
STA247

Pierre's Guide

Fall 2022

Contents

Chapter 1	Introduction to Probability	Page 2
1.1	Sample Space	2
	Events — 3 • Basic Operations on Events — 3	
1.2	Introduction to Probability	4
1.3	The Inclusion-Exclusion Principle	5
1.4	Probability as Relative Frequency	6
1.5	Conditional Probability	6
	Closed Form of Conditional Probability — 7 • Law of Total Probability — 7	
1.6	Bayes' Rule	7
1.7	Independent Events	8
Chapter 2	Counting	Page 10
2.1	Fundamental Principle of Counting	10
2.2	Permutations	11
2.3	Indistinguishable Objects	11
2.4	Combinations	12
Chapter 3	Random Variables	Page 13
3.1	Introduction to Random Variables	13
3.2	PMF of Discrete Random Variables	14
3.3	Characteristics of Random Variables	14
	Expected Values — 14 • Variance — 15 • Properties of Expectation and Variance — 15	
3.4	Cumulative Distribution Function	16

Chapter 1

Introduction to Probability

Probability encompasses being able to predict the likelihood of an event occurring. To understand what this means we must understand how to describe the domain of these probabilities. We call these, “Random Experiments.”

Definition 1.0.1: Random Experiment

A process with trackable outcomes, often called an experiment. This process is repeatable. The possible outcomes are known, yet the outcome of a specific experiment is not known

Example 1.0.1

- Rolling five dice and observing their sum of top-facing numbers.
- Choosing a card from a standard deck and observing its suite.
- Observing the time between lightning strikes.

1.1 Sample Space

Events are how we categorize observations. The sample space is the category containing only all the possible outcomes of a Random Experiment. Formally defined as follows:

Definition 1.1.1: Sample Space

The set of all possible/outcomes and results from a random experiment. Often denoted was Ω or S . The elements in the samples space are **are determined by the outcome of interest**.

Example 1.1.1

Example Smaple Spaces of the previous example:

- The minimum value of the sum of five dice is 5, and the maximum is $6 \cdot 5 = 30$. So we have that $S = [5, 30] \cap \mathbb{N}$
- The 4 different suites are all the elements of S
- The time between two things happening can be represented in seconds. Since this value is continuous

we have that $S = \mathbb{R}^+$

Note:-

In STA247 Ω is used quite often in definitions while S is often used in questions.

1.1.1 Events

The definition of an event follows that of sample space:

Definition 1.1.2: Event

An event is a subset of the sample space. A **simple event** is an event that includes only one element of the sample space. Conversely, a **compound event** is an event consisting of multiple elements.

Note:-

Instead of directly defining some set of simple observations, events are usually described with some statement directly addressing some subset of the sample space. An example being an experiment on the number of a rolled dice, and the event being an even parity.

1.1.2 Basic Operations on Events

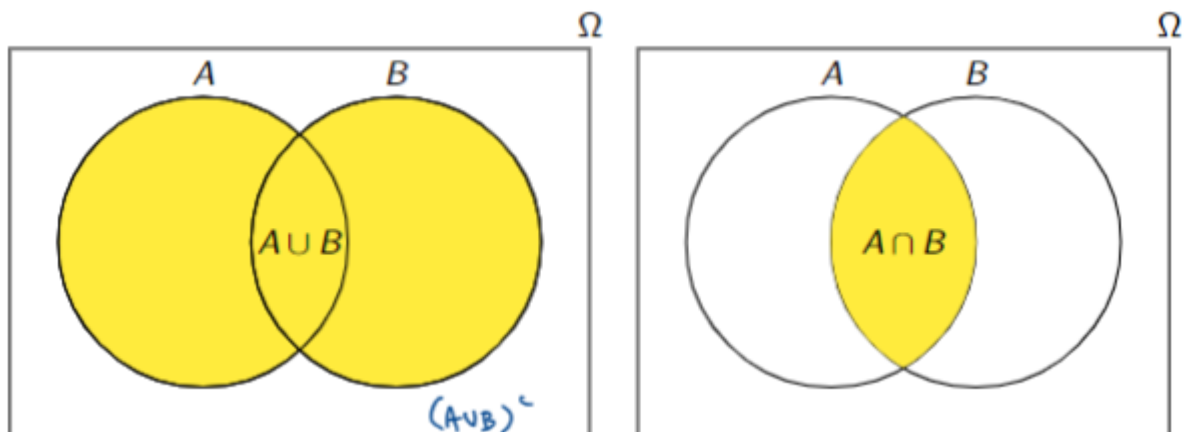
The basic operations on events usually encompass their set operations. However the most important operations, complement, relates to an events property of being a subset. The basic operations are as follows:

