

LIHORNE.COM

STAT 230

PROBABILITY



DR. JIWEI ZHAO • FALL 2013 • UNIVERSITY OF WATERLOO

 Last Revision: December 4, 2013

Table of Contents

1	Introduction to Probability	1
1.1	Definitions of Probability	1
1.2	Solutions to Problems on Chapter 1	2
2	Mathematical Probability Models	2
2.1	Sample Spaces and Probability	2
2.2	Solutions to Problems on Chapter 2	4
3	Probability - Counting Techniques	4
3.1	Counting Arguments	4
3.2	Review	8
3.3	Tutorial 1	9
3.4	Solutions to Problems on Chapter 3	11
4	Conditional	12
4.1	General Methods	12
4.2	Rules for Unions of Events	13
4.3	Intersections of Events and Independence	13
4.4	Conditional Probability	13
4.5	Multiplication and Partition Rules	14
4.6	Tutorial 2	15
4.7	Problems on Chapter 4	17
5	Probability Distributions	18
5.1	Solutions to Problems on Chapter 5	18
5.2	Tutorial 3	24
6	Chapter 7	27
6.1	Tutorial 4	27
6.2	Solutions to Problems on Chapter 7	30

7 Chapter 8	35
7.1 Solutions to Problems on Chapter 8	35
7.2 Tutorial 5	45
8 Continuous Probability Distributions	48
8.1 Solutions to Problems on Chapter 9	48

Abstract

These notes are intended as a resource for myself; past, present, or future students of this course, and anyone interested in the material. The goal is to provide an end-to-end resource that covers all material discussed in the course displayed in an organized manner. If you spot any errors or would like to contribute, please contact me directly.

1 Introduction to Probability

Why is it important to learn about Probability and Statistics? There are many mistakes made currently in the industry, for example, watch the Ted Talk on "How juries are fooled by statistics.". For example, consider the case of Sally Clark. For Wednesday, understand what a sample space is, an event is, and a probability distribution is.

1.1 Definitions of Probability

Definition 1.1 (Randomness). Uncertainty or **randomness** (i.e. variability of results) is usually due to some mixture of at least two factors including:

1. Variability in populations consisting of animate or inanimate objects (e.g., people vary in size, weight, blood type etc.)
2. Variability in processes or phenomena (e.g., the random selection of 6 numbers from 49 in a lottery draw can lead to a very large number of different outcomes)

Definition 1.2 (sample space). We refer to the set of all possible distinct outcomes to a random experiment as the sample space (usually denoted by S). Groups or sets of outcomes of possible interest, subsets of the sample space, we will call events.

Then we might define probability in three different ways:

Definition 1.3 (Probability (Classic), Relative Frequency, Subjective Probability). The classical definition: The **probability** of some event is

$$\frac{\text{number of ways the event can occur}}{\text{number of outcomes in } S}$$

provided all points in the sample space S are equally likely. For example, when a die is rolled the probability of getting a 2 is $\frac{1}{6}$ because one of the six faces is a 2.

The **relative frequency** definition: The probability of an event is the (limiting) proportion (or fraction) of times the event occurs in a very long series of repetitions of an experiment or process. For example, this definition could be used to argue that the probability of getting a 2 from a rolled die is $\frac{1}{6}$.

The **subjective probability** definition: The probability of an event is a measure of how sure the person making the statement is that the event will happen. For example, after considering all available data, a weather forecaster might say that the probability of rain today is 30% or 0.3.

There are problems with all three of these definitions. The first being that "equally likely" is not defined, the second being unfeasible, and third being hard to agree on. Thus we create a mathematical model with a set of axioms to better understand probability.

Definition 1.4 (probability model).

- a sample space of all possible outcomes of a random experiment is defined.
- a set of events, subsets of the sample space to which we can assign probabilities, is defined.
- a mechanism for assigning probabilities (numbers between 0 and 1) to events is specified.

1.2 Solutions to Problems on Chapter 1

- 1.1 The chance of head or tails; repetitive dice throwing in grade six; stock forecasts.
- 1.2 Claim on car insurance would be subjective based on someone's basic opinion, or relative considering how often other people have claims. A meltdown at nuclear power plant would be subjective since it doesn't happen often, isn't one of a set of probabilities, and is based on opinion or skepticism. A person's birthday is classical since there are 12 possibilities and April is 1 of 12.
- 1.3 Lottery draws are classical in that a ticket is one of several million, thus there is a value for $P(x)$ calculated using the formula above. Small businesses may use probability, or decision, trees to save money and time when deciding upon which areas of their company require an extensive review due to fraud or human error. There is subjective probability in understanding disease transmission, as well relative frequency since lots of data and samples can be taken. Public opinion polls can be subjectively evaluated, for example a presidential election (a type of public opinion poll) typically has forecasts by experts in politics and media.
- 1.4 Positions of small particles in space involve quantum physics, for example a particle must follow Heisenberg's Uncertainty Principle. The velocity of an object dropped from the leaning tower of Pisa does not depend on anything (if not acted on by another force). Values of stocks are not deterministic at all (though people try to think of them as such). Currency changes value steadily but it is not deterministic.

2 Mathematical Probability Models

2.1 Sample Spaces and Probability

Consider some repeatable phenomenon with certain events or outcomes labelled A_1, A_2, \dots, A_n are needed. We call this phenomenon an **experiment** and repetitions of it as a **trial**. The probability of an event if A is $0 \leq P(A) \leq 1$.

Definition 2.1 (Sample Space). A **sample space** S is a set of distinct outcomes for an experiment or process, with the property that in a single trial, one and only one of these outcomes occurs.

The outcomes that make up the sample space may sometimes be called **sample points** or just points on occasion.

Example 2.1. Roll a 6-sided die, and define the events $a_i =$ top face is i , for $i = 1, 2, 3, 4, 5, 6$. Then we can take a sample space as $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$. Rather than defining this sample space we could have also defined $S = \{E, O\}$ where E is the event that an even number comes up, and O the event that an odd number turns up. Both sample spaces satisfy the definition.

However, we try to choose sample spaces that are indivisible, so the first is preferred.

Definition 2.2. A **discrete** sample space is one that consists of a finite or countably infinite set of simple events. Recall that a countably infinite sequence is one that can be put in one-one correspondence with positive integers.

Definition 2.3. A non-discrete sample space is one that is not discrete.

Example 2.2. The sample space $S = \mathbb{Z}$ is discrete, $S = \mathbb{R}$ is not, but $S = \{1, 2, 3\}$ is.

Definition 2.4. An event in a discrete sample space is a subset $A \subset S$. If the event is indivisible so it contains only one point, e.g. $A_1 = \{a_1\}$ we call it a **simple event**. An event A made up of two or more simple events such as $A = \{a_1, a_2\}$ is called a compound event.

We will simplify $P(\{a_1\})$ to mean $P(a_1)$.

Definition 2.5. Let $S = \{a_1, a_2, a_3, \dots\}$ be a discrete sample space. Then **probabilities** $P(a_i)$ are numbers attached to the a_i 's ($i = 1, 2, 3, \dots$) such that the following two conditions hold:

- (1) $0 \leq P(a_i) \leq 1$
- (2) $\sum_i P(a_i) = 1$

The above function $P(a_i)$ is called a **probability distribution on S** .

Definition 2.6. The probability $P(A)$ of an event A is the sum of the probabilities for all the simple events that make up A or $P(A) = \sum_{a \in A} P(a)$.

Probability theory does not choose values for a_i but rather defines probability to be restricted to the above two mathematical rules. All else is experiment to mimic the real world values.

Example 2.3. For a dice we would assign $P(a_i)$ equal to $\frac{1}{6}$ for $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$. So, $P(E) = P(a_2, a_4, a_6) = 3 \times \frac{1}{6} = \frac{1}{2}$.

The typical approach to a given problem of these types is the following:

- (1) Specify a sample space S .
- (2) Assign numerical probabilities to the simple events in S .
- (3) For any compound event A , find $P(A)$ by adding the probabilities of all the simple events that make up A .

Example 2.4. Draw 1 card from a standard well-shuffled deck. Find the probability of finding a club.

Solution Let $S = \{ \text{spade, heart, diamond, club} \}$. S then has four sample points and one of them is club, so $P(\text{club}) = \frac{1}{4}$. We could also have defined $S = \{ \text{all cards} \}$ where a club event is $A = \{ \text{all clubs} \}$ and $P(A) = \sum P(\text{one club}) = \frac{1}{52} + \dots + \frac{1}{52} = \frac{13}{52} = \frac{1}{4}$.

Definition 2.7. The **odds** in favour of an event A is the probability the event occurs divided by the probability it does not occur or $\frac{P(A)}{1-P(A)}$. The odds against the event is the reciprocal.

When treating objects in an experiment as distinguishable leads to a different answer from treating them as identical, the points in the sample space for identical objects are usually not "equally likely" in terms of their long run relative frequencies. It is generally safer to pretend objects can be distinguished even when they can't be, in order to get equally likely sample points.

2.2 Solutions to Problems on Chapter 2

- 2.1 a) $S = \{a_1b_1, a_1b_2, a_1b_3, a_1b_4, a_2b_1, a_2b_2, a_2b_3, a_2b_4, a_3b_1, a_3b_2, a_3b_3, a_3b_4, a_4b_1, a_4b_2, a_4b_3, a_4b_4\}$
 b) The probability that both students ask the same prof is $P(A)$ where $A = \bigcup_{i=1}^4 a_i b_i$. $P(A) = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$.
- 2.2 a) $S = \{HHH, HHT, HTH, THH, HTT, TTH, THT, TTT\}$
 b) $P(TT) = P(\{HTT, TTH\}) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$
- 2.3 $S = \{(1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 3), (2, 4), (2, 5), (3, 1), (3, 2), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4)\}$ then the probability of finding numbers that differ by 1 is $P(\{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3), (4, 5), (5, 4)\}) = \frac{4}{10}$.
- 2.4 a) AB is for A's letter in B's envelope.
 $S = \{(WW, XX, YY, ZZ), (WW, XX, YZ, ZY), (WW, XZ, YY, ZX), (WZ, XX, YY, ZW), (WW, XY, YX, ZZ), (WY, XX, YW, ZZ), (WX, XW, YY, ZZ), (WW, XY, YZ, ZX), (WW, XZ, YX, ZY), (WY, XX, YZ, ZW), (WZ, XX, YW, ZY), (WX, XY, YY, ZZ), (WY, XZ, YY, ZW), (WX, XY, YW, ZZ), (WY, XW, YX, ZZ), (WY, XZ, YW, ZX), (WY, XW, YZ, ZX), (WY, XZ, YX, ZW), (WX, XY, YZ, ZW), (WX, XZ, YW, ZY), (WX, XW, YZ, ZY), (WZ, XW, YX, ZY), (WZ, XY, YW, ZX), (WZ, XY, YX, ZW)\}$
 c) $\frac{1}{4}, \frac{3}{8}, \frac{1}{4}, 0$
- 2.5 a)
 c)
- 2.6 $P(\text{test positive has disease}) = \frac{18}{1000}$, $P(\text{has the disease}) = \frac{2}{100}$

TODO : Finish these problems

3 Probability - Counting Techniques

Definition 3.1. A **uniform distribution** over the set $\{a_1, \dots, a_n\}$ is when the probability of each event is $\frac{1}{n}$. That is, each event has equal probability. If a compound event A contains r points, then $P(A) = \frac{r}{n}$.

3.1 Counting Arguments

There are two useful rules for counting.

1. The **Addition Rule**: Suppose we can do job 1 in p ways and job 2 in q ways. Then we can do either job 1 **OR** job 2, but not both, in $p + q$ ways.
2. The **Multiplication Rule**: Suppose we can do job 1 in p ways and, for each of these ways, we can do job 2 in q ways. Then we can do both job 1 **AND** job 2 in $p \times q$ ways.

This interpretation using AND and OR occur throughout probability. For example, if we are picking two numbers from $\{1, 2, 3, 4, 5\}$ with uniform distribution. Then the probability of having at least one even can be found like so: we need the first to be even AND the second odd OR the first odd AND the second even. Then,

$$P(\text{one number is even}) = \frac{2 \times 3 + 3 \times 2}{5 \times 5} = \frac{12}{25}$$

Note that the phrases **at random** and **uniformly** are often used to mean that all of the points in a sample space are equally likely.

Example 3.1. Suppose a course has 4 sections with no limit on how many can enrol in each section. Three students each pick a section at random. The probability that they are all in the same section is $\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) \times 4 = \frac{1}{16}$. The probability that they all end up in different sections is $\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) \times 12 = \frac{3}{16}$. The probability that nobody selects section 1 is $\left(\frac{3}{4}\right)\left(\frac{3}{4}\right)\left(\frac{3}{4}\right) \times = \frac{27}{64}$

Remark 3.1. In many problems the sample space is a set of arrangements or sequences. These are called **permutations**. A key step in the argument is to be sure to understand what it is that you are counting. It is helpful to invent a notation for the outcomes in the sample space and the events of interest.

Example 3.2. Suppose the first six letters are arranged at random to form a six-letter word (an arrangement) - we must use each letter once only. The sample space

$$S = \{abcdef, abcdef, \dots, fedcba\}$$

has a large number of outcomes and, because we formed the word "at random", we assign the same probability to each. To count the number of words in S , count the number of ways that we can construct such a word, each way corresponds to a unique word. For each word we follow a pattern of having six choices, then five, and so until until we're left with one. Thus there are $6 \times 5 \times 4 \times 3 \times 2 \times 1$ ways of constructing words; 720 ways.

Now consider events such as A ; the second letter is e or f . So there are now $5 \times 2 \times 4 \times 3 \times 2 \times 1 = 240$ ways. So,

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} = \frac{240}{720} = \frac{1}{3}$$

Generally, there are $n!$ arrangements of length n using each symbol once and only once, and $(n-r+1)!$ arrangements of length r using each symbol at most once. This product is denoted $n^{(r)}$ (" n to r factors").

There is an approximation to $n!$ called Stirling's formula which states

$$n! \rightarrow n^n e^{-n} \sqrt{2\pi n} \text{ as } n \rightarrow \infty$$

Example 3.3. A pin number of length 4 is formed by randomly selecting with replacement 4 digits from the set $\{0, 1, 2, \dots, 9\}$.

A: The probability that the pin number is even is the probability that the last digit is even.

$$P(A) = \frac{5 \times 10^3}{10^4} = \frac{1}{2}$$

B: The probability that the pin only has even digits is

$$P(B) = \frac{5^4}{10^4}$$

C: That all digits are unique is

$$P(C) = \frac{10^{(4)}}{10^4} = \frac{10 \times 9 \times 8 \times 7}{10^4} = \frac{63}{125}$$

D: That the pin number contains at least one 1 is

$$P(D) = 1 - \frac{9^4}{10^4} = \frac{3439}{10000}$$

Definition 3.2 (complement). For some general event A , the **complement** of A denoted by \bar{A} is the set of all outcomes in S which are not in A . It is often easier to count outcomes in the complement rather than in the event itself.

