Section 8.6: Bernoulli Experiments and Binomial Distribution

We have already learned how to solve problems such as “if a person randomly guesses the answers to 10 multiple choice questions, what is the probability that they will get all 10 correct?” or “what is the probability that they will get none correct?”. We can even answer the questions “what is the probability that they will get at least one correct?” and “what is the probability that they will get at least one wrong?” using the complement rule. There are a number of events on the spectrum between getting all of the questions correct and getting all of them wrong. In this section, we will answers questions such as “what is the probability that the person will get exactly three correct?” or what is the probability that they will get at least three correct?”.

The situation above fits the criteria for a type of experiment called a **Bernoulli Experiment** with (in this case) 10 trials. These experiments involve repeated independent trials of an experiment with 2 outcomes called success and failure for the purposes of abstraction. The criteria for a Bernoulli experiment with n trials are shown below.

Bernoulli Experiment with n Trials

- The experiment is repeated a fixed number of times (n times).
- Each trial has only two possible outcomes success and failure. The possible outcomes are exactly the same for each trial.
- The probability of success remains the same for each trial. We use p for the probability of success (on each trial) and $q = 1 - p$ for the probability of failure.
- The trials are independent (The outcome of previous trials has no influence on the outcome of the next trial).
- We are interested in the random variable X where $X =$ the number of successes. Note the possible values of X are $0, 1, 2, 3, \ldots, n$.

Examples

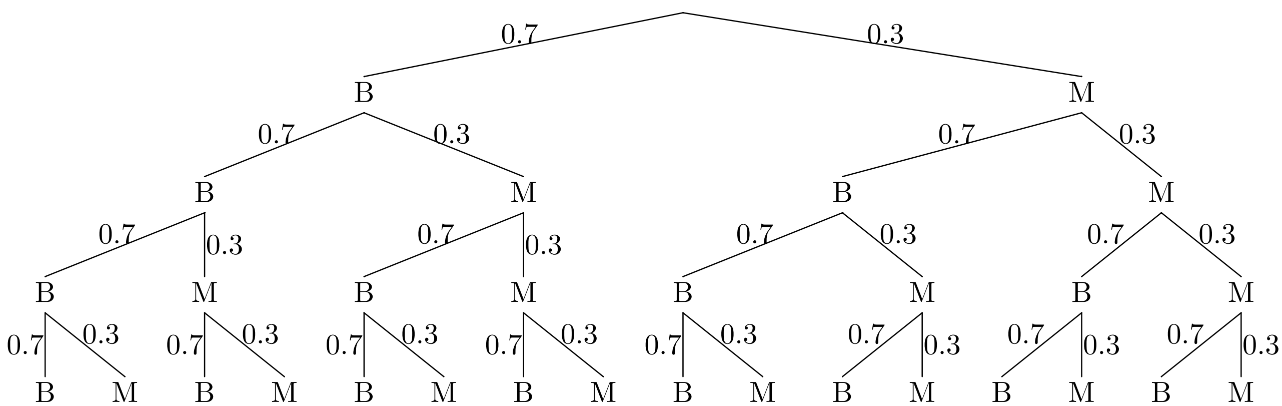
1. We flip a coin 12 times and count the number of heads. Here $n = 12$. Each flip is a trial. It is reasonable to assume the trials are independent. Each trial has two outcomes heads (success) and tails (failure). The probability of success on each trial is $p = 1/2$ and the probability of failure is $q = 1 - 1/2 = 1/2$. We are interested in the variable X which counts the number of successes in 12 trials. This is an example of a Bernoulli Experiment with 12 trials.
2. A basketball player takes four independent free throws with a probability of 0.7 of getting a basket on each shot. The number of baskets made is recorded. Here each free throw is a trial and trials are assumed to be independent. Each trial has two outcomes basket (success) or no basket (failure). The probability of success is $p = 0.7$ and the probability of failure is $q = 1 - p = 0.3$. We are interested in the variable X which counts the number of successes in 4 trials. This is an example of a Bernoulli experiment with 4 trials.
3. An urn contains 6 red marbles and 4 blue marbles. Five marbles are drawn from the urn without replacement and the number of red marbles is observed. We might let a trial here consist of drawing a marble from the urn and let success be getting a red. However, this is **not a Bernoulli experiment** since the trials are not independent (because the mix of reds and blues changes on each trial since we do not replace the marble) and the probability of success and failure vary from trial to trial.