Complement: The complement of an event A , denoted A^C , is the event where none of the observations of A occur. This can be also thought of as $A^C = S - A$, but is not often expressed this way.

Intersection: The intersection of two events, A and B , is the set of outcomes that are both in A and B . This is expressed as $A \cap B$ and often just AB .

Union: The union of two events, A and B , is the set of outcomes that are in either A or B , including those which are in both. This is denoted as $A \cup B$.

This can be visualized diagrammatically with venn diagrams:



The definition of the intersection of events brings us to an important terminology of events:

Definition 1.1.3: Mutually Exclusive

We say two events are **mutually exclusive** if $A \cap B = \emptyset$. Meaning the events can not occur simultaneously as an outcome of the random experiment. Synonymously denoted **disjoint** events.

The following laws of set relationships are useful when simplifying probabilities in the future:

Commutative Law : $A \cup B = B \cup A$

Associative Law : $(A \cup B) \cup C = A \cup (B \cup C)$ (Same for intersection)

Distributive Law : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

DeMorgan's Laws of Union : $(A \cup B)^C = A^C \cap B^C$

DeMorgan's Laws of Intersection : $(A \cap B)^C = A^C \cup B^C$

1.2 Introduction to Probability

Probability is a mathematical discipline that studies the likelihood of events occurring. It is used to model and analyze random phenomena, such as the roll of a dice or the outcome of a coin toss. Probability is defined as follows:

Definition 1.2.1: Probability

For any random experiment with sample space S , and event $A \subseteq S$, the **probability** of event A is denoted as $P(A)$. Where $P : S \rightarrow [0, 1]$, is a function which assigns a value to an event. This value corresponds to the chance that an event A occurs

Probability theory is based on a set of axioms, or fundamental principles, that provide a consistent framework for calculating probabilities and making predictions. In this section, we will delve into the axioms of probability and their implications for the mathematical treatment of random events.

The axioms of probability specify the rules for assigning probabilities to events. They are designed to capture the basic properties of probability, such as the fact that the probability of an event occurring must be a number between 0 and 1, inclusive. The axioms also provide a foundation for more advanced concepts in probability theory, such as conditional probability and independence. By studying the axioms of probability, we can gain a better understanding of the mathematical principles underlying the analysis of random events.

Definition 1.2.2: The Three Axioms of Probability

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. For a set of disjoint events, A_1, A_2, \dots, A_n , in Ω :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

Example 1.2.1 (Using the Three Axioms in a Proof)

Proof. $P(A) = 1 - P(A^C)$ By the definition of complement we have $A^C \cup A = \Omega$. Following this we get

$P(A \cup A^C) = P(\Omega)$. Using the third axiom we can get that $P(A) + P(A^C) = P(\Omega) \rightarrow P(A) = P(\Omega) - P(A^C)$. Finally using the second axiom of probability we have that $P(A) = 1 - P(A^C)$. \odot

1.3 The Inclusion-Exclusion Principle

when we are dealing with events that are not mutually exclusive, the basic principle of counting does not apply and we need a more sophisticated approach. This is where the inclusion-exclusion principle comes in.

The inclusion-exclusion principle is a powerful tool that allows us to calculate the probability of the union of two or more events in a precise and efficient way. It is based on the idea of "inclusion" and "exclusion" of elements, and it states that if we have a collection of events, then the probability of the union of these events is equal to the sum of the probabilities of the individual events, minus the probability of the intersection of any two events, plus the probability of the intersection of any three events, and so on. This alternating pattern of inclusion and exclusion allows us to account for the overcounting and undercounting of elements that occurs when we simply add up the probabilities of the individual events.

