Introduction to Probability and Statistics

Aakash Jog

2015-16

Contents

1 Lecturer Information **v**

2 Instructor Information **v**

3 Recommended Reading **v**

I Basics of Probability 1

1 Terminology **1**

2 Basic Laws **1**

3 Axioms of Probability **2**

4 Basics of Combinatorics **4**

5 Conditional Probability **11**

6 Bayes' Theorem **14**

7 Independent Events **16**

II Discrete Random Variables 19

1 Discrete Random Variables **19**

2 Probability Mass Function **19**

3 Variance **24**

III Continuous Random Variables 28

1 Cumulative Distribution Function 28

2 Continuous Random Variable 29

3 Expectation and Variance 31

IV Special Distributions 33

1 Bernoulli and Binomial Random Variables 33

2 Poisson Random Variables 36

2.1 Assumptions for Poisson Distributions for Events Over a Period of Time 39

3 Geometric Random Variables 41

4 Negative Binomial Random Variable 44

5 Hypergeometric Random Variable 47

6 Uniform Random Variable 48

7 Exponential Random Variable 50

8 Normal Distribution 51

9 de Moivre-Laplace Limit Theorem 54

V Combination of Random Variables 56

1 Expected Value of Sums of Random Variables 56

VI Jointly Distributed Random Variables 58

1 Joint Cumulative Probability Distribution Function 58

2 Joint Continuous Variables 59

3 Independent Random Variables 62

4 Conditional Random Variables **65**

5 Properties of Expectation **66**

6 Properties of Variance **70**

1 Lecturer Information

Dr. Galit Ashkenazi-Golan

E-mail: galit.ashkenazi@gmail.com

2 Instructor Information

Liran Mendel

E-mail: liran.mendel@gmail.com

3 Recommended Reading

1. Sheldon M. Ross: A First Course in Probability Pearson Prenticce Hall, 8th Edition, 2010.
2. Bertsekas, Dimitri P. and Tsitsikis, John N., Introduction to Probability. Athena Science, 2nd editions, 2008.
3. Montgomery, D.C and Runger, G.C. and Hubele, N.F. Engineering Statistics. Wiley & Sons, NY, 4th Edition, 2007.

Part I
Basics of Probability

1 Terminology

Definition 1 (Experiment). A situation with uncertain results is called an experiment.

Definition 2 (Sample space). The set of all possible outcomes of an experiment is called the sample space. It is denoted by Ω or S.

Definition 3 (Event). Any subset A of the sample space is called an event.

Definition 4 (Intersection of sets). Let A and B be two events of sample space Ω. The set of all outcomes that are both in A and B is called the intersection of A and B. It is denoted by $A \cap B$.

Definition 5 (Union of sets). Let A and B be two events of sample space Ω. The set of all outcomes that are in either of A and B is called the union of A and B. It is denoted by $A \cup B$.

Definition 6 (Complement of set). Let A be an event of sample space Ω. The set of all outcomes that are not in A, but are in Ω is called the complement of A. It is denoted by $\overline{A}$ or A^{c}.

Definition 7 (Mutually exclusive events). Two events A and B are said to be mutually exclusive, if

$$A \cap B = \emptyset$$

2 Basic Laws

Law 1 (Commutative Laws).

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

Law 2 (Associative Laws).

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

Law 3 (Distributive Laws)**.**

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

Law 4 (De Morgan's Laws)**.**

$$\overline{A_1 \cup \cdots \cup A_n} = \overline{A_1} \cap \cdots \cap \overline{A_n}$$
$$\overline{A_1 \cap \cdots \cap A_n} = \overline{A_1} \cup \cdots \cup \overline{A_n}$$

Proof.

$$\begin{aligned}
&\omega \in \overline{A_a \cup \cdots \cup A_n} \\
&\iff \omega \notin A_{\cup} \cdots \cup A_n \\
&\iff \omega \notin A_1 \text{ and } \ldots \text{ and } \omega \notin A_n \\
&\iff \omega \in \overline{A_1} \text{ and } \ldots \text{ and } \omega \in \overline{A_n} \\
&\iff \omega \in \overline{A_1} \cap \cdots \cap \overline{A_n}
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\omega \in \overline{A_a \cap \cdots \cap A_n} \\
&\iff \omega \notin A_{\cap} \cdots \cap A_n \\
&\iff \omega \notin A_1 \text{ or } \ldots \text{ or } \omega \notin A_n \\
&\iff \omega \in \overline{A_1} \text{ or } \ldots \text{ or } \omega \in \overline{A_n} \\
&\iff \omega \in \overline{A_1} \cup \cdots \cup \overline{A_n}
\end{aligned}$$

□

3 Axioms of Probability

Definition 8 (Probability)**.** The probability of an event E is defined to be a function which satisfies the three basic axioms. It is denoted by $P(E)$.

Axiom 1.

$$0 \leq \mathrm{P}(E) \leq 1$$

Axiom 2.

$$\mathrm{P}(\Omega) = 1$$

Axiom 3. *For any sequence of mutually exclusive events $A_1, \dots,$*

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

Theorem 1.

$$\mathrm{P}(\emptyset) = 0$$

Theorem 2. *For a finite collection of mutually exclusive event $A_1, \dots, A_n$,*

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

Theorem 3.

$$\mathrm{P}\left(\overline{A}\right) = 1 - \mathrm{P}(A)$$

Proof.

$$A \cap \overline{A} = \emptyset$$

Therefore, A and $\overline{A}$ are mutually exclusive. Therefore,

$$\begin{aligned} \mathrm{P}(A) + \mathrm{P}\left(\overline{A}\right) &= \mathrm{P}\left(A \cup \overline{A}\right) \\ &= \mathrm{P}(\Omega) \\ &= 1 \\ \therefore \mathrm{P}\left(\overline{A}\right) &= 1 - \mathrm{P}(A) \end{aligned}$$

□

Theorem 4.

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

Definition 9 (Symmetric sample spaces)**.** A sample space is said to be symmetric if the probabilities of all $\omega \in \Omega$ are the same.

4 Basics of Combinatorics

Theorem 5. *The number of combinations of k objects out of n, without repetition is*

$$\begin{aligned}\binom{n}{k} &= {}^{n}\mathrm{C}_{k} \\ &= \frac{n!}{(n-k)!k!}\end{aligned}$$

Theorem 6.

$$\begin{aligned}{}^{n}\mathrm{P}_{k} &= k!{}^{n}\mathrm{C}_{k} \\ &= \frac{n!}{(n-k)!}\end{aligned}$$

The number of permutations of k objects out of n, without repetition is

Exercise 1.
8 books are to be arranged on 2 shelves, of capacities 3 and 5 respectively. Out of the 8 books, 2 books are special. Find the probability that the two special books end up on the same shelf.

Solution 1.
Let the special books be placed first.
If the first special book is placed on the longer shelf, then it has 5 available positions, and the second special book has 4 available positions.
If the first special book is placed on the shorter shelf, then it has 3 available positions, and the second special book has 2 available positions.
In either case, the number of ways of arranging the remaining 6 books in the remaining positions is 6!.
Therefore, the total number of arrangements satisfying the conditions is $(5 \cdot 4 \cdot 6! + 3 \cdot 2 \cdot 6!)$. The total number of arrangements is 8!. Therefore, the required probability is $\frac{5 \cdot 4 \cdot 6! + 3 \cdot 2 \cdot 6!}{8!}$.

Solution 1.
Let the places for the special books be allotted first.
If the first special book is assigned a place on the longer shelf, then it has 5 available positions, and the second special book has 4 available positions.
If the first special book is assigned a place on the shorter shelf, then it has 3 available positions, and the second special book has 2 available positions.
The total number of arrangements is 8!. Therefore, the required probability is $\frac{5 \cdot 4 + 3 \cdot 2}{8 \cdot 7}$.

Solution 1.
If the special books are to be placed on the longer shelf, the possible combinations are $\binom{5}{2}$.
If the special books are to be placed on the shorter shelf, the possible combinations are $\binom{3}{2}$.
To total possible arrangements are $\binom{8}{2}$.
Therefore, the required probability is $\frac{\binom{5}{2}+\binom{3}{2}}{\binom{8}{2}}$.

Exercise 2.
Ben is going to celebrate the beginning of the year of the dragon. He lives close to two pubs. The probability that he would go to pub A is 0.5 and the probability that the would go to pub B is 0.4. In addition, the probability that he would go to at least one of the two venues is 0.8.

1. What is the sample space?
2. What is the probability that he would go to both pubs?
3. What is the probability that he would go to exactly one pub?

Solution 2.
1. Let A be the event that he would go to pub A, and let B be the event that he goes to pub B. Therefore,

$$S = \{A \cap B^c, A^c \cap B, A \cap B, A^c \cap B^c\}$$

2. The probability that he would go to both pubs is

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= 0.5 + 0.4 - 0.8 \\ &= 0.1 \end{aligned}$$

3. The probability that he would go to exactly one pub is

$$\begin{aligned} P\left((A \cup B) \setminus (A \cap B)\right) &= P(A \cup B) - P(A \cap B) \\ &= 0.8 - 0.1 \\ &= 0.7 \end{aligned}$$

Exercise 3.
A lady crosses three traffic signals, with red and green lights only, on the way to her dog's hairdresser.
The probabilities of encountering 0, 1, and 2 red lights are 0.4, 0.1, 0.2 respectively.
Find the probabilities of

1. Encountering at least one red light.
2. Encountering at least one green light.
3. Encountering an odd number of red lights.

Solution 3.
1. The required probability is

$$\begin{aligned} P(1 \text{ red}) + P(2 \text{ red}) + P(3 \text{ red}) &= 1 - P(0 \text{ red}) \\ &= 1 - 0.4 \\ &= 0.6 \end{aligned}$$

2. The required probability is

$$\begin{aligned} P(1 \text{ green}) + P(2 \text{ green}) + P(3 \text{ green}) &= P(0 \text{ red}) + P(1 \text{ red}) + P(2 \text{ red}) \\ &= 0.4 + 0.1 + 0.2 \\ &= 0.7 \end{aligned}$$

3. The required probability is

$$\begin{aligned} P(1 \text{ red}) + P(3 \text{ red}) &= 1 - P(0 \text{ red}) - P(2 \text{ red}) \\ &= 1 - 0.4 - 0.2 \\ &= 0.4 \end{aligned}$$

Exercise 4.
5 cards are taken out randomly from a 52 card deck. Consider the following events.

1. A: All cards are with numbers higher than 10.
2. B: All cards are hearts.
3. C: All cards have different numbers.
4. D: All cards are consecutive numbers.

Assuming ace to have value 1, find the probabilities of A, B, C, and D.

Solution 4.

$$|\Omega| = \binom{52}{5}$$

There are 12 cards with numbers higher than 10. Therefore,

$$|A| = \binom{12}{5}$$

Therefore,

$$\begin{aligned} \mathrm{P}(A) &= \frac{|A|}{|\Omega|} \\ &= \frac{\binom{12}{5}}{\binom{52}{5}} \end{aligned}$$

There are 13 heart cards. Therefore,

$$|B| = \binom{13}{5}$$

Therefore,

$$\begin{aligned} \mathrm{P}(B) &= \frac{|B|}{|\Omega|} \\ &= \frac{\binom{13}{5}}{\binom{52}{5}} \end{aligned}$$

The number of ways of selecting 5 different numbers out of 13 is $\binom{13}{5}$. For each of the selected number, there are 4 cards, of which exactly one has to be selected. Therefore,

$$|C| = \binom{13}{5}4^5$$

Therefore,

$$\mathrm{P}(C) = \frac{\binom{13}{5}4^5}{\binom{52}{5}}$$

There are 9 sequences of consecutive numbers. Each of the numbers have 4 corresponding cards each. Therefore,

$$|D| = 9 \cdot 4^5$$

Therefore,

$$\mathrm{P}(D) = \frac{9 \cdot 4^5}{\binom{52}{3}}$$

Exercise 5.
A die is tossed 3 times. Consider the following events.

1. A: The sum of all three numbers is even.

Find the probability of A.

Solution 5.
Every time the die is rolled, there as 6 possible outcomes. Therefore,

$$|\Omega| = 6^3$$

For the sum of three numbers to be even, exactly 0 or 2 of them must be odd.
There is 1 combination for all three numbers to be even. Each of these even numbers has 3 options. Therefore, the total number of combinations following the restriction in 3^3.
There are 3 combinations for exactly 2 numbers to be odd. Each of the odd numbers has 3 option, and the even number has 3 options. Therefore, the total number of combinations following the restriction in 3^3.
Therefore,

$$|A| = 4 \cdot 3^3$$

Therefore,

$$\mathrm{P}(A) = \frac{4 \cdot 3^3}{6^3}$$

Exercise 6.
There are n students in a classroom. Assuming 365 days in a year, what is the probability that at least two of them share the same birthday, ignoring the year?

