The Binomial Distribution

January 27, 2021

Contents

The Binomial Distribution

When you flip a coin there are only two possible outcomes - heads or tails. This is an example of a **dichotomous** event. Other examples are getting an answer right vs. wrong on a test, catching vs. missing a bus, or eating vs. not eating your vegetables. A roll of a dice, on other hand, is not a dichotomous event since there are six possible outcomes.

If you flip a coin repeatedly, say 10 times, and count up the number of heads, this number is drawn from what's called a **binomial distribution**. Other examples are counting the number of correct answers on an exam, or counting the number of days that your ten year old eats his vegetables at dinner. Importantly, each event has to be independent, so that the outcome of one event does not depend on the outcomes of other events in the sequence.

We can define a binomial distribution with three parameters:

P is the probability of a 'successful' event. That is the event type that you're counting up - like 'heads' or 'correct answers' or 'did eat vegetables'. For a coin flip, P = 0.5. For guessing on a 4-option multiple choice test, P = 1/4 = .25. For my ten year old eating his vegetables, P = 0.05.

N is the number of repeated events.

k is the number of 'successful' events out of N.

The probability of obtaining k successful events out of N, with probability P is:

$$\frac{N!}{k!(N-k)!}P^k(1-P)^{N-k}$$

where N! = $N(N-1)(N-2)...$, or N 'factorial'.

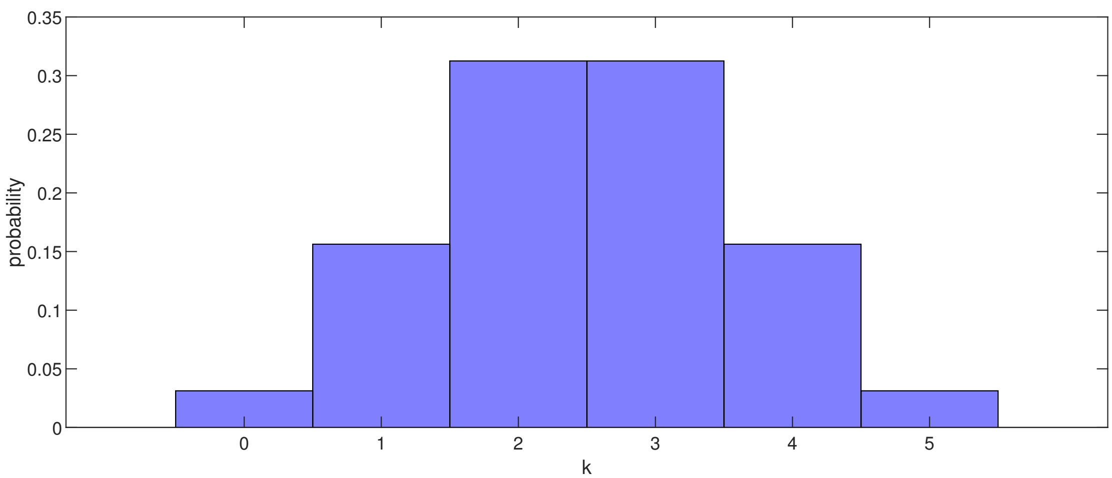
For example, if you flip a fair coin (P=0.5) 5 times, the probability of getting 2 heads is:

$Pr(k=2) = \frac{5!}{2!(5-2)!}(0.5)^2(1-0.5)^{(5-2)} = (10)(0.5^2)(0.5)^3 = 0.3125$

Our textbook, the table handout, and our Excel spreadsheet gives you this number, where the columns are for different values of P and the rows are different values of k:

n	**k**	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.116	0.0778	0.0503	0.0313
	1	0.2036	0.328	0.3915	0.4096	0.3955	0.3601	0.3124	0.2592	0.2059	0.1562
	2	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3364	0.3456	0.3369	0.3125
	3	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.1811	0.2304	0.2757	0.3125
	4	0	0.0005	0.0022	0.0064	0.0146	0.0283	0.0488	0.0768	0.1128	0.1562
	5	0	0	0.0001	0.0003	0.001	0.0024	0.0053	0.0102	0.0185	0.0313

We can plot this binomial frequency distribution as a bar graph:



The shape of the probability distribution for N=5 should look familiar. It looks normal! More on this later.

We just calculated the probability of getting exactly 2 heads out of 5 coin flips. What about the probability of calculating 2 **or more** heads out of 5? It's not hard to see that:

$Pr(k >= 2) = Pr(k=2) + Pr(k=3) + Pr(k=4) + Pr(k=5)$

Using the table we can see that

$Pr(k >= 2) = 0.3125 + 0.3125 + 0.1562 + 0.0313 = 0.8125$

These 'cumulative' binomial problems are common enough that I've provided a page in the Excel spreadsheet and a table in the handout that provides the cumulative binomial probabilities. Here's how to use the cumulative binomial spreadsheet for Pr(k>=2) for N = 5:

n	**k**	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**
5	0	1	1	1	1	1	1	1	1	1	1
	1	0.2262	0.4095	0.5563	0.6723	0.7627	0.8319	0.884	0.9222	0.9497	0.9688
	2	0.0226	0.0815	0.1648	0.2627	0.3672	0.4718	0.5716	0.663	0.7438	0.8125
	3	0.0012	0.0086	0.0266	0.0579	0.1035	0.1631	0.2352	0.3174	0.4069	0.5
	4	0	0.0005	0.0022	0.0067	0.0156	0.0308	0.054	0.087	0.1312	0.1875
	5	0	0	0.0001	0.0003	0.001	0.0024	0.0053	0.0102	0.0185	0.0313

Notice how the first row are all 1's. That's because you will always get zero or more positive events, so $Pr(k >= 0) = 1$.