Definition 1.3.1: Inclusion-Exclusion Principle

For any set of events, A_1, A_2, \dots, A_n , the probability of any of these events occurring is:

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) \\ &\quad - \sum_{i < j} P(A_i \cap A_j) \\ &\quad + \dots \\ &\quad + (-1)^{r+1} \sum_{i_1 < i_2 < \dots < i_r} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_r}) \\ &\quad + \dots \\ &\quad + (-1)^{n+1} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}) \end{aligned}$$

Example 1.3.1

Suppose we have two events, A and B , with probabilities $P(A) = 0.4$ and $P(B) = 0.6$. We want to find the probability of the union of these events, which we can denote as $P(A \cup B)$.

Using the inclusion-exclusion principle, we can write:

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

To calculate the probability of the intersection of A and B , $P(A \cap B)$, we need to know whether the events are mutually exclusive or not. If they are mutually exclusive, then $P(A \cap B) = 0$, and the probability of the union is simply the sum of the probabilities of the individual events:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.6 - 0 = 1.0$$

However, if the events are not mutually exclusive, then we need to know the probability of the intersection in order to calculate the probability of the union. For example, suppose $P(A \cap B) = 0.2$. Then the probability of the union is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.6 - 0.2 = 0.8$$

This shows that the probability of the union of two events is equal to the sum of the probabilities of the individual events, minus the probability of the intersection of the events

1.4 Probability as Relative Frequency

This final section prefaces counting and relates to the probability of some event occurring, given the probability of any outcome is equally likely to any other outcome. This can be summarized with the following

Definition 1.4.1: Probability as Relative Frequency

If the probability of all outcomes in Ω are equal, then the probability of any event is the ratio of the size of the event to the size of Ω . Formally described as: IF: $\forall e_1, e_2 \in \Omega. P(e_1) = P(e_2)$
THEN: For all subsets of Ω , A :

$$P(A) = \frac{|A|}{|\Omega|}$$

Note:-

Relative Frequency can also mean the number of outcomes of A being true given some number of trials. In this case the relative frequency of A is the ratio of the number of occurrences of A and the number of trials.

1.5 Conditional Probability

Considering different events allows us to have cases where two events occur. From the section on intersection, given two events, A and B , they are considered to have occurred at the same time if any outcome from $A \cap B$ has occurred. Sometimes we want to consider the probability that an event has occurred, given that we know that another event has occurred. This concept may seem abstract; however, it is simplified by understanding the probability of an event occurring given we know another event occurred by just imagining the probability of an event occurring given a new sample space. This new sample space is all the outcomes of the event we know occurred. This concept is called conditional probability.

Definition 1.5.1: Conditional Probability

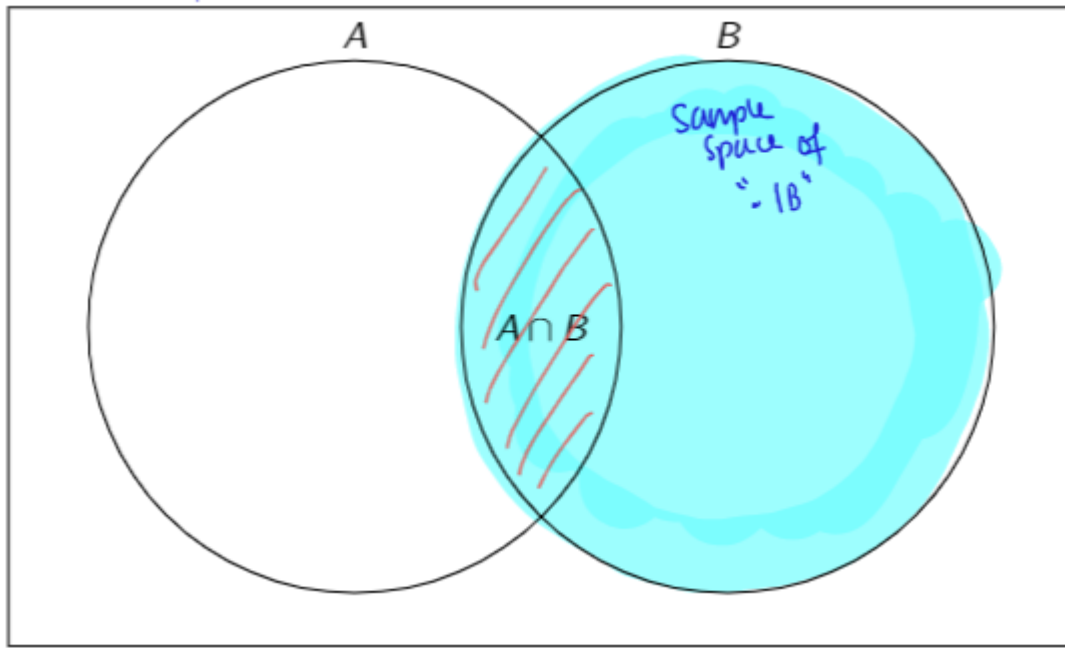
Let there be a random experiment with sample space Ω . Let A, B be events, $A \subseteq \Omega \wedge B \subseteq \Omega$, where $P(B) > 0$. We denote the conditional probability of A occurring given B as:

