

V. PROBABILITY THEORY AND THE BINOMIAL DISTRIBUTION

1. Probability definitions
2. Rules of counting
3. Rules of Probability
4. Probability distributions
5. Binomial probability distribution

Probability Definitions

- **Trial** - *any operation that results in the collection of observations, whose outcomes cannot be predicted in advance with certainty.* A trial is also referred to as an experiment. For example, the act of flipping a coin, rolling a die, or drawing a card from a deck is each a trial.
- **Outcome** - *each distinct result of a trial.* For example, we may obtain a “heads” on the flip of a coin, a “2” on the roll of a die, or an “Ace of Spades” on the draw of a card.
- **Sample space** - *the set of all possible outcomes, denoted S .* For example, the sample space for flipping a coin is:

$$S = \{H, T\}$$

for rolling a die is:

$$S = \{1, 2, 3, 4, 5, 6\}$$

and for drawing a card is all 52 cards. The sample space for flipping two coins is:

$$S = \{HH, HT, TH, TT\}$$

- **Event** - *any collection of outcomes, or any subset of S .* We often refer to the occurrence of an event on a given trial as a *success*. The probability of an event A is defined as:

$$p(A) = \frac{\# \text{ of observations favoring event } A}{\text{total } \# \text{ of observations in sample space}}$$

- **Complement of an event** - *the mutually exclusive and exhaustive set of all possibilities of an event not occurring.* We often refer to the occurrence of the complement of an event on a given trial as a failure. The probability of the complement of an event A is denoted:

$$p(\text{not } A) = p(\bar{A}) = p(A') = 1 - p(A)$$

The sum of an event and its complement is always one: $p(A) + p(A') = 1$.

- Let's work through some examples of events.
 - Let's say we flip a single coin. The sample space is $S = \{H, T\}$. The event of obtaining “heads” is $p(H) = 1/2 = 0.50$.
 - Let's say we roll a die. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. The event of rolling a “one” is $p(1) = 1/6 = 0.17$. The event of rolling an “even number” is $p(\text{even}) = 3/6 = 0.50$.
 - Let's say we are drawing cards from a deck. The sample space for this example is all 52 cards. The event of drawing the “jack of diamonds” is $p(\text{jack of diamonds}) = 1/52 = 0.02$. The event of drawing an “ace” is $p(\text{ace}) = 4/52 = 0.08$.

- The notion of the *probability* of an event deals with events over the long run, or over an infinite number of trials. The fact that we refer to an infinite number of trials means that a probability is a theoretical concept. It does not imply that if we flip a coin twice, for example, we will always get “heads” on one trial and “tails” on the other, since each outcome occurs with probability $1/2$. In the terminology of statistics, then, we refer to a probability as a “limiting relative frequency.” This is in contrast to a proportion or relative frequency from a distribution of outcomes in a finite sample. In practice, however, probability is often used in place of proportion.

Rules of Counting

- **Basic counting rule:** the total number of possible outcomes from n independent trials is $k_1 \times k_2 \times k_3 \dots k_n$ where k indicates the number of possible outcomes for each of the n independent trials. A special case of the basic counting rule is when we have the same trial repeated numerous times. In this case, k is the same for each of the n trials, so the number of possible outcomes is k^n .

– It is often useful to create a tree diagram, which allows you to visualize all possible outcomes.

- **Permutation rule:** the number of possible *ordered* arrangements of r objects from a group of n objects (notice that we are now talking about *objects* rather than *trials*). It is important to remember that with permutations, order matters. Each different ordering is treated as a separate, independent outcome.

with replacement: $P_r^n = n^r$

without replacement: $P_r^n = \frac{n!}{(n-r)!}$

– A special case of the permutation rule exists when we are interested in knowing how many ordered arrangements of all n objects we can obtain (i.e., $r = n$). In this case, n different objects can be arranged $n!$ different ways. This formula tells us that the first position may be taken by any one of the n objects. Having filled this position we can then fill the second position by any of the $n - 1$ remaining objects, the third position by any of the remaining $n - 2$ objects, and so on. This is an example of sampling without replacement.

- **Combination rule:** the number of possible *unordered* arrangements of r objects from a collection of n objects. This is an important difference between a permutation and a combination.

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

- Now let's put these counting skills to use with some examples. To answer each of these problems, you must first ask yourself if order matters. If order does not matter, you use the combination rule. If order does matter, you must then ask yourself if you are sampling with or without replacement.

