Statistics for Scientists – CSC261

Probability

Dr DEBAJYOTI PAL

SCHOOL OF INFORMATION TECHNOLOGY,
KMUTT

Where We’re Going (Objectives)

- Probability as a Measure of Uncertainty
- Basic Rules for Finding Probabilities
- Probability as a Measure of Reliability for an Inference

Events, Sample Spaces, and Probability

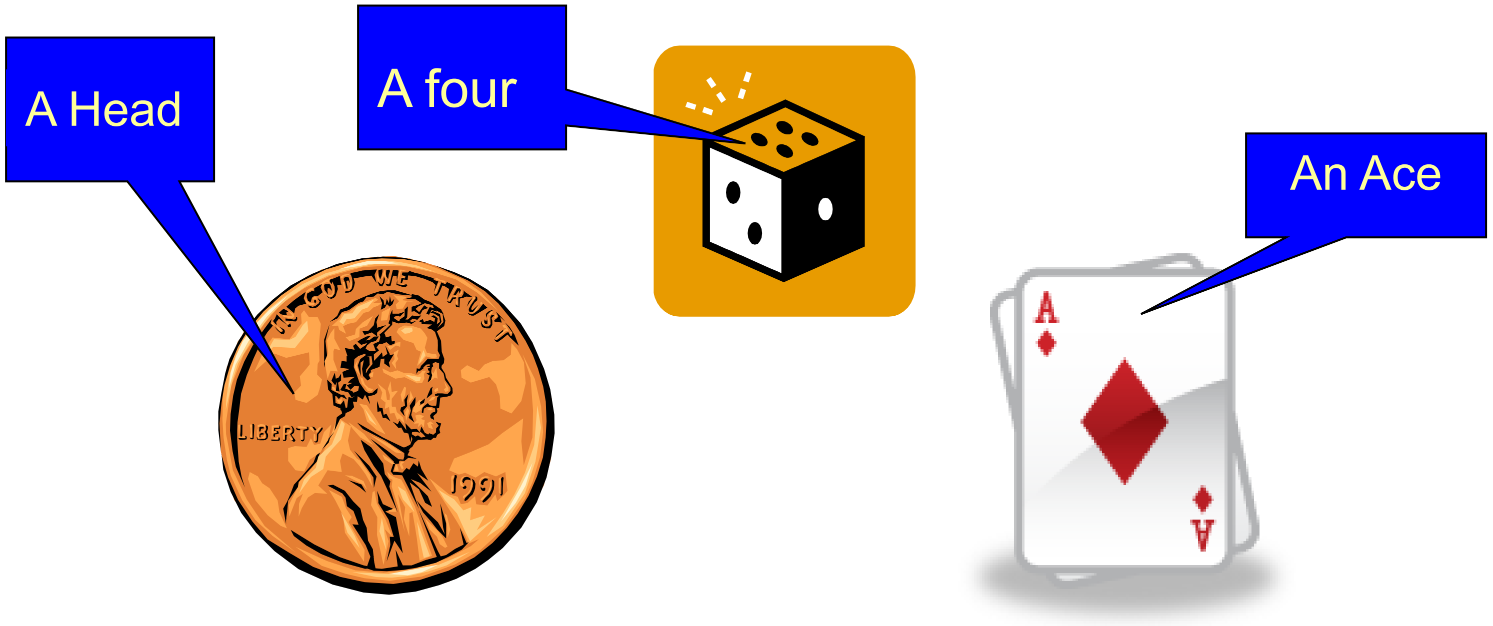
- Experiment – process of observation that leads to a single outcome with no predictive certainty
- Sample point – most basic outcome of an experiment
- Sample Space – a listing of all sample points for an experiment
- Event – a specific collection of sample points
- Example:
    - Experiment – tossing 2 coins
    - A Sample Point – HT
    - Sample Space – S: {HH, HT, TH, TT}
    - Event – Observation of at least one head → {HH, HT, TH}
- Sample point probability – relative frequency of the occurrence of the sample point

Events, Sample Spaces and Probability

- An **experiment** is an act or process of observation that leads to a single outcome that cannot be predicted with certainty.

Events, Sample Spaces and Probability

- A **sample point** is the most basic outcome of an experiment.



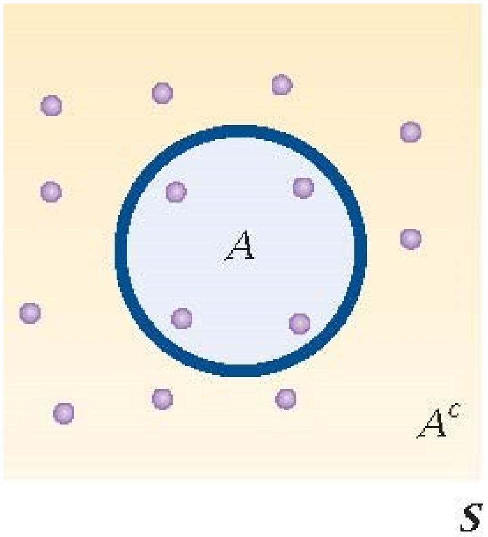
Events, Sample Spaces and Probability

- A **sample space** of an experiment is the collection of all sample points.
  - Roll a single die:

S: {1, 2, 3, 4, 5, 6}

Events, Sample Spaces and Probability

- **Sample points and sample spaces** are often represented with Venn diagrams.



Events, Sample Spaces and Probability

- Probability Rules for Sample Points

  - All probabilities must be between 0 and 1, inclusive.

$$0 \leq p_i \leq 1$$

  - The probabilities of *all* the sample points must sum to 1.

$$\sum_{i=1}^{n} p_i = 1$$

Events, Sample Spaces and Probability

- Probability Rules for Sample Points

- All probabilities must be between 0 and 1

$$0 \le p_i \le 1$$

0 indicates an impossible outcome and 1 a certain outcome.