Example 3.4. Suppose now that we're selecting pins again, but this time without replacement.

A: The probability that the pin number is even is the probability that the last digit is even.

$$P(A) = \frac{5 \times 9 \times 8 \times 7}{10^{(4)}} = \frac{1}{2}$$

B: The probability that the pin only has even digits is

$$P(B) = \frac{5 \times 4 \times 3 \times 2}{10^{(4)}} = \frac{1}{42}$$

C: That the pin number contains 1 is

$$P(D) = 1 - \frac{9^{(4)}}{10^{(4)}} = \frac{2}{5}$$

In some problems, the outcomes in the sample space are subsets of a fixed sized. Here we look at counting such subsets. Suppose we want to randomly select a subset of 3 digits from the set of digits 0 - 9. Let us now however consider order to be irrelevant, and numbers cannot be picked twice. Then we end up with

$$m = \frac{10^{(3)}}{3!} = 120$$

arrangements of digits.

Definition 3.3. We use the combinatorial symbol $\binom{n}{r}$ to denote the number of subsets of size r that can be selected from a set of n objects. If we denote m to be the number of subsets of size r that can be selected from n things, then $m \times r! = n^{(r)}$ and so

$$\binom{n}{r} = \frac{n^{(r)}}{r!}$$

Example 3.5. Suppose a box contains 10 balls of which 3 are red, 4 are white and 3 are green. A sample of 4 balls is selected at random without replacement. Find the probability of the events

E: the sample contains 2 red balls

F: the sample contains 2 red, 1 white, and 1 green ball

G: the sample contains 2 or more red balls

For simplicity, define 0,1,2 being red, 3,4,5,6 being white, and 7,8,9 being green. Then we construct a uniform probability model

$$S = \{\{0, 1, 2, 3\}, \dots, \{6, 7, 8, 9\}\}$$

Now for E , we take the number of ways we can take 2 red balls of the three AND take 2 other-coloured balls from the other seven

$$P(E) = \frac{\binom{3}{2} \binom{7}{2}}{\binom{10}{4}} = \frac{3}{10}$$

For F it is the same idea, we multiply the number of ways of finding each ball. For G we take the sum of the ways of finding 2 red balls and 3 red balls.

$$P(G) = \frac{\binom{3}{2}\binom{7}{2}}{\binom{10}{4}} + \frac{\binom{3}{3}\binom{7}{1}}{\binom{10}{4}}$$

Here are some important properties of the combinatorial function n choose r :

1. $n^{(r)} = \frac{n!}{(n-r)!} = n(n-1)^{(r-1)}$ for $r \geq 1$
2. $\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n^{(r)}}{r!}$
3. $\binom{n}{r} = \binom{n}{n-r}$ for all $r = 0, 1, \dots, n$.
4. If we define $0! = 1$, then the formulas above make sense for $\binom{n}{0} = \binom{n}{n} = 1$
5. $(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \dots + \binom{n}{r}x^n$ (binomial theorem)

Example 3.6. Suppose the letters of the word STATISTICS are arranged at random. Find the probability of the event G that the arrangement begins with and ends with S . The sample space is $S = \{SSSTTTIIAC, SSSTTTIICA, \dots\}$. First consider the length of this set (the number of ways of spelling STATISTICS)

$$\binom{10}{3}\binom{7}{3}\binom{4}{2}\binom{2}{1}\binom{1}{1} = n$$

Now the number of ways of spelling it with two S's fixed at either end

$$\frac{\binom{8}{3}\binom{5}{2}\binom{3}{1}\binom{2}{1}\binom{1}{1}}{\binom{10}{3}} = k$$

The quotient $\frac{k}{n} = \frac{1}{15}$ is our answer to the probability that the words start and end with S.

Remark 3.2. The number of arrangements when some symbols are alike generally, for n symbols of type $i \in \{1, \dots, k\}$ with $\sum_{i=1}^k n_i = n$, then the number of arrangements using all of the symbols is

$$\binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \binom{n-n_1-n_2}{n_3} \times \dots \times \binom{n_k}{n_k} = \frac{n!}{n_1!n_2!\dots n_k!}$$

Example 3.7. Suppose we make a random arrangement of length 3 using letters from the set $\{1, b, c, d, e, f, g, h, i, j\}$, then what is the probability of the event B that the letters are in alphabetic order if (a) the letters are selected without replacement, and (b) if the letters are selected with replacement.

There are $10^{\binom{3}{3}}$ ways of picking 3 letter words from this set, and $\binom{10}{3}$ ways of picking them in alphabetical order, thus for part (a) we simply take the quotient of the two.

$$P(B) = \frac{\binom{10}{3}}{10^{\binom{3}{3}}}$$

With replacement implies three cases, one for each number of identical letters in a sequence of 3 letters. In case 1, there are 3 identical letters and there are obviously only 10 ways of achieving this, for case 2 there are 2 identical letters and so there are 6 possibilities of arrangements, but only two are identical, so we multiply the choices of two letters by two. The third case is the same as part (a). Thus

$$P(B) = \frac{10 + \binom{10}{2} + \binom{10}{3}}{10^3} = \frac{11}{50}$$

Note. While $n^{(r)}$ only has a physical interpretation when n and r are positive integers with $n \geq r$, it still has meaning when n is not a positive integer, as long as r is a non-negative integer. In general, we can define $n^{(r)} = n(n-1)\cdots(n-r+1)$. For example:

$$(-2)^{(3)} = (-2)(-2-1)(-2-2) = -24$$

$$1.3^{(2)} = (1.3)(1.3-1) = 0.39$$

Note that in order for $\binom{n}{0} = \binom{n}{n} = 1$ we must define

$$n^{(0)} = \frac{n!}{(n-0)!} = 1 \quad \text{and} \quad 0! = 1.$$

Also $\binom{n}{r}$ loses its physical meaning when n is not a non-negative $\geq r$ but we can use

$$\binom{n}{r} = \frac{n^{(r)}}{r!}$$

to define it when n is not a positive integer but r is. For example,

$$\binom{\frac{1}{2}}{3} = \frac{(\frac{1}{2})^{(3)}}{3!} = \frac{1}{16}$$

Also, when n and r are non-negative integers and $r > n$ notice that $\binom{n}{r} = \frac{n^{(r)}}{r!} = \frac{n(n-1)\cdots(0)\cdots}{r!} = 0$

3.2 Review of Useful Series and Sums

Recall the following series and sums.

1. Geometric Series

$$a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a(1-r^n)}{1-r} \quad \text{for } r \neq 1$$

If $|r| < 1$, then

$$a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a}{1-r}$$

2. Binomial Theorem

$$(1+a)^n = 1 + \binom{n}{1}a + \binom{n}{2}a^2 + \cdots + \binom{n}{n}a^n = \sum_{x=0}^n \binom{n}{x}a^x$$

3. Binomial Theorem (general)

$$(1+a)^n = \sum_{x=0}^{\infty} \binom{n}{x}a^x \quad \text{if } |a| < 1$$

Proof. Recall from Calculus the Maclaurin's series which says that a sufficiently smooth function $f(x)$ can be written as an infinite series using an expansion around $x = 0$,

$$f(x) = f(0) + \frac{f'(0)}{1}x + \frac{f''(0)}{2!}x^2 + \cdots$$

provided that this series is convergent. In this case, with $f(a) = (1 + a)^n$, $f'(0) = n$, $f''(0) = n(n - 1)$ and $f^{(r)}(0) = n^{(r)}$. Substituting,

$$f(a) = 1 + \frac{n}{1}a + \frac{n(n-1)}{2!}a^2 + \cdots + \frac{n^{(r)}}{r!}a^r + \cdots = \sum_{x=0}^{\infty} \binom{n}{x} a^x$$

It is not hard to show that this converges whenever $|a| < 1$. □

4. Multinomial Theorem, a generalization of the binomial theorem is

$$(a_1 + a_2 + \cdots + a_k)^n = \sum \frac{n!}{x_1! x_2! \cdots x_k!} a_1^{x_1} a_2^{x_2} \cdots a_k^{x_k}.$$

with the summation over all x_1, x_2, \dots, x_k with $\sum x_i = n$.

5. Hypergeometric Identity

$$\sum_{x=0}^{\infty} \binom{a}{x} \binom{b}{n-x} = \binom{a+b}{n}.$$

There will not be an infinite number of terms if a and b are positive integers since the terms become 0 eventually. For example

$$\binom{4}{5} = \frac{4^{(5)}}{5!} = 0$$

6. Exponential Series

$$e^x = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

7. Special series involving integers:

$$\begin{aligned} \sum_{x=0}^n x &= \frac{n(n+1)}{2} \\ \sum_{x=0}^n x^2 &= \frac{n(n+1)(2n+1)}{6} \\ \sum_{x=0}^n x^3 &= \left[\frac{n(n+1)}{2} \right]^2 \end{aligned}$$

3.3 Tutorial 1

Example 3.8. Suppose one's birthday is equally likely to be in one of the 12 months. For a group of 10 students, find the probability that

A = "None of them has a birthday in January or December"

B = "All of them have birthdays in different months"

C = "All birthdays in the same month"

D = "2 of them in one month, 3 in another, and the remaining 5 in different months"

Consider a box model where we have 12 slots and 10 objects to fill the the slots with. So,

$$P(A) = \frac{10^{10}}{12^{10}}$$

$$P(B) = \frac{12^{(10)}}{12^{10}}$$

$$P(C) = \frac{12}{12^{10}}$$

$$P(D) = \frac{\binom{12}{1}\binom{10}{2}\binom{11}{1}\binom{8}{3}10^{(5)}}{12^{10}}$$

Consider now a new case $E =$ "5 in one month and remaining 5 in another month."

$$P(E) = \frac{\binom{12}{1}\binom{10}{5}\binom{11}{1}\binom{5}{5}}{12^{10}} \text{ not correct, we need to divide by 2}$$

We take a different approach, reserving two months to begin with, so

$$P(E) = \frac{\binom{12}{2}\binom{10}{5}\binom{5}{5}}{12^{10}}$$

Example 3.9. Four numbers are randomly selected from $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ without replacement and form a sequence.

A = "The sequence is a 4-digit number"

B = "It is a 4-digit even number"

C = "It is a 4-digit even number bigger than 4000"

Our solution,

$$P(A) = \frac{\binom{9}{1}9^{(3)}}{10^{(4)}} = \frac{9}{10}$$

$$P(B) = \frac{\binom{5}{1}\binom{5}{1}8^{(2)} + \binom{4}{1}\binom{4}{1}8^{(2)}}{10^{(4)}}$$

$$P(C) = \frac{\binom{3}{1}\binom{5}{1}8^{(2)} + \binom{3}{1}\binom{4}{1}8^{(2)}}{10^{(4)}}$$

Example 3.10. Roll a regular die 4 times. $A =$ "The sum of the 4 numbers appeared is 10"

$$S = \{(1, 1, 1, 1), (1, 1, 1, 2), \dots, (6, 6, 6, 6)\}$$

and $n = 6^4$. So,

$$A = \{(2, 3, 3, 2), (1, 4, 1, 4), \dots\}$$

The number of ways to have 4 numbers add up to 10, each number is from $\{1, 2, 3, 4, 5, 6\}$ is the same as the number of ways to put 10 identical balls into 4 boxes, and each box has at least one ball, at most 6 balls.

$$P(A) = \frac{\binom{9}{3} - 4}{6^4}$$

Another way to solve this problem is to consider $(x + x^2 + x^3 + x^4 + x^5 + x^6)^4$, we need the coefficient in front of x^10 in the expansion. We can find this using binomial theorem and it describes the same thing. So,

$$x^4(1 + x + x^2 + x^3 + x^4 + x^5)^4 = x^4 \left(\frac{1 - x^6}{1 - x} \right)^4 = x^4(1 - x^6)^4(1 - x)^{-4}$$

3.4 Solutions to Problems on Chapter 3

3.1 (a) $P(a) = \frac{\binom{4}{1}6^{(5)}}{7^{(4)}} = \frac{4}{7}$

(b) $P(b) = \frac{5 \times 5^{(5)}}{7^{(6)}}$

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

3.2(a) i. $P = \frac{(n-1)^r}{n^r}$

(a) ii. $P = \frac{n^{(r)}}{n^r}$

3.3 (a) $P(a) = \frac{6^{(4)}}{6^4} = \frac{5}{18}$

(b) $P(b) = \frac{6 \times \binom{6}{2}}{6^4} = \frac{5}{72}$ (6 destinations for 6 spots split among groups of 2)

3.4 $P(H) = \frac{\binom{4}{2} \binom{12}{4} \binom{36}{7}}{\binom{52}{13}}$ (36 since 16 cards not applicable as "others")

3.5 (a) $P(\text{STATISTICS}) = \frac{1}{\frac{10!}{3!3!2!1!1!1!1!}} = \frac{1}{50400}$

(b) $P(\text{"same letter at each end"}) = \frac{8 \binom{7}{3} \binom{4}{2} \binom{2}{1} + 8 \binom{7}{3} \binom{4}{2} \binom{2}{1} + \binom{8}{3} \binom{5}{3} \binom{2}{1} \binom{1}{1}}{50400} = \frac{7}{45}$

3.6 (a) $P(a) = \frac{1}{6}$ (take any 3 digits, the probability there are $3^{(3)}$ arrangements, and only one is strictly increasing)

(b) $P(b) = P(a) \times \frac{10^{(3)}}{10^3}$

3.7 $P(A) = \frac{365^{(r)}}{365^r}$

3.8 (a) $P(a) = \frac{1}{n}$ (not in order, just pick a single key out of n)

(b) $P(b) = \frac{2}{n}$

3.9 $P(A) = \frac{1+3+5+\dots+(2n-1)}{\binom{2n+1}{3}}$

3.10 (a) i. $\frac{6}{10000}$ (6 ways to arrange this group, refer to Remark 3.2)

ii. $\frac{4^{(4)}}{10000} = \frac{24}{10000}$, better odds to pick four non-repeating digits

(b) The operator loses money if the prize money exceeds $\$1 \times 10000$, that is, if there are more than 20 prize winners. We need the probability that four digits are picked with no repeating digits. $P(b) = \frac{10^{(4)}}{10000}$. The amount of money the operator gives out is $\$500 \times |S|$ where S is the set of all possible rearrangements of the four digits drawn.

