The Binomial Distribution

Paul Johnson

June 2013

1 Binomial Distribution

1.1 Bernoulli model for a single coin flip

If someone conducts a coin flip to decide "Yes" or "No", she is conducting an exercise that simulates a "Bernoulli process." Obviously, if the chance of a "Yes" (or "True" or "Success" or whatever you call it) is π and we code the outcomes 1 and 0 (for Yes and No), then the Bernoulli variable is very easy to understand. It has an expected value of

$$E[x_i] = \pi \cdot 1 + (1-\pi) \cdot 0 = \pi \tag{1}$$

and variance

$$\begin{aligned} V[x_i] &= \pi(1-\pi)^2 - (1-\pi)(0-\pi^2) \\ &= \pi(1-\pi) \end{aligned} \tag{2}$$

1.2 Binomial Terminology

N **observations** (or "tests" or "trials") of a random process that can give only 2 possible answers, such as "1" and "0", "yes" or "no", and "success" or "failure."

k **successes** Out of N observations, suppose there are k "successes" (and, obviously, $N-k$ "failures).

π **probability of success fixed** The chance of each outcomes is fixed across all experiments. Let π represent the chance of a "success" (or "heads" or "yes") and $(1-\pi)$ is the chance of "failure" (or "tails" or "no").

The paradigmatic example of a binomial distribution would be a series of coin flips. There are 2 possible outcomes, "heads" or "tails", and the chance of "heads" is fixed.

The binomial distribution represents the number of "successes" that will be observed in N experiments. The set of possible outcomes is thus

$$X = \{0, 1, 2, 3, \ldots, N\} \tag{3}$$

The binomial distribution describes the chances of observing k successes out of N trials, with the probability of success fixed at π.

$$Prob(k|N, \pi) \tag{4}$$

1.3 Probability Mass Function

The Binomial probability mass function is:

$$Prob(k|N,\pi) = \frac{N!}{(N-k)!k!}\pi^k(1-\pi)^{N-k} \tag{5}$$

Suppose the chance of having a boy baby is 0.63 for all women in a community. If 437 women have babies, what is the probability that there will be 200 boys?

Inserting N and π into the previous expression, the chance of k successes is seen to be:

$$Prob(k|437, 0.63) = \frac{437!}{(437-k)!k!}(0.63)^k(1-\pi)^{437-k} \quad (6)$$

For $k = 200$, the probability is:

$$P(200|437, 0.63) = 7.626893 \times 10^{-34} \quad (7)$$

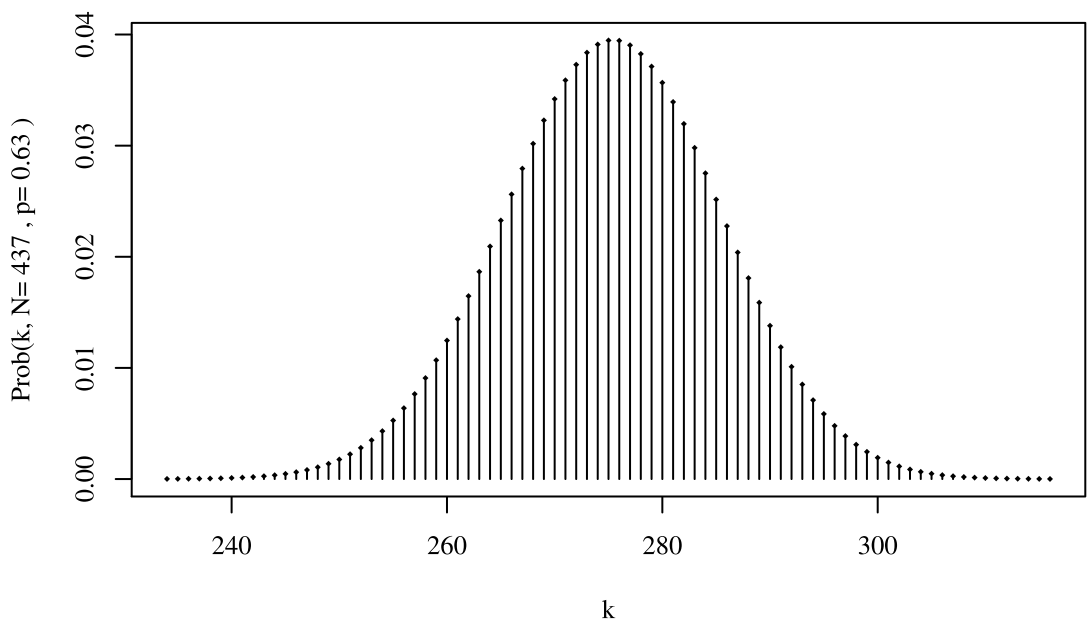
That's a very small number (Recall, $10^{-34} = 1/10^{34}$).