Example: guessing on an exam

Suppose you're taking a multiple choice exam in PHYS 422, "Contemporary Nuclear and Particle Physics". There are 10 questions, each with 4 options. Assume that you have no idea what's going on and you guess on every question. What is the probability of getting 5 or more answers right?

Since there are 4 options for each multiple choice question, the probability of guessing and getting a single question right is P $= \frac{1}{4} = 0.25$. The probability of getting 5 or more right is $Pr(k >= 5)$. We can find this answer in the cumulative binomial distribution table with N = 10, k = 5 and P = 0.25:

n	**k**	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**
10	0	1	1	1	1	1	1	1	1	1	1
	1	0.4013	0.6513	0.8031	0.8926	0.9437	0.9718	0.9865	0.994	0.9975	0.999
	2	0.0861	0.2639	0.4557	0.6242	0.756	0.8507	0.914	0.9536	0.9767	0.9893
	3	0.0115	0.0702	0.1798	0.3222	0.4744	0.6172	0.7384	0.8327	0.9004	0.9453
	4	0.001	0.0128	0.05	0.1209	0.2241	0.3504	0.4862	0.6177	0.734	0.8281
	5	0.0001	0.0016	0.0099	0.0328	0.0781	0.1503	0.2485	0.3669	0.4956	0.623
	6	0	0.0001	0.0014	0.0064	0.0197	0.0473	0.0949	0.1662	0.2616	0.377
	7	0	0	0.0001	0.0009	0.0035	0.0106	0.026	0.0548	0.102	0.1719
	8	0	0	0	0.0001	0.0004	0.0016	0.0048	0.0123	0.0274	0.0547
	9	0	0	0	0	0	0.0001	0.0005	0.0017	0.0045	0.0107
	10	0	0	0	0	0	0	0	0.0001	0.0003	0.001

So there's about a 8 percent chance of getting 5 or more questions right if you're guessing. Better not take that class.

Example for when $P > 0.5$: Counting wrong answers

Notice that the values of P only go up to 0.5. What is P>0.5? For example, on that physics exam, what is the probability of getting 7 or more **wrong** out of the 10 questions. Now the probability of a 'successful' event is $1 - 0.25 = 0.75$.

The trick to problems with P>0.5 is to turn the problem around so that a 'successful' event has probability of 1-P. For our example, the probability of getting k=7 or more wrong is the same as the probability of getting N-k = 10-7 = 3 or fewer right.

So we can rephrase the problem: What is $Pr(k <= 3)$ with N=10 and P = 0.25?

We can find this using the binomial table spreadsheet:

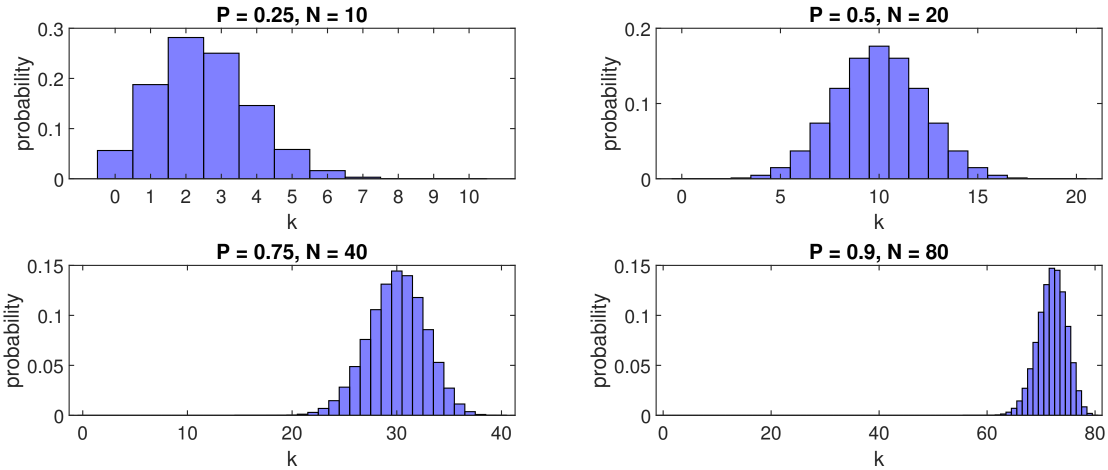
n	**k**	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.006	0.0025	0.001
	1	0.3151	0.3874	0.3474	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
	2	0.0746	0.1937	0.2759	0.302	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
	3	0.0105	0.0574	0.1298	0.2013	0.2503	0.2668	0.2522	0.215	0.1665	0.1172
	4	0.001	0.0112	0.0401	0.0881	0.146	0.2001	0.2377	0.2508	0.2384	0.2051
	5	0.0001	0.0015	0.0085	0.0264	0.0584	0.1029	0.1536	0.2007	0.234	0.2461
	6	0	0.0001	0.0012	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
	7	0	0	0.0001	0.0008	0.0031	0.009	0.0212	0.0425	0.0746	0.1172
	8	0	0	0	0.0001	0.0004	0.0014	0.0043	0.0106	0.0229	0.0439
	9	0	0	0	0	0	0.0001	0.0005	0.0016	0.0042	0.0098
	10	0	0	0	0	0	0	0	0.0001	0.0003	0.001

