Lab 4: Probability

Goals

- Get practice solving probability problems

Question #1 A bag has 3 black marbles and 4 white marbles. You reach into the bag and draw a marble, and then draw another marble without replacing the first one.

a. What is the chance that the first draw is a white marble?

b. If the first draw was a white marble, what is the chance that the second draw is a black marble?

c. What is the chance that both marbles that you have removed are the same color as each other?

Question #2 In order for an annual plant to flower, it must germinate and then survive for several weeks. A group of botanists estimate that 70% of seeds will germinate in a particular grassland. One quarter of the plants that germinate will survive until they flower.

What is the chance that a randomly selected seed from this grassland will flower?

Question #3 (This is question 19 from the text): DNA is made up of A, C, G, and T. Different chromosomal regions have different frequencies of these nucleotides. Assume that in the first region that you study, the four nucleotides have equal frequencies. In a second region, the nucleotide frequencies are 20% A, 30%C, 30%G, and 20%T.

a. If you choose a nucleotide randomly from each region, what is the probability that both nucleotides will be the same base?

b. Assume that nucleotides occur independently within each region and you random sample a three-nucleotide sequence from each of the 2 regions. What is the chance that these two triples are identical?

Question #4 (Assignment problem # 25 on page 125) A seed randomly blows around a variable habitat: If it lands on high-quality soil it has a 0.8 chance of survival. If it lands on medium-quality soil it has a 0.3 chance of survival. A low-quality soil gives it only a 0.1 chance of survival. The three soil types (high, medium, and low) are present in proportions of 30:20:50, respectively. The probability of landing in a soil type is simply the proportion of the environment that is that habitat type.

a. Draw a probability tree to determine the probabilities of survival in all circumstances.

b. What is the probability of survival of a seed (assuming that a seed lands)?

c. Assume that a seed has a 0.2 chance of dying before it lands. What is the overall probability of survival?

Question #5: Territorial males of a species of fish are more obvious to females and to predators. In a particular stream, 60% of these territorial males will die before the breeding season. Another set of "cryptic" males are smaller and do not defend territories. Only one quarter of these cryptic males will die before the breeding season. Most females prefer to mate with territorial males. If a territorial male is able to survive until breeding season, there is a 70% chance that he will fertilize a clutch. Only half of the cryptic males who survive until the breeding season will fertilize a clutch of eggs. You can assume that a male cannot fertilize more than one clutch and that all males die after the breeding season ends.

(a) Circle One: True or False: The events "the male dies before breeding season" and "the male fertilizes a clutch of eggs" are mutually exclusive.

(b) Circle One: True or False: The events "the male survives until breeding season" and "the male fertilizes a clutch of eggs" are independent.

(c) If you find a (live) territorial male at the beginning of the **breeding** season (after all of the seasonal mortality events have occurred), what is the probability that he will fertilize a clutch of eggs?

(d) Imagine that you are able to track a cryptic male for his entire lifetime (starting at the beginning of the year **before** any mortality events occur). What is the probability that he would fertilize a clutch of eggs?

Question #6

A disease is found in 8% of the population. There is a diagnostic test for the disease, but the test has a false positive rate of 10% (in other words 10% of healthy people will have positive result for the test), and a false negative rate of 5% (someone with the disease will not have a positive result for the test 5% of the time).

a. Which of the following statements is a correct summary of the data given in the problem (more than one can be correct!):

(a) $\Pr(\text{disease} \mid \text{pos. test}) = 0.08$

(b) $\Pr(\text{disease}) = 0.08$

(c) $\Pr(\text{pos. test} \mid \text{disease}) = 0.1$

(d) $\Pr(\text{pos. test} \mid \text{not disease}) = 0.1$

(e) $\Pr(\text{neg. test} \mid \text{disease}) = 0.1$

(f) $\Pr(\text{neg. test} \mid \text{disease}) = 0.05$

b. We would like to know the probability that someone with a positive test result actually has the the disease.

(a) Write down in notational form, the probability that you are being asked to calculate.

(b) Calculate the probability.