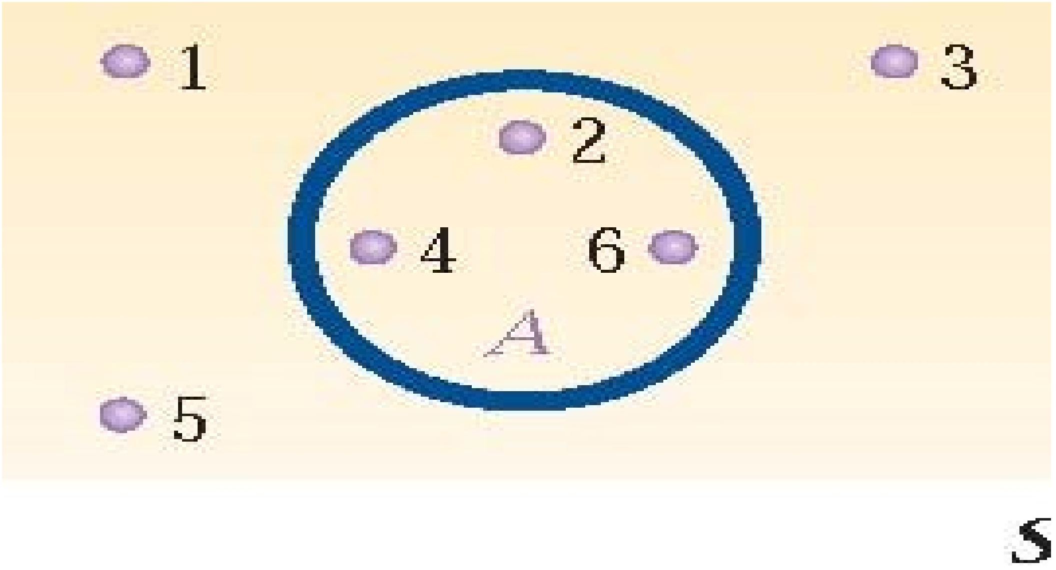
- The probabilities of *all* the sample points must sum to 1

$$\sum_{i=1}^{n} p_i = 1$$

***Something** has to happen.*

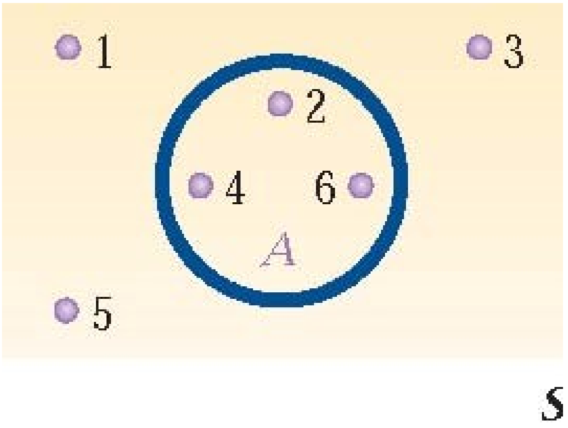
Events, Sample Spaces and Probability

- An **event** is a specific collection of sample points:
  - Event *A*: Observe an even number of a die rolling.



Events, Sample Spaces and Probability

- The **probability of an event** is the sum of the probabilities of the sample points in the sample space for the event.
  - Event *A*:
    - Observe an even number.
  - P(*A*) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2

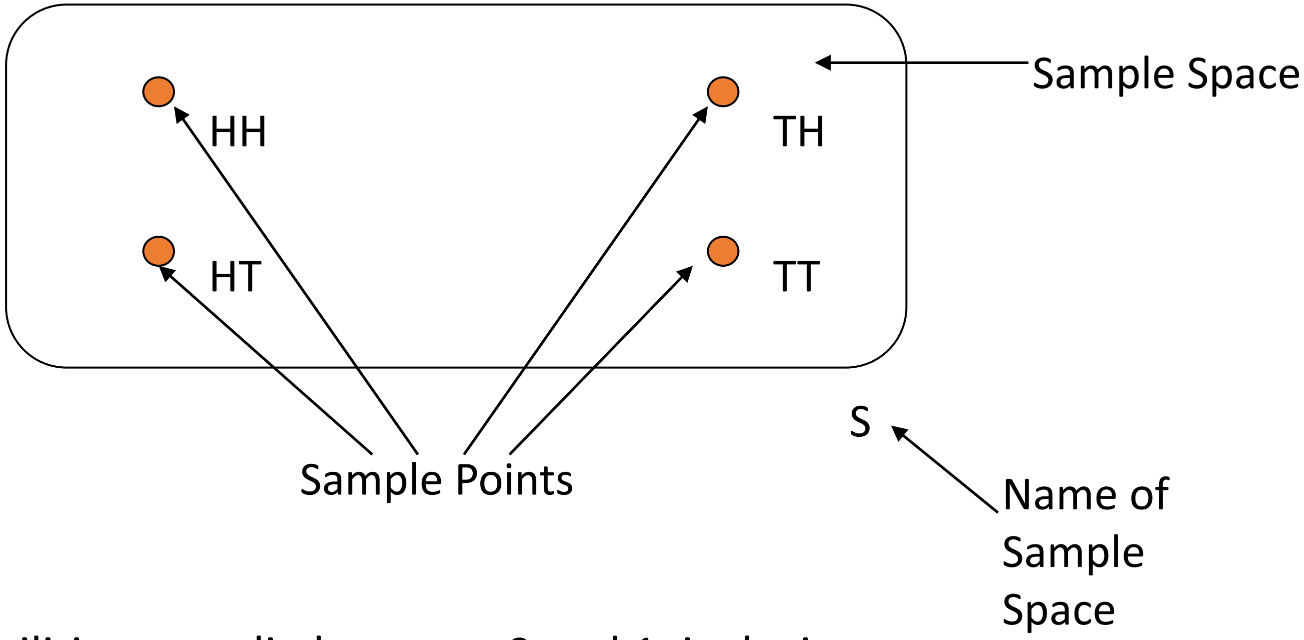


Events, Sample Spaces and Probability

- Calculating Probabilities for Events
  - Define the experiment.
  - List the sample points.
  - Assign probabilities to sample points.
  - Collect all sample points in the event of interest.
  - The sum of the sample point probabilities is the event probability.

Events, Sample Spaces and Probability

- Venn Diagram



- Sample Point Probabilities must lie between 0 and 1, inclusive
- The sum of all sample point probabilities must be one

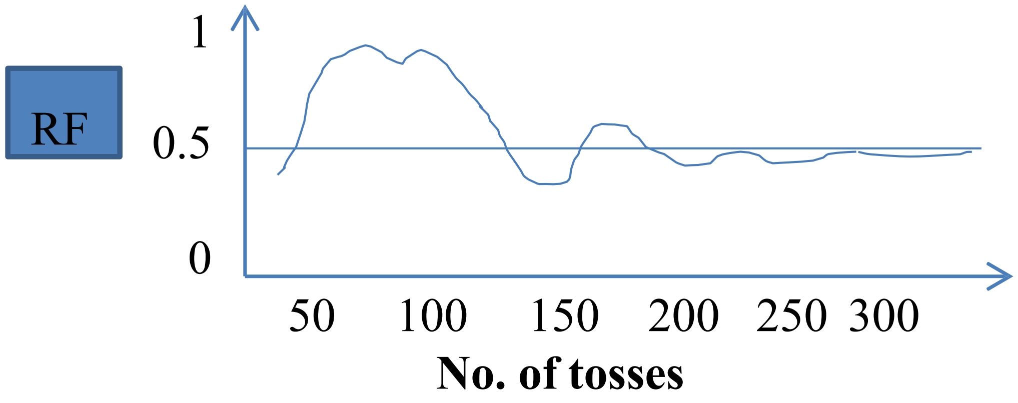
Three Types of Probability

1. Observed/relative frequency approach
2. Classical approach
3. Subjective approach

1. Observation/Relative Frequency (RF) :

This method uses the **relative frequencies of past occurrences** as probabilities.
How often something has **happened** in past - we **predict** future

More trails , greater accuracy: Tossing a fair coin for 300 times. In first 100 tosses prob **is far from** 0.5, but approaches 0.5 as we increase number of toss. RF becomes **stable** as no. of tosses **become large**.