4. An urn contains 6 red marbles and 4 blue marbles. A marble is drawn at random from the urn, its color is noted and then it is replaced. Five marbles are drawn from the urn in this way (with replacement) and the number of red marbles is observed. This is repeated five times. **This is a Bernoulli experiment** where each time we draw a marble from the urn constitutes one trial. Trials are independent since we draw randomly from the urn and the probability of success (getting a red) is the same on each trial $p = 6/10$ since we replace the marble after each draw. We are interested in the number of successes in five trials of this experiment.

Our next goal is to calculate the probability distribution for the random variable X, where X counts the number of successes in a Bernoulli experiment with n trials. We will start with a simple example for which a tree diagram can be drawn, in fact we have already looked at a specific case of this example when we studied tree diagrams.

Example A basketball player takes 4 independent free throws with a probability of 0.7 of getting a basket on each shot. Let X = the number of baskets he gets.

(a) Use the tree diagram below to find the probability that he gets exactly 2 baskets or $P(X = 2)$.
B = gets a basket, M = misses.



$P(X = 2) = C(4, 2)(0.7)^2(0.3)^2 = 0.2646$

In general we have the following:

If X is the number of success' in a Bernoulli experiment with n independent trials, where the probability of success is p in each trial (and the probability of failure is then $q = 1 - p$), then

$$\boxed{Pr(X = k) = C(n, k)p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}}$$

for $k = 0, 1, 2, \cdots, n$.

[We can see why this is true if we visualize a tree diagram for the n independent trials. The number of paths with exactly k success(out of n trials) is $C(n, k)$ and the probability of every such path equals $p^k q^{n-k}$. The event that $X = k$ can result from any one of these outcomes(paths), hence the $P(X = k)$ is the sum of the probabilities of all paths with exactly k successes which is $C(n, k)p^k q^{n-k}$.]

Example A basketball player takes 4 independent free throws with a probability of 0.7 of getting a basket on each shot. Let X = the number of baskets he gets. Write out the full probability distribution for X.

X	**P(X)**
0	
1	
2	
3	
4	

X	**P**(X)	
0	$C(4, 0)(0.7)^0(0.3)^4$	0.0081
1	$C(4, 1)(0.7)^1(0.3)^3$	0.0756
2	$C(4, 2)(0.7)^2(0.3)^2$	0.2646
3	$C(4, 3)(0.7)^3(0.3)^1$	0.4116
4	$C(4, 4)(0.7)^4(0.3)^0$	0.2401

Note $0.0081+0.0756+0.2646+0.4116+0.2401 = 1$

Binomial Random Variables For a Bernoulli experiment with n trials, let X denote the number of successes in the n trials, where the probability of success in each trial is p. The distribution of random variable X is called a **binomial distribution** with parameters n and p. The **expected value** of X is $E(X) = np$ and the **standard deviation of** X is $\sigma(X) = \sqrt{npq}$ where $q = 1 - p$.

Example If a basketball player takes 8 independent free throws with a probability of 0.7 of getting a basket on each shot, what is the probability that he gets exactly 6 baskets?

$C(8, 6)(0.7)^6(0.3)^2 = 0.29647548$

Example A student is given a multiple choice exam with 10 questions, each question has five possible answers. The student guesses each of the answers randomly.

(a) What is the probability that the student will get exactly 6 of the questions correct?

$n = 10$, $k = 6$, $p = \frac{1}{5} = 0.2$, so $C(10, 6)(0.2)^6(0.8)^4 = 0.005505024$

(b) What is the probability that the student will get at least 6 of the questions correct and will pass the class?