3.11 (a) $P(a) = \frac{\binom{6}{2} \binom{19}{4}}{25}$ (note to self: remember you have 19 left after taking 6 away, not 23 kiddo)

(b) We could estimate there were 15 deer (extrapolate experiment three times to meet 6 tagged deer)

$$3.12 \quad (a) \quad P(a) = \frac{1}{\binom{49}{6}}$$

$$(b) \quad P(b) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}}$$

$$(c) \quad P(c) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}}$$

$$(d) \quad P(d) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}}$$

$$3.13 \quad (a) \quad \text{There are 2 Jacks, left, so } P(a) = 1 - \frac{\binom{48}{3}}{\binom{50}{3}}$$

$$(b) \quad 1 - \frac{\binom{45}{2}}{\binom{47}{2}}$$

$$(c) \quad \frac{\binom{2}{2}\binom{48}{3}}{\binom{50}{5}} = \frac{\binom{48}{3}}{\binom{50}{5}}$$

π The binomial theorem states that

$$(1 + a)^n = \sum_{x=0}^n \binom{n}{x} a^x$$

So,

$$n(1 + a)^{n-1} = \sum_{x=0}^n x \binom{n}{x} a^{x-1} \quad (\text{Differentiate})$$

$$na(1 + a)^{n-1} = \sum_{x=0}^n x \binom{n}{x} a^x$$

$$n \left(\frac{p}{1-p} \right) \left(1 + \left(\frac{p}{1-p} \right) \right)^{n-1} = \sum_{x=0}^n x \binom{n}{x} \left(\frac{p}{1-p} \right)^x \quad \left(a = \left(\frac{p}{1-p} \right) \right)$$

$$np = \sum_{x=0}^n x \binom{n}{x} p^x p^{n-x}$$

4 Probability Rules and Conditional Probability

4.1 General Methods

Recall that a [probability model](#) consists of a [Sample Space](#) S , a set of events or subsets of the sample space to which we can assign probabilities and a mechanism for assigning these probabilities. So the probability for some event A is the sum of the probabilities of simple events in A . Thus we have these rules:

- **Rule 1** $P(S) = 1$
- **Rule 2** For any event A , $0 \leq P(A) \leq 1$.
- **Rule 3** If A and B are two events with $A \subseteq B$, $P(A) \leq P(B)$.

Now consider since I am not going to make venn diagrams right now a venn diagram with two circles A and B inside a rectangle S . Then, S is the sample space and A and B are events. $A \cup B$ is everything in either circle, $A \cap B$ is the intersection of both circles. As well, \bar{A} is everything outside of A and is defined as the [complement](#).

Theorem 4.1 (De Morgan's Laws).

- $\overline{A \cup B} = \bar{A} \cap \bar{B}$
- $\overline{A \cap B} = \bar{A} \cup \bar{B}$

4.2 Rules for Unions of Events

In addition to the last two rules, we have the following

- **Rule 4a** (probability of unions) $P(A \cup B) = P(A) + P(B) - P(AB)$
- **Rule 4b** (probability of union of three events)

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

- **Rule 4c** (inclusion-exclusion principle)

$$P(A_1 \cup \dots \cup A_n) = \sum_i P(A_i) - \sum_{i < j} P(A_i A_j) + \sum_{i < j < k} P(A_i A_j A_k) - \sum_{i < j < k < l} P(A_i A_j A_k A_l) + \dots$$

Definition 4.1 (mutually exclusive). Events A and B are **mutually exclusive** if $AB = \varphi$ (the empty set).

Since **mutually exclusive** events A and B have no common points, $P(AB) = P(\varphi) = 0$.

- **Rule 5a** (unions of mutually exclusive events) Let A and B be **mutually exclusive** events. Then $P(A \cup B) = P(A) + P(B)$.
- **Rule 5b** (generalized 5a) If A_1, \dots, A_n are **mutually exclusive** events, then $P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$.
- **Rule 6** (probability of complements) $P(A) = 1 - P(\bar{A})$.

4.3 Intersections of Events and Independence

Definition 4.2 (independent). Events A_1, A_2, \dots, A_n are **independent** if and only if $P(A_{i_1}, A_{i_2}, \dots, A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$ for all sets (i_1, i_2, \dots, i_k) of distinct subscripts chosen from $(1, 2, \dots, n)$. If they are not independent, we call the events **dependent**.

Example 4.1. Suppose you're tossing a dice, are the events that your first toss is a 3 and the sum of two tosses is 7 **independent**? $P(A) = 1/6$, $P(B) = 1/6$, and $P(AB) = 1/36$, so yes.

Theorem 4.2. If A and B are independent events, then so too are \bar{A} and \bar{B} .

4.4 Conditional Probability

In some situations we'll want to know what the probability of A is while knowing that B has already occurred. We let the symbol $P(A|B)$ represent this probability.

Definition 4.3 (conditional probability).

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

if $P(B) \neq 0$.

Example 4.2. If a fair coin is tossed 3 times, find the probability that if at least 1 head occurs, then exactly 1 head occurs.

Let A be the event that we obtain 1 head, and B be the event that at least one head occurs. We want $P(A|B)$. So

$$\begin{aligned} P(A|B) &= \frac{P(AB)}{P(B)} \\ &= \frac{\frac{3}{8}}{\frac{7}{8}} \\ &= \frac{3}{7} \end{aligned}$$

4.5 Multiplication and Partition Rules

- **Product Rule** Let A, B, C, D, \dots be arbitrary events in a sample space. Assume that $P(A) > 0, P(AB) > 0$, and $P(ABC) > 0$. Then

$$\begin{aligned} P(AB) &= P(A)P(B|A) \\ P(ABC) &= P(A)P(B|A)P(C|AB) \\ P(ABCD) &= P(A)P(B|A)P(C|AB)P(D|ABC) \end{aligned}$$

and so on.

- **Partition Rule** Let A_1, \dots, A_k be a partition of the sample space S into disjoint (mutually exclusive) events, that is

$$A_1 \cup A_2 \cup \dots \cup A_k = S \quad \text{and} \quad A_i \cap A_j = \varphi \text{ if } i \neq j$$

Let B be an arbitrary event in S . Then

$$\begin{aligned} P(B) &= P(BA_1) + P(BA_2) + \dots + P(BA_k) \\ &= \sum_{i=1}^k P(B|A_i)P(A_i) \end{aligned}$$

Example 4.3. In an insurance portfolio 10% of the policy holders are in Class A_1 (high risk), 40% are in Class A_2 (medium risk), and 50% are in Class A_3 (low risk). The probability there is a claim on a Class A_1 policy in a given year is .10; similar probabilities for Classes A_2 and A_3 are .05 and .02. Find the probability that if a claim is made, it is made on a Class A_1 policy.

Solution.

For a randomly selected policy, let B be the event that a policy has a claim, and A_i be the event that the policy is of class i for $i = 1, 2, 3$. We are asked to find $P(A_1|B)$. Note that,

$$P(A_1|B) = \frac{P(A_1B)}{P(B)}$$

and that

$$P(B) = P(A_1B) + P(A_2B) + P(A_3B).$$

We are told that

$$P(A_1) = 0.10, P(A_2) = 0.40, P(A_3) = 0.50$$

and that

$$P(B|A_1) = 0.10, P(B|A_2) = 0.05, P(B|A_3) = 0.02.$$

Thus,

$$P(A_1B) = P(A_1)P(B|A_1) = 0.01$$

$$P(A_2B) = P(A_2)P(B|A_2) = 0.02$$

$$P(A_3B) = P(A_3)P(B|A_3) = 0.01$$

Therefore $P(B) = 0.04$ and $P(A_1|B) = 0.01/0.04 = 0.25$.

Theorem 4.3 (Bayes Theorem).

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

4.6 Tutorial 2

Example 4.4. Independent games are to be played between teams X and Y , with team X having probability 0.6 to win each game.

(i) A = "Team X wins a best-of-seven series with a sweep" (4:0)

(ii) B = "Team X wins a best of seven series."

Solution. Let B_i = "Team X wins game i " ($1 \leq i \leq 7$). For any game $P(B_i) = 0.6$. Note that B_1, B_2, \dots, B_7 are independent.

(i) $P(B_1B_2B_3B_4) = P(B_1)P(B_2)P(B_3)P(B_4) = 0.6^4$

(ii) D_i = "Team X wins the series in i games" for $i = 4, 5, 6, 7$.

$P(B) = P(D_4 \cup D_5 \cup D_6 \cup D_7) = P(D_4) + P(D_5) + P(D_6) + P(D_7)$. Now,

- $P(D_4) = 0.6^4$

-

$$\begin{aligned} P(D_5) &= P(B_1B_2B_3\bar{B}_4B_5 \cup B_1B_2\bar{B}_3B_4B_5 \cup B_1\bar{B}_2B_3B_4B_5 \cup \bar{B}_1B_2B_3B_4B_5) \\ &= P(B_1B_2B_3\bar{B}_4B_5) + \dots + P(\bar{B}_1B_2B_3B_4B_5) \\ &= 4 \times 0.4 \times 0.6^4 \end{aligned}$$

$$\text{Then } P(B) = 0.6^4 + \binom{4}{1}0.4 \times 0.6^4 + \binom{5}{2}0.4^2 \times 0.6^4 + \binom{6}{3}0.4^3 \times 0.6^4$$

Example 4.5. There are n letters L_1, \dots, L_n and n envelopes E_1, \dots, E_n and L_i and E_i are intended to be matched randomly, put letters into envelopes, one letter per envelope. Find the probability that

- A = "At least one match"
- B = "No matches"
- C = "The 1st r letters and the r envelopes matche dbut none of the others match"

- $D =$ "Exactly r matches"

Solution. $A_i =$ " L_i matches E_i " for $i = 1, 2, \dots, n$. Are A_1, \dots, A_n independent? No. Are A_i and A_j disjoint? No.

$$\begin{aligned}
 P(A) &= P(A_1 \cup A_2 \cup \dots \cup A_n) \\
 &= \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n) \\
 &= nP(A_1) - \binom{n}{2} P(A_1 A_2) + \binom{n}{3} P(A_1 A_2 A_3) + \dots + (-1)^{n-1} P(A_1 \dots A_n) \\
 &= n \frac{1}{n} - \binom{n}{2} \frac{1}{n(n-1)} + \binom{n}{3} \frac{1}{n(n-1)(n-2)} + \dots + (-1)^{n-1} \frac{1}{n!} \\
 &= 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots + (-1)^{n-1} \frac{1}{n!} \\
 &= \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!}
 \end{aligned}$$

End of Chapter 4:

- 4.12 An experiment has three possible outcomes A, B, C with respective probabilities p, q, r , where $p + q + r = 1$. The experiment is repeated until either outcome A or outcome B occurs. Show that A occurs before B with probability $\frac{p}{p+q}$.

Solution. On a single trial: A, B, C . $P(A) = p$, $P(B) = q$, $P(c) = r$, and $p + q + r = 1$. Under repeated trials, $A_i =$ "The i -th trial has result A ", similarly for B_i and C_i . Note that $P(A_i) = p$ still, and similarly for the other two since these are repeated trials. Note that the results from different trials are independent of each other. For example, A_1 and B_2 are independent, so is C_1 and A_3 .

Now, the probability that A occurs before B includes trials $A_1 \cup C_1 A_2 \cup C_1 C_2 A_3 \cup \dots$,

$$\begin{aligned}
 P(A \text{ occurs before } B) &= P(A_1 \cup C_1 A_2 \cup C_1 C_2 A_3 \cup \dots) \\
 &= P(A_1) + P(C_1 A_2) + P(C_1 C_2 A_3) + \dots \\
 &= p + rp + r^2 p + \dots \\
 &= p(1 + r + r^2 + r^3 + \dots) \\
 &= \frac{p}{1-r} \\
 &= \frac{p}{p+q}
 \end{aligned}$$

k possible outcomes: D_1, D_2, \dots, D_k . Find $p_i = P(D_i)$ for $i = 1, 2, \dots, k$ and $\sum_{i=1}^k p_i = 1$. Find $P(D_1 \text{ occurs before } D_2) = \frac{p_1}{p_1+p_2}$

- 4.13 $A =$ "you win the game", and $B_i =$ "the sum of the two numbers from rolling two dice" $i = 2, 3, 4, \dots, 11, 12$. We want $P(A) = \sum_{i=2}^{12} P(B_i)P(A|B_i)$. How do we find B_i ? Well consider $P(B_2) = \frac{1}{36}$, $P(B_3) = \frac{2}{36}$. Now the tricky part is the conditional probability. For example, $P(A|B_7) = 1$. $P(A|B_2) = 0$, and $P(A|B_5) = \frac{P(B_5)}{P(B_5)+P(B_7)}$

4.7 Problems on Chapter 4

4.1 Use [De Morgan's Laws](#) for the second last one.

4.2 $P(A) = \frac{10}{10^3}, P(B) = \frac{10^{(3)}}{10^3}, P(C) = \frac{9^3}{10^3}, P(D) = \frac{5^3}{10^3}, P(E) = \frac{2 \times 5^3}{10^3}.$

4.3 We want $P(A|\bar{B})$ which is $\frac{P(A\bar{B})}{P(\bar{B})} = \frac{P(A) - P(AB)}{1 - P(B)} = \frac{P(A) - P(B)P(A|B)}{1 - P(B)}$

- 4.4 (a) It's the probability of not getting a 1 to the 8th power. $(1 - P(1))^8$
 (b) It's the probability of not getting a 2 to the 8th power. $(1 - P(2))^8$
 (c) Same idea, but for both. $(1 - P(1) - P(2))^8$
 (d) $1 - P((a)) - P((b)) + P((c))$

4.5 $P(A \cup B) = P(A) + P(B) - P(AB) = P(A) + P(B) - P(A)P(B)$ (independent)

4.6 Let A, B, C be the events that A, B, C get the right answer. Then we want to find $P(\bar{C}|X)$ where X is the event that two people get the right answer. The probability that two people get the right answer is $P(\bar{C}|X) = \frac{P(\bar{C}X)}{P(X)} = \frac{P(ABC)}{P(X)} = \frac{P(A)P(B)P(\bar{C})}{P(A)P(B)P(\bar{C}) + P(A)P(B)P(C) + P(A)P(\bar{B})P(C)}$

4.7 Let X be the event that a customer pays with credit card then

- (a) $\binom{5}{3} 0.7^3 0.3^2$
 (b) $\binom{4}{2} 0.7^3 0.3^2$

4.9 Let D be the probability that one randomly selected person out of this population is diseased. Then $P(D)$ is the union of the probabilities at given A , A is diseased, and given B , B is diseased, and given C , C is diseased which is 0.041. Then the probability that at least one of ten are diseased is one minus the probability ten people selected have no disease. Thus $P(\text{"at least one diseased of 10"}) = 1 - (1 - 0.041)^{10}$.

4.10 The probability that A sweeps is $P(A_H A_H A_A A_A) = 0.7^2 0.5^2$, and in five is $P(B_A A_H A_A A_A A_A \cup A_H B_A A_A A_A A_A \cup A_H A_H B_H A_A A_A \cup A_H A_H A_A B_H A_A) = 0.175$, probability it doesn't go to six games is the sum of these two plus the same idea for B winning in four or five.