– Most license plates have $r = 3$ letters (from $n = 26$) followed by $r = 3$ numbers (from $n = 10$) with the possibility of repeating values (i.e., with replacement). We can use the permutation rule with replacement to determine that there are $(26^3)(10^3) = 17,576,000$ combinations.

– A standard combination lock has $n = 36$ numbers, $r = 3$ numbers in the combination, and no repeated numbers (i.e., without replacement). We can use the permutation rule without replacement to determine that there are

$$\frac{36!}{(36-3)!} = \frac{36!}{33!} = 42,840$$

possible combination sequences. So, the chance that someone will guess your locker combination on the first try is 1 in 42,840, or a probability of .0000233!

– When dealing cards, the order in which the cards are dealt does not matter, thus we will use the combination rule. The number of possible hands with five cards, in which $n = 52$ and $r = 5$, is

$$\frac{52!}{5!(52-5)!} = \frac{52!}{5!47!} = \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1} = 2,598,960$$

– Suppose we have a bag of four marbles, each of a different color (red, orange, green, blue). How many *ordered* arrangements of these marbles are possible? We have $n! = 4! = 4 \times 3 \times 2 \times 1 = 24$ permutations of the four marbles:

ROGB	ORGB	GROB	BGRO
ROBG	ORBG	GRBO	BGOR
RBOG	OBRG	GBOR	BRGO
RBGO	OBGR	GBRO	BROG
RGBO	OGBR	GORB	BORG
RGOB	OGRB	GOBR	BOGR

- Now let's say that we only want to draw two marbles. In this case we have:

$$\frac{n!}{(n-r)!} = \frac{4!}{(4-2)!} = \frac{(4 \times 3 \times 2 \times 1)}{(2 \times 1)} = 12$$

ordered arrangements of any two marbles from four that are possible. This provides us with the following possible outcomes:

RO	OR	BO	GO
RB	OB	BR	GR
RG	OG	BG	GB

- Suppose that we are interested in the number of unordered arrangements of the marbles that are possible. In this case, ROGB is treated the same as the other twenty-three arrangements. We can see that we should obtain only one unordered arrangement of these four marbles. We can use the combination formula to obtain it:

$$\frac{4!}{4!(4-4)!} = \frac{4!}{4!} = 1$$

If I were to draw two marbles at a time out, how many combinations are possible? We have

$$\frac{4!}{2!(4-2)!} = \frac{4!}{2!2!} = \frac{24}{4} = 6$$

unordered arrangements possible. Let's see why this is so:

RO/OR	BG/GB
RG/GR	BO/OB
RB/BR	GO/OG

Rules of Probability

- **Bounding rule:** $0 \leq p(A) \leq 1$. This is the simple rule that a probability must be bound by zero and one. In other words, it is impossible to obtain a probability that is negative or greater than one.
- **Addition rule for alternative events:** $p(A \cup B) = p(A) + p(B) - p(A \cap B)$. This rule is concerned with the probability of observing *either* event A *or* event B, or the *union* of events A and B.
 - **Mutually exclusive events** – two events that cannot occur at the same time. If two events are mutually exclusive, we have a special case of the addition rule: $p(A \cup B) = p(A) + p(B)$.
- **Multiplication rule for compound events:** $p(A \cap B) = p(A) \times p(B | A)$. This rule is concerned with the probability of observing events A and B *simultaneously*, or the *intersection* of events A and B. This is also referred to as a joint probability and may also be written as $p(A \text{ then } B)$.
- **Statistically independent events** – the probability of one event is not affected by the occurrence of another event. In other words, the conditional probability of one event given that the other event has occurred is equal to the unconditional probability of the event, or $p(B | A) = p(B)$. We refer to this as sampling *with replacement*. If two events are statistically independent, we have a special case of the multiplication rule: $p(A \cap B) = p(A) \times p(B)$.
- **Conditional probability** – the probability of one event occurring given that another event has occurred, denoted $p(B | A)$ in this equation. There are many different types of questions that I can ask that involve conditional probabilities. For example, what is the likelihood that someone will be incarcerated given that they commit a violent offense? What is the chance someone will be injured in a car crash given that they don't wear a seat belt? What is the probability that a person will be convicted as an adult if they are convicted as a juvenile? What is the probability that someone will relapse into drug use conditional on participating in a drug program?
- Not all compound events involve replacement. For example, replacement is not an issue with tossing a coin or rolling dice. This is because dice and coins have no memory. Each trial is independent of the one that occurs before and after it. For example, no matter how many heads you flip in a row, the probability of obtaining a tails on the next flip will always be 0.5. Drawing cards, on the other hand, you must specify whether you are sampling with or without replacement.
 - Let's work through some examples using coin flips, dice rolls, and card decks.