Limitation: We need **sufficient no. of experiments** or observations before conclusion.

Three Types of Probability

2. Classical approach

Prob of an event = (no. of outcomes where the event occurs)/(total number of possible outcomes)

P(H) = 1/(1+1) = 1/2

Total possible outcomes

P(5) = 1/6 for the dice rolling example

Classical prob is also called **apriori probability** because we don't need to perform experiments.

3. Subjective probability : Based on belief, experience, when event has occurred once or few times.

Because most higher-level social and managerial decisions are concerned with **specific**, **unique** satiations, rather than with a **long series of identical** situation, **decisions makers** use this prob.

Activity 1: Retirement policy is to be presented to top management. To know the support of the policy a manger conducts a poll.

	Machinists	inspector
Strongly support	9	10
Mildly support	11	3
Undecided	2	2
Mildly oppose	4	8
Strongly oppose	4	7
	30	30

a)What is the prob that a **machinist** randomly selected from the polled group **mildly supports** the package.
b) What is the prob that an **inspector** randomly selected from the polled group **is undecided**.
c)What is the prob that a worker (machinist or inspector) randomly selected from the polled group **strongly or mildly supports** the package.
d) What prob **estimates** are these.

Activity 2: Classify the following probability estimates as to their type (classical, observed, or subjective):

(a) The probability of scoring on a penalty shot in ice hockey is 0.47.

(b) The probability that the current Mayor will resign is 0.85.

(c) The probability of rolling two sixes with two dice is 1/36.

(d) The probability that a president elected in a year ending in zero will die in office is 7/10.

(e) The probability that you will go Europe this year is 0.14.

Mutually Exclusive Events

- Mutually exclusive events
  - Events that **cannot** occur simultaneously

Example: Drawing one card from a deck of cards

A = queen of diamonds; B = queen of clubs

  - Events A and B are mutually exclusive

Collectively Exhaustive Events

- Collectively exhaustive events
  - One of the events must occur
  - The set of events covers the entire sample space

example:

A = aces; B = black cards; C = diamonds; D = hearts

  - Events A, B, C and D are collectively exhaustive (but not mutually exclusive – **an ace may also be a heart**)
  - Events B, C and D are collectively exhaustive and also mutually exclusive

Ex: Give a collectively exhaustive list of the possible outcomes of **two dice**.

Events, Sample Spaces and Probability

- What is the probability of rolling an eight in a single toss of a pair of dice?
  - Experiment is toss of pair of dice
- Probability of rolling an 8 = 1/36+1/36+1/36+1/36+1/36 = 5/36= **0.14**

1,1 1/36	1,2 1/36	1,3 1/36	1,4 1/36	1,5 1/36	1,6 1/36
2,1 1/36	2,2 1/36	2,3 1/36	2,4 1/36	2,5 1/36	2,6 1/36
3,1 1/36	3,2 1/36	3,3 1/36	3,4 1/36	3,5 1/36	3,6 1/36
4,1 1/36	4,2 1/36	4,3 1/36	4,4 1/36	4,5 1/36	4,6 1/36
5,1 1/36	5,2 1/36	5,3 1/36	5,4 1/36	5,5 1/36	5,6 1/36
6,1 1/36	6,2 1/36	6,3 1/36	6,4 1/36	6,5 1/36	6,6 1/36

Ex: What is the probability for each of the following totals in the rolling of two dice: 1,2,5,6,7,10, and 11.

Events, Sample Spaces and Probability

- **Tree Diagrams**
- Sample spaces can also be described graphically with **tree diagrams.**
  - When a sample space can be constructed in several steps or stages, we can represent each of the *n1* ways of completing the first step as a branch of a tree.
  - Each of the ways of completing the second step can be represented as *n2* branches starting from the ends of the original branches, and so forth.

