Section 2.7 – Binomial Probability Distribution

Definition

A binomial probability distribution results from a procedure that meets all the following ***requirements***:

1. The procedure has a fixed number of trials.
2. The trials must be independent. (The outcome of any individual trial doesn't affect the probabilities in the other trials.)
3. Each trial must have all outcomes classified into two categories (commonly referred to as success and failure).
4. The probability of a success remains the same in all trials.

Notation for Binomial Probability Distributions

S and *F* (*success* and *failure*) denote the two possible categories of all outcomes; *p* and *q* will denote the probabilities of *S* and *F*, respectively, so

$P(\text{S}) = p$ (p = probability of success)

$P(\text{F}) = 1 - p = q$ (q = probability of failure)

n denotes the fixed number of trials.

x denotes a specific number of successes in *n* trials, so *x* can be any whole number between 0 and *n*, inclusive.

p denotes the probability of success in *one* of the *n* trials.

q denotes the probability of failure in one of the *n* trials.

***P*(*x*)** denotes the probability of getting exactly *x* successes among the *n* trials.

Example

Consider an experiment in which 5 offspring peas are generated each having the green/yellow combination of genes for pod color. The probability of an offspring pea with a green pod is 0.75. That is, $P(\text{green pod}) = 0.75$. Suppose we want to find the probability that exactly 3 of the 5 offspring peas have a green pod.

a) Does this procedure result in a binomial distribution?

b) If this procedure result in a binomial distribution, identify the values of *n, x, p*, and *q*.

Solution

a) 1. The number of trials (5) is fixed.

2. The 5 trials are independent, because the probability of any offspring pea having a green pod is not affected by the outcome of any other offspring pea.

3. Each of the 5 trials has 2 categories of outcomes: The pea has a green pod or it does not.

4. For each offspring pea, the probability that it has a green pod is 0.75, and that probability remains the same for each of the 5 peas.

b) Having concluded that the given procedure does result in a binomial distribution, we now proceed to identify the values of *n, x, p*, and *q*.

1. With 5 offspring peas, we have $n = 5$.
2. We want the probability of exactly 3 peas with green pods, so $x = 3$.
3. The probability of success (getting a pea with a green pod) for one selection is 0.75, so $p = 0.75$.
4. The probability of failure (not getting a green pod) is 0.25, so $q = 0.25$..

Important Hints

- Be sure that *x* and *p* both refer to the same category being called a success.
- When sampling without replacement, consider events to be independent if $n < 0.05N$.

Example

Which of the following are binomial experiments?

a) A player rolls a pair of fair die 10 times. The number *X* of 7's rolled is recorded.

b) The 11 largest airlines had an on-time percentage of 84.7% in November, 2001 according to the Air Travel Consumer Report. In order to assess reasons for delays, an official with the FAA randomly selects flights until she finds 10 that were not on time. The number of flights *X* that need to be selected is recorded.

c) In a class of 30 students, 55% are female. The instructor randomly selects 4 students. The number *X* of females selected is recorded.

Solution

a) Binomial experiment

b) Not a binomial experiment – not a fixed number of trials.

c) Not a binomial experiment – the trials are not independent.

Method 1: Using the Binomial Probability Formula

$$P(x) = \frac{n!}{(n-x)!\,x!} \cdot p^x \cdot q^{n-x} \quad for \quad x = 0,\ 1,\ 2,\ \ldots,\ n$$

$$P(x\ success) = C_{n,x}\, p^x q^{n-x} \quad or \quad P(x\ success) = C_{n,x}\, p^x (1-p)^{n-x}$$

where

n = number of trials

x = number of successes among *n* trials

p = probability of success in any one trial

q = probability of failure in any one trial ($q = 1 - p$)

Example

The probability of an offspring pea with a green pod is 0.75, use the binomial probability formula to find the probability that exactly 3 peas with green pods when of the 5 offspring peas have a green pod when 5 offspring peas are generated. That is, find $P(3)$ given that $n = 5$, $x = 3$, $p = 0.75$, and $q = 0.25$.

Solution

$$P(x=3) = {}_5C_3 \cdot (0.75)^3 \cdot (0.25)^{5-3}$$
$$= 0.263671875$$

The probability of getting exactly 3 peas with green pods among 5 offspring peas is 0.264,

TI-83/84 PLUS Press **2nd VARS** (to get **DISTR,** which denotes "distributions"), then select the option identified as **binompdf(.** Complete the entry of **binompdf(n, p, x)** with specific values for *n*, *p*, and *x*, then press **ENTER.** The result will be the probability of getting *x* successes among *n* trials.

