Probability

Sravani Vadlamani

QMB 3200

09/01/2020

Probability

- Chapter 3 from Textbook

Probability

- The measure of uncertainty associated with events and their outcomes
- Uncertainty is also called randomness
- In the early days, probability was associated with gambling or game of chance
- Examples of events
  - Weather
  - Coin toss
  - Die roll
  - Poker
  - Stocks
- The possible outcomes are known but unknown is the specific outcome that will happen

Random Experiment

- Process that leads to one of several possible outcomes
- Process of generating an observation, outcome or simple event

Coin toss	Heads, Tails
Die Roll	1, 2,3,4,5,6,
Grades	A, B, C, D, F

Properties of Random Experiment

- Repetitions of same experiment produces same or different individual outcomes
- Outcomes cannot be predicted with certainty
- Drop a penny from 100 feet
  - Time to reach the floor – deterministic physics
  - Heads or tails is probability

Simple Event

- Basic outcome of an experiment that cannot be decomposed into simpler outcomes
- Example
  - Heads or Tails in a coin toss
  - Rolling 1 on a die

Event

- Collection of one or more simple events in a sample space
- Example
  - Die Roll

Number 4 will be rolled

Simple Event

An odd number (1,3 or 5) will be rolled

Event

Sample Space

- Set of all possible outcomes of an experiment
- Example
  - Coin Toss
    - { H, T}
  - Die Roll
    - {1, 2, 3, 4, 5, 6}
  - Choosing a card from a deck of cards

Ace ♥	2 ♥	3 ♥	4 ♥	5 ♥	6 ♥	7 ♥	8 ♥	9 ♥	10 ♥	Jack ♥	Queen ♥	King ♥
Ace ♦	2 ♦	3 ♦	4 ♦	5 ♦	6 ♦	7 ♦	8 ♦	9 ♦	10 ♦	Jack ♦	Queen ♦	King ♦
Ace ♠	2 ♠	3 ♠	4 ♠	5 ♠	6 ♠	7 ♠	8 ♠	9 ♠	10 ♠	Jack ♠	Queen ♠	King ♠
Ace ♣	2 ♣	3 ♣	4 ♣	5 ♣	6 ♣	7 ♣	8 ♣	9 ♣	10 ♣	Jack ♣	Queen ♣	King ♣

Examples

- Experiment – Toss 2 Coins
- Sample Space
  - {HH, HT, TH, TT}

1Head & 1 Tail	HT, TH
Heads on 1st Coin	HH, HT
At least 1 Head	HH, HT, TH
Heads on Both	HH

Sample Space Exercises

- Picking 2 marbles from a bag that contains Red and Black marbles

Sample Space Exercises

- Pick a card from a stack of cards that spell POLYTECHNIC

Sample Space Exercises

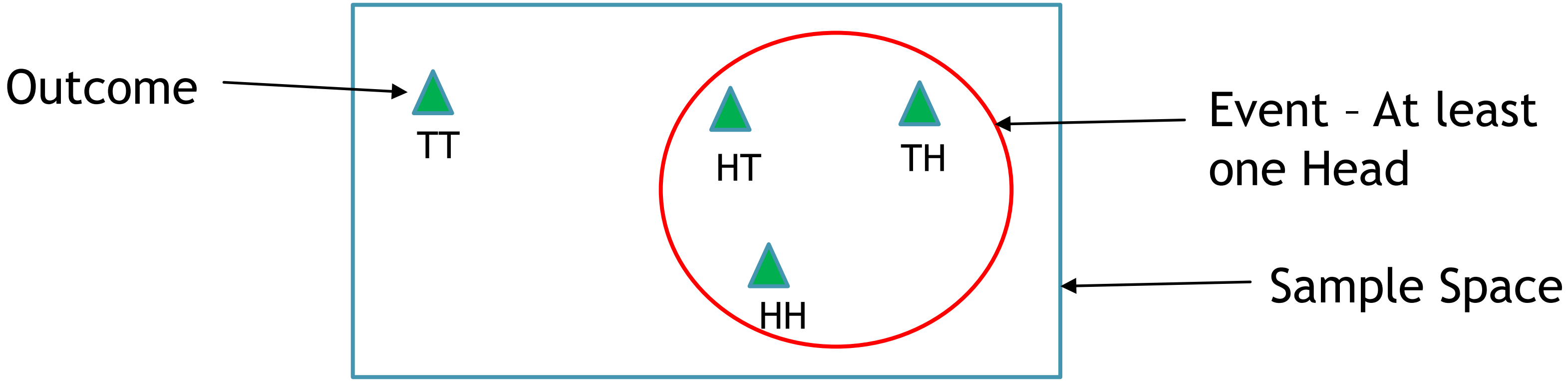
- You are hungry while on campus. The cafeteria has the option to choose from small, medium or large pizza with cheese or pepperoni. What are the different combinations you have?