$$* p(H \cup T) = p(H \cup H') = \frac{1}{2} + \frac{1}{2} = 1.0$$

$$* p(HH \cup TT) = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = 0.50$$

$$* p(H \cap T) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25$$

$$* p(HH \cap TT) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16} = 0.06$$

$$* p(H \cap H \cap H \cap H \cap H) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{32} = 0.031$$

$$* p(1 \cup 6) = \frac{1}{6} \times \frac{1}{6} = \frac{2}{6} = 0.33$$

$$* p(1 \cup 2 \cup 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = 0.50$$

$$* p(1 \cap 6) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} = 0.028$$

$$* p(1 \cap 2 \cap 3) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216} = 0.004$$

$$* p(\text{ace} \cup \text{king}) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = 0.16$$

$$* p(\text{ace} \cup \text{diamond}) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = 0.31$$

$$* \text{ with replacement: } p(\text{ace} \cap \text{ace}) = \frac{4}{52} \times \frac{4}{52} = \frac{16}{2704} = 0.006$$

$$* \text{ without replacement: } p(\text{ace} \cap \text{ace}) = \frac{4}{52} \times \frac{3}{51} = \frac{12}{2652} = 0.005$$

- A contingency table provides a way to illustrate the joint distribution of two variables. A contingency table is often referred to as a “cross-tab.” We have already worked with contingency tables, but I have just not defined them as such up to this point. The dimension of a contingency table is defined by the number of categories in the two variables. Contingency tables provide a useful way to practice the addition and multiplication rules. The simplest contingency table is a cross-tab of two binary variables, which we refer to as a 2x2 contingency table.

- The following contingency table is a cross-tab of employment status with a six-item self-report delinquency index. Let’s consider some examples using these data.

Delinquent	Employed		Total
	No	Yes	
No	3642	3046	6688
Yes	955	1291	2246
Total	4597	4337	8934

- And here are some equations to compute probabilities:
 - $p(E) = \frac{4337}{8934} = 0.485$
 - $p(D) = \frac{2246}{8934} = 0.251$
 - $p(\text{ND} \cup \text{NE}) = \frac{6688}{8934} + \frac{4597}{8934} - \frac{3642}{8934} = \frac{7643}{8934} = 0.855$
 - $p(\text{D} \cup \text{E}) = \frac{2246}{8934} + \frac{4337}{8934} - \frac{1291}{8934} = \frac{5292}{8934} = 0.592$
 - $p(\text{E} \cup \text{NE}) = \frac{4337}{8934} + \frac{4597}{8934} = \frac{8934}{8934} = 1.00$
 - $p(\text{D} \cap \text{E}) = \frac{2246}{8934} \times \frac{1291}{2246} = \frac{1291}{8934} = 0.145$
 - $p(\text{D} | \text{E}) = \frac{1291}{4337} = 0.298$
 - $p(\text{D} | \text{NE}) = \frac{955}{4597} = 0.208$

- The following contingency table shows the joint distribution of school performance with delinquency. This is a 2x4 contingency table.

Delinquent	School Performance				Total
	A's & B's	B's & C's	C's & D's	D's & F's	
No	1878	1537	708	78	4201
Yes	1290	1679	1127	252	4348
Total	3169	1679	1127	330	8549

- And here are more equations to compute probabilities:

$$- p(B \cup B) = \frac{3168}{8549} + \frac{3216}{8549} = \frac{6384}{8549} = 0.747$$

$$- p(C \cup F) = \frac{1835}{8549} + \frac{330}{8549} = \frac{2165}{8549} = 0.253$$

$$- p(D | A) = \frac{1290}{3168} = 0.407$$

$$- p(D | B) = \frac{1679}{3216} = 0.522$$

$$- p(D | C) = \frac{1127}{1835} = 0.614$$

$$- p(D | F) = \frac{252}{330} = 0.764$$

$$- p(A \cap N) = \frac{3168}{8549} \times \frac{1878}{3168} = \frac{1878}{8549} = 0.220$$