If we had asked for the probability of 300 boys, we would find:

$$P(300|437, 0.63) = 0.0001122501 \quad (8)$$

I've done some "hunting and pecking" with this distribution to find out which values of k are most likely. The outcomes with noticable chances are between 240 and 310, as indicated in Figure 1.

Figure 1: Binomial with N=437 and p=0.63



1.3.1 Aside: The "Binomial Coefficient"

The formula for the probability of k successes is often presented using "N choose k" notation.

$$Prob(k|N, \pi) = \binom{N}{k} \pi^k(1-\pi)^{N-k} \quad (9)$$

$\binom{N}{k}$ is the number of different ways to get k "successes" and $N - k$ failures out of N tests. It is called "the binomial coefficient" because it plays a part in the "binomial formula" that is used in algebra.

$$\binom{N}{k} = \frac{N!}{(N-k)!k!} \quad (10)$$

2 Central Tendency and Dispersion

Suppose $x \sim Binomial(N, \pi)$.

2.1 Expected Value.

The expected value is:

$$E(x) = \pi \cdot N \tag{11}$$

It seems obvious to me that this is correct. If we flip a coin 10 times and the chance of a $Head$ is π, it seems reasonable to expect $\pi \cdot 10$ $Heads$.

There is a simple way to demonstrate that. And there's also the hard way.

Let's take the easy way first. Think of the outcome, the number of successes, as a sum of 0's and 1's. For instance, the observed sample:

$$0, 1, 1, 0, 1, 1, 0, 0 \ldots, 1, 0 \tag{12}$$

is really just a realization of Bernoulli trials, and the number of successes is just the sum of those trials, as in

$$x_1 + x_2 + x_3 + \ldots + x_{N-1} + x_N \tag{13}$$

Those are "statistically independent" samples of size 1, and each one has probability of success equal to π. So, considering just one "event" in isolation, the chance is π of observing a 1 and $(1-\pi)$ chance of observing 0. So the expected value of that one draw is

$$E[x_1] = \pi \cdot 1 + (1-\pi) \cdot 0 = \pi \tag{14}$$

So of you think of the Binomial as the sum of N of those experiments,

$$\begin{aligned} E[x_1 + x_2 + \ldots x_N] &= E[x_1] + E[x_2] + \ldots + E[x_n] \\ &= \pi + \pi + \ldots + \pi \\ &= N \cdot \pi \end{aligned} \tag{15}$$

Now the "hard way." Recall the $E(x)$ is defined as the probability weighted sum of outcomes:

$$E(x) = \sum_{i=0}^{N} Prob(X_i = x_i | N, \pi) x_i.$$

Inserting the Binomial probability model

$$E(x) = \sum_{i=0}^{N} \binom{N}{x_i} \pi^{x_i} (1-\pi)^{N-x_i} x_i = \sum_{i=0}^{N} \frac{N!}{x_i!(N-x_i)!} \pi^{x_i} (1-\pi)^{N-x_i} x_i \tag{16}$$

In case you wonder how that can be simplified to $\pi \cdot N$, you can read the proof in the Wikipedia: `http://en.wikipedia.org/wiki/Binomial_distribution`

2.2 Variance

And the variance is:

$$Var(x) = \pi(1-\pi)N \tag{17}$$

If you take the easy route, consider just one draw, x_1, in isolation. Its variance is

$$\begin{aligned} Var[x_1] &= \pi(1-E[x_1])^2 + (1-\pi)(0-E[x_1])^2 \\ &= \pi(1-\pi)^2 - (1-\pi)(-\pi)^2 \\ &= \pi(1-2\pi+\pi^2) + \pi^2 - \pi^3 \\ &= \pi - 2\pi^2 + \pi^3 + \pi^2 - \pi^3 \\ &= \pi - \pi^2 = \pi(1-\pi) \end{aligned} \tag{18}$$

The Binomial distribution is a sum of N of those variables, and they are all statistically independent of each other. Thus, the law for calculating the variance of a sum of terms applies.

$$\begin{aligned} Var[x_1 + x_2 + \ldots x_N] &= Var[x_1] + Var[x_2] + \ldots + Var[x_N] \\ &\quad + \{a\, lot\, of\, Covariances\, between\, x_i\, and\, x_j\} \end{aligned} \tag{19}$$

All the covariances are 0, because all of the draws are statistically independent. Hence the problem is solved.

$$\begin{aligned} Var[x_1 + x_2 + \ldots x_N] &= \pi(1-\pi) + \pi(1-\pi) + \ldots + \pi(1-\pi) \\ &= \pi(1-\pi)N \end{aligned}$$

3 Is the Binomial Approximately Normal?

The Binomial distribution is a discrete distribution. The number of successes can only take on integer values. Thus, OBVIOUSLY, it is never going to be exactly Normal, since the normal is defined on a continuum.