Sample Space Exercises

- Determine your sample space when you roll a die and toss a coin at the same time

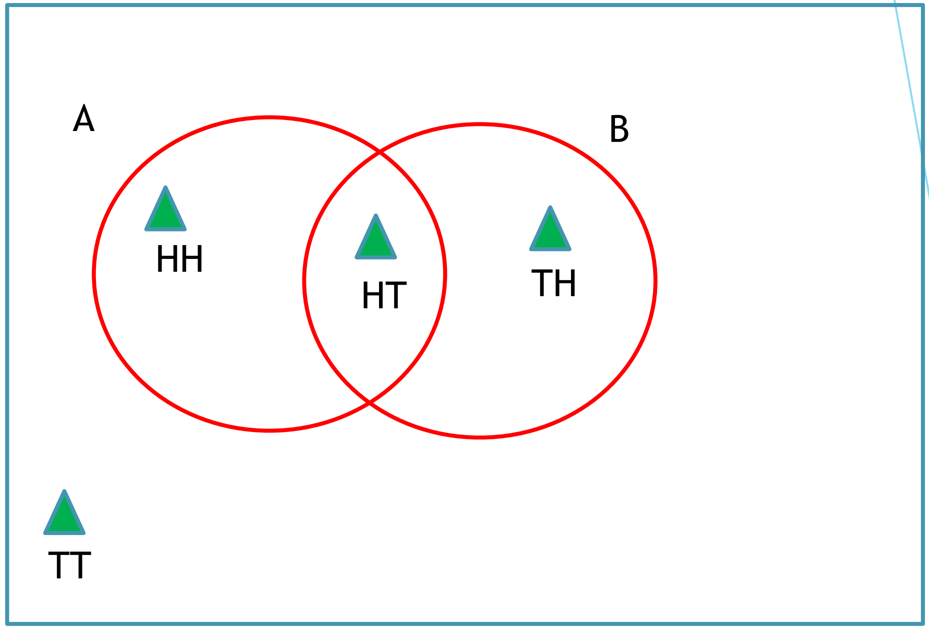
Venn Diagram Visualization

- Experiment – Toss 2 Coins
- Sample Space {HH, HT, TH, TT}



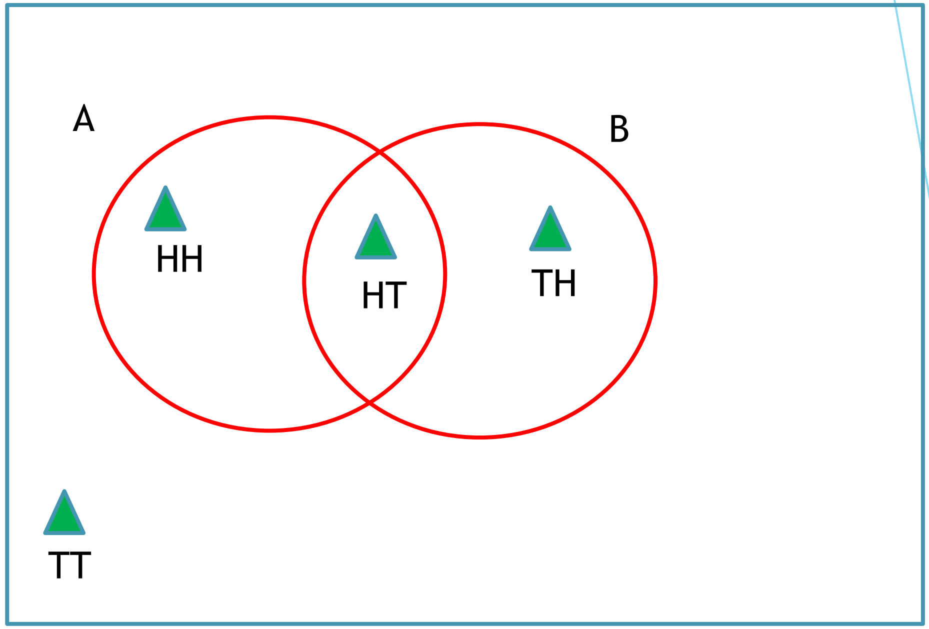
Venn Diagram Visualization

- Experiment – Toss 2 Coins
- Sample Space {HH, HT, TH, TT}
- A – Heads on 1st Coin { HH, HT}
- B – 1 Head and 1 tail {HT, TH}



Venn Diagram Visualization

- Experiment – Toss 2 Coins
- Sample Space {HH, HT, TH, TT}
- A – Heads on 1st Coin { HH, HT}
- B – 1 Head and 1 tail {HT, TH}



UNION – Either A or B
$A \cup B = \{ HH, HT, TH\}$

INTERSECTION – Both A and B
$A \cap B = \{HT\}$

COMPLEMENT – event does not occur
A' or $\overline{A}$ or $A^c = \{ TT, TH\}$
B' or $\overline{B}$ or $B^c = \{ HH, TT\}$

$$\overline{A \cup B} = ?$$

$$A \cup \overline{A} = ?$$

$$B \cup \overline{B} = ?$$

Properties of Outcome

- The outcomes of a random experiment must be
  - Exhaustive – All possible outcomes must be included.
    - Die roll ~~{1, 2, 3, 4, 5}~~
    - Die roll {1, 2, 3, 4, 5, 6}
  - Mutually Exclusive – No two outcomes can occur at the same time
    - Cannot roll 2 numbers at the same time on a die
    - A person cannot be both Male and Pregnant

Definition of Probability

- Probability of given event A
  - Denoted as P(A)
  - $P(A) = \frac{n(A)}{n(S)}$
  - $n(A)$ = Total number of outcomes in an event
  - $n(S)$ = Total number of ALL outcomes in an event

Probability Example

- Probability of Rolling 5
  - $n(A) = 1$
    - *number of times you can get a 5 when you roll a die once*
  - $n(S) = 6$
    - *number of all possible outcomes*
  - $P(A) = \frac{n(A)}{n(S)} = \frac{1}{6}$