So the probability of getting 7 or more wrong out of 10 is the same as the probability of getting fewer than 3 right, which is $0.0563 + 0.1877 + 0.2816 + 0.2503 = 0.7759$

The Normal Approximation to the Binomial

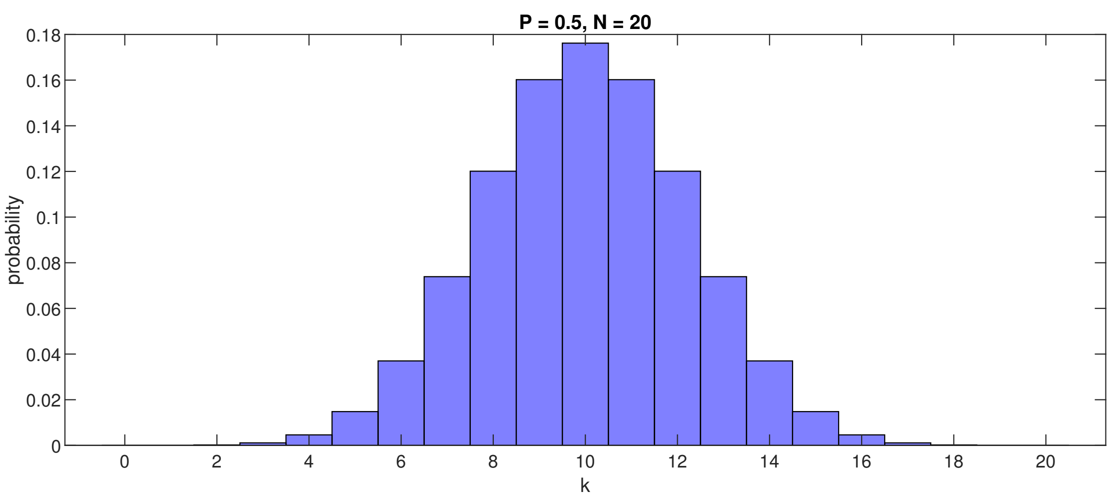
The table in the book goes up to N=15, and the Excel spreadsheet goes up to N=20. But what about higher values of N?

Here are some examples of binomial probability distributions for different values of n and P:



There's that familiar bell curve! It turns out that the discrete binomial probability distribution can be approximated by the continuous normal distribution with a known mean and standard deviation. The binomial distribution becomes more 'normal' with larger values of N and values of P closer to 0.5. A good rule is that the binomial distribution is very close to normal for N>=20.

Let's look more closely at the probability distribution for P = 0.5 and N = 20:



The mean of the normal distribution is intuitive. If you have 20 coin flips, each with probability 0.5, then the average number of heads should be (20)(0.5) = 10. So:

$\mu = NP = 10$

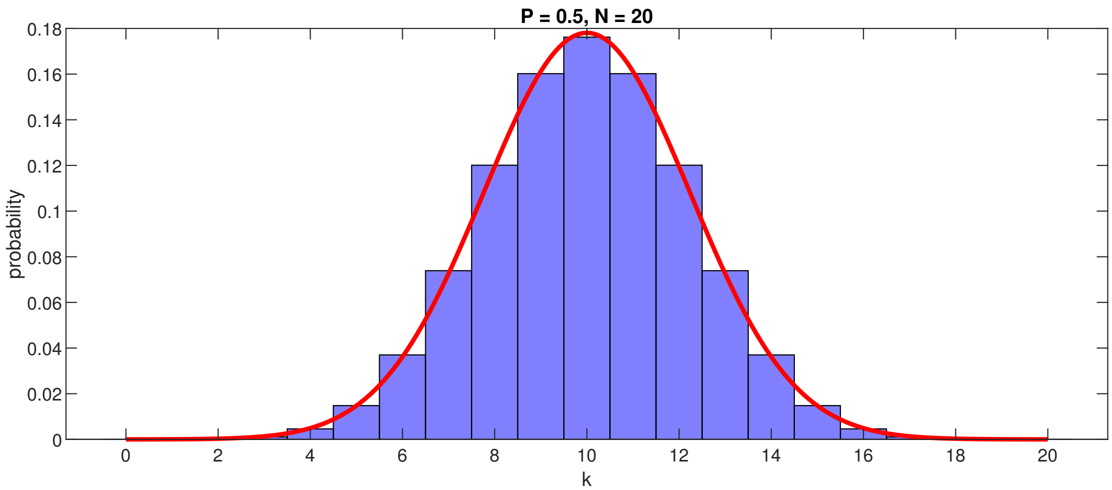
The standard deviation is:

$\sigma = \sqrt{(N)(P)(1-P)}$

Which is for N = 20 is:

$\sqrt{(20)(0.5)(1-0.5)} = 2.24$

Here's that normal distribution drawn on top of the binomial distribution for N=20:

