Chapter 1: Probability

We see probabilities almost every day in our real lives. Most times you pick up the newspaper or read the news on the internet, you encounter probability. There is a 65% chance of rain today, or a pre-election poll shows that 52% of voters approve of a ballot

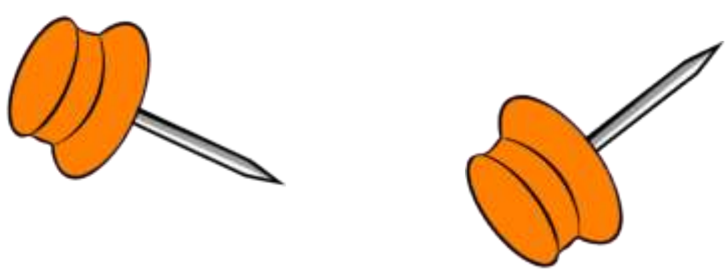
of the idea's students think they know about probability are incorrect. This is one area of item, are examples of probabilities. Did you ever wonder why a flush beat a full house in poker? It's because the probability of getting a flush is smaller than the probability of getting a full house. Probabilities can also be used to make business decisions, figure out insurance premiums, and determine the price of raffle tickets.

If an experiment has only three possible outcomes, does this mean that each outcome has a 1/3 chance of occurring? Many students who have not studied probability would answer yes. Unfortunately, they could be wrong. The answer depends on the experiment. Many math where their intuition is sometimes misleading. Students need to use experiments or mathematical formulas to calculate probabilities correctly.

Section 1.1: Basic Probabilities and Probability Distributions; Three Ways to Define Probabilities

Toss a thumb tack one time. Do you think the tack will land with the point up or the point down?

Figure 1.1.1: Which Way Will a Tack Fall?



(Thumbtack, n.d.)

We cannot predict which way the tack will land before we toss it. Sometimes it will land with the point up and other times it will land with the point down. Tossing a tack is a random experiment since we cannot predict what the outcome will be. We do know that there are only two possible outcomes for each trial of the experiment: lands point up or lands point down. If we repeat the experiment of tossing the tack many times, we might be able to guess how likely it is that the tack will land point up.

able to guess how likely it is that the tack will land point up.

A **random experiment** is an activity or an observation whose outcome cannot be predicted ahead of time

A **trial** is one repetition of a random experiment.

The **sample space** is the set of all possible outcomes for a random experiment.

An **event** is a subset of the sample space.

Do you think chances of the tack landing point up and the tack landing point down are the same? This is an example where your intuition may be wrong. Having only two possible outcomes does not mean each outcome has a 50/50 chance of happening. In fact, we are going to see that the probability of the tack landing points up is about 66%.

To begin to answer this question, toss a tack 10 times. For each toss record whether the tack lands point up or point down.

Table 1.1.2: Toss a Tack Ten Times

Trial	1	2	3	4	5	6	7	8	9	10
Up/Down	Up	Down	Up	Up	Down	Up	Up	Down	Up	Down

The random experiment here is tossing the tack one time. The possible outcomes for the experiment are that the tack lands point up or the tack lands point down so the sample space is S = {point up, point down}. We are interested in the event E that the tack lands point up, E = {point up}.

Based on our data we would say that the tack landed point up six out of 10 times. The fraction, $\frac{6}{10}$, is called the relative frequency. Since $\frac{6}{10} = 0.60$ we would guess that probability of the tack landing points up is about 60%.

Let's repeat the experiment by tossing the tack ten more times.

Table 1.1.3: Toss a Tack Ten More Times

Trial	1	2	3	4	5	6	7	8	9	10
Up/Down	Up	Up	Up	Down	Down	Up	Up	Down	Up	Up

This time the tack landed point up seven out of 10 times or 70% of the time. If we tossed the tack another 10 times, we might get a different result again. The probability of the tack landing point up refers to what happens when we toss the tack many, many times. Let's toss the tack 150 times and count the number of times it lands point up. Along the way we will look at the proportion of landing point up.

Table 1.1.4: Toss a Tack Many Times

Trials	Number Up	Total Number Of Up	Total Number Of Trials	Proportion
0-10	6	6	10	6/10=0.60
11-20	7	13	20	13/20=0.65
21-30	8	21	30	21/30=0.70
31-40	6	27	40	27/40=0.68
41-50	8	35	50	35/50=0.70
51-60	8	43	60	43/60=0.72
61-70	4	47	70	47/70=0.67
71-80	7	54	80	54/80=0.68
81-90	5	59	90	59/90=0.66
91-100	6	65	100	65/100=0.65
101-110	10	75	110	75/110=0.68
111-120	5	80	120	80/120=0.67
121-130	8	88	130	88/130=0.68
131-140	4	92	140	92/140=0.66
141-150	7	99	150	99/150=0.66

When we have a small number of trials the proportion varies quite a bit. As we start to have more trials the proportion still varies but not by as much. It appears that the proportion is around 0.66 or 66%. We would have to do about 100,000 trials to get a better approximation of the actual probability of the tack landing point up.

The tack landed point up 99 out of 150 trials. The probability P of event E is written as:

$$P(E) = \frac{\text{\# of trials with point up}}{\text{total number of trials}} = \frac{99}{150} \approx 0.66$$

We would say the probability that the tack lands point up is about 66%.

Equally Likely Outcomes:

In some experiments all the outcomes have the same chance of happening. If we roll a fair die the chances are the same for rolling a two or rolling a five. If we draw a single card from a well shuffled deck of cards, each card has the same chance of being selected. We call outcomes like these equally likely. Drawing names from a hat or drawing straws are other examples of equally likely outcomes. The tack tossing example did not have equally likely outcomes since the probability of the tack landing point up is different than the probability of the tack landing point down.

An experiment has **equally likely outcomes** if every outcome has the same probability of occurring.

For equally likely outcomes, the **probability of outcome A**, *P(A)*, is:

$$P(A) = \frac{\text{number of ways for A to occur}}{\text{total number of outcomes}}.$$

Round Off Rule: Give probabilities as a fraction or as a decimal number rounded to three decimal places.

Figure 1.1.5: Deck of Cards

(Pine, 2007)

Example 1.1.1: Simple Probabilities with Cards

Draw a single card from a well shuffled deck of 52 cards. Each card has the same chance of being drawn so we have equally likely outcomes. Find the following probabilities:

a. P(card is red)

$$P(\text{card is red}) = \frac{\text{number of red cards}}{\text{total number of cards}} = \frac{26}{52} = \frac{1}{2}$$

The probability that the card is red is $\frac{1}{2}$.

b. P(card is a heart)

$$P(\text{card is a heart}) = \frac{\text{number of hearts}}{\text{total number of cards}} = \frac{13}{52} = \frac{1}{4}$$

The probability that the card is a heart is $\frac{1}{4}$.

c. P(card is a red 5)

$$P(\text{card is a red 5}) = \frac{\text{number of red fives}}{\text{total number of cards}} = \frac{2}{52} = \frac{1}{26}$$

The probability that the card is a red five is $\frac{1}{26}$.

Example 1.1.2: Simple Probabilities with a Fair Die

Roll a fair die one time. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Find the following probabilities.

a. P(roll a four)

$$P(\text{roll a four}) = \frac{\text{number of ways to roll a four}}{\text{total number of ways to roll a die}} = \frac{1}{6}$$

The probability of rolling a four is $\frac{1}{6}$.

b. P(roll an odd number)

The event roll an odd number is $E = \{1, 3, 5\}$.

$$P(\text{roll an odd number}) = \frac{\text{number of ways to roll an odd number}}{\text{total number of ways to roll a die}} = \frac{3}{6} = \frac{1}{2}$$

The probability of rolling an odd number is $\frac{1}{2}$.

c. P(roll a number less than five)

The event roll a number less than five is F = {1, 2, 3, 4}.

$$P(\text{roll a number less than five}) = \frac{\text{number of ways to roll number less than five}}{\text{total number of ways to roll a die}} = \frac{4}{6} = \frac{2}{3}$$

The probability of rolling a number less than five is $\frac{2}{3}$.

Example 1.1.3: Simple Probability with Books

Figure 1.1.6: Books on a Shelf

(Bookshelf, 2011)

A small bookcase contains five math, three English and seven science books. A book is chosen at random. What is the probability that a math book is chosen?

Since the book is chosen at random each book has the same chance of being chosen and we have equally likely events.

$$P(\text{math book}) = \frac{\text{number of ways to choose a math book}}{\text{total number of books}} = \frac{5}{15} = \frac{1}{3}$$

The probability a math book was chosen is $\frac{1}{3}$.

Three Ways of Finding Probabilities:

There are three ways to find probabilities. In the tack tossing example we calculated the probability of the tack landing point up by doing an experiment and recording the outcomes. This was an example of an empirical probability. The probability of getting a red jack in a card game or rolling a five with a fair die can be calculated from mathematical formulas. These are examples of theoretical probabilities. The third type of probability is a subjective probability. Saying that there is an 80% chance that you will go to the beach this weekend is a subjective probability. It is based on experience or guessing.

A **theoretical probability** is based on a mathematical model where all outcomes are equally likely to occur.

An **empirical probability** is based on an experiment or observation and is the relative frequency of the event occurring.

A **subjective probability** is an estimate (a guess) based on experience or intuition.

Complements:

If there is a 75% chance of rain today, what are the chances it will not rain? We know that there are only two possibilities. It will either rain or it will not rain. Because the sum of the probabilities for all the outcomes in the sample space must be 100% or 1.00, we know that

P(will rain) + P(will not rain) = 100%.

Rearranging this we see that

P(will not rain) = 100% - P(will rain) = 100% - 75% = 25%.

The events E = {will rain} and F = {will not rain} are called complements.

The **complement** of event E, denoted by $\bar{E}$, is the set of outcomes in the sample space that are not in the event E. The probability of $\bar{E}$ is given by $P\left(\bar{E}\right) = 1 - P\left(E\right)$.

Example 1.1.4: Complements with Cards

Draw a single card from a well shuffled deck of 52 cards.

a. Look at the suit of the card. Here the sample space S = {spades, clubs, hearts, diamonds}. If event E = {spades} the complement $\bar{E}$ = {clubs, hearts, diamonds}.

b. Look at the value of the cards. Here the sample space is S = {A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K}. If the event E = {the number is less than 7} = {A, 2, 3, 4, 5, 6} the complement $\bar{E}$ = {7, 8, 9, 10, J, Q, K}.

Example 1.1.5: Complements with Trains

A train arrives on time at a particular station 85% of the time. Does this mean that the train is late 15% of the time? The answer is no. The complement of E = {on time} is not $\bar{E}$ = {late}. There is a third possibility. The train could be early. The sample space is S = {on time, early, late} so the complement of E = {on time} is $\bar{E}$ = {early or late}. Based on the given information we cannot find P(late) but we can find P(early or late) = 15%.

Impossible Events and Certain Events:

Recall that $P(A) = \dfrac{\text{number of ways for A to occur}}{\text{total number of outcomes}}$. What does it mean if we say the probability of the event is zero? $P(A) = \dfrac{\text{number of ways for A to occur}}{\text{total number of outcomes}} = 0$. The only way for a fraction to equal zero is when the numerator is zero. This means there is no way for event A to occur. A probability of zero means that the event is impossible.

What does it mean if we say the probability of an event is one?
$P(A) = \dfrac{\text{number of ways for A to occur}}{\text{total number of outcomes}} = 1$. The only way for a fraction to equal one is if the numerator and denominator are the same. The number of ways for A to occur is the same as the number of outcomes. There are no outcomes where A does not occur. A probability of 1 means that the event always happens.

$P(A) = 0$ means that A is impossible
$P(A) = 1$ means that A is certain

Probability Distributions:

A probability distribution (probability space) is a sample space paired with the probabilities for each outcome in the sample space. If we toss a fair coin and see which side lands up, there are two outcomes, heads and tails. Since the coin is fair these are equally likely outcomes and have the same probabilities. The probability distribution would be $P(\text{heads}) = 1/2$ and $P(\text{tails}) = 1/2$. This is often written in table form:

Table 1.1.7: Probability Distribution for a Fair Coin

Outcome	Heads	Tails
Probability	1/2	1/2

A **probability distribution** for an experiment is a list of all the possible outcomes and their corresponding probabilities.

Example 1.1.6: Probabilities for the Sum of Two Fair Dice

In probability problems when we roll two dice, it is helpful to think of the dice as being different colors. Let's assume that one die is red and the other die is green. We consider getting a three on the red die and a five on the green die different than getting a five on the red die and a three on the green die. In other words, when we list the outcomes the order matters. The possible outcomes of rolling two dice and looking at the sum are given in Table 3.1.8.

Table 1.1.8: All Possible Sums of Two Dice

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

Table 1.1.9: Probability Distribution for the Sum of Two Fair Dice

Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
Reduced Probability	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Example 1.1.7: Valid and Invalid Probability Distributions

Are the following valid probability distributions?

a. **Table 1.1.10:**

Outcome	**A**	**B**	**C**	**D**	**E**
Probability	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

This is a valid probability distribution. All the probabilities are between zero and one inclusive and the sum of the probabilities is 1.00.

b. **Table 1.1.11:**

Outcome	**A**	**B**	**C**	**D**	**E**	**F**
Probability	0.45	0.80	-0.20	-0.35	0.10	0.20

This is not a valid probability distribution. The sum of the probabilities is 1.00, but some of the probabilities are not between zero and one, inclusive.

c. **Table 1.1.12:**

Outcome	**A**	**B**	**C**	**D**
Probability	0.30	0.20	0.40	0.25

This is not a valid probability distribution. All of the probabilities are between zero and one, inclusive, but the sum of the probabilities is 1.15 not 1.00.

Odds:

Probabilities are always numbers between zero and one. Many people are not comfortable working with such small values. Another way of describing the likelihood of an event happening is to use the ratio of how often it happens to how often it does not happen. The ratio is called the odds of the event happening. There are two types of odds, odds for and odds against. Casinos, race tracks and other types of gambling usually state the odds against an event happening.

If the probability of an event *E* is *P(E)*, then the **odds for event *E***, *O(E)*, are given by:

$$O(E)=\frac{P(E)}{P(\bar{E})} \qquad \text{OR} \qquad O(E)=\frac{\text{number of ways for E to occur}}{\text{number of ways for E to not occur}}$$

Also, the **odds against event *E***, are given by:

$$O(\bar{E})=\frac{P(\bar{E})}{P(E)} \qquad \text{OR} \qquad O(E)=\frac{\text{number of ways for E to not occur}}{\text{number of ways for E to occur}}$$

Example 13.1.8: Odds in Drawing a Card

A single card is drawn from a well shuffled deck of 52 cards. Find the odds that the card is a red eight.

There are two red eights in the deck.

$$P(\text{red eight}) = \frac{2}{52} = \frac{1}{26}$$

$$P(\text{not a red eight}) = \frac{50}{52} = \frac{25}{26}.$$

$$O(\text{red eight}) = \frac{P(\text{red eight})}{P(\text{not a red eight})} = \frac{\frac{1}{26}}{\frac{25}{26}} = \frac{1}{\cancel{26}} \cdot \frac{\cancel{26}}{25} = \frac{1}{25}.$$

The odds of drawing a red eight are 1 to 25. This can also be written as 1:25.

Note: Do not write odds as a decimal or a percent.

Example 1.1.9: Odds in Roulette

Many roulette wheels have slots numbered 0, 00, and 1 through 36. The slots numbered 0 and 00 are green. The even numbered slots are red and the odd numbered slots are black. The game is played by spinning the wheel one direction and rolling a marble around the outer edge the other direction. Players bet on which slot the marble will fall into. What are the odds the marble will land in a red slot?

There are 38 slots in all. The slots 2, 4, 6, …, 36 are red so there are 18 red slots. The other 20 slots are not red.

$$P(\text{red}) = \frac{18}{38} = \frac{9}{19}$$

$$P(\text{not red}) = 1 - \frac{9}{19} = \frac{19}{19} - \frac{9}{19} = \frac{10}{19}$$

$$O(\text{red}) = \frac{P(\text{red})}{P(\text{not red})} = \frac{\frac{9}{19}}{\frac{10}{19}} = \frac{9}{\cancel{19}} \cdot \frac{\cancel{19}}{10} = \frac{9}{10}$$

The odds of the marble landing in a red slot are 9 to 10. This can also be written as 9:10.

Example 1.1.10: Odds Against an Event

Two fair dice are tossed and the sum is recorded. Find the odds against rolling a sum of nine.

The event *E*, roll a sum of nine is: *E* = {(3, 6), (4, 5), (5, 4), (6, 3)}

There are 36 ways to roll two dice and four ways to roll a sum of nine. That means there are 32 ways to roll a sum that is not nine.

$$P(\text{sum is nine}) = \frac{4}{36} = \frac{1}{9}$$

$$P(\text{sum is not nine}) = \frac{32}{36} = \frac{8}{9}$$

$$O(\text{against sum is nine}) = \frac{P(\text{sum is not nine})}{P(\text{sum is nine})} = \frac{\frac{8}{9}}{\frac{1}{9}} = \frac{8}{\cancel{9}} \cdot \frac{\cancel{9}}{1} = \frac{8}{1}$$

The odds against rolling a sum of nine are 8 to 1 or 8:1.

We can also find the probability of an event happening based on the odds for the event. Saying that the odds of an event are 3 to 5 means that the event happens three times for every five times it does not happen. If we add up the possibilities of both we get a sum of eight. So the event happens about three out of every eight times. We would say the probability is 3/8.

If the odds favoring event *E* are *a* to *b*, then:

$P(E) = \dfrac{a}{a+b}$ and $P(\bar{E}) = \dfrac{b}{a+b}$.

