MATH 323 - Tutorial 3 Questions

Tyrel Stokes

February 15, 2022

1. Two six-sided fair die are tossed. Assume that the die are independent of each other.

a) If at least one 3 was rolled, what is the probability that both were rolled 3's?

b) Suppose this time that the die sum to 7. What is the probability that a 3 was rolled?

c) If it is known that at least one of the dice is 3, what is the probability that their faces sum to 7?

d) Challenge: Suppose we keep rolling a pair of dice and sum their faces. What is the probability that we roll a 3 before we roll a 7? Assume that the game stops as soon as the dice sum to either 3 or 7, but otherwise continues indefinitely.

2. Workers and Shirkers:

A student is working on a group assignment in a class with 12 other students. Suppose that 2 of the other students are known to be shirkers (that is they neglect their responsibilities), but that the remaining 10 are known hard workers (called workers throughout). The teacher tells the student to reach into a hat containing the names of the twelve other students and randomly select 2 names which will make up the students work group.

- Suppose that if no shirkers are selected, the probability of getting an A on the assignment is 0.95.
- If 1 shirker is selected, the probability of getting an A is 0.6.
- If 2 shirkers are selected, the probability of getting an A is 0.1.

What is the probability that:

a) The student does not get an A on the assignment?

b) Given that the student didn't get an A, the probability that precisely 1 shirker was selected?

c) Given that the student didn't get an A, the probability that at least 1 shirker was selected?

3. Athletes are routinely tested for the use of performance-enhancing drugs (PEDs). When a testis performed, an athlete provides 2 blood samples (an A sample and a B sample). If the A sample tests positive, then the B sample is also tested. If the B sample is also positive,the athlete is considered to have failed the drug test.

Suppose an athlete is selected at random and two identical, random blood samples (A and B) are obtained. Let $T_i, i \in \{A, B\}$ be the event that the ith sample tests positive. Let C be the event that their truly are PEDs present in the samples. Suppose the test is quite accurate, in that it correctly indicates the **presence** of drugs in 99.5% of samples and correctly indicates the absence of drugs in 98% of samples. Suppose that only 1 athlete per 1000 is actually taking PEDs which would show up in the samples.

a) What is the probability that the first test is positive?

b) What is the probability that drugs are actually present in the sample if the A sample tests positive?

c) What is the the probability that both tests are positive? Suppose that the tests are independent.

d) What is the probability that drugs are actually present in the sample if both the A and B samples test positive?

Challenge:

e) Suppose instead that the tests are not independent, but independent conditional on C. (This hasn't been covered in class so do not worry if you have not heard of this concept. For those looking to go further in statistics, this is a concept that will come up but again not something you need to know for this class.). For the purposes of this question, all you need to know is that conditional independence here implies that $P(T_A \cap T_B|C) = P(T_A|C)P(T_B|C)$. What is the probability under the conditional independence assumption that both tests are positive? (Think about how this compares to the case that the tests are fully independent and why)

3. Athletes are routinely tested for the use of performance-enhancing drugs (PEDs). When a testis performed, an athlete provides 2 blood samples (an A sample and a B sample). If the A sample tests positive, then the B sample is also tested. If the B sample is also positive,the athlete is considered to have failed the drug test.

Suppose an athlete is selected at random and two identical, random blood samples (A and B) are obtained. Let $T_i, i \in \{A, B\}$ be the event that the ith sample tests positive. Let C be the event that their truly are PEDs present in the samples. Suppose the test is quite accurate, in that it correctly indicates the **presence** of drugs in 99.5% of samples and corrrectly indicates the absence of drugs in 98% of samples. Suppose that only 1 athlete per 1000 is actually taking PEDs which would show up in the samples.

a) What is the probability that the first test is positive?

Solutions:

From the question, we are given several conditional probabilities. The first step of these kinds of probabilities is to identify the probability of interest and take stock of the givens in the word problem. In this case $P(T_A)$ is the probability of interest. We are given that $P(T_A|C) = 0.995$ (correctly identifies positive cases) and that $P(T_A^c|C^c) = 0.98$ (correctly identifies negative cases). Finally, we are given that overall, $P(C) = \frac{1}{1000}$.

