Worksheets

Stephanie Ranft S2459825

October 8, 2019

Statistics 1A
PSBE1-08

1 Week 4

1. You throw a pair of dice and sum the values shown face up, e.g. if you throw and 2 and a 4, you have a sum of $X = 6$.
   (a) Is this a discrete or continuous probability distribution? Explain your answer, and give the name of it if you know it.
   (b) What is the probability that your sum $X = 2$ (expressed as a fraction)?
   (c) What is the probability that your sum $X = 5$ (expressed as a fraction)?
   (d) What is the probability that at least one of the dice shows a 5 (expressed as a fraction)?
   (e) What is the probability that your sum $X = 12$ (expressed as a fraction)?
   (f) Complete the table below.

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = x)$											

Table 1

   (g) What is the probability of having at least 10 (expressed as a fraction)?
   (h) You roll the first die and get a 5. After you roll the second die, what is the probability of having at least 10 (expressed as a fraction)?
   (i) What is the expected value and variance?
   (j) Let $Y = X - 7$. What is the expected value and variance of Y?

2. In a Child Health Survey conducted in a school, 1567 children answered "yes" and 433 children answered "no" in response to the question: "Do you drink milk?"
   (a) Estimate the probability that a randomly selected child drinks milk.
   (b) Estimate the probability that a child does not drink milk.
   (c) Of those who drink milk, 40% like it with chocolate. Estimate the probability that a randomly chosen child likes to drink chocolate milk.

3. In a town of 10,000 people, 400 have beards (all men), 4000 are adult men, 5 of the townspeople are murderers. All 5 murderers are men and 4 of the murderers have beards.

   Suppose you go to this town and select a towns person at random.

   - Let A be the event that the person turns out to be one of the five murderers.
   - Let B be the event the person is bearded.
   - Let C be the event the person is an adult male.

   Find $\mathbb{P}(A)$, $\mathbb{P}(A \mid B)$, $\mathbb{P}(B \mid A)$, $\mathbb{P}(A \mid B^c)$, $\mathbb{P}(A \mid C)$, $\mathbb{P}(A \mid C^c)$.

4. Disease X is a disease affecting about 1 percent of the population. A test for Disease X will test positive on all afflicted with the disease and will also test positive for 5% of the population who do not have the disease.

   (a) Draw a tree diagram to display this information.

   (b) Fill in the table.

	Positive	Negative
Disease		
No disease		

Table 2

   (c) What is the probability a randomly chosen person has disease X?

   (d) What is the probability a randomly chosen person will test positive for the disease?

   (e) Suppose you test positive for the disease. What is the probability you don't actually have the disease? (This is the conditional probability that you don't have the disease given you tested positive for it).

   (f) What would be the probability that you test positive for the disease twice given that you don't have disease X?

5. A die is thrown 3 times. What is the probability of

   (a) No 5's showing up?

   (b) One 5 showing up?

   (c) Three 5's showing up?

6. Hospital records show that of patients suffering from a certain disease, 75% die of it. What is the probability that of 6 randomly selected patients, 4 will recover?

7. In the old days, there was a probability of 0.8 of success in any attempt to make a telephone call. (This often depended on the importance of the person making the call, or the operator's curiosity!)

   (a) Calculate the probability of having 7 successes in 10 attempts.

   (b) Can we use the normal distribution to approximate this? Why or why not? If not, how can we modify our distribution so that we can, and what is the resulting mean and standard deviation for this normal approximation?

8. A (blindfolded) marksman finds that on the average he hits the target 4 times out of 9. If he fires 45 shots, what is the probability of

   (a) more than 2 hits?

   (b) at least 3 misses?

   (c) Can we use the normal distribution to approximate this? Why or why not? If not, how can we modify our distribution so that we can, and what is the resulting mean and standard deviation for this normal approximation?

9. The ratio of boys to girls at birth in Singapore is quite high at 1.09:1. What proportion of Singaporian families with exactly 6 children will have at least 3 boys (ignoring the probability of twins, or multiple births).