4.12 look up

4.13 $A = \text{"you win the game"}$, and $B_i = \text{"the sum of the two numbers from rolling two dice" } i = 2, 3, 4, \dots, 11, 12$. We want $P(A) = \sum_{i=2}^{12} P(B_i)P(A|B_i)$. How do we find B_i ? Well consider $P(B_2) = \frac{1}{36}$, $P(B_3) = \frac{2}{36}$. Now the tricky part is the conditional probability. For example, $P(A|B_7) = 1$, $P(A|B_2) = 0$, and $P(A|B_5) = \frac{P(B_5)}{P(B_5) + P(B_7)}$. So,

$$P(A) = 0 + 0 + \frac{P(B_4)^2}{P(B_4) + P(B_7)} + \frac{P(B_5)^2}{P(B_5) + P(B_7)} + \frac{P(B_6)^2}{P(B_6) + P(B_7)} + P(B_7) + \frac{P(B_8)^2}{P(B_8) + P(B_7)} + \frac{P(B_9)^2}{P(B_9) + P(B_7)} + \frac{P(B_{10})^2}{P(B_{10}) + P(B_7)} + P(B_{11}) + 0 = 0.492929292$$

4.14 (a) The probability that a student says yes is the 0.2 times the chance that he or she was born in one of the two months added to 0.8 times the chance they have ever cheated on an exam p . Thus

$$P((a)) = \frac{1}{30} + \frac{4p}{5}$$

(b) Suppose x of n students answer yes, then set $P((a)) = \frac{x}{n}$. So,

$$p = \frac{5x - \frac{1}{6}n}{4n}$$

(c) The proportion of students that answer yes responding to question B is

$$\frac{\frac{4p}{5}}{P((a))} = \frac{4p}{\frac{1}{6} + 4p} = \frac{24p}{1 + 24p}$$

4.16 (a) $\frac{1}{5} \frac{3}{5} \frac{1}{5} = \frac{3}{125}$

(b) $\frac{1}{10} \frac{1}{10} \frac{8}{10}$ is the minimum.

4.17 (a) $P(A|Spam) = \frac{P(Spam)P(A|Spam)}{P(A)} = \frac{0.5 \cdot 0.2}{0.5(0.2+0.001)}$

(b) $P(A|NotSpam)$

4.18 (a) $1 - \frac{P(A_1|NotSpam) \cdots (P(A_n|NotSpam))}{P(A_1|Spam) \cdots P(A_n|Spam)}$

(b) Same as above, use $1 - P(A_3)$.

(c)

Note. Sorry to anyone using these notes that I haven't created a better summary of 4 or anything for 5, I'll try to get back on track after the midterm.

Note. Good questions to ask:

- Can I use the negation? [De Morgan's Laws](#), remember to take 1 - "neg".
- Is there a product I don't know? Use product rule. ($P(AB) = P(A)P(B|A)$)
- Probability of a union? Take sum of parts, less their various intersections.
- Mutually exclusive? Intersection is 0.
- Independent? Probability of product is product of probabilities of individual parts of initial product
- Conditional probability? Turn into fraction of product over probability of given part.
- Do I know probabilities of various condition parts but not main part? Partition rule; $P(B) = P(B|A_1) + \cdots + P(B|A_n)$

5 Probability Distributions

5.1 Solutions to Problems on Chapter 5

Quick tip, when the feature of independence is mentioned in the question, we're probably talking about either Bernoulli trials (Binomial, Negative Binomial), or Poisson process. Hypergeometric is NOT independent.

5.1

5.2

5.3

5.4 Let $P(X = x) = p(1 - p)^x$, then the probability function on R is a function on $\frac{X}{4}$,

$$\begin{aligned} P(R = r) &= P(X = r) + P(X = (r + 4)) + P(X = (r + 8)) + \dots \\ &= p(1 - p)^r + p(1 - p)^{r+4} + p(1 - p)^{r+8} + \dots \\ &= p(1 - p)^r(1 + p(1 - p)^4 + p(1 - p)^8 + \dots) \end{aligned}$$

The sum inside the brackets is a geometric series similar to

$$z + z^4 + z^8 + \dots = \frac{z}{1 - z^4}$$

So,

$$P(R = r) = \frac{p(1 - p)^r}{1 - p(1 - p)^4}$$

5.5 (a) Event X that our friend takes 10 tickets to have 4 winners with probability of each trial is 0.3 has a negative binomial distribution. We want 6 failures before 4th success.

$$f(x) = \binom{x + 4 - 1}{x} (0.3)^4 (0.7)^x$$

then ($x = 6$ failures)

$$P(X = 6) = 0.0800483796$$

(b) We have a binomial distribution, we want $x = 0$. 15 events, probability of success trial is $1/9$.

$$f(x) = \binom{15}{x} \frac{1}{9^x} \left(\frac{8}{9}\right)^{15-x} \implies P(X = 0) = 0.170888$$

(c) Same idea, 60 cups not 15, 1 or 0 not just 0.

$$P(X \leq 1) = \binom{60}{1} \frac{1}{9} \left(\frac{8}{9}\right)^{59} + \binom{60}{0} \left(\frac{8}{9}\right)^{60} = 7.2488 \times 10^{-3}$$

5.6 (a) One minus probability this guy wins no prizes. Probability of a win is $\frac{500}{500000} = 0.001$. Binomial distribution, ten trials, 0 successes.

$$P(x \geq 1) = 1 - P(X = 0) = 1 - \binom{10}{0} (0.001)^0 (1 - 0.001)^{10} = 9.955 \times 10^{-3} \approx 0.010$$

(b) Same thing, 2000 tickets.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - 0.001)^{2000} = 0.8648$$

5.7 (a) Straightforward. $\frac{40}{150} = \frac{4}{15}$

(b) Hypergeometric distribution, men are successes, women are failures. 150 people, 74 men, 76 women, 12 people being picked.

$$P(Y = y) = \frac{\binom{74}{y} \binom{150-74}{12-y}}{\binom{150}{12}}$$

(c) Plug in numbers.

$$P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 0.176$$

5.8 This is classic Poisson Distribution from Poisson Process. We have a rate, a timeline, a period and some number of events in that period. Both units in months so we're good.

$$P(X \geq 4) = 1 - \sum_{k=0}^3 P(X = k) = 1 - e^{-\lambda t} \left(\sum_{k=0}^3 \frac{(\lambda t)^k}{k!} \right) = 0.9989497$$

5.9 Fish are more interesting than bacteria, so we use Poisson (translate: fish) process. We have a rate $\frac{1}{20}$

(a) Classic. V is volume of sample.

$$P(X = x) = \frac{(\lambda V)^x e^{-\lambda V}}{x!} \implies P(X = 2) = \frac{\left(\frac{1}{20} \cdot 10\right)^2 e^{-\left(\frac{1}{20} \cdot 10\right)}}{2!} = 0.7581633$$

(b) Exactly the same procedure, use 1 - probability of finding none, and 1 instead of 10.

$$P(X \geq 1) = 0.488$$

(c) We use binomial distribution on poisson distribution. p is probability of no fish present in sample.

$$p = \frac{(10\lambda)^0 e^{-10\lambda}}{0!} = e^{-10\lambda}$$

$$P(Y = y) = \binom{10}{y} p^y (1-p)^{10-y} = \binom{10}{y} (e^{-10\lambda})^y (1 - e^{-10\lambda})^{10-y}$$

(d)

$$p = \frac{3}{10} \implies \lambda = 0.12$$

5.10 (a) One minus probability of no claims,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\left(\frac{8}{100} \cdot 1\right)^0 e^{-\left(\frac{8}{100} \cdot 1\right)}}{0!} = 0.07688365$$

(b) Probability of no claims for 20 is binomial on top of above,

$$P(X = x) = \binom{20}{x} (0.07688365)^x (1 - 0.07688365)^{20-x}$$

Set $x = 0$ for first part of this question. For second part, calculate $P(X \geq 2)$.

5.11 80% of months have no failures.

(a) Straightforward.

$$P(X = 5) = \binom{7}{5} 0.8^5 0.2^2 = 0.2753$$

(b) Negative binomial.

$$P(X = 7) = \binom{2+5-1}{2} 0.2^2 0.8^5 = 0.1966$$

- (c) So for this part we think about it in terms of Poisson Distribution. So $P(X = x)$ is the probability that a month has x power failures,

$$P(X = 0) = 0.8$$

from the question, and $t = 1$ for one month, so

$$0.8 = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t} = e^{-\lambda(1)} \implies \lambda = 0.223$$

Then,

$$P(X > 1) = 1 - (P(X = 0) + P(X = 1)) = 1 - (e^{-0.223} + (0.223)e^{-0.223}) = 0.0215$$

• 5.12

- (a) First note that $\lim_{N \rightarrow \infty} \frac{r}{N} = \frac{r}{N}$

$$\begin{aligned} \lim_{N \rightarrow \infty} f(x) &= \lim_{N \rightarrow \infty} \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}} \\ &= \binom{n}{x} \lim_{N \rightarrow \infty} \frac{r!(N-r)!n!}{(r-x)!(N-r-n+x)!N!} \\ &= \binom{n}{x} \lim_{N \rightarrow \infty} \frac{[r(r-1) \cdots (r-x+1)][(N-r)(N-r-1) \cdots (N-r-n+x+a)]}{N(N-1) \cdots (N-n+1)} \end{aligned}$$

For $x = 0, 1, 2, \dots, n$. It must be the case that $r \rightarrow \infty$ since $p = \frac{r}{N}$ is fixed. We expect that $N-r \leq N-n$ and n could be ignored as r, N go to infinity. Let $q = \frac{1}{p} = \frac{N}{r}$. Then,

$$\begin{aligned} f(x) &= \binom{n}{x} \frac{[r(r-1) \cdots (r-x+1)][(q-1)r((q-1)r-1) \cdots ((q-1)r-n+x+1)]}{qr(qr-1) \cdots (qr-n+1)} \\ &= \binom{n}{x} \left(\frac{1}{q}\right)^x \left(\frac{q-1}{q}\right)^{n-x} \\ &= \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

- 5.13 (a) Call this event X . We have a rate λ per hectare and measure which is n hectare plots, and we're looking for at least k of some item. So it's Poisson process. We are looking for one minus the probability that all of the hectare plots have less than k spruce budworms. That probability is the n -th power of the probability that a single hectare plot has less than k spruce budworms, which is the sum from $i = 0$ to $i = k-1$ of the poisson distribution over one hectare (so $\mu = 1 \cdot \lambda$) for i . Therefore we conclude,

$$P(X) = 1 - \left(\sum_{i=0}^{k-1} \frac{(\lambda \cdot 1)^i e^{-\lambda \cdot 1}}{i!} \right)^n$$

5.14 So we're essentially told at the beginning here that $\lambda = 20$, and $p = 0.2$.

- (a) Probability of two sales in 5 calls is straightforward Binomial Distribution,

$$\binom{5}{2} 0.2^2 \cdot 0.8^3 = 0.2048$$

(b) This is negative binomial distribution, 8 fails up to reach 2 successes.

$$\binom{2+6-1}{1} 0.2^2 \cdot 0.8^6 = 0.0734$$

(c) This situation occurs only if calls 4,5,6,7 are failures. Thus our probability is (*) the ways that the first three can yield one success, then (**) four failures, then (***) one success, over all possibilities where it takes 8 trials for 2 successes (which we found in part b).

$$\frac{\underbrace{\binom{3}{1}}_{(*)} \times 0.2 \times 0.8^2 \times \underbrace{0.8^4}_{(**)} \times \underbrace{0.2}_{(***)}}{P((b))} = \frac{0.03145728}{0.0734} = 0.428$$

(d) This is Poisson, we have to convert minutes to hours, λ is per hour, so just use $\frac{1}{4}$ for 15 minutes.

$$\frac{(\frac{1}{4} \times \lambda)^3 e^{-\lambda}}{3!} = \frac{(\frac{1}{4} \times 20)^3 e^{-\frac{20}{4}}}{3!} = 0.14037$$

5.15 Collection of $N = 105$ objects, $n = 35$ of type A and $r = 70$ of type B . We choose objects without replacement until 8 of type B have been selected. Let X be the number of type A withdrawn. This is Hypergeometric distribution. Think about it this way: how many ways can you choose x objects of type A , and also 7 objects of type B ? (7 here since the last one is always of type B to get our 8th success) Well, $\binom{35}{x} \times \binom{70}{7}$. All possibilities are $\binom{105}{7+x}$ because you're taking 7 of type B for sure, and then x of type A . Now finally all we need is the probability that the last one is an 8. Well, that's simply the number of objects of type B left after our first 7 are taken, 63, over the total number less the x and 7 we already took.

$$f(x) = P(X = x) = \frac{\binom{35}{x} \binom{70}{7}}{\binom{105}{7+x}} \left(\frac{70-7}{105-7-x} \right)$$

5.16 (a) Basic Poisson questions, just convert seconds to hours.

$$\frac{(540 \cdot \frac{1}{120})^{11} e^{-(540 \cdot \frac{1}{120})}}{11!} = 4.26438 \times 10^{-3}$$

and

$$1 - \left(\sum_{k=0}^{10} \frac{(540 \cdot \frac{1}{120})^k e^{-(540 \cdot \frac{1}{120})}}{k!} \right) = 6.668672 \times 10^{-3}$$

(b) This is binomial on the first part of part (a)

$$\binom{20}{2} (0.004264)^2 (1 - 0.004264)^{18} = 3.19877 \times 10^{-3}$$

(c) (i) It's not quite the same as part (b), this time we're using negative binomial distribution to find the probability of only 11 successes in 1399 trials, then the 1400th being golden.

$$\binom{1399}{11} (0.004264)^{11} (1 - 0.004264)^{1388} \cdot 0.004264$$

- (ii) Check page 90 of textbook for "use the Poisson distribution with $\mu = np$ as a close approximation to the binomial distribution $Bi(n, p)$ in processes for which n is large and p is small."

$$\frac{(0.004264 \times 1399)^{11} e^{-(0.004264 \times 1399)}}{11!} \times 0.004264 = 9.331 \times 10^{-5}$$

- 5.17 (a) The probability of a single sheet of glass with no bubbles is

$$\frac{(1.2 \times 0.8)^0 e^{-(1.2 \times 0.8)}}{0!}$$

Therefore,

$$f(x) = P(X = x) = \binom{n}{x} (e^{-0.96})^x (1 - e^{-0.96})^{n-x}$$

- (b)

$$\begin{aligned} \frac{1}{2} &\leq e^{-\lambda \times 0.8} \\ \frac{\ln(2)}{0.8} &\geq \lambda \\ \lambda &\leq 0.866 \end{aligned}$$

- 5.18 $4k^2 = 1 \implies k = 0.5$, then just draw the same little table for $f(x)$, and add up $f(3)$ and $f(4)$.

