

Quiz 6 (probability quiz 2)

1. Mammograms and ultrasounds are used to test for breast cancer. In a certain country, there were 4500 women who went for both tests in 2016.

| | Positive ultrasound | Negative ultrasound |
|--------------------|---------------------|---------------------|
| Positive mammogram | 100 | 200 |
| Negative mammogram | 300 | 3900 |

For a woman randomly selected from these 4500 women, is having a positive mammogram independent from having a positive ultrasound?

- a) Yes
- b) No
- c) Cannot be determined from the given information

Solution: The answer is b). For the randomly selected woman, let M be the event that she has a positive mammogram and let U be the event that she has a positive ultrasound. Then $P(M \text{ and } U) = 100/4500 = 0.02$, while $P(M) = (100 + 200)/4500 = 0.07$ and $P(U) = (100 + 300)/4500 = 0.09$.

Since $P(M \text{ and } U) = 0.02$ is not equal to $P(M) * P(U) = 0.07 * 0.09 = 0.0063$, that is, having a positive mammogram and having a positive ultrasound are not independent events.

2. Bob took a quiz consisting of 500 questions with two options: Yes and No. Bob did not prepare for the quiz. Flustered, he hastily reached for a fair coin in his wallet and started to toss this coin for answers. Whenever the coin landed heads, he shaded Yes; whenever the coin landed tails, he shaded No.

It turned out that 40% of the questions have Yes as the correct answer, and the rest have No as the correct answer. Among those questions that Bob shaded Yes, roughly how many percent did he get correct?

- a) 30%
- b) 40%
- c) 50%
- d) 60%
- e) 70%

Solution: The answer is b). Regard a question as a unit in an experimental study conducted by Bob. Imagine that a question is male if the correct answer is Yes and that a question is female if the correct answer is No. The coin tosses randomly assign these questions to treatment (shaded Yes) or control (shaded No), which are likely to be similar, in any given respect, to the entire group of units in the study.

In terms of our analogy, this means that whether Yes is the correct answer is not associated to whether Bob actually shaded Yes. Thus the percentage of correct answers among those questions that Bob shaded Yes is equal to the percentage of the questions having Yes as the correct answer, which is 40%.

3. Felicia observed getting “6 spots” 3 times out of 6 rolls of a six-sided die. Eve suggested that the die is biased towards “6 spots”. In order to quantitatively evaluate Eve’s suggestion, Felicia set up the null hypothesis that $P(\text{“6 spots”}) = 1/6$.

Which of the following choices has exactly those events that are equally or more extreme than the observed event?

- a) Getting “6 spots” 0, 1, or 2 times
- b) Getting “6 spots” 0, 1, 2, or 3 times
- c) Getting “6 spots” 3, 4, 5, or 6 times
- d) Getting “6 spots” 4, 5, or 6 times

Solution: The answer is c). If the null hypothesis is true, Felicia expects to get “6 spots” once. If she had observed it 4 times, she would be even more skeptical about the null, and even more with 5 or 6 times.

4. Colorectal cancer has a base rate of 4%. A fecal occult blood test used to detect colorectal cancer has sensitivity 77% and specificity 57%. Among those who tested negative, roughly what percentage has colorectal cancer?

- a) 2%
- b) 3%
- c) 4%
- d) 5%

Solution: The answer is a). Suppose 10000 people did the fecal occult blood test.

| | Tests positive | Tests negative | Row sum |
|-----------------------|----------------|----------------|---------|
| Has colorectal cancer | 308 | 92 | 400 |
| No colorectal cancer | 4128 | 5472 | 9600 |
| Column sum | 4436 | 5564 | 10000 |

From the table, among those who test negative, roughly $92/5564 = 2\%$ has colorectal cancer.

5. There is a new home test kit for HIV detection. This test kit is known to be 99% accurate (ie. its specificity and sensitivity are both 99%.)

Which of the following statements are correct?

- (I) Among those with HIV, roughly 1% tests negative.
- (II) Among those who test negative, roughly 1% has HIV.

- a) Only (I)
- b) Only (II)
- c) Both (I) and (II)
- d) Neither (I) nor (II)

Solution: The answer is c). As the sensitivity is 99%, hence, by definition, among those with HIV, roughly 99% tests positive. Thus among those with HIV, roughly 1% tests negative, i.e. (I) holds.

Similarly, as the specificity is 99%, hence, by definition, among those who test negative, roughly 99% does not have HIV. Thus among those who test negative, roughly 1% has HIV, i.e. (II) holds.