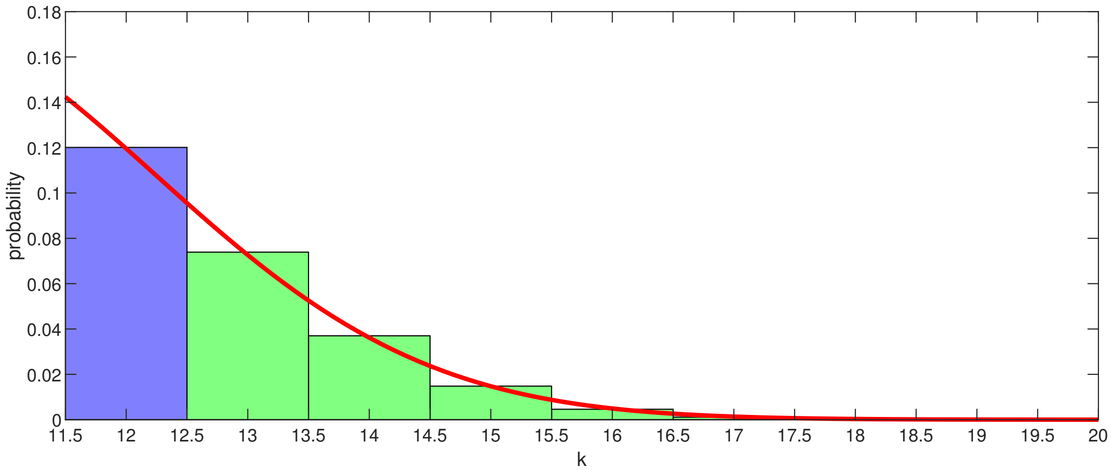


You can see that it's a pretty good fit. The fit gets better with larger values of N, and for values of P that don't get too far away from 1 or -1.

Since the fit is good, we can use the z-table to estimate probabilities from binomial distribution problems for appropriate values of N and P.

For example, if you flip a coin 20 times, what is the probability of obtaining 13 or more heads?

Let's zoom in on the figure above, coloring the events that we are counting in green:



Using the binomial table, the actual probability of obtaining 13 or more heads is:

Pr(k=13) + Pr(k=14) + ... + Pr(k = 20) =

0.0739 + 0.037 + 0.0148 + 0.0046 + 0.0011 + 0.0002 + 0 + 0 = 0.1316

Looking at the figure above, notice that the widths of each bar is 1 unit. The area of each bar (height times width) is therefore equal to the probability of that event. That means that **Pr(k>=13) is equal to the sum of the areas of the green bars**.

So, to approximate the area of the green bars with the normal distribution (the red curve), we need to find the **area under the red curve that covers the same range as the green bars**.

Look closely, the green bar at k = 13 covers the range from 12.5 to 13.5. It follows that to approximate the area of the green bars, we need to **find the area under the normal distribution above k = 12.5**.

Since we know the mean and standard deviation of this normal distribution, we can find the z-score:

$z = \frac{x-\mu}{\sigma} = \frac{12.5-10}{2.24} = 1.12$

Using the z-table:

Pr(z>1.12) = 0.1314

This is pretty close to the actual answer of 0.1316.

Left hander example

Of the 152 students in our class that took the survey, 7 reported themselves as left-handed. If 10% of the population is left-handed, what is the probability that 7 or **fewer** people in will be left handed in a random sample of 152 people?

The normal distribution that best approximates the distribution of left-handed people in a sample size of 152 will have a mean of:

$\mu = NP = (152)(0.1) = 15.2$

The standard deviation of:

$\sigma = \sqrt{(N)(P)(1-P)} = \sqrt{(152)(0.1)(1-0.1)} = 3.6986$

To convert our value of 7 left-handers to a z-score, we need to include the bar that ranges from 6.5 to 7.5. So this time we need add 0.5 to 7 and calculate Pr(x<=7.5).

$z = \frac{x-\mu}{\sigma} = \frac{7.5-15.2}{3.6986} = -2.08$

Pr(z < -2.08) = 0.0188.

So there is a about a 2 percent chance that on any given year we'd have 7 or fewer left-handers out of 152 students if the overall population is 10 percent left-handed.

The Binomial Hypothesis Test

Seahawks example

We can use our knowledge of the binomial distribution to make statistical inferences. For example, in 2019 the Seattle Seahawks football team won 11 games out of 16. Is this a better team than 'average'? Use an α value of 0.05. In other words, is the probability of winning 11 or more games out of 16 less than 0.05 under the null hypothesis that there is a 50/50 chance of winning each game?

To test this, we calculate $Pr(k >= 11)$ for N = 16 and P = 0.5. Since N<20, we'll use the Cumulative Binomial table:

n	k	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**
16	0	1	1	1	1	1	1	1	1	1	1
	1	0.5599	0.8147	0.9257	0.9719	0.99	0.9967	0.999	0.9997	0.9999	1
	2	0.1892	0.4853	0.7161	0.8593	0.9365	0.9739	0.9902	0.9967	0.999	0.9997
	3	0.0429	0.2108	0.4386	0.6482	0.8029	0.9006	0.9549	0.9817	0.9934	0.9979
	4	0.007	0.0684	0.2101	0.4019	0.595	0.7541	0.8661	0.9349	0.9719	0.9894
	5	0.0009	0.017	0.0791	0.2018	0.3698	0.5501	0.7108	0.8334	0.9147	0.9616
	6	0.0001	0.0033	0.0235	0.0817	0.1897	0.3402	0.51	0.6712	0.8024	0.8949
	7	0	0.0005	0.0056	0.0267	0.0796	0.1753	0.3119	0.4728	0.634	0.7728
	8	0	0.0001	0.0011	0.007	0.0271	0.0744	0.1594	0.2839	0.4371	0.5982
	9	0	0	0.0002	0.0015	0.0075	0.0257	0.0671	0.1423	0.2559	0.4018
	10	0	0	0	0.0002	0.0016	0.0071	0.0229	0.0583	0.1241	0.2272
	11	0	0	0	0	0.0003	0.0016	0.0062	0.0191	0.0486	0.1051
	12	0	0	0	0	0	0.0003	0.0013	0.0049	0.0149	0.0384
	13	0	0	0	0	0	0	0.0002	0.0009	0.0035	0.0106
	14	0	0	0	0	0	0	0	0.0001	0.0006	0.0021
	15	0	0	0	0	0	0	0	0	0.0001	0.0003
	16	0	0	0	0	0	0	0	0	0	0

So there's about a 11 percent chance of winning 11 games or more by chance.