- 5.19 (a)

$$\begin{aligned} P(Y \geq y) &= P(Y = y) + P(Y = y + 1) + \dots \\ &= p(1-p)^y + p(1-p)^{y+1} + \dots \\ &= p(1-p)^y (1 + (1-p) + (1-p)^2 + \dots) \\ &= \frac{p(1-p)^y}{1 - (1-p)} \\ &= \frac{p(1-p)^y}{p} \\ &= (1-p)^y \end{aligned}$$

Then,

$$\begin{aligned} P(Y \geq s+t | Y \geq s) &= \frac{P(Y \geq s+t)}{P(Y \geq s)} \\ &= \frac{(1-p)^{s+t}}{(1-p)^s} \\ &= (1-p)^t \\ &= P(Y \geq t) \end{aligned}$$

- (b) Since $p \leq 1$, and $(1-p)$ is therefore less than one, the probability function is decreasing. The most probable value of Y is 0. $P(Y = 0) = p$.

(c) Call this event X ,

$$\begin{aligned} P(X) &= P(Y = 0) + P(Y = 3) + P(Y = 6) + \dots \\ &= p(1-p)^0 + p(1-p)^3 + p(1-p)^6 + \dots \\ &= p(1 + (1-p)^3 + (1-p)^6 + \dots) \\ &= \frac{p}{1 - (1-p)^3} \end{aligned}$$

(d)

$$\begin{aligned} P(R = r) &= p(1-p)^r + p(1-p)^{r+3} + p(1-p)^{r+6} + \dots \\ &= p(1-p)^r(1 + (1-p)^3 + (1-p)^6 + \dots) \\ &= \frac{p(1-p)^r}{1 - (1-p)^3} \end{aligned}$$

5.2 Tutorial 3

Recall the Poisson approximation to Binomial, for n large and p small, $\mu = np$ (so $p = \frac{\mu}{n}$). Then

$$\binom{n}{x} p^x (1-p)^{n-x} \approx \frac{\mu^x e^{-\mu}}{x!}$$

Example 5.1. For example, consider $X \sim Bi(100, 0.02)$

i. Find the probability of $P(X = 0)$, $P(X = 1)$, and $P(X = 2)$

$$P(X = 0) = \binom{100}{0} (0.02)^0 (0.98)^{100} = 0.1326$$

$$P(X = 1) = \binom{100}{1} (0.02)^1 (0.98)^{99} = 0.2707$$

$$P(X = 2) = \binom{100}{2} (0.02)^2 (0.98)^{98} = 0.2734$$

ii. Use Poisson approximation to find probabilities in i. ($\mu = 100 \cdot 0.02 = 2$)

$$P(X = 0) = \frac{\mu^0 e^{-\mu}}{0!} = 0.135$$

$$P(X = 1) = \frac{\mu^1 e^{-\mu}}{1!} = 0.2707$$

$$P(X = 2) = \frac{\mu^2 e^{-\mu}}{2!} = 0.2707$$

Example 5.2. Say $X \sim NB(2, 0.02)$.

- Find probability $P(X = 99)$

- Use Poisson approx to find $P(X = 99)$ Let

$$f(x) = P(X = x) = \binom{x+k-1}{x} p^k (1-p)^x$$

Then

$$\begin{aligned} P(X = 99) &= \binom{99+2-1}{99} 0.02^2 \cdot 0.98^{99} \\ &= \binom{100}{99} (0.02)^2 (0.98)^{99} \\ &= \binom{100}{1} (0.02)(0.98)^{99} \cdot 0.02 \\ &= \frac{\mu^1}{1!} e^{-\mu} \cdot 0.02 \end{aligned}$$

Example 5.3. Traffic accident occurs following a poisson process. It is known that 10% of days have no accidents.

- Let x be the number of accidents in a day. Find distribution of X .
- Find the probability that there are two accidents in a day.
- Find the probability that there are four accidents in two days.
- Find the probability that there are at least 2 accidents in a day.
- Find the probability that, among a week of 7 days, there are exactly 3 days which have at least 2 accidents per day.

Then for i., $X \sim Poi(\mu)$, and $\mu = \lambda t$ not available. However, we know that $x = 0$ means no accidents, and then

$$P(X = 0) = e^{-\mu} = 0.1 \implies \mu = 2.3$$

So, $X \sim Poi(2.3)$.

For ii., $P(X = 2) = \frac{\mu^2}{2!} e^{-\mu}$.

For iii., let y be the number of accidents in 2 days. Then $y \sim Poi(2.3 \times 2) = Poi(4.6)$. So,

$$P(y = 4) = \frac{4.6^4}{4!} e^{-4.6}$$

For iv., Let x be the number of accidents in a day, $X \sim Poi(2.3)$. Then

$$\begin{aligned} p &= P(X \geq 2) \\ &= 1 - [f(0) + f(1)] \\ &= 1 - [e^{-\mu} + \mu e^{-\mu}] \\ &= 0.67 \end{aligned}$$

For v., let y be the number of days among a week of seven days that have at least 2 accidents. Then $y \sim Bi(7, 0.67)$. $P(Y = 3) = \binom{7}{3} 0.67^3 0.33^4$.

For vi. Find the probability that we have to observe a total of 10 days to see the third day with greater than or equal to 2 accidents in the day. Let y be the days required to see 3rd day with ≥ 2 accidents.

$$P(y = 10) = \binom{9}{2} (0.67)^3 (0.33)^7$$

(could also use y the number of failures before 3rd success, and then $y \sim NB(3, 0.67)$)

Now vii, if there are a total of 13 accidents in a week of 7 days, what's the probability that there are 6 accidents in the first 3 days of the week?

Let y be the number of accidents in 7 days. Let x_1 be the number of accidents in the first 3 days. Let x_2 be the number of accidents in the last 4 days. Then,

- $y \sim Poi(\mu)$ for $\mu = 2.3 \times 7$.
- $x_1 \sim Poi(\mu_1)$ for $\mu_1 = 2.3 \times 3$.
- $x_2 \sim Poi(\mu_2)$ for $\mu_2 = 2.3 \times 4$.

Then

$$\begin{aligned} P(x = 6 | y = 13) &= \frac{P(x_1 = 6 \cap y = 13)}{P(y = 13)} \\ &= \frac{P(x_1 = 6 \cap x_2 = 7)}{P(y = 13)} \\ &= \frac{P(x_1 = 6)P(x_2 = 7)}{P(y = 13)} \\ &= \frac{\frac{\mu_1^6 e^{-\mu_1}}{6!} \cdot \frac{\mu_2^7 e^{-\mu_2}}{7!}}{\frac{\mu^{13} e^{-\mu}}{13!}} \\ &= \binom{13}{6} \left(\frac{3}{7}\right)^6 \left(\frac{4}{7}\right)^7 \end{aligned}$$

6 Chapter 7

6.1 Tutorial 4

(1) Moment generating function:

$$M(t) = E(e^{tx}), \quad t \in (-h, h), h > 0$$

If X is discrete with probability function $f(x)$,

$$M(t) = \sum_x e^{tx} f(x)$$

Then,

$$- X \sim Bi(n, p), M(t) = (pe^t + q)^n, \text{ any } t$$

$$- X \sim Geo(p), M(t) = \frac{p}{1-qe^t}, t < -\ln(q)$$

$$- X \sim Poi(\mu), M(t) = e^{\mu(e^t-1)}, \text{ any } t$$

(2) X and Y are two discrete random variables with probability functions $f_x(x)$ and $f_y(y)$, MGFs that are $M_x(t)$ and $M_y(t)$, $t \in (-h, h)$, $h > 0$. Then,

$$f_x(x) = f_y(x) \quad \text{for all } x$$

if and only if

$$M_x(t) = M_y(t) \quad \text{for } t \in (-h, h)$$

Two random variables have the same distribution if and only if they have the same moment generating function.

Example 6.1. If X has Moment Generating Function,

$$M(t) = 0.5e^t + 0.2e^{3t} + 0.3e^{5t}$$

find the probability function of X

$$X \sim f(x) : \begin{array}{c|ccc} x & 1 & 3 & 5 \\ \hline f(x) & 0.5 & 0.2 & 0.3 \end{array}$$

(3) X_n has probability function $f_n(x)$, moment generating function $M_n(x)$; X has probability function $f(x)$, moment generating function $M(t)$

If $M_n(t) \rightarrow M(t)$, $t \in (-h, h)$ as $n \rightarrow \infty$. Then,

$$f_n(x) \rightarrow f(x) \text{ for all } x, \text{ as } n \rightarrow \infty$$

Example 6.2. Let $p = \frac{\mu}{n}$. Use the MGF technique to show that

$$\binom{n}{x} p^x (1-p)^{n-x} \rightarrow \frac{\mu^x}{x!} e^{-\mu}$$

when $n \rightarrow \infty$.

$$X_n \sim Bi(n, p) : M_n(t) = (pe^t + q)^n$$

$$X \sim Poi(\mu) : M(t) = e^{\mu(e^t - 1)}$$

Then

$$\begin{aligned} (pe^t + q)^n &= (pe^t + 1 - p)^n \\ &= \left(1 + \frac{\mu}{n}(e^t - 1)\right)^n \\ &= \left(1 + \frac{c}{n}\right)^n \end{aligned} \quad x = \mu(e^t - 1)$$

And $(n \rightarrow \infty) \rightarrow e^c = e^{\mu(e^t - 1)}$

(4) If $y = aX + b$ for some constants a and b , then

$$M_y(t) = e^{bt} M_x(at)$$

special case:

If $y = X + b$, then $M_y(y) = e^{bt} M_x(t)$. Since by definition,

$$M_x(t) = E(e^{tx})$$

$$\begin{aligned} M_y(t) &= E(e^{ty}) \\ &= E(e^{t(aX+b)}) \\ &= E(e^{bt} e^{atX}) \\ &= e^{bt} E(e^{atX}) \\ &= e^{bt} M_x(at) \end{aligned}$$

Question 7.7 from the book: Suppose that n people take a blood test for a disease, where each person has probability p of having the disease, independent of the other person. To save time and money, blood samples from k people are pooled and analyzed together. If none of the k persons has the disease then the test will be negative, but otherwise it will be positive. If the pooled test is positive then each of the k persons is tested separately (so $k + 1$ tests are done in that case).

(a) Let X be the number of tests required for a group of k people, show that

$$E(X) = k + 1 - k(1 - p)^k$$

Let X be the number of tests required for a group of k persons, then $\frac{x}{f(x)} \mid \frac{1}{(1-p)^k} \frac{k+1}{1-(1-p)^k}$ Then

$$\begin{aligned} E(x) &= 1 \times (1-p)^k + (k+1) \times (1 - (1-p)^k) \\ &= k + 1 - k(1-p)^k \end{aligned}$$

(b) What is the expected number of tests required for n/k groups of k people each? If $p = 0.01$, evaluate this for cases $k = 1, 5, 10$.

We have that

$$\begin{aligned}\frac{n}{k} \cdot E(x) &= \frac{n}{k} \left(k + 1 - k(1-p)^k \right) \\ &= n + \frac{n}{k} - n(1-p)^k \quad (*)\end{aligned}$$

Question 1: Show that $(*) \approx n \left(kp + \frac{1}{k} \right)$

Question 2: $(*)$ is minimized at $k = p^{-1/2}$

- (c) Show that if p is small, the expected number of tests in part (b) is approximately $n(kp + k^{-1})$, and is minimized for $k \approx p^{-1/2}$.

Since p is small,

$$\begin{aligned}(1-p)^k &= 1 - kp + \frac{k(k-1)}{2!}p^2 + \dots \\ &\approx 1 - kp\end{aligned}$$

Then,

$$\begin{aligned}n + \frac{n}{k} - n(1-p)^k &\approx n + \frac{n}{k} - n(1 - kp) \\ &= n \left(\frac{1}{k} + kp \right) \\ &= n \left(\left(\frac{1}{\sqrt{k}} - \sqrt{k}\sqrt{p} \right)^2 + 2\sqrt{p} \right)\end{aligned}$$

This is minimized when the squared term is 0, so

$$\begin{aligned}\frac{1}{\sqrt{k}} &= \sqrt{k}\sqrt{p} \\ k &= p^{-1/2}\end{aligned}$$

Question 7.11 from the book: Find the distributions that correspond to the following moment-generating functions:

- (a) $M(t) = \frac{1}{3e^{-t}-2}$ for $t < \ln(3/2)$

Then,

$$\begin{aligned}M(t) &= \frac{e^t}{3 - 2e^t} \\ &= \frac{e^t}{3} \underbrace{\left(\frac{1}{1 - \frac{2}{3}e^t} \right)}_{M_x(t)}\end{aligned}$$

then $M_x(t) : X \sim Geo\left(\frac{1}{3}\right)$. Then,

$$M_y(t) = e^t M_x(t)$$

Then $y = x + 1$: The total number of trials to the get the first success.

$$f(y) = P(Y = y) = q^{y-1}p, \quad y = 1, 2$$

(b) $M(t) = e^{2(e^t-1)}$ for $t < \infty$

Question 7.13 from the book: Let X be a random variable in the set $\{0, 1, 2\}$ with moments $E(X) = 1$ and $E(X^2) = \frac{3}{2}$.

Note. When you have a random variable $X \sim f(x)$ we notice

x		0	1	2
$f(x)$		p_0	p_1	p_2

then

$$E(X) = 1 : 0 \times p_0 + 1 \times p_1 + 2 \times p_2 = 1$$

and

$$E(X^2) = \frac{3}{2} : 0^2 \times p_0 + 1^2 \times p_1 + 2^2 \times p_2 = \frac{3}{2}$$

and

$$p_0 + p_1 + p_2 = 1$$

6.2 Solutions to Problems on Chapter 7

7.1 The mean multiplies each x by its $f(x)$, so

$$\frac{11}{40} \times 1 + \sum_{x=2}^6 \frac{x}{2x} = \frac{11}{40} + \frac{5}{2} = 2.775$$

The variance is then $E((X - 2.775)^2)$ which is

$$Var(X) = (1 - 2.775)^2 \frac{11}{40} + \sum_{x=2}^6 \frac{(x - 2.775)^2}{2x} = 2.574375$$

7.2 $E(X)$ is the expected winnings based on the expected number of tosses, $g(X)$.

$$E(g(X)) = 0.5^1 2^1 + 0.5^2 2^2 + 0.5^3 2^3 + 0.5^4 2^4 + 0.5^5 2^5 + \sum_{x>5} 0.5^x (-256) = 5 + \frac{-256 \left(\frac{1}{2}\right)^6}{1 - \frac{1}{2}} = -3$$

7.3 Expected cost per person if a population is tested for the disease using the inexpensive test followed, if necessary, by the expensive test. First let $X = 1$ if the initial test is negative, $X = 2$ if the initial test is positive. Let $g(x)$ be the total cost of testing the person, then

$$E(g(X)) = g(1)f(1) + g(2)f(2)$$

Then

$$P(X = 2) = P(D) + P(R|\bar{D})P(\bar{D}) = 0.02 + 0.05 \times (1 - 0.02) = 0.069$$

Then

$$E(g(X)) = 110 \times 0.069 + 10 \times (1 - 0.069) = \$16.90$$

7.4 $P(\text{condition}) = P(C) = 0.02$. $P(A|C) = 0.8$ and $P(A|\bar{C}) = 0.05$. Also, $P(B|C) = 1$ and $P(B|\bar{C}) = 1$. Then,

(a) We want

$$\sum_{i=1}^{150} P(C) = 3$$

(b) Let $g(X)$ be the expected total cost for the same X as last time, then

$$P(X = 1) = P(A|C)P(C) + P(A|\hat{C})P(\hat{C}) = 0.065$$

Thus,

$$E(g(X)) = 101 \times 0.065 + 1 \times (1 - 0.065) = 7.5$$

So for 2000 people, that's $7.5 \times 2000 = \$15000$. Then the expected number of detected cases are the ones where someone has the condition, and is verified by A , because if they are verified by A and don't have it, obviously the second test B will confirm they are alright.

$$2000 \times 0.08 \times 0.02 = 32$$

7.5 (a) Expected "winnings" if you bet a dollar is

$$E(X) = \left(\frac{18}{37}\right) \times 1 - \left(\frac{19}{37}\right) \times 1$$

multiply by ten for ten plays and we get $-\$ \frac{10}{37}$.

