# Mathematical Statistics 

Gregg Waterman

Oregon Institute of Technology


This work is licensed under the Creative Commons Attribution 4.0 International license. The essence of the license is that

## You are free to:

- Share copy and redistribute the material in any medium or format
- Adapt remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

## Under the following terms:

- Attribution You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

## Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to the web page below.

To view a full copy of this license, visit https://creativecommons.org/licenses/by/4.0/legalcode

## Contents

3 The Binomial and Normal Distributions ..... 57
3.1 The Binomial Distribution ..... 58
3.2 The Standard Normal Distribution ..... 62
3.3 Normal Distributions ..... 64
3.4 Normal Approximation to the Binomial Distribution ..... 66
3.5 Chapter 3 Exercises ..... 68
H Solutions to Exercises ..... 141
H. 3 Chapter 3 Solutions ..... 141

## 3 The Binomial and Normal Distributions

## Performance Criteria:

3. (a) Determine whether an experiment is a Bernoulli process. If it is not, tell why not. If it is, give the probability of a success, when possible.
(b) Correctly denote the binomial distribution function and cumulative binomial distribution function associated with a Bernoulli process. Give, and distinguish between, the parameters and the variable for the distribution.
(c) Compute probabilities using both the binomial probability distribution and cumulative binomial probability distribution.
(d) Compute the expected value and variance for a binomial distribution.
(e) Compute probabilities from the standard normal distribution.
(f) Determine a range of values of the random variable $Z$ of the standard normal distribution having a given probability.
(g) Compute probabilities from a normal distribution.
(h) Determine a range of values of a normal random variable having a given probability.
(i) Use the normal distribution to approximate the binomial probabilities when appropriate.

In the last chapter you studied some simple examples of discrete and continuous probability distributions. Most of those distributions I made up solely for the purposes of learning the concepts of random variables, probability functions, and the expected value and variance of a random variable.

Discrete distributions arise in practice whenever we are counting things, like a the number of heads when flipping a coin a certain number of times, or the number of hits on a web site during a given period of time. Continuous distributions are used when we are dealing with measurements, usually of time or dimensions in one, two or three dimensions.

In this chapter we will look at what are perhaps the most representative examples of discrete and continuous distributions, the binomial distribution and the normal distribution. In reality, each of these is not a single distribution, but a family of distributions that are distinguished from each other by certain parameters.

It is important to distinguish between parameters and variables. In some sense, parameters are variables that vary from situation to situation, but once the specific situation has been determined they become constants. We will say that variables are things that can still vary once a specific situation has been determined. For example, suppose we are flipping a coin a fixed number of times, and we want to discuss the probabilities of getting different numbers of heads. The number of times we will flip the coin is a parameter that will be determined ahead of time. Once the number of flips has been determined, the number of heads can still vary, so it is a variable.

### 3.1 The Binomial Distribution

## Performance Criteria:

3. (a) Determine whether an experiment is a Bernoulli process. If it is not, tell why not. If it is, give the probability of a success, when possible.
(b) Correctly denote the binomial distribution function and cumulative binomial distribution function associated with a Bernoulli process. Give, and distinguish between, the parameters and the variable for the distribution.
(c) Compute probabilities using both the binomial probability distribution and cumulative binomial probability distribution.
(d) Compute the expected value and variance for a binomial distribution.

A Bernoulli process is an experiment satisfying the following conditions:

- The experiment consists of a fixed number $n$ repeated trials, called Bernoulli trials.
- For each trial there are two outcomes, generally referred to as "success" or "failure".
- The probability of success is the same for each trial; we will denote it by p.
- The trials are independent.