Example 1.1.11: Finding the Probability from the Odds

A local little league baseball team is going to a tournament. The odds of the team winning the tournament are 3 to 7. Find the probability of the team winning the tournament.

$$P(\text{winning}) = \frac{3}{3+7} = \frac{3}{10} = 0.3$$

Set Definitions

- A set can be defined as a collection of objects. Sets are generally denoted by capital letters as: A, B, C, ...
- The individual objects forming the set are called "**elements**" or "**members**". They are generally denoted by lower case letters as: a,b, c,...
- If an element g belongs to a set G, we write:

$$g \in G \tag{1}$$

Otherwise, we say g is not a member of G, we write:

$$g \notin G \tag{2}$$

- A set is specified by the content of two braces: $\{\cdot\}$.
- **Representation of sets:**
  - Tabular method: the elements are enumerated explicitly. For example:

$A = \{3,4,5,6\}$.

  - Rule method: the content of the set is specified using a rule. This representation is more convenient when the set is large. For example:

$$G = \{g \mid g \text{ is an integer and } 3 \leq g \leq 6\} \tag{3}$$

Such that

- **Countable and uncountable sets:** A set is called to be **"countable"** if its elements can be put in one-to-one correspondence with the integers 1,2,..etc.Otherwise, it is called "**uncountable"**.
- **Empty set:** A set G is said to be empty, if it has no elements. It is also called null set and it is denoted by $\emptyset$.
- **Finite and infinite sets:** A finite set is either empty set or has elements that can be counted, with the counting process terminating. If a set is not finite it is called infinite.
- **Subset:** Given two sets A and B, if every element of A is also an element of B, A is said to be contained in A. A is known as a subset of B. We write:

$$A \subseteq B \tag{4}$$

is contained

- **Proper subset:** If at least one element in $B$is not in A, then A is a proper subset of B, denoted by

$$A \subset B \tag{5}$$

- **Disjoint sets:** If two sets A and B have no common elements, then they are called disjoint or mutually exclusive.

Example 1: Let us consider the following four sets:

$A = \{ 1,3,5,7\}$ $\qquad$ $D = \{ 0 \}$

$B = \{ 1,2,3, \dots .\}$ $\qquad$ $E = \{ 2,4,6,8,10,12,14\}$

$C = \{c|c \text{ } is \text{ } real \text{ } and \text{ } 0.5 < c \leq 8.5\}$ $\qquad$ $F = \{f|f \text{ } is \text{ } real \text{ } and -5 < f \leq 12\}$

Illustrate the previous concepts using the sets A, B, C, D, E, F.

Set A	Relation between set A and set B	Set B
Tabular or rule countable or uncountable Finite or infinite Empty	(subset, proper subset, mutually exclusive) ⟺	Tabular or rule Finite or infinite countable or uncountable Empty

Solution:

- **The set A is tabularly specified, countable, and finite.**
- **Set A is contained in sets B, C and F.**
- **The set B is tabularly specified and countable, but is infinite.**
- **Set C is rule-specified, uncountable, and infinite.**

- **Sets D and E are countably finite.**
- **Set F is uncountably infinite.**
- $\boldsymbol{C \subset F,\ D \subset F,}$ **E** $\boldsymbol{\subset B.}$
- **Sets *B* and *F* are not sub sets of any of the other sets or of each other.**
- **Sets *A*, *D* and *E* are mutually exclusive of each other.**

Universal set: The set of all elements under consideration is called the universal set, denoted S. All sets (of the situation considered) are subsets of S.

If we have a set S with n elements, then there are 2^n subsets.

In case of rolling die, the universal set is $S = \{1,2,3,4,5,6\}$ and the number of subsets is 2^6=64 subsets.

Example 2: Determine the subsets of the following universal set $S = \{ 1,2,3,4\}$

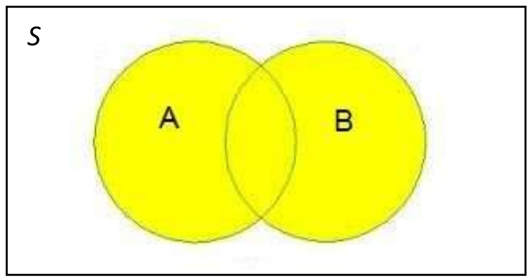
Solution: The universal set is $\boldsymbol{S = \{1, 2, 3, 4\}}$ and the number of subsets is $\boldsymbol{2^4}$=16 subsets.

1	$\emptyset$	9	{2,3}
2	{1}	10	{2,4}
3	{2}	11	{3,4}
4	{3}	12	{1,2,3}
5	{4}	13	{1, 3,4}
6	{1,2}	14	{1,2,4}
7	{1,3}	15	{2,3,4}
8	{1,4}	16	{1,2,3,4}

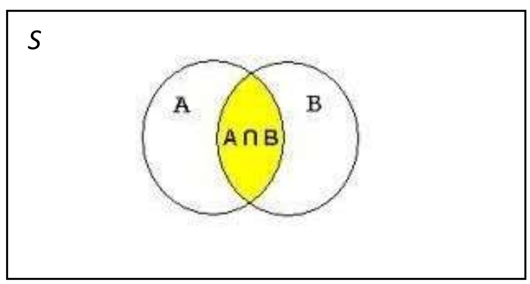
Set Operations

- **Venn diagram:** is a graphical representation of sets to help visualize sets and their operations.

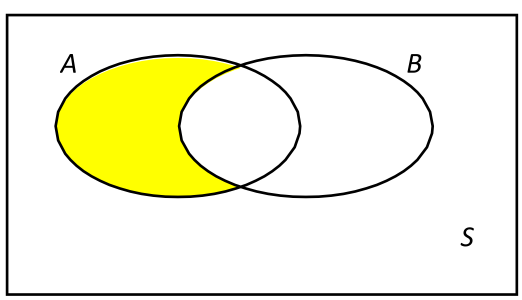
- **Union:** set of all elements that are members of A or B or both and is denoted by $A \cup B$.



- **Intersection:** set of all elements which belong to both $A$and B and is denoted by $A \cap B$.

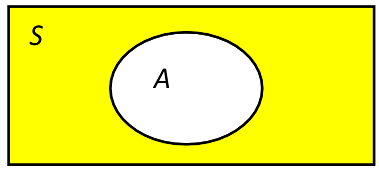


- **Difference:** Set consisting of all elements in A which are not in B and is denoted as $A - B$



- **Complement:** The set composed of all members in $S$and not in A is the complement of A and denoted A^C. Thus

$$A^C = S - A \tag{6}$$



It is easy to see that $\emptyset^C = S,\ S^C = \emptyset,\ A \cup A^C = S,\ \ and\ A \cap A^C = \emptyset$

Example 3: Let us illustrate these concepts on the following four sets

$S = \{\ a \mid a$ is an integer and $1 < a \leq 12\}$

$A = \{1,3,5,12\}$

$B = \{2,6,7,8,9,10,11\}$

$C = \{1,3,4,6,7,8\}$

Solution:

- **Unions and intersections**

$\boldsymbol{A \cup B = \{1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12\}}$ $\boldsymbol{A \cap B = \emptyset}$

$\boldsymbol{A \cup C = \{1, 3, 4, 5, 6, 7, 8, 12\}}$ $\boldsymbol{A \cap C = \{1, 3\ \}}$

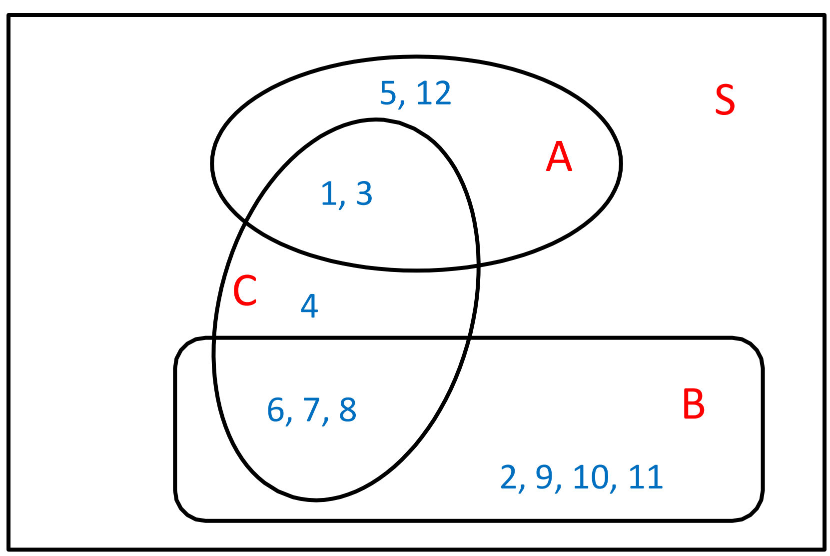
$\boldsymbol{B \cup C = \{1, 2, 3, 4, 6, 7, 8, 9, 10, 11\}}$ $\boldsymbol{B \cap C = \{6, 7, 8\}}$

- **Complements**

$\boldsymbol{A^C = \{2, 4, 6, 7, 8, 9, 10, 11\}}$

$\boldsymbol{B^C = \{1, 3, 4, 5, 12\}}$

$\boldsymbol{C^C = \{2, 5, 9, 10, 11, 12\}}$



- **Algebra of sets:**
  - ✓ **Commutative law:** $A \cap B = B \cap A$

    $A \cup B = B \cup A$
  - ✓ **Distributive law:** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

    $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  - ✓ **Associative law:** $(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$

    $(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$
  - ✓ **De Morgan's Law :** $(A \cup B)^C = A^C \cap B^C$

    $(A \cap B)^C = A^C \cup B^C$

Probability definition and axioms

- Let A an event defined on the sample space S. The probability of the event A denoted as $P(A)$ is a function that assigns to A a real number such that:

✓ **Axiom1:** $P(A) \geq 0$ (7)

✓ **Axiom2:** $P(S) = 1$ (8)

Certain event

✓ **Axiom3:** if we have N events $A_n, n = 1, 2, \dots, N$ defined on the sample space S, and having the propriety: $A_m \cap A_n = \emptyset$ for $m \neq n$ (mutually exclusive events). Then:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) \tag{9}$$

Or $$P(\cup_{n=1}^{N} A_n) = \sum_{n=1}^{N} P(A_n) \tag{10}$$

Some Properties:

- For every event A, its probability is between 0 and 1:

$$0 \leq P(A) \leq 1 \tag{11}$$

- The probability of the impossible event is zero

$$P(\emptyset) = 0 \tag{12}$$

- If $\bar{A}$ is the complement of A, then:

$$P(A^C) = 1 - P(A) \tag{13}$$

- **To model a real experiment mathematically, we shall:**
  - Define the sample space.
  - Define the events of interest.
  - Assign probabilities to the events that satisfy the probability axioms.

Section 2: Combining Probabilities with "And" and "Or"

Many probabilities in real life involve more than one outcome. If we draw a single card from a deck, we might want to know the probability that it is either red or a jack. If we look at a group of students, we might want to know the probability that a single student has brown hair and blue eyes. When we combine two outcomes to make a single event, we connect the outcomes with the word "and" or the word "or." It is very important in probability to pay attention to the words "and" and "or" if they appear in a problem. The word "and" restricts the field of possible outcomes to only those outcomes that simultaneously satisfy more than one event. The word “or” broadens the field of possible outcomes to those that satisfy one or more events.

Example 2.1: Counting Students

Figure 2.1: College Classroom

(Colwell, 2013)

Suppose a teacher wants to know the probability that a single student in her class of 30 students is taking either Art or English. She asks the class to raise their hands if they are taking Art and counts 13 hands. Then she asks the class to raise their hands if they are taking English and counts 21 hands. The teacher then calculates

$$P(\text{Art or English}) = \frac{13+21}{30} = \frac{33}{30}.$$

The teacher knows that this is wrong because probabilities must be between zero and one, inclusive. After thinking about it she remembers that nine students are taking both Art and English. These students raised their hands each time she counted, so the teacher counted them twice. When we calculate probabilities, we have to be careful to count each outcome only once.

Mutually Exclusive Events:

An experiment consists of drawing one card from a well shuffled deck of 52 cards. Consider the events *E*: the card is red, *F*: the card is a five, and *G*: the card is a spade. It is possible for a card to be both red and a five at the same time but it is not possible for a card to be both red and a spade at the same time. It would be easy to accidentally count a red five twice by mistake. It is not possible to count a red spade twice.

Two events are **mutually exclusive** if they have no outcomes in common.

Example 2.2: Mutually Exclusive with Dice

Two fair dice are tossed and different events are recorded. Let the events E, F and G be as follows:

E = {the sum is five} = {(1, 4), (2, 3), (3, 2), (4, 1)}

F = {both numbers are even} = {(2, 2), (2, 4), (2, 6), (4, 2), (4, 4), (4, 6), (6, 2), (6, 4), (6, 6)}

G = {both numbers are less than five} = {(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (2, 4), (3, 1), (3, 2), (3, 3), (3, 4), (4,1), (4, 2), (4, 3), (4,4)}

a. Are events E and F mutually exclusive?

   Yes. E and F are mutually exclusive because they have no outcomes in common. It is not possible to add two even numbers to get a sum of five.

b. Are events E and G mutually exclusive?

   No. E and G are not mutually exclusive because they have some outcomes in common. The pairs (1, 4), (2, 3), (3, 2) and (4, 1) all have sums of 5 and both numbers are less than five.

c. Are events F and G mutually exclusive?

   No. F and G are not mutually exclusive because they have some outcomes in common. The pairs (2, 2), (2, 4), (4, 2) and (4, 4) all have two even numbers that are less than five.

Addition Rule for "Or" Probabilities:

The addition rule for probabilities is used when the events are connected by the word "or". Remember our teacher in Example 3.2.1 at the beginning of the section? She wanted to know the probability that her students were taking either art or English. Her problem was that she counted some students twice. She needed to add the number of students taking art to the number of students taking English and then subtract the number of students she counted twice. After dividing the result by the total number of students she will find the desired probability. The calculation is as follows:

$$P(\text{art or English}) = \frac{\text{\# taking art + \# taking English - \# taking both}}{\text{total number of students}}$$
$$= \frac{13+21-9}{30}$$
$$= \frac{25}{30} \approx 0.833$$

The probability that a student is taking art or English is 0.833 or 83.3%.

When we calculate the probability for compound events connected by the word "or" we need to be careful not to count the same thing twice. If we want the probability of drawing a red card or a five, we cannot count the red fives twice. If we want the probability a person is blonde-haired or blue-eyed, we cannot count the blue-eyed blondes twice. The addition rule for probabilities adds the number of blonde-haired people to the number of blue-eyed people then subtracts the number of people we counted twice.

Addition Rule for "Or" Probabilities
If A and B are any events then, $P(A \text{ or } B) = P(A) + P(B) – P(A \text{ and } B)$.
If A and B are mutually exclusive events then $P(A \text{ and } B) = 0$, so then
$P(A \text{ or } B) = P(A) + P(B)$.

Example 2.3: Additional Rule for Drawing Cards

A single card is drawn from a well shuffled deck of 52 cards. Find the probability that the card is a club or a face card.

There are 13 cards that are clubs, 12 face cards (J, Q, K in each suit) and 3 face cards that are clubs.

$$P(\text{club or face card}) = P(\text{club}) + P(\text{face card}) - P(\text{club and face card})$$
$$= \frac{13}{52} + \frac{12}{52} - \frac{3}{52}$$
$$= \frac{22}{52} = \frac{11}{26} \approx 0.423$$

The probability that the card is a club or a face card is approximately 0.423 or 42.3%.

Example 2.4: Addition Rule for Tossing a Coin and Rolling a Die

An experiment consists of tossing a coin then rolling a die. Find the probability that the coin lands head up or the number is five.

Let H represent heads up and T represent tails up. The sample space for this experiment is S = {H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6}.

There are six ways the coin can land heads up, {H1, H2, H3, H4, H5, H6}.

There are two ways the die can land on five, {H5, T5}.

There is one way for the coin to land heads up and the die to land on five, {H5}.

$$\begin{aligned} P(\text{heads or five}) &= P(\text{heads}) + P(\text{five}) - P(\text{both heads and five}) \\ &= \frac{6}{12} + \frac{2}{12} - \frac{1}{12} \\ &= \frac{7}{12} \approx 0.583 \end{aligned}$$

The probability that the coin lands heads up or the number is five is approximately 0.583 or 58.3%.

Example 2.5: Addition Rule for Satisfaction of Car Buyers

Two hundred fifty people who recently purchased a car were questioned and the results are summarized in the following table.

Table 2.2: Satisfaction of Car Buyers

	Satisfied	**Not Satisfied**	**Total**
New Car	**92**	**28**	**120**
Used Car	**83**	**47**	**130**
Total	**175**	**75**	**250**

Find the probability that a person bought a new car or was not satisfied.

$$\begin{aligned} P(\text{new car or not satisfied}) &= P(\text{new car}) + P(\text{not satisfied}) - P(\text{new car and not satisfied}) \\ &= \frac{120}{250} + \frac{75}{250} - \frac{28}{250} = \frac{167}{250} \approx 0.668 \end{aligned}$$

The probability that a person bought a new car or was not satisfied is approximately 0.668 or 66.8%.

Joint and conditional probability

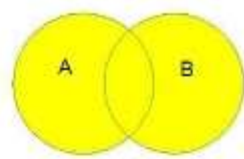
Joint probability

- When two events A and B have some elements in common (not mutually exclusive), then axiom3 cannot be applied.
- The probability $P(A \cap B)$ called the joint probability for the events A and B, which intersect, in sample space.

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) \tag{14}$$

Equivalently:

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



Conditional probability

- Given some event B with nonzero probability $P(B) > 0$
- We defined, the conditional probability of an event A given B, by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{15}$$

- $P(A|B)$ is the probability that A will occur given that B has occurred.
- If the occurrence of event B has no effect on A, we say that A and B are independent events. In this case,

$$P(A|B) = P(A) \tag{16}$$

Which means that:

$$P(A \cap B) = P(A)P(B) \tag{17}$$

Example: Conditional Probability for Drawing Cards without Replacement

Two cards are drawn from a well shuffled deck of 52 cards without replacement. Find the following probabilities.

a. The probability that the second card is a heart given that the first card is a spade.