Probability Distributions

- Earlier, we talked about frequency distributions, in which we listed the frequency, proportion, and percent of each value that we observe in sample data. Recall that a proportion or relative frequency is something that we actually observe in a sample. This means that a frequency distribution is an *empirical* distribution. A probability distribution is something akin to a frequency distribution, except that it is constructed from probability theory rather than from observed data in a sample. Recall that a probability or limiting relative frequency is something that we don't actually observe but would *expect* to observe over the long run. This means that a probability distribution is a theoretical distribution, in other words, it is a frequency distribution for an infinite number of trials.
- Let's consider an example. Suppose that I flip a single coin. The sample space is $S = \{H, T\}$, so we know that, theoretically, $p(H) = 0.50$ and $p(T) = 0.50$ in an infinite number of coin flips. Now suppose that I flip the coin ten times and record the number of times that I get heads and the number of times that I get tails. Then I flip the coin one hundred times, then one thousand, then ten thousand, and so on.

Outcome	f1	f2	f3	f4	f5
Heads	0	4	45	490	4990
Tails	1	6	55	510	5010
Number of Flips	1	10	100	1000	10000

- You can see what happens as I increase the number of flips. Slowly, the empirical frequency distribution is converging on the theoretical probability distribution, so that if I were to continue flipping thousands more times, I will get closer and closer to the theoretical limit of 0.50. Once I reach an infinite number of flips, in fact, I will obtain exactly 0.50. Unfortunately (depending upon your perspective), we will never actually flip a coin an infinite number of times, which is why probability distributions are theoretical.
- The fact that we don't actually observe probability distributions with empirical data does not make them any less useful. We don't have to observe a probability distribution in order to know its properties. Two very useful properties for our purposes are the mean and variance (or standard deviation) of a probability distribution. These are denoted $\mu = E(X)$ and $\sigma^2 = V(X)$, respectively. This knowledge forms an essential link for hypothesis testing and statistical inference.

Binomial Probability Distribution

- The best introduction to the binomial probability distribution is by way of example.
 - Suppose that we are looking at families with three children. We can treat each birth as an independent trial with two possible outcomes. With this in mind, how many different boy-girl sequences are possible? Using the basic counting rule, we can determine that there are $2^3 = 8$ possible outcomes. Notice that I did not use the permutation or combination rule to determine how many different boy-girl combinations are possible. This is because there is not a predetermined number of boys and girls that we are arranging in particular sequences.
 - Now, it is also the case that the sex of one child is independent of the sex of the other children. Nationally, boys are only slightly more likely to be born than girls, such that $p(\text{boy}) = 0.52$ and $p(\text{girl}) = 0.48$. Since we are dealing with independent trials, we can easily calculate the probability of obtaining a particular sequence of boys and girls.

Sequence	Probability
BBB	$(.52)(.52)(.52) = .141$
BBG	$(.52)(.52)(.48) = .130$
BGB	$(.52)(.48)(.52) = .130$
GBB	$(.48)(.52)(.52) = .130$
BGG	$(.52)(.48)(.48) = .120$
GBG	$(.48)(.52)(.48) = .120$
GGB	$(.48)(.48)(.52) = .120$
GGG	$(.48)(.48)(.48) = .111$

- We can use this information to answer interesting probability questions.
 - + $p(1 \text{ girl}) = p(BBG) + p(BGB) + p(GBB) = 0.130 + 0.130 + 0.130 = 0.390$
 - + $p(1 \text{ boy}) = p(BGG) + p(GBG) + p(GGB) = 0.120 + 0.120 + 0.120 = 0.360$
 - + $p(2+ \text{ girls}) = p(BGG) + p(GBG) + p(GGB) + p(GGG) = 0.111 + 0.120 + 0.120 + 0.120 = 0.471$
 - + $p(2+ \text{ boys}) = p(BBB) + p(BBG) + p(BGB) + p(GBB) = 0.141 + 0.130 + 0.130 + 0.130 = 0.531$
- Notice how for each of these probability questions, the specific order of births is not important.
 - We can arrive at the same answer to our probability questions by using what is called a binomial probability distribution. A binomial distribution is a probability distribution for a particular kind of discrete variable, specifically, one that has only two outcomes on each trial (what we call dichotomous outcomes). The outcomes of binomial trials are labeled “successes” and “failures.” We denote the probability of obtaining a success on any given trial as p , and the probability of a failure on any given trial as $q = 1 - p$. There are n independent trials. A binomial distribution gives us the (theoretical) probability of observing r successes in n trials. Importantly, the specific sequence of successes and failures is unimportant, meaning we are interested in knowing the probability that one girl is born into a family of three children. It is not important to us whether she is born first, second, or third, only that one of the three children is a girl.
 - The binomial distribution combines the multiplication rule with the combination rule. Let’s return to our sex composition problem. Suppose we want to know $p(1 \text{ girl})$. We will label a success as the event that we obtain a girl on any given “trial.” By this definition, $p(\text{success}) = p(\text{girl}) = p = 0.48$. First, since each trial is independent, we can use the multiplication rule to determine that the probability of one specific sequence is $p(GBB) = (0.48)(0.52)(0.52) = 0.48^1 \times 0.52^2 = 0.130$. But this only gives us the probability that the first child is a girl and the next two boys, $p(\text{girl} \cap \text{boy} \cap \text{boy})$. Second, since we are not interested in order-specific sequences, we can use the combination rule to determine that the number of possible sequences of one girl and two boys

