

Unit 21: Binomial Distributions



PREREQUISITES

Students should have a background in basic probability and random variables. These topics are covered in Units 18 – 21.

ADDITIONAL TOPIC COVERAGE

Additional information on this topic can be found in *The Basic Practice of Statistics*, Chapter 13, Binomial Distributions.

ACTIVITY DESCRIPTION

Students can work on this activity either individually or in small groups. For this activity, students revisit the context of inheritance of a trait from parents. Instead of the context of sickle cell disease used in the video, the context switches to eye color. Students simulate the outcomes for 30 families with four children in which both parents have brown eyes but carry a recessive gene for blue eyes. Success is a child with blue eyes.

Two methods for simulating the data are provided in the activity. However, you may decide to ask students to use another method. Here are some additional methods:

- Students could have containers with 100 slips of paper, 25 marked success and 75 marked failure.
- Students could use rand on their graphing calculators.
- Students could use statistical software to generate random data from the Bernoulli distribution with $p = 0.25$.

MATERIALS

Two coins of different denominations per student or group of students or access to Excel.
(Optional, depending on choice of method for collecting data)

After students have completed the data collection process and have compiled their data in question 2(a), they should hand in a copy of their first three columns of Table 21.2. Ask them to write their names on their tables. That way you can return them for use in Unit 28, Inference for Proportions. Once you have the individual/group data, combine the results to get a table of class data. Students will need the class data for question 4.

Question 2(b) asks students to find the theoretical probabilities for a binomial distribution with $n = 4$ and $p = 0.25$. This provides an opportunity for students to learn how to use technology to find these probabilities. Statistical software and Excel (BINOM.DIST) have built-in binomial distribution functions. In addition, the TI-84 has a built-in binomial function (binompdf). An alternative is to have students do an online search for binomial calculators. You could also instruct them to use the formulas in the Content Overview to calculate the probabilities and then to check their results using technology.

Students should save their individual/group data and class data from this activity for use in the activity in Unit 28, Inference for Proportions. (Save copies in case students lose theirs.)

THE VIDEO SOLUTIONS

1. Success and failure.
2. Free throws are always shot from the same distance. There is no defensive pressure during play. The probability of making the shot is based on individual player's shooting skills.
3. The probability is 0.25, or 25%. Parents' genetic makeup never changes from child to child so the probability is the same for each child born to these parents.
4. $\mu = np$.
5. The four conditions are listed below.
 1. There are a fixed number of n trials or observations.
 2. The trials are independent.
 3. The trials end in one of two possible outcomes: Success (S) or Failure (F).
 4. The probability of success, p , is the same for all trials.

UNIT ACTIVITY SOLUTIONS

1. Sample answer (student data will differ):

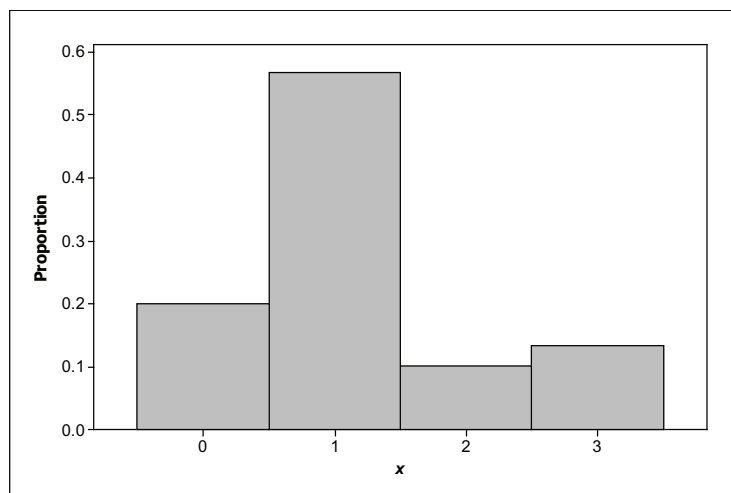
Sample	Trial #				x
	1	2	3	4	
1	F	F	F	S	1
2	F	F	F	F	0
3	F	S	S	F	2
4	F	F	S	F	1
5	F	F	F	F	0
6	S	F	F	F	1
7	S	F	F	F	1
8	S	S	S	F	3
9	F	F	F	S	1
10	F	S	F	F	1
11	S	F	F	F	1
12	S	F	S	S	3
13	F	F	F	F	0
14	S	S	S	F	3
15	F	S	F	F	1
16	F	F	F	S	1
17	F	S	F	S	2
18	F	F	S	F	1
19	F	S	F	S	2
20	F	F	S	F	1
21	F	S	F	F	1
22	F	F	F	S	1
23	F	F	S	F	1
24	S	F	S	F	2
25	F	F	F	S	1
26	F	F	F	F	0
27	F	S	S	S	3
28	F	F	F	F	0
29	F	F	F	F	0
30	F	S	F	F	1