Notice that $C \cup C^c$ form a partition, thus we can use the law of total probability.

$$\begin{aligned} P(T_A) &\stackrel{TLP}{=} P(T_A|C)P(C) + P(T_A|C^c)P(C^c) \\ &= P(T_A|C)P(C) + (1 - P(T_A^c|C^c))(1 - P(C)) \\ &= .995 * .001 + 0.02 * .999 \\ &= 0.021 \end{aligned}$$

b) What is the probability that drugs are actually present in the sample if the A sample tests positive?

Solutions:

Here the probability of interest is $P(C|T_A)$. When faced with a conditional probability like this, we typically have two options. Bayes rule or use the direct definition. Which is easier will depend (usually) on whether it is easier to find (or access if it is given) the so-call "inverse probability" (in this case $P(C|T_A) \to P(T_A|C)$) or the intersection (in this case $P(C \cap T_A)$). In this case, due to the information given before, bayes rule will be easier.

$$\begin{aligned} P(C|T_A) &\stackrel{Bayes}{=} \frac{P(T_A|C)P(C)}{P(T_A)} \\ &= \frac{.995 * .001}{0.021} \\ &= 0.047 \end{aligned}$$

Where we used the law of total probality to find the denominator in part a.

Think about what this answer means! Even though the test is very accurate in a sense, it still doesn't give us much information about whether or not the athlete is actually on banned substances. In this case, this is largely die to the fact that $P(C)$ is so small. If the rate of banned substance use was much higher, the information from a positive test would be much greater.

c) What is the the probability that both tests are positive? Suppose that the tests are independent.

Solutions:

$$
\begin{aligned}
P(T_A \cap T_B) &= P(T_A)P(T_B) \quad [\text{ Independence }] \\
&= 2 * P(T_A) \quad [\text{ Identical }] \\
&= (0.21)^2 \quad [\text{ From part a) }] \\
&\approx 0.000441
\end{aligned}
$$

Challenge:

d) Suppose instead that the tests are not independent, but independent conditional on C. (This hasn't been covered in class so do not worry if you have not heard of this concept. For those looking to go further in statistics, this is a concept that will come up but again not something you need to know for this class.). For the purposes of this question, all you need to know is that conditional independence here implies that $P(T_A \cap T_B|C) = P(T_A|C)P(T_B|C)$. What is the probability under the conditional independence assumption that both tests are positive? (Think about how this compares to the case that the tests are fully independent and why)

Solutions:

$$
\begin{aligned}
P(T_A \cap T_B) &\overset{TLP}{=} P(T_A \cap T_B|C)P(C) + P(T_A \cap T_B|C^c)P(C^c) \\
&\overset{\text{cond. ind.}}{=} P(T_A|C)P(T_B|C)P(C) + P(T_A|C^c)P(T_B|C^c)P(C^C) \\
&\overset{ident.}{=} (P(T_A|C))^2 P(C) + (P(T_A|C^c))^2(1 - P(C)) \\
&= (.995)^2 \times (0.001) + (1 - .98)^2(1 - .001) \\
&\approx 0.014
\end{aligned}
$$

Notice that compared to when we assumed independence that the probability of $P(T_A \cap T_B)$ is higher (i.e $0.014 > 0.000441$)

e) What is the probability that drugs are actually present in the sample if both the A and B samples test positive?

Solutions:

$$
\begin{aligned}
P(C|T_A \cap T_B) &\overset{Bayes}{=} \frac{P(T_A \; capT_B|C)P(C)}{P(T_A \cap T_B)} \\
&\overset{cond.ind.}{=} \frac{P(T_A|C)P(T_B|C)P(C)}{P(T_A \cap T_B)} \\
&= \frac{(0.995)^2(0.001)}{0.014} \\
&\approx 0.712
\end{aligned}
$$