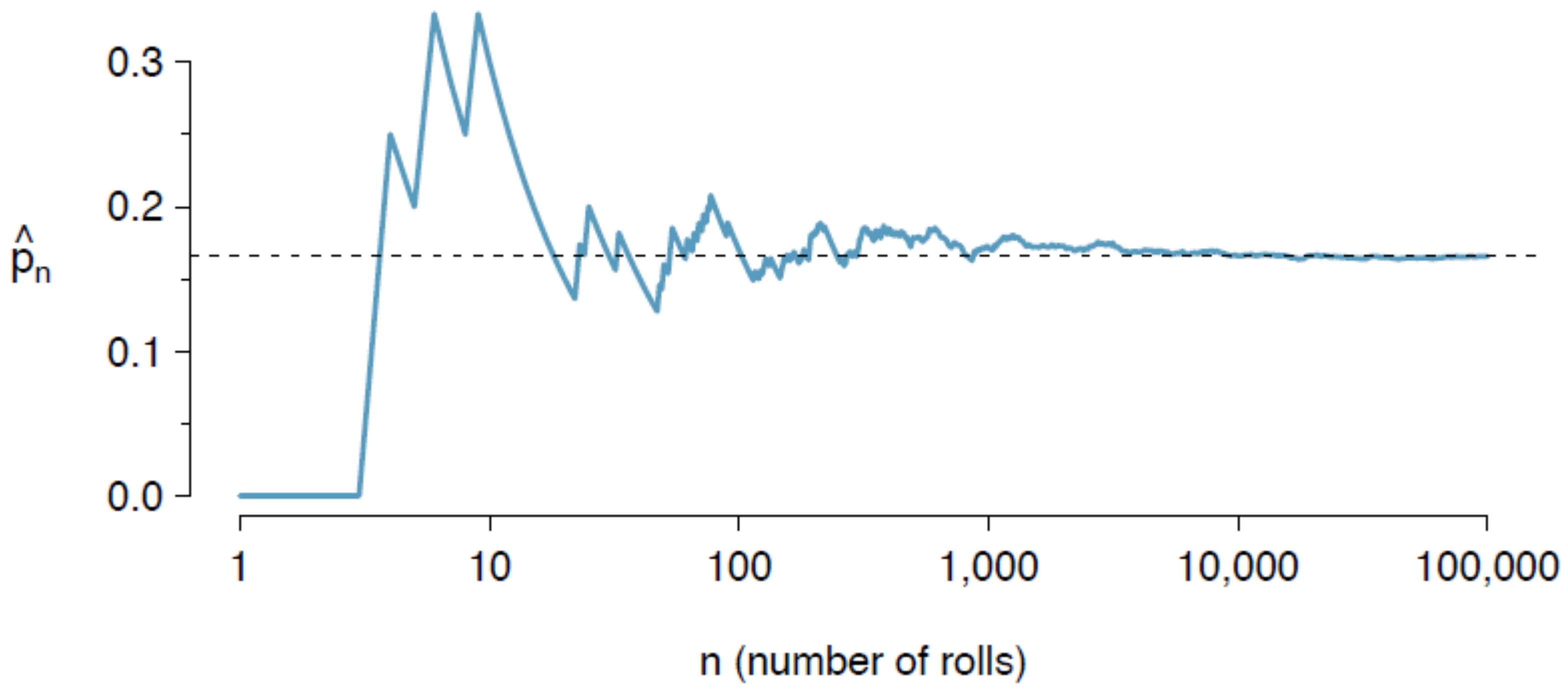
Properties of Probability

- Probability is the likelihood that an event will occur
- Probability of an event must be between 0 and 1 (inclusive)
  - $0 \leq P(A) \leq 1$
- Sum of probabilities of all mutually exclusive and collective events is 1
  - $P(A) + P(B) + P(C) = 1$

Law of Large Numbers

- $\widehat{p_n}$ = proportion of outcomes that are 1 after first n rolls of die
- As n increases, $\widehat{p_n}$ will converge to the probability of rolling a 1 i.e. p = 1/6
- When the experiment is repeated a very large number of times, the proportion of occurrences with a particular outcome converges to the probability p of that outcome (OR) the average over a large number of trials will converge to the expected value as the number of trials get large

Law of Large Numbers



Disjoint Outcomes

- Cannot happen at the same time
- A coin toss cannot yield both head and tail at the same time
- How to calculate probability?

- **ADDITION RULE**
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- For disjoint outcomes $P(A \cap B) = 0$
- $P(A \cup B) = P(A) + P(B)$

- $P(H \text{ or } T) = P(H) + P(T)$
- $P(1 \text{ or } 2 \text{ on a die}) = P(1) + P(2)$