Solution 6.
There are $\binom{365}{1}$ options for each person's birthday. Therefore,

$$|\Omega| = 365^n$$

Let A be the event that everyone has distinct birthdays.
If $n > 365$, at least two persons must share a birthday.
If $n \leq 365$,

$$|A| = {}^{365}\mathrm{P}_n$$

Therefore,

$$\begin{aligned} \mathrm{P}(A) &= \frac{{}^{365}\mathrm{P}_n}{365^n} \\ &= \frac{\frac{365!}{(365-n)!}}{365^n} \end{aligned}$$

Therefore, the probability of at least two persons sharing a birthday is

$$\begin{aligned} \mathrm{P}\left(\overline{A}\right) &= 1 - \mathrm{P}(A) \\ &= 1 - \frac{\frac{365!}{(365-n)!}}{365^n} \end{aligned}$$

Exercise 7.
A president, a treasurer, and a secretary, all different, are to be chosen from a club consisting of 10 people. How many different choices of office bearers are possible if

1. There are no restrictions.
2. Alice and Bob cannot serve together.
3. Charlie and David can serve together or not at all.
4. Eve must be an officer.
5. Frank can serve only if he is the president.

Solution 7.

1. The total possible combinations are ${}^{10}\mathrm{P}_3$.

2. If neither Alice nor Bob are office bearers, there are ${}^{8}\mathrm{P}_3$ possible combinations. If one of Alice and Bob is an office bearer, there are three possible posts for the selected person. The number of combinations for the rest of the posts are ${}^{8}\mathrm{P}_2$. Therefore, the total number of combinations are ${}^{8}\mathrm{P}_3 + 3{}^{8}\mathrm{P}_2$.

3. If both Charlie and David are chosen, the number of combinations are $3 \cdot 2 \cdot {}^{8}\mathrm{P}_1$. If neither Charlie nor David are chosen, the number of combinations are ${}^{8}\mathrm{P}_3$. Therefore, the total number of combinations are $3 \cdot 2 \cdot {}^{8}\mathrm{C}_1 + {}^{8}\mathrm{P}_3$.

4. There are three possible posts for Eve. Therefore, the total number of combinations are $3 \cdot {}^{9}\mathrm{P}_2$.

5. If Frank is the president, the number of combinations are ${}^{9}\mathrm{P}_2$. If Frank is not the president, the number of combinations are ${}^{9}\mathrm{P}_3$. Therefore, the total number of combinations are ${}^{9}\mathrm{P}_2 + {}^{9}\mathrm{P}_3$.

Exercise 8.

a different balls are divided randomly into n different cells. Find the probability that all cells are non-empty when

1. $a = n$
2. $a = n + 1$

Solution 8.

1.

$$\begin{aligned} |S| &= n^a \\ &= n^n \end{aligned}$$

If all cells are to be non-empty, the number of combinations are $n!$. Therefore, the probability is $\frac{n!}{n^n}$.

2.

$$\begin{aligned} |S| &= n^a \\ &= n^{n+1} \end{aligned}$$

The number of combinations to select 2 balls is $\binom{n+1}{2}$. Let these two balls be glued together and be treated as one.

The number of arrangements of these n balls are $n!$.
Therefore, the total number of combinations are $\binom{n+1}{2}n!$.
Therefore, the probability is $\frac{\binom{n+1}{2}}{n!}$.

5 Conditional Probability

Definition 10. For two events A and B in sample space Ω, where $\mathrm{P}(B) > 0$, the conditional probability, i.e. the probability that A will occur after B has already occurred is defined as

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

Exercise 9.
A die is rolled once. Consider the following events.

1. A: The result is even.

2. B: The result is higher than 3.

What is the probability that the result is even, if it is known that result is higher than 3?

Solution 9.

$$\begin{aligned}
\mathrm{P}(A|B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\
&= \frac{\mathrm{P}\left(\{2,4,6\} \cap \{4,5,6\}\right)}{\mathrm{P}\left(\{4,5,6\}\right)} \\
&= \frac{\mathrm{P}\left(\{4,6\}\right)}{\mathrm{P}\left(\{4,5,6\}\right)} \\
&= \frac{\frac{1}{3}}{\frac{1}{2}} \\
&= \frac{2}{3}
\end{aligned}$$

Exercise 10.
A coin is flipped twice. What is the probability of getting 'Heads' on both flips, given that the first flip results in 'Heads'.

Solution 10.

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Let A be the event of getting two 'Heads'. Therefore,

$$A = \{(H, H)\}$$

Let B be the event that the first flip results in 'Heads'. Therefore,

$$B = \{(H, T), (H, H)\}$$

Therefore,

$$\begin{aligned} \mathrm{P}(A|B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\ &= \frac{\frac{1}{4}}{\frac{1}{2}} \\ &= \frac{1}{2} \end{aligned}$$

Exercise 11.
A coin is flipped twice. What is the probability of getting 'Heads' on both flips, given that at least one flip results in 'Heads'.

Solution 11.

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Let A be the event of getting two 'Heads'. Therefore,

$$A = \{(H, H)\}$$

Let B be the event that at least one flip results in 'Heads'. Therefore,

$$B = \{(H, T), (T, H), (H, H)\}$$

Therefore,

$$\begin{aligned} \mathrm{P}(A|B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\ &= \frac{\frac{1}{4}}{\frac{3}{4}} \\ &= \frac{1}{3} \end{aligned}$$

Theorem 7.

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \prod_{k=1}^{n} \mathrm{P}\left(A_k | \bigcap_{l=1}^{k-1} A_l\right)$$

Exercise 12.
A deck of cards is randomly divided into four stacks of 13 cards each. Find the probability that each stack has exactly one ace.

Solution 12.
Let A_1 be the event that $\mathtt{A}\spadesuit$ is in any one of the stacks. Therefore,

$$\mathrm{P}(A_1) = 1$$

Let A_2 be the event that $\mathtt{A}\spadesuit$ and $\mathtt{A}\heartsuit$ are in different stacks. Therefore,

$$\begin{aligned}\mathrm{P}(A_2|A_1) &= \frac{\mathrm{P}(A_1 \cap A_2)}{\mathrm{P}(A_1)} \\ &= \frac{39}{51}\end{aligned}$$

Let A_3 be the event that $\mathtt{A}\spadesuit$, $\mathtt{A}\heartsuit$, and $\mathtt{A}\diamondsuit$ are in different stacks. Therefore,

$$\begin{aligned}\mathrm{P}(A_3|A_1 \cap A_2) &= \frac{\mathrm{P}(A_1 \cap A_2 \cap A_3)}{\mathrm{P}(A_1 \cap A_2)} \\ &= \frac{26}{50}\end{aligned}$$

Let A_4 be the event that $\mathtt{A}\spadesuit$, $\mathtt{A}\heartsuit$, $\mathtt{A}\diamondsuit$, and $\mathtt{A}\clubsuit$ are in different stacks. Therefore,

$$\begin{aligned}\mathrm{P}(A_4|A_1 \cap A_2 \cap A_3) &= \frac{\mathrm{P}(A_1 \cap A_2 \cap A_3 \cap A_4)}{\mathrm{P}(A_1 \cap A_2 \cap A_3)} \\ &= \frac{13}{49}\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{P}(A_1 \cap A_2 \cap A_3 \cap A_4) &= \mathrm{P}(A_1)\,\mathrm{P}(A_2|A_1)\,\mathrm{P}(A_3|A_1 \cap A_2)\,\mathrm{P}(A_4|A_1 \cap A_2 \cap A_3) \\ &= (1)\left(\frac{39}{51}\right)\left(\frac{26}{50}\right)\left(\frac{13}{49}\right)\end{aligned}$$

Exercise 13.
A deck of cards is randomly divided into four stacks of 13 cards each. Find the probability, using combinatorics, that each stack has exactly one ace.

Solution 13.

$$|\Omega| = \binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}$$

Let A be the event that each stack has exactly one ace. Therefore, each stack has one ace, and 12 non-ace cards. Therefore,

$$|A| = \left(\binom{4}{1}\binom{48}{12}\right)\left(\binom{3}{1}\binom{36}{12}\right)\left(\binom{2}{1}\binom{24}{12}\right)\left(\binom{1}{1}\binom{12}{12}\right)$$

Therefore,

$$\begin{aligned} \mathrm{P}(A) &= \frac{|A|}{|\Omega|} \\ &= \frac{\left(\binom{4}{1}\binom{48}{12}\right)\left(\binom{3}{1}\binom{36}{12}\right)\left(\binom{2}{1}\binom{24}{12}\right)\left(\binom{1}{1}\binom{12}{12}\right)}{\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}} \end{aligned}$$

6 Bayes' Theorem

Definition 11 (Division)**.** A set of events $A_1, \dots, A_n$ is called a division of the sample space Ω, if

$$\bigcup_{k=1}^{n} A_k = \Omega$$

and

$$A_i \cap A_j = \emptyset$$

for $i \neq j$.

Theorem 8 (Bayes' Theorem)**.** *Given a division $A_1, \dots, A_n$ of the sample space Ω, and an event B in Ω,*

$$\mathrm{P}(B) = \sum_{i=1}^{n} \mathrm{P}(A_i)\,\mathrm{P}(B|A_i)$$

Exercise 14.
A chocolate factory has three production lines.
50% of the production is milk chocolate, out of which 1% is defective.
30% of the production is dark chocolate, out of which 2% is defective.
20% of the production is white chocolate, out of which 0.5% is defective.
If a chocolate bar is picked randomly, what is the probability that it is defective?

Solution 14.
Let A_1, A_2, A_3 be the events that selected chocolate bar is made of milk, dark, white chocolate, respectively.
Let B be the event that the selected chocolate bar is defective.
Therefore,

$$\begin{aligned}
\mathrm{P}(B) &= \sum_{i=1}^{3} \mathrm{P}(A_i)\,\mathrm{P}(B|A_1) \\
&= ((0.5)(0.01)) + ((0.3)(0.02)) + ((0.2)(0.005)) \\
&= 0.005 + 0.006 + 0.001 \\
&= 0.012
\end{aligned}$$

Exercise 15.
In a certain stage of a criminal investigation, the inspector in charge is 60% convinced of the guilt of a certain suspect. Suppose that a new piece of evidence which shows that the criminal is bald, is uncovered. If 20% of the population is bald, and if the suspect is bald, how certain should that inspector be of the guilt of the suspect?

Solution 15.
Let G be the event that the suspect is guilty.
Let B be the event that the suspect is bald.
Therefore, the event that the suspect is both bald and guilty, is $G|C$.
Therefore,

$$\begin{aligned}
\mathrm{P}(G|B) &= \frac{\mathrm{P}(G \cap B)}{\mathrm{P}(B)} \\
&= \frac{0.6}{(1)(0.6) + (0.4)(0.2)}
\end{aligned}$$

Exercise 16.
What is the probability that when a deck of cards is dealt in a game of bridge, the $\heartsuit$s will be dealt such that Alice gets 3, Bob gets 4, Charlie gets 2, David gets 4.

Solution 16.
Let E_{Alice} be the event of Alice getting the required number of $\heartsuit$s.
Let E_{Bob} be the event of Bob getting the required number of $\heartsuit$s.
Let E_{Charlie} be the event of Charlie getting the required number of $\heartsuit$s.
Let E_{David} be the event of David getting the required number of $\heartsuit$s.
Therefore,

$$\begin{aligned}
\mathrm{P}(E_{\text{Alice}}) &= \frac{{}^{13}\mathrm{C}_3\,{}^{39}\mathrm{C}_{10}}{{}^{52}\mathrm{C}_{13}} \\
\mathrm{P}(E_{\text{Bob}}|E_{\text{Alice}}) &= \frac{{}^{10}\mathrm{C}_4\,{}^{29}\mathrm{C}_9}{{}^{39}\mathrm{C}_{13}} \\
\mathrm{P}(E_{\text{Charlie}}|E_{\text{Alice}} \cap E_{\text{Bob}}) &= \frac{{}^{6}\mathrm{C}_2\,{}^{20}\mathrm{C}_{11}}{{}^{26}\mathrm{C}_{13}} \\
\mathrm{P}(E_{\text{David}}|E_{\text{Alice}} \cap E_{\text{Bob}} \cap E_{\text{Charlie}}) &= \frac{{}^{4}\mathrm{C}_4\,{}^{9}\mathrm{C}_9}{{}^{13}\mathrm{C}_{13}}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{P}(E_{\text{Alice}} \cap E_{\text{Bob}} \cap E_{\text{Charlie}} \cap E_{\text{David}}) = &\quad \mathrm{P}(E_{\text{Alice}}) \\
&\times \mathrm{P}(E_{\text{Bob}}|E_{\text{Alice}}) \\
&\times \mathrm{P}(E_{\text{Charlie}}|E_{\text{Alice}} \cap E_{\text{Bob}}) \\
&\times \mathrm{P}(E_{\text{David}}|E_{\text{Alice}} \cap E_{\text{Bob}} \cap E_{\text{Charlie}}) \\
= &\frac{{}^{13}\mathrm{C}_3\,{}^{39}\mathrm{C}_{10}}{{}^{52}\mathrm{C}_{13}} \frac{{}^{10}\mathrm{C}_4\,{}^{29}\mathrm{C}_9}{{}^{39}\mathrm{C}_{13}} \frac{{}^{6}\mathrm{C}_2\,{}^{20}\mathrm{C}_{11}}{{}^{26}\mathrm{C}_{13}} \frac{{}^{4}\mathrm{C}_4\,{}^{9}\mathrm{C}_9}{{}^{13}\mathrm{C}_{13}}
\end{aligned}$$

7 Independent Events

Definition 12 (Two independent events)**.** Two events, A and B, are said to be independent if

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$$

Theorem 9.

$$\mathrm{P}(A|B) = \mathrm{P}(A)$$

if and only if A and B are independent.

