Quiz 2

Stats 60/160

August 5, 2020

Important note - Throughout the quiz, show your work. You can get partial credit even if the final answer is wrong.

1 Rolling two dice

A pair of fair dice with 6 numbers is thrown and the numbers are recorded. Suppose the two dice are different.

(a) What is the sample space?

Answer The sample space is the collection of all possible basic outcomes of an experiment. In this problem, there are 36 different outcomes in total, and we can enumerate them as $(1,1),(1,2),\ldots,(1,6)$, and so on, till $(6,1),(6,2),\ldots,(6,6)$.

(b) What is the probability that the smaller number on the two dice is 2? (i.e. one is 2 and the other is larger than or equal to 2)

Answer Enumerate all outcomes in the sample space where the smaller is 2:

$$(2,2),(2,3),(2,4),(2,5),(2,6),(3,2),(4,2),(5,2),(6,2),$$

there's a total of 9 outcomes. Each of them are equally likely, and this gives the probability $9/36 = 1/4$.

2 Drawing marbles

A box contains 15 red marbles and 5 blue marbles.

(a) 10 marbles are drawn without replacement. What is the chance at least 2 of them will be red?

Answer The probability is 1 because there are only 5 blue marbles, so to draw 10 marbles without replacement, at least 2 will be red.

(b) 10 marbles are drawn with replacement. What is the chance at least 2 of them will be red?

Answer We can use the complement rule:

$$\begin{aligned}\Pr(\text{at least 2 red}) &= 1 - \Pr(\text{no red}) - \Pr(1 \text{ red}) \\ &= 1 - \left(\frac{5}{20}\right)^{10} - 10 \times \left(\frac{5}{20}\right)^{9} \left(\frac{15}{20}\right) \\ &= 0.9999.\end{aligned}$$

(c) Keep drawing marbles from the box with replacement, what's the chance the first time you see a red marble is the 3rd draw?

Answer

$$\begin{aligned}\Pr(\text{Red appear first on 3rd draw}) &= \Pr(\text{first two draws are blue AND third draw is red}) \\ &= \Pr(\text{first draw is blue}) \times \Pr(\text{second draw is blue}) \times \Pr(\text{third draw is red}) \\ &= \left(\frac{5}{20}\right) \times \left(\frac{5}{20}\right) \times \left(\frac{15}{20}\right) \\ &= 0.047.\end{aligned}$$

In the second equality we use the multiplication rule because results from each draw are independent.

3 Measuring blood pressure

Systolic blood pressure measurement is intrinsically variable, and assume the measurement is normally distributed, centered at the actual blood pressure, and its the standard deviation is about 10% the actual value. Suppose Lisa's blood pressure is 130mm Hg.

(a) What's the chance that her blood pressure is measured to be above 140 mm Hg?

Answer The standard deviation of her measurement is 13 mmHg. 140 is 0.77 SD above the mean, from the normal table the area is about 56.2%, which means the chance to see a higher measurement is $(1 - 0.562)/2 = 0.219$.

(b) Suppose her blood pressure is measured 4 times. What's the chance that at least 3 out of 4 times the measurement is higher than 140mm Hg?

Answer

$$\begin{aligned}&\Pr(\text{above 140mmHg at least 3 out of 4 times}) \\ &= \Pr(\text{above 140mmHg 3 out of 4 times}) + \Pr(\text{above 140mmHg every time}) \\ &= \binom{4}{3} 0.219^3 (1 - 0.219) + \binom{4}{4} 0.219^4 \\ &= 0.035.\end{aligned}$$

You can also use the binomial density formula

$$\text{dbinom}(3; 4, 0.219) + \text{dbinom}(4; 4, 0.219) = 0.035.$$

4 Statistical Independence

Throw a die with both color and numbers:

$$(\text{Yellow}, 1), (\text{Yellow}, 2), (\text{Yellow}, 3), (\text{Green}, 4), (\text{Green}, 5), (\text{Green}, 6).$$

Is the color independent of numbers on the die?