Without replacement means that the first card is set aside before the second card is drawn and we assume the first card is a spade. There are only 51 cards to choose from for the second card. Thirteen of those cards are hearts.

It's important to notice that the question only asks about the second card.

$$P(\text{2nd heart} \mid \text{1st spade}) = \frac{13}{51}$$

The probability that the second card is a heart given that the first card is a spade is $\frac{13}{51}$.

b. The probability that the first card is a face card and the second card an ace.

Notice that this time the question asks about both of the cards.

There are 12 face cards out of 52 cards when we draw the first card. We set the first card aside and assume that it is a face card. Then there are four aces out of the 51 remaining cards. We want to draw a face card <u>and</u> an ace so use multiplication.

$$P(\text{1st face card and 2nd ace}) = \frac{12}{52} \cdot \frac{4}{51} = \frac{48}{2652} \approx 0.018$$

The probability that the first card is a face card and the second card an ace is approximately 0.018 or 1.8%.

c. The probability that one card is a heart and the other a club.

There are two ways for this to happen. We could get a heart first and a club second or we could get the club first and the heart second.

$$
\begin{aligned}
P(\text{heart and club}) &= P(\text{heart 1st and club 2nd or club 1st and heart 2nd}) \\
&= P(\text{heart 1st and club 2nd}) + P(\text{club 1st and heart 2nd}) \\
&= \frac{13}{52}\cdot\frac{13}{51} + \frac{13}{52}\cdot\frac{13}{51} \\
&\approx 0.127
\end{aligned}
$$

The probability that one card is a heart and the other a club is approximately 0.127 or 12.7%.

Example: Conditional Probability for Rolling Dice

Two fair dice are rolled and the sum of the numbers is observed. What is the probability that the sum is at least nine if it is known that one of the dice shows a five?

Solution: Since we are given that one of the dice shows a five this is a conditional probability. List the pairs of dice with one die showing a five. Be careful not to count (5,5) twice.

$$\{(1,5), (2,5), (3,5), (4,5), (5,5), (6,5), (5,1), (5,2), (5,3), (5,4), (5,6)\}$$

List the pairs from above that have a sum of at least nine.

$$\{(4,5), (5,5), (6,5), (5,4), (5,6)\}$$

There are 11 ways for one die to show a five and five of these ways have a sum of at least nine.

$$P(\text{sum at least } 9 \mid \text{one die is a } 5) = \frac{5}{11}$$

The probability that the sum is at least nine if it is known that one of the dice shows a five is $\frac{5}{11}$.

Multiplication Rule for "And" Probabilities: Any Events
For events A and B, $\boldsymbol{P(A \cap B) = P(A) \cdot P(B \mid A)}$

Conditional Probability
For events A and B, $\boldsymbol{P(B \mid A) = \frac{P(A \cap B)}{P(A)}}$

Independent Events

- Two events A and B are said to be independent if the occurrence of one event is not affected by the occurrence of the other. That is:

$$P(A|B) = P(A) \tag{24}$$

And we also have

$$P(B|A) = P(B) \tag{25}$$

Since

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \Longrightarrow P(A\cap B) = P(A)P(B) \qquad \textbf{(joint occurrence, intersection)} \tag{26}$$

- - Note that for mutually exclusive events $P(A\cap B) = 0$

Therefore, for $P(A) \neq 0,\ P(B) \neq 0$, A and B cannot be both mutually exclusive ($A \cap B = \emptyset$), and independent ($A \cap B \neq \emptyset$).

Example:

One card is selected from 52 card deck. Define events

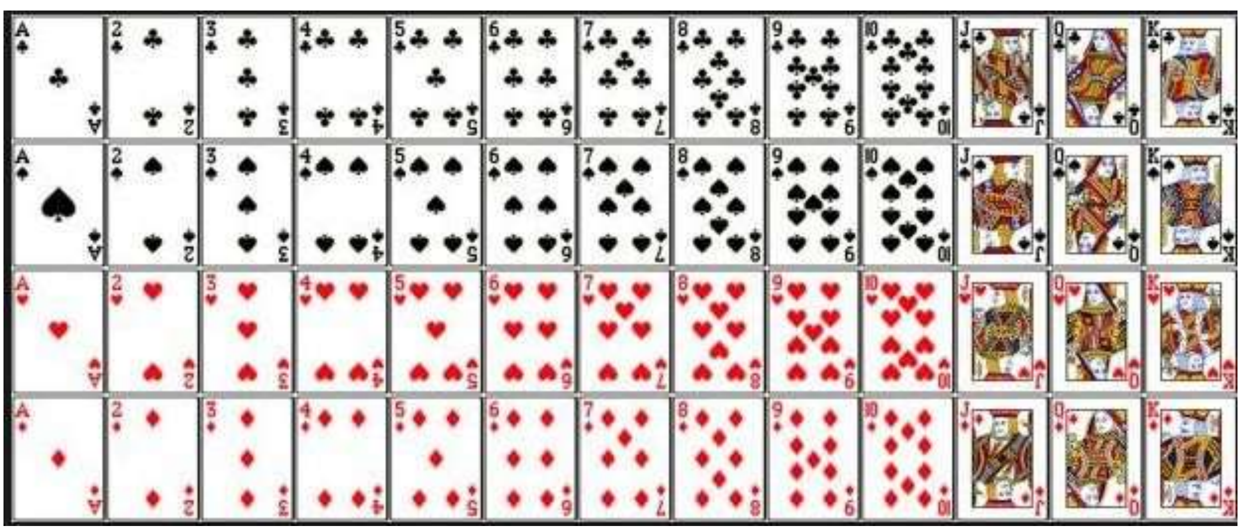
$A =$ "*select a king*", $B =$ *select Jack or queen*" *and* $C =$ "*select a heart*"

Find:

a) $P(A), P(B), P(C)$

b) $P(A\cap B), P(B\cap C), P(A\cap C)$.

c) Are the events independent?



Solution:

$a)\ P(A) = \frac{4}{52}, P(B) = \frac{8}{52}, P(C) = \frac{13}{52}$

It is not possible to simultaneously select a king and a jack or a queen.

$b)\ P(A \cap B) = 0, P(A \cap C) = \frac{1}{52}, P(B \cap C) = \frac{2}{52}$

We determine whether A, B and C are independent by pairs.

$c)\ P(A \cap B) = 0 \neq P(A).P(B) \Longrightarrow A$ and B are not independent.

$P(A \cap C) = \frac{1}{52} = P(A).P(C)) \Longrightarrow A$ and C are independent.

$P(B \cap C) = \frac{2}{52} = P(B).P(C)) \Longrightarrow B$ and C are independent.

In case of multiple events, they are said to be independent if all pairs are independent and

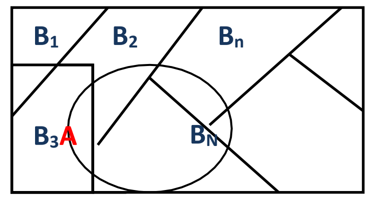
$$P(A_1 \cap A_2 \cap A_3) = P(A_1).P(A_2).P(A_3)$$

Total probability

- Suppose we are given n mutually exclusive events $B_n, n = 1, \dots, N$ such that:

$$\bigcup_{n=1}^{N} B_n = S \tag{18}$$

and $B_m \cap B_n = \emptyset$ for $m \neq n$



- The **total probability of** an event A defined on the sample space S can be expressed in terms of conditional probabilities as follows:

$$P(A) = \sum_{n=1}^{N} P(A|B_n)P(B_n) \tag{19}$$

Proof: since $A = A \cap S = A \cap (\bigcup_{n=1}^{N} B_n) = \bigcup_{n=1}^{N} (A \cap B_n)$

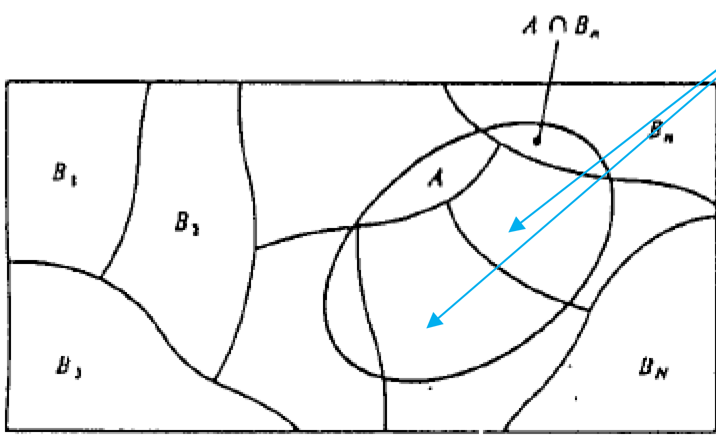
As shown in the diagram are mutually exclusive events; therefore

$$P(A) = P\left[\bigcup_{n=1}^{N} (A \cap B_n)\right] = \sum_{n=1}^{N} P(A \cap B_n)$$

Since $P(A \cap B_n) = P(A|B_n)P(B_n)$

$A \cap B_m$ *and* $A \cap B_n$

Are mutually exclusive



Bayes' Theorem:

- The Bayes rule expresses a conditional probability in terms of other conditional probabilities, we have:

$$P(B_n|A) = \frac{P(B_n \cap A)}{P(A)} \tag{20}$$

$$P(A|B_n) = \frac{P(A \cap B_n)}{P(Bn)} \tag{21}$$

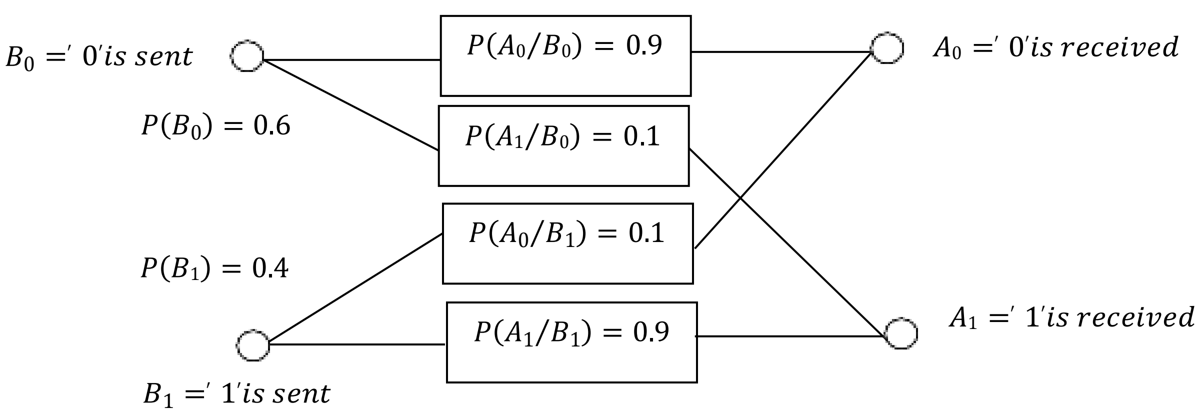
Therefore, one form of the Bayes theorem given by equating these two expressions:

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A)} \tag{22}$$

Which can be written also as another form:

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A|B_1)P(B_1)+, \ldots\ldots P(A|B_N)P(B_N)} \tag{23}$$

Example: A binary Communication system is described as:



Find:

a) $P(A_0)$ ('0' is received).

b) $P(A_1)$ ('1' is received).

c) $P(B_0/A_0), P(B_0/A_1), P(B_1/A_0), P(B_1/A_1)$.

Solution:

Total probability

$a)\ P(A_0) = P(A_0|B_0)\cdot P(B_0) + P(A_0|B_1)\cdot P(B_1)$

$= 0.9(0.6) + 0.1(0.4) = 0.58$

$b)\ P(A_1) = P(A_1|B_0)\cdot P(B_0) + P(A_1|B_1)\cdot P(B_1)$

$= 0.1(0.6) + 0.9(0.4) = 0.42$

Note that A_0 and A_1 are mutually exclusive and $P(A_0) + P(A_1) = 1$

$$c)\ \ P(B_0|A_0) = \frac{P(A_0|B_0)\cdot P(B_0)}{P(A_0)} = \frac{0.9(0.6)}{0.58} = 0.931$$

Bayes' Theorem

$$P(B_0|A_1) = \frac{P(A_1|B_0)\cdot P(B_0)}{P(A_1)} = \frac{0.1(0.6)}{0.42} = 0.143$$

$$P(B_1|A_0) = \frac{P(A_0|B_1)\cdot P(B_1)}{P(A_0)} = \frac{0.1(0.4)}{0.58} = 0.069$$

$$P(B_1|A_1) = \frac{P(A_1|B_1)\cdot P(B_1)}{P(A_1)} = \frac{0.9(0.4)}{0.42} = 0.857$$

Counting Methods

Recall that $P(A) = \frac{\text{number of ways for } A \text{ to occur}}{\text{total number of outcomes}}$ for theoretical probabilities. So far the problems we have looked at had rather small total number of outcomes. We could easily count the number of elements in the sample space. If there are a large number of elements in the sample space, we can use counting techniques such as permutations or combinations to count them.

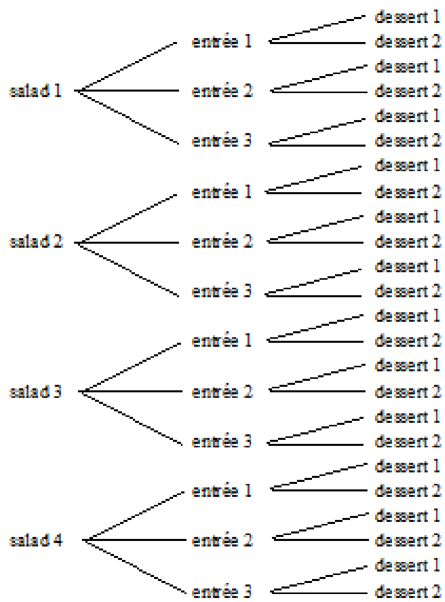
Multiplication Principle and Tree Diagrams:

The simplest of the counting techniques is the multiplication principle. A tree diagram is a useful tool for visualizing the multiplication principle.

Example: Multiplication Principle for a Three Course Dinner

Let's say that a person walks into a restaurant for a three-course dinner. There are four different salads, three different entrees, and two different desserts to choose from. Assuming the person wants to eat a salad, an entrée and a desert, how many different meals are possible?

Figure: Tree Diagram for Three-Course Dinner



Looking at the tree diagram we can see that the total number of meals is 4 x 3 x 2 = 24 meals.

Multiplication Principle: If there are n_1 ways to of choosing the first item, n_2 ways of choosing the second item after the first item is chosen, n_3 ways of choosing the third item after the first two have been chosen, and so on until there are n_k ways of choosing the last item after the earlier choices, then the total number of choices overall is given by $n_1 n_2 n_3 \dots n_k$.

Example: Multiplication Principle for Lining up People

Let's look at the number of ways that four people can line up. We can choose any of the four people to be first. Then there are three people who can be second and two people who can be third. At this point there is only one person left to be last. Using the multiplication principle there are 4 x 3 x 2 x 1 = 24 ways for four people to line up.

This type of calculation occurs frequently in counting problems so we have some notation to simplify the problem.

The **factorial** of n, read "n factorial" is $n! = n(n-1)(n-2)\dots(2)(1)$.
By definition, $0! = 1$.

Example: Factorials

$5! = 5 \; x \; 4 \; x \; 3 \; x \; 2 \; x \; 1 \; = \; 120$

$8! = 8 \; x \; 7 \; x \; 6 \; x \; 5 \; x \; 4 \; x \; 3 \; x \; 2 \; x \; 1 \; = \; 40320$

Factorials get very large very fast. $20! = 2.43 \times 10^{18}$ and $40! = 8.16 \times 10^{47}$. 70! is larger than most calculators can handle.

The multiplication principle may seem like a very simple idea but it is very powerful. Many complex counting problems can be solved using the multiplication principle.

Example: Multiplication Principle for License Plates

Some license plates in Arizona consist of three digits followed by three letters. How many license plates of this type are possible if?

a. both digits and letters can be repeated?
There are 10 digits (0, 1, 2, 3, …, 9) and 26 letters.

$$\underbrace{(\underline{10}\cdot\underline{10}\cdot\underline{10})}_{\text{digits}} \cdot \underbrace{(\underline{26}\cdot\underline{26}\cdot\underline{26})}_{\text{letters}} = 17{,}576{,}000 \text{ license plates}$$

b. letters can be repeated but digits cannot?

$$\underbrace{(\underline{10}\cdot\underline{9}\cdot\underline{8})}_{\text{digits}} \cdot \underbrace{(\underline{26}\cdot\underline{26}\cdot\underline{26})}_{\text{letters}} = 12{,}654{,}720 \text{ license plates}$$

c. the first digit cannot be zero and both digits and letters can be repeated?

$$\underbrace{(\underline{9}\cdot\underline{10}\cdot\underline{10})}_{\text{digits}} \cdot \underbrace{(\underline{26}\cdot\underline{26}\cdot\underline{26})}_{\text{letters}} = 15{,}818{,}400 \text{ license plates}$$

d. neither digits nor numbers can be repeated.

$$\underbrace{(\underline{10}\cdot\underline{9}\cdot\underline{8})}_{\text{digits}} \cdot \underbrace{(\underline{26}\cdot\underline{25}\cdot\underline{24})}_{\text{letters}} = 11{,}232{,}000 \text{ licenseplates}$$

Permutations:

Consider the following counting problems. 1) In how many ways can three runners finish a race? 2) In how many ways can a group of three people be chosen to work on a project? What is the difference between these two problems? In the first problem the order that the runners finish the race matters. In the second problem the order in which the three people are chosen is not important, only which three people are chosen matters.