Outcomes not Disjoint

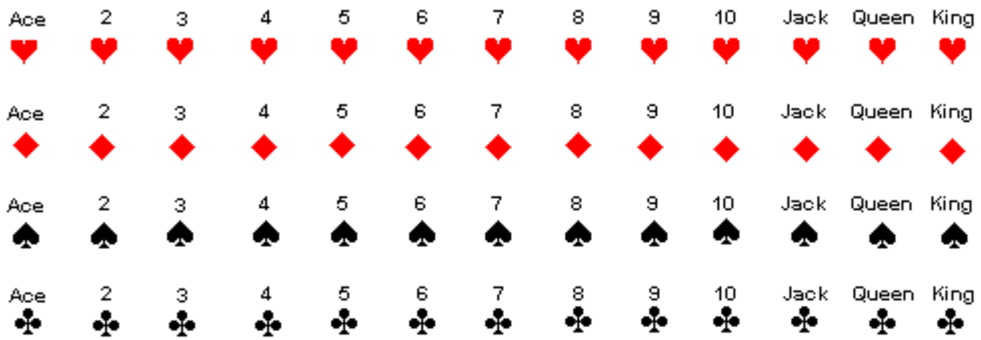
- Can happen at the same time
- Need to account for double counting
- Probability of drawing a heart card or a face card
- How to calculate probability?

- **GENERAL ADDITION RULE**

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cap B) > 0$
- $P(A \cup B) < P(A) + P(B)$

Outcomes not Disjoint

- Example
- Probability of drawing a heart or a face card



Outcomes not Disjoint

- Example
- Probability of drawing a red card or a jack from a standard deck of cards

Probability Distributions

- A table/ graph of all disjoint outcomes and their associated probabilities
- Rules
  - Outcomes must be disjoint
  - Probability of each outcome should be between 0 and 1
  - Sum of probabilities of all outcomes should be 1

Probability Distributions

	Household Incomes (by %)			
Data Set	0-25	26-50	51-100	101+
a.	.18	.34	.33	.16
b.	.38	$\boxed{-.27}$	.52	.37
c.	.28	.27	.27	.16

Probability Distributions

- Sum of two dice

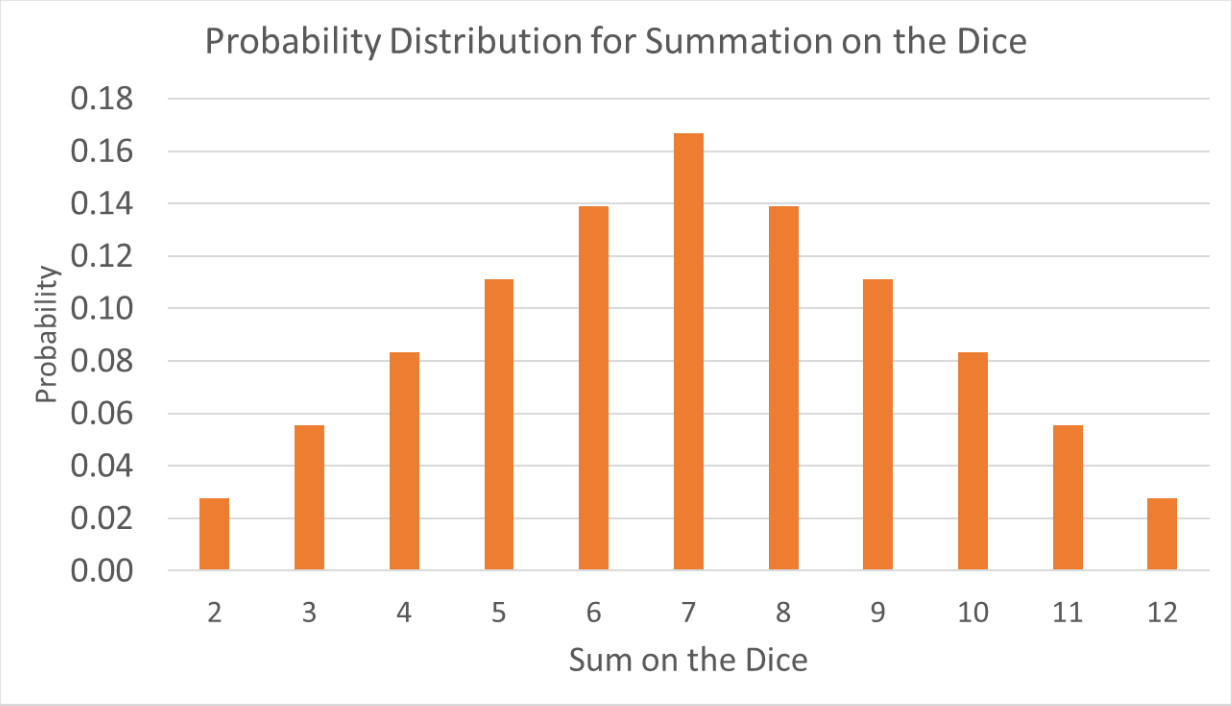
Sum	Probability
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Probability Distributions