Recall, the Central Limit Theorem states that the mean of any variable tends to be normally distributed, as the size of the sample tends to infinity. In section 2.1, I asked the reader to consider the Binomial as the sum of N variables, each of which is 0 or 1. Reasoning from the CLT, we should expect the distribution of the Binomial will tend toward Normality. That is indeed the case, as we shall now see with some illustrations.

If N is "pretty big" and if π is in the "middle range," then the distribution of the Binomial appears to be rather similar to a Normal distribution. Consider a very-Normal looking case, a large sample of of 2000 draws for which the success on each is 0.50 in Figure 2.

On the contrary, when the sample is small, the discreteness of the observed values is more stark and the appearance is not all that Normal. The probability of outcomes in Bin(4, 0.50) is illustrated in Figure 3.

When the number of draws is small, the appearance is decidedly not Normal when the probability of success is small (or large). Consider the case in which the probability is 0.15 in Figure 4.

However, as the CLT would lead us to expect, the observed Binomial outcomes are quite a bit more Normal in appearance when the sample is large. A model in which there are 2000 observations with probability of success 0.15 has an expected value of 300 and a standard deviation of 15 (calculated as $\sqrt{0.15(1-0.15)/N}$). See Figure 5.

In fact, even when the probability of success on one trial is very small, say just 0.01, the distribution of the observed number of successes is rather symmetric and unimodal, as illustrated in Figure .

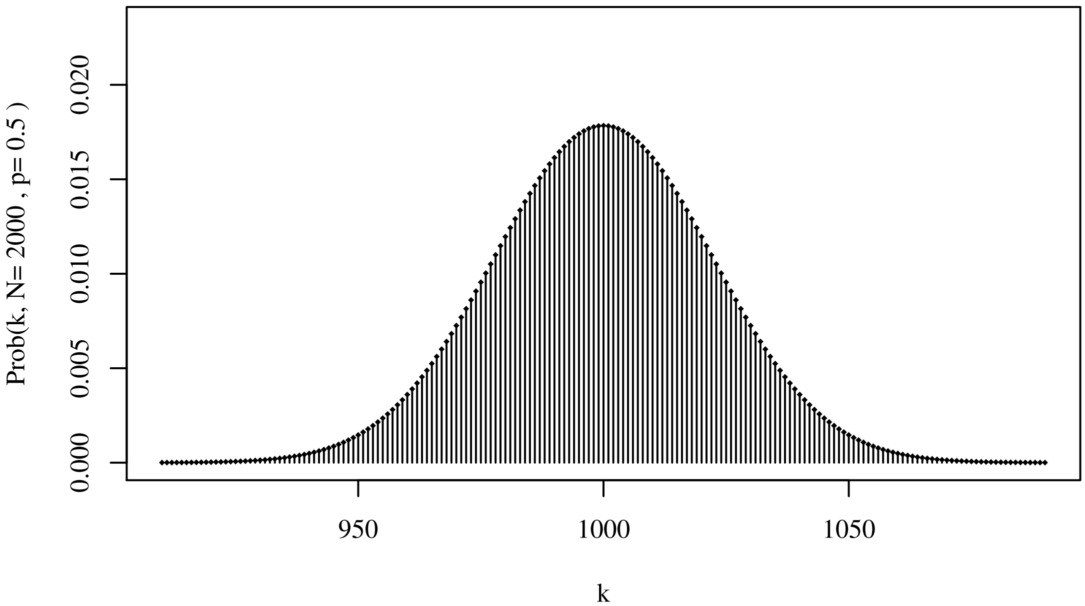
4 Derivation of the Binomial Probability Formula

Over the years, I have worked really hard to develop an explanation for the probability mass function.

$$Prob(k|N,\pi) = \begin{pmatrix} N \\ k \end{pmatrix} \pi^k (1-\pi)^{N-k}$$

Perhaps you are willing to just accept that result, but I've worked too hard on this to just throw it away. So I'm leaving it here in case you wonder where the formula comes from.

Figure 2: Binomial with N=2000 and p=0.50



4.1 The $\pi^k(1-\pi)^{N-k}$ part is pretty obvious.

What is the probability of observing 7 coin flips with 5 $Heads$ and 2 $Tails$:

$$Prob(H,H,H,H,H,T,T) = \pi\cdot\pi\cdot\pi\cdot\pi\cdot\pi\cdot(1-\pi)\cdot(1-\pi) \tag{20}$$

$$\pi^5(1-\pi)^{7-5}$$

4.2 What about the Binomial Coefficient?

The Binomial coefficient $\left(\begin{array}{c} N \\ k \end{array}\right)$ is pronounced "N choose k". The number of ways to choose k successes out of N Bernoulli trials is

$$\left(\begin{array}{c} N \\ k \end{array}\right) = \frac{N!}{(N-k)!k!} \tag{21}$$

Where does that come from?

