Homework 5 - Probability

Christopher Munoz

3.35 Consider the population of voters described in Example 3.6. Suppose that there are N = 5000 voters in the population, 40% of whom favor Jones. Identify the event favors Jones as a success S. It is evident that the probability of S on trial 1 is .40. Consider the event B that S occurs on the second trial. Then B can occur two ways: The first two trials are both successes or the first trial is a failure and the second is a success. Show that P(B) = .4. What is P(B| the first trial is S)? Does this conditional probability differ markedly from P(B)?

$$P(B) = P(SS) + P(FS) = 0.4^2 + 0.6(0.4) = 0.4$$

3.37 In 2003, the average combined SAT score (math and verbal) for college-bound students in the United States was 1026. Suppose that approximately 45% of all high school graduates took this test and that 100 high school graduates are randomly selected from among all high school grads in the United States. Which of the following random variables has a distribution that can be approximated by a binomial distribution? Whenever possible, give the values for n and p.

a The number of students who took the SAT

Not Binomial

b The scores of the 100 students in the sample

Not Binomial

c The number of students in the sample who scored above average on the SAT

Binomial with n = 100, p = students who scored above 1026

d The amount of time required by each student to complete the SAT

Not Binomial

e The number of female high school grads in the sample

Not Binomial

3.39 A complex electronic system is built with a certain number of backup components in its subsystems. One subsystem has four identical components, each with a probability of .2 of failing in less than 1000 hours. The subsystem will operate if any two of the four components are operating. Assume that the components operate independently. Find the probability that

a exactly two of the four components last longer than 1000 hours.

$$\begin{aligned} n &= 4 \\ p &= .2 \\ P(Y = K) &= C(n, k)p^k(1-p)^{n-k} \\ P(Y = 2) &= C(4, 2) * 0.2^2(.8)^{4-2} \\ &= 6 * .2^2 * .8^2 = 0.1536 \end{aligned}$$

b the subsystem operates longer than 1000 hours.

$$P(Y \geq 2) = 0.1536 + 0.4096 + 0.4096 = 0.9728$$

3.41 A multiple-choice examination has 15 questions, each with five possible answers, only one of which is correct. Suppose that one of the students who takes the examination answers each of the questions with an independent random guess. What is the probability that he answers at least ten questions correctly?

$$P(Y \geq 10) = \sum_{k=10}^{15} \binom{15}{k} (0.2)^k (0.8)^{15-k} = 0.000$$

(Used the Table)

3.42 Refer to Exercise 3.41. What is the probability that a student answers at least ten questions correctly if a for each question, the student can correctly eliminate one of the wrong answers and subsequently answers each of the questions with an independent random guess among the remaining answers?

$$\begin{aligned} n &= 15 \\ p &= .25 \\ P(Y \geq 10) &= 1 - 1.000 = 0.000 \end{aligned}$$

b he can correctly eliminate two wrong answers for each question and randomly chooses from among the remaining answers?

$$\begin{aligned} n &= 15 \\ p &= 1/3 \\ P(Y \geq 10) &= 1 - P(Y \leq 9) = 0.00085 \end{aligned}$$

3.43 Many utility companies promote energy conservation by offering discount rates to consumers who keep their energy usage below certain established subsidy standards. A recent EPA report notes that 70% of the island residents of Puerto Rico have reduced their electricity usage sufficiently to qualify for discounted rates. If five residential subscribers are randomly selected from San Juan, Puerto Rico, find the probability of each of the following events: a All five qualify for the favorable rates.

$$\begin{aligned} p &= 0.7 \\ n &= 5 \\ P(Y = 5) &= C(5, 5)0.7^5 * (1 - 0.7)^{5-5} = 0.7^5 = 0.16807 \end{aligned}$$

b At least four qualify for the favorable rates.

$$\begin{aligned} P(Y \geq 4) = P(Y = 4) + P(Y = 5) &= 0.16807 + 5 * 0.7^4 * (0.3)^1 \\ &= 0.16807 + 36015 \\ &= 0.52822 \end{aligned}$$

3.44 A new surgical procedure is successful with a probability of p. Assume that the operation is performed five times and the results are independent of one another. What is the probability that a all five operations are successful if p = .8?