If you spend 10 on a single play, your expected winnings are

$$E(X) = \frac{18}{37} \times 10 - \frac{19}{37} \times (10)$$

reduce by the 10 you put in again and find it's the same number from part (a).

(b) The probability that your winnings are positive if you bet one dollar ten times is the probability that you win at least 6, 7, 8, 9, or 10 games

$$\sum_{x=6}^{10} \binom{10}{x} \left(\frac{18}{37}\right)^x \left(\frac{19}{37}\right)^{10-x}$$

the other result is straightforward, the probability you win one game.

7.6 The probability of each can be calculated by multiplying the probability of getting it on a single wheel for all three wheels, thus

$$E(g(X)) = 20 \times \left(\frac{2}{10}\right) \left(\frac{6}{10}\right) \left(\frac{2}{10}\right) + 10 \times \left(\frac{4}{10}\right) \left(\frac{3}{10}\right) \left(\frac{3}{10}\right) + 5 \times \left(\frac{4}{10}\right) \left(\frac{1}{10}\right) \left(\frac{5}{10}\right)$$

which is \$0.94.

7.7 (a) There are two cases, either no one has it or at least one person has it. Probability of no one having it is $(1-p)^k$, in which case 1 test is done. The probability of at least one person having it is $1 - (1-p)^k$ in

which case $k + 1$ tests are done.

$$\begin{aligned} E(X) &= (1-p)^k + (1 - (1-p)^k)(k+1) \\ &= (1-p)^k + k+1 - (1-p)^k(k+1) \\ &= k+1 - k(1-p)^k \end{aligned}$$

(b) Just multiply by the number of groups from part(a). $n + \frac{n}{k} + n(1-p)^k$.

(c) Since p is small,

$$\begin{aligned} (1-p)^k &= 1 - kp + \frac{k(k-1)}{2!}p^2 + \dots \\ &\approx 1 - kp \end{aligned}$$

Then,

$$\begin{aligned} n + \frac{n}{k} - n(1-p)^k &\approx n + \frac{n}{k} - n(1 - kp) \\ &= n \left(\frac{1}{k} + kp \right) \\ &= n \left(\left(\frac{1}{\sqrt{k}} - \sqrt{k}\sqrt{p} \right)^2 + 2\sqrt{p} \right) \end{aligned}$$

This is minimized when the squared term is 0, so

$$\begin{aligned} \frac{1}{\sqrt{k}} &= \sqrt{k}\sqrt{p} \\ k &= p^{-1/2} \end{aligned}$$

7.8 The probability a radio is defective is 0.05. In a single carton of n radios, we can find the expected value for the amount that the manufacturer will pay to the reailer, it is

$$E(g(x)) = 59.50n - 25 - 200 \cdot E(X^2)$$

Find $E(X^2)$ using the fact that $Var(X) = E(X^2) + E(X)^2$, and that for a binomial distribution $Var(X) = np(1-p)$ and $E(X) = np$. Plug in to the above equation and find max using methods from Calculus.

7.9 (a) By definition,

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \sum_{x \geq 0} e^{tx} f(x) \\ &= \sum_{x \geq 0} e^{tx} p(1-p)^x \\ &= p \sum_{x \geq 0} (e^t(1-p))^x \\ &= \frac{p}{1 - e^t(1-p)} \end{aligned}$$

with the constraint that $e^t(1-p) < 1$.

(b)

$$\begin{aligned} E(X) &= M'(0) \\ &= \frac{-p}{(1 - (1-p))^2} (-1)(1-p) \\ &= \frac{1-p}{p} \end{aligned}$$

$$\begin{aligned} E(X^2) &= M''(0) \\ &= \frac{(1-p)^2 + p(1-p)}{p^2} \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ &= \frac{(1-p)^2 + p(1-p)}{p^2} - \frac{(1-p)p}{p^2} \\ &= \left(\frac{1-p}{p}\right)^2 \end{aligned}$$

(c)

7.10 (a) If X is the number of comparisons needed, then since in general $E(X) = \sum_{i=1}^n E(X|A_i)P(A_i)$, it's clear that

$$C_n = \sum_{i=1}^n E(X|\text{initial pivot if the } i\text{th smallest number}) \left(\frac{1}{n}\right)$$

(b) Once we have a pivot, we need to split the remaining $n-1$ numbers into 2 groups (one comparison each), and then solve the two sub-problems; one of size $i-1$ and one of size $n-i$, respectively. So,

$$E(X|\text{initial pivot if the } i\text{th smallest number}) = n-1 + C_{i-1} + C_{n-i}$$

The recursion comes from plugging in this result to the formula from part a.

$$C_n = n-1 + \frac{2}{n} \sum_{k=1}^{n-1} C_k, \quad n = 2, 3, \dots$$

(c) So,

$$\begin{aligned}
 C_{n+1} &= n + \frac{2}{n+1} \sum_{k=1}^n C_k && n = 2, 3, \dots \\
 (n+1)C_{n+1} &= (n+1)n + 2 \sum_{k=1}^n C_k \\
 &= (n+1)n + 2C_n + 2 \sum_{k=1}^{n-1} C_k \\
 &= (n+1)n + 2C_n + n(C_n - n + 1) \\
 &= 2n + (n+2)C_n && n = 1, 2, \dots
 \end{aligned}$$

(d) From part c,

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \frac{C_{n+1}}{n+1} &= \lim_{n \rightarrow \infty} \frac{2n + (n+2)C_n}{(n+1)^2} \\
 &= \lim_{n \rightarrow \infty} \frac{2n + nC_n + 2C_n}{n^2 + 2n + 1} \\
 &= \lim_{n \rightarrow \infty} \frac{2n + n \left(n - 1 + \frac{2}{n} \sum_{k=1}^{n-1} C_k \right) + 2 \left(n - 1 + \frac{2}{n} \sum_{k=1}^{n-1} C_k \right)}{n^2 + 2n + 1} \\
 &= \lim_{n \rightarrow \infty} \frac{2n + \left(n^2 - n + 2 \sum_{k=1}^{n-1} C_k \right) + \left(2n - 2 + \frac{4}{n} \sum_{k=1}^{n-1} C_k \right)}{n^2 + 2n + 1} \\
 &= \lim_{n \rightarrow \infty} \frac{3n + n^2 + \left(1 + \frac{4}{n} \right) \sum_{k=1}^{n-1} C_k}{n^2 + 2n + 1}
 \end{aligned}$$

7.11 Look above

7.12

$$\begin{aligned}
 M(t) &= E(e^{tx}) \\
 &= \sum_{x=a}^b e^{tx} \frac{1}{b-a+1} \\
 &= \frac{1}{b-a+1} \sum_{x=a}^b e^{tx} \\
 &= \frac{1}{b-a+1} (e^{at} + e^{(a+1)t} + \dots + e^{bt}) \\
 &= \frac{1}{b-a+1} \left(\frac{e^{(b+1)t} - e^{at}}{e^t - 1} \right)
 \end{aligned}$$

If $a = b$ then

$$M(t) = e^{at}$$

If $b = a + 1$ then

$$M(t) = \frac{1}{2}e^{at}(1 + e^t)$$

7.13 First, $\frac{x}{f(x)} \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline & p_0 & p_1 & p_2 \end{array}$ Then we know that

$$E(X) = 1 : 0 \times p_0 + 1 \times p_1 + 2 \times p_2 = 1$$

and

$$E(X^2) = \frac{3}{2} : 0^2 \times p_0 + 1^2 \times p_1 + 2^2 \times p_2 = \frac{3}{2}$$

and

$$p_0 + p_1 + p_2 = 1$$

Solving this system of equations gets us $p_0 = \frac{1}{4}$, $p_1 = \frac{1}{2}$ and $p_2 = \frac{1}{4}$. Then,

(a) $M(t) = 0.25 + 0.5e^t + 0.25e^{2t}$

(b) $M^{(k)}(0)$ for $k = 1, 2, 3, 4, 5, 6$,

$$M'(0) = 1, M''(0) = \frac{3}{2}, M^{(3)}(0) = \frac{5}{2}, M^{(4)}(0) = \frac{9}{2}, M^{(5)}(0) = \frac{17}{2}, M^{(6)}(0) = \frac{33}{2}$$

(c) See above

(d) For given values of $E(X) = \mu_1$, $E(X^2) = \mu_2$, there is a unique solution to $p_0 + p_1 + p_2 = 1$, $p_1 + 2p_2 = \mu_1$, and $p_1 + 4p_2 = \mu_2$. Then,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & \mu_1 \\ 0 & 1 & 4 & \mu_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 0 & 1 + \frac{\mu_2}{2} - \frac{3\mu_1}{2} \\ 0 & 1 & 0 & 2\mu_2 - \mu_1 \\ 0 & 0 & 1 & \frac{1}{2}\mu_2 - \mu_1 \end{pmatrix}$$

7 Chapter 8

7.1 Solutions to Problems on Chapter 8

8.1 (a) We check if $f(x, y) = f_1(x)f_2(y)$ for all (x, y) . Consider $(0, 0)$. Then $f(0, 0) = .15 \neq f_1(0)f_2(0) = (0.15 + 0.35)(0.15 + 0.1 + 0.5)$.

(b) $P(X > Y) = f(1, 0) + f(2, 0) + f(2, 1) = 0.3$ and $P(X = 1|Y = 0) = \frac{f(1,0)}{f_2(0)} = \frac{1}{3}$.

8.2 (a) Since $X \sim \text{Poisson}(\mu_1 = 0.10)$ and $Y \sim \text{Poisson}(\mu_2 = 0.5)$ we have

$$P(X = x) = \frac{\mu_1^x e^{-\mu_1}}{x!} \quad P(Y = y) = \frac{\mu_2^y e^{-\mu_2}}{y!}$$

If X and Y are independent, then

$$P(X = x, Y = y) = \frac{\mu_1^x e^{-\mu_1}}{x!} \frac{\mu_2^y e^{-\mu_2}}{y!}$$

parameterizing for $P(T = t) = P(X + Y = t) = \sum_x P(X = x, Y = t - x)$ then we have

$$\begin{aligned}
 P(T = t) &= \sum_x \frac{\mu_1^x e^{-\mu_1}}{x!} \frac{\mu_2^{t-x} e^{-\mu_2}}{(t-x)!} \\
 &= 0.05^t e^{-0.15} \sum_x \frac{1}{x!(t-x)!} \left(\frac{0.01}{0.005}\right)^x \\
 &= 0.05^t e^{-0.15} \frac{1}{t!} \sum_x \frac{t!}{x!(t-x)!} \left(\frac{0.01}{0.005}\right)^x \\
 &= 0.05^t e^{-0.15} \frac{1}{t!} \sum_x \binom{t}{x} \left(\frac{0.01}{0.005}\right)^x \\
 &= \frac{0.05^t e^{-0.15}}{t!} \left(1 + \frac{0.1}{0.05}\right)^t \\
 &= \frac{e^{-0.15} 0.15^t}{t!}
 \end{aligned}$$

Then we want $P(t > 1)$, it's just

$$1 - (P(T = 1) + P(T = 0)) \approx 0.0102$$

(b) The mean and variance of a poisson distribution are both μ , so we get 0.15 for both.

8.4 Let's say the events are A, O , and T , and $P(A) = p, P(O) = q, P(T) = 1 - p - q$.

(a) We want $P(X = x, Y = y)$. This is a case of negative binomial distribution, the first combinatorial term is just the number of ways that we can arrange the x of type A , y of type O , and 9 of type T before the tenth T .

$$f(x, y) = \frac{(x + y + 9)!}{x!y!9!} p^x q^y (1 - p - q)^{10}$$

(b) The conditional probability function:

$$\begin{aligned}
 f(y|x) &= \frac{f(x, y)}{f_1(x)} \\
 &= \frac{\frac{(x+y+9)!}{x!y!9!} p^x q^y (1-p-q)^{10}}{f_1(x)}
 \end{aligned}$$

Now,

$$\begin{aligned}
 f_1(x) &= \sum_{y \geq 0} \frac{(x+y+9)!}{x!y!9!} p^x q^y (1-p-q)^{10} \\
 &= \sum_{y \geq 0} \frac{(x+y+9)!}{x!y!9!} p^x q^y (1-p-q)^{10} \\
 &= \frac{p^x (1-p-q)^{10}}{x!9!} \sum_{y \geq 0} \frac{(x+y+9)!}{y!} q^y \\
 &= \frac{p^x (1-p-q)^{10}}{x!9!} (x+9)! \sum_{y \geq 0} \binom{x+y+9}{y} q^y \\
 &= \frac{p^x (1-p-q)^{10} \binom{x+9}{x}}{(1-q)^{x+10}} \sum_{y \geq 0} \binom{x+y+10-1}{y} q^y (1-q)^{x+10} \\
 &= \binom{x+9}{x} \left(\frac{p}{1-q} \right)^x \left(\frac{1-p-q}{1-q} \right)^{10}
 \end{aligned}$$

Then

$$\begin{aligned}
 f(y|x) &= \frac{f(x,y)}{f_1(x)} \\
 &= \frac{\frac{(x+y+9)!}{x!y!9!} p^x q^y (1-p-q)^{10}}{\binom{x+9}{x} \left(\frac{p}{1-q} \right)^x \left(\frac{1-p-q}{1-q} \right)^{10}} \\
 &= \binom{x+y+9}{y} q^y (1-q)^{x+10}
 \end{aligned}$$

8.7 On any draw, of the 5 yellow balls you choose x , and of the 3 remaining you choose $2-x$. Then, of the remaining yellow balls you choose $y-x$ and then $2+x-y$ of the remaining.

$$f(x,y) = \frac{\binom{5}{x} \binom{3}{2-x} \binom{5-x}{y-x} \binom{1+x}{2+x-y}}{\binom{8}{2} \binom{6}{2}}$$

for part (b), to determine if they are not independent we need to find a case where $f(x,y) \neq f_1(x)f_2(y)$. Note $f_1(0) \neq 0$, $f_2(3) \neq 0$ but $f(0,3) = 0$.