$$P(A|B)$$

We define **conditional probability** with the following:

1. $\forall b \in B. P(b|B) \geq 0$
2. $\forall b_\omega \in \Omega - B. P(b_\omega|B) = 0$
3. $\sum_{b \in B} P(b|B) = 1$
4. $P(A|B) = \sum_{a \in A} P(a|B)$

We can envision this concept as a venn diagram. The shade blue area is the sample space of the conditional probability, and the crossed area is the event $A|B$. $P(A|B)$ can be envisioned with:



1.5.1 Closed Form of Conditional Probability

The definition of conditional probability lets us express it with the following:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ where } P(B) > 0$$

1.5.2 Law of Total Probability

The law of total probability allows us to break down probabilities with conditional probabilities so that, sometimes, they are easier to work with.

Definition 1.5.2: Law of Total Probability

If B_1, B_2, \dots, B_k is a collection of mutually exclusive and exhaustive partition of the sample space, then for any event A we have:

$$P(A) = \sum_{i=1}^k P(A|B_i) \cdot P(B_i)$$

1.6 Bayes' Rule

Using the Law of Total Probability with our definition of conditional probability we can expand and use Bayes' Rule:

Definition 1.6.1: Bayes' Rule

Let B_1, B_2, \dots, B_k form partition of the sample space. Let $A \subseteq \Omega$ be an event. Bayes' Rule states:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{i=1}^k P(A|B_i) \cdot P(B_i)}$$

This rule is very abstract and initially confusing on how to use. However, it is very practical once understood.

Example 1.6.1

A ball is drawn at random from an urn containing one red and one white ball. If the white ball is drawn, it is put back into the urn. If the red ball is drawn, it is returned to the urn together with two more red balls. Then a second draw is made.

First we will consider the probability that the red ball was drawn on both draws. We can represent this with two events, R_1 and R_2 , the event that the red ball is drawn on the first draw and the event that the red ball is drawn on the second draw. We can represent our final probability as $P(R_1) \cdot P(R_2|R_1) = \frac{1}{2} + \frac{3}{4} = \frac{3}{8}$. This is easy to consider and we do not need to expand the definition of a conditional probability to understand it.

Consider the probability of the first ball being chosen red given we know the second ball was red. This is represented as $P(R_1|R_2)$. We can understand this by expanding the definition and using Bayes' Rule:

$$\begin{aligned} P(R_1|R_2) &= \frac{P(R_1 \cap R_2)}{P(R_2)} \\ &= \frac{P(R_1 \cap R_2)}{P(R_2|(\Omega - R_1)) \cdot P(\Omega - R_1) + P(R_2|R_1) \cdot P(R_1)} \\ &= \frac{\frac{3}{8}}{\frac{1}{4} + \frac{3}{8}} \\ &= \frac{3}{5} \end{aligned}$$

1.7 Independent Events

The definition of the independence of an event can be defined and better understood using conditional probability. The informal definition of two events being independent is: given we know one event has occurred, the probability of the other occurring has not changed. Expressed with conditional probability:

Definition 1.7.1: Independent Events

Let A and B be two events of some random experiment. A and B being independent is equivalent to saying:

$$P(A|B) = P(A) \wedge P(B|A) = P(B)$$

This assumes that $P(A)$ and $P(B)$ are both greater than 0. An equivalent statement is:

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

If A and B are not independent we consider them dependent.