1. Determine which of the following experiments are Bernoulli processes; for those that are not, tell why. For those that are, describe what a "success" is, and give the probability of a success if you can.
(a) An urn contains seven blue marbles and ten red marbles. Three marbles are drawn, without replacement. The number of blue marbles drawn is noted.
(b) A die is rolled 5 times, and each time the number on the die is noted.
(c) Three marbles are drawn, with replacement, from the same urn as used for (a). The number of blue marbles drawn is noted.
(d) Every day on your way to school you pass though a particular stoplight. Assume that the light is not set to recognize the approach of vehicles; it simply remains green for a period of 60 seconds, yellow for 5 seconds, then red for 75 seconds, over and over again. (This is in the direction you pass through it.) Assume also that the time at which you approach the light each day, relative to its cycle, is random. For one week (assuming also that you go to school all five of those days) the number of times you encounter a green light there is recorded.
(e) Consider the situation described in (d), but suppose that the light senses the approach of vehicles and changes in some way related to those approaches.
(f) The Acme Company makes Widgets using parts obtained from Fly-By-Night Machining (FBNM). Out of each batch of 1000 parts Acme receives from Fly-By-Night, they test 10 parts (with replacement) and note how many of them are defective.
(g) A coin is flipped repeatedly until heads is obtained. The number of flips it takes to get a head is recorded.

For any Bernoulli experiment there is a random variable $X$, the number of successes. This is called a binomial random variable, and is perhaps the most common discrete random variable for applications. The probability distribution for this random variable is called the binomial distribution. It is actually a family of distributions; there is one such distribution for every value $1,2,3, \ldots$ of $n$ and for each real number $p$ in the interval $[0,1]$. The numbers $n$ and $p$ are the parameters of the distribution. Note again that for a particular Bernoulli process, $n$ and $p$ are FIXED quantities. $x$ is the only variable.

Instead of writing $f(x)$ for the binomial distribution, we write it as $b(x ; n, p)$. The $b$ is the name of the function, $b$ for binomial. $x$ is the variable, and the two symbols after the semi-colon are the parameters for the particular binomial distribution being used.
$\diamond$ Example 3.1(a): Consider the experiment consisting of flipping a coin five times in a row, with the random variable $X$ assigning to each outcome of the number of heads obtained. This experiment is a Bernoulli process -

There are $n=5$ trials, and the probability of success on any trial is $p=\frac{1}{2}$. We would denote the probability distribution by $b(x ; 5,0.5)$ or $b\left(x ; 5, \frac{1}{2}\right)$.
2. Write the notation for the probability distribution associated with any of the experiments from Exercise 1 that ARE Bernoulli processes. If either $n$ or $p$ cannot be determined from the information given, just use those letters for that parameter.

So now it boils down to this: What are the values of the function $b(x ; n, p)$ ? Those values could be determined on an experiment-by-experiment basis. But that is not necessary. Earlier we actually derived the following.

## Binomial Distribution

Suppose that a Bernoulli process has $n$ trials and the probability of success is $p$. The probability density function for the associated Bernoulli random variable is the binomial distribution

$$
b(x ; n, p)=\binom{n}{x} p^{x} q^{n-x}, \quad x=0,1,2, \ldots, n,
$$

where $q=1-p$.

The right way to think of the binomial distribution is not that it is one distribution, but rather a family of distributions, each with a different combination of the two parameters $n$ and $p$. Sometimes we will refer to this as a two-parameter family of distributions. Note that $p$ is the probability of a success, and its exponent $x$ in the probability density function is the number of successes. Similarly, $q$ is the number of failures and its exponent $n-x$ is the number of failures.
$\diamond$ Example 3.1(b): Give the probability distribution for the experiment of flipping a coin five times in a row (see Example 3.1(b)), as both an algebraic expression and as a table of values.

Again, $n=5$ and $p=\frac{1}{2}$, so the probability distribution is

$$
b\left(x ; 5, \frac{1}{2}\right)=\binom{5}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{5-x}=\frac{5!}{x!(5-x)!}\left(\frac{1}{2}\right)^{5}, \quad x=0,1,2, \ldots, 5
$$

Letting $x$ take the integer values from zero to five gives us

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $b\left(x ; 5, \frac{1}{2}\right)$ | $\frac{1}{32}$ | $\frac{5}{32}$ | $\frac{10}{32}$ | $\frac{10}{32}$ | $\frac{5}{32}$ | $\frac{1}{32}$ |

