NAME:

STAT 503 – Statistical Methods for Biology

Homework 3

24 Points (26 available). Due at 11:59 PM on Monday, Sept 28, 2020

**Please use complete sentences unless the question is marked with an asterisk (\*).**

1. The relationship between squirrels and oak trees is referred to as a *conditional mutualism* because it can shift from mutualistic (both parties benefit) to predatory (the squirrels benefit at the trees' expense) as a result of changing ecological conditions. Squirrels eat acorns (the seeds of oak trees), but they do not eat all of the acorns as soon as they are collected. Instead, they bury most of the acorns in the ground and store them for later. The majority of these buried acorns are eventually eaten. However, if a squirrel buries an acorn and then never returns to eat it (and if nothing else discovers the acorn and eats it), then the buried acorn will be protected from cold, desiccation (water-loss), and fire. As a result, it will have a much better chance of surviving and establishing a new seedling than it would have if it had never been collected in the first place.

If squirrels bury and then abandon enough acorns, then oak trees tree will produce more seedlings in systems that have squirrels than in systems without squirrels. This would mean that the squirrels act as mutualists. On the other hand, if fewer acorns are buried and then abandoned, the production of seedlings will be higher in the absence of squirrels. In that case, the squirrels function as seed predators instead of mutualists.

Suppose that squirrels collect 80% of the acorns produced by a particular population of oak trees. They bury 70% of the acorns that they collect, and the remaining acorns that were collected are eaten immediately. In addition, only 1% of the collected-and-buried acorns are abandoned. The squirrels eventually dig up and eat the rest of the buried acorns. In one year, 0.2% of acorns that are not collected by squirrels survive to become seedlings (the others become desiccated and die). In contrast, 25% of the acorns that were collected, buried, and then abandoned become seedlings.

In the following questions, let be the event that an acorn is collected by a squirrel, let *B* indicate that an acorn is buried by a squirrel, and let indicate that an acorn is abandoned. In addition, let be the event that an acorn survives to become a seedling. Finally, let a prime () indicate the complement of an event (e.g., is the event that an acorn is *not* collected), and let two letters next to each other denote an intersection. For example, is the event that an acorn is collected and buried.

**Please report you answers to 5 decimal places (dropping trailing zeroes is okay).**

1. \*[3.5 points] The following table lists a set of probabilities. All of these values were given in the problem above, or are readily determined by logic (e.g., if an acorn is never collected by a squirrel, then it cannot be buried by a squirrel, so ). Please give the value for each probability and briefly explain its biological interpretation, as shown in line for .

|  |  |  |
| --- | --- | --- |
| Symbol | Value | Interpretation |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  | 0 | probability that an acorn will be buried given that it is not collected (if the acorn is not collected, it cannot be buried) |
|  |  |  |
|  |  |  |

1. \*[3 points] Use the probabilities from question 1a to draw a tree diagram that shows all of the potential outcomes for a single acorn. **Your tree will have 8 different end points, IF you end a branch as soon as you know the acorn's fate.** You will know its fate when the acorn becomes a seedling, when it is eaten, or when you encounter a probability of zero (0), indicating that the combination of events on the current branch is impossible.

Your diagram should use the event notations and probabilities listed in the table for question 1a, as well as their complements. It should also include the probability for each path through the tree (see Lecture 3.3 slide 7 as an example). You can either use a computer to draw the diagram (powerpoint is useful for this kind of figure) or you may draw the diagram by hand and insert a photo of it. If you use a photo, please make sure it is legible. Landscape format is okay.