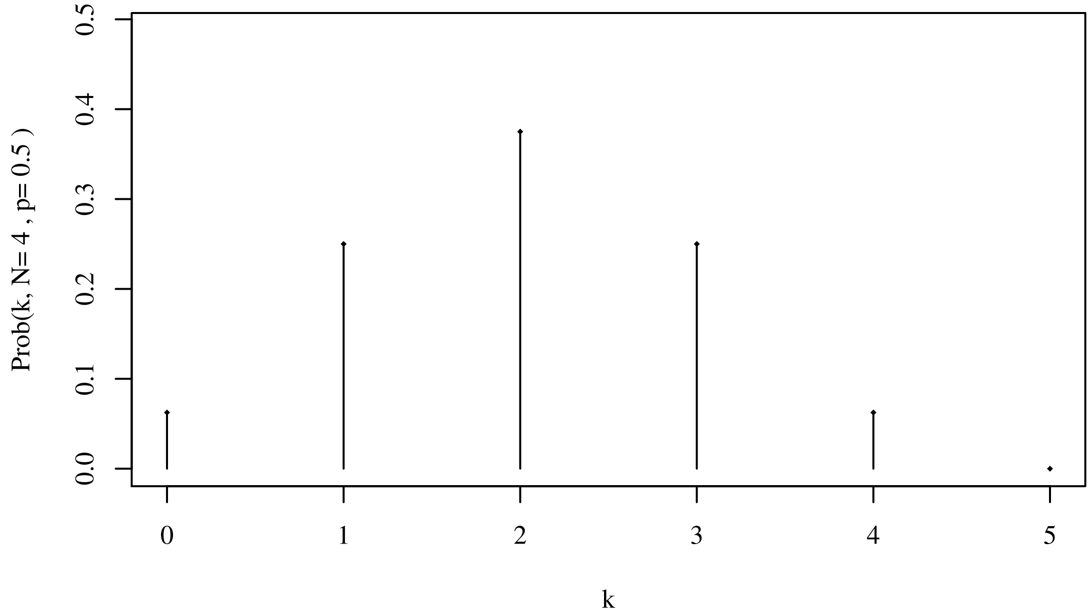
4.2.1 Step 1: We only really care about the number of successes, not the ordering.

$$\begin{array}{c} Prob(H,H,H,H,H,T,T) = Prob(T,T,H,H,H,H,H) = Prob(T,H,T,H,H,H,H) \\ = Prob(any\,ordering\,with\,5\,H\;and\,2\,T) \end{array}$$

We want the probability of 5 $Heads$ out of 7coin flips, we don't care if they happen in the beginning, middle or end.

That means when I want "the chances of 5 $Heads$ out of 7 flips", I need to go through and count up all of the different ways to get 5.

Figure 3: Binomial with N=4 and p=0.50



It turns out that calculating that is not hard, but it is much easier to explain in person with chalk and a blackboard than it is to writed it down in a clear, understandable way (and still survive the scrutiny of mathematicians).

4.2.2 A couple of examples.

I think I have a fool-proof illustration of the Binomial probability model using these $N = 3$. The possible number of "heads" in a series of 3 coin flips is 0, 1, 2, 3. The N choose k notation gives the number of different ways in which these can be obtained.

There is only 1 way to obtain 0 "Heads" out of 3 flips. We would have to observe Tails 3 times, (T, T, T).

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \frac{3!}{(3-0)!0!} = \frac{3!}{3!} = 1 \tag{22}$$

Next, consider 1 $Head$ and 2 $Tails$.

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \frac{3!}{(3-1)!1!} = \frac{3 \cdot 2 \cdot 1}{2 \cdot 1} = 3 \tag{23}$$

The 3 possible ways to end up with 1 $Head$ and 2 $Tails$ are $(H, T, T)(T, T, H)(H, T, T)$.