Think about this probability, 0.1051. If this number is really small it means one of two things - either the Seahawks are an average team that had a very lucky season, or they are actually a better than average team.

This is the logic of hypothesis testing, which we'll discuss in detail later on this quarter. We've set up a hypothesis about the world, called the 'null hypothesis', which is that the Seahawks are an average team. We then calculate the probability of our observation if we assume that the null hypothesis is true. The lower the probability, the less confident we are about the null hypothesis. Let's conclude an arbitrary criterion for this probability, say 0.1051. If our observed probabilty is lower than 0.05 then we'll consider our results as evidence that the null hypothesis is false.

For our example, since $0.1051 > 0.05$, we cannot conclude that the 2019 Seahawks were a better than average team.
We'll learn later that probabilty that we calculated, 0.1051 is called a 'p-value', and the number we compare it to, 0.05 is called 'alpha', or α.

Mariners example

How about the 2019 Seattle Mariners baseball team? They finished the 2019 with a record of 68 wins and 94 losses. What is the probability of losing 94 or more games?

Since there were N= 162 games we'll have to use the normal approximation to the binomial distribution. With P = 0.5, the number of games that an average team loses (or wins) will be distributed approximately normally with a mean of:

$\mu = NP = (162)(0.5) = 81$

and a standard deviation of:

$\sigma = \sqrt{(162)(0.5)(1-0.5)} = 6.364$

Since we're finding the probability of 94 **or more** losses, we have to **subtract 0.5** from 94 when calculating the z-score. (Note that if we wanted to find the probability of 94 **or fewer** losses we'd **add 0.5** to 94 when calculating the z-score). So, for 94 or more losses, the z-score for losing 94 or more games is:

$z = \frac{(93.5-81)}{6.364} = 1.96$

The probability of losing 94 or more games is therefore:

$Pr(k >= 94) = Pr(z > 1.96) = 0.025$

Since $0.025 < 0.05$, we conclude that yes, the 2019 Mariners were a worse than average team.

Computing Binomial Probabilities in R

Computing binomial probabilities is really easy in R - much easier than using the table. The function 'binom.test' does everything for you. Note, the test always finds the 'exact' answer, rather than the normal approximation, even if the number of trials is greater than 20.

The R commands shown below can be found here: BinomialDistribution.R

```
# BinomialDistribution.R

# Calculating binomial probabilities is easy in R.  The function 'binom.test' takes in
# three variables: (1) K, the number of successful outcomes, (2) N, the total number of trials,
# and (3) P, the probability of a successful outcome on any given trial.
#
# A fourth argument can be 'alternative = "less"', or 'alternative = "greater"', depending on
# whether you want Pr(x<k) or Pr(x>k)
#
# If you just want a binomial probability, that is, the probability of getting
# exactly k successful events out of N tries, you can use the function 'dbinom'
# where the three values go in the same order as 'binom.test'.
```

```
# Example: Given 10 flips of a fair 50/50 coin, what is the probability of
# getting exactly 6 heads?
dbinom(6,10,.5)
[1] 0.2050781

# Example: Given 10 flips of a fair 50/50 coin, what is the probability of obtaining 6 or more he
out <- binom.test(6,10,.5,
           alternative = "greater")
out <- binom.test(5,10,.5,
                  alternative = "less")
# The result can be found in the field 'p.value':
print(out$p.value)
[1] 0.6230469

# Example: If you guess on a 20 question multiple choice test where each question has 5 possible
# answers, what is the probability of getting 4 or less correct?
out <- binom.test(4,20,1/5,
                  alternative = "less")
print(out$p.value)
[1] 0.6296483

# Example: If a basketball player has a 2 out 3 chance of making a free throw on any given try,
# and all tries are independent, what is the probability of making 7 or more out of 10?
out <- binom.test(7,10,2/3,
                  alternative = "greater")
print(out$p.value)
[1] 0.5592643

# This is the same as:
out <- binom.test(10-7,10,1-2/3,alternative = "less")
print(out$p.value)
[1] 0.5592643

# what do binomial distributions look like?
N <- 20
k <- seq(0,N)  # 'sequence' of numbers counting from 0 to N
P <- .5
p <- dbinom(k,N,P)

barplot(p)
# To make the plot prettier, and to add labels:
barplot(height = p,                   # probabilities
        names = k,                    # x axis values
        col = 'blue',                 # color
        xlab = 'k',                   # x label
        ylab = 'Pr(k)',               # y label
        main = sprintf('N = %d',N),   # title
        space = 0)                    # spacing between bars
```

30 Problems

It's problem time. The following 30 questions are all binomial distribution problems. Use the binomial table or cumulative binomial table if N is less than or equal to 20, and use the normal approximation to the binomial if N is greater than 20.