You could also enter **binompdf(n, p)** to get a list of *all* of the probabilities corresponding to $x = 0, 1, 2, \ldots, n$. You could store this list in L2 by pressing **STO** → **L2.** You could then manually enter the values of 0, 1, 2, . . . , *n* in list L1, which would allow you to calculate statistics (by entering **STAT, CALC,** then **L1, L2**) or view the distribution in a table format (by pressing **STAT,** then **EDIT**).

The command **binomcdf** yields *cumulative* probabilities from a binomial distribution. The command **binomcdf(n, p, x)** provides the sum of all probabilities from $x = 0$ through the specific value entered for *x*.

Example

The fast food chain McDonald's has a brand name recognition rate of 95% around the world. Assuming that we randomly select 5 people, use the table to find the following

a) The probability that exactly 3 of the 5 people recognize McDonald's

b) The probability that the number of people who recognize McDonald's is 3 or fewer

n	*x*	*P(x)* =.95
5	0	0+
4	1	0+
3	2	0.001
2	3	0.021
1	4	0.204
0	5	0.774

Solution

a) $P(x=3) = .021$

b) $P(3 \text{ or fewer}) = P(3) + P(2) + P(1) + P(0)$
$$= 0.021 + 0.001 + 0 + 0$$
$$= 0.02259$$

Phrase	***Math Symbol***
"more than" ***or*** "greater than"	$>$
"fewer than" ***or*** "less than"	$<$
"no more than" ***or*** "at most" ***or*** "less than or equal to	$\le$
"at least" ***or*** "no less than" ***or*** "greater than or equal to"	$\ge$
"exactly" ***or*** "equals" ***or*** "is"	$=$

Example

According to the Experian Automotive, 35% of all car-owning households have three or more cars.

a) In a random sample of 20 car-owning households, what is the probability that exactly 5 have three or more cars?

b) In a random sample of 20 car-owning households, what is the probability that less than 4 have three or more cars?

c) In a random sample of 20 car-owning households, what is the probability that at least 4 have three or more cars?

Solution

a) $P(5) = {}_{20}C_5(.35)^5(1-.35)^{20-5}$ $\boxed{=0.1272}$

b) $P(X<4) = P(0)+P(1)+P(2)+P(3)$

$$= {}_{20}C_0(.35)^0(.65)^{20} + {}_{20}C_1(.35)^1(.65)^{19} + {}_{20}C_2(.35)^2(.65)^{18} + {}_{20}C_3(.35)^3(.65)^{17}$$

$\boxed{=0.0444}$

c) $P(X \ge 4) = 1 - P(X \le 3)$

$= 1 - .0444$

$\boxed{=0.9556}$

Example

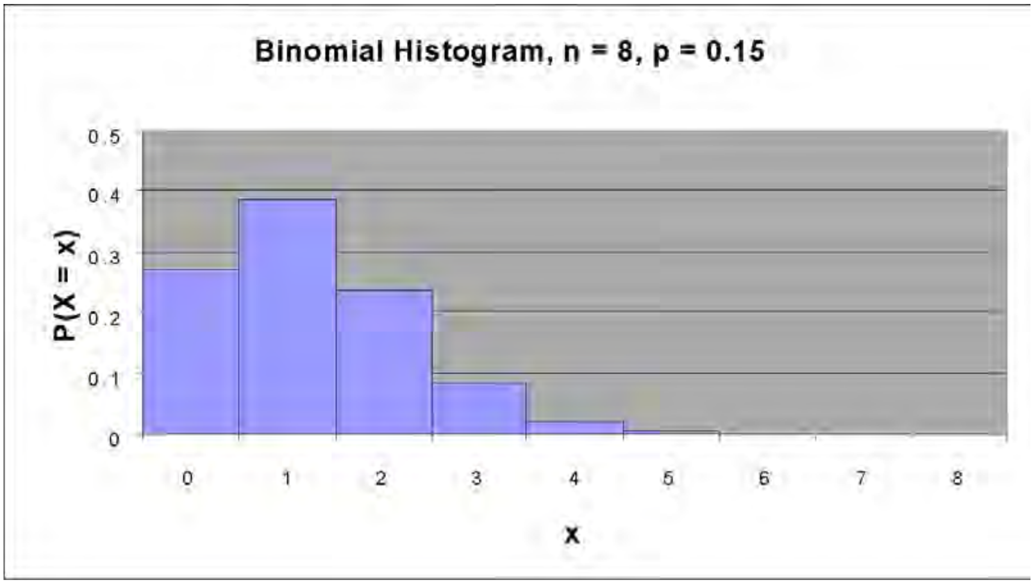
a) Construct a binomial probability histogram with $n = 8$ and $p = 0.15$.