3. (a) Three marbles are drawn, with replacement, from an urn containing seven blue and ten red marbles. The number of blue marbles drawn is noted. (This is the experiment from Exercise 1(c).) Give the probability distribution for the experiment, giving all probabilities in fractional form.
(b) Give the probability distribution for the experiment from Exercise 1(d), again giving all probabilities in fractional form.
4. Refer to Exercise 1(f). If Fly-By-Night Machining claims that 985 out of every 1000 parts that they manufacture are not defective, what is the probability that in a sample of 10 such parts, Acme will find
(a) exactly one defective part?
(b) exactly four defective parts?
(c) at most 1 defective part?
(d) one, two, three or four defective parts?
5. (a) Write out a table of values for $b\left(x ; 3, \frac{1}{5}\right)$, giving all probabilities in exact form.
(b) The cumulative binomial distribution with parameters $n$ and $p$ is denoted by $B(x ; n, p)$. Give the piecewise definition of $B\left(x ; 3, \frac{1}{5}\right)$.
(c) Suppose that this experiment had five trials instead of three. How many successes would you then expect? How many successes would you expect if there were 15 trials? How many would you expect if there were three trials?
(d) Find $\mu$ and $\sigma^{2}$ for this probability distribution. Does the value of $\mu$ you obtain this way agree with your last answer to (c)?

Theorem 3.1: The mean and variance for the binomial distribution $b(x ; n, p)$ are

$$
\mu=n p, \quad \sigma^{2}=n p q
$$

6. Verify that the equation given above for $\sigma^{2}$ gives the correct value for the distribution from Exercise 5(d).
7. Consider a Bernoulli process with 3 trials.
(a) How many successes are possible; that is, what values does the random variable $X$ have?
(b) Write a simplified expression for $b(x ; n, p)$, in terms of $p$ and $q$, for each possible value of $X$.
(c) What is the sum of the expressions you found in (b), and why?
(d) Expand the expression $(p+q)^{3}$. What do you notice?
(e) What is the value of $p+q$, assuming these are the $p$ and $q$ used in the binomial distribution? What must the value of $(p+q)^{3}$ then be?

The figure shown below is Pascal's triangle, which you are probably familiar with. Note how it relates to the above exercise.


### 3.2 The Standard Normal Distribution

## Performance Criteria

3. (e) Compute probabilities from the standard normal distribution.
(f) Determine a range of values of the random variable $Z$ of the standard normal distribution having a given probability.

A very important class of functions in mathematics are those of the form $f(x)=C e^{-k x^{2}}$. When $C=\frac{1}{\sqrt{2 \pi}}$ and $k=\frac{1}{2}$ this function is called the standard normal distribution. Traditionally the letter $Z$ has been used for the continuous random variable for this distribution.

## Standard Normal Distribution

A random variable $Z$ has the standard normal distribution if it has the probability density function

$$
f(z)=\frac{1}{\sqrt{2 \pi}} e^{-\frac{z^{2}}{2}}
$$

This distribution doesn't really describe any phenomena that are seen in the "real world", but it is closely related to a family of distributions called normal distributions, which we will look at in the next section.

You will show later that the expected value of the standard normal distribution is zero, and its variance (and standard deviation) is one. The other normal distributions have other means and variances. A normal distribution with mean $\mu$ and variance $\sigma^{2}$ will be denoted $n(x ; \mu, \sigma)$; like the binomial distribution, the normal distributions are a two-parameter family of distributions. So the standard normal distribution is $n(z ; 0,1)$.

By the definition of a continuous probability distribution,

$$
P(a \leq Z \leq b)=\int_{a}^{b} \frac{1}{\sqrt{2 \pi}} e^{-\frac{z^{2}}{2}} d z .
$$

The problem here is that the function $\frac{1}{\sqrt{2 \pi}} e^{-\frac{z^{2}}{2}}$ can't be integrated analytically, since it has no anti-derivative. Therefore it must be done numerically. This can be done by many calculators and a variety of computer software. Prior to the advent of this technology, people used (or still use!) a table of values of the cumulative distribution $N(z ; 0,1)$, which were of course computed numerically as well. I have provided you with such a table in Appendix B; it gives values of $N(z ; 0,1)$ in increments of 0.01 for the variable $z$. By Theorem 3.5 we then have

$$
P(a \leq Z \leq b)=\int_{a}^{b} \frac{1}{\sqrt{2 \pi}} e^{-\frac{z^{2}}{2}} d z=N(b ; 0,1)-N(a ; 0,1) .
$$

$\diamond$ Example 3.2(a): Use the standard normal distribution table to find $N(1.32 ; 0,1)$.
First we note that $1.32=1.3+0.02$. Now look at the table that has positive numbers down the left hand side of the table and find the row starting with 1.3. Go across the top of the table to the column that has .02 at the top and find the number in that column that is also in the row that you located with the value 1.3. Try this; you should get $N(1.32 ; 0,1)=0.9066$.

