SUPPLEMENTARY EXERCISES, CHAPTER 4

AN INTRODUCTION TO STATISTICAL MODELING

S4.1. Researchers conjecture that dyslexia (an information-processing impediment) and handedness are related. To test their conjecture, they randomly select the files of 10,000 elementary school children. Of these, 987 are left handed and 9013 right handed, while 543 are categorized as dyslexic. Of the 543 dyslexic children, 274 are left handed.

a. Suppose L denotes the event that a randomly selected elementary school child is left handed. Based on the given data, what would you estimate the probability of L, $P(L)$, to be? Be sure to justify your answer.

b. If M is the event a randomly selected elementary school child is both left handed and dyslexic, estimate the probability of M.

c. If K is the event a randomly selected elementary school child is neither left handed nor dyslexic, estimate the probability of K.

S4.2. The random variable Y has density

$$\begin{aligned} p_Y(y) &= c/y^3,\ y > 2 \\ &= 0, \text{ otherwise.} \end{aligned}$$

a. What is the value of c?

b. What proportion of the population modeled by this probability density function takes values below 3?

c. How relatively likely (e.g., twice as likely) is it that Y occurs in an interval about $y = 4$ of width dy compared with occurring in an interval about $y = 8$ of width dy, where dy is infinitesimally small?

S4.3. Quartz crystals for digital watches must vibrate within tight specification limits in order to be suitable for production.

A quartz crystal manufacturer claims to have a 1% rate of defectives. Approximate the probability of a shipment of 10,000 having 125 or more defectives if the manufacturer's claim is true. Be sure to justify your approximation.

S4.4. It seems as if many, if not most, retail prices end in a 9 (e.g., $29.99). To see if this is true, students in a business course took a random sample of 100 retail prices, and found that 19 ended in a 9. Is this convincing evidence that more retail prices end in a 9 than might be expected by chance? Justify your answer using a probability argument. (Hint: Assume a 9 is no more likely to occur than any other integer. Then compute the approximate probability of obtaining 19 or more 9s in a random sample of 100 retail prices.)

S4.5. As played at Atlantic city, a roulette wheel has 38 slots, numbered 1-36, 0, and 00. One of the bets you can make is that an odd number (i.e., one of 1,3,...,35) comes up. If an odd number does come up, the house pays you $1 for a $1 bet, so your gain is $1. If 0 or 00 come up, you lose half your bet, and if an even number comes up, you lose the entire bet. Suppose Y denotes your gain on a $1 bet that an odd number comes up. Find the probability mass function of Y.

S4.6. As we have all seen recently, counting votes is not an exact science. Suppose the error rate in counting votes is 0.1%. If 6,000,000 votes are cast in an election, and if correct counting of any vote is independent of correct counting of any other vote,

a. What is the distribution of Y, the total number of mis-counted votes?

b. How many mis-counted votes do you expect?

c. Approximate the probability there are 5800 or fewer mis-counted votes in the election.

S4.7. A multiple choice exam contains 10 questions, each with 5 choices, only one of which is correct. If a student just guesses on every question, what is the probability mass function of Y, the number of the 10 questions he answers correctly?

S4.8. Historically, a proportion 0.001 of monthly bills sent to residential customers by a utility company contains at least one error.

a. A random sample of 15000 residential bills from the most recent six-month billing period is taken. If the billing error rate conforms to the historical proportion, and if presence of errors in different bills are independent, what kind of distribution model describes the number of the 15000 bills with errors?

b. Of the 15000 bills in the random sample, 25 have at least one error. Is this convincing evidence that the proportion of bills with errors has increased from 0.01? Use a large-sample approximation to help you decide.

S4.9. In the game of ski-ball, players roll a ball up a ramp and onto a disk of radius 5 feet. The disk is divided into 5 sections by raised concentric circles of radii 1, 2, 3, 4, and 5 feet. The player gets 50 points for getting the ball in the smallest circle, 40 points for getting it in the sector between the smallest and second smallest circle, and 30, 20, and 10 points for getting it in the next three sectors. Assume the ball falls on the 5 foot disk at random (meaning there is no skill involved, which is exactly how I play!).

a. Obtain the distribution of Y, the number of points earned on a single roll.

b. What is the probability of getting at least 30 points on a single roll?

S4.10. Diameter specifications for mylar sheaths produced by the Ebco Company are 15 ± 2 mm. Measurements taken from a large number of production units suggest the diameters Y of the population of sheaths follow a normal distribution with mean 15.1 and standard deviation 0.9.

(a) What is the population proportion of sheaths that fall within spec?

(b) The value of the $N(15.1, 0.9^2)$ density curve at 16.2 is one half its value at 15.1. Relate this fact to the relative likelihood of finding a sheath with a diameter close to 15.1 compared with finding one with diameter close to 16.2.

S4.11. Newly-picked oranges are sorted into four classifications, which we will label 1 (highest) through 4 (lowest). Let the random variable Y denote the classification of a randomly selected orange. Experience shows that Y has probability distribution $p_Y(y) = c/y,\ y = 1, 2, 3, 4$.

(a) Find the value of c.

(b) Sketch the population histogram of this random variable.

(c) Use the information from (a) to estimate how many of the next 10,000 oranges harvested will be given one of the top two classifications.

S4.12. A manufacturer of audio CDs claims the rate of defective disks is 0.015% (i.e., a proportion 0.00015 of all disks is defective.)

(a) Approximate the probability that in a shipment of 100,000 disks, 25 or more are defective.

(b) Justify the approximation you made in (a).

S4.13. The eccentricity of a mouse tracking ball is a measure of how out-of-round it is. A perfectly round ball has value of 1, an out-of round ball has a value greater than 1. Suppose the eccentricity of a population of mouse tracking balls is measured by the random variable Y having density curve

$$\begin{aligned} p(y) &= 0,\ y < 1, \\ &= c/y^5,\ y \geq 1 \end{aligned}$$

(a) Find the value of the constant c for the density curve.

(b) A mouse tracking ball is considered unacceptable if its eccentricity is greater than 10. What proportion of the mouse tracking balls are unacceptable due to excessive eccentricity?

(c) How much more likely is it to find a mouse tracking ball with eccentricity close to 5 than to find one with eccentricity close to 10?

S4.14. Based on historical data, Egbert's travel time from home to work, in minutes, is well modeled by a normal distribution with mean 19 and variance 9. If Egbert leaves home thirteen minutes ahead of an appointment at work, what is the probability he is late for the appointment?