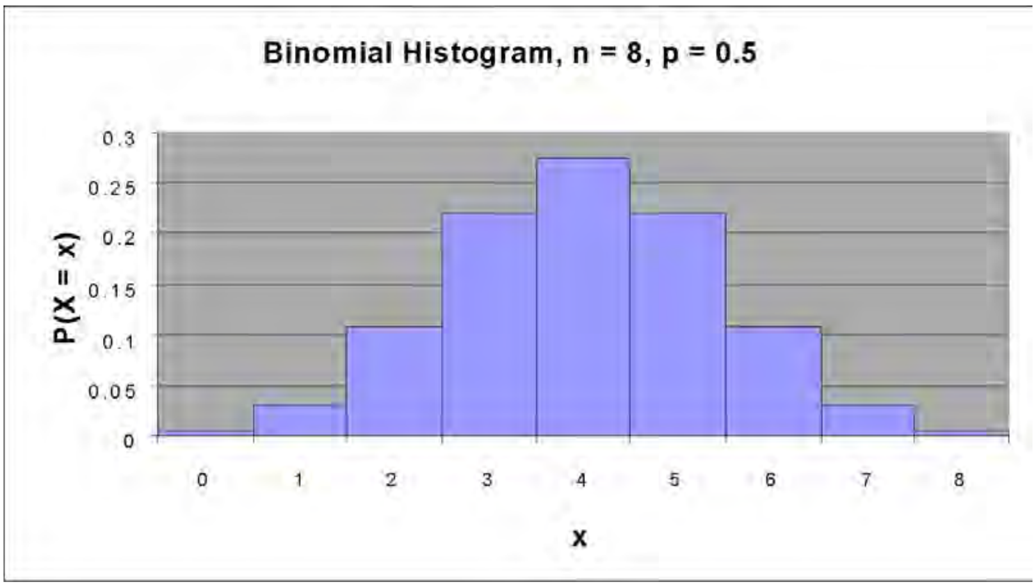
b) Construct a binomial probability histogram with $n = 8$ and $p = 0.5$.

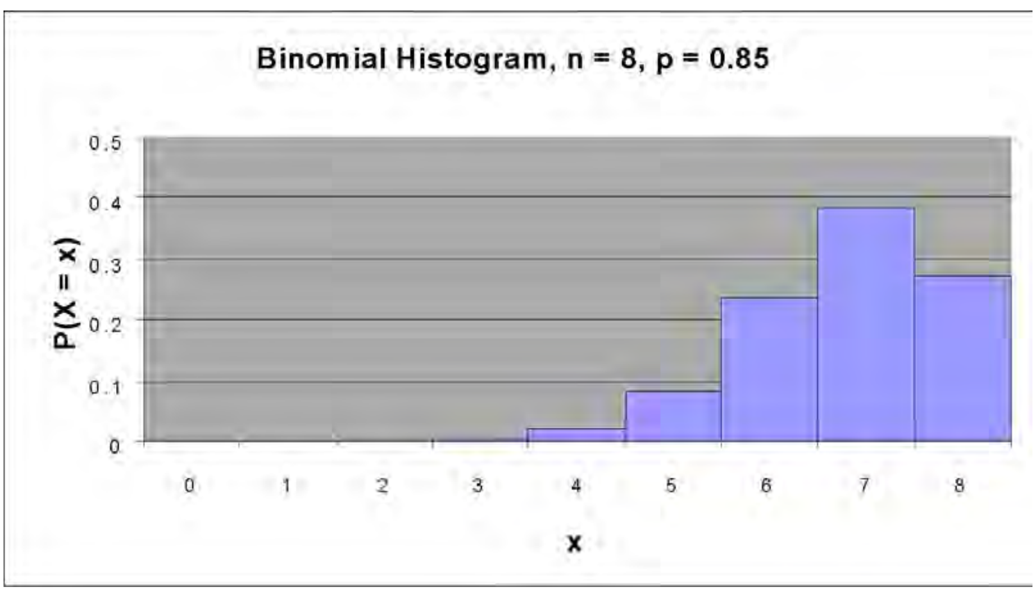
c) Construct a binomial probability histogram with $n = 8$ and $p = 0.85$.