1) For P = 0.05 and N = 14, find $Pr(k >= 2)$

2) For P = 0.85 and N = 32, find $Pr(k <= 29)$

3) For P = 0.75 and N = 26, find $Pr(k <= 19)$

4) For P = 0.5 and N = 7, find $Pr(k >= 4)$

5) For P = 0.35 and N = 5, find $Pr(k <= 1)$

6) For P = 0.65 and N = 22, find $Pr(k >= 21)$

7) For P = 0.8 and N = 42, find $Pr(k >= 33)$

8) For P = 0.6 and N = 9, find $Pr(k <= 8)$

9) For P = 0.85 and N = 33, find $Pr(k >= 29)$

10) For P = 0.05 and N = 15, find $Pr(k >= 2)$

11) For P = 0.6 and N = 18, find $Pr(k >= 4)$

12) For P = 0.35 and N = 45, find $Pr(k <= 11)$

13) For P = 0.7 and N = 44, find $Pr(k >= 38)$

14) For P = 0.1 and N = 31, find $Pr(k >= 1)$

15) For P = 0.25 and N = 40, find $Pr(k <= 13)$

16) For P = 0.45 and N = 36, find $Pr(k <= 14)$

17) For P = 0.35 and N = 16, find $Pr(k <= 5)$

18) For P = 0.85 and N = 24, find $Pr(k >= 15)$

19) For P = 0.85 and N = 7, find $Pr(k >= 3)$

20) For P = 0.3 and N = 26, find $Pr(k >= 12)$

21) For P = 0.25 and N = 25, find $Pr(k <= 6)$

22) For P = 0.1 and N = 43, find $Pr(k >= 1)$

23) For P = 0.25 and N = 20, find $Pr(k >= 2)$

24) For P = 0.45 and N = 26, find $Pr(k >= 16)$

25) For P = 0.7 and N = 7, find $Pr(k >= 4)$

26) For P = 0.15 and N = 38, find $Pr(k <= 7)$

27) For P = 0.8 and N = 33, find $Pr(k >= 28)$

28) For P = 0.7 and N = 34, find $Pr(k >= 27)$

29) For P = 0.25 and N = 41, find $Pr(k <= 12)$

30) For P = 0.5 and N = 19, find $Pr(k >= 3)$

30 Answers

1) For P = 0.05 and N = 14, find $Pr(k >= 2)$
Since $N <= 20$ use the binomial table.
$Pr(k >= 2) = 0.1229 + 0.0259 + ... + 0 = 0.1529$

Using R:
```
out<-binom.test(2,14,0.05,alternative = "greater")
print(out$p.value)
[1] 0.1529856
```

2) For P = 0.85 and N = 32, find $Pr(k <= 29)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (32)(0.85) = 27.2$
$\sigma = \sqrt{(32)(0.85)(1 - 0.85)} = 2.0199$
$z = \frac{(29.5-27.2)}{2.0199} = 1.14$
$Pr(k <= 29.5) = Pr(z <= 1.14) = 0.8729$

Using R we'll run the 'exact' test:
```
out<-binom.test(29,32,0.85,alternative = "less")
print(out$p.value)
[1] 0.878194
```

3) For P = 0.75 and N = 26, find $Pr(k <= 19)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (26)(0.75) = 19.5$
$\sigma = \sqrt{(26)(0.75)(1 - 0.75)} = 2.2079$
$z = \frac{(19.5-19.5)}{2.2079} = 0$
$Pr(k <= 19.5) = Pr(z <= 0) = 0.5$

Using R we'll run the 'exact' test:
```
out<-binom.test(19,26,0.75,alternative = "less")
print(out$p.value)
[1] 0.4846068
```

4) For P = 0.5 and N = 7, find $Pr(k >= 4)$

Since $N <= 20$ use the binomial table.
$Pr(k >= 4) = 0.2734 + 0.1641 + 0.0547 + 0.0078 = 0.5$

Using R:
```
out<-binom.test(4,7,0.5,alternative = "greater")
print(out$p.value)
[1] 0.5
```

5) For P = 0.35 and N = 5, find $Pr(k <= 1)$
Since $N <= 20$ use the binomial table.
$Pr(k <= 1) = 0.116 + 0.3124 = 0.4284$

Using R:
```
out<-binom.test(1,5,0.35,alternative = "less")
print(out$p.value)
[1] 0.428415
```

6) For P = 0.65 and N = 22, find $Pr(k >= 21)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (22)(0.65) = 14.3$
$\sigma = \sqrt{(22)(0.65)(1-0.65)} = 2.2372$
$z = \frac{(20.5-14.3)}{2.2372} = 2.77$
$Pr(k >= 20.5) = Pr(z >= 2.77) = 0.0028$

Using R we'll run the 'exact' test:
```
out<-binom.test(21,22,0.65,alternative = "greater")
print(out$p.value)
[1] 0.0009837097
```

7) For P = 0.8 and N = 42, find $Pr(k >= 33)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (42)(0.8) = 33.6$
$\sigma = \sqrt{(42)(0.8)(1-0.8)} = 2.5923$
$z = \frac{(32.5-33.6)}{2.5923} = -0.42$
$Pr(k >= 32.5) = Pr(z >= -0.42) = 0.6628$

Using R we'll run the 'exact' test:

```
out<-binom.test(33,42,0.8,alternative = "greater")
print(out$p.value)
[1] 0.6756239
```

8) For P = 0.6 and N = 9, find $Pr(k <= 8)$
Since $N <= 20$ use the binomial table.
With $P > 0.5$ we need to switch the problem to
P = 1-0.6 = 0.4, N = 9, $Pr(k >= 9 - 8) = Pr(k >= 1)$
$Pr(k >= 1) = 0.0605 + 0.1612 + ... + 0.0003 = 0.9898$

Using R:

```
out<-binom.test(8,9,0.6,alternative = "less")
print(out$p.value)
[1] 0.9899223
```