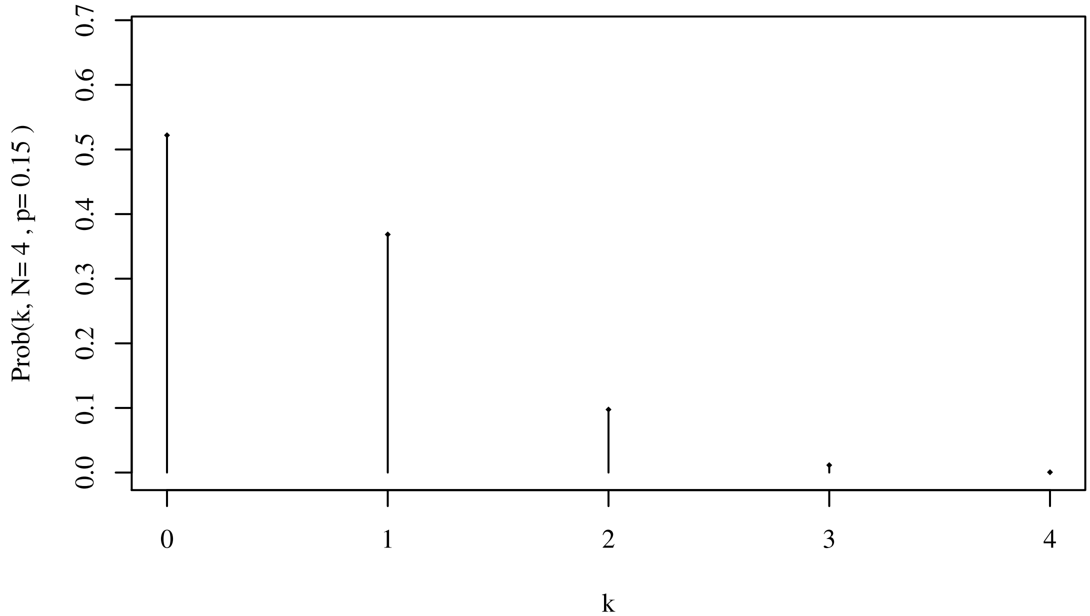
Next, consider 2 $Heads$ and 1 $Tail$.

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \frac{3!}{(3-2)!2!} = \frac{3 \cdot 2 \cdot 1}{2 \cdot 1} = 3 \tag{24}$$

The 3 possible ways to end up with 2 $Heads$ and 1 $Tail$ are $(H, H, T)(H, T, H)(T, H, H)$.

Finally, consider that there is only one way to get 3 $Heads$:

Figure 4: Binomial with N=4 and p=0.150



$$\left(\begin{array}{c} 3 \\ 3 \end{array} \right) = \frac{3!}{(3-3)!3!} = \frac{3!}{3!} = 1 \tag{25}$$

There are 8 possible outcomes in a 3 coin-flip experiments, then, and the chances of each 3-tuple are summarized in the following table:

Outcome	*Probability*
(H, H, H)	π^3
(H, H, T)	$\pi^2(1-\pi)$
(H, T, H)	$\pi^2(1-\pi)$
(H, T, T)	$\pi(1-\pi)^2$
(T, H, H)	$\pi^2(1-\pi)$
(T, H, T)	$\pi(1-\pi)^2$
(T, T, H)	$\pi(1-\pi)^2$
(T, T, T)	$(1-\pi)^3$

(26)

We have 8 sets of $3-tuples$, but we don't really need that many. Note that the probability of (H, H, T) is the same as the probability of (H, T, H) or (T, H, H).

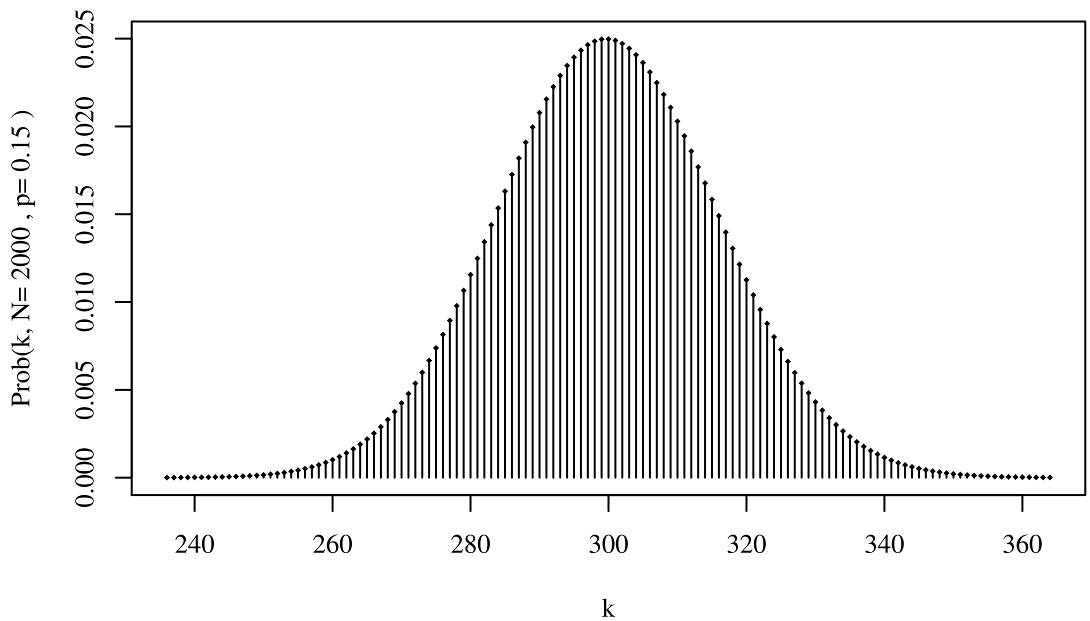
The Binomial distribution groups those together. We need to collect together the outcomes in which there are 0, 1, 2 , and 3 outcomes. When we group together the outcomes with a certain number of successes, we end up with $\left(\begin{array}{c} N \\ k \end{array} \right)$ of each type.

