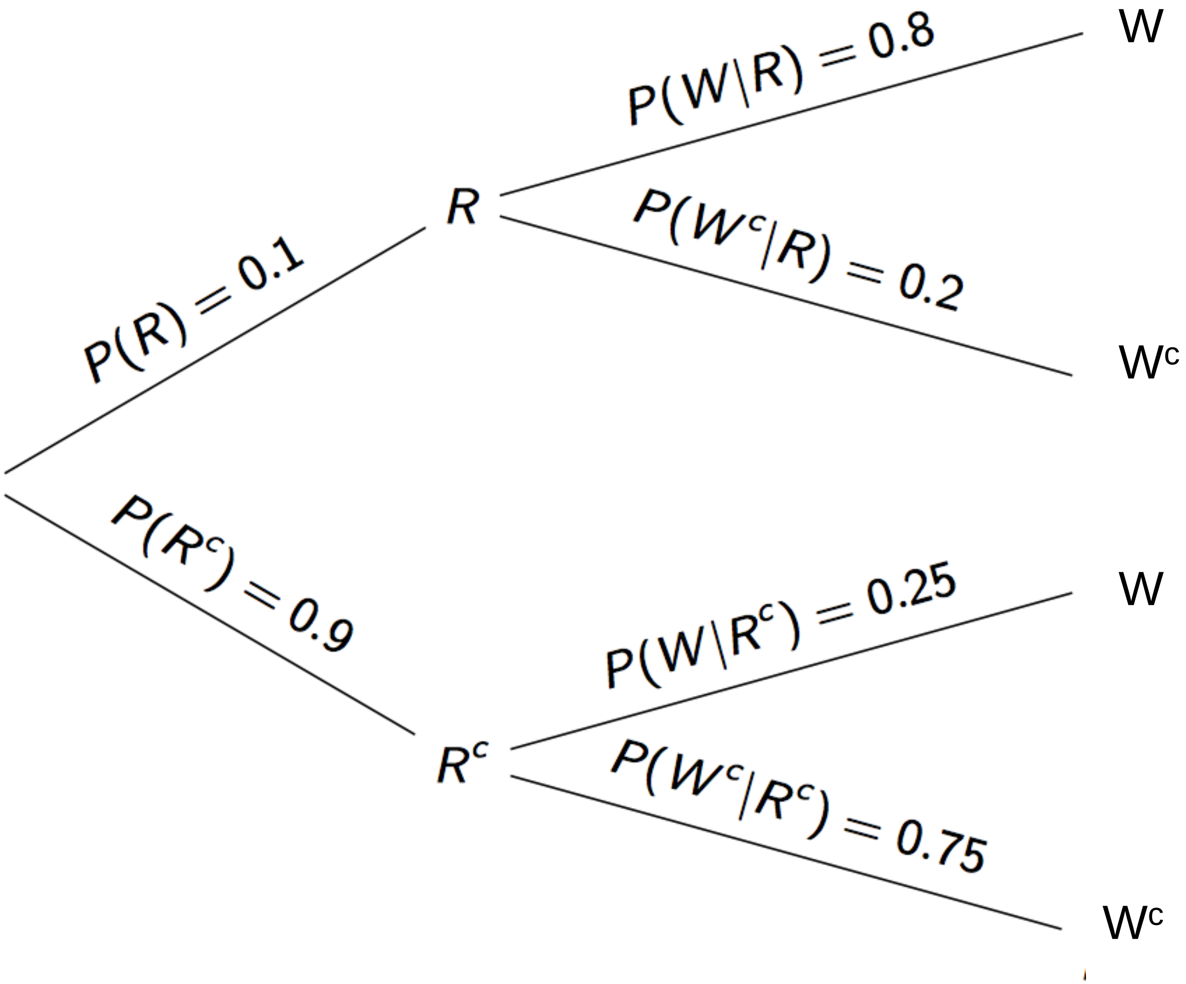
# Probability Tree

Similarly, the probability that it rains tomorrow <u>and</u> Anita goes for a run =
$P(R \cap W) = 0.1 \times 0.8 = 0.08$



# Probability Tree

Here are all the joint probabilities:




$P(R \cap W) = 0.1 \times 0.8$

$P(R \cap W^c) = 0.1 \times 0.2$

$P(R^c \cap W) = 0.9 \times 0.25$

$P(R^c \cap W^c) = 0.9 \times 0.75$

# The Multiplication Rule

This is known as the multiplication rule:

- For two events: P(A∩B) = P(A)P(B|A) = P(B)P(A|B)

- For three events: What is P(A∩B∩C)?