Theorem 10. *If A and B are independent, then so are $\overline{A}$ and B, A and $\overline{B}$, $\overline{A}$ and $\overline{B}$*

Exercise 17.
Two fair dice are rolled.
Let A be the event that the sum of the results of the dice is 6.
Let B be the event that the result of the first die is 4.
Let C be the event that the sum of the results of the dice is 7.
Which of the possible pairs of the events are independent?

Solution 17.

$$
\begin{aligned}
\mathrm{P}(A) &= \frac{5}{36} \\
\mathrm{P}(B) &= \frac{1}{6} \\
\mathrm{P}(C) &= \frac{6}{36} \\
\mathrm{P}(A \cap B) &= \frac{1}{36} \\
\mathrm{P}(B \cap C) &= \frac{1}{36} \\
\mathrm{P}(A \cap C) &= 0
\end{aligned}
$$

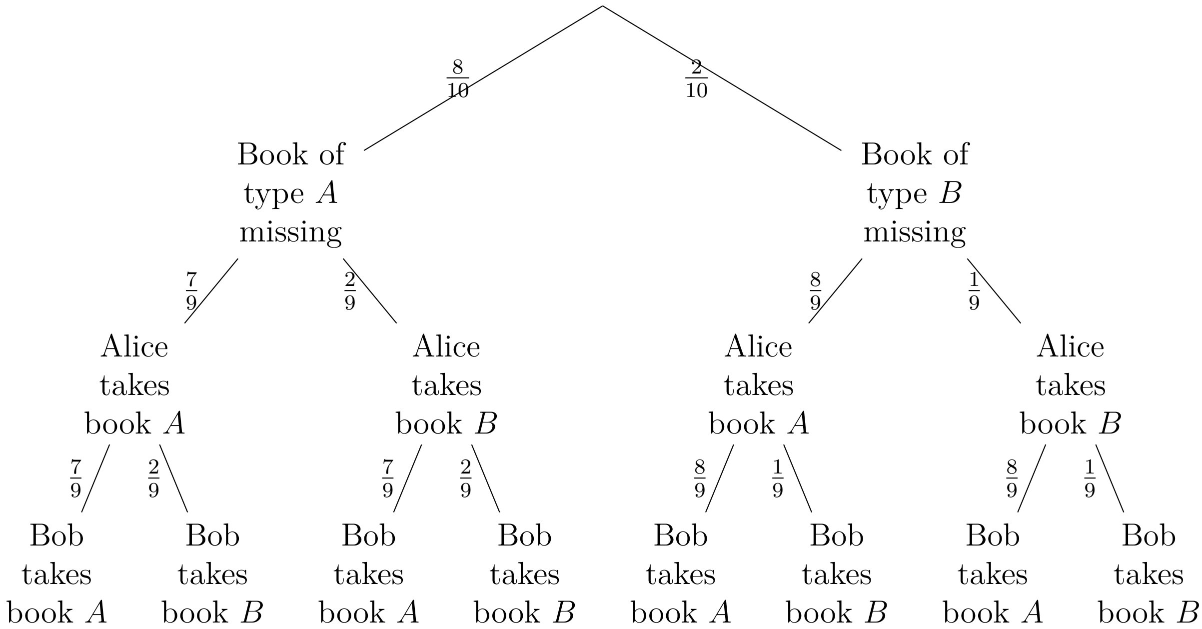
Therefore,

$$
\begin{aligned}
\mathrm{P}(A \cap B) &\neq \mathrm{P}(A)\,\mathrm{P}(B) \\
\mathrm{P}(B \cap C) &= \mathrm{P}(B)\,\mathrm{P}(C) \\
\mathrm{P}(C \cap A) &\neq \mathrm{P}(C)\,\mathrm{P}(A)
\end{aligned}
$$

Therefore, only B and C are independent.

Exercise 18.
There are 10 books in a library, 8 of type A and 2 of type B. On a day when one of the books is missing, Alice enters the library, randomly takes a book, reads it and returns it. After Alice returns the book, Bob enters the library and randomly takes a book.
Let E be the event that Alice took a book of type A.
Let F be the event that Bob took a book of type A.
Are E and F independent?

Solution 18.



$$
\begin{aligned}
\mathrm{P}(E) &= \left(\frac{8}{10}\right)\left(\frac{7}{9}\right) + \left(\frac{2}{10}\right)\left(\frac{8}{9}\right) \\
&= \frac{8}{10} \\
\mathrm{P}(F) &= \left(\frac{2}{10}\right)\left(\frac{8}{9}\right)\left(\frac{8}{9}\right) + \left(\frac{2}{10}\right)\left(\frac{1}{9}\right)\left(\frac{8}{9}\right) + \left(\frac{8}{10}\right)\left(\frac{7}{9}\right)\left(\frac{7}{9}\right) + \left(\frac{8}{10}\right)\left(\frac{2}{9}\right)\left(\frac{7}{9}\right) \\
&= \frac{8}{10}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{P}(E \cap F) &= \left(\frac{2}{10}\right)\left(\frac{8}{9}\right)\left(\frac{8}{9}\right) + \left(\frac{8}{10}\right)\left(\frac{7}{9}\right)\left(\frac{7}{9}\right) \\
&\neq \mathrm{P}(E)\,\mathrm{P}(F)
\end{aligned}
$$

¡++¿ Therefore, the events are not independent.

Definition 13. Three events, A, B, and C, are said to be independent if

$$
\begin{aligned}
\mathrm{P}(A \cap B \cap C) &= \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C) \\
\mathrm{P}(A \cap B) &= \mathrm{P}(A)\,\mathrm{P}(B) \\
\mathrm{P}(B \cap C) &= \mathrm{P}(B)\,\mathrm{P}(C) \\
\mathrm{P}(C \cap A) &= \mathrm{P}(C)\,\mathrm{P}(A)
\end{aligned}
$$

Part II

Discrete Random Variables

1 Discrete Random Variables

Definition 14 (Random variable)**.** A function $X : \Omega \to \mathbb{R}$ which maps points from the sample space to the real line is called a random variable.

Exercise 19.
Three balls are to be randomly selected, without replacement, from an urn containing 20 balls numbered 1 to 20. If Alice bets that at least one of the balls drawn has a number as large as or larger than 17, what is the probability that Alice wins the bet?

Solution 19.
Let X be the largest number selected.
Therefore, X is a random variable which has a value from $\{3, \dots, 20\}$.
Let the value of the highest valued ball be i. Therefore, the number of ways to select the remaining two balls is ${}^{i-1}\mathrm{C}_2$.
Therefore, the probability of the value of the highest valued ball being i is

$$\mathrm{P}(X = i) = \frac{{}^{i-1}\mathrm{C}_2}{{}^{20}\mathrm{C}_3}$$

Therefore,

$$\begin{aligned}\mathrm{P}(X \geq 17) &= \mathrm{P}(X = 17) + \mathrm{P}(X = 18) + \mathrm{P}(X = 19) + \mathrm{P}(X = 20) \\ &= \frac{{}^{16}\mathrm{C}_2 + {}^{17}\mathrm{C}_2 + {}^{18}\mathrm{C}_2 + {}^{19}\mathrm{C}_2}{{}^{20}\mathrm{C}_3}\end{aligned}$$

2 Probability Mass Function

Definition 15 (Discrete random variable)**.** A random variable that can have at most a countable number of possible values is said to be discrete.

Definition 16 (Probability mass function)**.** A function which gives the probability of a discrete random variable X having value x is called the probability mass function of X. It is denoted as $\mathrm{P}(X = x)$.

Exercise 20.
The probability mass function of a random variable X is given by

$$\mathrm{P}(X = i) = \frac{c\lambda^i}{i!}$$

where $i \in \mathbb{W}$, and $\lambda > 0$.
Find

1. $\mathrm{P}(X = 0)$

2. $\mathrm{P}(X > 2)$

Solution 20.
1.

$$\begin{aligned} \sum_{i=0}^{\infty} \mathrm{P}(X = i) &= 1 \\ \therefore \sum_{i=0}^{\infty} \frac{c\lambda^i}{i!} &= 1 \\ \therefore c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} &= 1 \\ \therefore ce^{\lambda} &= 1 \\ \therefore c &= e^{-\lambda} \end{aligned}$$

Therefore,

$$\mathrm{P}(X = i) = \frac{e^{-\lambda}\lambda^i}{i!}$$

Therefore,

$$\begin{aligned} \mathrm{P}(X = 0) &= \frac{e^{-\lambda}\lambda^0}{0!} \\ &= e^{-\lambda} \end{aligned}$$

2.

$$\begin{aligned}\sum_{i=0}^{\infty} \mathrm{P}(X=i) &= 1\\ \therefore \sum_{i=0}^{\infty} \frac{c\lambda^i}{i!} &= 1\\ \therefore c\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} &= 1\\ \therefore ce^{\lambda} &= 1\\ \therefore c &= e^{-\lambda}\end{aligned}$$

Therefore,

$$\mathrm{P}(X=i) = \frac{e^{-\lambda}\lambda^i}{i!}$$

Therefore,

$$\begin{aligned}\mathrm{P}(X>2) &= 1-\mathrm{P}(X\le 2)\\ &= 1-\mathrm{P}(X=2)-\mathrm{P}(X=1)-\mathrm{P}(X=0)\\ &= 1-\frac{e^{-\lambda}\lambda^2}{2!}-\frac{e^{-\lambda}\lambda^1}{1!}-\frac{e^{-\lambda}\lambda^0}{0!}\end{aligned}$$

Definition 17. If X is a discrete random variable having a probability mass function $P(X=x)$, then the expectation or the expected value of X is defined to be

$$\mathrm{E}[X] = \sum_{\left\{x \mid \mathrm{P}(X=x)>0\right\}} x\,\mathrm{P}(X=x)$$

Exercise 21.
Find $\mathrm{E}[X]$ where X is the outcome of rolling a fair die.

Solution 21.

$$\begin{aligned}\mathrm{E}[x] &= \sum_{x=1}^{6} \frac{1}{6}x\\ &= 3.5\end{aligned}$$

Exercise 22.
A class of 120 students is driven in 3 buses to a symphonic perfomance. There are 36 students in the first bus, 14 in the second, and 44 in the third bus. When the buses arrive, a student is randomly chosen. Let X denote the number of students on the bus of the chosen student. Find $\mathrm{E}[X]$.

Solution 22.

$$\begin{aligned}
\mathrm{P}(X = 36) &= \frac{36}{120} \\
\mathrm{P}(X = 40) &= \frac{40}{120} \\
\mathrm{P}(X = 44) &= \frac{44}{120}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}[X] &= 36\left(\frac{36}{120}\right) + 40\left(\frac{40}{120}\right) + 44\left(\frac{44}{120}\right) \\
&= 40.2667
\end{aligned}$$

Exercise 23.
X has the following distribution.

$$\begin{aligned}
\mathrm{P}(X = -1) &= 0.2 \\
\mathrm{P}(X = 0) &= 0.5 \\
\mathrm{P}(X = 1) &= 0.3
\end{aligned}$$

Find $\mathrm{E}\left[x^2\right]$.

Solution 23.
Let

$$Y = X^2$$

Therefore,

$$\begin{aligned}
\mathrm{P}(Y = 0) &= \mathrm{P}(X = 0) \\
&= 0.5 \\
\mathrm{P}(Y = 1) &= \mathrm{P}(X = -1) + \mathrm{P}(X = 1) \\
&= 0.5
\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{E}[Y] &= \mathrm{E}\left[X^2\right] \\ &= (0)(0.5) + (1)(0.5) \\ &= 0.5\end{aligned}$$

Theorem 11. *If X is a discrete random variable that takes on one of the values x_i, where $i \in \mathbb{N}$, with probability mass function $\mathrm{P}(X = x_i)$. Then, for any real valued function g,*

$$\mathrm{E}\left[g(x)\right] = \sum_i g(x_i)\,\mathrm{P}(x = x_i)$$

Exercise 24.
A product that is sold seasonally yeilds a net profit of b for each unit sold and a net loss of l for each unit left unsold when the season ends. The number of units of the product that are sold at a specific store during any season is a random variable X, with probability mass function such that

$$\mathrm{P}(X = i) = \mathrm{P}(i)$$

where $i \in \mathbb{N}$. If the store must stock this product in advance, determine the number of units the store should stock, so as to maximize its profit.