$C(10, 6)(0.2)^6(0.8)^4 + C(10, 7)(0.2)^7(0.8)^3 + C(10, 8)(0.2)^8(0.8)^2 + C(10, 9)(0.2)^9(0.8)^1 + C(10, 10)(0.2)^{10}(0.8)^0 = 0.0063693824.$

(c) What is the expected number of correct answers that will result from this student's strategy? What is the expected standard deviation?

$E(X) = np = 10 \cdot 0.2 = 2$, $\sigma(X) = \sqrt{npq} = \sqrt{10 \cdot 0.2 \cdot 0.8} = \sqrt{1.6} \approx 1.2649110641$

Example Assume that the Phillies and the Yankees are in the world series, that the Phillies have a 3/5 chance of winning any given game, and that the games are independent experiments. What is the probability of a 7 game series. **Note** A seven game series will occur only when each team wins 3 of the first 6 games.

A seven game series will occur whenever the Phillies win exactly 3 of the first 6 games. The probability of this is $C(6, 3)(0.6)^3(0.4)^3 = 0.27648$.

It is also true that a seven game series will occur whenever the Yankees win exactly 3 of the first 6 games. The probability of this is $C(6, 3)(0.4)^3(0.6)^3$.

Example (Quality Control) The Everlasting Lightbulb company produces lightbulbs, which are packaged in boxes of 20 for shipment. Tests have shown that 4% of their lightbulbs are defective.

(a) What is the probability that a box, ready for shipment, contains exactly 3 defective lightbulbs?

$C(20, 3)(0.04)^3(0.96)^{17} = 0.0364498535.$

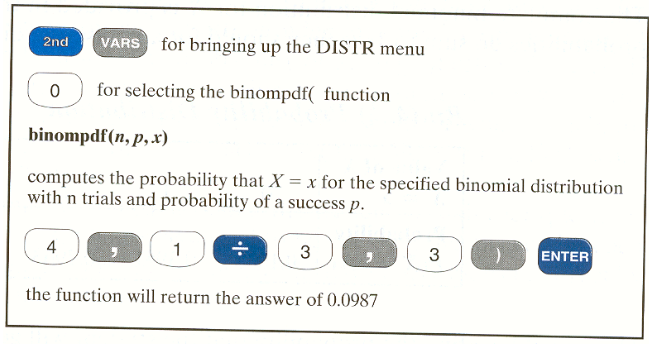
(b) What is the probability that the box, contains 3 or more defective lightbulbs?

$1 - \left(C(20,2)(0.04)^2(0.96)^{18} + C(20,1)(0.04)^1(0.96)^{19} + C(20,0)(0.04)^0(0.96)^{20}\right) = 1 - 0.9561372094 = 0.0438627906.$

We can also compute the expected number of defective bulbs, $E(X) = 20 \cdot 0.04 = 0.8$, and the expected standard deviation, $\sigma(X) = \sqrt{20 \cdot 0.04 \cdot 0.96} \approx 0.876356092.$

Extras: Using Your Calculator

Let X be a binomial random variable where $n = 4$ and $p = 1/3$. We can use a TI-83 calculator to calculate $P(x = 3)$


2nd VARS for bringing up the DISTR menu

0 for selecting the binompdf(function

binompdf(*n*, *p*, *x*)

computes the probability that $X = x$ for the specified binomial distribution with n trials and probability of a success p.