1. Find each probability.
(a) $P(Z<-1.86)$
(b) $P(Z \geq 2.26)$
(c) $P(-1.54 \leq Z \leq-0.13)$
(d) $P(-2.09 \leq Z \leq 1.27)$

Sometimes we will know a probability and wish to find a value of the random variable $Z$ from the given probability. In a sense we are reversing the process from the previous example .
$\diamond$ Example 3.2(b): Find a value $z$ such that $P(Z \leq z)=0.27$.
Recalling that $P(Z \leq z)=N(z ; 0,1)$, we look in the body of the standard normal distribution table for the value that is nearest to 0.27 . That number lies between 0.2709 and 0.2676 , with 0.2709 being the closer number. We then go to the left end of the row and the top of the column to find $z$, finding that $z=-0.61$.
2. For each of the following, find the value of $z$ such that
(a) $P(Z \leq z)=0.58$
(b) $P(Z \geq z)=0.71$
(c) $P(-z \leq Z \leq z)=0.37$
(d) $P(0 \leq Z \leq z)=0.42$

Just use the probability in the standard normal distribution table that is closest to the given probability, as in the example above.

### 3.3 Normal Distributions

## Performance Criteria:

3. (g) Compute probabilities from a normal distribution.
(h) Determine a range of values of a normal random variable having a given probability.

Replacing $z$ with $\frac{x-\mu}{\sigma}$ in the standard normal distribution gives us the family of normal distributions, each of which can be shown to have expected value $\mu$ and variance $\sigma$.

## Normal Distribution

A random variable $X$ has the normal distribution if it has the probability density function

$$
\begin{equation*}
n(x ; \mu, \sigma)=\frac{1}{\sigma \sqrt{2 \pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2}} \tag{1}
\end{equation*}
$$

Here $\mu$ and $\sigma$ are parameters, and the distributions in this two-parameter family model probabilities for many applied situations. Sizes and weights of objects are often normally distributed.

When computing probabilities for these distributions we run into the same problem as we did in the previous section, for the standard normal distribution. Howeve $r$, if we have tables or software that allows us to evaluate integrals for the standard normal distribution, we can use those tables for other normal distributions as well. This is because if we make the substitution $z=\frac{x-\mu}{\sigma}$ we get

$$
P(a \leq X \leq b)=P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) .
$$

$\diamond$ Example 3.3(a): Suppose that you know that actual resistances of 10 ohm resistors from a particular manufacturer are normally distributed, with mean 9.97 ohms and standard deviation 0.03 ohms. Find the probability that a randomly selected resistor has resistance between 9.95 and 10.05 ohms.

Letting $X$ represent the random variable that assigns to a randomly selected resistor its resistance, its probability distribution is then $n(x ; 9.97,0.03)$. To find the desired probability, we first compute the values of $Z$ that correspond to $x=9.95$ and $x=10.05$ :

$$
z=\frac{9.95-9.97}{0.03}=-0.67 \quad \text { and } \quad z=\frac{10.05-9.97}{0.03}=2.67 .
$$

(Here we have rounded the $Z$ values to the nearest hundredth because that is the precision of the standard normal distribution tables.) We now have

$$
\begin{aligned}
P(9.95 \leq X \leq 10.05) & =P(-0.67 \leq Z \leq 2.67) \\
& =N(2.67 ; 0,1)-N(-0.67 ; 0,1) \\
& =0.9962-0.2514
\end{aligned}
$$