is

$$\frac{3!}{1!(3-1)!} = \frac{3!}{2!} = 3$$

Third, we simply take the product of the combination and multiplication rules to obtain $p(1 \text{ girl}) = 3 \times 0.130 = 0.390$.

- The binomial formula tells us the probability of obtaining r successes in n independent trials. It is calculated as follows:

$$p(r) = C_r^n p^r q^{n-r} = \binom{n}{r} p^r q^{n-r}$$

Note that this formula, as described above, simply combines information about the combination rule and the multiplication rule. In other words, it is the product of the number of possible sequences with the probability of obtaining one specific sequence. We can use this formula to construct the full probability distribution of sex composition. I will label a success as the event that a girl is born, there are $n = 3$ independent trials, with $p = p(\text{success}) = 0.48$ and $q = p(\text{failure}) = 1 - 0.48 = 0.52$. The probability distribution simply ranges the r successes over the n trials.

#of Girls (r)	C_r^n	$p^r q^{n-r}$	$p(r)$
0	1	$.48^0 \times .52^3 = .141$	$1 \times .141 = .141$
1	3	$.48^1 \times .52^2 = .130$	$3 \times .130 = .390$
2	3	$.48^2 \times .52^1 = .120$	$3 \times .120 = .360$
3	1	$.48^3 \times .52^0 = .111$	$1 \times .111 = .111$
Total	8		1.002



- Recall that a probability distribution illustrates the probability associated with a given outcome if the trials were repeated an infinite number of times. Thus, in any sample of families with three children, the relative frequency distribution of the number of girls may look different. But as the number of families increases infinitely, the relative frequency distribution will converge to the probability distribution.
 - We can use the information that the probability distribution provides to determine the mean μ , of this distribution. For the binomial distribution, $E(X) = \mu = np = (3)(0.48) = 1.44$. This tells us that the mean number of girls in 3-children families, where $p(\text{girl}) = 0.48$ is 1.44. The variance is computed as $V(X) = \sigma^2 = npq = (3)(0.48)(0.52) = 0.749$, with standard deviation $\sigma = 0.865$.
 - Let's consider another concrete example. We are interested in determining the probability that a person will run a red light. Let's say that prior research in Charlotte-Mecklenburg County has found that the probability that a car passing through a particularly dangerous intersection will run a red light is 0.60. We refer to the event that someone runs a red light as a success, and the associated probability of a success p . This $p = 0.60$ in this example. The event that a person does not run a red light is a failure, with probability $q = 0.40$. Let's say that we sit at this intersection, select 5 random cars, and write down whether each one runs a red light or not. The number of cars that we observe represents the number of trials, so $n = 5$. We find in our study that 4 cars run a red light. The number of successes is referred to as r , so in this example $r = 4$.
 - We can use the binomial formula to determine the probability of observing these 4 successes out of 5 trials. In other words, we can determine the probability of observing 4 cars out of 5 running a red light. First, we use the multiplication rule to determine the probability of obtaining a sequence of 4 successes and 1 failure in that order. Since these trials are statistically independent of each other, we can use the special case of the multiplication rule to obtain:

$$p(\text{SSSSF}) = p(S) \times p(S) \times p(S) \times p(S) \times p(F) = p^r q^{n-r} = 0.60^4 \times 0.40^1 = 0.052$$