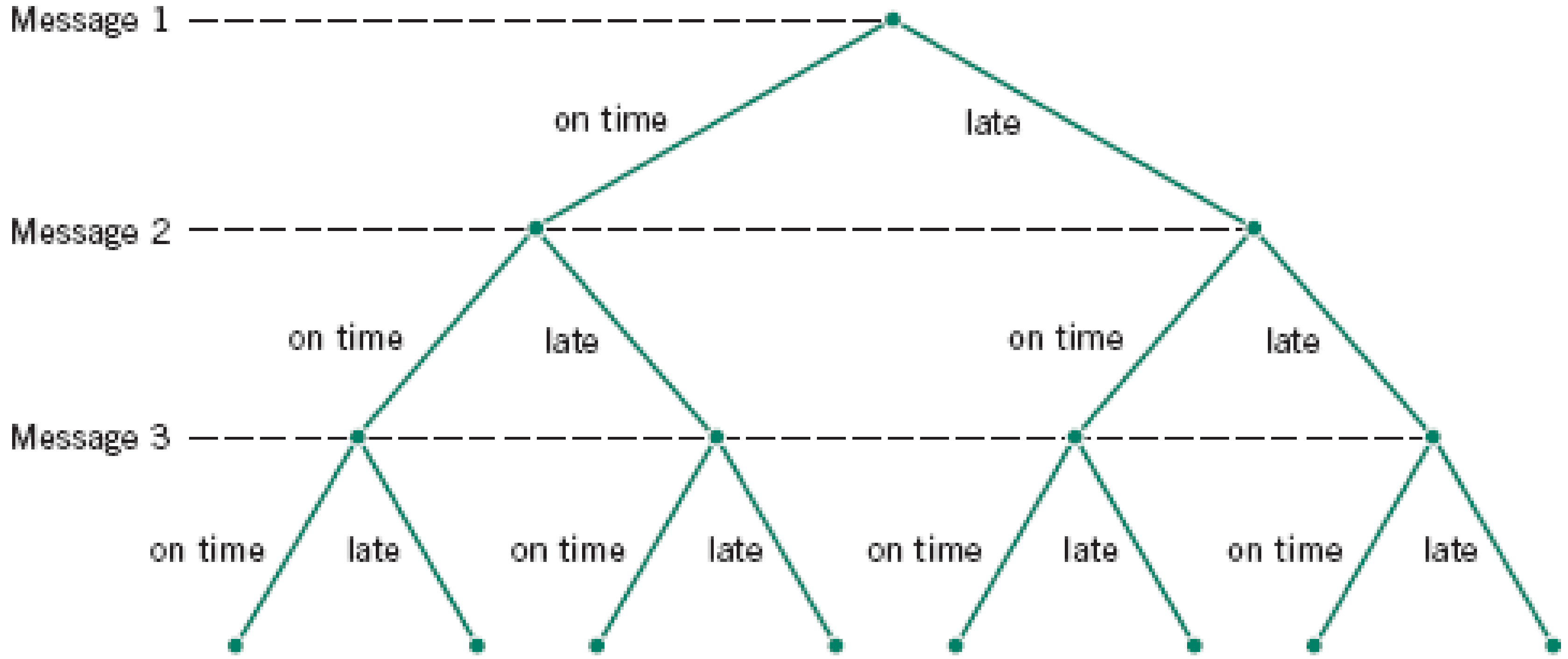
Events, Sample Spaces and Probability

Example

Each message in a digital communication system is classified as to whether it is received within the time specified by the system design. If three messages are classified, use a tree diagram to represent the sample space of possible outcomes.

Each message can be either received on time or late. The possible results for three messages can be displayed by eight branches.

Events, Sample Spaces and Probability



Tree diagram for three messages.

Unions and Intersections

Compound Events

Made of two or more other events

Union

$A \cup B$

Either A or B, or both, occur

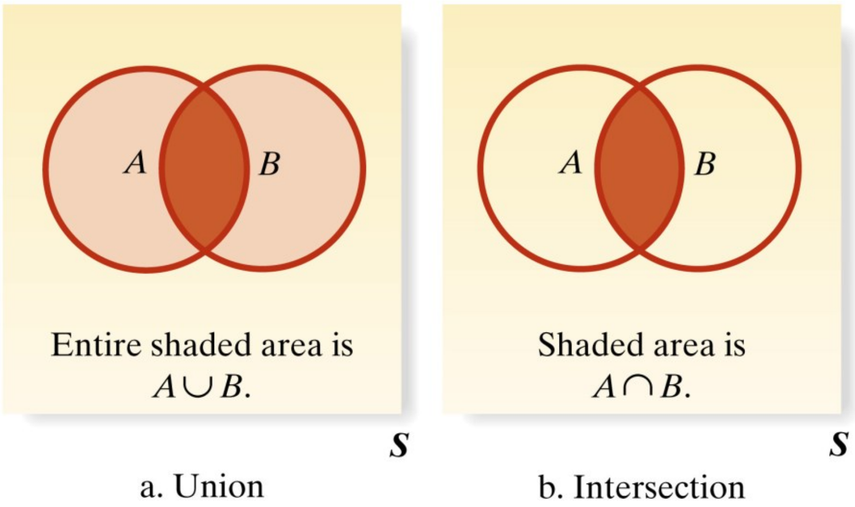
Intersection

$A \cap B$

Both A and B occur

Unions and Intersections

- Compound Event – a composition of 2 or more events
- Can be the result of a union or intersection of events



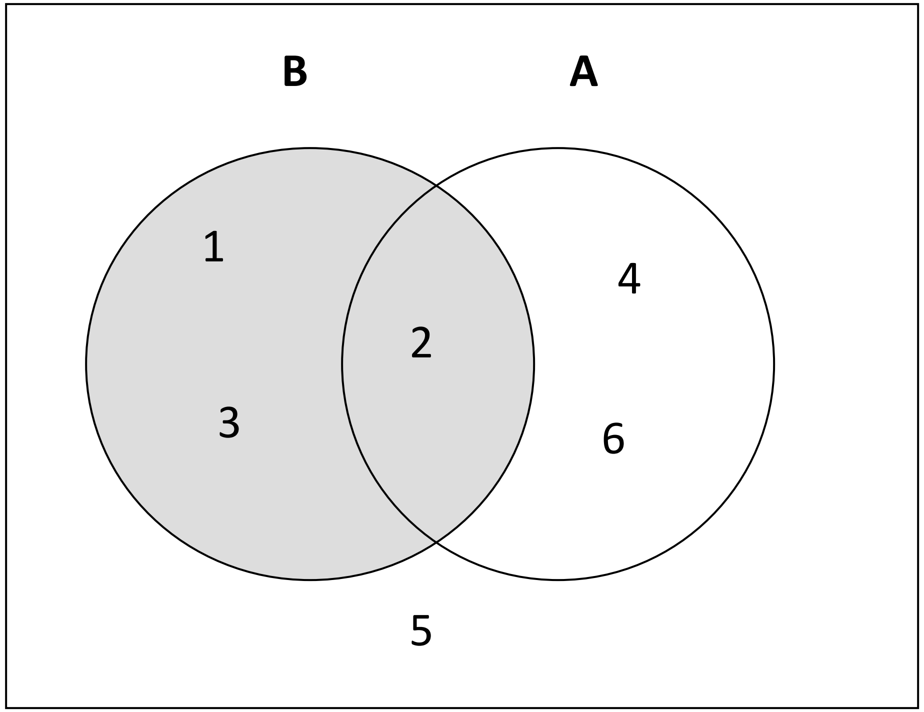
Complementary Events

- The **complement** of any event *A* is the event that *A* does not occur, A^C.