1. Replacement times for CD players are normally distributed, with a mean of 5.3 years and a standard deviation of 1.7 years.
(a) If you just bought a CD player, what is the probability that you will have to replace it within 6 years?
(b) Find the probability that a CD player will need replacement sometime between 3 and 7 years after it is purchased.
(c) Find the probability that a CD player will last at least 8 years before it needs replacement.
2. The lengths of pregnancies are normally distributed, with a mean of 268 days and a standard deviation of 15 days. A baby is considered to be premature if it is born at least three weeks early. What is the probability that a randomly selected baby will have been premature?
3. Suppose that we want to know how long the most reliable $20 \%$ of all CD players can be expected to last. In other words, find the length of time such that there is a probability of 0.20 that a CD player will last at least that long. Think about trying to solve this without the following hints, then use them if needed.
(a) Sketch a normal distribution and shade an area at the right end that appears to have an area of about 0.20 .
(b) Use the Standard Normal Distribution Table to find the $z$-score that corresponds to that area.
(c) Solve the equation $z=\frac{x-\mu}{\sigma}$ to find the $x$-value that corresponds to the $z$ that you just found. That is the answer!
4. The most unreliable $10 \%$ of CD players will need replacement within at most how many years?
5. The middle $60 \%$ of all CD players will last how many years?
6. Show that the distribution (1) has mean $\mu$. You will need to use the fact that the standard normal distribution has mean zero.

### 3.4 Normal Approximation to the Binomial Distribution

## Performance Criteria:

3. (i) Use the normal distribution to approximate the binomial probabilities when appropriate.

The normal distribution is the "classic" continuous distribution, and the binomial distribution is the classic discrete distribution. In this section you will use Excel to explore a relationship between the two distributions. Of course we can't simply compare values of the binomial distribution with values of the normal distribution, since the normal distribution is a density, and its values do not actually represent probabilities.

1. (a) Consider a Bernoulli process with 10 trials. What are the possible numbers of successes? Put these values in the A column of an Excel spreadsheet, beginning in cell A2. Remember that these are the values $x$ of a random variable $X$.
(b) Suppose that the probability of a success is 0.1 . In column $B$, next to each number of successes in column A, put the probability of that number of successes. It is suggested that you use the built-in binomial distribution function that Excel has.
(c) Label the two columns at their tops, in the empty cells there.
(d) Repeat part (b) for $p=0.3$ and $p=0.5$, putting the results in columns D and F respectively. (Yes, this leaves some empty columns.) Label them appropriately, maybe changing your labeling of column $B$ if necessary.
2. The mean and standard deviation for the binomial distribution are

$$
\mu=n p, \quad \sigma=\sqrt{n p q}
$$

Find the mean and standard deviation for each of the distributions from Exercise 1.
3. We will now consider a normal distribution whose mean and standard deviation are those you computed for the binomial distribution with $p=0.1$. We want to compare this distribution with the corresponding binomial distribution, but we have to consider the following: Since the normal distribution is continuous, the probability at any point is zero, so the probabilities of the values of $X$ that you listed in column A of your spreadsheet will all be zero! Instead, in column C we want $P(x-0.5<X<x+0.5)$ for the value of $x$ in the corresponding cell of column A. Do this; you should not necessarily expect the values in column $C$ to correspond with those in column $B$.
4. Repeat Exercise 3 in columns E and G of your spreadsheet, but for the probabilities $p=0.3$ and $p=0.5, \quad$ respectively.
5. Find $B(4.5 ; 10,0.3)-B(3.5 ; 10,0.3)$ and $N(4.5 ; 3, \sqrt{2.1})-N(3.5 ; 3, \sqrt{2.1})$.
6. Print your spreadsheet. Look at how the binomial and normal distributions compare for all three probability values. Think about it a little. For which of the three binomial distributions does the normal distribution most closely approximate the binomial distribution? Turn in your spreadsheet (with your name on it).
7. Go to Sheet 2 of your worksheet and find the binomial probabilities for a Bernoulli process with $n=50$ and $p=0.1$. Then find the corresponding normal distribution probabilities. How do the values compare in this case?
8. Summarize what you have seen. For what value or values of $p$ does the normal distribution best approximate the binomial distribution? For what value or values of $n$ does the normal distribution best approximate the binomial distribution?