4 , 1 ÷ 3 , 3) ENTER

the function will return the answer of 0.0987

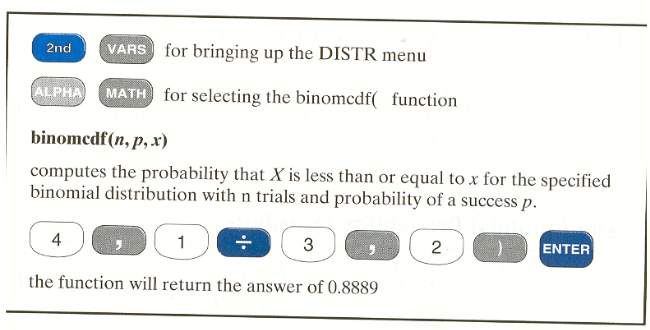
There are two choices for the binomial distribution on your calculator, binomialpdf and binomialcdf.

binomialpdf(4, 1/3, 3) gives $P(x = 3)$, when x is binomial with $n = 3$, $p = 1/3$ and $k = 3$.

binomialcdf(4, 1/3, 3) gives $P(x \leq 3)$ when x is binomial with $n = 3, p = 1/3$ and $k = 3$.

pdf means probability density function whereas cdf means cumulative density function.

We can also calculate $P(x \leq 2)$ when $n = 4$, and $p = \frac{1}{3}$ as follows:


2nd VARS for bringing up the DISTR menu

ALPHA MATH for selecting the binomcdf(function

binomcdf(*n*, *p*, *x*)

computes the probability that X is less than or equal to x for the specified binomial distribution with n trials and probability of a success p.

4 , 1 ÷ 3 , 2) ENTER

the function will return the answer of 0.8889

The directions for different models of TI calculators can be found online.

Although taking a random sample for polling from a population is much line drawing marbles from an urn without replacement, if the population is very large, the chemistry of the population does not change very much from trial to trial if we do not replace the subjects. Therefore we can use the binomial distribution as a good approximation to the distribution of the random variable X in the following experiment.

Example Suppose that the voting population in Utopia is 300 million and 60% of the voting population intend to vote for Melinda McNulty in the next election. We take a random sample of size 100 from the same voting population and ask each person chosen whether they will vote for Melinda McNulty in the next election or not. Let X be the number of yes' in our sample. The possible values of X(The number of successes) are $0, 1, 2, 3, \cdots, 100$. $n = 100$, $p = 0.6$, $q = 1 - p = 0.4$. You may use your calculator to calculate the following:

(a) What is $P(X = 60)$?

(b) What is $P(X \leqslant 20)$?

(c) What is $P(X > 70)$?

(d) What is $P(X < 50)$?

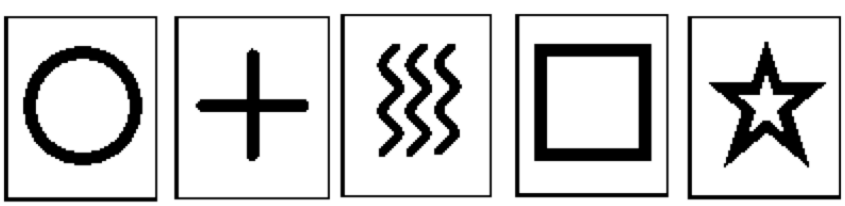
(e) What is $P(50 \leqslant X \leqslant 60)$?

$P(X = 60) = 0.081219145;$
$P(X \leqslant 20) = 3.42043584166077E - 016$
$P(X > 70) = 0.0147753182$
$P(X < 50) = 0.0167616865$
$P(50 \leqslant X \leqslant 60) = 0.5211629726$

Extras : ESP Test `Zener Cards`

One controversial test for ESP involves using a deck of Zener cards. This deck consists of 5 copies of the 5 cards shown below. The tester (sender) shuffles the 25 cards thoroughly, looks at the one on top of the deck and "sends" the information to the "receiver" (the person being tested for E.S.P.), without letting them see the card of course. This process is repeated many times. Many of the early results about the test were controversial because of flaws in how the test was conducted and miscalculation of probabilities `Zener Cards Skepticism`.

In this online version: `Zener Test`, you have to guess 25 cards, each of which is selected randomly by the computer prior to your guess. The actual card will be shown after you click on your chosen symbol. To show evidence of ESP, you need to score at least 10 correct guesses (hits). If the selection of the card is random, this is a Bernoulli Experiment with 25 trials and a probability of $p = 0.2$ of success(correct guess) in each.