2. a. See solution to (b).

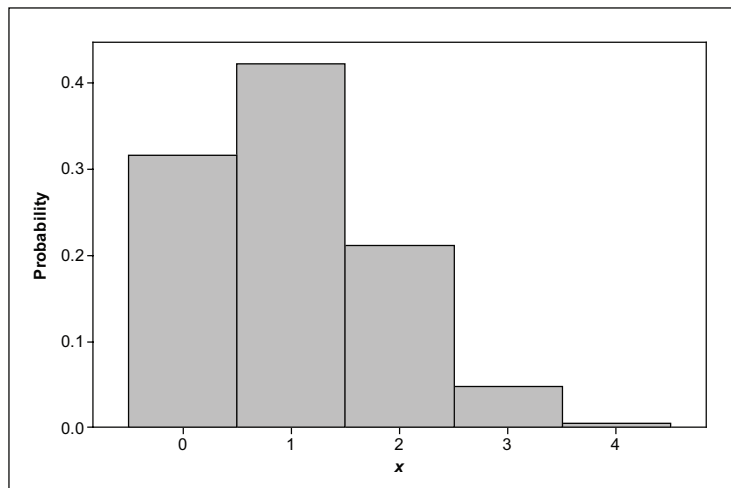
b. Sample answer:

Number of Successes, x	Number of Families with x successes	Proportion of Families with x Successes	Theoretical Probability
0	6	0.2	0.316
1	17	0.567	0.422
2	3	0.1	0.211
3	4	0.133	0.047
4	0	0	0.004

3. a. Sample answer:



b.

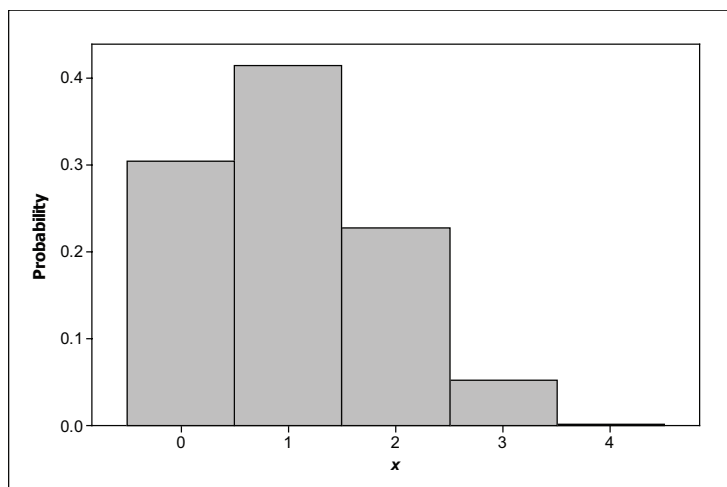


c. Sample answer: They don't look a lot alike. However, the tallest bar is above 1 and the second tallest is above 0 on both graphs.

4. Sample class data is based on 480 samples (from 16 students):

Number of Successes, x	Number of Families with x successes	Proportion of Families with x Successes	Theoretical Probability
0	146	0.304	0.316
1	199	0.415	0.422
2	109	0.227	0.211
3	25	0.052	0.047
4	1	0.002	0.004

a. Sample answer (based on sample class data):



b. The histogram of the class data closely resembles the probability histogram in 3(b).