For each histogram, comment on the shape of the distribution.

Solution


Binomial Histogram, n = 8, p = 0.15
P(X = x)
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8
x


Binomial Histogram, n = 8, p = 0.5
P(X = x)
0.3
0.25
0.2
0.15
0.1
0.05
0
0 1 2 3 4 5 6 7 8
x


Binomial Histogram, n = 8, p = 0.85
P(X = x)
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8
x

Mean, Variance, and Standard Deviation

We consider important characteristics of a binomial distribution including center, variation and distribution. That is, given a particular binomial probability distribution we can find its mean, variance and standard deviation.

A strong emphasis is placed on *interpreting* and *understanding* those values.

	Discrete Probability Distribution	***Binomial Distribution***
Mean	$\mu = \sum [x \cdot P(x)]$	$\mu = np$
Variance	$\sigma^2 = \left[\sum x^2 \cdot P(x)\right] - \mu^2$	$\sigma^2 = npq$
Std. Dev.	$\sigma = \sqrt{\left[\sum x^2 \cdot P(x)\right] - \mu^2}$	$\sigma = \sqrt{npq}$

Where

n = number of fixed trials
p = probability of ***success*** in one of the ***n*** trials
q = probability of ***failure*** in one of the ***n*** trials

Interpretation of Results

It is especially important to interpret results. The range rule of thumb suggests that values are unusual if they lie outside of these limits:

$$\text{Maximum usual values} = \boldsymbol{\mu + 2\sigma}$$
$$\text{Minimum usual values} = \boldsymbol{\mu - 2\sigma}$$

Example

Find the mean and standard deviation for the numbers of peas with green pods when groups of 5 offspring peas are generated. Assume that there is 0.75 probability that an offspring pea has a green pod.

Solution

Given: $n = 5;\quad p = 0.75;\quad q = 1 - p = 1 - 0.75 = 0.25$

Mean: $\mu = np = 5(0.75) = 3.75 \approx 3.8$

Standard deviation: $\sigma = \sqrt{npq} = \sqrt{(5)(0.75)(0.25)} = 0.9375 \approx 1$

Example

Mendel generated 580 offspring peas. He claimed that 75% or 4.35, of them would have green pods. The actual experiment resulted in 428 peas with green pods.

a) Assuming that groups of 580 offspring peas are generated find the mean and standard deviation for the numbers of peas with green pods.

b) Use the range rule of thumb to find the minimum usual number and the maximum usual number of peas with green pods. Based on those numbers, can we conclude that Mendel's actual result of 428 peas with green pods is unusual? Does this suggest that Mendel's value of 75% wrong?

Solution

a) ***Given***: $n = 580; \quad p = 0.75; \quad q = 0.25$

Mean: $\mu = np = (580)(0.75) \underline{= 435}$

Standard deviation: $\sigma = \sqrt{npq} = \sqrt{(580)(0.75)(0.25)} \underline{\approx 10.4}$

b) *Maximum usual values* $= \mu + 2\sigma = 435 + 2(10.4) \underline{= 455.8}$

Minimum usual values $= \mu - 2\sigma = 435 - 2(10.4) \underline{= 414.2}$

If Mendel generated groups of 580 offspring peas and if his 75% rate is correct, the numbers of peas with green pods should usually fall between 414.2 and 455.8.
Mendel actually got 428 peas with green pods, and that value does fall within the range of usual values, so the experimental results are consistent with the 75% rate.
The results do not suggest Mendel's claimed rate of 75% is wrong.

Exercises *Section* 2.7 – Binomial Probability Distribution

1. 20 different Senators are randomly selected from the 100 Senators in the current Congress, and each was asked whether he or she is in favor of abolishing estate taxes. Does this procedure result in a binomial distribution, if it is not binomial, identify at least one requirement that is not satisfied?

2. 15 different Governors are randomly selected from the 50 Governors in the currently office and the sex of each Governor is recorded. Does this procedure result in a binomial distribution, if it is not binomial, identify at least one requirement that is not satisfied?

3. 200 statistics students are randomly selected and each asked if he or she owns a TI-84 Plus calculator. Does this procedure result in a binomial distribution, if it is not binomial, identify at least one requirement that is not satisfied?