A: {Toss an even number}
A^C: {Toss an odd number}

B: {Toss a number ≤ 3}
B^C: {Toss a number ≥ 4}

$A \cup B = \{1,2,3,4,6\}$
$[A \cup B]^C = \{5\}$
(Neither A nor B occur)



Complementary Events

$$P(A) + P(A^C) = 1$$

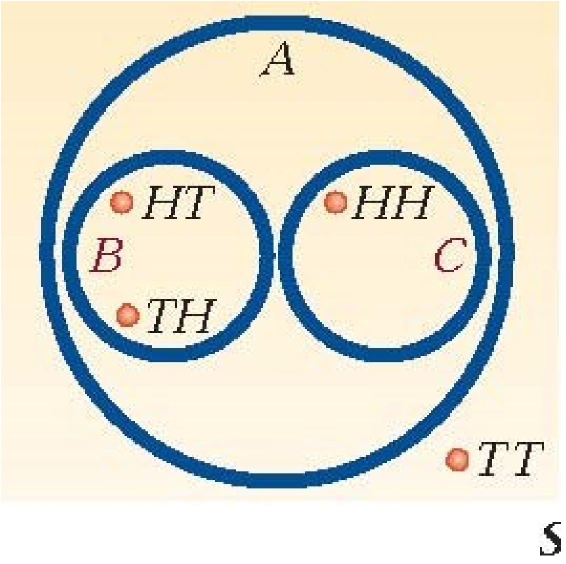
$$P(A) = 1 - P(A^C)$$

$$P(A^C) = 1 - P(A)$$

Complementary Events

A: {At least one head on two coin flips}

A^C: {No heads}



$$A:\{HH, HT, TH\}$$

$$A^C:\{TT\}$$

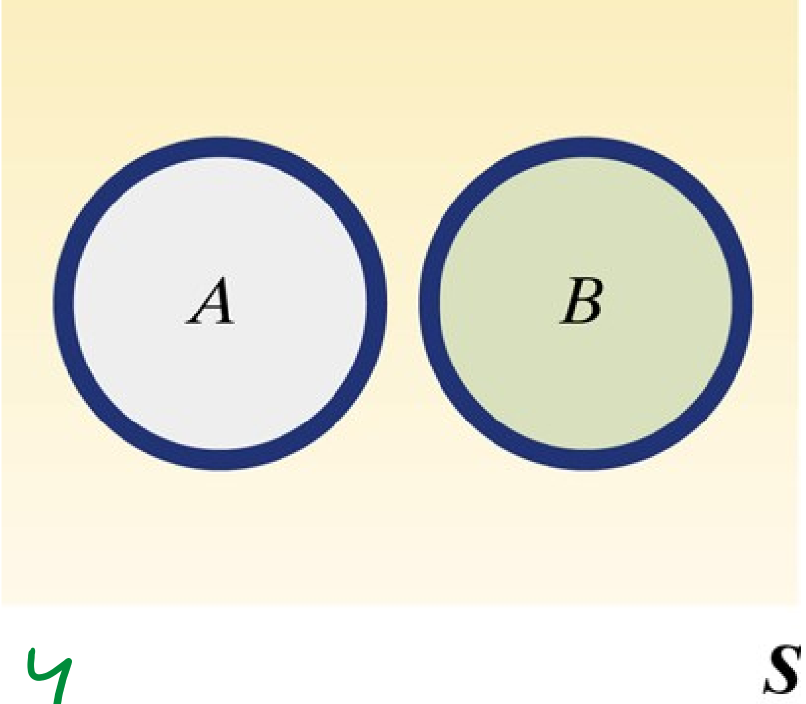
$$P(A) = 1/4 + 1/4 + 1/4 = 3/4$$

$$P(A^C) = 1/4$$

$$P(A) = 1 - P(A^C) = 1 - 1/4 = 3/4$$

The Additive Rule and Mutually Exclusive Events

- Mutually Exclusive Events – Events are mutually exclusive if they share no sample points.



$$P(A \cap B) = 0$$

The Additive Rule and Mutually Exclusive Events

- The Additive Rule for Mutually Exclusive Events

$$P(A \cup B) = P(A) + P(B)$$

Here, $P(A \cap B)$ is 0.

Unions and Intersections

- Event A – A New Jersey Birth Mother is white
- Event B – A New Jersey Birth Mother was a teenager when giving birth

$$P(A) = .79$$

$$P(B) = .09$$

$$P(A \cap B) = .05$$

TABLE 3.4 Percentage of New Jersey Birth Mothers in Age–Race Classes

	Race	
Maternal Age (years)	White	Black
≤17	2%	2%
18–19	3%	2%
20–29	41%	12%
≥30	33%	5%

Source: Reichman, N. E., and Pagnini, D. L. "Maternal age and birth outcomes: Data from New Jersey." *Family Planning Perspectives*, Vol. 29, No. 6, Nov./Dec. 1997, p. 269 (adapted from Table 1).

$$P(A \cup B) = .02 + .03 + .41 + .33 + .02 + .02 = .83$$

The Additive Rule

- The Additive Rule

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) =$$

$$.79 + .09 - .05 = .83$$

TABLE 3.4 **Percentage of New Jersey Birth Mothers in Age–Race Classes**

	Race	
Maternal Age (years)	White	Black
≤17	2%	2%
18–19	3%	2%
20–29	41%	12%
≥30	33%	5%

Source: Reichman, N. E., and Pagnini, D. L. "Maternal age and birth outcomes: Data from New Jersey." *Family Planning Perspectives*, Vol. 29, No. 6, Nov./Dec. 1997, p. 269 (adapted from Table 1).

Conditional Probability

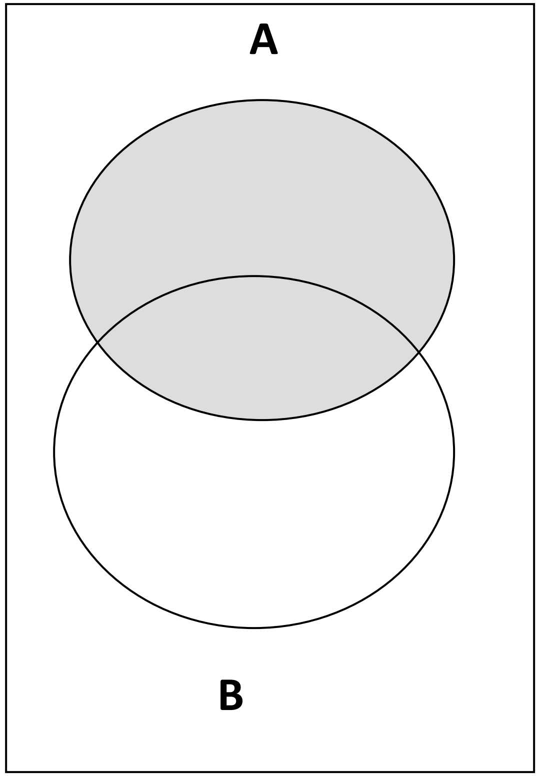
- Conditional Probability – the probability that event A occurs given that event B occurs