A **permutation** is an arrangement of a set of items.
The number of permutations of n items taking r at a time is given by: $P(n,r) = \frac{n!}{(n-r)!}$

Note: Many calculators can calculate permutations directly. Look for a function that looks like ${}_nP_r$ or $P(n,r)$.

Example: Permutation for Race Cars

Let's look at a simple example to understand the formula for the number of permutations of a set of objects. Assume that 10 cars are in a race. In how many ways can three cars finish in first, second and third place? The order in which the cars finish is important. Use the multiplication principle. There are 10 possible cars to finish first. Once a car has finished first, there are nine cars to finish second. After the second car is finished, any of the eight remaining cars can finish third. 10 x 9 x 8 = 720. This is a permutation of 10 items taking three at a time.

Using the permutation formula:

$$P(10,3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{1098\cancel{7}6\cancel{5}43\cancel{2}1}{76\cancel{5}4\cancel{3}2\cancel{1}} = 1098 = 720$$

Using the multiplication principle:

$\underline{10} \bullet \underline{9} \bullet \underline{8} = 720$

There are 720 different ways for cars to finish in the top three places.

Example: Permutation for Orchestra Programs

The school orchestra is planning to play six pieces of music at their next concert. How many different programs are possible?

This is a permutation because they are arranging the songs in order to make the program.

Using the permutation formula:

$$P(6,6) = \frac{6!}{(6-6)!} = \frac{6!}{0!} = \frac{720}{1} = 720$$

Using the multiplication principle:

$\underline{6} \cdot \underline{5} \cdot \underline{4} \cdot \underline{3} \cdot \underline{2} \cdot \underline{1} = 720$

There are 720 different ways of arranging the songs to make the program.

Example: Permutation for Club Officers

The Volunteer Club has 18 members. An election is held to choose a president, vice-president and secretary. In how many ways can the three officers be chosen?

The order in which the officers are chosen matters so this is a permutation.

Using the permutation formula:

$$P(18,3)=\frac{18!}{(18-3)!}=\frac{18!}{15!}=18\cdot 17\cdot 16=4896$$

Note: All digits in 18! in the numerator from 15 down to one will cancel with the 15! in the denominator.

Using the multiplication principle:

$$\underset{\text{Pres.}}{\underline{18}}\cdot\underset{\text{V.P.}}{\underline{17}}\cdot\underset{\text{Sec.}}{\underline{16}}=4896$$

There are 4896 different ways the three officers can be chosen.

Another notation for permutations is nPr. So, $P(18,3)$ can also be written as ${}_{18}P_3$. Most scientific calculators have an nPr button or function.

Combinations:

Example: Formula for Combinations

Choose a committee of two people from persons A, B, C, D and E. By the multiplication principle there are $5 \cdot 4 = 20$ ways to arrange the two people.

AB	AC	AD	AE	BA	BC	BD	BE	CA	CB
CD	CE	DA	DB	DC	DE	EA	EB	EC	ED

Committees AB and BA are the same committee. Similarly, for committee's CD and DC. Every committee is counted twice. $\frac{20}{2} = 10$ so, there are 10 possible different committees.

Now choose a committee of three people from persons A, B, C, D and E. There are $5 \cdot 4 \cdot 3 = 60$ ways to pick three people in order. Think about the committees with persons A, B and C. There are $3! = 6$ of them.

ABC ACB BAC BCA CAB CBA

Each of these is counted as one of the 60 possibilities but they are the same committee. Each committee is counted six times so there are $\frac{60}{6} = 10$ different committees.

In both cases we divided the number of permutations by the number of ways to rearrange the people chosen.

The number of permutations of n people taking r at a time is *P(n,r)* and the number of ways to rearrange the people chosen is *r!*. Putting these together we get

$$\frac{\text{\# permutations of n items taking r at a time}}{\text{\# ways to arrange r items}} = \frac{P(n,r)}{r!} = \frac{\frac{n!}{(n-r)!}}{\frac{r!}{1}}$$

$$= \frac{n!}{(n-r)!} \cdot \frac{1}{r!}$$

$$= \frac{n!}{(n-r)!r!}$$

A **combination** is a selection of objects in which the order of selection does not matter.
The number of combinations of n items taking r at a time is: $C(n,r) = \frac{n!}{r!(n-r)!}$

Note: Many calculators can calculate combinations directly. Look for a function that looks like ${}_nC_r$ *or* $C(n,r)$.

Example: Combination for Picking Books

A student has a summer reading list of eight books. The student must read five of the books before the end of the summer. In how many ways can the student read five of the eight books?

The order of the books is not important, only which books are read. This is a combination of eight items taking five at a time.

$$C(8,5) = \frac{8!}{5!(8-5)!} = \frac{8!}{5!3!} = \frac{876\cancel{54321}}{\cancel{54321}321} = \frac{87\cancel{6}}{\cancel{32}1} = 87 = 56$$

There are 56 ways to choose five of the books to read.

Example: Combination for Halloween Candy

A child wants to pick three pieces of Halloween candy to take in her school lunch box. If there are 13 pieces of candy to choose from, how many ways can she pick the three pieces?

This is a combination because it does not matter in what order the candy is chosen.

$$\boldsymbol{C(13,3) = \frac{13!}{3!\,(13-3)!} = \frac{13!}{3!\,10!} = \frac{131211\cancel{10987654321}}{321\cancel{10987654321}}}$$

$$\boldsymbol{= \frac{131211}{321} = \frac{1716}{6} = 286}$$

There are 286 ways to choose the three pieces of candy to pack in her lunch.

Note: The difference between a combination and a permutation is whether order matters or not. If the order of the items is important, use a permutation. If the order of the items is not important, use a combination.

Example: Permutation or Combination for Bicycle Serial Numbers

A serial number for a particular model of bicycle consists of a letter followed by four digits and ends with two letters. Neither letters nor numbers can be repeated. How many different serial numbers are possible?

This is a permutation because the order matters.

Use the multiplication principle to solve this. There are 26 letters and 10 digits possible.

$\underline{26}\cdot\underline{10}\cdot\underline{9}\cdot\underline{8}\cdot\underline{7}\cdot\underline{25}\cdot\underline{24}=78{,}624{,}000$

There are 78,624,000 different serial numbers of this form.

Example: Permutation or Combination for Choosing Men and Women

A class consists of 15 men and 12 women. In how many ways can two men and two women be chosen to participate in an in-class activity?

This is a combination since the order in which the people is chosen is not important.

Choose two men: $C(15,2) = \frac{15!}{2!(15-2)!} = \frac{15!}{2!13!} = 105$

Choose two women: $C(12,2) = \frac{12!}{2!(12-2)!} = \frac{12!}{2!10!} = 66$

We want 2 men <u>and</u> 2 women so multiply these results.

$105(66) = 6930$

There are 6930 ways to choose two men and two women to participate.

Probabilities Involving Permutations and Combinations:

Now that we can calculate the number of permutations or combinations, we can use that information to calculate probabilities.

Example: Probability with a Combination for Choosing Students

There are 20 students in a class. Twelve of the students are women. The names of the students are put into a hat and five names are drawn. What is the probability that all of the chosen students are women?

This is a combination because the order of choosing the students is not important.

$P(\text{all females}) = \frac{\#\text{ ways to pick 5 women}}{\#\text{ ways to pick 5 students}}$

The number of ways to choose 5 women is $C(12,5) = 792$

The number of ways to choose 5 students is $C(20,5) = 15,504$

$$P(\text{all females}) = \frac{\#\text{ ways to pick 5 women}}{\#\text{ ways to pick 5 students}} = \frac{792}{15,504} = 0.051$$

The probability that all the chosen students are women is 0.051 or 5.1%.

Example: Probability with a Permutation for a Duck Race

The local Boys and Girls Club holds a duck race to raise money. Community members buy a rubber duck marked with a numeral between 1 and 30, inclusive. The box of 30 ducks is emptied into a creek and allowed to float downstream to the finish line. What is the probability that ducks numbered 5, 18 and 21 finish in first, second and third, respectively?

This is a permutation since the order of finish is important.

$$P(5, 18 \text{ \& } 21 \text{ finish 1st, 2nd \& 3rd}) = \frac{\#\text{ ways 5, 18 \& 21 finish 1st, 2nd \& 3rd}}{\#\text{ ways any ducks can finish 1st, 2nd \& 3rd}}$$

There is only one way that the ducks can finish with #5 in first, #18 in second and #21 in third.

The number of ways any ducks can finish in first, second and third is

$P(30,3) = 24,360$

$$P(5, 18 \text{ \& } 21 \text{ finish 1st, 2nd \& 3rd}) = \frac{\#\text{ ways 5, 18 \& 21 finish 1st, 2nd \& 3rd}}{\#\text{ ways any ducks can finish 1st, 2nd \& 3rd}}$$
$$= \frac{1}{24,360} \approx 4.10x10^{-5}$$

The probability that ducks numbered 5, 18 and 21 finish in first, second and third, respectively, is approximately 0.000041 or 0.0041%.

Example: Probability with a Permutation for Two-Card Poker Hands

A poker hand consists of two cards. What is the probability that the poker hand consists of two jacks or two fives?

It's not possible to get two jacks and two fives at the same time so these are mutually exclusive events.

The number of ways to get two jacks is $C(4,2) = 6$.

The number of ways to get two fives is $C(4,2) = 6$

The number of ways to get two jacks or two fives is 6 + 6 = 12.

The total number of ways to get a 2-card poker hand is $C(52,2) = 1326$

$$P(2 \text{ jacks or } 2 \text{ fives}) = \frac{\text{number of ways to get 2 jacks or 2 fives}}{\text{number of ways to choose 2 cards}} = \frac{12}{1326} \approx 0.009$$

The probability of getting two jacks or two fives is approximately 0.009 or 0.9%.

Example: Probability with a Combination for Rotten Apples

A basket contains 10 good apples and two bad apples. If a distracted shopper reaches into the basket and picks three apples without looking, what is the probability he gets one bad apple?

This is a combination since the order in which the apples were picked is not important. He picks three apples total. If one apple is bad the other two must be good. Find the probability of one bad apple and two good apples.

$$P(\text{one bad and two good apples}) = \frac{\#\text{ ways to get one bad and two good apples}}{\#\text{ ways to get three apples}}$$

The number of ways to get one bad and two good apples is

$C(2,1) \cdot C(10,2) = 2 \cdot 45 = 90$

The number of ways to get three apples is $C(10,3) = 120$

$$P(\text{one bad and two good apples}) = \frac{\#\text{ ways to get one bad and two good apples}}{\#\text{ ways to get three apples}}$$

$= \frac{90}{120} = 0.75$The probability of getting one bad apple out of three apples is 0.75 or 75%.

RANDOM VARIABLE: A random variable X represents a single experiment/trial. It consists of outcomes with a probability for each. Or in other words a random variable is a variable which assigns a real value to each outcome of a random experiment.

For example: Let E be the random experiment consisting of two tosses of a coin

$$S = \{HH, HT, TH, TT\}$$

We may define the random variable X which denotes the number of heads $(0,1\ or\ 2)$.

$$X = \{2,1,1,0\}$$

2.1 THEOREMS ON RANDOM VARIABLES

THEOREM 1 If $X_1\ and\ X_2$ are random variables and if $C_1\ and\ C_2$ are constants then $C_1X_1, C_2X_2, X_1 + X_2, C_1X_1 + C_2X_2, X_1 - X_2$ are also random variables.

THEOREM 2 If X is a random variable, then $\frac{1}{X}, |X|$ are also random variables.

THEOREM 3 If $X_1\ and\ X_2$ are random variables, then $\max[X_1, X_2]$ and $\min[X_1, X_2]$ are also random variables.

THEOREM 4 If X is a random variable and f is a continuous function, then $f(x)$ is a random variable.

2.2 DISTRIBUTION FUNCTION

Let X be a random variable, then the function $F_x(x) = F(x) = P(X \leq x) = P\{w: x(w) \leq x\}, -\infty < x < \infty$ is called the distribution function of X.

2.3 PROPERTIES OF DISTRIBUTION

1. If $F(x)$ is the distribution of the random variable X and if $a < b$, then

$$P(a < X < b) = F(b) - F(a)$$

2. If $F(x)$ is the distribution function of one-dimensional random variable X then

(i) $0 \leq F(x) \leq 1$.

(ii) $F(x) \leq F(y), if\ x < y$.

3. If $F(x)$ is the distribution function of one-dimensional random variable X then

$$F(\infty) = \lim_{x \to \infty} F(x) = 1$$

And

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$$

2.4 DISCRETE RANDOM VARIABLE

If a random variable X takes at most a countable number of values or countably infinite number of values, it is called a *discrete random variable.* In other words, a real valued function defined on a discrete sample space is called a *discrete random variable.*

Examples of discrete random variables

- **The outcome of a roll of a die may only take on the integer values from 1 to 6.**
- **The number of floods per year at a point on a river can only take on integer values.**
- **The number of children of a family.**
- **The number of students in a class.**

2.5 PROBABILITY MASS FUNCTION (PMF)

Suppose X is an one-dimensional discrete random variable taking at most a countably infinite number of values $x_1, x_2, \ldots$ with each possible outcome x_i, we associate a number p_i, $P(X = x_i) = p(x_i) = p_i$ called the probability of x_i.

The function $p(x_i), i = 1,2, \ldots$ satisfying the conditions

(i) $p(x_i) \geq 0, \forall i.$

(ii) $\sum_{i=1}^{\infty} p(x_i) = 1.$

is called the probability mass function or probability function of the random variable X. The collection of pairs $\{x_i, p_i\}$, $i = 1,2,3, \ldots$ is called the probability distribution of the random variable X.

Note: The set of values which X takes is called the ***spectrum*** of the random variable.

2.6 DISCRETE DISTRIBUTION FUNCTION

The distribution function of the random variable X with PMF $p(x_i), i = 1,2,3, \ldots$ is defined as

$$F(x_i) = \sum_{i: x_i \leq x} p(x_i)$$

Note:

(i) $p(x_i) = P(X = x) = F(x_i) - F(x_{i-1})$, **where F is the distribution function of the random variable X.**

(ii) **Mean of the random variable X is**

$$E(X) = \sum_x xp(x)$$

(iii) **Variance of the random variable X is**

$$Var(X) = \sum x^2 p(x) - \left[\sum xp(x)\right]^2$$

EXAMPLE 2.1 If X is a discrete random variable having the probability distribution

$X = x$	1	2	3
$P(X = x)$	k	$2k$	k

Find $P(X \le 2)$.

Solution

We know that

$$\sum P(X = x) = 1 \Longrightarrow P(X = 1) + P(X = 2) + P(X = 3) = 1 \Longrightarrow k + 2k + k = 1$$

$$\Longrightarrow 4k = 1 \Longrightarrow k = \frac{1}{4}$$

$$P(X \le 2) = P(X = 1) + P(X = 2) = k + 2k = 3k \Longrightarrow P(X \le 2) = 3 \cdot \frac{1}{4} = \frac{3}{4}$$

EXAMPLE 2.2 If X is a discrete random variable having the PMF

x	-1	0	1
$P(x)$	k	$2k$	$3k$

Find $P(X \ge 0)$.

Solution

We know that

$$\sum P(X=x)=1 \Longrightarrow P(X=-1)+P(X=0)+P(X=1)=1$$

$$\Longrightarrow k+2k+3k=1 \Longrightarrow 6k=1 \Longrightarrow k=\frac{1}{6}$$

$$P(X\geq 0)=P(X=0)+P(X=1)=2k+3k=5k \Longrightarrow P(X\geq 0)=5\cdot\frac{1}{6}=\frac{5}{6}$$

EXAMPLE 2.3 If X is a discrete random variable with the following probability distribution

x	1	2	3	4
$P(x)$	a	$2a$	$3a$	$4a$

Find $P(2<X<4)$.

Solution

We know that

$$\sum P(X=x)=1 \Longrightarrow P(X=1)+P(X=2)+P(X=3)+P(X=4)=1$$

$$\Longrightarrow a+2a+3a+4a=1 \Longrightarrow 10a=1 \Longrightarrow a=\frac{1}{10}$$

$$P(2<X<4)=P(X=3)=3a=3\cdot\frac{1}{10}=\frac{3}{10}$$

EXAMPLE 2.4 If the probability distribution of X is given as:

x	1	2	3	4
$P(x)$	0.4	0.3	0.2	0.1

Find $P(\frac{1}{2}<X<\frac{7}{2}|\,X>1)$.

Solution

By definition

$$P\left(\frac{1}{2} < X < \frac{7}{2} \middle| X > 1\right) = \frac{P\left(\frac{1}{2} < X < \frac{7}{2} \cap X > 1\right)}{P(X > 1)}$$

$$= \frac{P\left(1 < X < \frac{7}{2}\right)}{P(X > 1)} = \frac{P(X = 2) + P(X = 3)}{1 - P(X \le 1)} = \frac{P(X = 2) + P(X = 3)}{1 - P(X = 1)} = \frac{0.5}{0.6} = \frac{5}{6}$$

EXAMPLE 2.5 A random variable X has the following probability function

x	0	1	2	3	4
$P(x)$	k	$3k$	$5k$	$7k$	$9k$

Find

(i) The value of k.

(ii) $P(X < 3)$ and $P(0 < X < 4)$.

(iii) The distribution function of X.