$$\begin{aligned} p &= 0.8 \\ n &= 5 \\ P(Y = 5) &= C(5, 5) * 0.8^5 * (0.2)^0 = 0.32768 \end{aligned}$$

b exactly four are successful if p = .6?

$$P(Y = 4) = C(5, 4) * 0.6^4 * 0.4 = 0.2592$$

c less than two are successful if p = .3?

$$\begin{aligned} P(Y < 2) = P(Y = 1) + P(Y = 0) &= C(5, 1) * 0.3^1 * 0.7^4 + C(5, 0) * 0.3^0 * 0.7^5 \\ &= 0.52822 \end{aligned}$$

3.45 A fire-detection device utilizes three temperature-sensitive cells acting independently of each other in such a manner that any one or more may activate the alarm. Each cell possesses a probability of p = .8 of activating the alarm when the temperature reaches 100 ∘ Celsius or more. Let Y equal the number of cells activating the alarm when the temperature reaches 100 ∘ .

a Find the probability distribution for Y .

$$\begin{aligned} n &= 3 \\ p &= .8 \\ P(Y = 1) = C(3, 1) * 0.8^1 * (0.2)^2 &= 0.096 \\ P(Y = 2) = C(3, 2) * 0.8^2 * (0.2)^1 &= 0.384 \\ P(Y = 3) = C(3, 3) * 0.8^3 * (0.2)0 &= 0.512 \end{aligned}$$

b Find the probability that the alarm will function when the temperature reaches 100

$$1 - P(Y - 0) = 1 - C(3, 0) * 1 * (0.2)^3 = 0.992$$

∘ .

3.46 Construct probability histograms for the binomial probability distributions for n = 5, p = .1, .5, and .9. (Table 1, Appendix 3, will reduce the amount of calculation.) Notice the symmetry for p = .5 and the direction of skewness for p = .1 and .9

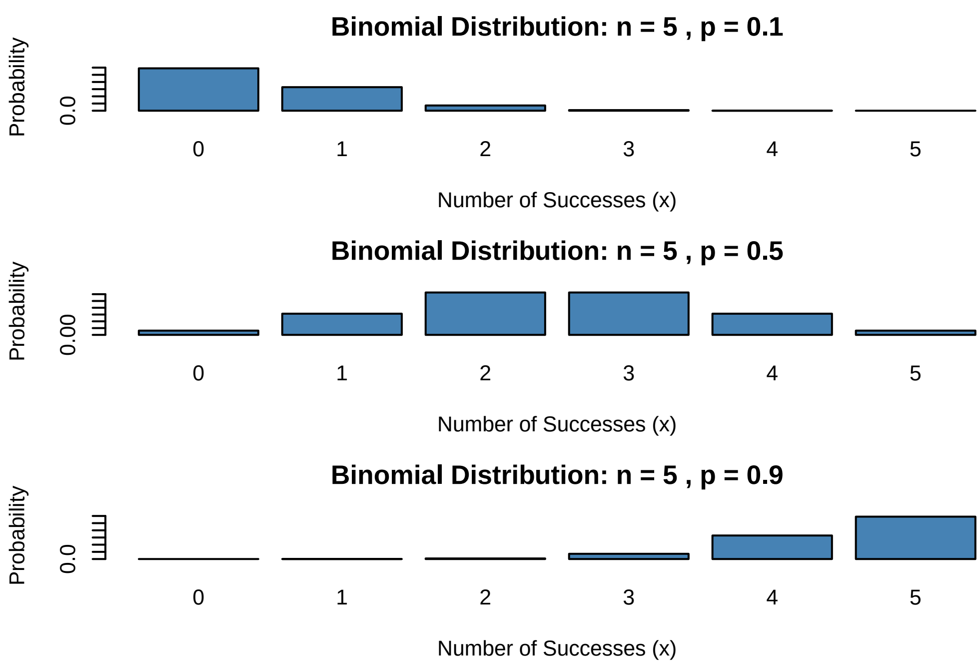
```r
n <- 5
p_values <- c(0.1, 0.5, 0.9)

par(mfrow = c(3, 1), mar = c(4, 4, 3, 2))

for (p in p_values) {
  x <- 0:n
  prob <- dbinom(x, size = n, prob = p)

  # Create barplot (histogram style)
```

```
  barplot(prob,
          names.arg = x,
          main = paste("Binomial Distribution: n =", n, ", p =", p),
          xlab = "Number of Successes (x)",
          ylab = "Probability",
          col = "steelblue",
          border = "black",
          ylim = c(0, max(prob) * 1.1))
}
```




```
par(mfrow = c(1, 1))
```