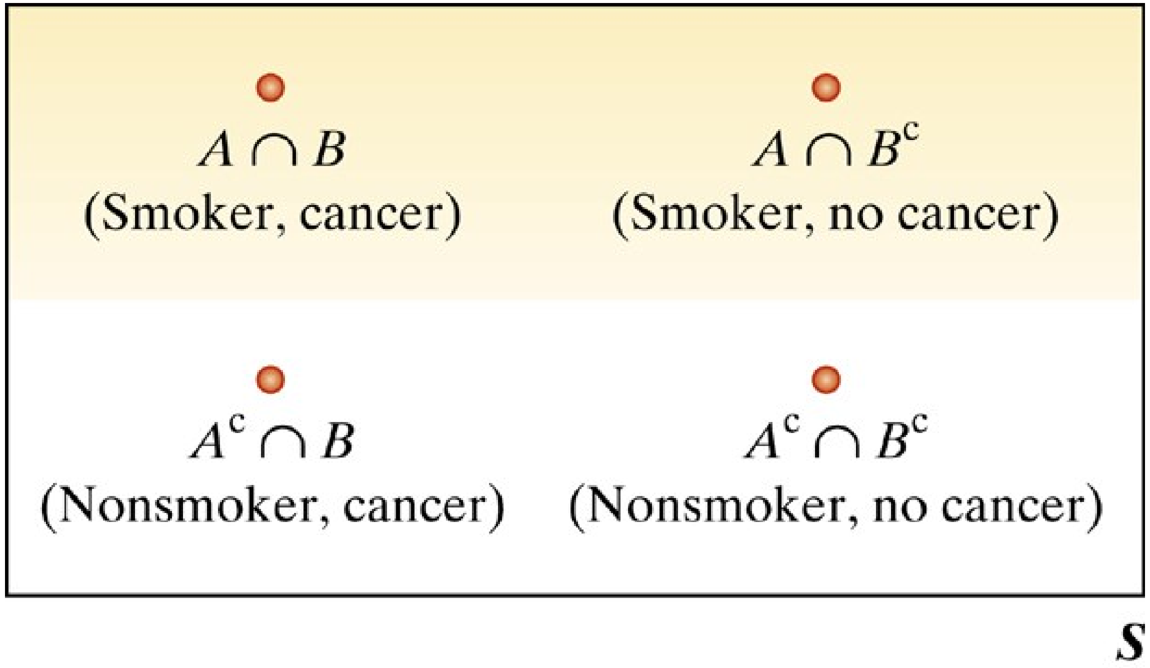
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Conditional probability works with a reduced sample space, the space that contains B and $A \cap B$

Conditional Probability

- Sample space





Probabilities of Smoking and Developing Cancer of Individuals

		Develops **Cancer**	
		Yes, B	No, B^c
Smoker	Yes, A	0.05	0.20
	No, A^c	0.03	0.72

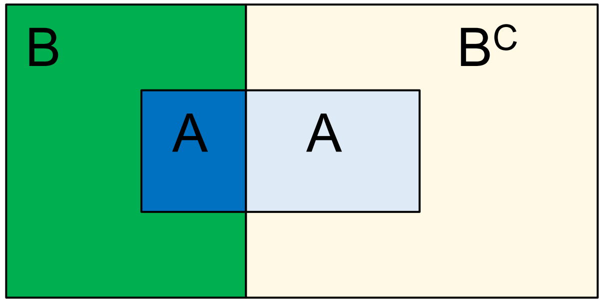
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.05}{0.25} = 0.20$$

Conditional Probability

- Additional information may have an impact on the probability of an event.
    - P(Rolling a 6) is one-sixth (unconditionally).
    - If we know an even number was rolled, the probability of a 6 goes up to one-third.

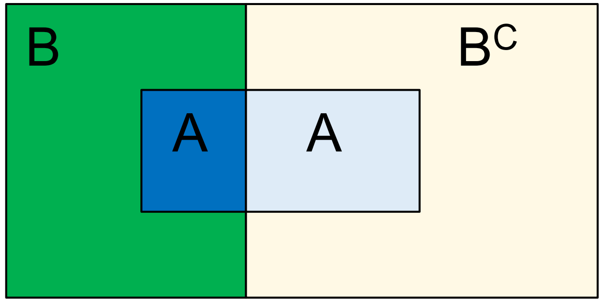
Conditional Probability

- The sample space is reduced to only the conditioning event.
- To find P(A), once we know B has occurred (i.e., *given* B), we ignore B^C (including the A region within B^C).



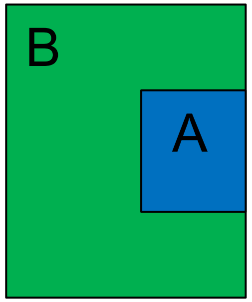
Conditional Probability

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



Conditional Probability

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



Conditional Probability

- Event A – cause of complaint is appearance
- Event B – complaint occurred during guarantee period

TABLE 3.6 Distribution of Product Complaints

	Reason for Complaint			
	Electrical	Mechanical	Appearance	Totals
During Guarantee Period	18%	13%	32%	63%
After Guarantee Period	12%	22%	3%	37%
Totals	30%	35%	35%	100%

$$P(A \cap B) = .32$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.32}{.63} = .51$$

Conditional Probability

- 55% of sampled executives had cheated at golf (event A).
  - $P(A) = .55$
- 20% of sampled executives had cheated at golf *and* lied in business (event B).
  - $P(A \cap B) = .20$
- What is the probability that an executive had lied in business, *given* s/he had cheated in golf, $P(B|A)$?

Conditional Probability

- P(A) = .55
- P(A ∩ B) = .20
- What is P(B|A)?

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

$$P(B|A) = \frac{.20}{.55} = .364$$

The Multiplicative Rule and Independent Events

- Example:

10 welfare workers interview prospective food stamp recipients and are randomly checked by their supervisor for illegal deductions. 3 of the workers have regularly given illegal deductions to applicants. *What's the probability that 2 of the workers chosen have given illegal deductions?*

The Multiplicative Rule and Independent Events

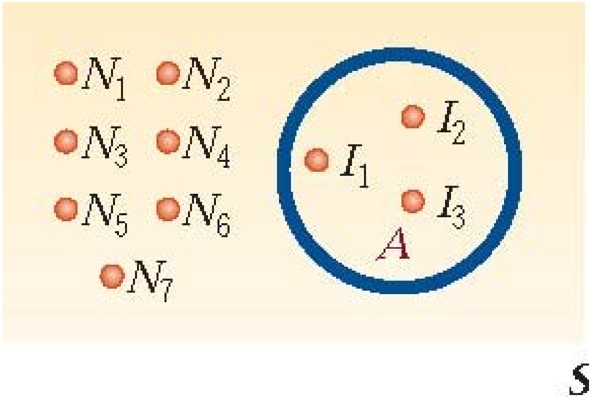
Assume three of ten workers give illegal deductions