The probability of observing 3 $Tails$ ($3Failures$) is

$$Prob(T, T, T) =$$

$$Prob(0|3, p) = 1 \cdot (1-\pi)^3 = \left(\begin{array}{c} 3 \\ 0 \end{array} \right) \pi^0 (1-\pi)^3 \tag{27}$$

Figure 5: Binomial with N=2000 and p=0.150



And the chance of 1 *Head* is:

$$Prob(H,T,T) + Prob(T,H,T) + Prob(T,T,H) =$$
$$= 3 \cdot \pi^1 (1-\pi)^2 = \left(\begin{array}{c} 3 \\ 1 \end{array} \right) \pi^1 (1-\pi)^2 \qquad (28)$$

The probability of observing 2 *Heads* out of 3 flips (2 *Successes* out of 3 tests) is:

$$Prob(H,H,T) + Prob(H,T,H) + Prob(T,H,H) =$$

$$Prob(2|3,p) \quad = \quad 3 \cdot \pi^2 (1-\pi) = \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \pi^2 (1-\pi) \qquad (29)$$

The probability of observing 3 *Heads* (3 *Successes*) is

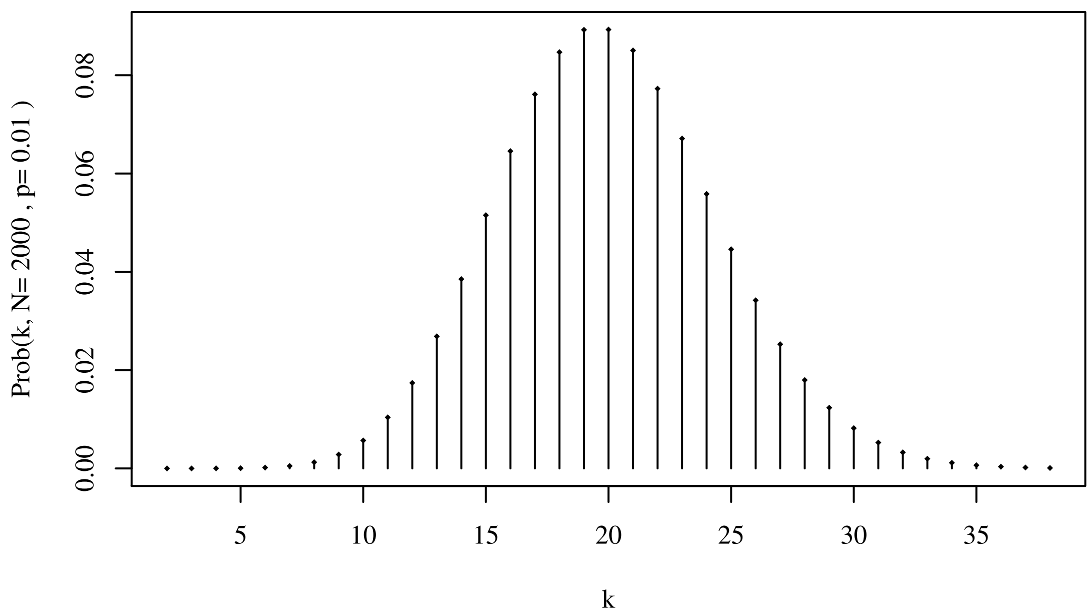
$$Prob(H,H,H) =$$
$$Prob(3|3,p) = 1 \cdot \pi^3 (1-\pi)^0 = \left(\begin{array}{c} 3 \\ 3 \end{array} \right) \pi^3 (1-\pi)^0 \qquad (30)$$

Note that if we add up those 4 expressions, the result is equal to 1.0. That means we have found a probability distribution on the set of possible combinations of *Heads* and *Tails* with 3 experiments. And that was the whole point of this from the beginning.

How many ways are there to get 5 *Heads* during 10 coin flips? Suppose $N = 10$ and $k = 5$.

$$\left(\begin{array}{c} 10 \\ 5 \end{array} \right) = \frac{10!}{(10-5)!5!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5!}{5! \cdot 5!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{5 \cdot 4 \cdot 3 \cdot 2} = 252 \qquad (31)$$

Figure 6: Binomial with N=2000 and p=0.01



Why don't you go ahead and make me a list of all possible ways to draw 5 *Heads* out of 10 coin flips. Start with:

$$(H, H, H, H, H, T, T, T, T, T)$$

You owe me 251 more vectors.