Solution

(i) We know that

$$\sum p(x) = 1 \Longrightarrow P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1$$

$$\Longrightarrow k + 3k + 5k + 7k + 9k = 1 \Longrightarrow k = \frac{1}{25}$$

(ii) $P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$

$$= k + 3k + 5k = \frac{9}{25}$$

(iii) $P(0 < X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= 3k + 5k + 7k = 15k = \frac{15}{25}$$

(iv) The distribution function of X is

$$F(x) = 0, x < 0$$
$$= \frac{1}{25}, 0 \le x < 1$$
$$= \frac{4}{25}, 1 \le x < 2$$
$$= \frac{9}{25}, 2 \le x < 3$$
$$= \frac{16}{25}, 3 \le x < 4$$

$$= 1, x \geq 4$$

HOMEWORK

1) After a coin is tossed two times, if X is the number of the heads, find the probability distribution of X.
2) If $P(X = x) = \begin{cases} kx, & x = 1, 2, 3, 4, 5 \\ 0, & otherwise \end{cases}$ represents a probability distribution, find
   - (i) k.
   - (ii) $P(X$ being a prime number$)$.
   - (iii) $P\left(\frac{1}{2} < X < \frac{5}{2} \mid X > 1\right)$.
   - (iv) The distribution function.
3) Suppose that the random variable assumes three values $0, 1$ and 2 with probabilities $\frac{1}{3}, \frac{1}{6}$ and $\frac{1}{2}$ respectively. Obtain the distribution function of X
4) If $P(X = x) = \begin{cases} \frac{x}{15}, & x = 1, 2, 3, 4, 5 \\ 0, & otherwise \end{cases}$ find
   - (i) $P(X = 1\ or\ 2)$.
   - (ii) $P\left(\frac{1}{2} < X < \frac{5}{2} \mid X > 1\right)$.
5) If the probability mass function of a random variable is given by $P(X = r) = kr^3, r = 1, 2, 3, 4$, find
   - (i) The value of k.
   - (ii) $P\left(\frac{1}{2} < X < \frac{5}{2} \mid X > 1\right)$.
   - (iii) The mean and variance of X.
   - (iv) The distribution function of X.
6) Find $E(X)$ and $Var(X)$ for the examples $(2.1, 2.2, 2.3, 2.3, 2.4, 2.5)$ and the questions $(1, 2, 3, 4)$.

2.7 CONTINUOUS RANDOM VARIABLE

A random variable X is said to be continuous if it can take all possible values between certain limits. In other words, a random variable is said to be continuous when its different values cannot be put in one to one correspondence with a set of positive integers. Examples of continuous random variable are height, weight, age, etc.

Note: The sample space of the continuous random variable must be continuous and cannot be discrete. In most of the practical problems, continuous random variable represent measured data, such as all possible heights, weights, temperature, etc. whereas discrete random variables represent count data such as the number of defectives in a sample and so on.

2.8 PROBABILITY DENSITY FUNCTION (PDF)

Consider the small interval $\left(\boldsymbol{x}-\frac{\Delta x}{2}, \boldsymbol{x}+\frac{\Delta x}{2}\right)$ of length Δx round the point x. Let $f(x)$ be any continuous function of x so that $f(x)dx$ represents the probability that x falls in the infinitesimal interval $\left(\boldsymbol{x}-\frac{\Delta x}{2}, \boldsymbol{x}+\frac{\Delta x}{2}\right)$, which is denoted by $\boldsymbol{P}\left(\boldsymbol{x}-\frac{\Delta x}{2} \leq \boldsymbol{x} \leq +\frac{\Delta x}{2}\right) = \boldsymbol{f(x)dx}$.

Let $f(x)dx$ represent the area bounded by the curve $y = f(x)$. x axis and the ordinates at the points $\boldsymbol{x}-\frac{\Delta x}{2}$ and $\boldsymbol{x}+\frac{\Delta x}{2}$. The function $f(x)$ so defined is known as probability density function or density function of the random variable X.

The probability density function of a random variable X denoted by $f(x)$ has the following properties:

(i) $f(x) \geq 0, \forall x \in R$.

(ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

(iii) $P(a < X < b) = \int_{a}^{b} f(x)\, dx$.

Note: In case of continuous random variables, the probability at a point is always zero, i.e. $P(X = a) = 0$ for all possible values of a.

2.9 CUMULATIVE DISTRIBUTION FUNCTION (CDF)

The cumulative distribution $F(x)$ of a continuous random variable X with PDF $f(x)$ is given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\, dx, -\infty < x < \infty$$

Note: $P(a < x < b) = F(b) - F(a)$.

The relation between the CDF and PDF is

$$f(x) = \frac{d}{dx}F(x)$$

If X is a continuous random variable with PDF $f(x)$, then

$$Mean = E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$Var(X) = E(X^2) - [E(X)]^2$$

EXAMPLE 2.6 Verify whether $f(x) = \begin{cases} |x|, & -1 \le x \le 1 \\ 0, & elsewhere \end{cases}$ can be the PDF of a continuous random variable.

Solution

For $f(x)$ to be a PDF, it should satisfy

$$(i) f(x) = |x| \ge 0, \forall x \qquad (ii) \int_{-\infty}^{\infty} f(x)\,dx = 1$$

Given:

$(i)\ f(x) = |x| \ge 0, \forall x$

$(ii)\ \int_{-\infty}^{\infty} f(x)\,dx = \int_{-1}^{1} |x|\,dx = 2\int_{0}^{1} x dx = 2\left[\frac{x^2}{2}\right]_0^1 = 1$

Therefore, $f(x)$ can be the PDF of X.

EXAMPLE 2.7 A random variable X has the PDF $f(x)$ given by

$$f(x) = \begin{cases} cxe^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

Find the value of c and CDF of x.

Solution If $f(x)$ is a PDF, then

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \Longrightarrow \int_{0}^{\infty} cxe^{-x}\,dx = 1 \Longrightarrow c[xe^{-x} - e^{-x}]_0^{\infty} = 1$$

$$\Longrightarrow c(0+1) = 1 \Longrightarrow c = 1$$

$$f(x) = xe^{-x}, x > 0$$

The CDF of X

$$F(x) = P(X \le x) = \int_0^x xe^{-x}\,dx = [xe^{-x} - e^{-x}]_0^x = (-xe^{-x} - e^{-x}) - (0-1)$$

$$F(x) = 1 - (1+x)e^{-x}, x > 0$$

$$= 0, otherwise$$

HOMEWORK

1) **A continuous random variable X follows the probability law $f(x) = ax^2, 0 \le x \le 1$. Determine a and find the probability that x lies between $\frac{1}{4}$ and $\frac{1}{2}$.**
2) **If the PDF of a random variable X is $f(x) = \frac{x}{2}, 0 \le x \le 2$. Find $P(X > 1.5 | X > 1)$.**
3) **If $f(x) = kx^2, 0 < x < 3$ is to be the density function, find the value of k.**
4) **The cumulative distribution of X is $F(x) = \frac{x^3+1}{9}, -1 < x < 2$ and $= 0$ otherwise. Find $P(0 < X < 1)$.**
5) **The CDF of X is given by $F(x) = \begin{cases} 0, & x > 0 \\ x^2, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$. Find the PDF of X and obtain $P(X > 0.75)$.**
6) **Find $E(X)$ and $Var(X)$ for the examples $(2.6, 2.7)$ and the questions $(1, 2, 3)$.**

2.10 PROPERTIES OF EXPECTATION AND VARIANCE

1. If N is a random variable, then *E{aX + b) —— aE(X) + b.*

Proof: By definition

E{aX + b) —— (*ax* + *b*) *p{x)*

= a up(x) + *b* *p(x)*

—— *aE(X)* -F *b* . p(x) = 1]

.. *E(aX + b) —— aE(X) + b*

Note: *E{X* -F *Y) —— E{X)* -1- *E{Y).*

2. If N is a random variable, then Our(nd -1- *b)* —— n^2 Kar(I).

Proof: Let *Y* —— *aX* -1- *b, E{Y) —— E(aX* -F *b) —— aE{X)* -F *b*

.. *Y — E{Y) ——* (nd -F *b) — [aE(X)* -F *b]* —— n[N — *E(X)]*

[F — £(F)]^2 = a^2[N — *E{X)]*^{2} (F — Y) ^{2} = a^2(I — I) ^{2}

$$E(Y-\bar{Y})^2 = E[a^2(X-\bar{X})^2] = a^2E[(X-\bar{X})^2]$$

Var(Y) —— n^2 **far(I)**

-- **far(at +*b*)**-- a ^{2}Uar(I) [*Var{b)* – 0]

3. It can easily be proved that if N and *Y* are independent, then

Our(nd + *bY)* —— a ^{2}*Var tX)* -F *b*^{2}*Var{Y)*

4. IfI and *Y* are independent random variables, then *E{XY) —— E{X-) E{Y).*

5. IfI and *Y* are any two random variables such that *Y E* X, then *E{Y) E E(X).*

Proof: Given *Y E X* *Y — X E* 0

i.e. *X — Y E* 0

.. *E(X — Y) E* 0 *E(X) — E(Y) E* 0

i.e. $E(X) \quad E(Y) \quad E(Y)\ E\ E(X)$

2.11 COVARIANCE (A, *Y)*

The covariance of the two random variables is denoted by

$$p(x, \mathrm{y}) = pqp — Cov\{X, Y)$$

Which is defined as

$$Cov(X,Y) — E\{XY) – E\{X)E\{Y)$$

Important note:

(i) If N and Y are independent random variables, then $Cov\{X,Y)$ — 0.

(ii) If N and Y are any two random variables, then

$$Var(aX\ k\ bY) — \mathrm{n}^2 \mathrm{far(N)}\ 4\text{-}\ \mathrm{b}^2 \mathrm{Par(r)}\ k\ 2ab\ Cov(X, Y)$$

(iii) $Cov\{aX \text{ -1- } b, cY + d)$ — $ab\ Cov\{X, Y).$

EXAMPLE 2.8 The number of hardware failures of a computer system in a week of operations has the following PMF:

Numbers of failures	0	1	2	3	4	5	6
probability	**0. 18**	**0. 28**	**0. 25**	**0. 18**	**0. 16**	**0. 04**	**0. 01**

Find the mean of the number of failures in a week.

Solution We know that

$$E\{X) — \overset{6}{\underset{a=0}{}}\ x\ p(x)$$

$$= 0 \times 0.18 \text{ —F } 1 \times 0.28 \text{ —F } 2 \times 0.25 \text{ -1- } 3 \times 0.18 \text{ —F } 4 \times 0.06 + 5 \times 0.04 \text{ —s } 6 \times 0.01$$

$$E(X) = 1.82$$

EXAMPLE 2.9 If N and Y are independent random variables with means 2,3 and variance 1,2 respectively, find the mean and variance of the random variable Z — 2J — 5r.

Solution Given

$$E(X) = 2, £(\mathrm{F}) — 3, \mathrm{far(I)} = 1, Var(Y) — 2, \mathrm{Z} — 2\mathrm{N} – 5\mathrm{P}$$

$$.\ E\{Z)\ \text{-}\ \mathrm{E}(2\mathrm{X} – 5Y) — 2E\{X) – 5E\{Y) — (2 \times 2) – (5 \times 3) = —11$$

If A and Y are independent, then

$$Our(at + bY) — a\ ^2Var\,tX) + b^2Var\{Y)$$

- $far(2J — 5F) — 4Kar(I) + 25Par(r) —— (4 \text{ x } 1) + (25 \text{ x } 2) = 54.$

EXAMPLE 2.10 A fair coin is tossed three times. Let A be the number of tails appearing. Find the probability distribution of I and also calculate *E{X).*

Solution Given N denotes the number of tails. Since the coin is tossed 3 times, the number of tails may be 0,1,2 or 3

$$it(S) —— 2^3 = 8$$

$$S —— [HHH, HHT, HTH, HTT, THH, TTH, THT, TTT\ j$$

The probability distribution is

	0	1	2	3
$p(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Now,

$$E(X) = \sum_{i=0}^{3} x\,p(x) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2}$$

HOMEWORK

1) Given the following probability distribution of X

3 0 1 2 3

0 05 0.10 0.30 0 0.30 0 5 0 10

Compute

(i) *E(X).*

(ii) *$E(X^2)$.*

(iii) **r(2X** + 3).

(iv) Our(2X -1- 3).

2) Suppose that the random variable X is equal to the number of hits obtained by a certain baseball player in his next 3 bats and if P(X = 1) = 0. 3, P(X - 2) = 0. 2 and P(X = 0) = 3P(X = 3), find *E(X).*

3) The cumulative distribution function (CDF) of a random variable X is *F{x)* —— 1 — (I + *z)e* ^, z > 0. Find the probability density function of X, mean and variance.

4) If $dF —— kz^2 e*^dz, z \quad 0$, find k, mean and variance.

5) Let X be a random variable with $E(X) = 10$ and Our(X) = 25. Find the positive values of n and b such that $Y —— aX — b$ has expectation 0 and variance 1.

6) For the triangular distribution $f(x) = \begin{cases} x, & 0 < x E 1 \\ 2 — z, & 1 ñ z < 2 \\ 0, & otherwise \end{cases}$. Find the mean and variance.

7) Suppose the duration x in minutes of long-distance calls from your home, follows exponential law with PDF $f(x) = e$, for z > 0 and 0 otherwise. Find

(i) f•(X > 5).

(ii) P(3 N X N 6).

(iii) Mean of X.

(iv) Variance of X.

2.12 MOMENTS

If X is a random variable which is discrete or continuous, the moments about the origin denoted by μ_r' is defined as

$$\mu_r' = E(X^r), for\ r = 1,2,3, \ldots$$

The moments about the mean or central moments denoted by μ_r is defined as

$$\mu_r = E[(X - \bar{X})^r], for\ r = 1,2,3, \ldots$$

If X is a discrete random variable which can assume any of the values $x_1, x_2, \ldots, x_n$ with respective probabilities $p(x_1), p(x_2), \ldots, p(x_n)$, then

$$\mu_r = E(X^r) = \sum_{r=1}^{\infty} x^r\, p(x)$$

And

$$\mu_r = E[(X - \bar{X})^r] = \sum_{r=1}^{\infty} (x - \bar{x})^r\, p(x_r)$$

If X is a continuous random variable with PDF $f(x)$, then

$$\mu_r' = \int_{-\infty}^{\infty} x^r\, f(x)dx, r = 1,2,3, \ldots$$

And

$$\mu_r = \int_{-\infty}^{\infty} (x - \bar{x})^r\, f(x)dx, r = 1,2,3, \ldots$$

2.13 Relation between Moments about Origin and Moments about Mean $\bar{X}$

By definition

$$\begin{aligned}\boldsymbol{\mu_r} &= \boldsymbol{E[(X - \bar{X})^r]} \\ &= \boldsymbol{E[X^r - rC_1X^{r-1}\bar{X} + rC_2X^{r-2}\bar{X}^2 - rC_3X^{r-3}\bar{X}^3 + \cdots} \\ &\quad \boldsymbol{+ (-1)^{r-1}rC_{r-1}X\bar{X}^{r-1} + (-1)^r\bar{X}^r]}\end{aligned}$$

$$\boldsymbol{= E(X^r) - rE(X^{r-1})\bar{X} + \frac{r(r-1)}{2!}E(X^{r-2})\bar{X}^2 - \frac{r(r-1)(r-2)}{3!}E(X^{r-3})\bar{X}^3 + \cdots + (-1)^r\bar{X}^r}$$

$\because \bar{X} = E(X) = \mu'_1$, we have

$$\mu_r = \mu'_r - r\mu'_{r-1}\mu'_1 + \frac{r(r-1)}{2!}\mu'_{r-2}\mu'^2_1 - \frac{r(r-1)(r-2)}{3!}\mu'_{r-3}\mu'^3_1 + \cdots + (-1)^r\mu'^r_1$$

$$E(X-\bar{X}) = \mu_1 = \mu'_1 - \mu'_1 = 0$$

⟹ The first moment about the mean is always zero.

$$E(X-\bar{X})^2 = \mu_2 = \mu'_2 - 2\mu'^2_1 + \mu'^2_1$$

i.e.

$$\mu_2 = \mu'_2 - \mu'^2_1$$

$$E(X-\bar{X})^3 = \mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + \frac{6}{2!}\mu'_1\mu'^2_1 - \mu'^3_1$$

$$= \mu'_3 - 3\mu'_2\mu'_1 + 3\mu'^3_1 - \mu'^3_1$$

i.e.

$$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1$$

similarly,

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'^2_1\mu'_2 - 3\mu'^4_1$$

And so on.

2.14 Relation between Moments about any point A and Moments about Mean $\bar{X}$

We know that

$$\mu'_r = E[(X-A)^r]$$

Putting $r = 1$,

$$\mu'_1 = E(X-A) = \bar{X} - A \Rightarrow Mean = \bar{X} = \mu'_1 + A$$

Putting $r = 2$,

$$\mu'_2 = \mu_2 + \mu'^2_1$$

Similarly, we get

$$\mu'_3 = \mu_3 + 3\mu_2\mu'_1 + \mu'^3_1$$

$$\mu'_4 = \mu_4 + 4\mu_3\mu'_1 + 6\mu_2\mu'^2_1 + \mu'^4_1, etc$$

EXAMPLE 2.11 The first four moments of the distribution about the value 4 of the variable are $-1.5, 17, -30$ and 108 respectively. Find the mean and moments about the mean.

Solution Given $\mu'_1 = -1.5, \mu'_2 = 17, \mu'_3 = -30, \mu'_4 = 108$ and $A = 4$

$$Mean = \bar{X} = A + \mu'_1 = 4 + (-1.5) = 2.5$$

The moments about mean are

$$\mu_1 = 0$$

$$\mu_2 = \mu'_2 - \mu'^2_1 = 17 - (-1.5)^2 = 14.75$$

$$\mu_3 = \mu'_3 - 3\mu_2\mu'_1 + 2\mu'^3_1 = -30 + 76.5 - 6.75 = 39.75$$

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'^2_1\mu'_2 - 3\mu'^4_1 = 108 - 180 + 229.5 - 15.1875 = 142.313$$

HOMEWORK

The first four moments of a distribution about $X = 4$ are $1, 4, 10$ and 45 respectively. Show that the mean is 5, variance is 3, $\mu_3 = 0$ and $\mu_4 = 26$.