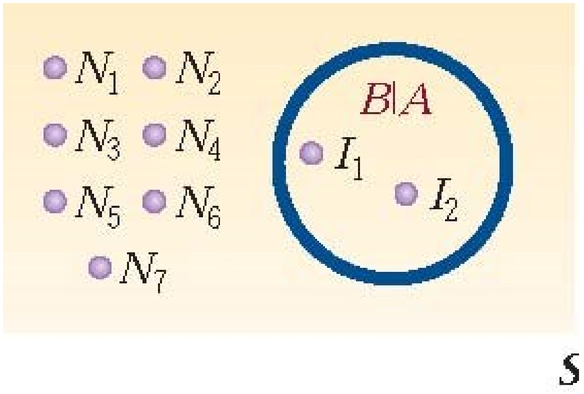
- Event A: {First worker selected gives an illegal deduction}
- Event B: {Second worker selected gives an illegal deduction}
  - $P(A) = .1 + .1 + .1 = .3$
- $P(B|A)$ has only nine sample points, and two targeted workers, since we selected one of the targeted workers in the first round:
  - $P(B|A) = 1/9 + 1/9 = .11 + .11 = .22$
- The probability that both of the first two workers selected will have given illegal deductions
  - $P(A \cap B) = P(B|A)P(A) = (0.22)(0.3) = .066…$
  - Or, first chosen illegal deductions = 3/10