Recall, however, that this represents the probability of obtaining this particular *ordered* arrangement of outcomes. So, we then use the combination rule to obtain the number of possible sequences of 4 successes out of 5 trials, computed as:

$$\frac{5!}{4!(5-4)!} = \frac{5!}{4!1!} = 5$$

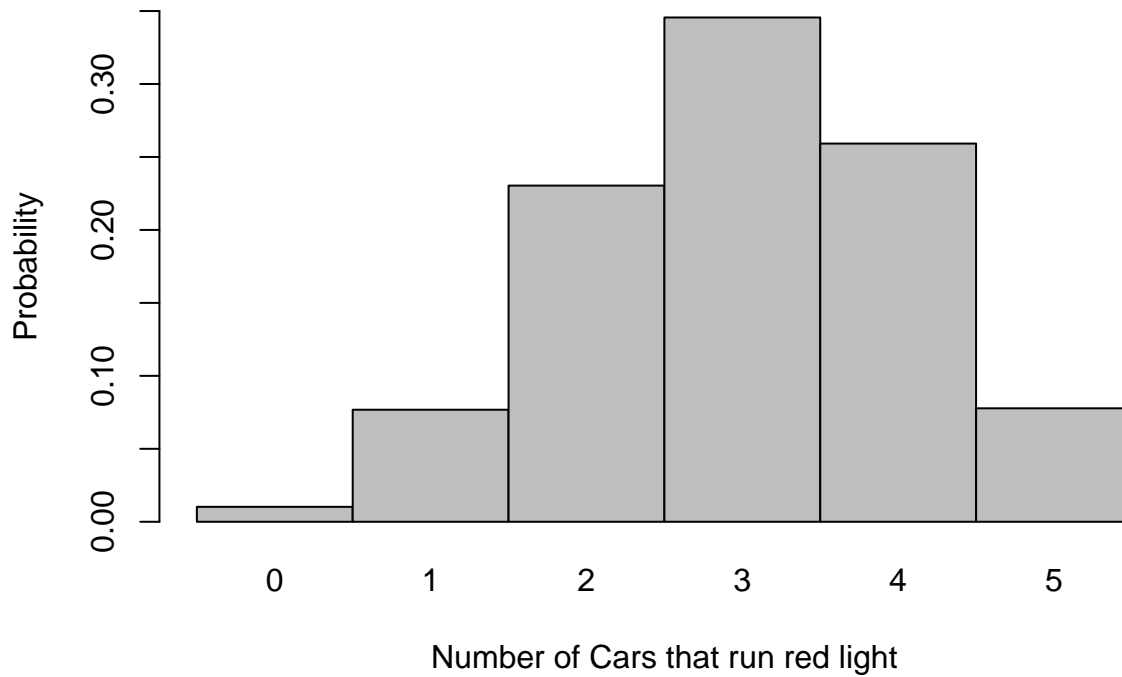
Third, we multiply these two quantities together to obtain the probability we are interested in:

$$p(r = 4) = 5 \times 0.052 = 0.259$$

- Let's also construct the full probability distribution, where $n = 5$, $p = 0.60$, and $q = 0.40$.
 - The moments of this probability distribution are $E(X) = np = (5)(0.60) = 3.0$ and $V(X) = npq = (5)(0.60)(0.40) = 1.200$ with $\sigma = 1.095$.

r	C_r^n	$p^r q^{n-r}$	$p(r)$
0	1	$.6^0 \times .4^5 = .010$	$1 \times .010 = .010$
1	5	$.6^1 \times .4^4 = .015$	$5 \times .015 = .075$
2	10	$.6^2 \times .4^3 = .023$	$10 \times .023 = .230$
3	10	$.6^3 \times .4^2 = .035$	$10 \times .035 = .350$
4	5	$.6^4 \times .4^1 = .052$	$5 \times .052 = .260$
5	1	$.6^5 \times .4^0 = .078$	$1 \times .078 = .078$
Total	32		1.003