Note: In R you can confirm that calculation. See `?Special` and then run `choose(10,5)`.

4.2.3 Step 2: Generalize the previous reasoning.

The previous section demonstrates that finding for $N = 3$, but for higher values of N it would not be too practical to write out that same argument. We need a result for all N and k.

Recall that $\begin{pmatrix} N \\ k \end{pmatrix}$ is the number of ways to get

- k successes
- out of N tests

without concerning ourselves about the order in which the successes and failures occur.

Represent the N successes that might result in N coin flips as $(H_1, H_2, H_3, \ldots, H_N)$. The flips are numbered by the test on which they are observed.

1. Question: How many ways can we order N items? That is the same as asking for the number of ordered sets of N items that can be created out of a set of N items.

   Answer:

   $$N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot 2 \cdot 1 = N! \tag{32}$$

   You can pick N different items on your first pick, but only $N-1$ on your second, $N-2$ on your third, and so forth.

Example: $X = \{H_1, H_2, H_3\}$, so $N = 3$.

The claim is those can be ordered in $N! = 3 \cdot 2 \cdot 1 = 6$ ways.

List them out to verify:

$$(H_1, H_2, H_3)\,(H_1, H_3, H_2)\,(H_2, H_1, H_3)\,(H_2, H_3, H_1)\,(H_3, H_1, H_2)\,(H_3, H_2, H_1) \tag{33}$$

2. Repeat that same exercise, but only pick k successes out of the N possible successes. That is, we draw one from the set of N outcomes, and then one from the remining $(N-1)$, then one from the remaining $(N-2)$ until we have taken out k successes. The total number of ways to draw those k successes is:

$$N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-k+1) \tag{34}$$

This can be represented as

$$\frac{N!}{(N-k)!} \tag{35}$$

See why?

$$\frac{N!}{(N-k)!} = \frac{N \cdot (N-1) \cdot \ldots \cdot (N-k+1) \cdot (N-k) \cdot (N-k-1) \cdot (N-k-2) \ldots \cdot 2 \cdot 1}{(N-k) \cdot (N-k-1) \cdot (N-k-2) \cdot \ldots \cdot 2 \cdot 1} \tag{36}$$

$$\frac{N \cdot (N-1) \cdot \ldots (N-k+1) \ldots \cancel{(N-k)} \cdot \cancel{(N-k-1)} \cdot \ldots \cdot \cancel{2}\cancel{1}}{\cancel{(N-k)} \cdot \cancel{(N-k-1)}\cancel{(N-k-2)} \cdot \ldots \cdot \cancel{2} \cdot \cancel{1}} \tag{37}$$

$$= N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-k+1) \tag{38}$$

This gives us the number of ways you can have k successes in N experiments when you take the order of successes into account.

3. We still have too many outcomes because we are still treating an outcome like (H_1, T, H_3, T, H_5) as if it were a different thing than (H_3, T, H_1, T, H_5). There are $k!$ different ways in which the 3 victories might be ordered. To obtain the final result, we divide the number of ordered N-tuples that have k successes by $k!$:

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

4. Another illustration with 3 coin flips. Let's apply that approach to calculate the number of ways in which we can obtain 2 $Heads$.

Suppose we lay out the 3 possible $Heads$ that might be observed:

$$\{H_1, H_2, H_3\}$$

We can draw 2 items from this list in $3 \cdot 2$ methods:

$$\begin{array}{cc} (H_1, H_2) & (H_2, H_1) \\ (H_1, H_3) & (H_3, H_1) \\ (H_2, H_3) & (H_3, H_2) \end{array} \tag{39}$$

When we fail to obtain a $Head$, then we must be observing a $Tail$, the $3-tuples$ would be:

$$\begin{array}{cc} (H_1, H_2, T) & (H_2, H_1, T) \\ (H_1, T, H_3) & (H_3, T, H_1) \\ (T, H_2, H_3) & (T, H_3, H_2) \end{array} \tag{40}$$

Group together the similar sets, and we end up with 3 possible types of outcomes with 2 $Heads$ out of 3 flips:

$$\begin{array}{c} (H, H, T) \\ (H, T, H) \\ (T, H, H) \end{array} \tag{41}$$

5 Multinomial

Binomial is based on 2 possible outcomes with probabilities π_1 and π_2. And the outcome of a Binomial experiment is not just one number of successes x, but it is really a pair, (x_1, x_2), which represent x_1 successes and x_2 failures. We don't usually think of it that way because $\pi_2 = 1 - \pi_1$, so we don't need to keep track of 2 separate π's. And $x_2 = N - x_1$, so we don't think of a Binomial as generating a 2-tuple like $(x, N - x)$, *but we could.* And that's important.