Solution 24.
Let the number of units sold by X.
Let the number of units stocked by s.
Therefore, the total profit is

$$\pi(s) = \begin{cases} bX - (s - X)l & ; \quad X \leq s \\ bs & ; \quad X > s \end{cases}$$

Therefore,

$$\begin{aligned}\mathrm{E}\left[\pi(s)\right] &= \sum_{i=0}^{s} (bi - (s-i)l)\,\mathrm{P}(X = i) + \sum_{i=s+1}^{\infty} sb\,\mathrm{P}(X = i) \\ &= (b + l)\sum_{i=0}^{s} i\,\mathrm{P}(X = i) - sl\sum_{i=0}^{s}\mathrm{P}(X = i) + sb\left(1 - \sum_{i=0}^{s}\mathrm{P}(X = i)\right) \\ &= (b + l)\sum_{i=0}^{s} i\,\mathrm{P}(X = i) - (b + l)s\sum_{i=0}^{s}\mathrm{P}(X = i) + sb \\ &= sb + (b + l)\sum_{i=0}^{s}(i - s)\,\mathrm{P}(X = i)\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}\left[\pi(s+1)\right] &= b(s+1) + (b+l)\sum_{i=0}^{s+1}(i-s-1)\,\mathrm{P}(X=i) \\
&= b(s+1) + (b+l)\sum_{i=0}^{s}(i-s-1)\,\mathrm{P}(X=i)
\end{aligned}$$

Therefore,

$$\mathrm{E}\left[\pi(s+1)\right] - \mathrm{E}\left[\pi(s)\right] = b - (b+l)\sum_{i=0}^{s}\mathrm{P}(X=i)$$

Therefore, increasing the stock is profitable when

$$\sum_{i=0}^{s}\mathrm{P}(X=i) < \frac{b}{b+l}$$

3 Variance

Definition 18 (Variance)**.** The variance of a random variable X is defined to be

$$\begin{aligned}
\mathrm{V}(x) &= \mathrm{E}\left[(X-\mathrm{E}[X])^2\right] \\
&= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2
\end{aligned}$$

Exercise 25.
Calculate $\mathrm{V}(X)$ where X represents the outcome of rolling a fair die.

Solution 25.

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{x=1}^{6}\frac{1}{6}x \\
&= \frac{7}{2} \\
\mathrm{E}\left[X^2\right] &= \sum_{x=1}^{6}\frac{1}{6}x^2 \\
&= \frac{91}{6}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= \frac{91}{6} - \frac{49}{4} \\
&= \frac{35}{12}
\end{aligned}$$

Theorem 12.

$$\mathrm{V}(aX+b) = a^2\,\mathrm{V}(X)$$

Proof.

$$\begin{aligned}
\mathrm{V}(aX+b) &= \mathrm{E}\left[(aX+b-\mathrm{E}[aX+b])^2\right] \\
&= \mathrm{E}\left[(aX+b-a\,\mathrm{E}[X]-b)^2\right] \\
&= \mathrm{E}\left[a^2\,(X-\mathrm{E}[X])^2\right] \\
&= a^2\,\mathrm{E}\left[(x-\mathrm{E}[x])^2\right] \\
&= a^2\,\mathrm{V}(x)
\end{aligned}$$

□

Exercise 26.

You have a coin with probability p of getting 'Heads'. You flip this coin twice. For each flip, if the result is 'Heads', you win \$30. If the result is 'Tails', you lose \$20.

Let X be your profit in the game.

1. What is the sample space?
2. Describe the probability mass function of X.
3. Describe the cumulative distribution function of X.
4. What is the expected value of X?
5. What is the value of p upto which you would agree to participate in the game?
6. What is $\mathrm{V}(X)$?
7. What is $\sigma(X)$?

Solution 26.

1.

$$
\begin{aligned}
S &= \{(30+30), (30-20), (-20-20)\} \\
&= \{60, 10, -40\}
\end{aligned}
$$

2.

$$
\mathrm{P}(X=x) = \begin{cases} p^2 & ; \quad x = 60 \\ 2p(1-p) & ; \quad x = 10 \\ (1-p)^2 & ; \quad x = -40 \end{cases}
$$

3.

$$
\mathrm{F}_(X) = \begin{cases} 0 & ; \quad x < -40 \\ (1-p)^2 & ; \quad -40 \le x < 10 \\ (1-p)^2 + 2p(1-p) & ; \quad 10 \le x < 60 \\ (1-p)^2 + 2p(1-p) + p^2 & ; \quad x \le 60 \end{cases}
$$

4.

$$
\begin{aligned}
\mathrm{E}[X] &= (-40)(1-p)^2 + (10)\left(2p(1-p)\right) + (60)\left(p^2\right) \\
&= 100p - 40
\end{aligned}
$$

5. The game should be played as long as the expected value of X is non negative. Therefore,

$$
\begin{aligned}
\mathrm{E}[X] &\ge 0 \\
\therefore 100p - 40 &\ge 0 \\
\therefore p &\ge 0.4
\end{aligned}
$$

6.

$$
\begin{aligned}
\mathrm{E}\left[x^2\right] &= (-40)^2(1-p)^2 + (10)^2 2p(1-p) + (60)^2 p^2 \\
&= 5000p^2 - 3000p + 1600 \\
\mathrm{E}[x]^2 &= (100p-40)^2 \\
&= 10000p^2 - 8000p + 1600
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= 5000p(1-p)
\end{aligned}
$$

7.

$$\begin{aligned}\sigma(X) &= \sqrt{\mathrm{V}(X)} \\ &= \sqrt{5000p(1-p)}\end{aligned}$$

Part III

Continuous Random Variables

1 Cumulative Distribution Function

Definition 19 (Cumulative distribution function)**.** The function

$$\mathrm{F}_X = \mathrm{P}(X \leq x)$$

for $-\infty < x \leq x$, is called the cumulative distribution function of the variable X.

Exercise 27.
Let X be the result of a die roll. Plot the cumulative distribution function of X.

Solution 27.

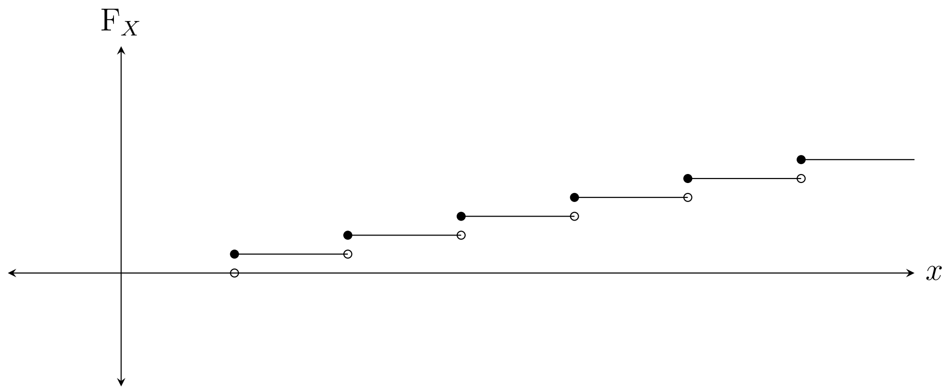


Figure 1: Cumulative Distribution Function for a Die Roll

Theorem 13. *Let* F_X *be the cumulative distribution function of a random variable* X*. Then,*

1. F_X *is a non-decreasing function.*
2. $\lim\limits_{b\to\infty} \mathrm{F}_X(b) = 1.$
3. $\lim\limits_{b\to-\infty} \mathrm{F}_X(b) = 0.$
4. F_X *is right continuous, i.e., the function is equal to its right hand limit.*

2 Continuous Random Variable

Definition 20 (Continuous random variable)**.** A random variable X is said to be a continuous random variable if there exists a function f such that

$$\mathrm{F}_X(x) = \int\limits_{-\infty}^{x} \mathrm{f}(t)\,\mathrm{d}t$$

Definition 21. Let

$$\mathrm{F}_X(x) = \int\limits_{-\infty}^{x} \mathrm{f}(t)\,\mathrm{d}t$$

Then, $\mathrm{f}(t)$ is called the probability density function of x.

Exercise 28.
Let X be a continuous variable whose probability density function is

$$\mathrm{f}(x) = \begin{cases} c\left(4x - 2x^2\right) & ; \quad 0 < x < 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

1. Find c.
2. Find $\mathrm{P}(X > 1)$.

Solution 28.
1. As $\mathrm{f}(x)$ is a probability density function,

$$\begin{aligned} 1 &= \int\limits_{-\infty}^{\infty} \mathrm{f}(x)\,\mathrm{d}x \\ &= \int\limits_{0}^{2} c\left(4x^2 - 2x^2\right)\mathrm{d}x \\ &= \frac{8c}{3} \end{aligned}$$

Therefore,

$$c = \frac{3}{8}$$

2.

$$\begin{aligned}
\mathrm{P}(X > 1) &= \int_{1}^{\infty} \mathrm{f}(x)\,\mathrm{d}x \\
&= \int_{1}^{2} \frac{3}{8}\left(4x - 2x^2\right)\mathrm{d}x \\
&= \frac{1}{2}
\end{aligned}$$

Exercise 29.
The amount of time in hours that a computer functions before breaking down has the distribution

$$\mathrm{f}(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & ; \quad x \geq 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

What is the probability that the computer functions for more than 50 but less than 150 hours?

Solution 29.

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} \mathrm{f}(x)\,\mathrm{d}x \\
&= \lambda \int_{0}^{\infty} e^{-\frac{x}{100}}\,\mathrm{d}x \\
&= 100\lambda
\end{aligned}$$

Therefore,

$$\lambda = \frac{1}{100}$$

Therefore,

$$\begin{aligned}
\mathrm{P}(50 < x < 150) &= \int_{50}^{150} \lambda e^{-\frac{x}{100}}\,\mathrm{d}x \\
&= \int_{50}^{150} \frac{1}{100} e^{-\frac{x}{100}}\,\mathrm{d}x \\
&\approx 0.384
\end{aligned}$$

3 Expectation and Variance

Definition 22. Let X be a continuous random variable. Then, the expectation is defined as

$$\mathrm{E}[X] = \int\limits_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

Exercise 30.
Find the expectation of X is the probability density function is given to be

$$\mathrm{f}(x) = \begin{cases} 2x & ; \quad 0 < x < 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Solution 30.

$$\begin{aligned} \mathrm{E}[X] &= \int\limits_{-\infty}^{\infty} x\,\mathrm{f}(x)\,\mathrm{d}x \\ &= \int\limits_{0}^{1} 2x\,\mathrm{d}x \\ &= \frac{2}{3} \end{aligned}$$

Exercise 31.
The probability density function of X is given by

$$\mathrm{f}(x) = \begin{cases} 1 & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find $\mathrm{E}\left[e^X\right]$.

Solution 31.
Let

$$Y = e^X$$

Therefore,

$$\begin{aligned}
\mathrm{F}_Y(y) &= \mathrm{P}(Y \leq y) \\
&= \mathrm{P}\left(e^X \leq y\right) \\
&= \mathrm{P}(X \leq \ln y) \\
&= \int\limits_{-\infty}^{\ln y} \mathrm{f}_X(x) \, \mathrm{d}x \\
&= \int\limits_{0}^{\ln y} \mathrm{d}x \\
&= \ln y
\end{aligned}$$

Therefore, differentiating,

$$\begin{aligned}
\mathrm{f}_Y(y) &= \frac{\mathrm{d}\mathrm{F}_Y(y)}{\mathrm{d}y} \\
&= \frac{\mathrm{d} \ln y}{\mathrm{d}y} \\
&= \frac{1}{y}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}\left[e^X\right] &= \mathrm{E}[Y] \\
&= \int\limits_{-\infty}^{\infty} y \, \mathrm{f}_Y(y) \, \mathrm{d}y \\
&= \int\limits_{1}^{e} y \frac{1}{y} \, \mathrm{d}y \\
&= e - 1
\end{aligned}$$

Theorem 14. *If X is a continuous random variable, then for any real function g,*

$$\mathrm{E}\left[g(x)\right] = \int\limits_{-\infty}^{\infty} g(x) f(x) \, \mathrm{d}x$$

Part IV
Special Distributions

1 Bernoulli and Binomial Random Variables

Definition 23 (Bernoulli random variable)**.** A random variable X is said to be a Bernoulli random variable if its probability mass function is given by

$$\mathrm{P}(X = 0) = 1 - p$$
$$\mathrm{P}(X = 1) = p$$

Theorem 15. *For a Bernoulli random variable,*

$$\mathrm{E}[x] = p$$
$$\mathrm{V}(x) = p(1 - p)$$

Definition 24 (Binomial random variable)**.** Consider n independent trials, each of which has a probability of success p, and probability of failure $1-p$. If X represents the number of successes that occur in the n trials, then X is said to be a binomial random variable with parameters (n, p). It is denoted as

$$X \sim \mathrm{Bin}(n, p)$$

Theorem 16. *For a binomial random variable,*

$$X \sim \mathrm{Bin}(n, p)$$

the probability mass function is

$$\mathrm{P}(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$$

Exercise 32.
A die is rolled 5 times. What is the probability that the result is 6, 3 times?

Solution 32.
Let X be the number of times 6 appears.
Therefore,

$$X \sim \mathrm{Bin}\left(5, \frac{1}{6}\right)$$

Therefore,

$$\begin{aligned} \mathrm{P}(X = i) &= {}^{n}\mathrm{C}_{i} p^{i} (1-p)^{n-i} \\ \therefore \mathrm{P}(X = 3) &= {}^{5}\mathrm{C}_{3} \left(\frac{1}{6}\right)^{3} \left(\frac{5}{6}\right)^{3} \end{aligned}$$

Exercise 33.

A player bets on a number from 1 to 6, both including. Three dice are then rolled. If the number bet on by the player appears i times where $i = 1, 2, 3$, he wins i units. If the number bet on by the player does not appear on any of the dice, he loses 1 unit.
A game is considered to be fair if the expected value for the player is at least 0. Is this game fair towards the player?