Drivers are Bad All Over the Place

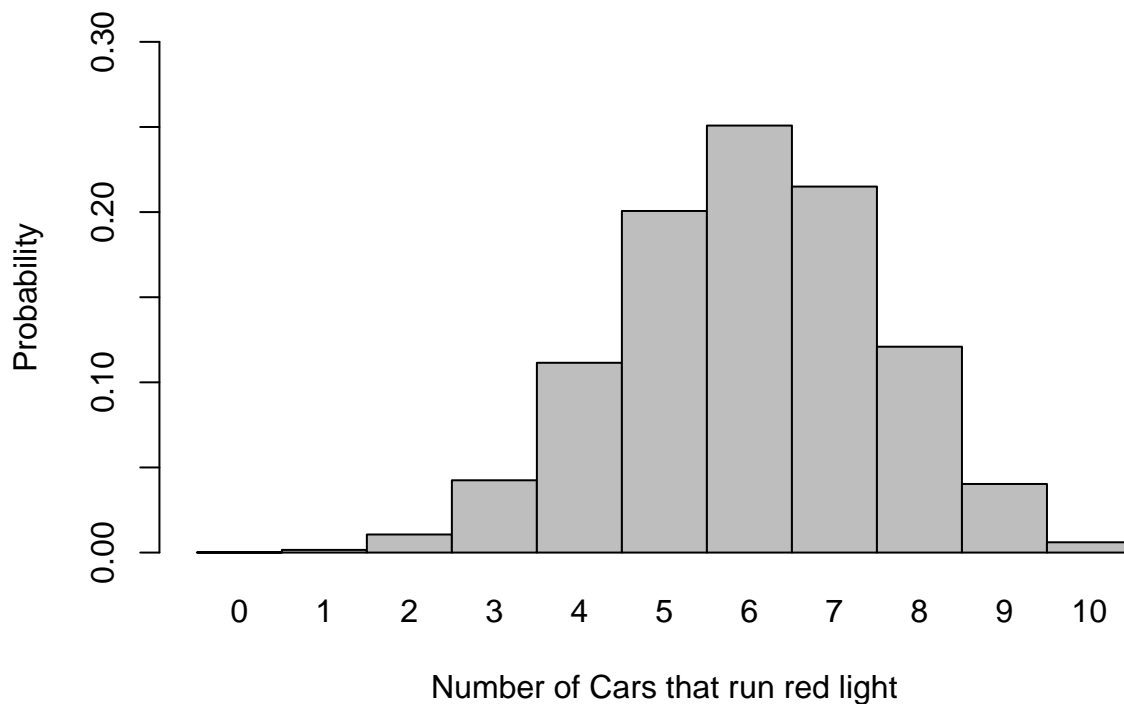


- Let's say that the example we just did took place a year ago, before Charlotte-Mecklenburg County installed red light cameras. We want to find out if these cameras have helped to reduce the number of people that run red lights. The probability of a success on any given trial has not changed, so $p = 0.60$ and conversely, $q = 0.40$. This time, we return to the same intersection, select 10 random cars passing through the intersection, and record whether each one stops for the red light or runs it. This means that $n = 10$. Let's say that this time around we observe only 2 cars that run red lights, so $r = 2$. Let's build this new probability distribution to see what it looks like.

r	$p(r) = C_r^n p^r q^{n-r}$	cp
0	$(1)(.6^0)(.4^{10}) = .000$.000
1	$(10)(.6^1)(.4^9) = .002$.002
2	$(45)(.6^2)(.4^8) = .011$.013
3	$(120)(.6^3)(.4^7) = .042$.055
4	$(210)(.6^4)(.4^6) = .111$.166
5	$(252)(.6^5)(.4^5) = .202$.368
6	$(210)(.6^6)(.4^4) = .250$.618
7	$(120)(.6^7)(.4^3) = .215$.833
8	$(45)(.6^8)(.4^2) = .121$.954
9	$(10)(.6^9)(.4^1) = .040$.994
10	$(1)(.6^{10})(.4^0) = .006$	1.000

- The moments of this probability distribution are $E(X) = np = (10)(0.60) = 6.0$ and $V(X) = npq = (10)(0.60)(0.40) = 2.40$, with $\sigma = 1.549$.

Red Light Cameras and Bad Drivers



- So, what is the probability that we would observe only two cars running a red light? The distribution tells us that this probability is 0.011. What is the probability that the number of cars that run red lights is in the tail; in other words, what is the probability that two or fewer cars run red lights? The

cumulative probability is 0.013, very remote. Do you think there is sufficient evidence in this sample to conclude that the installation of red light cameras was successful in curbing red-light running? Yes, we can say with some confidence that red light cameras have indeed been effective in reducing the number of cars that run red lights.