1. [2 points] Conceptually, how does the marginal probability that an acorn will not be buried, , differ from the conditional probability that it will not be buried, given that it was collected, ? In addition, please calculate the values for each of these two probabilities.
2. \*[3 points] Based on your tree, what is the probability that an acorn will survive to become a seedling,
   1. Given that the acorn is collected by a squirrel
   2. Given that it is not collected by a squirrel
   3. Overall
3. \*[1 point] Let represent the probability that an acorn will become a seedling if there are no squirrels in the system (i.e., if ). Assuming that all other probabilities remain the unchanged, please calculate .
4. [1.5 point] Based on the results from question 1d and 1.e, were the squirrels functioning as mutualists or as predators in this year (for the interaction to be mutualistic, the trees have to do better with squirrels than without squirrels)? Please explain the reason for your interpretation.
5. [1.5 point] Suppose that in the next year, the squirrels abandon 1.8% of their buried acorns instead of 1%, but survival of uncollected acorns also increases to 0.3%. Assuming that all other probabilities in the tree remain unchanged, how does this increase affect the interaction?
6. \*[1.5 points] You are walking through the woods and find an oak seedling. **Using your answers in question 1.d**, what is the probability that this seedling came from an acorn that was collected by a squirrel (i.e., find ). Compare this result with your answer in question 1f. How can these both be true?
7. Female song sparrows (*Melospiza melodia*) prefer males with larger song repertoires, probably because the number of different songs in a male's repertoire is positively associated with several indicators of his genetic fitness, including physical body condition and immune function (Pfaff *et al*. 2007, *Proceedings of the Royal Society, Series B: Biological Sciences* 274: 2035-2040). Suppose that you collect data on the repertoire sizes of all of the male song sparrows with territories in Horticulture Park (i.e., you take a census), and obtain the relative frequency distribution in Table 1.

**Table 1:** Repertoire sizes of male song sparrows in Hort Park.

|  |  |
| --- | --- |
| Number of songs | Proportion of males |
| 4 | 0.032 |
| 5 | 0.144 |
| 6 | 0.216 |
| 7 | 0.209 |
| 8 | 0.129 |
| 9 | 0.113 |
| 10 | 0.087 |
| 11 | 0.028 |
| 12 | 0.042 |
| Total | 1.000 |

Please use this information to answer the following questions. Show your work for all calculations. **For this question, you may round to 3 decimal places**.

1. [1 point] Let the random variable represent the repertoire size of a randomly selected male from this population.
   1. Is discrete or continuous? Please explain your answer.
   2. \*Given your answer in Question 2a(i), does the second column in Table 1 describe a probability mass function (), or a probability density function ()?
2. \*[1 point] If you were to randomly sample a single male sparrow from this population, how many songs would you expect (or predict) him to have in his repertoire?
3. [2 points] If you calculated it correctly, your answer in Question 2b will not be an integer. Explain why this is not a problem (hint: does the answer in Question 2b really describe an individual sample unit?).
4. Not all distributions in nature are symmetric or unimodal. Some are highly skewed. For example, when a sick individual coughs or sneezes, pathogenic viruses or bacteria can be dispersed inside aerosolized droplets. Fernandez *et al*. (2019, <https://doi.org/10.1098/rsif.2018.0779>) assessed the survival time of *E. coli* MRE162 (a bacteria) in these droplets, yielding a heavy-tailed, right-skewed distribution of survival time similar to the one shown in Figure 1. The domain for this distribution is . Suppose that its density function is given by the equation,

**Figure 1:** Probability density function for survival of *E. coli* in bioaerosols.

Please use this equation to answer the following questions.  **Round to 3 decimal places.**

* 1. \* [1 point] In this problem, what does the random variable describe?
  2. \* [1 point] Use the probability density function provided above to find the cumulative distribution function, . Please show your work. At a minimum, show (i) the symbolic equation for the integral that sets up the CDF calculation, (ii) how values are substituted into the equation, (iii) the indefinite integral, and (iii) the solution. **The solution will be an equation, not a number; it will have an in it.**

Lecture 3.4 and the Calculus Review Notes in the cheat sheets section of the website may be helpful. If you would like to, you may also use an automated solver such as a graphing calculator or Wolfram Alpha to calculate the indefinite integral (the solver is at <https://www.wolframalpha.com/>; for examples, see: [https://www.wolframalpha.com/ examples/mathematics/calculus-and-analysis/integrals/](https://www.wolframalpha.com/%20examples/mathematics/calculus-and-analysis/integrals/)). Be aware that the free version of Wolfram Alpha will only give you the indefinite integral.

* 1. Assume that is measured in hours. Use the CDF that you calculated in part b to find the following values. Please show your work and round to 3 decimal places.
     1. \* [1 point] Find the 95th percentile of survival time for *E. coli* MRE162 under the tested conditions. Hints:
        1. If , then
        2. You can check your answer by plugging it into the CDF and confirming that you get the correct proportion that you started with.
     2. \* [1point] Give a biological interpretation for the number calculated in question 3c(ii). What does it mean?
     3. \*[1 point] Find the **percentage** of *E. coli* that survive for more than 2 days.