Solution 33.

Let X be the player's winnings.
Let Y be the number of times the number the player bet on appeared. Therefore,

$$Y \sim \mathrm{Bin}\left(3, \frac{1}{6}\right)$$

Therefore,

$$\begin{aligned}
\mathrm{P}(X=-1) &= \mathrm{P}(Y=0) \\
&= {}^3\mathrm{C}_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^3 \\
&= \frac{125}{216} \\
\mathrm{P}(X=1) &= \mathrm{P}(Y=1) \\
&= {}^3\mathrm{C}_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^2 \\
&= \frac{75}{216} \\
\mathrm{P}(X=2) &= \mathrm{P}(Y=2) \\
&= {}^3\mathrm{C}_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^1 \\
&= \frac{15}{216} \\
\mathrm{P}(X=3) &= \mathrm{P}(Y=3) \\
&= {}^3\mathrm{C}_3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^0 \\
&= \frac{1}{216}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}[X] &= (-1)\left(\frac{125}{216}\right) + (1)\left(\frac{75}{216}\right) + (2)\left(\frac{15}{216}\right) + (3)\left(\frac{1}{216}\right) \\
&= -\frac{17}{216}
\end{aligned}$$

Therefore, as the expected value of the winnings is less than 0, the game is not fair towards the player.

Theorem 17. *Let*

$$\begin{aligned}
X &\sim \mathrm{Bin}(n,p) \\
Y &\sim \mathrm{Bin}(n-1,p)
\end{aligned}$$

Then,

$$\mathrm{E}\left[X^k\right] = np\,\mathrm{E}\left[(Y+1)^{k-1}\right]$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[X^k\right] &= \sum_{i=0}^{n} i_k \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i_k \binom{n}{i} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-1}
\end{aligned}$$

Let

$$j = i - 1$$

Therefore,

$$\begin{aligned}
\mathrm{E}\left[X^k\right] &= np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \\
&= \mathrm{P}(Y = j) \\
&= np\, \mathrm{E}\left[(Y+1)^{k-1}\right]
\end{aligned}$$

□

2 Poisson Random Variables

Definition 25 (Poisson Random Variables)**.** A random variable X that takes on whole number values is said to be a Poisson random variable with parameter λ if for some $\lambda > 0$,

$$\mathrm{P}(X = i) = \frac{e^{-\lambda}\lambda^i}{i!}$$

where $i \in \mathbb{W}$.
It is denoted as

$$X \sim \mathrm{Poi}(p)$$

Theorem 18. *Let*

$$X \sim \mathrm{Bin}(n, p)$$

If $n \to \infty$, the probability distribution of X is a Poisson distribution.

Proof. Let

$$X \sim \mathrm{Bin}(n, p)$$

Therefore,

$$\begin{aligned} \mathrm{P}(X = i) &= \binom{n}{i} p^i (1-p)^{n-i} \\ &= \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \end{aligned}$$

Let

$$\lambda = np$$

Therefore,

$$\begin{aligned} \mathrm{P}(X = i) &= \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{(n)(n-1)\dots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{\left(1-\frac{\lambda}{n}\right)^2}{\left(1-\frac{\lambda}{n}\right)^i} \end{aligned}$$

$$\therefore \lim_{n\to\infty} \mathrm{P}(X = i) = \frac{\lambda^i e^{-\lambda}}{i!}$$

□

Theorem 19. *Let*

$$X \sim \mathrm{Poi}(\lambda)$$

Then,

$$\mathrm{E}[X] = \lambda$$

Proof.

$$\begin{aligned} \mathrm{E}[X] &= \sum_{i=0}^{\infty} i \frac{e^{-\lambda}\lambda^i}{i!} \\ &= \lambda \sum_{i=1}^{\infty} \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} \end{aligned}$$

Let

$$j = i - 1$$

Therefore,

$$\mathrm{E}[X] = \lambda \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!}$$

As $\sum\limits_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!}$ represents the total probability of X,

$$\sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!} = 1$$

Therefore,

$$\mathrm{E}[X] = \lambda$$

□

Theorem 20. *Let*

$$X = \mathrm{Poi}(\lambda)$$

Then,

$$\mathrm{V}(X) = \lambda$$

Proof.

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \sum_{i=0}^{\infty} i^2 \frac{e^{-\lambda}\lambda^i}{i!}\\ &= \lambda \sum_{i=1}^{\infty} i \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!}\end{aligned}$$

Let

$$j = i - 1$$

Therefore,

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \lambda \sum_{j=0}^{\infty} (j+1) \frac{e^{-\lambda}\lambda^j}{j!}\\ &= \lambda \left(\sum_{j=0}^{\infty} j \frac{e^{-\lambda}\lambda^j}{j!} + \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!} \right)\\ &= \lambda \left(\mathrm{E}\left[\mathrm{Poi}(\lambda)\right] + \sum_{j=0}^{\infty} \mathrm{P}\left(\mathrm{Poi}(\lambda) = j\right) \right)\\ &= \lambda(\lambda + 1)\end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\ &= \lambda \end{aligned}$$

□

Exercise 34.
Consider an experiment that consists of counting the number of α particles given off in a second by a gram of radioactive material. If it is known that on average, 3.2 such α particles are emitted, what is the probability that no more than 2 α particles will be emitted?

Solution 34.
Let X be the number of α particles emitted. Therefore,

$$X \sim \mathrm{Poi}(3.2)$$

Therefore,

$$\begin{aligned} \mathrm{P}(X \leq 2) &= \mathrm{P}(X = 0) + \mathrm{P}(X = 1) + \mathrm{P}(X = 2) \\ &= \frac{e^{-3.2}\lambda^0}{0!} + \frac{e^{-3.2}\lambda^1}{1!} + \frac{e^{-3.2}\lambda^2}{2!} \\ &\approx 0.3799 \end{aligned}$$

2.1 Assumptions for Poisson Distributions for Events Over a Period of Time

1. The probability that an event occurs in an interval of length h is $\lambda h + \mathrm{o}(h)$.
2. For a small enough h, the probability that two or more events occur in an interval of length h is small, i.e. is $\mathrm{o}(h)$.
3. The number of events in intervals that are not overlapping are independent.

Exercise 35.
Consider that earthquakes occur with the assumptions of Poisson distributions, with $\lambda = 2$, and with a week as a unit of time.

1. Find the probability that at least three earthquakes occur during the next two weeks.

2. Given that three earthquakes occurred in the last four weeks, what is the probability that exactly one of them occurred in the last week.

Solution 35.

1. Let X be the number of earthquakes occurring in two weeks. Therefore,

$$X \sim \text{Poi}(4)$$

Therefore,

$$\begin{aligned}
\text{P}(X \geq 3) &= 1 - (\text{P}(X = 0) + \text{P}(X = 1) + \text{P}(X = 2)) \\
&= 1 - \left(\frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{21} \right) \\
&= 1 - 13e^{-4}
\end{aligned}$$

2. Let Y be the number of earthquakes in the last four weeks.
Let Z_1 be the number of earthquakes in the first three of the last four weeks.
Let Z_2 be the number of earthquakes in the last week.
Therefore,

$$\begin{aligned}
Y &\sim \text{Poi}(8) \\
Z_1 &\sim \text{Poi}(6) \\
Z_2 &\sim \text{Poi}(2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{P}(Z_2 = 1 | Y = 3) &= \frac{\text{P}(Z_2 = 1 \cap Y = 3)}{\text{P}(Y = 3)} \\
&= \frac{\text{P}(Z_2 = 1 \cap Z_1 = 2)}{\text{P}(Y = 3)}
\end{aligned}$$

Therefore, as Z_1 and Z_2 belong to non-overlapping intervals, the events $Z_2 = 1$ and $Z_1 = 2$ are independent. Therefore,

$$\begin{aligned}
\text{P}(Z_2 = 1 | Y = 3) &= \frac{\text{P}(Z_2 = 1)\,\text{P}(Z_1 = 2)}{\text{P}(Y = 3)} \\
&= \frac{\frac{e^{-2}2^1}{1!}\frac{e^{-6}6^2}{2!}}{\frac{e^{-8}8^3}{3!}} \\
&= \left({}^3\text{C}_1\right) \left(\frac{2}{8}\right)^1 \left(\frac{6}{8}\right)^2
\end{aligned}$$

3 Geometric Random Variables

Definition 26 (Geometric random variable)**.** Suppose that independent trials, each having probability of success $0 < p < 1$, are performed until a success occurs. If X is the number of trials required, then X is said to have a geometric distribution. It is denoted as

$$X \sim \mathrm{Geo}(p)$$

The probability distribution if X is

$$\mathrm{P}(X = n) = (1-p)^{n-1}p$$

Exercise 36.

Alice eats cookies one after another until she finds and a chocolate cookie. For each cookie, the probability of the cookie being a chocolate cookie is $\frac{1}{10}$.

1. What is the probability that Alice eats more than 3 cookies?

2. Given that Alice has already eaten 5 cookies, and has not found a chocolate cookie, what is the probability that she will eat at least 8 more cookies?

Solution 36.

1.

$$\begin{aligned}
\mathrm{P}(X > 3) &= \sum_{k=4}^{\infty} \mathrm{P}(X = k) \\
&= \sum_{k=4}^{\infty} \left(1 - \frac{1}{10}\right)^{k-1} \left(\frac{1}{10}\right) \\
&= \left(1 - \frac{1}{10}\right)^{3} \sum_{j=1}^{\infty} \left(1 - \frac{1}{10}\right)^{j-1} \left(\frac{1}{10}\right) \\
&= \left(\frac{9}{10}\right)^{3} \left(\frac{1}{10}\right) \left(\frac{1}{1 - \frac{9}{10}}\right) \\
&= \left(\frac{9}{10}\right)^{3}
\end{aligned}$$

2.

$$
\begin{aligned}
P(X \geq 13|X > 5) &= P(X > 12|X > 5) \\
&= \frac{P(X > 12 \cap X > 5)}{P(X > 5)} \\
&= \frac{P(X > 12)}{P(X > 5)} \\
&= \frac{\left(\frac{9}{10}\right)^{12}}{\left(\frac{9}{10}\right)^{5}} \\
&= \left(\frac{9}{10}\right)^{7} \\
&= P(X > 7)
\end{aligned}
$$

Therefore, the fact that Alice has already eaten 5 cookies does not affect the probability of her eating at least 8 more cookies.

Theorem 21. *Let*

$$X \sim \text{Geo}(p)$$

Then,

$$E[X] = \frac{1}{p}$$

Proof.

$$E[X] = \sum_{i=1}^{\infty} i(1-p)^{i-1}p$$

Let

$$q = 1 - p$$

Therefore,

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=1}^{\infty} i q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1+1) q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1) q^{i-1} p + \sum_{i=1}^{\infty} q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1) q^{i-1} p + \frac{p}{1-q} \\
&= \sum_{i=1}^{\infty} (i-1) q^{i-1} p + \frac{p}{1-(1-p)} \\
&= \sum_{i=1}^{\infty} (i-1) q^{i-1} p + 1 \\
&= \sum_{j=0}^{\infty} j q^{j} p + 1 \\
&= q \sum_{j=0}^{\infty} j q^{j-1} p + 1 \\
&= q\,\mathrm{E}[X] + 1 \\
\therefore \mathrm{E}[X](1-q) &= 1 \\
\therefore \mathrm{E}[X] &= \frac{1}{1-q} \\
&= \frac{1}{p}
\end{aligned}$$

□

Theorem 22. *Let*

$$X \sim \mathrm{Geo}(p)$$

Then,

$$\mathrm{V}(X) = \frac{1-p}{p^2}$$

Proof. Let

$$q = 1 - p$$

$$
\begin{aligned}
\mathrm{E}\left[X^2\right] &= \sum_{i=1}^{\infty} i^2 q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1+i)^2 q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1)^2 q^{i-1} p + \sum_{i=1}^{\infty} 2(i-1) q^{i-1} p + \sum_{i=1}^{\infty} q^{i-1} p \\
&= \sum_{j=0}^{\infty} j^2 q^j p + 2 \sum_{j=1}^{\infty} j q^j p + 1 \\
&= q\,\mathrm{E}\left[X^2\right] + 2q\,\mathrm{E}[X] + 1 \\
\therefore p\,\mathrm{E}\left[X^2\right] &= \frac{2q}{p} + 1 \\
\therefore \mathrm{E}\left[X^2\right] &= \frac{2q+p}{p^2} \\
&= \frac{q+1}{p^2}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= \frac{q+1}{p^2} - \frac{1}{p^2} \\
&= \frac{q}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}
$$

$\square$

4 Negative Binomial Random Variable

Definition 27 (Negative binomial random variable)**.** Suppose that independent trials, each having probability of success $0 < p < 1$, are performed until a total of r successes are accumulated. If X is the number of trials required, then X is said to be a negative binomial random variable.
It is denoted as

$$X \sim \mathrm{NB}(r, p)$$

The last trial must necessarily result in a success, and there must be $r-1$ more success in the first $n-1$ trials. Therefore, the probability distribution of X is

$$\begin{aligned}
\mathrm{P}(X=n) &= \binom{n-1}{r-1} p^{r-1}(1-p)^{n-r} p \\
&= \binom{n-1}{r-1} p^{r}(1-p)^{n-r}
\end{aligned}$$