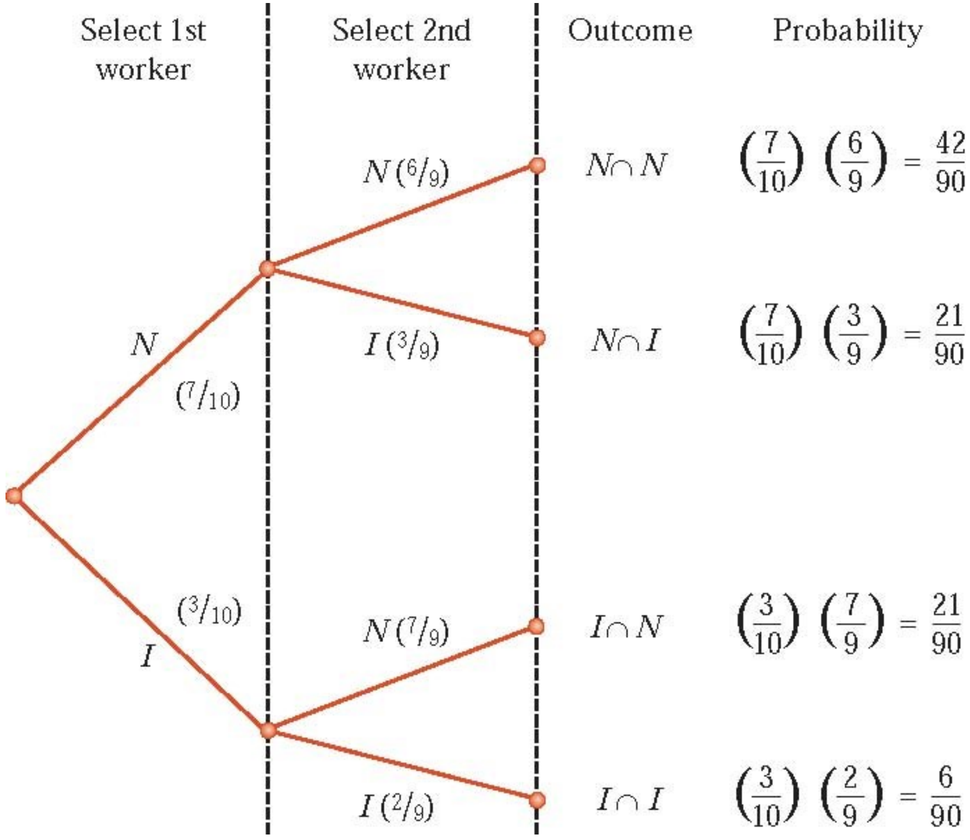
    second one is 2/9,

    thus the chance is $(3/10)*(2/9) = 6/90 = 0.66…$

The Multiplicative Rule and Independent Events



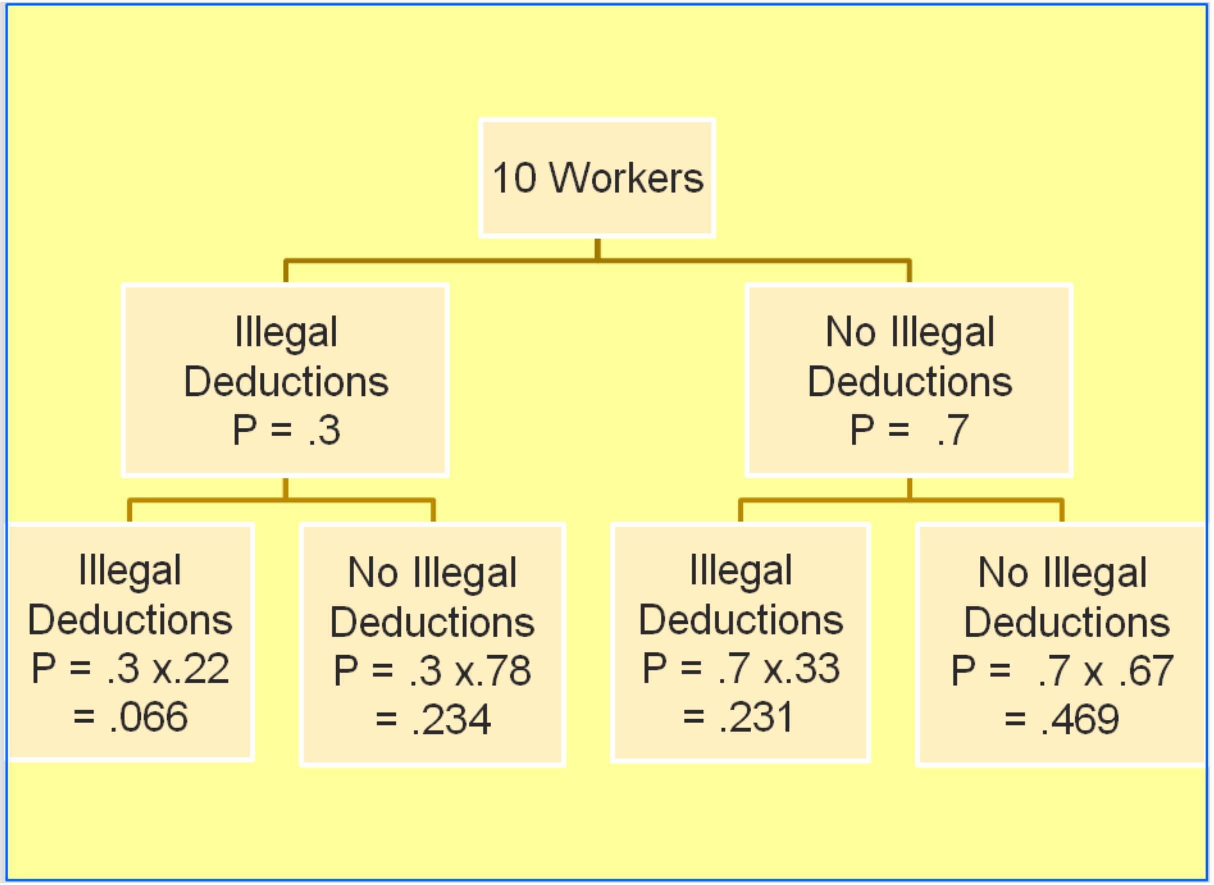




The Multiplicative Rule and Independent Events

First selected worker

Second selected worker



A Tree Diagram

The Multiplicative Rule and Independent Events

- Dependent Events
  - $P(A|B) \neq P(A)$
  - $P(B|A) \neq P(B)$

Mutually exclusive events are dependent: $P(B|A) = 0$

- Independent Events
  - $P(A|B) = P(A)$
  - $P(B|A) = P(B)$

Since $P(B|A) = P(B)$,

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(A)P(B) \end{aligned}$$

The Multiplicative Rule and Independent Events

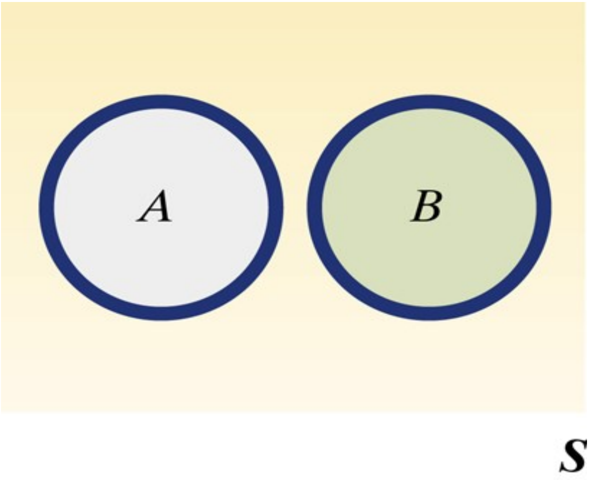
- Events A and B are independent if the occurrence of one does not alter the probability of the other occurring

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B)$$

- If A and B are independent events

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

Are these two events independent?



$$P(A \cap B) = 0$$

No, because $P(A|B) \neq P(A)$.

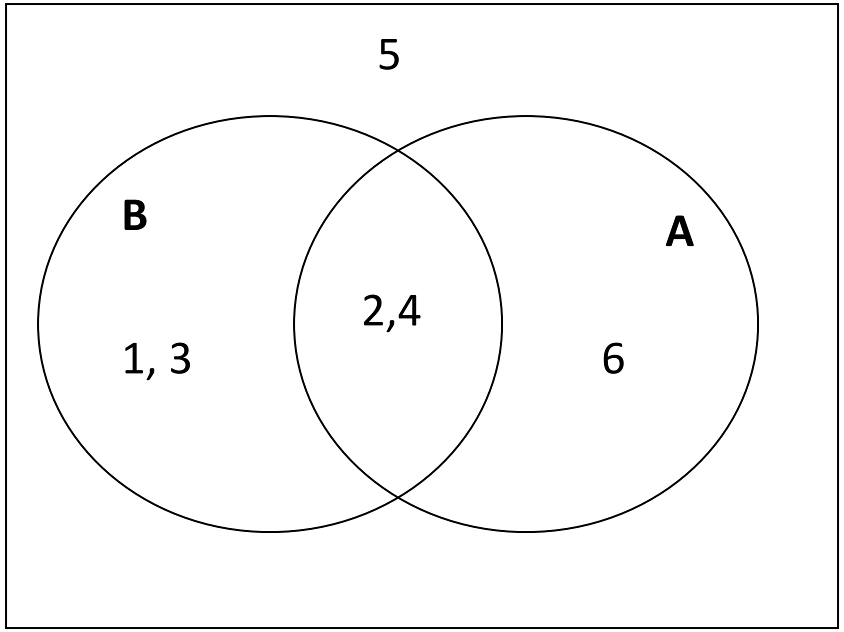
Example:

Given two events of tossing a fair die as:

A = {Observe an even number.}

B = {Observe a number less than or equal to 4.}

Are *A* and *B* independent events?



Yes,

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

$$\frac{1/3}{2/3} = \frac{1}{2} = P(A)$$

The Chain Rule

$$P(A \cap B \cap ... \cap Z) = P(A \mid B \cap C \cap ... \cap Z) *$$
$$P(B \mid C \cap D \cap ... \cap Z) * ...$$
$$P(Y \mid Z) * P(Z)$$

For example, with 3 variable, we'll have:

$$P(A \cap B \cap C) = P(A \mid B \cap C) * P(B \mid C) * P(C)$$
$$\because P(A \cap B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} * \frac{P(B \cap C)}{P(C)} * P(C)$$

Activity 1

Number of children	0	1	2	3	4	5	6 or more
Proportion of having this many children	0.05	0.10	0.30	0.25	0.15	0.10	0.05

What is the prob of a randomly chosen family having 4 or more children?

Activity 2: The employees have selected five representatives to represent them to management. A spokesperson is to be selected.

Gender	Age
Male	30
Male	32
Female	45
Female	20
Male	40

What is the prob the spokesperson will be *either* female *or* over 35?

Activity 3: An inspector of the Alaska pipeline has the task of comparing the **reliability** of **two pumping stations**. Each station is susceptible to **two kinds of failure**: **pump** failure and **leakage**. When either (or both) occur, the station must be shut down. The data at hand indicate that the following probabilities prevail:

Station	P (Pump failure)	P (Leakage)	P (Both)
1	0.07	0.10	0
2	0.09	0.12	0.06

Which station has the higher probability of being shut down?

Activity 4: Record of 45 years of a jail where prisoners tried to escape.

Attempted Escapes	Winter	Spring	Summer	Fall
0	3	2	1	0
1-5	15	10	11	12
6-10	15	12	11	16
11-15	5	8	7	7
16-20	3	4	6	5
21-25	2	4	5	3
More than 25	2	5	4	2
Total	45	45	45	45

What is the prob that in a year **selected at random**, the number of escapes was **between 16 and 20 during the winter.**
What is the prob that **more than 10** escapes were during **summer**.
What is the prob that **between 11 and 20 escapes** were attempted **during a randomly chosen season**.