- Let's think of some other uses of the binomial probability distribution.
 - Suppose that a baseball player has only a 1/100 chance of missing any game due to illness or injury. What is the probability that he will play in 2,132 straight games (Cal Ripken)? We will use

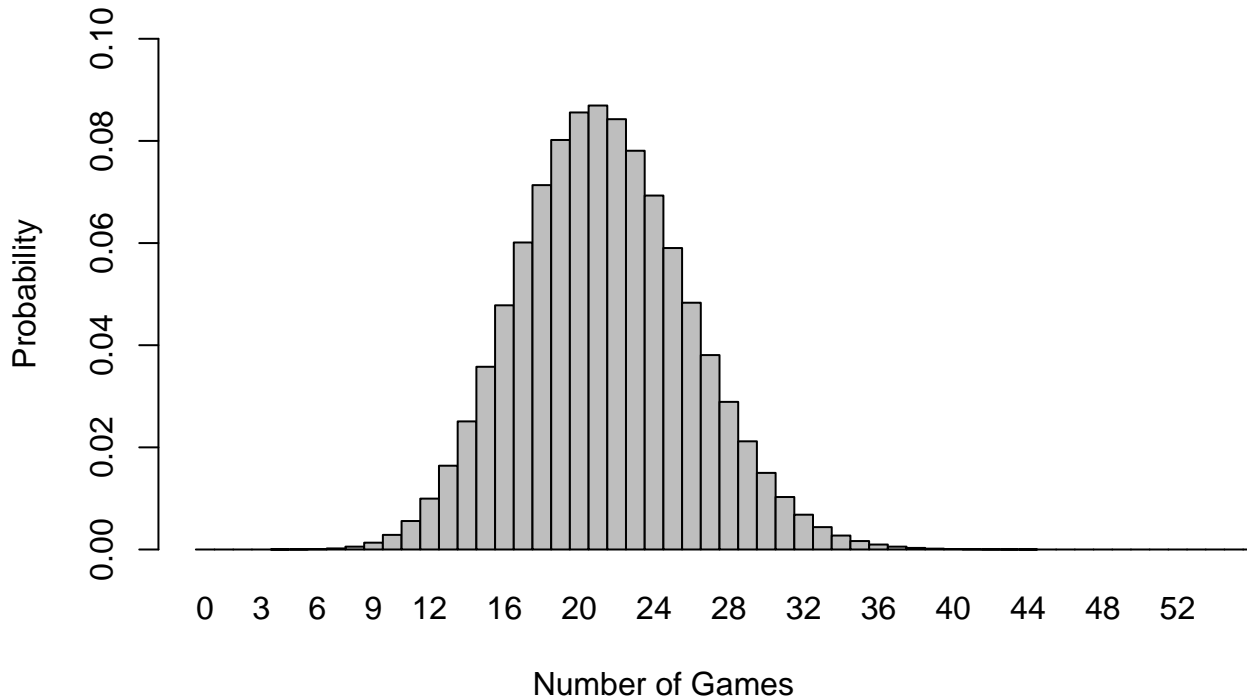
$$p = p(\text{miss game}) = 1/100 = 0.01$$

with $n = 2132$. We are looking for:

$$p(r = 0) = C_0^{2132}(0.01)^0(0.99)^{2132} = 0.99^{2132} = 4.95e - 10 = .000000000495$$

or 1 in 2,021,925,353. Another way to look at this is to ask, what is the probability that a player will miss at least one game due to illness or injury? We can use the complement rule to determine that $p(r \geq 1) = 1 - p(r = 0) = 1 - 4.95e - 10 = 1$. We would expect Ripken to miss $E(X) = (2132)(0.01) = 21.32$ games, on average, with $V(X) = (2132)(0.01)(0.99) = 21.107$ and $\sigma = 4.594$. The probability distribution is provided below.

Why They Called Cal Ripken the "Iron Man"



- Suppose your roommate takes a multiple-choice quiz with 10 questions and 4 answers per question. Your roommate gets 8 correct and brags that he guessed on every question. What is the probability that you would guess correctly on 8 or more questions? We will use $p = p(\text{correct}) = \frac{1}{4} = 0.25$, with $n = 10$. We are looking for:

$$p(r \geq 8) = p(r = 8) + p(r = 9) + p(r = 10) =$$

$$(45)(0.25)^8(0.75)^2 + (10)(0.25)^9(0.75)^1 + (1)(0.25)^{10}(0.75)^0 = 4.15802 \times 10^{-4}$$

or 1 in 2,405. On average, you would expect to guess $E(X) = (10)(0.25) = 2.5$ answers correctly, with $V(X) = (10)(0.25)(0.75) = 1.875$ and $\sigma = 1.369$.

My Lying Friend