Answer No. We can check the definition: the chance of drawing a "Yellow" is 1/2 and the chance of drawing the number 1 is 1/6, the chance of drawing (Yellow, 1) is 1/6, so they are not independent because

$$\Pr((\text{Yellow}, 1)) \neq \Pr(\text{Yellow}) \times \Pr(1).$$

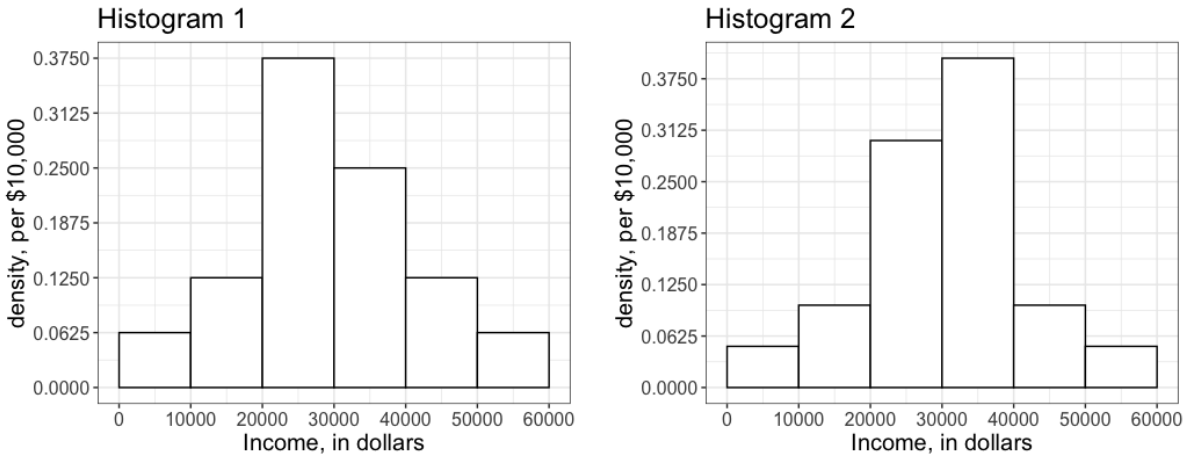
In fact, the color is determined by the number.

5 Drawing a histogram

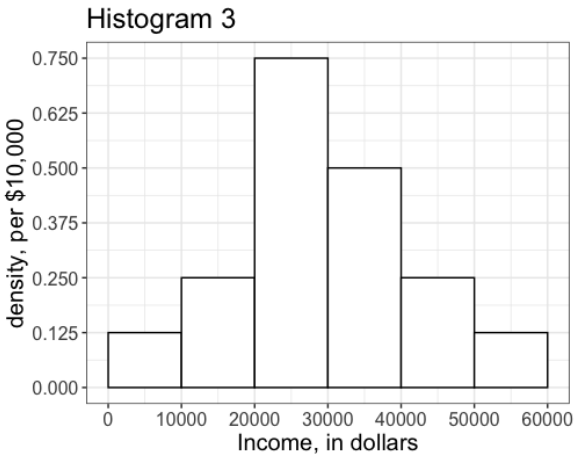
Consider the following income data Which of the following is the probability density histogram of this

Income range	# households
0 - 9,999	5
10,000 - 19,999	10
20,000 - 29,999	30
30,000 - 39,999	20
40,000 - 49,999	10
50,000 - 59,999	5

data? Explain your answer.



Answer The first histogram is correct. The second histogram peaks in the interval 30,000 - 40,000, which is incorrect, and the area of the third histogram does not add up to 1.


Histogram 3
0.750
0.625
0.500
0.375
0.250
0.125
0.000
density, per $10,000
0
10000
20000
30000
40000
50000
60000
Income, in dollars