Theorem 23.

$$n\binom{n-1}{r-1} = r\binom{n}{r}$$

Theorem 24. *Let*

$$X \sim \mathrm{NB}(r,p)$$

Then,

$$\mathrm{E}[X] = \frac{r}{p}$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[X^k\right] &= \sum_{n=r}^{\infty} n^k \binom{n-1}{r-1} p^r (1-p)^{n-r} \\
&= \frac{r}{p} \sum_{n=r}^{\infty} n^{k-1} \binom{n}{r} p^{r+1} (1-p)^{n-r} \\
&= \frac{r}{p} \sum_{m=r+1}^{\infty} (m-1)^{k-1} \binom{m-1}{r} p^{r+1} (1-p)^{m-(r+1)}
\end{aligned}$$

Let

$$Y \sim \mathrm{NB}(r+1,p)$$

Therefore,

$$\mathrm{E}\left[X^k\right] = \frac{r}{p}\,\mathrm{E}\left[(Y-1)^{k-1}\right]$$

Therefore,

$$\mathrm{E}[X] = \frac{r}{p}$$

□

Theorem 25. *Let*

$$X \sim \mathrm{NB}(r, p)$$

Then,

$$\mathrm{E}[X] = \frac{r}{p}$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[X^k\right] &= \sum_{n=r}^{\infty} n^k \binom{n-1}{r-1} p^r (1-p)^{n-r} \\
&= \frac{r}{p} \sum_{n=r}^{\infty} n^{k-1} \binom{n}{r} p^{r+1} (1-p)^{n-r} \\
&= \frac{r}{p} \sum_{m=r+1}^{\infty} (m-1)^{k-1} \binom{m-1}{r} p^{r+1} (1-p)^{m-(r+1)}
\end{aligned}$$

Let

$$Y \sim \mathrm{NB}(r+1, p)$$

Therefore,

$$\mathrm{E}\left[X^k\right] = \frac{r}{p} \mathrm{E}\left[(Y-1)^{k-1}\right]$$

Therefore,

$$\begin{aligned}
\mathrm{E}[X] &= \frac{r}{p} \\
\mathrm{E}\left[X^2\right] &= \frac{r}{p} \mathrm{E}[Y-1] \\
&= \frac{r}{p} \left(\frac{r+1}{p} - 1\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= \frac{r(1-p)}{p^2}
\end{aligned}$$

□

Exercise 37.
Find the expected value and variance of the number of time one must throw a die until the outcome 1 has occurred four times.

Solution 37.
Let X be the number of times the die must be thrown for 1 to occur four times. Therefore,

$$X \sim \mathrm{NB}\left(4, \frac{1}{6}\right)$$

Therefore,

$$\begin{aligned}
\mathrm{E}[X] &= \frac{r}{p} \\
&= \frac{4}{\frac{1}{6}} \\
&= 24 \\
\mathrm{V}(X) &= \frac{r(1-p)}{p^2} \\
&= \frac{4\left(1-\frac{1}{6}\right)}{\left(\frac{1}{6}\right)^2} \\
&= 120
\end{aligned}$$

5 Hypergeometric Random Variable

Definition 28 (Hypergeometric random variable)**.** Suppose that a sample of size n is to be chosen randomly and without replacement from a population of N, of which m possess a particular characteristic, and the other $N-m$ do not. If X is the number of individuals in the selected sample, then X is said to be a hypergeometric random variable.
It is denoted as

$$X \sim \mathrm{HG}(n, N, m)$$

The probability distribution of X is

$$\mathrm{P}(X = i) = \frac{\binom{n}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$$

Theorem 26. *Let*

$$X \sim \mathrm{HG}(n, N, m)$$

Then,

$$\mathrm{E}[X] = \frac{nm}{N}$$

Theorem 27. *Let*

$$X \sim \mathrm{HG}(n, N, m)$$

Then,

$$\mathrm{V}(X) = n\frac{m}{N}\left(1 - \frac{m}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Exercise 38.
A buyer of electrical component buys electrical components in lots of size 10. It is his policy to inspect 3 components randomly from a lot, and to accept the lot only if all 3 are non-defective. If 30% of the lots have 4 defective components and 70% of the lots have 1 defective components, what is the proportion of the lots that the purchaser rejects.

Solution 38.
Let A be the event that the buyer accepts the lot. Therefore,

$$\begin{aligned}
\mathrm{P}(A) &= \mathrm{P}(4 \text{ defective})\,\mathrm{P}(A|4 \text{ defective}) + \mathrm{P}(1 \text{ defective})\,\mathrm{P}(A|1 \text{ defective}) \\
&= \left(\frac{3}{10}\right)\left(\frac{{}^{4}\mathrm{C}_{0}{}^{6}\mathrm{C}_{3}}{{}^{10}\mathrm{C}_{3}}\right) + \left(\frac{7}{10}\right)\left(\frac{{}^{1}\mathrm{C}_{0}{}^{9}\mathrm{C}_{3}}{{}^{10}\mathrm{C}_{3}}\right) \\
&= \frac{54}{100}
\end{aligned}$$

Therefore, the buyer rejects 54% of the lots, i.e., he rejects 46% of the lots.

6 Uniform Random Variable

Definition 29 (Uniform Random Variable)**.** A random variable X is said to be a uniform random variable over the interval (a, b) if its probability density function is

$$\mathrm{f}(x) = \begin{cases} \frac{1}{b-a} & ; \quad a < x < b \\ 0 & ; \quad \text{otherwise} \end{cases}$$

It is denoted as

$$X \sim \mathrm{U}(a, b)$$

Theorem 28. *The cumulative distribution function of a uniform random variable X is*

$$\mathrm{F}_X(x) = \begin{cases} 0 & ; \quad x < a \\ \frac{x-a}{b-a} & ; \quad a \leq x \leq b \\ 1 & ; \quad b < x \end{cases}$$

Theorem 29. *Let*

$$X \sim \mathrm{U}(a, b)$$

Then,

$$\mathrm{E}[X] = \frac{a+b}{2}$$

Proof.

$$\begin{aligned} \mathrm{E}[X] &= \int\limits_{-\infty}^{\infty} x \, \mathrm{f}(x) \, \mathrm{d}x \\ &= \int\limits_{a}^{b} x \frac{1}{b-a} \, \mathrm{d}x \\ &= \frac{a+b}{2} \end{aligned}$$

□

Theorem 30. *Let*

$$X \sim \mathrm{U}(a, b)$$

Then,

$$\mathrm{V}(X) = \frac{(b-a)^2}{12}$$

Proof.

$$\begin{aligned} \mathrm{E}\left[X^2\right] &= \int\limits_{-\infty}^{\infty} x^2 \, \mathrm{f}\left(x^2\right) \mathrm{d}x \\ &= \int\limits_{a}^{b} x^2 \frac{1}{b-a} \\ &= \frac{b^3 - a^3}{3(b-a)} \\ &= \frac{a^2 + ab + b^2}{3} \end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

□

7 Exponential Random Variable

Definition 30 (Uniform Random Variable)**.** A random variable X is said to be a exponential random variable over the interval (a, b) if its probability density function is

$$\mathrm{f}(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

It is denoted as

$$X \sim \mathrm{Exp}(\lambda)$$

Theorem 31. *The cumulative distribution function of a exponential random variable X with parameter λ is*

$$\mathrm{F}_X(x) = \begin{cases} 1 - e^{-\lambda x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

Theorem 32. *Let*

$$X \sim \mathrm{Exp}(\lambda)$$

Then,

$$\mathrm{E}[X] = \frac{1}{\lambda}$$

Theorem 33. *Let*

$$X \sim \mathrm{Exp}(\lambda)$$

Then,

$$\mathrm{E}\left[X^n\right] = \frac{n!}{\lambda^n}$$

Theorem 34. *Let*

$$X \sim \mathrm{Exp}(\lambda)$$

Then,

$$\mathrm{V}(X) = \frac{1}{\lambda^2}$$

8 Normal Distribution

Definition 31 (Normal Random Variable)**.** A random variable X is said to be a normal random variable if its probability density function is

$$\mathrm{f}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ and σ^2 are parameters.
If

$$\begin{aligned} \mu &= 0 \\ \sigma^2 &= 1 \end{aligned}$$

then X is said to be a standard normal random variable.
The cumulative distribution function of a standard normal X is denoted by $\Phi(x)$.

Theorem 35. *Let X be a standard normal random variable. Then,*

$$\Phi(-x) = 1 - \Phi(x)$$

Theorem 36. *If X is normally distributed with parameters μ and σ^2, then $Y = ax + b$ where $a > 0$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$.*

Proof.

$$\begin{aligned} \mathrm{F}_Y(y) &= \mathrm{P}(Y \leq y) \\ &= \mathrm{P}(ax + b \leq y) \\ &= \mathrm{P}\left(x \leq \frac{y-b}{a}\right) \\ &= \mathrm{F}_X\left(\frac{y-b}{a}\right) \end{aligned}$$

Therefore, differentiating,

$$\begin{aligned}
\mathrm{f}_Y(y) &= \frac{1}{a}\,\mathrm{f}_X\left(\frac{y-b}{a}\right)\\
&= \frac{1}{\sqrt{2\pi}a\sigma}e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}\\
&= \frac{1}{\sqrt{2\pi}a\sigma}e^{-\frac{(y-b-a\mu)^2}{2(a\sigma)^2}}
\end{aligned}$$

□

Theorem 37. *If X is normally distributed with parameters μ and σ^2, then $Z = \frac{X-\mu}{\sigma}$ is normally distributed with parameters 0 and 1.*

Exercise 39.
Let X be a normal random variable with parameters

$$\begin{aligned}
\mu &= 3\\
\sigma^2 &= 9
\end{aligned}$$

Find

1. $\mathrm{P}(2 < X < 5)$
2. $\mathrm{P}(X > 3)$

Solution 39.
1. Let

$$Z = \frac{X-\mu}{\sigma}$$

Therefore, Z is a standard normal random variable.
Therefore,

$$\begin{aligned}
\mathrm{P}(2 < X < 5) &= \mathrm{P}\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right)\\
&= \mathrm{P}\left(-\frac{1}{3} < Z < \frac{2}{3}\right)\\
&= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)\\
&= \Phi\left(\frac{2}{3}\right) - \left(1 - \Phi\left(\frac{1}{3}\right)\right)\\
&\approx 0.3779
\end{aligned}$$

2. Let

$$Z = \frac{X - \mu}{\sigma}$$

Therefore, Z is a standard normal random variable.
Therefore,

$$\begin{aligned}
\mathrm{P}(X > 3) &= \mathrm{P}\left(\frac{X-3}{3} > \frac{3-3}{3}\right) \\
&= \mathrm{P}(Z > 0) \\
&= 0.5
\end{aligned}$$

Theorem 38. *Let Z be a standard normal random variable. Then,*

$$\mathrm{E}[X] = 0$$

Proof.

$$\begin{aligned}
\mathrm{E}[Z] &= \int\limits_{-\infty}^{\infty} z\,\mathrm{f}(z)\,\mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} ze^{-\frac{z^2}{2}}\,\mathrm{d}x \\
&= 0
\end{aligned}$$

□

Theorem 39. *Let Z be a standard normal random variable. Then,*

$$\mathrm{V}(X) = 1$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[Z^2\right] &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}}\,\mathrm{d}z \\
&= 1
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= 1 - 0 \\
&= 1
\end{aligned}$$

□

Theorem 40. *Let X be a normal random variable. Then,*

$$E[X] = \mu$$

Proof. Let

$$Z = \frac{X - \mu}{\sigma}$$

Therefore,

$$X = Z\sigma + \mu$$

Therefore,

$$\begin{aligned} E[X] &= E[Z]\sigma + \mu \\ &= \mu \end{aligned}$$

□

Theorem 41. *Let X be a normal random variable. Then,*

$$V(X) = \sigma^2$$

Proof. Let

$$Z = \frac{X - \mu}{\sigma}$$

Therefore,

$$X = Z\sigma + \mu$$

Therefore,

$$\begin{aligned} V(X) &= \sigma^2 V(Z) \\ &= \sigma^2 \end{aligned}$$

□

9 de Moivre-Laplace Limit Theorem

Theorem 42 (de Moivre-Laplace Limit Theorem)**.** *Let S_n be the number of successes that occur when n independent trials, each with probability of success p, are performed. Then, for any $a < b$,*

$$\lim_{n\to\infty} P\left(a < \frac{S_n - np}{\sqrt{np(1-p)}} < b\right) = \Phi(b) - \Phi(a)$$

Exercise 40.
To determine the effectiveness of a certain diet in reducing the amount of cholesterol in the bloodstream, 100 people are put on the diet. A month later, the cholesterol is measured.
The nutritionist in charge of this experiment has decided to endorse this diet if at least 65% of the people have lower cholesterol levels than before.
What is the probability to endorse the diet if it has no effect on cholesterol levels?

