Section 5.2

5.43 Should you use the binomial distribution? In each of the following situations, is it reasonable to use a binomial distribution for the random variable $X$? Give reasons for your answer in each case. If a binomial distribution applies, give the values of $n$ and $p$.

(a) A poll of 200 college students asks whether or not you are usually irritable in the morning. $X$ is the number who reply that they are usually irritable in the morning.

(b) You toss a fair coin until a head appears. $X$ is the count of the number of tosses that you make.

(c) Most calls made at random by sample surveys don’t succeed in talking with a live person. Of calls to New York City, only one-twelfth succeed. A survey calls 500 randomly selected numbers in New York City. $X$ is the number of time that a live person is reached.

(d) You deal 10 cards from a shuffled deck and count the number $X$ of black cards.

5.44 Should you use the binomial distribution? In each of the following situations, is it reasonable to use a binomial distribution for the random variable $X$? Give reasons for your answer in each case.

(a) In a random sample of students in a fitness study, $X$ is the mean systolic blood pressure of the sample.

(b) A manufacturer of running shoes picks a random sample of the production of shoes each day for a detailed inspection. Today’s sample of 20 pairs of shoes includes one pair with a defect.

(c) A nutrition study chooses an SRS of college students. They are asked whether or not they usually eat at least five servings of fruits or vegetables per day. $X$ is the number who say that they do.

5.45 Typographic errors. Typographic errors in a text are either nonword errors (as when “the” is typed as “teh”) or word errors that result in a real but incorrect word. Spell-checking software will catch nonword errors but not word errors. Human proofreaders catch 70% of word errors. You ask a fellow student to proofread an essay in which you have deliberately made 10 word errors.

(a) If the student matches the usual 70% rate, what is the distribution of the number of errors caught? What is the distribution of the number of errors missed?

(b) Missing 4 or more out of 10 errors seems a poor performance. What is the probability that a proofreader who catches 70% of word errors misses 4 or more out of 10?
Solution

(a)  Distribution of errors caught $X$  

$X = B(10, 0.70)$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.0001</td>
</tr>
<tr>
<td>2</td>
<td>0.0014</td>
</tr>
<tr>
<td>3</td>
<td>0.0090</td>
</tr>
<tr>
<td>4</td>
<td>0.0368</td>
</tr>
<tr>
<td>5</td>
<td>0.1029</td>
</tr>
<tr>
<td>6</td>
<td>0.2001</td>
</tr>
<tr>
<td>7</td>
<td>0.2668</td>
</tr>
<tr>
<td>8</td>
<td>0.2335</td>
</tr>
<tr>
<td>9</td>
<td>0.1211</td>
</tr>
<tr>
<td>10</td>
<td>0.0282</td>
</tr>
</tbody>
</table>

(b)  $P[ Y \geq 4 ] = 0.2001 + 0.1029 + 0.0368 + 0.0090 + 0.0014 + 0.0001 + 0.0000 = 0.3504$

Note that this is the same as  $P[X \leq 6]$
5.47 Typographic errors. Return to the proofreading setting of Exercise 5.45.

(a) What is the mean number of errors caught? What is the mean number of errors missed? You see that these two means must add to 10, the total number of errors.

(b) What is the standard deviation \( \sigma \) of the number of errors caught?

(c) Suppose that a proofreader catches 90% of word errors, so that \( p = 0.9 \). What is \( \sigma \) in this case? What is \( \sigma \) if \( p = 0.99 \)? What happens to the standard deviation of a binomial distribution as the probability of a success gets close to 1?

Solution

(a) \( \mu_x = np = 10 \times 0.7 = 7 \)
\[ \mu_y = np = 10 \times 0.3 = 3 \]

(b) \( \sigma_x = \sqrt{np(1-p)} = \sqrt{10 \times 0.7 \times 0.3} = 1.449138 \)
\[ \sigma_y = \sqrt{np(1-p)} = \sqrt{10 \times 0.3 \times 0.7} = 1.449138 \]

(c) if \( p = 0.9 \Rightarrow \sigma_x = \sigma_y = \sqrt{10 \times 0.9 \times 0.1} = 0.948683 \)
if \( p = 0.99 \Rightarrow \sigma_x = \sigma_y = \sqrt{10 \times 0.99 \times 0.01} = 0.314643 \)

Standard deviation gets closer to 0.