9) For P = 0.85 and N = 33, find $Pr(k >= 29)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (33)(0.85) = 28.05$
$\sigma = \sqrt{(33)(0.85)(1 - 0.85)} = 2.0512$
$z = \frac{(28.5-28.05)}{2.0512} = 0.22$
$Pr(k >= 28.5) = Pr(z >= 0.22) = 0.4129$

Using R we'll run the 'exact' test:

```
out<-binom.test(29,33,0.85,alternative = "greater")
print(out$p.value)
[1] 0.4355177
```

10) For P = 0.05 and N = 15, find $Pr(k >= 2)$
Since $N <= 20$ use the binomial table.
$Pr(k >= 2) = 0.1348 + 0.0307 + ... + 0 = 0.171$

Using R:

```
out<-binom.test(2,15,0.05,alternative = "greater")
print(out$p.value)
[1] 0.1709525
```

11) For P = 0.6 and N = 18, find $Pr(k >= 4)$
Since $N <= 20$ use the binomial table.
With $P > 0.5$ we need to switch the problem to
P = 1-0.6 = 0.4, N = 18, $Pr(k <= 18 - 4) = Pr(k <= 14)$
$Pr(k <= 14) = 0.0001 + 0.0012 + ... + 0.0011 = 0.9999$

Using R:

```
out<-binom.test(4,18,0.6,alternative = "greater")
print(out$p.value)
[1] 0.9997852
```

12) For P = 0.35 and N = 45, find $Pr(k <= 11)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (45)(0.35) = 15.75$
$\sigma = \sqrt{(45)(0.35)(1 - 0.35)} = 3.1996$
$z = \frac{(11.5-15.75)}{3.1996} = -1.33$
$Pr(k <= 11.5) = Pr(z <= -1.33) = 0.0918$

Using R we'll run the 'exact' test:

```
out<-binom.test(11,45,0.35,alternative = "less")
print(out$p.value)
[1] 0.08955629
```

13) For P = 0.7 and N = 44, find $Pr(k >= 38)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (44)(0.7) = 30.8$
$\sigma = \sqrt{(44)(0.7)(1 - 0.7)} = 3.0397$
$z = \frac{(37.5-30.8)}{3.0397} = 2.2$
$Pr(k >= 37.5) = Pr(z >= 2.2) = 0.0139$

Using R we'll run the 'exact' test:

```
out<-binom.test(38,44,0.7,alternative = "greater")
print(out$p.value)
[1] 0.009975843
```

14) For P = 0.1 and N = 31, find $Pr(k >= 1)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (31)(0.1) = 3.1$
$\sigma = \sqrt{(31)(0.1)(1-0.1)} = 1.6703$
$z = \frac{(0.5-3.1)}{1.6703} = -1.56$
$Pr(k >= 0.5) = Pr(z >= -1.56) = 0.9406$

Using R we'll run the 'exact' test:

```
out<-binom.test(1,31,0.1,alternative = "greater")
print(out$p.value)
[1] 0.961848
```

15) For P = 0.25 and N = 40, find $Pr(k <= 13)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (40)(0.25) = 10$
$\sigma = \sqrt{(40)(0.25)(1-0.25)} = 2.7386$
$z = \frac{(13.5-10)}{2.7386} = 1.28$
$Pr(k <= 13.5) = Pr(z <= 1.28) = 0.8997$

Using R we'll run the 'exact' test:

```
out<-binom.test(13,40,0.25,alternative = "less")
print(out$p.value)
[1] 0.8967683
```

16) For P = 0.45 and N = 36, find $Pr(k <= 14)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (36)(0.45) = 16.2$
$\sigma = \sqrt{(36)(0.45)(1-0.45)} = 2.985$
$z = \frac{(14.5-16.2)}{2.985} = -0.57$
$Pr(k <= 14.5) = Pr(z <= -0.57) = 0.2843$

Using R we'll run the 'exact' test:

```
out<-binom.test(14,36,0.45,alternative = "less")
print(out$p.value)
[1] 0.2861312
```

17) For P = 0.35 and N = 16, find $Pr(k <= 5)$
Since $N <= 20$ use the binomial table.
$Pr(k <= 5) = 0.001 + 0.0087 + 0.0353 + 0.0888 + 0.1553 + 0.2008 = 0.4899$

Using R:

```
out<-binom.test(5,16,0.35,alternative = "less")
print(out$p.value)
[1] 0.4899636
```

18) For P = 0.85 and N = 24, find $Pr(k >= 15)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (24)(0.85) = 20.4$
$\sigma = \sqrt{(24)(0.85)(1-0.85)} = 1.7493$
$z = \frac{(14.5-20.4)}{1.7493} = -3.37$
$Pr(k >= 14.5) = Pr(z >= -3.37) = 0.9996$

Using R we'll run the 'exact' test:

```
out<-binom.test(15,24,0.85,alternative = "greater")
print(out$p.value)
[1] 0.9985174
```

19) For P = 0.85 and N = 7, find $Pr(k >= 3)$
Since $N <= 20$ use the binomial table.
With $P > 0.5$ we need to switch the problem to
P = 1-0.85 = 0.15, N = 7, $Pr(k <= 7-3) = Pr(k <= 4)$
$Pr(k <= 4) = 0.3206 + 0.396 + 0.2097 + 0.0617 + 0.0109 = 0.9989$

Using R:

```
out<-binom.test(3,7,0.85,alternative = "greater")
print(out$p.value)
[1] 0.9987784
```