This definition extends to multiple events. definition extends to another concept called mutual independence:

Definition 1.7.2: Independent Collection of Events

For a collection of n events, A_1, A_2, \dots, A_n , we consider they are independent if:

$$P(A_1 \cap \dots \cap A_n) = P()$$

Note:-

We consider a collection events **mutually independent** if any subset of the events are independent

Chapter 2

Counting

Counting is a fundamental concept in probability, as it allows us to determine the number of possible outcomes in a given probability experiment. Understanding how to count the number of possible outcomes is essential for calculating the likelihood of different events occurring.

Counting also plays a crucial role in understanding the relationships between different probability events. For example, knowing how to count the number of possible outcomes can help us understand how the probability of one event might be affected by the occurrence of another event.

In addition to its practical applications, counting is also a fundamental mathematical concept that is closely related to other areas of mathematics, such as combinatorics and graph theory. As such, a strong understanding of counting principles is important for anyone interested in pursuing a career in mathematics or a related field.

2.1 Fundamental Principle of Counting

When applying probability as a relative frequency it is important to be able to understand the number of possible outcomes in some experiment. Often we represent experiments as more complicated versions of others. One way we compliment experiments is adding stages. Meaning that we repeat the same experiment, or others, multiple times and record the combined outcome of every stage.

Definition 2.1.1: Fundamental Principle of Counting

If an experiment consists of m (ordered) stages with n_1 possible outcomes in stage 1, n_2 possible outcomes in stage 2, \dots , n_m possible outcomes in stage m , then the total number of possible outcomes is:

$$\prod_{i=1}^m n_i = n_1 \cdot n_2 \cdot \dots \cdot n_m$$

Example 2.1.1

If there are m ways to perform task 1 and n ways to perform task 2, then there are $m \times n$ ways to perform both tasks.

This principle is often used in combinatorics and probability to determine the number of possible outcomes in a given situation. For example, if you have 3 choices for one task and 4 choices for another task, then there are $3 \times 4 = 12$ possible combinations of the two tasks.

Note:-

The Fundamental Principle of Counting is specific to ordered stages. This means we know the order of sub-experiments we will use. To understand when the stages are not ordered we must first understand permutations and combinations.

2.2 Permutations

When counting the number of possible ways unordered stages can occur we have to consider all the possibilities by understanding all the different ways to order the stages in terms of the number of possible stages n , and, optionally, the number of stages we must consider, k .

Definition 2.2.1: Permutation - ${}_nP_k$

The number of ways to select an ordered subset of k items from a group of n distinct items is:

$${}_nP_k = \frac{n!}{(n-k)!}$$

Example 2.2.1

Suppose we have a bag containing 3 balls: red, green, and blue. If we draw 3 balls from the bag without replacement (meaning that we don't put the balls back in the bag after drawing them), the possible outcomes of the order of the balls are the permutations of the 3 balls. Using the formula above, we can calculate the number of permutations as follows:

$$P_3 = 3! = 3 \times 2 \times 1 = 6$$

Therefore, there are 6 possible outcomes for this experiment: red-green-blue, red-blue-green, green-red-blue, green-blue-red, blue-red-green, and blue-green-red.

2.3 Indistinguishable Objects

Some times we are trying to track the permutation of some objects, however we must be careful to not count duplicate permutations where only duplicate objects have been rearranged.

Example 2.3.1 (Indistinguishable Objects)

An urn contains 4 red balls, 1 yellow ball, 3 green balls, and 2 blue balls. How many different ways are there to arrange all 10 of these balls?