2.15 MOMENT GENERATING FUNCTION (MGF): $M_X(t)$

The moment generating function of a random variable denoted by $M_X(t)$ is defined as

$$M_X(t) = E(e^{tX})$$

First population moment: $E(X) = \frac{d}{dt} M_X(t)|_{t=0}$

Second population moment: $E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0}$

In general, the kth population moment is: $\boldsymbol{E(X^k) = \frac{d^k}{dt^k} M_X(t)|_{t=0}}$

And

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r'$$

If X is a discrete random variable with PMF $p(x)$, then

$$M_X(t) = E(e^{tX}) = \sum_X e^{tx}\, p(x)$$

If X is a continuous random variable with PDF $f(x)$, then

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Note: Moment generating function is used to calculate the higher moments.

The first four moments of the

which is discrete or continuous, the moments about the origin denoted by μ_r' is defined as

$$\mu_r' = E(X^r), for\ r = 1,2,3, \ldots$$

If X is a continuous random variable with PDF $f(x)$, then

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x)dx, r = 1,2,3, \ldots$$

2.16 THEOREM ON MOMENT GENERATING FUNCTION

THEOREM: $M_{ax}(t) = M_X(at), a$ being a constant.

Proof: By definition

$$M_{ax}(t) = E(e^{taX}) = E(e^{atX})$$

$$\therefore M_{ax}(t) = M_X(at)$$

THEOREM: The moment generating function of the sum of n independent random variables is equal to the product of their respective moment generating functions, i.e.

$$M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(t)M_{X_3}(t) \ldots M_{X_n}(t)$$

Proof: By definition

$$M_{X_1+X_2+X_3+\cdots+X_n}(t) = E[e^{t(X_1+X_2+X_3+\cdots+X_n)}]$$

$$= E(e^{tX_1})E(e^{tX_2})E(e^{tX_3}) \ldots E(e^{tX_n})$$

(since $X_1, X_2, X_3, \ldots, X_n$ are independent)

$$\therefore M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(t)M_{X_3}(t) \ldots M_{X_n}(t).$$

EXAMPLE 2.12 If X represents the outcome when a fair die is tossed, find the MGF of X and hence, find $E(X)$ and $Var(X)$.

Solution: When a fair die is tossed

$$P(X = x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$$

$$M_X(t) = \sum_{x=1}^{6} e^{tx} P(X = x)$$

$$= \frac{1}{6}\sum_{x=1}^{6} e^{tx}$$

$$= \frac{1}{6}(e^t + e^{2t} + e^{3t} + e^{4t} + e^{5t} + e^{6t})$$

$$E(X) = \frac{d}{dt} M_X(t)|_{t=0}$$

$$= \frac{1}{6}(e^t + 2e^{2t} + 3e^{3t} + 4e^{4t} + 5e^{5t} + 6e^{6t})|_{t=0}$$

$$= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2}$$

$$\therefore Mean = E(X) = \frac{7}{2}$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0}$$

$$= \frac{1}{6}(e^t + 4e^{2t} + 9e^{3t} + 16e^{4t} + 25e^{5t} + 36e^{6t})|_{t=0}$$

$$= \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{91}{6} - \frac{49}{4} = \frac{70}{24} = \frac{35}{12}$$

HOMEWORK

1) **Find the MGF of the random variable X whose probability function $P(X = x) = \frac{1}{2^x}, x = 1, 2, 3, \ldots$ Hence find its mean.**
2) **If the moments of a random variable X are defined by $E(X^r) = 0.6, r = 1, 2, \ldots$ show that $P(X = 0) = 0.4, P(X = 1) = 0.6$ and $P(X \geq 2) = 0$.**
3) **Find the first four moments about the origin for a random variable X having density function $f(x) = \frac{4x(9-x^2)}{81}, 0 \leq x \leq 3$.**
4) **Find the MGF of a random variable whose moments are $\mu_r = (r+1)!\,2^r$.**
5) **A random variable X has the PDF given by**

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0 & , x < 0 \end{cases}$$

Find

(i) **The MGF, and**
(ii) **The first four moments about the origin.**

3.1 DISCRETE RANDOM VARIABLES *X* AND *Y*

3.1.1 Joint Probability Mass Function of (X, Y)

- Let X and Y be ***discrete*** random variables. The joint probability distribution function is

$$f(x, y) = P(X = x, Y = y),$$

Where $x \; and \; y$ are possible values of X and Y respectively, f satisfies the following conditions:

1. $p(x, y) \geq 0$ for all x, y;

2.

$$\sum_{y} \sum_{x} f(x, y) = 1$$

.

- Let X and Y be ***continuous*** random variables. The joint probability density function satisfies the following conditions:

1. $f(x, y) \geq 0$ for $-\infty < x < \infty, -\infty < y < \infty$;

2.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dxdy = 1$$

.

3.1.2 Cumulative Distribution Function:

- For any random variables X and Y, the joint (cumulative) distribution function is given by $F(x, y) = P(X \leq x, Y \leq y)$.
- As X and Y are *discrete,*

$$P(a \leq X \leq b, c \leq Y \leq d) = \sum_{c \leq y \leq d} \sum_{a \leq x \leq b} f(x, y)$$

As X and Y are *continuous,*

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_{c}^{d} \int_{a}^{b} f(x, y) \, dxdy$$

- For any random variables X and Y with joint cumulative distribution function $F(x, y)$

1. $F(-\infty, \infty) = F(-\infty, y) = F(x, \infty) = 0$
2. $F(\infty, \infty) = 1$
3. If $b \geq a, d \geq c$ then

$$P(a \le X \le b, c \le Y \le d) = F(b,d) - F(a,d) + F(a,c) - F(b,c)$$

Example 1: Form a collection of 3 white balls, 2 black balls, and 1 red ball, 2 balls is to be randomly selected. Let X denote the number of white balls and Y the number of black balls.

(a) Find the joint probability distribution table of X and Y.
(b) $F(1,1)$ and $F(2,0)$.
Solution:

(a)

$$P(X = i, Y = j) = f(i,j) = \frac{\binom{3}{i}\binom{2}{j}\binom{1}{2-i-j}}{\binom{6}{2}}, 0 \le i,j \le 2; i+j = 1 \text{ or } 2;$$

For example,

$$P(X = 1, Y = 1) = f(1,1) = \frac{\binom{3}{1}\binom{2}{1}\binom{1}{0}}{\binom{6}{2}} = \frac{3 \cdot 2}{15} = \frac{6}{15}.$$

The joint probability distribution table is

	x			
y	0	1	2	Total
0	0	$3/15$	$3/15$	$\mathbf{6/15}$
1	$2/15$	$6/15$	0	$\mathbf{8/15}$
2	$1/15$	0	0	$\mathbf{1/15}$
Total	$\mathbf{3/15}$	$\mathbf{9/15}$	$\mathbf{3/15}$	1

(b) $F(1,1) = P(X_1 \le 1, X_2 \le 1) = f(0,0) + f(0,1) + f(1,0) + f(1,1) = \frac{0+2+3+6}{15} = \frac{11}{15}.$

$$F(2,0) = P(X_1 \le 2, X_2 \le 0) = f(0,0) + f(1,0) + f(2,0) = \frac{0+3+3}{15} = \frac{6}{15}$$

Example 2:

Let $f(x,y) = 2x, 0 \le x \le k; 0 \le y \le 1$; and $f(x,y) = 0$, otherwise.

(a) Find k.
(b) Find $F(0.7,0.5)$, $F(2,0)$ and $F(0.2,3)$.

Solution:

(a)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = \int_{0}^{1}\int_{0}^{k} 2xdx\,dy = \int_{0}^{1} [x^2]_0^k\,dy = \int_{0}^{1} k^2\,dy = k^2 = 1 \Leftrightarrow k = \pm 1$$

$$\Rightarrow k = 1.$$

(b)

$$F(0.7{,}0.5) = P(X \le 0.7, Y \le 0.5) = \int_{-\infty}^{0.5}\int_{-\infty}^{0.7} f(x,y)dxdy = \int_{0}^{0.5}\int_{0}^{0.7} 2xdxdy$$

$$= \int_{0}^{0.5} [x^2]_0^{0.7}\,dy = \int_{0}^{0.5} 0.49\,dy = 0.49 \cdot 0.5 = 0.245 F(2{,}0) = P(X \le 2, Y \le 0)$$

$$= \int_{-\infty}^{0}\int_{-\infty}^{2} f(x,y)\,dxdy = \int_{0}^{0}\int_{0}^{1} 2xdxdy$$

$$= \int_{0}^{0} [x^2]_0^1 dy = \int_{0}^{0} 1dy = 1 \cdot 0 = 0$$

$$F(0.2{,}3) = P(X \le 0.2, Y \le 3) = \int_{-\infty}^{3}\int_{-\infty}^{0.2} f(x,y)dxdy = \int_{0}^{1}\int_{0}^{0.2} 2x\,dxdy$$

$$= \int_{0}^{1} [x^2]_0^{0.2}\,dy = \int_{0}^{1} 0.04\,dy = 0.04$$

3.1.3 Marginal Probability Distribution (Density) Function:

- Let X and Y be ***discrete*** random variables. The marginal probability distribution function for X and Y are

$$f_1(x) = P(X = x) = \sum_y f(x, y),$$

and

$$f_2(y) = P(Y = y) = \sum_x f(x, y),$$

respectively.

- Let X and Y be ***continuous*** random variables. The marginal probability density function for X and Y are

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy,$$

and

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx,$$

respectively.

3.1.4 Conditional Probability Distribution (Density) Function:

- Let X and Y be ***discrete*** random variables. The conditional probability distribution function X given Y is

$$f_1(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x, y)}{f_2(y)},$$

provided that $f_2(y) > 0$. Similarly, the conditional probability distribution function Y given X is

$$f_2(y|x) = P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x,y)}{f_1(x)}$$

,

provided that $f_1(x) > 0$.

- Let X and Y be ***continuous*** random variables. The conditional probability density function X given Y is

$$f_1(x|y) = \begin{cases} \dfrac{f(x,y)}{f_2(y)}, & f_2(y) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Similarly, the conditional probability density function Y given X is

$$f_2(y|x) = \begin{cases} \dfrac{f(x,y)}{f_1(x)}, & f_1(x) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Example 1 (continue): From lecture (13)

(c) Find the marginal probability distribution function of X.

(d) Given one of chosen balls being black, what is the distribution for the number of balls being white?

Solution:

(c)

$$f_1(0) = f(0,0) + f(0,1) + f(0,2) = \frac{0+2+1}{15} = \frac{3}{15}$$

$$f_1(1) = f(1,0) + f(1,1) + f(1,2) = \frac{3+6+0}{15} = \frac{9}{15}$$

$$f_1(2) = f(2,0) + f(2,1) + f(2,2) = \frac{3+0+0}{15} = \frac{3}{15}$$

(d)

$$f_2(1) = f(0,1) + f(1,1) + f(2,1) = \frac{2+6+0}{15} = \frac{8}{15}$$

$$f_1(0|1) = P(X_1 = 0|X_2 = 1) = \frac{f(0,1)}{f_2(1)} = \frac{2/15}{8/15} = \frac{1}{4}$$

$$f_1(1|1) = P(X = 1|Y = 1) = \frac{f(1,1)}{f_2(1)} = \frac{6/15}{8/15} = \frac{3}{4}$$

$$f_1(2|1) = P(X=2|Y=1) = \frac{f(2,1)}{f_2(1)} = \frac{0}{8/15} = 0$$

Example 2 (continue): From lecture (13)

(c) Find the marginal probability density function for X and Y.

(d) Find the conditional probability density function for X given Y and the conditional probability density function for Y given X

Solution:

(c)

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 2x\,dy = 2x, 0 \le x \le 1;$$

.

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^1 2x\,dx = [x^2]_0^1 = 1, 0 \le y \le 1;$$

(d) For $0 \le x \le 1, 0 \le y \le 1$,

$$f_1(x|y) = \frac{f(x,y)}{f_2(y)} = 2x$$

and

$$f_2(y|x) = \frac{f(x,y)}{f_1(x)} = 1$$

.

HOMEWORK

1) **For the bivariate probability distribution of (X,Y) given below, find $P(X \le 1), P(Y \le 3), P(X \le 1, Y \le 3), P(X \le 1|Y \le 3), P(Y \le 3|X \le 1)$ and $P(X+Y \le 4)$.**

X \ Y	1	2	3	4	5	6
0	0	0	$1/32$	$2/32$	$2/32$	$3/32$
1	$1/16$	$1/16$	$1/8$	$1/8$	$1/8$	$1/8$
2	$1/32$	$1/32$	$1/64$	$1/64$	0	$2/64$

2) The joint probability mass function of (X,Y) is given by $P(x,y) = k(2x+3y), x = 0,1,2, y = 1,2,3.$ Find the marginal and conditional distributions for
   (i) $P(X = 2, Y \leq 2)$.
   (ii) $P(X \leq 1, Y = 3)$.
   (iii) $P(X = 2)$.
   (iv) $P(X \leq 2)$.
   (v) $P(X \leq 1 | Y \leq 2)$.
   (vi) $P(X = 0 | Y = 3)$.
3) If the joint PDF of $f(x,y) = k(1-x-y), 0 < x, y < \frac{1}{2}$, find k.
4) Find k if the joint PDF of a bivariate random variable (X,Y) is given by
$$f(x,y) = \begin{cases} k(1-x)(1-y), & if\ 0 < x < 4, 1 < y < 5 \\ 0, & otherwise \end{cases}$$
5) Let X and Y be continuous random variables with joint PDF
$$f(x,y) = \frac{3}{2}(x^2+y^2), 0 < x < 1, 0 < y < 1$$
   Find $f(x|y)$.
6) Let X and Y be continuous random variables with joint PDF
$$f(x,y) = \begin{cases} 2xy + \frac{3y^2}{2}, & 0 < x < 1, 0 < y < 1 \\ 0, & otherwise \end{cases}$$
   Find $P(X+Y < 1)$.

3.1.5 Expectation of Two-dimensional Random Variables

- If (X, Y) is a two-dimensional random variable, then the mean or expectation of (X, Y) is defined as follows

Case 1

When X and Y are discrete random variables, then

$$E(X) = \sum_{x_i} x_i \, p(x_i)$$

$$E(Y) = \sum_{y_i} y_i \, p(y_i)$$

Conditional expected values

$$E(X|Y) = \sum_{x_i} x_i \, p(X|Y)$$

$$E(Y|X) = \sum_{y_i} y_i \, p(Y|X)$$

And

$$E(XY) = \sum_{x_i} \sum_{y_i} x_i y_i \, p(x_i, y_i)$$

Case 2

If X and Y are continuous random variables, then

$$E(X) = \int_{-\infty}^{\infty} x f(x) \, dx$$

$$E(Y) = \int_{-\infty}^{\infty} y f(y) \, dy$$

Conditional expected values

$$E(X|Y) = \int_{-\infty}^{\infty} x \, f(x|y) dx$$

$$E(Y|X) = \int_{-\infty}^{\infty} y \, f(y|x) dy$$

And

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \, f(x, y) \, dxdy$$

Note:

i) If X and Y are independent random variables, then

$$E(X|Y) = E(X)$$

$$E(Y|X) = E(Y)$$

ii) $E[E(X|Y)] = E(X)$

$$E[E(Y|X)] = E(Y)$$

Example: Let X and Y are two random variables each having three values $-1, 0, 1$ and having the following joint probability distribution

		x		
y	-1	0	1	Total
-1	0	0.1	0.1	0.2
0	0.2	0.2	0.2	0.6
1	0	0.1	0.1	0.2
Total	0.2	0.4	0.4	1.0

Prove that X and Y have different expectations. Also prove that X and Y are uncorrelated and find $Var(X)$ and $Var(Y)$.

Solution:

From the table given

$$P(X = -1) = 0.2, P(X = 0) = 0.4, P(X = 1) = 0.4, P(Y = -1) = 0.2, P(Y = 0) = 0.6, P(Y = 1) = 0.2$$

$$E(X) = \sum xp(x) = (-1)(0.2) + (0)(0.4) + (1)(0.4) = 0.2$$

$$E(Y) = \sum yp(y) = (-1)(0.2) + (0)(0.6) + (1)(0.2) = 0$$

$\therefore X$ and Y have different expectations.

$$E(XY) = \sum xy\, p(x,y)$$

$$= (-1)(-1)(0) + (-1)(0)(0.1) + (-1)(1)(0.1) + (0)(-1)(0.2) + (0)(0)(0.2) + (0)(1)(0.2) + (1)(-1)(0) + (1)(0)(0.1) + (1)(1)(0.1) = 0$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - (0.2)(0) = 0$$

$\therefore X$ and Y are uncorrelated.

$$E(X^2) = \sum x^2\, p(x)$$

$$= (-1)^2(0.2) + (0)(0.4) + (1)^2(0.4)$$

$$= 0.2 + 0.4 = 0.6$$

$$E(Y^2) = \sum y^2\, p(y)$$

$$= (-1)^2(0.2) + (0)(0.6) + (1)^2(0.2)$$

$$= 0.2 + 0.2 = 0.4$$

$$Var(X) = E(X^2) - [E(X)]^2 = 0.6 - (0.2)^2 = 0.56$$

$$Var(Y) = E(Y^2) - [E(Y)]^2 = 0.4 - 0 = 0.4$$

Example: If the joint PDF is given by $f(x,y) = x + y, 0 \le x, y \le 1$. Find $E(XY)$.

Solution:

By definition

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dxdy$$

$$= \int_0^1\int_0^1 xy\,(x+y)dxdy$$

$$= \int_0^1\int_0^1 (x^2y + xy^2)\, dxdy$$

$$= \int_0^1 \left[\frac{x^3y}{3} + \frac{x^2y^2}{2}\right]_0^1 dy$$

$$= \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2}\right) dy$$

$$= \left[\frac{y^2}{6} + \frac{y^3}{6}\right]_0^1 = \frac{1}{3}$$

3.1.6 Karl Pearson Coefficient of Correlation

It is a numerical measure of intensity or degree of linear relationship between the two variables. Correlation coefficient between two random variables X and Y denoted by ρ_{XY} or r_{XY} and is defined as

$$r(X,Y) = r_{XY} = \rho(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$$

Where $\sigma_X = \sqrt{Var(X)}$ and $\sigma_Y = \sqrt{Var(Y)}$.