EXERCISE SOLUTIONS

1. a. Each roll is a trial. Outcomes from one trial do not affect another, so the trials are independent. The probability of success is $p = 1/6$. However, the number of trials is not fixed. So, this is not a binomial setting.

b. This is an example of a binomial setting (at least approximately). Each individual in the sample is a trial. The trials are independent because the sample was randomly selected. There are a fixed number of trials, $n = 10$. If an individual believes in extraterrestrial life, that is a success. The probability of success is 0.60. So x has a binomial distribution with $n = 10$ and $p = 0.6$.

c. This is not an example of a binomial setting. Success is drawing a red card. The probability of drawing the first red card is $1/2$. However, the probability of drawing the second red card is either $26/51$ or $25/51$ but not $1/2$. So, the probability changes as each card is drawn.

2. a. The probabilities are calculated as follows:

$$p(0) = \binom{4}{0} (0.374)^0 (0.626)^4 \approx 0.15$$

$$p(1) = \binom{4}{1} (0.374)^1 (0.626)^3 = (4)(0.374)^1 (0.626)^3 \approx 0.37$$

$$p(2) = \binom{4}{2} (0.374)^2 (0.626)^2 = (6)(0.374)^2 (0.626)^2 \approx 0.33$$

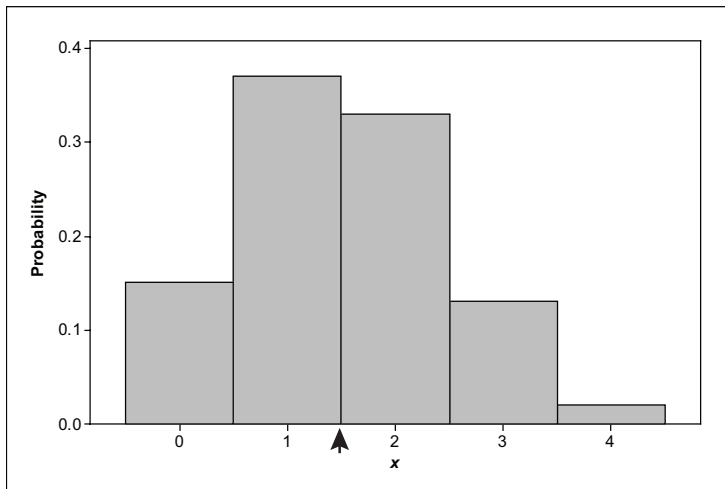
$$p(3) = \binom{4}{3} (0.374)^3 (0.626)^1 = (4)(0.374)^3 (0.626) \approx 0.13$$

$$p(4) = \binom{4}{4} (0.374)^4 (0.626)^0 = (0.374)^4 \approx 0.02$$

$$b. \mu = (0)(0.15) + (1)(0.37) + (2)(0.33) + (3)(0.13) + (4)(0.02) = 1.5$$

c. $\mu = (4)(0.374) = 1.496$. This answer is slightly less than my answer to (b). The discrepancy is due to the fact that the probabilities in (a) were rounded to two decimals.

d. The mean of 0.1496 is marked with an arrow.



3. a. The three probability distributions appear in the table below.