8.8 Hypergeometric,

(a)

$$f(x,y) = \frac{\binom{2}{x} \binom{1}{y} \binom{7}{3-x-y}}{\binom{10}{3}}$$

(b) Marginal probability functions are $f_1(x)$, $f_2(y)$, note that since

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

$$f_1(x) = \frac{\binom{2}{x} \binom{8}{3-x}}{\binom{10}{3}}$$

$$f_2(x) = \frac{\binom{1}{y} \binom{9}{3-y}}{\binom{10}{3}}$$

(c) Then $P(X = Y)$ is $f(x, x)$

$$f(x, x) = f(0, 0) + f(1, 1) = \frac{49}{120}$$

and $P(X = 1|Y = 0)$ is $f(1|0) = \frac{f(1,0)}{f_2(0)} = \frac{1}{2}$

8.9 (a) The marginal probability function of X is

$$f_1(x) = \sum_{y \geq 0} k \frac{2^{x+y}}{x!y!}$$

$$= \frac{2^x k}{x!} \sum_y \frac{2^y}{y!}$$

$$= \frac{2^x k}{x!} \left(1 + \frac{2}{1!} + \frac{2^2}{2!} \right)$$

$$= \frac{2^x k}{x!} e^2$$

(b) Evaluating k ...

$$\sum_x f_1(x) = 1$$

$$\sum_x \frac{k 2^x}{x!} e^2 = 1$$

$$k = \frac{1}{e^2 \sum_x \frac{2^x}{x!}}$$

$$k = e^{-4}$$

(c) Might as well check, note

$$f_2(y) = \frac{k 2^y}{y!} e^2$$

by symmetry, then

$$\begin{aligned}
 f(x, y) &= f_1(x)f_2(y) \\
 &= \frac{2^x}{x!e^2} \frac{2^y}{y!e^2} \\
 &= \frac{2^{x+y}}{x!y!e^4} \\
 &= k \frac{2^{x+y}}{x!y!} \\
 &= f(x, y)
 \end{aligned}$$

so yeah, independent

(d) $P(T = X + Y) = f(t)$ is

$$\begin{aligned}
 \sum_x f(x, t-x) &= \sum_x \frac{2^{x+t-x}}{x!(t-x)!e^4} \\
 &= \frac{2^t}{e^4} \frac{1}{t!} \sum_x \binom{t}{x} \\
 &= \frac{2^t}{t!e^4} (1+1)^t \\
 &= \frac{2^{2t}}{t!e^4}
 \end{aligned}$$

8.10 (a) Let Y represent number of Type A eveners that occur in a 1-min period. Let X be the total number of events in a 1 minute period. Then, $P(X = \text{type } A) = p$ and $X \sim \text{Poisson}(\mu)$.

$$\begin{aligned}
 P(Y = y) &= \sum_x P(Y = y|X = x)P(X = x) \\
 &= \sum_x f(y|x)f_1(x) \\
 &= \sum_x \binom{x}{y} p^y (1-p)^{x-y} \frac{e^{-\lambda} \lambda^x}{x!} \\
 &= e^{-\lambda} p^y \lambda^y \sum_x \binom{x}{y} \frac{(\lambda(1-p))^{x-y}}{x!} \\
 &= \frac{e^{-\lambda}}{y!} p^y \lambda^y \sum_x \frac{(\lambda(1-p))^{x-y}}{(x-y)!} \\
 &= \frac{e^{-\lambda} (p\lambda)^y}{y!} e^{\lambda(1-p)} \\
 &= \frac{(p\lambda)^y e^{-\lambda p}}{y!}
 \end{aligned}$$

$$(b) P(Y = y) = \frac{e^{-0.05 \times 3 \times 30} (0.05 \times 3 \times 30)^y}{y!}$$

$$1 - \sum_{y=0}^4 P(Y = y)$$

8.11 (a) Multinomial distribution...

$$P(10 \text{ of each type}) = \frac{40!}{10!10!10!10!} \left(\frac{3}{16}\right)^{10} \left(\frac{5}{16}\right)^{10} \left(\frac{5}{16}\right)^{10} \left(\frac{3}{16}\right)^{10}$$

(b)

$$\frac{40!}{16!24!} \left(\frac{8}{16}\right)^{40} = \binom{40}{16} \left(\frac{3+5}{16}\right)^{16} \left(\frac{5+3}{16}\right)^{24}$$

(c)

$$\begin{aligned} P(10 \text{ of 1} \mid 16 \text{ of 1 and 2}) &= \frac{\frac{40!}{10!6!24!} \left(\frac{3}{16}\right)^{10} \left(\frac{5}{16}\right)^6 \left(\frac{8}{16}\right)^{24}}{\frac{40!}{24!16!} \left(\frac{8}{16}\right)^{16} \left(\frac{8}{16}\right)^{24}} \\ &= \binom{16}{10} \left(\frac{3}{8}\right)^{10} \left(\frac{5}{8}\right)^6 \end{aligned}$$

8.12 (a)

$$\begin{aligned} E(X) &= \sum_x x f(x) \\ &= 0 \times .20 + \cdots + 8 \times .01 \\ &= \frac{44}{25} \end{aligned}$$

(b) First,

$$f(y|x) = \binom{x}{y} 0.5^y 0.5^{x-y}$$

So,

$$\begin{aligned} f(x, y) &= f(x) f(y|x) \\ &= f(x) \binom{x}{y} 0.5^y 0.5^{x-y} \\ &= f(x) \binom{x}{y} 0.5^x \end{aligned}$$

Then,

$$f_Y(y) = \sum_{x=y}^8 f(x, y)$$

So finally,

$$E(y) = \sum_y y f_Y(y) = \frac{1}{2} E(x) = 0.88$$

8.13 (a) This is multinomial distribution,

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = \frac{10!}{x_1! x_2! x_3! x_4! x_5! x_6!} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} p_5^{x_5} p_6^{x_6}$$

not helpful. Try again:

$$\begin{aligned} \text{Var}(T) &= \text{Var}(X + Y) \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \\ n(p + q)(1 - p - q) &= np(1 - p) + nq(1 - q) + 2\text{Cov}(X, Y) \\ \text{Cov}(X, Y) &= -npq \end{aligned}$$

It should be negative since as one goes up, the other goes down.

8.17 (a) First find them for X and Y .

$$E(X) = \frac{1}{2}(2) = 1, E(Y) = \frac{1}{2}(2) = 1$$

$$\text{Var}(X) = E(X) \frac{1}{2} = \frac{1}{2}, \text{Var}(Y) = \frac{1}{2}$$

Then,

$$E(U) = E(X) + E(Y) = 2, E(V) = 0$$

$$V(U) = V(X) + V(Y) = 1, V(V) = 1$$

(b)

$$\begin{aligned} \text{Cov}(U, V) &= \text{Cov}(X + Y, X - Y) \\ &= \text{Cov}(X, X) + (-1)\text{Cov}(X, Y) + \text{Cov}(Y, X) + (-1)\text{Cov}(Y, Y) \\ &= \text{Var}(X) - \text{Var}(Y) \\ &= 0 \end{aligned}$$

(c) Nah, $U = 0$ and $V = 1$ is a counterexample.

8.18 (a) Let X_i be whether or not question i was answered right (1 for yes 0 for no). Then,

(i)

$$Y = \sum_{i=1}^{100} X_i - \frac{1}{4} \left(100 - \sum_{i=1}^{100} X_i \right)$$

(ii) $E(X_i); V(X_i)$. Let $A_i =$ "know answer to question Q_i "

$$\begin{aligned} P(X_i = 1) &= P(A_i)P(X_i = 1|A_i) + P(\bar{A}_i)P(X_i = 1|\bar{A}_i) \\ &= p_i \times 1 + (1 - p_i) \times \frac{1}{5} \\ &= \frac{4}{5}p_i + \frac{1}{5} \end{aligned}$$

Then since $X = 0$ or 1 ,

$$E(X^2) = E(X) = 0 \times f(0) + 1 \times f(1) = \frac{4}{5}p_i + \frac{1}{5}$$

and then

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ &= \left(\frac{4}{5}p_i + \frac{1}{5}\right) - \left(\frac{4}{5}p_i + \frac{1}{5}\right)^2 \\ &= \frac{12}{25}p_i - \frac{16}{25}p_i^2 + \frac{4}{25} \end{aligned}$$

So we're set to find $E(Y)$ and $\text{Var}(Y)$

$$\begin{aligned} E(Y) &= E\left(\sum_{i=1}^{100} X_i - \frac{1}{4}\left(100 - \sum_{i=1}^{100} X_i\right)\right) \\ &= \sum_{i=1}^{100} \left(\frac{5}{4}E(X_i)\right) - 25 \\ &= \frac{5}{4}\left(\sum_{i=1}^{100} \frac{4}{5}p_i + \frac{1}{5}\right) - 25 \\ &= \sum_{i=1}^{100} p_i \end{aligned}$$

as required. Next,

$$\begin{aligned} \text{Var}(Y) &= \text{Var}\left(\sum_{i=1}^{100} X_i - \frac{1}{4}\left(100 - \sum_{i=1}^{100} X_i\right)\right) \\ &= \text{Var}\left(\sum_{i=1}^{100} \frac{5}{4}X_i\right) \\ &= \frac{25}{16}\sum \text{Var}(X_i) \\ &= \sum \frac{3}{4}p_i - p_i^2 + \frac{1}{4} \\ &= \sum p_i(1 - p_i) + \frac{100 - \sum p_i}{4} \end{aligned}$$

as required.

(b)

$$Y = \sum X_i$$

$$P(X_i = 1) = P(A_i)P(X_i = 1|A_i) = p_i$$

then

$$\begin{aligned} E(X_i) &= p_i \\ \text{Var}(X_i) &= E(X_i^2) - E(X_i)^2 \\ &= p_i - p_i^2 \\ &= p_i(1 - p_i) \end{aligned}$$

so we have

$$\begin{aligned} E(Y) &= E\left(\sum X_i\right) \\ &= \sum p_i \end{aligned}$$

and

$$\begin{aligned} \text{Var}(Y) &= \text{Var}\left(\sum X_i\right) \\ &= \sum \text{Var}(X_i) \\ &= \sum p_i(1 - p_i) \end{aligned}$$

8.19

$$\begin{aligned} \text{Cov}(X + Y, X - Y) &= \text{Cov}(X, X) - \text{Cov}(Y, Y) \\ &= \text{Var}(X) - \text{Var}(Y) \\ &= 1 - 2 \\ &= -1 \end{aligned}$$

8.20 (a) We know that $L = A + B + C$. So,

$$\text{Var}(L) = \text{Var}(A) + \text{Var}(B) + \text{Var}(C) = 0.6^2 + 0.8^2 + 0.7^2$$

so $\sigma = \sqrt{\sigma^2} = 1.22$.

(b) Do the same thing as before with the new standard deviation of B , and take the quotient of the it with answer from part (a).

8.21 X = total number of islands being cut off. Let

$$X_i = \begin{cases} 1 & \text{island } i \text{ being cut off} \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, 2, 3, 4, 5$$

(i) $X = X_1 + X_2 + X_3 + X_4 + X_5$. Now we want to find the expected value of X_i , the variance, and the probability that $X_i = 1$. So,

$$\begin{aligned} E(X_i) &= p_i \\ \text{Var}(X_i) &= p_i(1 - p_i) \\ p_i &= P(X_i = 1) \end{aligned}$$

however p_i varies for $i = 5$.

$$\begin{aligned} P(X_1 = 1) &= p^3 \\ P(X_5 = 1) &= p^4 \end{aligned}$$

now, it also varies if you cut off two bridges at a time (since some bridges are shared)

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j) \\ E(X_1 X_2) &= P(X_1 = 1, X_2 = 1) = p^5 \\ E(X_1 X_5) &= P(X_1 = 1, X_5 = 1) = p^6 \end{aligned}$$

So first off,

$$E(X) = E(X_1 + X_2 + X_3 + X_4 + X_5) = 4p^3 + p^4$$

so then

$$\begin{aligned} \text{Var}(X) &= \text{Var}(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= \text{Var}(X_1) + \cdots + \text{Var}(X_5) + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \\ &= 4p^3(1 - p^3) + p^4(1 - p^4) + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \end{aligned}$$

Now,

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

I forgot to finish this solution but the rest is pretty straightforward, I'll do it tomorrow morning, just find all the possible covariances and use that Var formula above

7.2 Tutorial 5

Example 7.1. Suppose $X \sim \text{Geo}(p)$ and $Y \sim \text{Geo}(p)$, X and Y are independent.

- Find the distribution of $T = X + Y$.
- Find the conditional distribution of X given $T = t$.
- $f_T(t) = P(T = t) = \binom{t+2-1}{t} p^2 q^t$. This is intuitive because geometric is the number of failures before a success, and adding two of them should be the number of failures before the second success.
- $P(X = x | T = t) = \frac{1}{t+1}$, $x = 0, 1, \dots, t$

Example 7.2. $X \sim \text{Poi}(\mu)$, $Y \sim \text{Bi}(n, p)$, X and Y are independent.

- Find $\text{Var}(2X - Y + 1)$
- Find $\text{Cov}(X + Y, X - Y)$
- $\text{Var}(2X - Y + 1) = 4\text{Var}(X) + (-1)^2\text{Var}(Y)$
- $\text{Cov}(X + Y, X - Y) = \text{Cov}(X, X) - \text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Cov}(Y, Y)$

Example 7.3. $(X_1, X_2, \dots, X_k) \sim \text{multinomial}(n; p_1, p_2, \dots, p_k)$

- Find $\text{Cov}(x_i, x_j)$ where $(i \neq j)$.
- Find the correlation coefficient ρ between x_i and x_j .

Note that $X_i \sim Bi(n, p_i)$. Similarly, $X_j \sim Bi(n, p_j)$. Also, $X_i + X_j \sim Bi(n, p_i + p_j)$. What about $X_i | X_k = t$? $X_i | X_k = t \sim Bi\left(n - t, \frac{p_i}{1 - p_k}\right)$. So,

(a) $Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$. But what is $E(X_i X_j)$? Let's approach it differently,

$Var(X_i + X_j) = Var(X_i) + Var(X_j) + 2Cov(X_i, X_j)$. From this equation we know all three Var terms since they are all binomially distributed. So,

$$\begin{aligned} Var(X_i + X_j) &= Var(X_i) + Var(X_j) + 2Cov(X_i, X_j) \\ n(p_i + p_j)(1 - p_i - p_j) &= np_i(1 - p_i) + np_j(1 - p_j) + 2Cov(X_i, X_j) \\ Cov(X_i, X_j) &= -np_i p_j \end{aligned}$$

Example 7.4. Randomly put n letters L_1, L_2, \dots, L_n into n envelopes E_1, E_2, \dots, E_n , one letter per envelope. If letter L_i is in envelope E_i , it's called a match. Let X be the total number of matches. Find $E(X)$ and $Var(X)$.