The correlation coefficient lies between -1 and 1 that is $-1 \le \rho \le 1$ and if the two random variables are independent then $Cov(X,Y) = 0$ but the converse need not be true.

HOMEWORK

1) **If $f(x,y) = 2x, 0 \le x \le y \le 1$, find $E(XY)$ and $E(Y)$.**
2) **The joint PDF of (X,Y) is given by**
$$f(x,y) = \begin{cases} 24xy, & x > 0, y > 0, x + y \le 1 \\ 0, & elsewhere \end{cases}$$
Find the conditional mean and variance of Y given X,
3) **The joint PDF of two random variables X and Y is given by**
$$f(y|x) = \frac{9(1+x+y)}{2(1+x)^4(1+y)^4}, 0 \le x < \infty, 0 \le y < \infty$$
Find the marginal distributions of X and Y, the conditional distribution of Y for $X = x$ and the expected value of this conditional distribution.
4) **Let X and Y be any two random variables and a, b be constants. Prove that $Cov(aX, bY) = ab\,Cov(X,Y)$.**
5) **If $Y = -2X + 3$, find $Cov(X,Y)$.**
6) **Two random variables X and Y have joint PDF**
$$f(x,y) = \begin{cases} \frac{xy}{96}, & 0 < x < 4, 1 < y < 5 \\ 0, & otherwise \end{cases}$$
Find $E(X), E(Y), E(XY), E(2X + 3Y), Var(X), Var(Y), Cov(X,Y)$. What can you infer from $Cov(X,Y)$.

4. Discrete Probability Distributions

4.1 Discrete Uniform Distribution

The random variable x is said to have a discrete uniform distribution with parameter n

$x \sim Un(n)$ if it is pmf is given by

$$f(x;n) = \begin{cases} \frac{1}{n}, & x = 1,2,\dots,n \\ 0, & \text{otherwise.} \end{cases}$$

Where n is a parameter and positive integer under this condition. It is clear that $f(x) \geq 0$ and that

$$\sum_{x=1}^{n} f(x;n) = \sum_{x=1}^{n} \frac{1}{n} = \frac{\sum_{x=1}^{n} 1}{n} = \frac{n(1)}{n} = 1$$

That is $f(x)$ is p.m.f and x is discrete random variable.

Mean of a uniform distribution

$$\mu = E(x) = \sum_{x=1}^{n} x\, f(x;n) = \sum_{x=1}^{n} x.\frac{1}{n} = \frac{\sum_{x=1}^{n} x}{n} = \frac{\frac{n(n+1)}{2}}{n} = \frac{n+1}{2}$$

$$E(x^2) = \sum_{x=1}^{n} x^2\, f(x;n) = \sum_{x=1}^{n} x^2.\frac{1}{n} = \frac{\frac{n(n+1)(2n+1)}{6}}{n} = \frac{(n+1)(2n+1)}{6}$$

The variance of a uniform distribution

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

Moment generating function of a uniform distribution

$$M_x(t) = \frac{e^t}{n}\left[\frac{1-e^{nt}}{1-e^{t}}\right].\, derive\ it\ (Homework)$$

Example: let $x \sim Un(100)$. Find $E(x), Var(1), M_x(t), E(x^2)$.

Solution:

Since $x \sim Un(100)$ so $n = 100$ that is mean $f(x) = \frac{1}{100}$

$$E(x) = \frac{n+1}{2} = \frac{101}{2} = 50.5$$

$$E(x^2) = \frac{(n+1)(2n+1)}{6} = \frac{(101)(201)}{6} = 3383.5$$

$$Var(1) = \frac{n^2-1}{12} = \frac{10000-1}{12} = 833.25$$

$$M_x(t) = \frac{e^t}{100}\left[\frac{1-e^{100t}}{1-e^t}\right]$$

Example: let $M_x(t) = \frac{et}{8}\left[\frac{1-e8t}{1-e^t}\right]$. Find $E(x), E(x^2), Var(x)$.

Solution: from the low of moment generating function $n = 8$ that is mean $f(x) = \frac{1}{8}$

$$E(x) = \frac{n+1}{2} = \frac{8+1}{2} = 4.5$$

$$E(x^2) = \frac{(n+1)(2n+1)}{6} = \frac{(8+1)(2(8)+1)}{6} = 25.5$$

$$Var(1) = \frac{n^2-1}{12} = \frac{64-1}{12} = 5.25$$

4.2 Bernoulli Distribution

The random variable x is said to have a Bernoulli distribution $x \sim Ber(p)$ with parameter p if it is p.m.f is given by

$$f(x;p) = p^x q^{1-x}$$

$$\sum f(x;p) = \sum_{x=0}^{1} p^x q^{1-x} = p^0 q^1 + p^1 q^0 = q + p = 1$$

is p.m.f.

Proof $E(x) = p$ in Bernoulli distribution

$$E(x) = \sum_{x=0}^{1} x\, f(x; p) = \sum_{x=0}^{1} x\, p^x q^{1-x} = 0. p^0 q^1 + 1. p^1 q^0 = p$$

Proof $E(x^2) = p$ in Bernoulli distribution

$$E(x^2) = \sum_{x=0}^{1} x^2\, f(x; p) = \sum_{x=0}^{1} x^2\, p^x q^{1-x} = 0^2. p^0 q^1 + 1^2. p^1 q^0 = p$$

Proof $Var(x) = pq$ in Bernoulli distribution

$$Var(x) = E(x^2) - [E(x)]^2 = p - p^2 = p(1-p) = pq \ (\because q = 1-p)$$

The moment generating function of Bernoulli distribution

$$M_x(t) = E(e^{tx}) = \sum e^{tx}\, f(x; p) = \sum_{x=0}^{1} e^{tx} p^x q^{1-x} = e^{0t} p^0 q^1 + e^t p q^0 = q + e^t p$$

Example: let $x \sim Ber(0.6)$. Find $f(x; p), E(x), E(x^2), Var(x), M_x(t)$.

Solution:

$$f(x; p) = p^x q^{1-x}$$

$p = 0.6, q = 1 - 0.6 = 0.4.$

$$\therefore f(x; p) = (0.6)^x (0.4)^{1-x}$$

$$E(x) = E(x^2) = p = 0.6$$

$$Var(x) = pq = (0.6)(0.4) = 0.24$$

$$M_x(t) = q + pe^t = 0.4 + 0.6e^t$$

Example: let $M_x(t) = \frac{1}{2} + \frac{1}{2} e^t$. Find $f(x; p), Var(x)$.

Solution:

$$f(x; p) = p^x q^{1-x} \quad (p = \frac{1}{2}, q = \frac{1}{2})$$

$$f(x; p) = (\frac{1}{2})^x (\frac{1}{2})^{1-x}$$

$$Var(x) = pq = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$

Example: let $x \sim Ber\left(\frac{1}{3}\right)$. Find $f(x;p), Var(x), M_x(t)$.

Solution:

$$f(x;p) = p^x q^{1-x}$$

$$p = \frac{1}{3}, q = 1 - \frac{1}{3} = \frac{2}{3}$$

$$f(x;p) = \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{1-x}$$

$$Var(x) = pq = \frac{1}{3}.\frac{2}{3} = \frac{2}{9}$$

$$M_x(t) = q + pe^t = \frac{2}{3} + \frac{1}{3}e^t$$

Example: let $x \sim Ber(0.3)$. Find $Var(x), M_x(t), E(x), E(x^2)$.

Solution:

$$Var(x) = pq = (0.3)(0.7) = 0.21$$

$$M_x(t) = q + pe^t = 0.7 + 0.3e^t$$

$$E(x) = E(x^2) = p = 0.3$$

4.3 Binomial Distribution

The probability mass function of the binomial distribution for random variable x is

$$f(x; n, p) = C_x^n p^x q^{n-x}, x = 0,1,2, \dots, n, \; x \sim Bin(n, p); q = 1 - p$$

x: number of successes

$n - x$: number of failures

n: number of trials (sample space)

$$C_x^n = \frac{n!}{x!\,(n-x)!}$$

Example: ***Calculate the chances (probability) of getting exactly two heads (in any order) on three tosses of a fair coin.***

Solution: We can use the above binomial formula to calculate desired probability. For this we can express the values as follows:

$p =$ characteristic probability or probability of success $= 0.5$
$q = (1 - \mathrm{p}) =$ probability of failure $= 0.5$
$x =$ number of successes desired $= 2$
$n =$ number of trials undertaken $= 3$

$$\text{Probability of 2 successes (heads) in 3 trials} = \frac{3!}{2!\,(3-2)!} 0.5^2 0.5^{(3-2)}$$

$$= \frac{3 \times 2 \times 1}{(2 \times 1)(1 \times 1)} 0.5^2 0.5^1$$

$$= 3 \times 0.25 \times 0.5 = 0.375$$

Thus, there is a 0.375 probability of getting two heads on three tosses of a fair coin.

Proof $E(x) = np$ in Binomial Distribution

$$E(x) = \sum x\, p(x) = \sum x\, f(x; n, p) = \sum x\, C_x^n p^x q^{n-x}$$

$$= \sum x \frac{n!}{x!\,(n-x)!} p^{x+1-1} q^{n-x}$$

$$= p \sum x \frac{n(n-1)!}{x(x-1)!\,(n-x)!} p^{x-1} q^{n-x}$$

$$= np \sum \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} q^{n-x}$$

Let $y = x - 1$ and let $n - 1 = m$

$x = y + 1,\ n - x = n - y - 1 = n - 1 - y = m - y$

$$= np \sum \frac{m!}{y!\,(m-y)!} p^{y} q^{m-y}$$

$$= np \sum C_y^m p^y q^{m-y} \ \ by\ binomial\ theorem$$

$$= np(p+q)^m$$

$$= np.1$$

$$\therefore E(x) = np$$

Proof $E(x^2) = n(n-1)p^2 + np$ in Binomial Distribution (Homework)

Proof $Var(x) = npq$ in Binomial Distribution

$$Var(x) = E(x^2) - [E(x)]^2 \quad \dots\dots\dots(1)$$

We have $E(x^2) = n(n-1)p^2 + np$ ……… (2)

And $E(x) = np$ ……… (3)

Now, we substitute (2) and (3) in (1)

$$Var(x) = n(n-1)p^2 + np - n^2p^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= -np^2 + np$$

$$= np - np^2$$

$$= np(1-p)$$

$$= npq$$

Moment generating function of Binomial Distribution

$$M_x(t) = E(e^{tx}) = \sum e^{tx} f(x; n, p)$$

$$= \sum e^{tx} C_x^n p^x q^{n-x}$$

$$= \sum C_x^n (e^t p)^x q^{n-x}$$

By using Binomial theorem

$$M_x(t) = (q + pe^t)^n$$

Homework: prove that $Var(x) = npq$ by using moment generating function.

Example: if $M_x(t) = \left(\frac{1}{2} + \frac{1}{2}e^t\right)^5$ find $f(x; n; p), E(x), E(x^2), Var(x)$.

Solution:

$$M_x(t) = (q + pe^t)^n$$

$$q = \frac{1}{2}, p = \frac{1}{2}, n = 5$$

$$f(x; n, p) = C_x^n p^x q^{n-x}$$

$$f(x; n, p) = C_x^5 \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x}$$

$$E(x) = np = 5\left(\frac{1}{2}\right) = \frac{5}{2} = 2.5$$

$$E(x^2) = n(n-1)p^2 + np$$

$$= 5(5-1)\left(\frac{1}{2}\right)^2 + 5\left(\frac{1}{2}\right) = \frac{20}{4} + \frac{5}{2} = \frac{10}{2} + \frac{5}{2} = \frac{15}{2}$$

$$E(x^2) = 7.5$$

$$Var(x) = npq = 5\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{5}{4} = 1.25$$

HOMEWORK

1) **Let $x \sim Bin\left(100, \frac{1}{3}\right)$. find $Var(x), f(x=3), M_x(t)$.**
2) **The probability that a patient recovers from disease is (0.4). if 15 people known this disease. What is probability**
   a) **At least 10 people survive.**
   b) **From 3 to 8.**
   c) **Exactly 5.**
3) **The manufacture indicates that defective rate of the device is 3%; the inspector randomly picks 20 items. What is the probability that three will be at least one defective items among these (20).**
4) **Approximately 10% of all people are left-handed. Consider group of fifteen people. Find the probability**
   a) **Non are left-handed.**
   b) **At least two left-handed.**
   c) **At most three left-handed.**

4.4 Poisson Distribution

The random variable x is said to have a Poisson distribution with parameter λ

$x \sim pos(\lambda)$ if it is pmf is given by

$$f(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!},\ x = 0,1,2,\ldots$$

Is $f(x;\lambda)$ p.m.f?

$$\sum_{x=0}^{\infty} f(x;\lambda) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!} = e^{-\lambda}e^{\lambda} = 1 \qquad \left(e^{x} = \sum_{n=0}^{\infty}\frac{x^{n}}{n!}\right)$$

The Poisson distribution is p.m.f.

Proof $E(x) = \lambda$ in Poisson distribution

$$E(x) = \sum x\, f(x;\lambda) = \sum x\, \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda}\sum x\, \frac{\lambda^{x+1-1}}{x!}$$

$$= e^{-\lambda}\lambda \sum \frac{x.\lambda^{x-1}}{x(x-1)!}$$

Let $x - 1 = y$

$$= e^{-\lambda}\lambda \sum \frac{\lambda^{y}}{y!}$$

$$= e^{-\lambda}\lambda e^{\lambda} = \lambda$$

Proof $E(x^2) = \lambda^2 + \lambda$ in Poisson distribution. (Homework)

Proof $Var(x) = \lambda = E(x)$

$$Var(x) = E(x^2) - [E(x)]^2 \ldots\ldots\ldots (1)$$

We have

$$E(x^2) = \lambda^2 + \lambda \ldots\ldots\ldots (2)$$

And

$$E(x) = \lambda \ldots\ldots\ldots (3)$$

Now, we substitute (2) and (3) in (1)

$$Var(x) = \lambda^2 + \lambda - \lambda^2 = \lambda = E(x)$$

The moment generating function of Poisson Distribution

$$M_x(t) = E(e^{tx}) = \sum e^{tx} f(x;\lambda) = \sum e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum \frac{(e^t\lambda)^x}{x!} = e^{-\lambda}. e^{e^t\lambda}$$

$$M_x(t) = e^{\lambda(e^t-1)}$$

Example: let $\boldsymbol{\lambda = 2}$ in Poisson distribution. Find $\boldsymbol{E(x), E(x^2), Var(x)}$ and find $\boldsymbol{p(x > 0), p(x = 2), p(x < 3)}$

Solution:

$$E(x) = Var(x) = \lambda = 2$$

$$E(x^2) = \lambda^2 + \lambda = 4 + 2 = 6$$

In Poisson distribution $x = 0,1,2, \ldots$

$$p(x > 0) = 1 - p(x \le 0)$$

$$= 1 - \frac{e^{-2}2^0}{0!} = 1 - e^{-2} = 1 - 0.135 = 0.8646$$

$$p(x = 2) = \frac{e^{-2}2^2}{2!} = \frac{4}{2} e^{-2} = 2e^{-2} = 0.27$$

$$p(x < 3) = p(x = 0) + p(x = 1) + p(x = 2)$$

$$= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!}$$

$$= e^{-2} + 2e^{-2} + 2e^{-2}$$

$$= 0.135 + 0.27 + 0.27$$

$$= 0.675$$

HOMEWORK

1) **Suppose the number of accidents occurring weekly on a particular stretch of a highway follow a Poisson distribution with mean 3. Calculate the probability that there is at least one accident this week.**
2) **The PMF of a random variable X is given by**

$$P(X = i) = \frac{C\lambda^i}{C!}, (i = 0, 1, 2, \dots)$$

Find (i) $P(X = 0)$ (ii) $P(X > 2)$

3) **If X and Y are independent Poisson variables such that $P(X = 1) = P(X = 2)$ and $P(Y = 2) = P(Y = 3)$. Find the variance of $(X - 2Y)$.**
4) **If X is a Poisson random variable with $P(X = 1) = P(X = 2)$. Find $P(X \geq 3)$.**
5) **The number of monthly breakdown of a computer is a random variable having a Poisson distribution with mean equal to 1.8. find the probability that this computer will function for a month with (i) only 1 breakdown (ii) at least 1 breakdown.**
6) **The atoms of a radioactive element are randomly disintegrating. If every gram of this element on average emits 3.9 alpha particles per second, what is the probability that during the next second the number of alpha particles emitted from 1 gram is (i) at most 6, (ii) at least 2, and (iii) at least 3 and at most 6.**
7) **A random variable has the moment generating function $e^{4(e^t - 1)}$. Find its mean, standard deviation, $P(\mu - 2\sigma < X < \mu + 2\sigma)$ and $P(\mu - 2\sigma < X < \mu)$.**

4.5 Geometric Distribution

The random variable x has geometric distribution $x \sim G(p)$ if the probability mass function has the form

$$f(x; p, q) = pq^{x-1}, \qquad x = 1,2, \dots$$

Where

p: probability of the first success.

q: probability of the first failure.

x: number of trials.

Is $P(x; p, q)$ p.m.f?

$$\sum_{x=1}^{\infty} f(x; p, q) = \sum_{x=1}^{\infty} pq^{x-1}$$

$$= \sum_{x=1}^{\infty} p\, q^x\, q^{-1}$$

$$= \frac{p}{q} \sum_{x=1}^{\infty} q^x$$

$$= \frac{p}{q}[q + q^2 + q^3 + \cdots]$$

$$= \frac{p}{q}. q[1 + q + q^2 + \cdots] \; Geometric\ series$$

$$= p.\frac{1}{1-q}, \qquad (1 - q = p)$$

$$= p.\frac{1}{p} = 1$$

The Geometric distribution is p.m.f.