(a) What is the expected number of correct answers in the above test if the person taking the test is guessing?

$E(X) = 25 \cdot (0.2) = 2.$

(b) Use your calculator to find $P(X \geqslant 10)$?

0.0173318695

(c) Approximately how many students in a class of 100 would you expect to get a score of 10 or greater on this test by randomly guessing?

$100 \cdot P(X \geqslant 10) = 100 \cdot 0.0173318695 = 1.73318695$ so 1 or 2.

Old Exam questions

Recall the notation

$$C(n,k) = \binom{n}{k}$$

1 An Olympic pistol shooter has a $\frac{2}{3}$ chance of hitting the target at each shot. Find the probability that he will hit exactly 10 targets in a game of 15 shots.

(a) $1-\binom{15}{10}\left(\frac{2}{3}\right)^{10}\left(\frac{1}{3}\right)^{5}$ (b) $\binom{15}{10}\left(\frac{2}{3}\right)^{10}\left(\frac{1}{3}\right)^{5}$ (c) $\left(\frac{2}{3}\right)^{10}$

(d) $\binom{15}{10}\left(\frac{2}{3}\right)^{5}\left(\frac{1}{3}\right)^{10}$ (e) $1 - \left(\frac{1}{3}\right)^{5}$

(b) is "hits exactly 10 targets out of 15 shots"

(d) is "hits exactly 5 targets out of 15 shots"

(a) is "does not hit exactly 10 targets out of 15 shots"

(c) is "hits 10 targets in his first 10 shots"

(e) is "does not miss all 5 of his first 5 shots"

2 A random variable X is the number of successes in a Bernoulli experiment with n trials, each with a probability of success p and a probability of failure q. the probability distribution table of X is shown below:

k	Pr(X = k)
0	$\frac{1}{81}$
1	$\frac{8}{81}$
2	$\frac{24}{81}$
3	$\frac{32}{81}$
4	$\frac{16}{81}$

Which of the following values of n, p, q give rise to this probability distribution?

($\underline{a}$) $n = 4,\ p = \frac{2}{3},\ q = \frac{1}{3}$ (b) $n = 4,\ p = \frac{1}{3},\ q = \frac{2}{3}$ (c) $n = 4,\ p = \frac{1}{6},\ q = \frac{5}{6}$

(d) $n = 5,\ p = \frac{1}{3},\ q = \frac{2}{3}$ (e) $n = 5,\ p = \frac{2}{3},\ q = \frac{1}{3}$

The listed probabilities add up to 1 so they are a probability distribution and therefore $n = 4$. $P(X = 0) = C(4, 0)p^0q^4 = q^4$ so $q = \frac{1}{3}$. Hence (a) is the right answer.

3 Peter is taking a quiz with 6 multiple choice questions. Each question has five options for the answer. Peter, who hasn't studied for the quiz, randomly guesses at each answer. Which of the following gives the probability that Peter gets 2 questions or less correct?

(a) $1 - \left[\binom{6}{0}(.2)^0(.8)^6 + \binom{6}{1}(.2)^1(.8)^5 + \binom{6}{2}(.2)^2(.8)^4\right]$

(b) $\binom{6}{2}(.2)^2(.8)^4$

(c) $\binom{6}{0}(.2)^0(.8)^6 + \binom{6}{1}(.2)^1(.8)^5 + \binom{6}{2}(.2)^2(.8)^4$

(d) $\binom{6}{3}(.2)^3(.8)^3$

(e) $1 - \left[\binom{6}{0}(.2)^0(.8)^6 + \binom{6}{1}(.2)^1(.8)^5\right]$

(b) is "gets exactly 2 right"

(d) is "gets exactly 3 right"

(c) is "gets 0,1 or 2 right"

(a) is "does not get 0, 1 or 2 right"

(e) is "'does not get 0 or 1 right"