What you should have realized is that the normal distribution is a better approximation of the binomial distribution when $p$ is close to 0.5 and when $n$ is large. A general rule of thumb that some people use is that the normal distribution is a good approximation of the binomial distribution when both

$$
n p \geq 10 \quad \text { and } \quad n(1-p) \geq 10 .
$$

In the past it has been useful to use the normal distribution as an approximation of the binomial distribution when computing probabilities for Bernoulli processes with large numbers of trials. With better computers and software, that is no longer necessary for most applications. There are some applications, however, where approximating a discrete distribution with a continuous distribution is useful.

### 3.5 Chapter 3 Exercises

1. A magazine article states that

Research from the National Highway Traffic Safety Administration shows that up to $80 \%$ of crashes can be attributed to driver inattentiveness.

Assume that whether or not a crash is caused by inattentiveness is a Bernoulli process. With probability as given above. For each of the probabilities asked for, give

- an expression involving the pdf $b$ that gives the desired probability
- an expression involving the cdf $B$ that gives the desired probability
- the probability, as a decimal rounded correctly to four places past the decimal

You can/should use your calculator, Excel or some other assistance to find the last of these. For the first two, use the notation given in the book.
(a) The probability that 7 of ten recent crashes were due to inattentiveness.
(b) The probability that 3 or 4 of five crashes are due to inattentiveness.
(c) The probability that 10 or fewer of 15 crashes are due to inattentiveness.
(d) the probability that 5 or more of eight crashes are due to inattentiveness.
2. In one year there are 427 crashes in a small town. How many of those would we expect to be due to inattentiveness? What concept that we've studied does this illustrate?
3. Suppose that 20 marbles are to be drawn, with replacement, from a top hat (think Cat in the Hat) containing three yellow marbles and two blue marbles. Create a table in Excel with three columns, one for the number $x$ of yellow marbles drawn, one for the appropriate pdf $b$ values and one for the appropriate cdf $B$ values. Label the top of each column with what it is and format the cells with numbers in them so that each probability is given to five places past the decimal.
4. Suppose that $100 \Omega$ (ohm) resistors from a certain manufacturer actually have a mean of $99.92 \Omega$ and standard deviation of $0.17 \Omega$. For each question below following, give each of the following, connected by equal signs:

- probability statement of the form $P($ something about $X)$, followed by
- the expression involving the cumulative distribution $N$ that gives the desired probability, followed by
- the desired probability, to four places past the decimal.

Find the probability that a randomly selected resistor has resistance
(a) over $100 \Omega$
(b) less than $99.7 \Omega$
(c) between 99.8 and $100.2 \Omega$
5. Give a probability statement of the form $P($ something about $Z)$ that is equivalent to your probability statement from part (c) of the previous exercise, followed by the expression involving the cumulative distribution $N$ that gives the desired probability, followed by the corresponding value(s) obtained from the Standard Normal Distribution Table, followed by the final answer.
6. Suppose that five resistors are to be selected from a batch of fifty. Find the probability that two of the five have resistances over $100 \Omega$. Indicate clearly, using appropriate notation, how you obtain your answer.

## H Solutions to Exercises

## H. 3 Chapter 3 Solutions

## Section 3.1

1. (a) Not a Bernoulli process, because the probability for success on each trial is different and because the trials are not independent.
(b) Not a Bernoulli process because there are more than two outcomes.
(c) Bernoulli process, success is drawing a blue and $p=\frac{7}{17}$.
(d) Bernoulli process, success is encountering a green light and $p=\frac{60}{140}=\frac{3}{7}$.
(e) Not a Bernoulli process, then number of trials is not fixed.
(f) Bernoulli process, success is testing a defective part, $p$ is unknown.
2. (b) $b\left(x ; 3, \frac{7}{17}\right)$
(c) $b\left(x ; 5, \frac{3}{7}\right)$
(d) $b(x ; 10, p)$
3. (a) $b\left(x ; 3, \frac{7}{17}\right)=\binom{3}{x}\left(\frac{7}{17}\right)^{x}\left(\frac{10}{17}\right)^{3-x}, \quad x=0,1,2,3$

| $x$ | 0 | 1 | 2 | 3 |
| :---: | :---: | :---: | :---: | :---: |
| $b\left(x ; 3, \frac{7}{17}\right)$ | $\frac{1000}{4913}$ | $\frac{2100}{4913}$ | $\frac{1470}{4913}$ | $\frac{343}{4913}$ |