Solution 40.
If the diet has no effect on cholesterol levels,

$$X \sim \text{Bin}\left(100, \frac{1}{2}\right)$$

Therefore,

$$\begin{aligned}
\mathrm{P}(X \geq 65) &= \mathrm{P}\left(\frac{X-(100)\left(\frac{1}{2}\right)}{\sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}} \geq \frac{65-(100)\left(\frac{1}{2}\right)}{\sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}}\right) \\
&\approx (Z \geq 2.9) \\
&\approx 0.0019
\end{aligned}$$

Part V
Combination of Random Variables

1 Expected Value of Sums of Random Variables

Theorem 43. *Let*

$$X = \sum_{i=1}^{n} X_i$$

where each X_i is a random variable, possibly of different distributions.
Then,

$$\mathrm{E}[X] = \sum_{i=1}^{n} \mathrm{E}[X_i]$$

Proof.

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{\omega \in \Omega} X(\omega)\,\mathrm{P}(\omega) \\
&= \sum_{\omega \in \Omega} \left(\sum_{i=1}^{n} X_i(\omega) \right) \mathrm{P}(\omega) \\
&= \sum_{i=1}^{n} \sum_{\omega}^{\Omega} X_i(\omega)\,\mathrm{P}(\omega) \\
&= \sum_{i=1}^{n} \mathrm{E}[X_i]
\end{aligned}$$

□

Exercise 41.
n people came to a winter party, carrying a coat each. While leaving the party, each person took a coat randomly. Let X be the number of people that left with their own coat. Find $\mathrm{E}[X]$.

Solution 41.
Let

$$X_i = \begin{cases} 1 & ; \quad \text{person } i \text{ got own coat} \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Therefore,

$$\mathrm{P}(X_i = 1) = \frac{1}{n}$$

Therefore,

$$\begin{aligned} \mathrm{E}[X] &= \sum_{i=1}^{n} \mathrm{E}[X_i] \\ &= \sum_{i=1}^{n} \frac{1}{n} \\ &= 1 \end{aligned}$$

Part VI

Jointly Distributed Random Variables

1 Joint Cumulative Probability Distribution Function

Definition 32. For any two random variables X and Y, the joint cumulative probability distribution function of X and Y is defined to be

$$\mathrm{F}_{X,Y}(a,b) = \mathrm{P}(X \leq a, Y \leq b)$$

where $a \in \mathbb{R}$ and $b \in \mathbb{R}$.

Theorem 44. *If the joint cumulative probability distribution function of X and Y is $\mathrm{F}_{X,Y}(a,b)$, then the probability distribution function of X is*

$$\mathrm{F}_X(a) = \lim_{b \to \infty} \mathrm{F}_{X,Y}(a,b)$$

Exercise 42.
3 balls are randomly selected from an urn containing 3 red, 4 white, and 5 blue balls.
Let X be the number of red balls chosen.
Let Y be the number of white balls chosen.
Find the joint probability mass function of X and Y.

Solution 42.

$$\mathrm{P}(X = x, Y = y) = \frac{{}^{3}\mathrm{C}_{x}\,{}^{4}\mathrm{C}_{y}\,{}^{5}\mathrm{C}_{3-x-y}}{{}^{12}\mathrm{C}_{3}}$$

Therefore,

	0	1	2	3	
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{10}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{10}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{10}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{10}{220}$
	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

2 Joint Continuous Variables

Definition 33 (Joint continuity and joint probability density function)**.** X and Y are said to be jointly continuous if there exists a function $\mathrm{f}(x, y)$ defined for all real x and y, such that for every set C of pairs of real numbers,

$$\mathrm{P}\left((x, y) \in C\right) = \iint\limits_{(x,y)\in C} \mathrm{f}(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

The function $\mathrm{f}(x, y)$ is called the joint probability density function of X and Y. The joint cumulative probability distribution function is

$$\begin{aligned}\mathrm{F}_{X,Y}(a, b) &= \mathrm{P}(-\infty \leq X \leq a, -\infty \leq Y \leq b)\\ &= \int\limits_{-\infty}^{b}\int\limits_{-\infty}^{a} \mathrm{f}(x, y)\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Therefore,

$$\mathrm{f}(a, b) = \frac{\partial^2}{\partial a \partial b} F(a, b)$$

Exercise 43.
The joint density function of X and Y is given by

$$\mathrm{f}(x, y) = \begin{cases} 2e^{-x}e^{-2y} & ; \quad 0 < x < \infty, 0 < y < \infty \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Compute

1. $\mathrm{P}(X > 1, Y < 1)$
2. $\mathrm{P}(X < Y)$
3. $\mathrm{P}(X < a)$

Solution 43.

1.

$$
\begin{aligned}
\mathrm{P}(X > 1, Y < 1) &= \int\limits_{-\infty}^{1} \int\limits_{1}^{\infty} \mathrm{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int\limits_{0}^{1} \int\limits_{1}^{\infty} 2e^{-x}e^{-2y}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int\limits_{0}^{1} 2e^{-2y} \left. -e^{-x} \right|_{x=1}^{x=\infty} \mathrm{d}y \\
&= e^{-1} \int\limits_{0}^{1} 2e^{-2y}\,\mathrm{d}y \\
&= e^{-1}\left(1 - e^{-2}\right)
\end{aligned}
$$

2.

$$
\begin{aligned}
\mathrm{P}(X < Y) &= \iint\limits_{(x,y):x<y} \mathrm{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint\limits_{(x,y):x<y} 2e^{-x}e^{-2y}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int\limits_{0}^{\infty} \int\limits_{0}^{y} 2e^{-x}e^{-2y}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int\limits_{0}^{\infty} 2e^{-2y}\left(1 - e^{-y}\right)\mathrm{d}y \\
&= \int\limits_{0}^{\infty} 2e^{-2y}\,\mathrm{d}y - \int\limits_{0}^{\infty} 2e^{-3y}\,\mathrm{d}y \\
&= 1 - \frac{2}{3} \\
&= \frac{1}{3}
\end{aligned}
$$

3.

$$\begin{aligned}
\mathrm{P}(X < a) &= \int_{-\infty}^{a} \int_{-\infty}^{\infty} \mathrm{f}(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{0}^{a} \int_{0}^{\infty} 2e^{-2y}e^{-x}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{0}^{a} e^{-x}\,\mathrm{d}x \\
&= 1 - e^{-a}
\end{aligned}$$

Exercise 44.
The joint density function of X and Y is given by

$$\mathrm{f}(x,y) = \begin{cases} e^{-(x+y)} & ; \quad 0 < x < \infty, 0 < y < \infty \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find the density function of the random variable $\frac{X}{Y}$.

Solution 44.

$$\begin{aligned}
\mathrm{F}_{\frac{X}{Y}}(a) &= \mathrm{P}\left(\frac{X}{Y} \leq a\right) \\
&= \iint\limits_{(x,y):\frac{x}{y}\leq a} \mathrm{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint\limits_{(x,y):\frac{x}{y}\leq a} e^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{0}^{\infty} \int_{0}^{ay} e^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{0}^{\infty} \left(1 - e^{-ay}\right) e^{-y}\,\mathrm{d}y \\
&= -e^{-y} + \left.\frac{e^{-(a+1)y}}{a+1}\right|_{0}^{\infty} \\
&= 1 - \frac{1}{a+1}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{f}_{\frac{X}{Y}}(a) &= \frac{\mathrm{dF}_{\frac{X}{Y}}(a)}{\mathrm{d}a} \\
&= \frac{1}{(a+1)^2}
\end{aligned}$$

3 Independent Random Variables

Definition 34 (Independent random variables)**.** Two random variables X and Y are said to be independent if

$$\mathrm{P}(X \in A, Y \in B) = \mathrm{P}(X \in A)\,\mathrm{P}(Y \in B)$$

where $A \subseteq \mathbb{R}$, $B \subseteq \mathbb{R}$.

Theorem 45. *X and Y are independent random variables if and only if*

$$\mathrm{F}_{X,Y}(a, b) = \mathrm{F}_X(a)\mathrm{F}_Y(a)$$

Proof. Let

$$\begin{aligned}
A &= (-\infty, a] \\
B &= (-\infty, b]
\end{aligned}$$

X and Y are independent if and only if

$$\begin{aligned}
&\mathrm{P}(X \in A, Y \in B) = \mathrm{P}(X \in A)\,\mathrm{P}(Y \in B) \\
\iff &\mathrm{P}(X \leq a, Y \leq b) = \mathrm{P}(X \leq a)\,\mathrm{P}(Y \leq b) \\
\iff &\mathrm{F}_{X,Y}(a, b) = \mathrm{F}_X(a)\mathrm{F}_Y(b)
\end{aligned}$$

□

Theorem 46. *If X and Y are discrete, then*

$$\mathrm{P}(X = x, Y = y) = \mathrm{P}(X = x)\,\mathrm{P}(Y = y)$$

for all (x, y), if and only if X and Y are independent.

Theorem 47. *If X and Y are continuous, then*

$$\mathrm{f}(x, y) = \mathrm{f}_X(x)\,\mathrm{f}_Y(y)$$

for all (x, y), if and only if X and Y are independent.

Exercise 45.
Let

$$
\begin{aligned}
X &\sim \mathrm{U}(0,1) \\
Y &\sim \mathrm{U}(0,1)
\end{aligned}
$$

Calculate the probability density function of $X+Y$

Solution 45.
Therefore,

$$
\begin{aligned}
\mathrm{F}_{X+Y}(t) &= \mathrm{P}(X+Y \le t) \\
&= \begin{cases} \frac{t^2}{2} & ; \quad 0 \le t \le 1 \\ 1 - \frac{(2-t)^2}{2} & ; \quad 1 \le t \le 2 \end{cases}
\end{aligned}
$$

Therefore,

$$
\mathrm{f}_{X+Y}(t) = \begin{cases} t & ; \quad 0 \le t \le 1 \\ 2-t & ; \quad 1 < t < 2 \\ 0 & ; \quad 2 \le t \end{cases}
$$

Theorem 48. *If X_i, for $i \in \mathbb{N}$ are independent normal random variables, with parameters μ_i and ${\sigma_i}^2$ respectively, then $\sum\limits_{i=1}^{n} x_i$ is a normal random variable with parameters $\sum\limits_{i=1}^{n} \mu_i$ and $\sum\limits_{i=1}^{n} {\sigma_i}^2$.*

Theorem 49. *Let*

$$
\begin{aligned}
X &= \mathrm{Poi}(\lambda_1) \\
Y &= \mathrm{Poi}(\lambda_2)
\end{aligned}
$$

be independent random variables. Then,

$$
X+Y \sim \mathrm{Poi}(\lambda_1+\lambda_2)
$$

Proof.

$$
\begin{aligned}
\mathrm{P}(X+Y=n) &= \sum_{k=0}^{n} \mathrm{P}(X=k, Y=n-k) \\
&= \sum_{k=0}^{n} \mathrm{P}(X=k)\,\mathrm{P}(Y=n-k) \\
&= \sum_{k=0}^{n} \frac{e^{-\lambda_1}{\lambda_1}^k}{k!}\frac{e^{-\lambda_2}{\lambda_2}^{n-k}}{(n-k)!} \\
&= e^{-\lambda_1-\lambda_2}\sum_{k=0}^{n} \frac{{\lambda_1}^k{\lambda_2}^{n-k}}{k!(n-k)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}{\lambda_1}^k{\lambda_2}^{n-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n} {}^n\mathrm{C}_k{\lambda_1}^k{\lambda_2}^{n-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n
\end{aligned}
$$

Therefore,

$$X+Y \sim \mathrm{Poi}(\lambda_1+\lambda_2)$$

□

Theorem 50. *Let*

$$
\begin{aligned}
X &\sim \mathrm{Bin}(n,p) \\
Y &\sim \mathrm{Bin}(m,p)
\end{aligned}
$$

be independent random variables. Then,

$$X+Y \sim \mathrm{Bin}(n+m,p)$$

Proof.

$$\begin{aligned}
\mathrm{P}(X+Y=k) &= \sum_{i=0}^{n} \mathrm{P}(X=i, Y=k-i) \\
&= \sum_{i=0}^{n} \mathrm{P}(X=i)\,\mathrm{P}(X=i)\,\mathrm{P}(Y=k-i) \\
&= \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\
&= p^k (1-p)^{n+m-k} \sum_{i=0}^{n} \binom{n}{i}\binom{m}{k-i} \\
&= p^k (1-p)^{n+m-k} \binom{n+m}{k}
\end{aligned}$$

Therefore,

$$X+Y \sim \mathrm{Bin}(n+m, p)$$

□

Theorem 51. *Let*

$$X_i \sim \mathrm{Geo}(p)$$

be independent random variables, for $i \in \mathbb{N}$. Then,

$$\sum_{i=1}^{n} X_i \sim \mathrm{NB}(n, p)$$

4 Conditional Random Variables

Theorem 52. *The conditional probability mass function of two discrete random variables X and Y is*

$$\mathrm{P}(X=x|Y=y) = \frac{\mathrm{P}(X=x, Y=y)}{\mathrm{P}(Y=y)}$$

Theorem 53. *The conditional probability density function of two continuous random variables X and Y is*

$$\mathrm{f}_{X|Y}(x|y) = \frac{\mathrm{f}(x,y)}{\mathrm{f}_Y(y)}$$

and the cumulative distribution function is

$$\mathrm{F}_{X|Y}(a|y) = \int_{-\infty}^{a} \mathrm{f}_{X|Y}(x|y)\,\mathrm{d}x$$

Exercise 46.

$$\mathrm{f}(x,y) = \begin{cases} \frac{e^{-\frac{x}{y}}e^{-y}}{y} & ; \quad 0 < x < \infty, 0 < y < \infty \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find $\mathrm{P}(X > 1|Y = y)$.