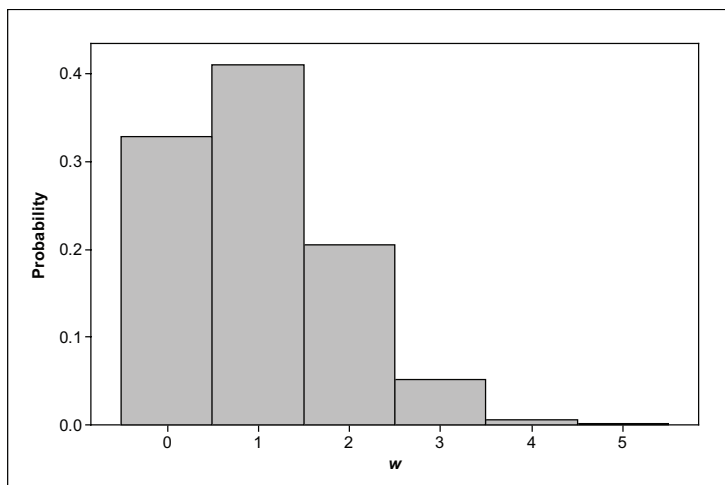
w	0	1	2	3	4	5
$p(w)$	0.3277	0.4096	0.2048	0.0512	0.0064	0.0003
$p(x)$	0.0313	0.1563	0.3125	0.3125	0.1563	0.0313
$p(y)$	0.0003	0.0064	0.0512	0.2048	0.4096	0.3277

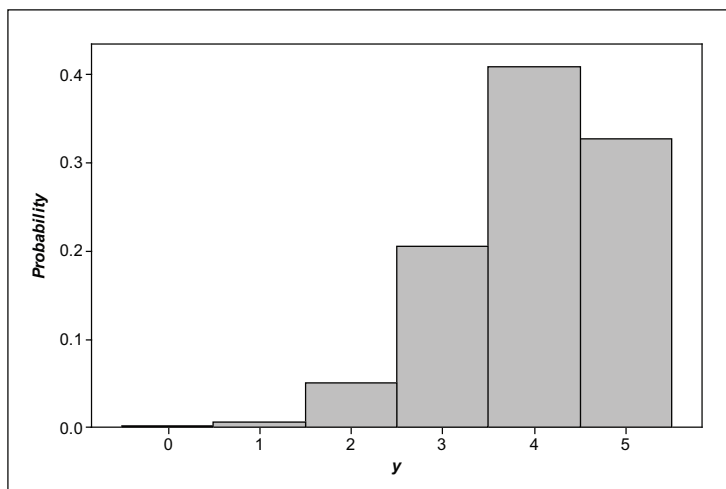
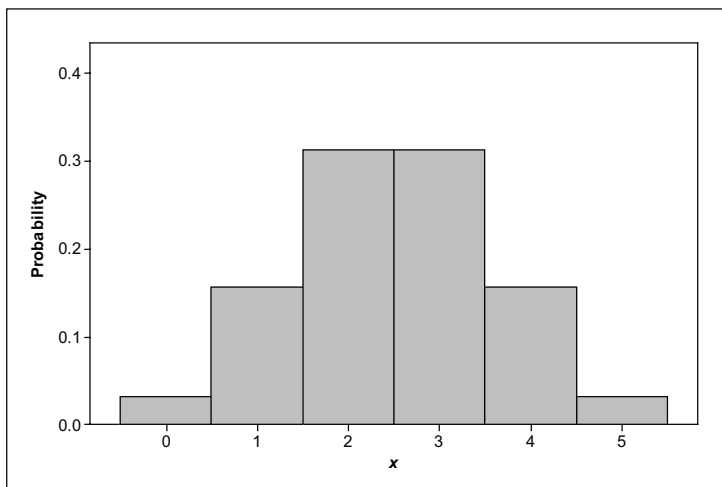
b. $\mu_w = (5)(0.2) = 1.0$; $\sigma_w^2 = (5)(0.2)(0.8) = 0.8$ and $\sigma_w \approx 0.894$

$\mu_x = (5)(0.5) = 2.5$; $\sigma_x^2 = (5)(0.5)(0.5) = 1.25$ and $\sigma_x \approx 1.118$

$\mu_y = (5)(0.8) = 4.0$; $\sigma_y^2 = (5)(0.8)(0.2) = 0.8$ and $\sigma_y \approx 0.894$

c.





d. When $p = 0.2$, the probability histogram is skewed to the right. When $p = 0.5$, the probability histogram is symmetric. When $p = 0.8$ the histogram is skewed to the left. The histograms corresponding to $p = 0.2$ and $p = 0.8$ are mirror images of each other about a vertical line drawn at their means. Hence, the values of the random variables are equally spread about their means for both distributions.