- Sum of two dice

Sum	Probability
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03



Independence

- Two events are independent if the outcome of one does not help determine the other
- Example
  - Probability of getting 2 5’s when two die are rolled

Multiplication Rule

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

Example

- About 9% of the people are left-handed. Suppose five people are selected at random.

a) What is the probability that all are right-handed?

b) What is the probability that all are left-handed?

c) What is the probability that not all of the people are right-handed?

Example

- Suppose whether you are right or left-handed is independent of gender i.e. knowing someone's gender does not provide useful information about their handedness and vice-versa. The proportion of the US population that is female is 50%. Three people are selected at random.

a) What is the probability that the first person is male and right-handed?

b) What is the probability that the first two people are male and right-handed?

c) What is the probability that the third person is female and left-handed?

Conditional Probability

- Probability of one event, given that another event has occurred.

$$\boxed{P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}} \rightarrow$$

The conditional probability of A given that B has occurred

$$\boxed{P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}} \rightarrow$$

The conditional probability of B given that A has occurred

P(A and B) = joint probability of A and B = $P(A \cap B)$

P(A) = marginal probability of A

P(B) = marginal probability of B

Conditional Probability

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both.

- Given a car has AC, what is the probability that it also has a CD player.
- P(CD|AC)

Conditional Probability

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both.

	CD	No CD	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

Conditional Probability

A = roll a pair of 5
B = sum =10,11,12

P(A) = ?
P(B) = ?
P(A|B) = ?
P(B|A) = ?

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Conditional Probability

A dataset contains a sample of 792 cases with two variables teen and parents as shown in the table below. The teen variable denotes if a teenager went to college immediately after high school. The parents variable takes the value degree if at least one parent of the teenager completed a college degree.

		Parents		
		Degree	No degree	Total
teen	College	231	214	445
	No College	49	298	347
	Total	280	512	792

Conditional Probability Example

		Parents		
		Degree	No degree	Total
teen	College	231	214	445
	No College	49	298	347
	Total	280	512	792

- If at least one parent of a teenager completed college degree, what is the chance that the teenager attended college right after high school?
- A teenager is randomly selected from the sample and she did not attend college right after high school. What is the probability that at least one of her parents has a college degree?
- Probability of a teen attending college?
- Probability of a teen attending college and parents did not?

Sampling with Replacement

- Put back what you just drew
- Example
  - A bag with 5 red, 3 blue and 2 orange candies
  - Probability of the first candy to be drawn be blue
  - P(Blue candy) = $\frac{3}{5+3+2} = 0.3$
  - If the first candy is blue, what is the probability of second to be blue if sampling with replacement?
    - P(2^{nd} blue candy | 1^{st} is blue) = $\frac{3}{5+3+2} = 0.3$

Sampling with Replacement

- If you drew an orange candy in the first draw, what is the probability of drawing a blue candy in the second draw?

- If drawing with replacement, what is the probability of drawing two red candies in a row?

Sampling with Replacement

- When drawing with replacement, the draws are independent.

Sampling without Replacement

- Do NOT put back what you just drew
- Example
  - A bag with 5 red, 3 blue and 2 orange candies
  - Probability of the first candy to be drawn be blue
  - P(Blue candy) = $\frac{3}{5+3+2} = 0.3$
  - If the first candy is blue, what is the probability of second to be blue if sampling with replacement?
    - P(2^{nd} blue candy | 1^{st} is blue) = $\frac{2}{9} = 0.22$

Sampling without Replacement

- When drawing with replacement, the draws are NOT independent.
- Important especially for small sample sizes.