Solution 46.

$$\begin{aligned} \mathrm{f}_{X|Y}(x|y) &= \frac{\mathrm{f}(x,y)}{\mathrm{f}_Y(y)} \\ &= \frac{\frac{e^{-\frac{x}{y}}e^{-y}}{y}}{e^{-y}\int\limits_0^\infty \left(\frac{1}{y}\right) e^{-\frac{x}{y}}\,\mathrm{d}x} \\ &= \frac{e^{-\frac{x}{y}}}{y} \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{P}(X > 1|Y = y) &= \int\limits_1^\infty \frac{1}{y}e^{-\frac{x}{y}}\,\mathrm{d}x \\ &= e^{-\frac{1}{y}} \end{aligned}$$

5 Properties of Expectation

Theorem 54. *If X and Y have a joint probability mass function $\mathrm{P}(X = x, Y = y)$, then*

$$\mathrm{E}\left[g(X,Y)\right] = \sum_{x=-\infty}^{\infty}\sum_{y=-\infty}^{\infty} g(x,y)\,\mathrm{P}(X = x, Y = y)$$

Theorem 55. *If X and Y have a joint probability density function $\mathrm{f}(x,y)$, then*

$$\mathrm{E}\left[g(X,Y)\right] = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} g(x,y)\,\mathrm{f}(x,y)$$

Exercise 47.
An accident occurs at a point X that is uniformly distributed on a road of length L. At the time of the accident, and ambulance is at a location Y that is also uniformly distributed on the same road.
Assuming that X and Y are independent, find the expected distance between the point of occurrence of the accident, and the position of the ambulance.

Solution 47.

$$
\begin{aligned}
X &\sim \mathrm{U}(0, L) \\
Y &\sim \mathrm{U}(0, L)
\end{aligned}
$$

Therefore,

$$
\mathrm{f}_X(x) = \begin{cases} \frac{1}{L} & ; \quad 0 < x < L \\ 0 & ; \quad \text{otherwise} \end{cases}
$$

$$
\mathrm{f}_Y(y) = \begin{cases} \frac{1}{L} & ; \quad 0 < x < L \\ 0 & ; \quad \text{otherwise} \end{cases}
$$

Therefore, as the variables are independent,

$$
\begin{aligned}
\mathrm{f}_{X,Y}(x, y) &= \mathrm{f}_X(x) \, \mathrm{f}_Y(y) \\
&= \begin{cases} \frac{1}{L^2} & ; \quad 0 < x < L, 0 < y < L \\ 0 & ; \quad \text{otherwise} \end{cases}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{E}\left[|X - Y|\right] &= \frac{1}{L^2} \int\limits_0^L \int\limits_0^L |x - y| \, \mathrm{d}y \, \mathrm{d}x \\
&= \frac{1}{L^2} \int\limits_0^L \left(\int\limits_0^x (x - y) \, \mathrm{d}y + \int\limits_x^L (y - x) \, \mathrm{d}y \right) \mathrm{d}x \\
&= \frac{1}{L^2} \int\limits_0^L \left(\frac{L^2}{2} + x^2 - xL \right) \mathrm{d}x \\
&= \frac{L}{3}
\end{aligned}
$$

Definition 35 (Sample)**.** Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having distribution function function F, and expected value μ. Such a sequence of random variables is said to constitute a sample from the distribution F.

Definition 36 (Sample mean)**.** Let $X_1, \ldots, X_n$ be a sample with F and μ.

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

is called the sample mean of the sample.

Theorem 56. *Let $X_1, \ldots, X_n$ be a sample with* F *and μ. Then*

$$\mathrm{E}\left[\overline{X}\right] = \mu$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[\overline{X}\right] &= \mathrm{E}\left[\frac{\sum\limits_{i=1}^{n} X_i}{n}\right] \\
&= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[X_i] \\
&= \frac{1}{n} n\mu \\
&= \mu
\end{aligned}$$

□

Exercise 48.
Suppose that there are N different types of coupons, and each time one obtains a coupon, it is equally likely to be any one of the N types. Find the expected number of coupons that one needs to collect before obtaining a complete set.

Solution 48.
Let X be the number of coupons collected.
Let X_i be the number of additional coupons that need to be collected after i distinct types have been collected in order to obtain another distinct type.

The probability that a coupon will be of a new type, when i distinct types have already been collected is $\frac{N-i}{N}$. Therefore,

$$P(X_i = k) = \left(\frac{i}{N}\right)^{k-1} \frac{N-i}{k}$$

Therefore,

$$X_i \sim \text{Geo}\left(\frac{N-i}{N}\right)$$

Therefore,

$$E[X_i] = \frac{N}{N-i}$$

Therefore,

$$\begin{aligned} E[X] &= \sum_{i=0}^{n} E[X_i] \\ &= \sum_{i=0}^{n} \frac{N}{N-i} \\ &= 1 + \frac{N}{N-1} + \cdots + \frac{N}{1} \\ &= N\left(1 + \cdots + \frac{1}{N-1} + \frac{1}{N}\right) \end{aligned}$$

Exercise 49.
A sequence of n 1s and m 0s is randomly permuted. Any consecutive string of 1s is said to constitute a run of 1s. Compute the mean number of such runs.

Solution 49.
Let

$$I_i = \begin{cases} 1 & ; \quad \text{run of 1s starts at the } i\text{th position} \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Let

$$R(1) = \sum_{i=1}^{n+m} I_i$$

Therefore,

$$\begin{aligned} \mathrm{E}\left[R(1)\right] &= \mathrm{E}\left[\sum_{i=1}^{n+m} I_i\right] \\ &= \sum_{i=1}^{n+m} \mathrm{E}[I_i] \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{E}[I_1] &= \mathrm{P}(1 \text{ in position } 1) \\ &= \frac{n}{n+m} \\ \mathrm{E}[I_i] &= \mathrm{P}(0 \text{ in position } i-1, 1 \text{ in position } i) \\ &= \frac{m}{n+m}\frac{n}{n+m-1} \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{E}\left[R(1)\right] &= \mathrm{E}[I_1] + (n+m-1)\,\mathrm{E}[I_i] \\ &= \frac{n}{n+m} + (n+m-1)\left(\frac{nm}{(n+m)(n+m-1)}\right) \end{aligned}$$

6 Properties of Variance

Definition 37 (Covariance)**.** Let X and Y be random variables. Then

$$\mathrm{Cov}(X,Y) = \mathrm{E}\Big[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])\Big]$$

is defined to be the covariance of X and Y.

Theorem 57.

$$\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$$

Proof.

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathrm{E}\Big[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])\Big] \\ &= \mathrm{E}\left[XY - \mathrm{E}[X]Y - X\,\mathrm{E}[Y] + \mathrm{E}[X]\,\mathrm{E}[Y]\right] \\ &= \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] - \mathrm{E}[X]\,\mathrm{E}[Y] + \mathrm{E}[X]\,\mathrm{E}[Y] \\ &= \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] \end{aligned}$$

□

Theorem 58.

$$\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{Cov}(X,Y) + \mathrm{V}(Y)$$

Proof.

$$\mathrm{V}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right]$$

Therefore,

$$\begin{aligned}
\mathrm{V}(X+Y) &= \mathrm{E}\left[(X + Y - \mathrm{E}[X+Y])^2\right] \\
&= \mathrm{E}\left[(X + Y - \mathrm{E}[X] - \mathrm{E}[Y])^2\right] \\
&= \mathrm{E}\left[(X - \mathrm{E}[X])^2 + 2\,(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y]) + (Y - \mathrm{E}[Y])^2\right] \\
&= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] + 2\,\mathrm{E}\left[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])\right] + \mathrm{E}\left[(Y - \mathrm{E}[Y])^2\right] \\
&= \mathrm{V}(X) + \mathrm{Cov}(X,Y) + \mathrm{V}(Y)
\end{aligned}$$

□

Definition 38 (Pearson corellation coefficient)**.**

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)\,\mathrm{V}(Y)}}$$

Theorem 59. *If X and Y are independent, then, for any function g and h,*

$$\mathrm{E}\left[g(X)h(Y)\right] = \mathrm{E}\left[g(X)\right]\mathrm{E}\left[h(Y)\right]$$

Theorem 60. *Let X and Y be jointly continuous with joint density function $\mathrm{f}\,x, y$. Then,*

$$\mathrm{E}\left[g(X)h(Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\,\mathrm{f}_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

As X and Y are independent,

$$\begin{aligned}
\mathrm{E}\left[g(X)h(Y)\right] &= \int_{-\infty}^{\infty} g(x)h(y)\,\mathrm{f}_X(x)\,\mathrm{f}_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} g(x)\,\mathrm{f}_X(x)\,\mathrm{d}x \int_{-\infty}^{\infty} h(y)\,\mathrm{f}_Y(y)\,\mathrm{d}y \\
&= \mathrm{E}\left[g(X)\right]\mathrm{E}\left[h(Y)\right]
\end{aligned}$$

Theorem 61. *If X and Y are independent, then*

$$\mathrm{Cov}(X, Y) = 0$$

However, the converse is not true.

Proof.

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$$

As X and Y are independent,

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathrm{E}[X]\,\mathrm{E}[Y] - \mathrm{E}[X]\,\mathrm{E}[Y] \\ &= 0\end{aligned}$$

□

Definition 39 (Uncorrelated random variables)**.** X and Y are said to uncorrelated if and only if

$$\mathrm{Cov}(X, Y) = 0$$

Theorem 62.

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$$

Theorem 63.

$$\mathrm{Cov}(X, X) = \mathrm{V}(X)$$

Theorem 64.

$$\mathrm{Cov}(aX, Y) = \mathrm{V}(X)$$

Theorem 65.

$$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j)$$

Theorem 66.

$$\mathrm{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{V}(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{i} \mathrm{Cov}(X_i, X_j)$$

Proof.

$$
\begin{aligned}
\mathrm{V}\left(\sum_{i=1}^{n} X_i\right) &= \mathrm{Cov}(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}\left(X_i, X_j\right) \\
&= \sum_{i=1}^{n} \mathrm{V}(X_i) + \sum_{i=1}^{n}\sum_{j\neq i} \mathrm{Cov}(X_i, X_j) \\
&= \mathrm{V}\left(\sum_{i=1}^{n} X_i\right) & = \sum_{i=1}^{n} \mathrm{V}(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{i} \mathrm{Cov}(X_i, X_j)
\end{aligned}
$$

□

Definition 40 (Sample variance)**.** Let $X_1, \ldots, X_n$ be a sample with expected value μ and variation σ^2.

$$
S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1}
$$

is called the sample variance of the sample.

Exercise 50.

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having expected value μ, and variance σ^2.

Let

$$
\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}
$$

Find $V\left(\overline{X}\right)$ and $\mathrm{E}\left[S^2\right]$.

Solution 50.

$$
\begin{aligned}
\mathrm{V}\left(\overline{X}\right) &= \mathrm{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{1}{n}\right)^2 \mathrm{V}\left(\sum_{i=1}^{n} X_i\right)
\end{aligned}
$$

As all X_i are independent,

$$\begin{aligned}
\mathrm{V}\left(\overline{X}\right) &= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathrm{V}(X_i) \\
&= \left(\frac{1}{n}\right)^2 n\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

Therefore,

$$\begin{aligned}
(n-1)S^2 &= \sum_{i=1}^{n} \left(X_i - \mu + \mu + \overline{X}\right)^2 \\
&= \sum_{i=1}^{n} (X_i - \mu)^2 + \sum_{i=1}^{n} \left(\overline{X} - \mu\right)^2 - 2\left(\overline{X} - \mu\right) \sum_{i=1}^{n} (X_i - \mu) \\
&= \sum_{i=1}^{n} (X_i - \mu)^2 + \sum_{i=1}^{n} \left(\overline{X} - \mu\right)^2 - 2\left(\overline{X} - \mu\right) \sum_{i=1}^{n} \left(n\overline{X} - n\mu\right) \\
&= \sum_{i=1}^{n} (X_i - \mu)^2 + \sum_{i=1}^{n} \left(\overline{X} - \mu\right)^2 - 2\left(\overline{X} - \mu\right) \sum_{i=1}^{n} n\left(\overline{X} - \mu\right) \\
&= \sum_{i=1}^{n} (X_i - \mu)^2 - n\left(\overline{X} - \mu\right)^2
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}\left[(n-1)S^2\right] &= \mathrm{E}\left[\sum_{i=1}^{n} (X_i - \mu)^2\right] - n\,\mathrm{E}\left[\left(\overline{X} - \mu\right)^2\right] \\
&= \sum_{i=1}^{n} \mathrm{V}(X_i) - n\,\mathrm{V}\left(\overline{X}\right) \\
&= n\sigma^2 - n\,\mathrm{V}\left(\overline{X}\right) \\
&= n\sigma^2 - n\frac{\sigma^2}{n} \\
&= (n-1)\sigma^2
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{V}\left(S^2\right) &= \mathrm{E}\left[\left(S^2 - \mathrm{E}\left[S^2\right]\right)^2\right] \\
&= \sigma^2
\end{aligned}$$