10. A manufacturer of metal pistons finds that on the average, 12% of his pistons are rejected because they are either oversize or undersize. What is the probability that a batch of 10 pistons will contain no more than 2 rejects? What about the probability that in this batch of 10, there will be at least 2 rejects?

1.1 Solutions

1. (a) Discrete. Not Binomial distribution.

   (b) $X = 2$ if you roll two 1's, i.e.

$$\mathbb{P}\left(\text{die one shows } 1 \cap \text{die two shows } 1\right) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

   This is because rolling of the die are **independent** of each other and the probability with a 6-sided die is $1/6$.

   (c) The sum $X = 5$ when I roll a 1 and a 4, a 2 and a 3, a 3 and a 2, or a 4 and a 1

$$\Longrightarrow \mathbb{P}\left(X = 5\right) = \underbrace{\frac{1}{6} \cdot \frac{1}{6}}_{\text{a 1 and a 4}} + \underbrace{\frac{1}{6} \cdot \frac{1}{6}}_{\text{a 2 and a 3}} + \underbrace{\frac{1}{6} \cdot \frac{1}{6}}_{\text{a 3 and a 2}} + \underbrace{\frac{1}{6} \cdot \frac{1}{6}}_{\text{a 4 and a 1}} = \frac{4}{36} = \frac{1}{9}.$$

   (d) The probability that one die shows a 5 (and the other shows any other number) is $2 \cdot 1/6 \cdot 5/6 = 10/36$, and the the probability that both dice show a 5 is $1/6 \cdot 1/6 = 1/36$. Then, the probability that at least one of the dice shows a 5 is the probability that one or both show a 5, i.e. $10/36 + 1/36 = 11/36$.
   Another thing to thing about:

$$\mathbb{P}\left(\text{at least one}\right) = 1 - \mathbb{P}\left(\text{none}\right).$$

   This intuitively makes sense! Then, the probability that neither of the dice shows a 5 is $5/6 \cdot 5/6 = 25/36$, and the probability of at least one is $1 - 25/36 = 11/36$.

   (e) The sum equals 12 when both dice show a 6, i.e. $1/6 \cdot 1/6 = 36$.

   (f)

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}\left(X = x\right)$	$1/36$	$1/18$	$1/12$	$1/9$	$5/36$	$1/6$	$5/36$	$1/9$	$1/12$	$1/18$	$1/36$

Table 3

   (g) The probability of having at least 10 is the sum of the probabilities of having 10, 11, or 12.

$$\mathbb{P}\left(X \geq 10\right) = \mathbb{P}\left(X = 10\right) + \mathbb{P}\left(X = 11\right) + \mathbb{P}\left(X = 12\right) = \frac{1}{12} + \frac{1}{18} + \frac{1}{36} = \frac{1}{6}.$$

   The probabilities are added in this way as the distribution is discrete!

   (h) This is a conditional probability question given by $\mathbb{P}\left(X = 10 \mid X \geq 6\right)$; as you have already rolled a 5 and the values on the dice start at 1, you know that your sum is at least 6.

$$\mathbb{P}\left(X = 10 \mid X \geq 6\right) = \frac{\mathbb{P}\left(X = 10 \cap X \geq 6\right)}{\mathbb{P}\left(X \geq 6\right)} = \frac{\mathbb{P}\left(X = 10\right)}{\mathbb{P}\left(X \geq 6\right)}.$$

   The probability $\mathbb{P}\left(X \geq 6\right)$ is calculated as

$$\mathbb{P}\left(X \geq 6\right) = \mathbb{P}\left(X = 6\right) + \mathbb{P}\left(X = 7\right) + \mathbb{P}\left(X = 8\right) + \mathbb{P}\left(X = 9\right) + \mathbb{P}\left(X \geq 10\right) = \frac{5}{36} + \frac{1}{6} + \frac{5}{36} + \frac{1}{9} + \frac{1}{6} = \frac{26}{36}.$$

   Therefore,

$$\mathbb{P}\left(X = 10 \mid X \geq 6\right) = \frac{1/12}{26/36} = \frac{3}{26}.$$

   (i) The expected value is the sum of the probabilities times the values, i.e.