5.53 Inheritance of blood types. Children inherit their blood type from their parents, with probabilities that reflect the parents’ genetic makeup. Children of Juan and Maria each have probability \( \frac{1}{4} \) of having blood type A and inherit independently of each other. Juan and Maria plan to have 4 children; let \( X \) be the number who have blood type A.

(a) What are \( n \) and \( p \) in the binomial distribution of \( X \)?

(b) Find the probability of each possible value of \( X \), and draw a probability histogram for this distribution.

(c) Find the mean number of children with type A blood, and mark the location of the mean on your probability histogram.

Solution

Work like 5.45 above
5.57 Shooting free throws. Since the mid-1960s, the overall free throw percent at all college levels, for both men and women, has remained pretty consistent. For men, players have shot roughly 69% with the season percent never falling below 67% or larger than 70%. Assume that 300,000 free throws will be attempted in the upcoming season.

(a) What is the mean and standard deviation of if the population proportion is \( p = 0.69 \)?

(b) Using the 68–95–99.7 rule, we’d expect \( \hat{p} \) to fall between what two percents about 95% of the time?

(c) Given the width of this interval in part (b) and the range of season percents, do you feel it reasonable to assume the population proportion has been the same over the last 50 seasons? Explain your answer.

Solution
(a) \( \mu_{\hat{p}} = p = 0.69 \)
\[ \sigma_{\hat{p}} = \sqrt{\frac{p*(1 - p)/n}{n}} = \sqrt{(0.69*(1-0.69)/300000)} = 0.000844 \]

(b) \( \mu_{\hat{p}} \pm 2\sigma_{\hat{p}} = 0.69 \pm 2(0.0.000844) = 0.69 \pm 0.001689 \)
i.e., in the interval (0.688311, 0.691689) or (68.83%, 69.17%)

(c) It is reasonable to assume that the population proportion over the last 50 years has not change.

5.58 Online learning. Recently the U.S. Department of Education released a report on online learning stating that blended instruction, a combination of conventional face-to-face and online instruction, appears more effective in terms of student performance than conventional teaching. You decide to poll the incoming students at your institution to see if they prefer courses that blend face-to-face instruction with online components. In an SRS of 400 incoming students, you find 294 prefer this type of course.

(a) What is the sample proportion who prefer this type of blended instruction?

(b) Suppose the population proportion for all students nationwide is 80%. What is the standard deviation of \( \hat{p} \)?

(c) Using the 68–95–99.7 rule, if you had drawn an SRS from the United States, you would expect \( \hat{p} \) to fall between what two percents about 95% of the time?

(d) Based on your result in part (a), do you feel the incoming students at your institution prefer this type of instruction more, less, or about the same as students nationally? Explain your answer.
5.61 A test for ESP. In a test for ESP (extrasensory perception), the experimenter looks at cards that are hidden from the subject. Each card contains either a star, a circle, a wave, or a square. As the experimenter looks at each of 20 cards in turn, the subject names the shape on the card.

(a) If a subject simply guesses the shape on each card, what is the probability of a successful guess on a single card? Because the cards are independent, the count of successes in 20 cards has a binomial distribution.

(b) What is the probability that a subject correctly guesses at least 10 of the 20 shapes?

(c) In many repetitions of this experiment with a subject who is guessing, how many cards will the subject guess correctly on the average? What is the standard deviation of the number of correct guesses?

(d) A standard ESP deck actually contains 25 cards. There are five different shapes, each of which appears on 5 cards. The subject knows that the deck has this makeup. Is a binomial model still appropriate for the count of correct guesses in one pass through this deck? If so, what are \( n \) and \( p \)? If not, why not?

5.67 Multiple-choice tests. Here is a simple probability model for multiple-choice tests.
Suppose that each student has probability \( p \) of correctly answering a question chosen at random from a universe of possible questions. (A strong student has a higher \( p \) than a weak student.) The correctness of an answer to a question is independent of the correctness of answers to other questions. Jodi is a good student for whom \( p = 0.88 \).

(a) Use the Normal approximation to find the probability that Jodi scores 85% or lower on a 100-question test.

(b) If the test contains 250 questions, what is the probability that Jodi will score 85% or lower?

(c) How many questions must the test contain in order to reduce the standard deviation of Jodi’s proportion of correct answers to half its value for a 100-item test?

(d) Laura is a weaker student for whom \( p = 0.72 \). Does the answer you gave in part (c) for the standard deviation of Jodi’s score apply to Laura’s standard deviation also?