Method 1. Use $f(x)$. Then,

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

This is hopelessly complicated. **Method 2.** Use indicator variables. Let

$$X_i = \begin{cases} 1 & \text{letter } L_i \text{ matches } E_i \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, 2, \dots, n$$

- i. $X = X_1 + X_2 + \dots + X_n$
- ii. X_1, \dots, X_n are not independent
- iii. $X_i \sim \text{Bernoulli}(p_i)$, so $f(0) = 1 - p_i$ and $f(1) = p_i$.

$$E(X_i) = p_i$$

$$Var(X_i) = p_i(1 - p_i)$$

$$\begin{aligned} p_i &= P(X_i = 1) \\ &= \frac{1}{n} \end{aligned}$$

Note that

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = n \times \frac{1}{n} = 1$$

Now,

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

now before we continue, note that:

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

$$\begin{aligned} E(X_i X_j) &= \sum_x \sum_y xyf(x, y) \\ &= f(1, 1) \\ &= P(X_i = 1, X_j = 1) \\ &= \frac{1}{n(n-1)} \end{aligned}$$

So,

$$\begin{aligned} V(X) &= V\left(\sum_{i=1}^n X_i\right) \\ &= \sum_{i=1}^n V(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \\ &= \underbrace{\sum_{i=1}^n \frac{1}{n} \left(1 - \frac{1}{n}\right)}_{1 - \frac{1}{n}} + 2 \binom{n}{2} \left[\frac{1}{n(n-1)} - \frac{1}{n^2}\right] \\ &= 1 \end{aligned}$$

Example 7.5. Question 8.21 from the book. The inhabitants of the beautiful and ancient city of Pentapolis live on 5 islands separated from each other by water. Bridges cross from one island to another as shown. On any day, a bridge can be closed, with probability p , for restoration work. Assuming that the 8 bridges are closed independently, find the mean and variance of the number of islands which are completely cut off because of restoration work.

X = total number of islands being cut off. Let

$$X_i = \begin{cases} 1 & \text{island } i \text{ being cut off} \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, 2, 3, 4, 5$$

- (i) $X = X_1 + X_2 + X_3 + X_4 + X_5$. Now we want to find the expected value of X_i , the variance, and the probability that $X_i = 1$. So,

$$\begin{aligned} E(X_i) &= p_i \\ \text{Var}(X_i) &= p_i(1 - p_i) \\ p_i &= P(X_i = 1) \end{aligned}$$

however p_i varies for $i = 5$.

$$\begin{aligned} P(X_1 = 1) &= p^3 \\ P(X_5 = 1) &= p^4 \end{aligned}$$

now, it also varies if you cut off two bridges at a time (since some bridges are shared)

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j) \\ E(X_1 X_2) &= P(X_1 = 1, X_2 = 2) = p^5 \\ E(X_1 X_5) &= P(X_1 = 1, X_5 = 1) = p^6 \end{aligned}$$

Example 7.6. Question 8.18 from the book. A multiple choice exam has 100 questions, each with 5 possible answers. One mark is awarded for a correct answer and $1/4$ mark is deducted for an incorrect answer. A particular student has probability p_i of knowing the correct answer to the i th question, independent of other questions.

- (a) Suppose that on a question where the student does not know the answer, he or she guesses randomly. Show that his or her total mark has mean $\sum p_i$ and variance $\sum p_i(1 - p_i) + \frac{(100 - \sum p_i)}{4}$.
- (b) Show that the total mark for a student who refrains from guessing also has mean $\sum p_i$, but with variance $\sum p_i(1 - p_i)$. Compare the variances when all p_i 's equal (i) .9, (ii) .5.

Let Y = total mark and let

$$X_i = \begin{cases} 1 & Q_i \text{ answered correctly} \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, 2, \dots, 100$$

- (i) $y = 1 \times \sum_{i=1}^{100} X_i - \frac{1}{4} \left(100 - \sum_{i=1}^{100} X_i \right)$
- (ii) $E(X_i); V(X_i)$

$$\begin{aligned} P(X_i = 1) &= P(A_i)P(X_i = 1|A_i) + P(\bar{A}_i)P(X_i = 1|\bar{A}_i) \\ &= p_i \times 1 + (1 - p_i) \times \frac{1}{5} \end{aligned}$$

Let A_i = "know answer to question Q_i "

8 Continuous Probability Distributions

8.1 Solutions to Problems on Chapter 9

9.1 $Y = \frac{4}{3}\pi \left(\frac{X}{2}\right)^3$. Plug in $X = 0.6$ and $X = 1.0$ to get the bounds $0.036\pi \leq y \leq \frac{\pi}{6}$. Next,

$$F(x) = \frac{x - 0.6}{0.4}$$

so then

$$\begin{aligned}
 F(y) &= P(Y \leq y) \\
 &= P\left(\frac{4}{3}\pi \left(\frac{X}{2}\right)^3 \leq y\right) \\
 &= P\left(X \leq \sqrt[3]{\frac{6y}{\pi}}\right) \\
 &= \frac{\sqrt[3]{\frac{6y}{\pi}} - 0.6}{0.4} \\
 &= \sqrt[3]{\frac{750y}{8\pi}} - \frac{3}{2}
 \end{aligned}$$

which implies

$$\begin{aligned}
 f(y) &= \frac{d}{dy}F(y) \\
 &= \frac{d}{dy} \sqrt[3]{\frac{750y}{8\pi}} - \frac{3}{2} \\
 &= \frac{5}{6} \left(\frac{6}{\pi}\right)^{\frac{1}{3}} y^{-\frac{2}{3}}
 \end{aligned}$$

9.2 (a)

$$\begin{aligned}
 \int_{-1}^1 f(x) &= 1 \\
 \int_{-1}^1 k(1-x^2)dx &= 1 \\
 kx - \frac{k}{3}x^3 \Big|_{-1}^1 &= 1 \\
 k &= \frac{3}{4}
 \end{aligned}$$

Then,

$$\begin{aligned}
 P(X \leq x) &= \int_{-1}^x \frac{3}{4}(1-x^2)dx \\
 &= \frac{3}{4}x - \frac{1}{4}x^3 \Big|_{-1}^x \\
 &= \frac{3}{4} \left(\frac{2}{3} + x - \frac{x^3}{3}\right)
 \end{aligned}$$

(b)

$$\begin{aligned}
 F(c) - F(-c) &= 0.95 \\
 \frac{3}{4} \left(\frac{2}{3} + c - \frac{x^3}{3} \right) - \frac{3}{4} \left(\frac{2}{3} - c - \frac{x^3}{3} \right) &= 0.95 \\
 c &= 0.811
 \end{aligned}$$

9.3 (a)

$$\begin{aligned}
 E(X) &= \int_x x f(x) \\
 &= \int_0^{\frac{1}{2}} 4x dx + \int_{\frac{1}{2}}^1 4(1-x) dx \\
 &= \frac{1}{2}
 \end{aligned}$$

$$\begin{aligned}
 \sigma^2 &= E(X^2) - E(X)^2 \\
 &= \int_x x^2 f(x) - \frac{1}{4} \\
 &= \frac{1}{24}
 \end{aligned}$$

9.4

$$F(x) = \frac{x}{20}$$

then since $X = 20Y - 10$,

$$F(y) = \frac{20y - 10}{20}$$

so

$$f(y) = 1$$

9.5 (a) we just need $\alpha + 1 > 0$. $\alpha > -1$.

(b)

$$\begin{aligned}
 P\left(X \leq \frac{1}{2}\right) &= \int_0^{\frac{1}{2}} (\alpha + 1)x^\alpha \\
 &= \left(\frac{1}{2}\right)^{\alpha+1}
 \end{aligned}$$

$$\begin{aligned} E(X) &= \int_0^1 xf(x) \\ &= \frac{\alpha + 1}{\alpha + 2} \end{aligned}$$

(c)

$$\begin{aligned} F(x) &= x^{\alpha+1} \\ F_T(t) &= P\left(\frac{1}{X} \leq t\right) \\ &= 1 - P\left(\frac{1}{t} > X\right) \\ &= 1 - \left(\frac{1}{t}\right)^{\alpha+1} \end{aligned}$$

so

$$\begin{aligned} f(t) &= (\alpha + 1)t^{-(\alpha+2)} \\ &= \frac{\alpha + 1}{t^{\alpha+2}} \end{aligned}$$

9.6 (a) In this case $\lambda = \frac{2}{5}$.

$$(1 - e^{-\lambda(5)})^3 = (1 - e^{-2})^3$$

(b)

$$P(\text{exceeds 4} | \text{exceeds 4}) = \frac{(1 - (1 - e^{-\lambda(5)}))}{(1 - (1 - e^{-\lambda(4)}))} = e^{-\frac{2}{5}}$$

9.7 We know that $E(X) = 1000$, therefore $1000 = \frac{1}{\lambda} \implies \lambda = \frac{1}{1000}$. So, the median lifetime x implies

$$F(x) = 0.5$$

So,

$$\begin{aligned} 0.5 &= 1 - e^{-\lambda x} \\ x &= -\frac{1}{\lambda} \ln\left(\frac{1}{2}\right) \\ &= 693.1472 \end{aligned}$$

9.8 $X \sim N(0.65, 0.01)$. Thus, $E(X) = 0.65$ and $Var(X) = 0.01$.(a) For each one just calculate the z -score, for example

$$P(X \geq 80\%) = 1 - P(X < 80\%) = 1 - F_X(0.8) = 1 - F_Z(1.5) = 1 - 0.93319 = 0.6681$$

where $F_X(x) = F_Z\left(\frac{x-\mu}{\lambda}\right)$

(b) \bar{X} is the average score in a group of 25 students. So $\bar{X} \sim N(\mu, \sigma^2/25)$

$$\begin{aligned} P(\bar{X} > 70\%) &= 1 - P(\bar{X} < 70\%) \\ &= 1 - P\left(\frac{0.7 - 0.65}{\sqrt{\frac{\sigma^2}{25}}}\right) \\ &= 1 - F_Z(2.5) \\ &= 1 - 0.99379 \\ &= 0.00621 \end{aligned}$$

(c) Introduce the random variable \bar{Y} = average score in another group of 25 students, $\bar{Y} \sim N(0.65, \sigma^2/25)$. So, $\bar{X} - \bar{Y} \sim N(0.65 - 0.65, \sigma^2/25 + \sigma^2/25)$. Then,

$$\begin{aligned} P(|\bar{X} - \bar{Y}| \geq 0.05) &= 1 - P(|\bar{X} - \bar{Y}| < 0.05) \\ &= 1 - P(|\bar{X} - \bar{Y}| < 0.05) \\ &= 1 - P\left(|Z| < \frac{0.05 - 0}{\sqrt{2\sigma^2/25}}\right) \\ &= 1 - (P(Z < 1.77) - P(Z < -1.77)) \\ &= 1 - (0.96164 - (1 - 0.96164)) \\ &= 0.0762 \quad \text{close enough?} \end{aligned}$$

9.9 Let X be the amount of water in a bottle in litres.

(a) $P(\text{less than 2 litres})$

$$\begin{aligned} P(X < 2) &= P\left(Z < \frac{2 - 2}{0.01}\right) \\ &= P(Z < 0) \\ &= 0.5 \end{aligned}$$

(b)

$$\begin{aligned} P(Z \leq d) &= P(Z \geq |d|) \\ &= 1 - P(Z \leq |d|) = 0.01 \\ \implies P(Z \leq |d|) &= 0.99 \\ |d| &= 2.3263 \\ d &= -2.3263 \end{aligned}$$

So,

$$\frac{2 - \mu}{0.01} = -2.326 \implies \mu = 2.023$$

9.10 Let X be the r.v. defined by length of the turbine shaft, and A_i the respective parts. So,

$$X = A_1 + A_2 + A_3 + A_4 \sim N(8.10 + 7.25 + 9.75 + 3.10, .22^2 + .20^2 + .24^2 + .20^2) = N(28.2, 0.186)$$

We need

$$\begin{aligned}
 P(|X - 28| < .26) &= P(27.74 < X < 28.26) \\
 &= P\left(\frac{27.74 - 28.2}{\sqrt{0.186}} < Z < \frac{28.26 - 28.2}{\sqrt{0.186}}\right) \\
 &= P(-1.07 < Z < 0.14) \\
 &= 0.55567 - (1 - 0.85769) \\
 &= 0.4134
 \end{aligned}$$

9.11 (a)

$$\begin{aligned}
 P(9.0 < X < 11.1) &= F_X(11.1) - F_X(9) \\
 &= F_Z\left(\frac{11.1 - 9.5}{2}\right) - F_Z\left(\frac{9 - 9.5}{2}\right) \\
 &= F_Z(0.8) - F_Z(-0.25) \\
 &= 0.78814 - (1 - 0.59871) \\
 &= 0.3868
 \end{aligned}$$

(b) $X + 4Y \sim N(9.5 + 4(-2.1), 2^2 + 4^2(0.75))$.

$$\begin{aligned}
 P(X + 4Y > 0) &= P\left(Z > \frac{0 - 1.1}{4}\right) \\
 &= P(Z < 0.275) \\
 &= 0.6083
 \end{aligned}$$

(c)

$$P(X > b) = 0.9 \implies P\left(Z > \frac{b - 9.5}{2}\right) = 0.9 \implies b = 6.94$$

9.12 (a)

$$\begin{aligned}
 P(A < l) &= P\left(Z < \frac{l - 1.05l}{\sqrt{0.0004l}}\right) \\
 &= P\left(Z < \frac{1 - 1.05}{\sqrt{0.0004}}\right) \\
 &= 0.062
 \end{aligned}$$

(b) Define $B = A_1 + A_2 + A_3 + \dots + A_{20}$ = volume of 20 randomly chosen bottles where $B \sim N(20(1.05), 20(0.0004))$,

$B \sim N(21, 0.008)$. Then $B - V \sim N(-1, 0.168)$ so we obtain

$$\begin{aligned} P(B \leq V) &= P(B - V \leq 0) \\ &= P\left(Z \leq \frac{0 - (-1)}{\sqrt{0.168}}\right) \\ &= P(Z \leq 2.44) \\ &= 0.9927 \end{aligned}$$

9.14 Let Y be the diameter of the eggs. X_i the wholesale of the i -th randomly selected egg, and \bar{X} the average wholesale selling price, $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$. We know that $\bar{X} = \mu = E(X_i)$ for large n .

$$\begin{aligned} E(X_i) &= \sum_{x_i=5}^7 x_i f(x_i) \\ &= 5P(Y < 37) + 6P(37 \leq Y \leq 42) + 7P(Y > 42) \\ &= 6.0292 \end{aligned}$$

Good luck on finals everyone. :)