$$\mu = \sum_{x=2}^{12} x \cdot \mathbb{P}\left(X = x\right) = \frac{2}{36} + \frac{1}{6} + \frac{1}{3} + \frac{5}{9} + \frac{5}{6} + \frac{7}{6} + \frac{40}{36} + 1 + \frac{5}{6} + \frac{11}{18} + \frac{1}{4} = 7.$$

   We compute the variance by look at the sum of product of the squared differences between the values and the mean, and the probabilities, i.e.

$$\sigma^2 = \sum_{x=2}^{1} 2\left(x - \mu\right)^2 \cdot \mathbb{P}\left(X = x\right) = \frac{25}{36} + \frac{8}{9} + \frac{3}{4} + \frac{4}{9} + \frac{5}{36} + 0 + \frac{5}{36} + \frac{4}{9} + \frac{3}{4} + \frac{8}{9} + \frac{25}{36} = 5 + \frac{5}{6} \approx 5.83.$$

   By taking the square root of the variance, we conclude that a person playing this game could expect a payout of $7 \pm 2.415 = (4.585, 9.415)$.

(j) The mean of Y is equal to the mean of X minus 7, i.e. zero. The variance of Y is the same as X.

2. (a) The proportion of milk drinkers $p = 1567/(1567 + 433) = 0.7835$.

   (b) The proportion of non- milk drinkers is $1 - p = 0.2165$.

   (c) This is a conditional probability, i.e. $\mathbb{P}(\text{likes chocolate} \mid \text{milk=“yes”}) = 0.40$, and we are looking for the probability that they drink milk **and** like chocolate, i.e.

   $\mathbb{P}(\text{likes chocolate} \cap \text{milk=“yes”}) = \mathbb{P}(\text{likes chocolate} \mid \text{milk=“yes”}) \cdot \mathbb{P}(\text{milk=“yes”}) = 0.4 \cdot 0.7835 = 0.3134.$

3.

$$\mathbb{P}(A) = \frac{5}{10000} = 0.0005 \qquad \mathbb{P}(A \mid B) = \frac{4}{400} = 0.01 \qquad \mathbb{P}(B \mid A) = \frac{4}{5} = 0.8$$

$$\mathbb{P}(A \mid B^c) = \frac{5-4}{10000-400} = 0.000104167 \qquad \mathbb{P}(A \mid C) = \frac{5}{4000} = 0.00125 \qquad \mathbb{P}(A \mid C^c) = 0$$

4. (a)

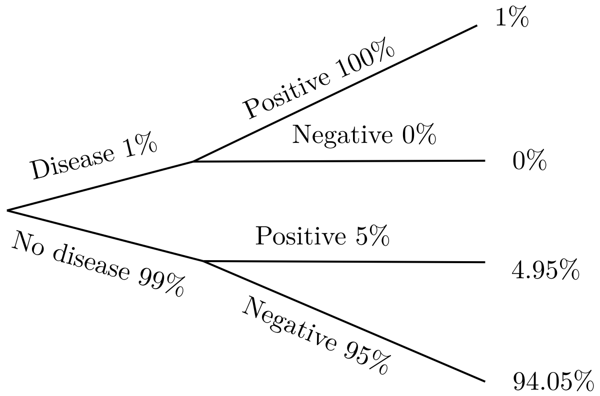

Figure 1

   (b)

	Positive	Negative
Disease	1%	0%
No disease	4.95%	94.05%

Table 4

   (c) $\mathbb{P}(D) = 0.01$.

   (d) $\mathbb{P}(P) = \mathbb{P}(P \mid D) \cdot \mathbb{P}(D) + \mathbb{P}(P \mid ND) \cdot \mathbb{P}(ND) = 1 \cdot 0.01 + 0.05 \cdot 0.99 = 0.0595$.

   (e) $\mathbb{P}(ND \mid P) = \frac{\mathbb{P}(ND \cap P)}{\mathbb{P}(P)} = \frac{\mathbb{P}(P|ND) \cdot \mathbb{P}(ND)}{\mathbb{P}(P)} = \frac{0.05 \cdot 0.99}{0.0595} \approx 0.832$.

   (f) We assume that taking the test are independent events, so $\mathbb{P}(P \text{ twice} \mid ND) = [\mathbb{P}(P \mid ND)]^2 = 0.025$.