3.48 A missile protection system consists of n radar sets operating independently, each with a probability of .9 of detecting a missile entering a zone that is covered by all of the units.
a If n = 5 and a missile enters the zone, what is the probability that exactly four sets detect the missile? At least one set?

$$
\begin{aligned}
n &= 5 \\
p &= 0.9 \\
P(Y=4) = C(5,4) * 0.9^4 * (0.1)^1 &= 0.32805 \\
P(Y \geq 1) = 1 - P(Y=0) &= 0.99999
\end{aligned}
$$

b How large must n be if we require that the probability of detecting a missile that enters the zone be .999?

$$n = 3, \text{ Refer to the table in back}$$

3.51 In the 18th century, the Chevalier de Mere asked Blaise Pascal to compare the probabilities of two events. Below, you will compute the probability of the two events that, prior to contrary gambling experience, were thought by de Mere to be equally likely.
a What is the probability of obtaining at least one 6 in four rolls of a fair die?

$$p = 1/6 = .166666$$
$$n = 4$$
$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - C(4,0) * (1/6)^0 * (5/6)^4 = 0.5177469136$$

b If a pair of fair dice is tossed 24 times, what is the probability of at least one double six?

$$p = 1/6 * 1/6 = 1/36$$
$$n = 24$$
$$P(Y \geq 1) = 1 - P(Y = 0) = C(24,0) * (1/36)^0 * (35/36)^2 4 = 0.4914038761$$

3.53 Tay-Sachs disease is a genetic disorder that is usually fatal in young children. If both parents are carriers of the disease, the probability that their offspring will develop the disease is approxi- mately .25. Suppose that a husband and wife are both carriers and that they have three children. If the outcomes of the three pregnancies are mutually independent, what are the probabilities of the following events?
a All three children develop Tay-Sachs.

$$p = .25$$
$$n = 3$$
$$P(Y = 3) = C(3,3) * .25^3 * .75^0 = 0.015625$$

b Only one child develops Tay-Sachs.

$$P(Y = 1) = C(3,1) * .25^1 * .75^2 = 0.421875$$

c The third child develops Tay-Sachs, given that the first two did not.

$$P(A|B) = .25$$

3.56 An oil exploration firm is formed with enough capital to finance ten explorations. The probability of a particular exploration being successful is .1. Assume the explorations are independent. Find the mean and variance of the number of successful explorations.

$$\begin{aligned} n &= 10 \\ p &= 0.1 \\ E(Y) &= np = 1 \\ V(Y) &= np(1-p) = 1(1-0.1) = 0.9 \end{aligned}$$

3.57 Refer to Exercise 3.56. Suppose the firm has a fixed cost of $20,000 in preparing equipment prior to doing its first exploration. If each successful exploration costs $30,000 and each unsuccessful exploration costs $15,000, find the expected total cost to the firm for its ten explorations.

$$\begin{aligned} P(Y) &= 20,000 + 30,000Y + 15,000(10-Y) \\ &= 20,000 + 30,000Y + 150,000 - 15,000Y \\ &= 170,000 + 15,000Y \\ E(1) &= 170,000 + 15,000(1) \\ &= \$185,000 \end{aligned}$$

3.66 Suppose that Y is a random variable with a geometric distribution. Show that a y p(y) = ∞ y=1 q y−1 p = 1.

$$\begin{aligned} \sum_{y=1}^{\infty} p(y) &= \sum_{y=1}^{\infty} q^{y-1}p \\ &= p\sum_{y=1}^{\infty} q^{y-1} \\ &= p(1+q+q^2+q^3+...) \\ &= p \cdot \frac{1}{1-q} \\ &= p \cdot \frac{1}{p} \\ &= 1 \end{aligned}$$

b p(y) p(y − 1) = q, for y = 2, 3, ... This ratio is less than 1, implying that the geomet- ric probabilities are monotonically decreasing as a function of y. If Y has a geometric distribution, what value of Y is the most likely (has the highest probability)?