4. Multiple choice questions on the SAT test have 5 possible answers (*a, b, c, d, e*), 1 of which is correct. Assume that you guess the answers to 3 such questions.

a) Use the multiplication rule to find the probability that the first 2 guesses are wrong and the third is correct. That is, find $P(WWC)$, where C denotes a correct answer and W denotes a wrong answer.

b) Beginning with *WWC*, make a complete list of the different possible arrangements of 2 wrong answers and 1 correct answer, then find the probability for each entry in the list.

c) Based on the proceeding results, what is the probability of getting exactly 1 correct answer when 3 guesses are made?

5. A psychology test consists of multiple choice questions, each having 4 possible answers (*a, b, c, d*), 1 of which is correct. Assume that you guess the answers to 6 such questions.

a) Use the multiplication rule to find the probability that the first 2 guesses are wrong and the last 4 guesses are correct. That is, find $P(WWCCCC)$, where C denotes a correct answer and W denotes a wrong answer.

b) Beginning with *WWCCCC*, make a complete list of the different possible arrangements of 2 wrong answers and 4 correct answers, then find the probability for each entry in the list.

c) Based on the proceeding results, what is the probability of getting exactly 4 correct answers when 6 guesses are made?

6. Use the Binomial Probability Table to find the probability of x success given the probability p of success on a single trial

$a)\ \ n=2,\quad x=1,\quad p=.30$ $\qquad b)\ \ n=5,\quad x=1,\quad p=0.95$

$c)\ \ n=15,\quad x=11,\quad p=0.99$ $\qquad d)\ \ n=14,\quad x=4,\quad p=0.60$

$e)\ \ n=10,\quad x=2,\quad p=0.05$ $\qquad f)\ \ n=12,\quad x=12,\quad p=0.70$

7. Use the Binomial Probability Formula to find the probability of x success given the probability p of success on a single trial

$a)\quad n=12,\quad x=10,\quad p=\frac{3}{4}$ $\qquad$ $b)\quad n=9,\quad x=2,\quad p=0.35$

$c)\quad n=20,\quad x=4,\quad p=0.15$ $\qquad$ $d)\quad n=15,\quad x=13,\quad p=\frac{1}{3}$

8. In the US, 40% of the population have brown eyes. If 14 people are randomly selected, find the probability that at least 12 of them have brown eyes. Is it unusual to randomly select 14 people and find that at least 12 of them have brown eyes? Why or why not?

9. When blood donors were randomly selected, 45% of them had blood that is Group *O*. The display shows that the probabilities obtained by entering the values of $n = 5$ and $p = 0.45$.

x	***P*(*x*)**
0	0.050328
1	0.205889
2	0.336909
3	0.275653
4	0.112767
5	0.018453

a) Find the probability that at least 1 of the 5 donors has Group *O* blood. If at least 1 Group *O* donor is needed, is it reasonable to expect that at least 1 will be obtained?

b) Find the probability that at least 3 of the 5 donors have Group *O* blood. If at least 3 Group *O* donors are needed, is it very likely to expect that at least 3 will be obtained?

c) Find the probability that all donors have Group *O* blood. Is it unusual to get 5 Group *O* donors from 5 randomly selected donors? Why or Why not?

d) Find the probability that at most 2 of the 5 donors have Group *O* blood.

10. There is 1% delinquency rate for consumers with FICO credit rating scores above 800. If a bank provides large loans 12 people with FICL scores above 800, what is the probability that at least one of them becomes delinquent? Based on that probability, should the bank plan on dealing with a delinquency?

11. Ten peas are generated from parents having the green/yellow pair of genes, so there is a 0.75 probability that an individual pea will have a green pod. Find the probability that among the 10 offspring peas, at least 9 have green pods. Is it unusual to get at least 9 peas with green pods when 10 offspring peas are generated? Why or why not?

12. You purchased a slot machine configured so that there is a $\frac{1}{2,000}$ probability of winning the jackpot on any individual trial. Although no one would seriously consider tricking the author, suppose that a guest claims that she played the slot machine 5 times and hit the jackpot twice

a) Find the probability of exactly 2 jackpots in 5 trials.

b) Find the probability of at least 2 jackpots in 5 trials.

c) Does the guest's claim of hitting 2 jackpots in 5 trials seem valid? Explain.

13. In a survey of 320 college graduates, 36% reported that they stayed on their first full-time job less than one year.

a) If 15 of those survey subjects are randomly selected without replacement for a follow-up survey, find the probability that 5 of them stayed on their first full-time job less than one year.

b) If part (a) is changed so that 20 different survey subjects are selected, explain why the binomial probability formula *cannot* be used.