5. This is a Binomial distribution where X is the number of 5's shown in three throws of the die.

$$\mathbb{P}(X = x) = \binom{3}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{3-x} = \frac{3!}{x!(3-x)!} \frac{5^{3-x}}{6^3}; \quad x = 0, 1, 2, 3.$$

   (a) Here, $x = 0$ so

$$\mathbb{P}(X = 0) = \frac{3!}{0!(3-0)!} \frac{5^{3-0}}{6^3} = \frac{125}{216}.$$

   (b) Here, $x = 1$ so

$$\mathbb{P}(X = 1) = \frac{3!}{1!(3-1)!} \frac{5^{3-1}}{6^3} = 3 \cdot \frac{25}{216} = \frac{75}{216}.$$

   (c) Here, $x = 3$ so

$$\mathbb{P}(X = 3) = \frac{3!}{3!(3-3)!} \frac{5^{3-3}}{6^3} = \frac{1}{216}.$$

6. This is a binomial distribution where X is the number of people (out of 6) who recover.

$$\mathbb{P}(X = x) = \binom{6}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{6-x} = \frac{6!}{x!(6-x)!}\frac{3^{6-x}}{4^6} \implies \mathbb{P}(X = 4) = \frac{6!}{4!(6-4)!}\frac{3^2}{4^6} \approx 0.03296.$$

7. (a)

$$\mathbb{P}(X = 7) = \binom{10}{7}(0.8)^7(0.2)^{10-7} \approx 0.20133.$$

(b) No, we require that np and $n(1-p)$ are both larger than 5. In order to achieve this, we need

$$0.8n \geq 5 \quad \text{and} \quad 0.2n \geq 5 \implies n \geq 25.$$

If the sample size is 25, we can approximate using a normal distribution with mean $25 \cdot 0.8 = 20$ and variance $25 \cdot 0.8 \cdot 0.2 = 4$ (standard deviation of 2).

8.

$$\mathbb{P}(X = x) = \binom{45}{x}\left(\frac{4}{9}\right)^x\left(\frac{5}{9}\right)^{45-x} = \frac{45!}{x!(45-x)!}\frac{4^x5^{45-x}}{9^45}.$$

(a)

$$\begin{aligned}\mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leq 1) = 1 - \mathbb{P}(X = 1) - \mathbb{P}(X = 0)\\ &= 1 - 45 \cdot \frac{4 \cdot 5^{44}}{9^45} - \frac{45 \cdot 44}{2}\frac{4^2 \cdot 5^{43}}{9^45} \approx 1\end{aligned}$$

(b) If we have at least 3 misses, then we have at most 42 hits, i.e. $\mathbb{P}(X \leq 42) \approx 1$.

(c) As $np = 20 > 5$ and $n(1-p) = 25 > 5$, we can use the normal distribution to approximate this, with mean $np = 20$ and standard deviation $\sqrt{np(1-p)} \approx 3.333$. The z-score for $x = 2$ is $(2-20)/3.33 = -5.4$ and for $x = 42$ is $(42-20)/3.33 = 6.60$, so this helps us make sense of the results for parts (a) and (b).

9. If the ratio of boys to girls is 1.09:1, then the proportion of boys born is $1.09/(1+1.09) \approx 0.5215311$. We can now structure this a binomial distribution where $n = 6$ and X is the number of boys:

$$\mathbb{P}(X \geq 3) = 1-\mathbb{P}(X \leq 2) = 1-\binom{6}{0}0.522^0(1-0.522)^6-\binom{6}{1}0.522^1(1-0.522)^5-\binom{6}{2}0.522^2(1-0.522)^4 \approx 0.6957.$$

10. The batch size $n = 10$ and X is the number of rejected pistons, with probability 0.12:

$$\mathbb{P}(X = x) = \binom{10}{x}0.12^x0.88^{10-x}.$$

If there are no more than 2 rejects, then

$$\mathbb{P}(X \leq 2) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \approx 0.89132.$$

If there are at least two rejects, then

$$\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X \leq 1) \approx 0.3417.$$