$$\begin{aligned}\frac{p(y)}{p(y-1)} &= \frac{q^{y-1}p}{q^{y-2}p} \\ &= \frac{q^{y-1}}{q^{y-2}} \\ &= q < 1\end{aligned}$$

3.67 Suppose that 30% of the applicants for a certain industrial job possess advanced training in com- puter programming. Applicants are interviewed sequentially and are selected at random from the pool. Find the probability that the first applicant with advanced training in programming is found on the fifth interview

$$\begin{aligned}p &= 0.3 \\ q &= 0.7 \\ Y &= 5 \\ P(Y=5) &= q^{y-1}p \\ &= (0.7)^4(0.3) \\ &= 0.2401 \times 0.3 \\ &= 0.07203\end{aligned}$$

3.71 Let Y denote a geometric random variable with probability of success p.
a Show that for a positive integer a
P(Y > a) = qa .

$$\begin{aligned}P(Y > a) &= \sum_{y=a+1}^{\infty} q^{y-1}p \\ &= p\sum_{y=a+1}^{\infty} q^{y-1} \\ &= p \cdot q^a \sum_{j=0}^{\infty} q^j \\ &= p \cdot q^a \cdot \frac{1}{1-q} \\ &= p \cdot q^a \cdot \frac{1}{p} \\ &= q^a\end{aligned}$$

b Show that for positive integers a and b,
P(Y > a + b|Y > a) = qb = P(Y > b).

This result implies that, for example, P(Y > 7|Y > 2) = P(Y > 5). Why do you think this property is called the memoryless property of the geometric distribution?

$$
\begin{aligned}
P(Y > a+b|Y > a) &= \frac{P(Y > a+b \cap Y > a)}{P(Y > a)} \\
&= \frac{P(Y > a+b)}{P(Y > a)} \\
&= \frac{q^{a+b}}{q^a} \\
&= q^b \\
&= P(Y > b)
\end{aligned}
$$

c In the development of the distribution of the geometric random variable, we assumed that the experiment consisted of conducting identical and independent trials until the first success was observed. In light of these assumptions, why is the result in part (b) "obvious"?

Since the trials are Independent and identicals, the previous results don't matter.

3.75 The probability of a customer arrival at a grocery service counter in any one second is equal to .1. Assume that customers arrive in a random stream and hence that an arrival in any one second is independent of all others. Find the probability that the first arrival

a will occur during the third one-second interval.

$$p = 0.1, \quad q = 0.9$$

$$
\begin{aligned}
P(Y = 3) &= q^{y-1}p \\
&= (0.9)^2(0.1) \\
&= 0.81 \times 0.1 \\
&= 0.081
\end{aligned}
$$

b will not occur until at least the third one-second interval.

$$
\begin{aligned}
P(Y \geq 3) = P(Y > 2) &= q^2 \\
&= (0.9)^2 \\
&= 0.81
\end{aligned}
$$

3.79 In responding to a survey question on a sensitive topic (such as "Have you ever tried marijuana?"), many people prefer not to respond in the affirmative. Suppose that 80% of the population have not tried marijuana and all of those individuals will truthfully answer no to your question. The remaining 20% of the population have tried marijuana and 70% of those individuals will lie. Derive the probability distribution of Y , the number of people you would need to question in order to obtain a single affirmative response

$$P(Y = y) = (0.94)^{y-1}(0.06), \quad y = 1, 2, 3, ...$$

Expected value:

$$E(Y) = \frac{1}{p} = \frac{1}{0.06} = 16.67 \text{ people}$$

3.81 How many times would you expect to toss a balanced coin in order to obtain the first head?

$$\begin{aligned} E(Y) &= \frac{1}{p} \\ &= \frac{1}{0.5} \\ &= 2 \text{ tosses} \end{aligned}$$