4. a. $b(200, 0.85)$

b. $\mu = (200)(0.85) = 170$; $\sigma = \sqrt{(200)(0.85)(0.15)} = \sqrt{25.5} \approx 5.05$

c. We can approximate the binomial distribution by the normal distribution with $\mu = 170$ and $\sigma = 5.05$. Using statistical software we find, $P(165 < x < 180) \approx 0.8151$.

REVIEW QUESTIONS SOLUTIONS

1. a. In this case, $n = 100$. The sample is randomly drawn so the trials are independent. Let the probability that a randomly selected college student routinely eats breakfast be p – we could estimate this from data, but this value is unknown. It is, however, the same for each of the students in the sample. There are two outcomes to this survey: the student either says yes or he/she does not say yes. So, x is a binomial distribution with $n = 100$ and p unknown.

b. This is not a binomial setting. We could view each accident as a trial, and an accident that involved alcohol as a success. However, the number of trials is not fixed; it varies from week to week.

c. This is a binomial setting. Random variable x has a binomial distribution with $n = 5$ and $p = 0.75$.

$$2. a. p(0) = \binom{5}{0} (0.066)^0 (0.934)^5 = (0.934)^5 \approx 0.71$$

$$b. p(1) = \binom{5}{1} (0.066)^1 (0.934)^4 = (5)(0.066)(0.934)^4 \approx 0.25$$

$$c. P(\text{no more than one has blood type O-}) = P(x = 0 \text{ or } x = 1) = p(0) + p(1) = 0.71 + 0.25 = 0.96$$

$$d. P(\text{at least one has blood type O-}) = 1 - p(0) \approx 1 - 0.71 = 0.29$$

$$3. a. p(0) = \binom{3}{0} (0.25)^0 (0.75)^3 \approx 0.42$$

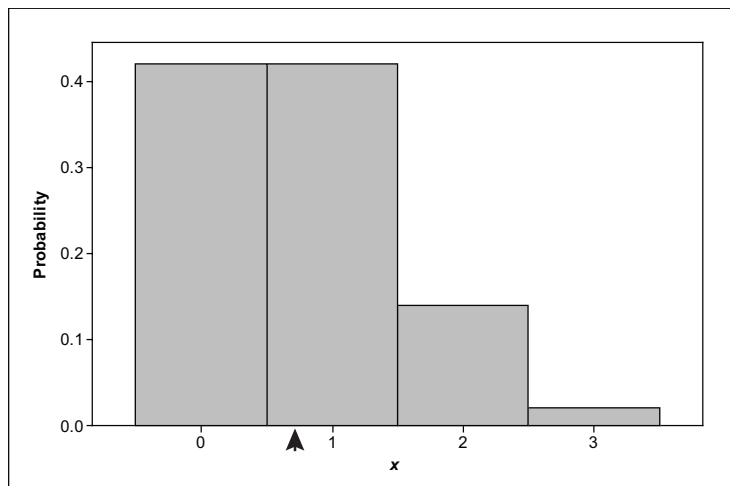
$$p(1) = \binom{3}{1} (0.25)^1 (0.75)^2 \approx 0.42$$

$$p(2) = \binom{3}{2} (0.25)^2 (0.75) \approx 0.14$$

$$p(3) = \binom{3}{3} (0.25)^3 (0.75)^0 \approx 0.02$$

b. $\mu = (3)(0.25) = 0.75$. This means that out of many, many families with three children where both parents are carriers, the average number of children per family who have sickle cell disease is 0.75.

c. The arrow on the horizontal axis marks the mean.



4. a. $b(100, 0.31)$

b. $\mu = (100)(0.31) = 31$; $\sigma = \sqrt{(100)(0.31)(0.69)} \approx 4.62$

c. Approximate the distribution with a normal distribution with mean 31 and standard deviation 4.62. Using technology, $P(x < 25) \approx 0.09702$.