This problem may seem simple. All we must do is find the permutation of 10 distinct objects, ${}_{10}P_{10}$. However, the issue is that these objects are not actual distinct. If we calculated the total number of outcomes this way we would end up with the following duplicate ordering:



To fix this we can try and consider every possible way each ordering of indistinct objects can repeat, and then divide our incorrect permutation with that. In this case there are 4 red balls, and hence $4!$ ways to order them.

Combining this for each set of colored balls we get that there are $4!3!1!2! = 288$ ways to order every permutation we counted earlier. Therefore we divide the previous answer to get $\frac{10!}{288} = 12600$ ways to order the 10 balls.

2.4 Combinations

Permutations describe an experiment where the stages, or objects, and the order of the stages, or objects we are tracking, is unclear. On the other hand there are cases when in a random experiment we known the order of the objects, or do not care, yet the possible objects are unclear.

Example 2.4.1

Consider the 5 letters, ABCDE and we want to select only three (unordered) letters. This means the selection ABC includes selections of BAC. How many unique groups of 3 letters are possible?

In this example it is obvious that this encompasses some sort of permutation of the letters, as we randomly select three. However, the issue is that calculating ${}_5P_3$ includes duplicates such as ABC and CBA . These two examples may seem different, but in the domain of our question, which ignores order, these instances are considered the same outcome. Instead we need to consider, just like for when calculating Indistinguishable objects, the number of duplicate permutations for every single permutation. For this we have to consider the number of ways to order the 3 letters. There are $3!$ ways to order three elements. This means that every permutation has $3!$ duplicates. So to find the answer we have to calculate:

$$\frac{{}_5P_3}{3!}$$

From this example we had to find the number of combinations of 3 elements where there are 5 to choose from. Unsurprisingly, this concept of counting is called, combinations.

Definition 2.4.1: Combinations - ${}_nC_k$

The number of ways to select an *unordered* subset of k items from a group of n **distinct** items without replacement is:

$$\binom{n}{k} = {}_nC_k = \frac{n!}{(n-k)!k!}$$

Note:-

The formula for combinations is derived from the formula for permutations. The extra $k!$ has to do with there being $k!$ ways to order every single combination.

Chapter 3

Random Variables

A random variable is a function that assigns a numerical value to each outcome of a random experiment. It is a way to represent the outcomes of a random process or experiment in a numerical form.

Random variables are an important concept in probability and statistics because they allow us to describe and analyze random processes and experiments. By assigning numerical values to the outcomes of a random experiment, we can use tools from mathematics and statistics to analyze the behavior and characteristics of the random process.

There are two main types of random variables: discrete and continuous. A discrete random variable is one that can take on only a finite or countably infinite number of values, such as the number of heads obtained when flipping a coin three times. A continuous random variable is one that can take on any value within a given range, such as the height of a person.

Understanding and working with random variables is a crucial skill for anyone interested in data analysis, statistical modeling, and other fields that involve the analysis of random processes and experiments.

3.1 Introduction to Random Variables

Definition 3.1.1: Random Variable

A **random variable** is a real-valued function that assigns a numerical value to each event in the sample space Ω arising from a random experiment. A random variable X is a real-valued function $X : \Omega \rightarrow \mathbb{R}$ such that for every $\omega \in \Omega$, $X(\omega) = x \in \mathbb{R}$. It is a mapping from the sample space to the real numbers, representing the outcome.

Example 3.1.1 (Using Random Variables)

Consider the experiment of tossing a coin with sample space $\Omega = \{H, T\}$. Let X be the outcome of a toss. We can define X such that when $X = 1$ corresponds to H and when $X = 0$ corresponds to T .

Note:-

The convention for random variables is to express them with capital letters and when interpreting their unknown value we use the lower case version. An example being for some random variable X we consider its possible probabilities as $P(X = x)$.

3.2 PMF of Discrete Random Variables

A discrete of a random variable X is one that can take on only a finite number or a countably infinite number of possible values x . The probability mass function of a discrete random variables assigns a probability to each value $x \in X$, such that we can understand the probability of that value occurring.