(b) $b\left(x ; 5, \frac{3}{7}\right)=\binom{5}{x}\left(\frac{3}{7}\right)^{x}\left(\frac{4}{7}\right)^{5-x}, \quad x=0,1,2,3,4,5$

$$
\begin{array}{ccccccc}
x & 0 & 1 & 2 & 3 & 4 & 5 \\
b\left(x ; 5, \frac{4}{7}\right) & \frac{1024}{16807} & \frac{3840}{16807} & \frac{5760}{16807} & \frac{4320}{16807} & \frac{1620}{16807} & \frac{243}{16807}
\end{array}
$$

4. (a) 0.13
(b) $9.7 \times 10^{-6}$
(c) 0.99
(d) 0.14
5. (a) $\begin{array}{ccccc}x & 0 & 1 & 2 & 3 \\ b\left(x ; 3, \frac{1}{5}\right) & \frac{64}{125} & \frac{48}{125} & \frac{12}{125} & \frac{1}{125}\end{array}$
(b)

$$
B\left(x ; 3, \frac{1}{5}\right)=\left\{\begin{array}{cl}
0 & \text { for } x<0 \\
\frac{64}{125} & \text { for } 0 \leq x<1 \\
\frac{112}{125} & \text { for } 1 \leq x<2 \\
\frac{124}{125} & \text { for } 2 \leq x<3 \\
1 & \text { for } x \geq 3
\end{array}\right.
$$

$\begin{array}{ll}\text { (c) } 1,3, \frac{3}{5} & \text { (d) } E(X)=0\left(\frac{64}{125}\right)+1\left(\frac{48}{125}\right)+2\left(\frac{12}{125}\right)+3\left(\frac{1}{125}\right)=\frac{3}{5} \\ \text { (e) } \sigma^{2}=E\left(X^{2}\right)-[E(X)]^{2}=\frac{105}{125}-\frac{9}{25}=\frac{12}{25}\end{array}$
(e) $\sigma^{2}=E\left(X^{2}\right)-[E(X)]^{2}=\frac{105}{125}-\frac{9}{25}=\frac{12}{25}$
6. $\quad \sigma^{2}=n p q=3\left(\frac{1}{5}\right)\left(\frac{4}{5}\right)=\frac{12}{25}$
7. (a) $0,1,2,3$,
(b) $b(0 ; n, p)=q^{3}, b(1 ; n, p)=3 p q^{2}, b(2 ; n, p)=3 p^{2} q, b(3 ; n, p)=p^{3}$
(c) $q^{3}+3 p q^{2}+3 p^{2} q+p^{3}=1$
(d) $(p+q)^{3}=(p+q)\left(p^{2}+2 p q+q^{2}\right)=p^{3}+3 p^{2} q+3 p q^{2}+q^{3}$
(e) $p+q=1$, so $(p+q)^{3}=1$

## Section 3.2:

1. (a) 0.0314
(b) $P(Z \geq 2.26)=1-p(Z \leq 2.26)=1-0.9881=0.0119$
(c) $P(-1.54 \leq Z \leq-0.13)=0.4483-0.0618=0.3865$
(d) $P(-2.09 \leq Z \leq 1.27)=0.8980-0.0183=0.8797$
2. (a) 0.20
(b) -0.55
(c) 0.48
(d) 1.41

## Section 3.3:

1. (a) $Z=\frac{6-5.3}{1.7}=0.41, \quad P(x \leq 6)=P(Z \leq 0.41)=0.6591$
(b) $P(3 \leq X \leq 7)=P(-1.35 \leq Z \leq 1.00)=0.8413-0.0885=0.7528$
(c) $P(X \geq 8)=1-P(X \leq 8)=1-P(Z \leq 1.59)=1-0.9441=0.0554$
2. $P(X \leq 247)=P(Z \leq-1.40)=0.0808$
3. 6.7 years
4. 3.1 years
5. 3.9-6.7 years