5.1 Extend the Binomial by thinking of m possible outcomes.

Next suppose there are 3 possible outcomes, say, "Win","Lose","Draw", and the probabilities of these are (π_1, π_2, π_3). All of those probabilies must sum to 1.0, so it is not really necessary to keep track of 3 different values. So many times, we just write $(\pi_1, \pi_2, 1 - \pi_1 - \pi_2)$. When we collect a sample of N observations from the "Win","Lose","Draw" distribution, the outcomes will be spread over 3 values in a vector, like (x_1, x_2, x_3) or, equivalently, $(x_1, x_2, N - x_1 - x_2)$.

Somebody comes along and says there are really 4 possible outcomes, "Win","Lose","Draw", and "Game Canceled". The chances of each one are $(\pi_1, \pi_2, \pi_3, 1 - \pi_1 - \pi_2 - \pi_3)$ and if we had a sample of N, they would be a vector like $(x_1, x_2, x_3, N - x_1 - x_2 - x_3)$.

Keep enumerating possibilities, eventually you should see a pattern emerging. We can have any number, say m possible outcomes. And we can list the probabilities for them,

$$\pi_1, \pi_2, \ldots, \pi_m \tag{42}$$

and we can hope to characterize the probability distribution of the vector

$$(x_1, x_2, \ldots, x_m) \tag{43}$$

There are N experiments altogether, but each one can reveal one of m outcomes with probabilities

An "outcome" is a vector of counts, a listing of how many outcomes of each type is observed. There are N items altogether, and they are divided among the m types.

Like the binomial, the probability is simply a reflection of the number of different ways a given set of counts can be observed. The element that makes the multinomial seem complicated is that there is an interaction across the counts for the different types. If a "freak" string of experiments led to only outcomes of type 1, then we would have an outcome vector of

$$(N, 0, 0, \ldots, 0)$$

If we want to suppose that there is 1 item in the last category, then one must be removed from the first one.

$$(N - 1, 0, 0, \ldots, 1)$$

We might draw another case that shows 2 observations in every single category

$$(2, 2, 2, \ldots, 2)$$

The main point is this. If the theory says the chances of the outcomes are $(\pi_1, \pi_2, \ldots, \pi_m)$, then we need a probability model that states the chances of observing any particular combination.

5.2 Explicit example with 3 possibilities.

Suppose we consider a particular outcome vector, like

$$(32, 10, 8)$$

As in the binomial case, we build a probability model in small steps. We have the chance of 32 things of type 1, which has to look something like

$$\begin{pmatrix} N \\ 32 \end{pmatrix} \pi_1^{32}$$

There are $\left(\begin{array}{c} N \\ 32 \end{array} \right)$ ways to arrange N things in which there are 32 things of type 1 in there, and π_1^{32} is the chance of getting one outcome with 32 things of type 1.

So you have already taken out 32 of the possible N outcomes.

Now consider the 10 outcomes in the second column. The "N" that's left-over after removing the 32 outcomes for the first column is $N - 32$. So the chance of getting 10 is

$$\left(\begin{array}{c} N-32 \\ 10 \end{array} \right) \pi_2^{10}$$

Consider the 8 in the third column. There are $N - 32 - 10$ outcomes left, and this column is grabbing 8 of them. So the chances of that are

$$\left(\begin{array}{c} N-32-10 \\ 8 \end{array} \right) \pi_3^{8}$$

Now multiply ALL OF THOSE TOGETHER because the overall probability of getting

$$(32, 10, 8, \ldots, 19)$$

has to be the product of the chance of getting each one separately. That is

$$P(X_1 = 32, X_2 = 10, X_3 = 8) = \left(\begin{array}{c} N \\ 32 \end{array} \right) \left(\begin{array}{c} N-32 \\ 10 \end{array} \right) \left(\begin{array}{c} N-32-10 \\ 8 \end{array} \right) \pi_1^{32} \pi_2^{10} \pi_3^{8} \tag{44}$$

5.3 Generalize to m possible outcomes

The Multinomial probability model is obtained by continuing that same procedure for each of m possible outcomes.

$$P(x_1, x_2, x_3, \ldots, x_m) = \left(\begin{array}{c} N \\ x_1 \end{array} \right) \left(\begin{array}{c} N-x_1 \\ x_2 \end{array} \right) \left(\begin{array}{c} N-x_1-x_2 \\ x_3 \end{array} \right) \cdots \left(\begin{array}{c} N-x_1 \ldots - x_{m-1} \\ x_m \end{array} \right) \pi_1^{x_1} \pi_2^{x_2} \pi_3^{x_3} \cdots \pi_m^{x_m} \tag{45}$$