20) For P = 0.3 and N = 26, find $Pr(k >= 12)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (26)(0.3) = 7.8$
$\sigma = \sqrt{(26)(0.3)(1-0.3)} = 2.3367$
$z = \frac{(11.5-7.8)}{2.3367} = 1.58$
$Pr(k >= 11.5) = Pr(z >= 1.58) = 0.0571$

Using R we'll run the 'exact' test:

```
out<-binom.test(12,26,0.3,alternative = "greater")
print(out$p.value)
[1] 0.06031255
```

21) For P = 0.25 and N = 25, find $Pr(k <= 6)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (25)(0.25) = 6.25$
$\sigma = \sqrt{(25)(0.25)(1-0.25)} = 2.1651$
$z = \frac{(6.5-6.25)}{2.1651} = 0.12$
$Pr(k <= 6.5) = Pr(z <= 0.12) = 0.5478$

Using R we'll run the 'exact' test:

```
out<-binom.test(6,25,0.25,alternative = "less")
print(out$p.value)
[1] 0.5610981
```

22) For P = 0.1 and N = 43, find $Pr(k >= 1)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (43)(0.1) = 4.3$
$\sigma = \sqrt{(43)(0.1)(1-0.1)} = 1.9672$
$z = \frac{(0.5-4.3)}{1.9672} = -1.93$
$Pr(k >= 0.5) = Pr(z >= -1.93) = 0.9732$

Using R we'll run the 'exact' test:

```
out<-binom.test(1,43,0.1,alternative = "greater")
print(out$p.value)
[1] 0.9892247
```

23) For P = 0.25 and N = 20, find $Pr(k >= 2)$
Since $N <= 20$ use the binomial table.
$Pr(k >= 2) = 0.0669 + 0.1339 + ... + 0 = 0.9757$

Using R:

```
out<-binom.test(2,20,0.25,alternative = "greater")
print(out$p.value)
[1] 0.9756874
```

24) For P = 0.45 and N = 26, find $Pr(k >= 16)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (26)(0.45) = 11.7$
$\sigma = \sqrt{(26)(0.45)(1 - 0.45)} = 2.5367$
$z = \frac{(15.5 - 11.7)}{2.5367} = 1.5$
$Pr(k >= 15.5) = Pr(z >= 1.5) = 0.0668$

Using R we'll run the 'exact' test:

```
out<-binom.test(16,26,0.45,alternative = "greater")
print(out$p.value)
[1] 0.06737017
```

25) For P = 0.7 and N = 7, find $Pr(k >= 4)$
Since $N <= 20$ use the binomial table.
With $P > 0.5$ we need to switch the problem to
P = 1-0.7 = 0.3, N = 7, $Pr(k <= 7 - 4) = Pr(k <= 3)$
$Pr(k <= 3) = 0.0824 + 0.2471 + 0.3177 + 0.2269 = 0.8741$

Using R:

```
out<-binom.test(4,7,0.7,alternative = "greater")
print(out$p.value)
[1] 0.873964
```

26) For P = 0.15 and N = 38, find $Pr(k <= 7)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (38)(0.15) = 5.7$
$\sigma = \sqrt{(38)(0.15)(1 - 0.15)} = 2.2011$
$z = \frac{(7.5 - 5.7)}{2.2011} = 0.82$
$Pr(k <= 7.5) = Pr(z <= 0.82) = 0.7939$

Using R we'll run the 'exact' test:

```
out<-binom.test(7,38,0.15,alternative = "less")
print(out$p.value)
[1] 0.7986452
```

27) For P = 0.8 and N = 33, find $Pr(k >= 28)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (33)(0.8) = 26.4$
$\sigma = \sqrt{(33)(0.8)(1 - 0.8)} = 2.2978$
$z = \frac{(27.5 - 26.4)}{2.2978} = 0.48$
$Pr(k >= 27.5) = Pr(z >= 0.48) = 0.3156$

Using R we'll run the 'exact' test:

```
out<-binom.test(28,33,0.8,alternative = "greater")
print(out$p.value)
[1] 0.3290296
```

28) For P = 0.7 and N = 34, find $Pr(k >= 27)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (34)(0.7) = 23.8$
$\sigma = \sqrt{(34)(0.7)(1 - 0.7)} = 2.6721$
$z = \frac{(26.5 - 23.8)}{2.6721} = 1.01$
$Pr(k >= 26.5) = Pr(z >= 1.01) = 0.1562$

Using R we'll run the 'exact' test:

```
out<-binom.test(27,34,0.7,alternative = "greater")
print(out$p.value)
[1] 0.1558404
```

29) For P = 0.25 and N = 41, find $Pr(k <= 12)$

Since $N > 20$ use the normal approximation and the z-table.
k will be distributed normally with:

$\mu = NP = (41)(0.25) = 10.25$
$\sigma = \sqrt{(41)(0.25)(1-0.25)} = 2.7726$
$z = \frac{(12.5-10.25)}{2.7726} = 0.81$
$Pr(k <= 12.5) = Pr(z <= 0.81) = 0.791$

Using R we'll run the 'exact' test:

```
out<-binom.test(12,41,0.25,alternative = "less")
print(out$p.value)
[1] 0.7944354
```

30) For P = 0.5 and N = 19, find $Pr(k >= 3)$
Since $N <= 20$ use the binomial table.
$Pr(k >= 3) = 0.0018 + 0.0074 + ... + 0 = 0.9997$

Using R:

```
out<-binom.test(3,19,0.5,alternative = "greater")
print(out$p.value)
[1] 0.9996357
```