Proof $E(x) = \frac{1}{p}$ in Geometric Distribution

$$E(x) = \sum x\, f(x; p, q) = \sum x\, pq^{x-1}$$

$$= p \sum x\, q^{x-1} \qquad \left(\frac{d}{dq} q^x = x q^{x-1}\right)$$

$$= p \sum \frac{d}{dq} q^x$$

$$= p \frac{d}{dq} \sum q^x$$

$$= p \frac{d}{dq} [q + q^2 + q^3 + \cdots]$$

$$= p \frac{d}{dq} q[1 + q + q^2 + \cdots] \quad (geometric\ series)$$

$$= p \frac{d}{dq} q \frac{1}{1-q}$$

$$= p \frac{d}{dq} \frac{q}{1-q}$$

$$= p \frac{(1-q)(1) - (q)(-1)}{(1-q)^2}$$

$$= p \frac{1 - q + q}{(1-q)^2}$$

$$= p\frac{1}{(1-q)^2} \qquad (1-q=p)$$

$$= p\frac{1}{p^2}$$

$$= \frac{1}{p}$$

Proof $E(x^2) = \frac{2q}{p^2} + \frac{1}{p}$ in Geometric Distribution. (Homework)

Proof $Var(x) = \frac{p}{q^2}$ in geometric distribution.

$$Var(x) = E(x^2) - [E(x)]^2 \dots\dots\dots (1)$$

We have

$$E(x^2) = \frac{2q}{p^2} + \frac{1}{p} \dots\dots\dots (2)$$

And

$$E(x) = \frac{1}{p} \dots\dots\dots (3)$$

Now, we substitute (2) and (3) in (1)

$$Var(x) = \frac{2q}{p^2} + \frac{1}{p} - \left(\frac{1}{p}\right)^2$$

$$= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{2q + p + 1}{p^2} \qquad (q = 1-p)$$

$$= \frac{2(1-p)+p+1}{p^2}$$

$$= \frac{2-2p+p+1}{p^2}$$

$$= \frac{1-p}{p^2} \quad (1-p=q)$$

$$= \frac{q}{p^2}$$

$\therefore Var(x) = \frac{q}{p^2}$

Moment generating function of Geometric Distribution

$$M_x(t) = E(e^{tx}) = \sum e^{tx} p q^{x-1}$$

$$= \sum e^{tx} p q^x \, q^{-1}$$

$$= \frac{p}{q} \sum e^{tx} q^x$$

$$= \frac{p}{q} \sum (e^t q)^x$$

$$= \frac{p}{q} [e^t q + e^{2t} q^2 + e^{3t} q^3 + \cdots]$$

$$= \frac{p}{q} e^t q [1 + e^t q + e^{2t} q^2 + \cdots] \qquad (geometric\ series)$$

$$= pe^t \frac{1}{1-e^t q} = \frac{pe^t}{1-e^t q}$$

$$\therefore M_x(t) = \frac{pe^t}{1-e^t q}$$

Example: let $M_x(t) = \frac{0.4e^t}{1-0.6e^t}$. find $f(x;p,q), Var(x), p(x \geq 2), p(x > 5)$

Solution:

$$p = 0.4, q = 0.6$$

$$f(x;p,q) = pq^{x-1}, \quad x = 1,2,\ldots$$

$$f(x;p,q) = (0.4)(0.6)^{x-1}$$

$$Var(x) = \frac{q}{p^2} = \frac{0.6}{(0.4)^2} = \frac{0.6}{0.16} = 3.75$$

$$p(x \geq 2) = \sum_{x=2}^{\infty} pq^{x-1} = \sum_{x=2}^{\infty} (0.4)(0.6)^{x-1}$$

$$= (0.4)(0.6)^1 + (0.4)(0.6)^2 + \cdots$$

$$= (0.4)(0.6)[1 + 0.6 + (0.6)^2 + \cdots] \ (geometric\ series)$$

$$= (0.4)(0.6)\frac{1}{1-0.6}$$

$$= (0.4)(0.6)\frac{1}{(0.4)} = 0.6$$

Another method

$$p(x \geq 2) = 1 - p(x < 2)$$

$$= 1 - p(x = 1)$$

$$= 1 - [pq^0] = 1 - [(0.4)(0.6)^0] = 1 - 0.4 = 0.6$$

$$p(x > 5) = \sum_{x=6}^{\infty} pq^{x-1} = \sum_{x=6}^{\infty} (0.4)(0.6)^{x-1}$$

$$= (0.4)(0.6)^5 + (0.4)(0.6)^6 + (0.4)(0.6)^7 + \cdots$$

$$= (0.4)(0.6)^5[1 + 0.6 + (0.6)^2 + \cdots] \ \ (geometric\ series)$$

$$= (0.4)(0.6)^5 \frac{1}{1 - 0.6}$$

$$= (0.4)(0.6)^5 \frac{1}{(0.4)} = (0.6)^5 = 0.07776$$

Another method

$$p(x > 5) = 1 - p(x \le 5)$$

$$= 1 - [p(x = 1) + p(x = 2) + p(x = 3) + p(x = 4) + p(x = 5)]$$

$$= 1 - [(0.4)(0.6)^0 + (0.4)(0.6)^1 + (0.4)(0.6)^2 + (0.4)(0.6)^3 + (0.4)(0.6)^4]$$

$$= 1 - (0.4)[1 + 0.6 + (0.6)^2 + (0.6)^3 + (0.6)^4]$$

$$= 1 - (0.4)(2.3056)$$

$$= 1 - 0.92224 = 0.07776$$

HOMEWORK

1) **Let $x \sim G\left(\frac{1}{4}\right)$. find $p(x > 3), p(x \le 2), E(x), E(x^2), Var(x), M_x(t)$.**
2) **If the probability that a certain kind of measuring device will show excessive drift is 0.05, what is the probability that the sixth of these measuring devices tested will be the first to show excessive drift?**
3) **If the probability that the target is destroyed on any one shot is 0.5, what is the probability that it will be destroyed on the sixth attempt.**
4) **If the probability that an applicant for a driver's license will pass road test on any given trial is 0.8, what is the probability that he will finally pass the test (i) on the fourth trial, (ii) fewer than 4 trials.**

4.6 Negative Binomial Distribution

The random variable x has negative binomial distribution $x \sim N\,Bin(r,p)$ if the probability mass function has the form

$$f(x;r,p) = C_x^{x+r-1} p^r q^x, \qquad x = 0,1,2,\ldots$$

$$E(x) = \frac{rq}{p}, Var(x) = \frac{rq}{p^2}$$

Moment generating function of Negative Binomial Distribution

$$M_x(t) = \left(\frac{p}{1-qe^t}\right)^r$$

Derive it. (Homework)

4.7 Hypergeometric Distribution

The random variable x has hepergeometric distribution $x \sim HG(N,m,n)$ if the probability mass function has the form

$$f(x;m,n,N) = \frac{C_x^m C_{n-x}^{N-m}}{C_n^N}$$

N: population size.

m: number of successes in population.

n: sample size.

x: number of successes in the sample.

$$E(x) = \frac{nm}{N}, \qquad Var(x) = \frac{nm}{N}\left(\frac{N-n}{N-1}\right)\left(\frac{N-m}{N}\right)$$

Example: A box contains 4 red and 10 blue balls, five balls are drawn at random without replacement from this box. What is the probability that two red balls are drawn?

Solution:

$$N = 4 + 10 = 14, \quad n = 5, \quad x = 2, \quad m = 4$$

$$p(x = 2) = \frac{C_2^4 C_{5-2}^{14-4}}{C_5^{14}} = 0.35964$$

HOMEWORK

1) **If $x \sim HG(14, 4, 5)$. find $p(x \geq 2), E(x), Var(x)$.**
2) **A deck of cards contains 20 cards, 6 red and 14 black. 5 cards are drawn random without replacement. What is probability that exactly 4 red card are drawn.**
3) **We have 20 televisions. 8 of which are defective, a sample is selected from it. Find the probability that $i)$ $one\ defective$ $ii)$ $at\ least\ 4\ defective.$**

5. Continuous Probability Distributions

5.1 Continuous Uniform Distribution

The random variable x is said to have uniform distribution $x \sim Un(a,b)$ if the probability mass function has the form

$$f(x;a,b) = \frac{1}{b-a}, a < x < b$$

Is the continuous uniform distribution p.d.f?

$$\int_a^b f(x;a,b)\,dx = \int_a^b \frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b dx$$

$$= \frac{1}{b-a}[x]_a^b = \frac{1}{b-a}.(b-a) = 1$$

Proof of $E(x) = \frac{b+a}{2}$ in continuous uniform distribution

$$E(x) = \int_a^b x\, f(x;a,b)\,dx$$

$$= \int_a^b x\,\frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b x\,dx$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{2(b-a)}(b^2-a^2)$$

$$= \frac{1}{2(b-a)}(b-a)(b+a) = \frac{b+a}{2}$$

$$E(x) = \frac{rq}{p}, Var(x) = \frac{rq}{p^2}$$

Proof of $E(x^2) = \frac{b^2+ab+a^2}{3}$ in continuous uniform distribution. (Homework)

Proof of $Var(x) = \frac{(b-a)^2}{12}$ in continuous uniform distribution

We have

$$E(x) = \frac{b+a}{2} \quad \dots\dots\dots(1)$$

$$E(x^2) = \frac{b^2+ab+a^2}{3} \quad \dots\dots\dots(2)$$

Now, substitute (1) and (2) in

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} = \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4}$$

$$= \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

Moment generating function of continuous uniform distribution

$$M_x(t) = E(e^{tx}) = \int_a^b e^{tx}\frac{1}{b-a}dx$$

$$= \frac{1}{b-a}\int_a^b e^{tx}\,dx$$

$$= \frac{1}{b-a}\left(\frac{1}{t}e^{tx}\right)_a^b$$

$$= \frac{1}{t(b-a)}(e^{bt} - e^{at})$$

$$M_x(t) = \frac{e^{bt}-e^{at}}{t(b-a)}$$

Example: let $0 < x < 30$. Find $f(x), E(x), Var(x)$ by using uniform distribution.

Solution:

Since $x \sim Un(0,30)$ so $a = 0, b = 30$

$$f(x) = \frac{1}{b-a} = \frac{1}{30-0}$$

$\therefore f(x) = \frac{1}{30}$

$$E(x) = \frac{b+a}{2} = \frac{30}{2}$$

$\therefore E(x) = 15$

$$Var(x) = \frac{(b-a)^2}{12} = \frac{900}{12}$$

$\therefore Var(x) = 75$

HOMEWORK

1) **Let $0 < x < 1$, find $f(x), E(x), Var(x), M_x(t)$ by using uniform distribution.**
2) **Let $-0.5 < x < 0.5$, find $f(x), E(x), Var(x), M_x(t)$ by using uniform distribution.**
3) **Let $x \sim Un(6,8)$, find $f(x), E(x), Var(x), M_x(t)$.**
4) **Let $x \sim Un(2,8)$, find $f(x), E(x), Var(x), M_x(t)$.**

5.2 Exponential Distribution

The random variable x has exponential distribution with parameter θ, $x \sim Exp(\theta)$

$$f(x;\theta) = \theta e^{-\theta x}$$

$$E(x) = \frac{1}{\theta}, Var(x) = \frac{1}{\theta^2}, M_x(t) = \frac{\theta}{\theta - t}$$

$(derivation\ of\ above\ laws\ is\ Homework)$

Example: If X has an exponential distribution with mean 2. Find $P(X < 1|X < 2)$.

Solution:

We know that mean of the exponential distribution is $\frac{1}{\theta}$

$\therefore$ Mean$= 2 \Longrightarrow \frac{1}{\theta} = 2 \Longrightarrow \theta = \frac{1}{2} = 0.5$

$$PDF = f(x) = 0.5e^{-(0.5)x}, x \geq 0$$

$$P(X < 1|X < 2) = \frac{P(X < 1 \cap X < 2)}{P(X < 2)} = \frac{P(X < 1)}{P(X < 2)}$$

$$P(X < 1) = \int_{-\infty}^{1} f(x)\, dx = \int_{0}^{1} 0.5\, e^{-0.5x} dx$$

$$= [-e^{-0.5x}]_0^1 = 0.393$$

$$P(X < 2) = \int_{0}^{2} f(x)\, dx = \int_{0}^{2} 0.5e^{-0.5x} dx$$

$$= [-e^{-0.5x}]_0^2 = 0.632$$

$$P(X < 1|X < 2) = \frac{P(X < 1)}{P(X < 2)} = \frac{0.393}{0.692} = 0.5679$$

Homework

1) **If X follows an exponential distribution with $P(X \leq 1) = P(X > 1)$, find the mean and variance.**
2) **The time (in hours) required to repair a watch is exponentially distributed with parameter $\theta = \frac{1}{2}$, (i) what is the probability that the repair time exceeds 2 hours? (ii) what is the probability that a repair takes 11 hours given that its duration exceeds 8 hours?**

5.3 Normal Distribution

The random variable x is said to have normal distribution $X\sim N(\mu,\sigma^2)$ if the probability density function has the form

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty$$

$$E(x) = \mu,\ Var(x) = \sigma^2$$

Moment generating function of normal distribution

$$M_x(t) = e^{\mu t+\frac{\sigma^2 t^2}{2}}$$

Standard Normal Distribution

We use this distribution to evaluate the value of $Z = \frac{X-\mu}{\sigma}$ which exists in normal distribution when $\mu = 0, \sigma^2 = 1$

$$Z = \frac{X-\mu}{\sigma}, x\sim N(0,1)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2}$$

$$E(x) = \mu = 0,\ \ Var(x) = \sigma^2 = 1, M_x(t) = e^{\frac{t^2}{2}}$$

Probabilities of $X\sim N(\mu,\sigma)$

- $\boldsymbol{X\sim N(\mu,\sigma) \Leftrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)}$
- $\boldsymbol{X \leq a \Leftrightarrow \frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma} \Leftrightarrow Z \leq \frac{a-\mu}{\sigma}}$
- $\boldsymbol{P(X \leq a) = P\left(Z \leq \frac{a-\mu}{\sigma}\right)}$

- $\boldsymbol{P(X \geq a) = 1 - P(X \leq a) = 1 - P\left(Z \leq \frac{a-\mu}{\sigma}\right)}$
- $\boldsymbol{P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)}$
- $\boldsymbol{P(X = a) = 0}$ **for every** $\boldsymbol{a}$**.**
- $\boldsymbol{P(X \leq \mu) = P(X \geq \mu) = 0.5}$

Example: let $x \sim N(0, 1)$. Find $p(x \leq -1.72)$.

Solution:

$$p(x \leq -1.72) = 1 - p(x < 1.72)$$

$$= 1 - 0.9573 = 0.0427$$

Example: X is a normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$ Find

$a)$ $P(x < 40)$

$b)$ $P(x > 21)$

$c)$ $P(30 < x < 35)$

Solution:

a)

$$P(x < 40) = P\left(\frac{x-\mu}{\sigma} < \frac{40-\mu}{\sigma}\right) = P\left(z < \frac{40-30}{4}\right) = P(z < 2.5) = 0.9938$$

b)

$$P(x > 21) = P\left(\frac{x-\mu}{\sigma} > \frac{21-\mu}{\sigma}\right) = P\left(z > \frac{21-30}{4}\right)$$

$$= P(z > -2.25) = 1 - P(z < -2.25) = 1 - 0.0122 = 0.9878$$

c)

$$P(30 < x < 35) = P\left(\frac{30-\mu}{\sigma} < \frac{x-\mu}{\sigma} < \frac{35-\mu}{\sigma}\right) = P\left(\frac{30-30}{4} < z < \frac{35-30}{4}\right)$$

$$= P(0 < z < 1.25) = P(z < 1.25) - P(z > 0)$$

$$= 0.8944 - 0.5 = 0.3944$$

Notation:

$$\boldsymbol{P(Z \geq Z_A) = A}$$

Result:

$$\boldsymbol{Z_A = -Z_{1-A}}$$

Example: let $Z \sim N(0,1)$ then

$$P(Z \geq Z_{0.025}) = 0.025$$

$$P(Z \geq Z_{0.95}) = 0.95$$

$$P(Z \geq Z_{0.90}) = 0.90$$

Example: let $Z \sim N(0,1)$ then

$$Z_{0.025} = 1.96$$

$$Z_{0.95} = -1.645$$

$$Z_{0.90} = -1.285$$

$$= 1 - 0.9573 = 0.0427$$

HOMEWORK

1) **Let $x \sim N(10, 4)$. Find $P(2 < x < 6)$**

2) **Let $x \sim N(4, 25)$. Find $P(x < 6), P(4 < x < 5)$.**

3) **A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?**

4) **For a certain type of computers, the length of time bewteen charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.**

5) **Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university?**

6) **The length of similar components produced by a company are approximated by a normal distribution model with a mean of 5 cm and a standard deviation of 0.02 cm. If a component is chosen at random**

   a) what is the probability that the length of this component is between 4.98 and 5.02 cm?

   b) what is the probability that the length of this component is between 4.96 and 5.04 cm?

7) **The length of life of an instrument produced by a machine has a normal distribution with a mean of 12 months and standard deviation of 2 months. Find the probability that an instrument produced by this machine will last**

a) less than 7 months.

b) between 7 and 12 months.

8) **The annual salaries of employees in a large company are approximately normally distributed with a mean of $50,000 and a standard deviation of $20,000.**

a) What percent of people earn less than $40,000?

b) What percent of people earn between $45,000 and $65,000?

c) What percent of people earn more than $70,000?

9) **The time taken to assemble a car in a certain plant is a random variable having a normal distribution of 20 hours and a standard deviation of 2 hours. What is the probability that a car can be assembled at this plant in a period of time**

a) less than 19.5 hours?

b) between 20 and 22 hours?