14. In a survey of 150 senior executives, 47% said that the most common job interview mistake is to have little or no knowledge of the company.

a) If 6 of those surveyed executives are randomly selected without replacement for a follow-up survey, find the probability that 3 of them said that the most common job interview mistake is to have little or no knowledge of the company.

b) If part (a) is changed so that 9 different surveyed executives are selected, explain why the binomial probability formula *cannot* be used.

15. In a Gallup poll of 1236 adults, it was found that 5% of those polled said that bad luck occurs after breaking a mirror. Based on these results, such randomly selected groups of 1236 adults will have a mean of 61.8 people with that belief, and a standard deviation of 7.7 people. What is the variance?

16. Random guesses are made for 50 SAT multiple choice questions, so $n = 50$ and $p = 0.2$.

a) Find the mean μ and standard deviation σ.

b) Use the range rule of thumb to find the minimum usual number and the maximum usual number.

17. In an analysis of test result from the YSORT gender selection method, 152 babies are born and it is assumed that boys and girls are equally likely, so $n = 152$ and $p = 0.5$.

a) Find the mean μ and standard deviation σ.

b) Use the range rule of thumb to find the minimum usual number and the maximum usual number.

18. In a Gallup poll of 1236 adults, it showed that 145% believe that bad luck follows if your path is crossed by a black car, so $n = 1236$ and $p = 0.14$.

a) Find the mean μ and standard deviation σ.

b) Use the range rule of thumb to find the minimum usual number and the maximum usual number.

19. The midterm exam in a nursing course consists of 75 true/false questions. Assume that an unprepared student makes random guesses for each of the answers.

a) Find the mean and standard deviation for the number of correct answers for such students.

b) Would it be unusual for a student to pass this exam by guessing and getting at least 45 correct answers? Why or why not?

20. The final exam in a nursing course consists of 100 multiple-choice questions. Each question has 5 possible answers, and only 1 of them is correct. An unprepared student makes random guesses for all of the answers.

a) Find the mean and standard deviation for the number of correct answers for such students.

b) Would it be unusual for a student to pass this exam by guessing and getting at least 60 correct answers? Why or why not?

21. In a test of the XSORT method of gender selection, 574 babies are born to couples trying to have baby girls, and 525 of those babies are girls.

a) If the gender-selection method has no effect and boys and girls are equally likely, find the mean and standard deviation for the numbers of girls born in groups of 574.

b) Is the result of 525 girls unusual? Does it suggest that the gender-selection method appears to be effective?

22. In a test of the YSORT method of gender selection, 152 babies are born to couples trying to have baby boys, and 127 of those babies are boys.

a) If the gender-selection method has no effect and boys and girls are equally likely, find the mean and standard deviation for the numbers of boys born in groups of 152.

b) Is the result of 127 boys unusual? Does it suggest that the gender-selection method appears to be effective?

23. A headline in USA Today states that "most stay at first job less than 2 years." That headline is based on a poll of 320 college graduates. Among those polled, 78% stayed at their full-time job less than 2 years.

a) Assuming that 50% is the true percentage of graduates who stay at their first job less than 2 years, find the mean and the standard deviation of the numbers of such graduates in randomly selected groups of 320 graduates.

b) Assuming that the 50% rate in part (*a*) is correct; find the range of usual values for the numbers of graduates among 320 who stay at their first job less than 2 years.

c) Find the actual number of surveyed who stayed at their first job less 2 years. Use the range of values from part (*b*) to determine whether that number is unusual. Does the result suggest that the headline is not justified?

d) This statement was given as part of the description of the survey methods used: "Alumni who opted-in to receive communications from Experience were invited to participate in the online poll, and 320 of them completed the survey." What does that statement suggest about the result?

24. In a study of 420,095 cell phone users in Denmark, it was found that 135 developed cancer of the brain or nervous system. If we assume that the use of cell phones has no effect on developing such cancer, then the probability of a person having such cancer is 0.000340.

a) Assuming that the cell phones have no effect on developing cancer, find the mean and the standard deviation of the numbers of people in groups of 420,095 that can be expected to have cancer of the brain or nervous system.

b) Based on the result from part (*a*), is it unusual to find that among 420,095 people, there are 135 cases of cancer of the brain or nervous system? Why or why not?

c) What do these results suggest about the publicized concern that cell phones are a health danger because they increase the risk of cancer of the brain or nervous system?

25. Mario's Pizza Parlor has just opened. Due to a lack of employee training, there is only a 0.8 probability that a pizza will be edible. An order for 5 pizzas has just been placed. What is the minimum number of pizzas that must be made in order to be at least 99% sure that there will be 5 that are edible?