Counting Rule #1

- To count the number of possible outcomes

If any one of k different mutually exclusive and collectively exhaustive events can occur on each of n trials, the number of possible outcomes is equal to

$$k^n$$

Counting Rule #2

- To count the number of possible outcomes

If there are k_1 events on the first trial, k_2 events on the second trial, … and k_n events on the n^{th} trial, the number of possible outcomes is

$$(k_1)(k_2)\ldots(k_n)$$

Counting Rule #2

Example:

You want to go to a park, eat at a restaurant, and see a movie. There are 3 parks, 4 restaurants, and 6 movie choices. How many different possible combinations are there?

Answer: (3)(4)(6) = 72 different possibilities

Counting Rule #3

- Arranging items in order

The number of ways that n items can be arranged in order is

$$n! = (n)(n - 1)(n-2)(n-3)\ldots(1)$$

Counting Rule #3

- Arranging items in order

Example:

Your restaurant has five menu choices for lunch. How many ways can you order them on your menu?

Answer: 5! = (5)(4)(3)(2)(1) = 120 different possibilities

Answer: 5! = (5)(4)(3)(2)(1) = 120 different possibilities

Counting Rule #4

- Permutations- Arranging x items selected from n items in order

$$\boxed{{}_{n}P_{x} = \frac{n!}{(n - X)!}}$$

Counting Rule #4

Example:

Your restaurant has five menu choices, and three are selected for daily specials. How many different ways can the specials menu be ordered?

$$nPx = \frac{n!}{(n-X)!} = \frac{5!}{(5-3)!} = \frac{120}{2} = 60$$

Counting Rule #5

- Combinations - Selecting x items selected from n items irrespective of the order

$$\boxed{{}_nC_x = \frac{n!}{X!(n-X)!}}$$

Counting Rule #5

Example:

Your restaurant has five menu choices, and three are selected for daily specials. How many different special combinations are there, ignoring the order in which they are selected?

$$ {}_{n}C_{x} = \frac{n!}{X!(n-X)!} = \frac{5!}{3!(5-3)!} = \frac{120}{(6)(2)} = 10 $$

Tree Diagram

- Tool to organize outcomes and probabilities
- Useful when processes occur in a sequence and each process is conditioned on its predecessor

Bayes Theorem

- Converting unknown conditional probability P(A|B) to one involving known conditional probability P(B|A)
- Generalization of the tree diagram.
- P(outcome A1 of variable 1 | outcome B of variable 2)

$$\frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_k)P(A_k)} \quad (2.55)$$

where A_2, A_3, ..., and A_k represent all other possible outcomes of the first variable.

Numerator =probability of both A1 and B happening

Denominator = marginal probability of B

Tree Diagram

- Consider a midterm and final for a statistics class. 13% of students earned an A on the midterm. Of those students who earned an A on the midterm, 47% earned an A on the final. 11% of the students who earned lower than an A on the midterm received an A on the final. You randomly pick a final exam and notice the student received an A. What is the probability that this student earned an A on the midterm?

Tree Diagram

- 0.5 percent of the population of an area is affected by a particular disease. A test is developed to detect the disease. This test gives a false positive 3% of the time and a false negative 2% of the time.
  - What is the probability that the test gives a positive result?
  - If a person's test turns out to be positive, what is the probability that he actually has the disease?

Example

- In Canada, 0.35% of women over 40 will develop breast cancer in any year. A common screening test for cancer is the mammogram but this is not perfect. In about 11% of the patients with breast cancer, the test gives a false negative which indicates a woman does not have breast cancer when she does have breast cancer. Similarly the test gives a false positive in 7% of the patients who do not have breast cancer which indicates these patients have breast cancer when they actually do not. If we tested a random woman over 40 for breast cancer using a mammogram and the test came back positive what is the probability that the patient actually has breast cancer?