Definition 3.2.1: Probability Mass Function of a Discrete Random Variable

Let X be a random variable. The **probability mass function** of X is such that for every $x \in X$ we assign a probability where:

1. $0 \leq P(X = x) \leq 1$
2. $\sum_{x \in X} P(X = x) = 1$

Example 3.2.1 (Finding a PMF)

A factory producing computer parts sends out a shipment of 10 parts of which 3 are defective. Find the probability mass function for the number of defectives a customer will get if the first customer randomly purchases 4 computer parts.

For this question we have to quantify the probability of each possible number of defective parts the customer receives. We know that this value will be between 0 – 3 as the customer buys 4 parts. We can think of this question as various combinations of parts, good and bad. Let D be the number of defective parts:

$$\begin{aligned}P(D = 0) &= \frac{\binom{7}{4}}{\binom{10}{4}} \\P(D = 1) &= \frac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}} \\P(D = 2) &= \frac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}} \\P(D = 3) &= \frac{\binom{7}{1}\binom{3}{3}}{\binom{10}{4}} \\P(D = d) &= \frac{\binom{7}{4-d}\binom{3}{d}}{\binom{10}{4}}\end{aligned}$$

Note:-

Discrete Distributions and their probability mass functions can be thought of as histograms/bar charts. This is expanded on in the Discrete Probability Distributions chapter.

3.3 Characteristics of Random Variables

3.3.1 Expected Values

The expected value lets us track the mean of Random Variables:

Definition 3.3.1: Expected Value

The long run/theoretical average. If a random experiment were to be conducted n times, then as $n \rightarrow \infty$ then the average of outcomes converges to the expected value. This is often denoted as μ . For any random variable X , its expected value is denoted and defined as:

$$E[X] = \mu = \sum_{x \in X} x \cdot P(X = x)$$

For any transformation of X , $g(X)$, we define the expected value of this transformation of X similarly:

$$E[g(X)] = \sum_{x \in X} g(x) \cdot P(X = x)$$

Note:-

$E[g(X)] \neq g(E[X])$ unless $g(X)$ is only a linear transformation.

3.3.2 Variance

The variance of a random experiment or random variable expressed the spread and variability of results:

Definition 3.3.2: Variance & Standard Deviation

The **variance** of X is defined to be

$$\sigma^2 = V(X) = E[(x - \mu)^2] = \sum_{x \in X} (x - \mu)^2 \cdot P(X = x)$$

Variance captures the spread in *units*². The standard deviation, $\sigma = \sqrt{\text{variance}}$ is the average spread, or distance from the mean, the random variable is.

3.3.3 Properties of Expectation and Variance

For any constants, a and b , and discrete random variables, X and Y , the following properties of expectation is true:

Corollary 3.3.1 Properties of Expectation

1. $E[a] = a$
2. $E[X + a] = E[X] + a$
3. $E[aX] = a \cdot E[X]$
4. $E[aX + b] = a \cdot E[X] + b$
5. $E[X + Y] = E[X] + E[Y]$
6. If X and Y are independent: $E[XY] = E[X] \cdot E[Y]$

For any constants, a and b , and discrete random variables, X and Y , the following properties of variance is true:

Corollary 3.3.2

1. $V(a) = 0$
2. $V(a + X) = V(X) = \sigma^2$
3. $V(aX) = a^2 \cdot V(X) = a^2 \cdot \sigma^2 = a^2 E[(X - \mu)^2]$
4. $V(aX + b) = a^2 \cdot V(X) = a^2 \cdot \sigma^2$
5. IF X and Y are independent $V(X + Y) = V(X) + V(Y)$

3.4 Cumulative Distribution Function

Definition 3.4.1

The cumulative distribution function (CDF) $F(x)$ of a discrete random variable with probability mass function $P(x)$ or $f(x)$ is a function that returns the cumulative (total) probability up to and including $X = x$.

$$F(b) = P(X \leq b) = \sum_{x \in \{x \leq b\}} P(x)$$

The domain of the CDF is always over the set of real numbers! As such, CDFs are often represented as a piecewise function.

Note:-

These notes do not include anything on distributions because I did not have time. Regardless those are best to be studied with practice problems.

Some graphics were taken from Karen Wong's lecture slides.