    - Treat (A ∩ B) as an event. Call this R
    - So P(A∩B ∩C) = P(R∩C) = P(R) P(C|R)
      = P(A ∩B)P(C|A ∩B)
    - But P(A ∩ B) = P(A)P(B|A)

    - So P(A∩B ∩C) = P(A) P(B|A) P(C|A ∩B)

- n events: Using induction we can prove that:

$$P(\cap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots A_{n-1})$$

# The Multiplication Rule

Example 1:

Three cards are drawn from an ordinary 52 card deck without replacement. What is the probability that there are no Face cards among them?

# The Multiplication Rule

Example 1 (contd.):

Three cards are drawn from an ordinary 52 card deck without replacement. What is the probability that there are no Face cards among them?

Face card : King, Queen or Jack

•Without replacement => cards don't go back into the deck after each draw.

•A1 : event that the first card is not a Face card
•A2 : event that the second card is not a Face card
•A3: event that the third card is not a Face card

•Remember: each suit has 3 Face cards and 10 Non-Face cards

•We want: $P(A1 \cap A2 \cap A3)$

# The Multiplication Rule

Example 1 (contd.):

Using multiplication rule:

$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2)$

$P(A_1) = 40/52$
$P(A_2|A_1) = 39/51$
$P(A_3|A_1 \cap A_2) = 38/50$

So $P(A_1 \cap A_2 \cap A_3) = \dfrac{(40 \times 39 \times 38)}{(52 \times 51 \times 50)} = 0.4471$

# The Multiplication Rule

Example 2:

I have two black balls, and one red ball. I pick two balls randomly without replacement.

- What is the probability that the first ball is black?
- What is the probability that the second ball is black?

# Example Problem

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

1. What is the sample space for the game?

2. What are the associated probabilities of the elements in the sample space? Do they sum to one?

3. Alice insists she should toss first. Why?

Note:

**Geometric Series:** $\sum_{k=0}^{\infty} x^k$

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \cdots \begin{cases} \dfrac{1}{1-x}, & \text{if } |x| < 1, \\ \text{diverges}, & \text{if } |x| \geq 1. \end{cases}$$

# Example Problem

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

1. What is the sample space?

Game ends when someone tosses a H. So sample space is {H, TH, TTH, TTTH, …}

2. What are the associated probabilities of the elements in the sample space? Do they sum to one?

$$\frac{1}{2}+\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3+\ldots=\frac{1/2}{1-\left(1/2\right)}$$

# Example Problem

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

3.Alice insists she should toss first. Why?

If Alice tosses first, sample space for her winning :

{H, TTH, TTTTH, …}

P(win | Alice tosses first) = $(1/2) + (1/2)^3 + (1/2)^5 + \dots$

$= 2/3$

If Alice tosses second, sample space for her winning :

{TH, TTTH, TTTTTH, …}

P(win | Alice tosses second) = $(1/2)^2 + (1/2)^4 + (1/2)^6 + \dots$

$= 1/3$

# Statistical Independence

# Statistical Independence

- Let's look at the experiment where two fair coins are tossed together.
- Suppose
  $H_1$: event that the dime lands a Head
  $H_2$: event that nickel lands a Head
- $P(H_1) = P(H_2) = ½$
- $P(H_1|H_2) =$ ½ $= P(H)$
- Knowing $H_2$ doesn't give me additional info on $H_1$
- So $H_1$ and $H_2$ are independent.
- In general,
- If $P(A|B) = P(A)$, we say the events A and B are independent.

# Statistical Independence

- The general multiplication rule gives us:
- $P(A \cap B) = P(A|B) \times P(B)$
- If A and B are independent, $P(A \cap B) = P(A) \times P(B)$
- This second equation is called the special multiplication rule for independent events.
- Independence is a symmetric property, so that if A is independent of B, then B is independent of A.

Therefore:

If two events A and B are independent,

$$P(A \cap B) = P(A)P(B)$$

# Statistical Independence

- If we want to see if two events are independent, it is best to not just read the context and answer.

- More reliable is to calculate both sides of any one of these equations to confirm equality, and show independence that way.

If two events A and B are independent,

- $P(A \cap B) = P(A)P(B)$
- $P(A|B) = P(A)$
- $P(B|A) = P(B)$

# Statistical Independence

Example: Car color preference by age among people who own one car.

| | | Blue | Red | Silver | **Total** |
|---|---|---|---|---|---|
| Age | 18-25 years | 12 | 15 | 8 | **35** |
| | 25-35 years | 5 | 3 | 12 | **20** |
| | 35-45 years | 2 | 0 | 23 | **25** |
| | **Total** | **19** | **18** | **43** | **80** |

(1) Name a few mutually exclusive events.

(2) Are any age-color combinations mutually exclusive?

(3) Are being 35-45 years old and driving a Silver car independent events?