

Week 5 Reading Questions

Mercy Melo (no collaborators)

Question 1: The sample space of randomly choosing 2 acorns from a small collection of 3 different oak species would be 9. You could choose either of the 3 species for your first acorn and then have an equal random chance of choosing any of the 3 species again for your 2nd acorn (see table below).

First Acorn	Second Acorn
Macrocarpa	Macrocarpa
	Rubra*
	Alba*
Rubra	Macrocarpa*
	Rubra
	Alba*
Alba	Macrocarpa*
	Rubra*
	Alba

Question 2: There are 3 ways to collect two acorns of the same species. You could collect 2 Macrocarpa, 2 Rubra, or 2 Alba.

Question 3: There are 6 ways to collect two acorns of different species (shown as asterisk in table above).

Question 4: The probability of one of your chosen acorns being *Q. alba* is $1/3 = 0.33$ because all three species can be found in equal distribution (and therefore you have an equal chance of picking up either of the three species).

Question 5: The probability of your 2nd chosen acorn being *Q. macrocarpa* would also be $1/3 = 0.33$. This is because there are a very large number of acorns on the ground, so the first acorn you picked up is likely not to affect the pool of acorns that you're choosing your 2nd acorn from (they are independent events).

Question 6: Since your 1st choice of acorn is independent of your 2nd choice and you already know what species your 1st choice is, the probability of your 2nd choice being *Q. alba* would still be $1/3 = 0.33$. Since the sample size is so large, the removal of your 1st acorn will have little effect on your probability of pulling the same species for your 2nd acorn.

Question 7: The probability of both chosen acorns being *Q. rubra* would be the product of both independent events occurring: $0.33 * 0.33 = 0.109$. Another way to look at this would be to describe this event as a single permutation within our sample space set of 9 permutations ($1/9 = 0.109$).

Question 8: Similar to question 7, the probability of choosing one *Q. alba* and one *Q. rubra* would be the product of each of their separate probabilities: $0.33 * 0.33 = 0.109$ but multiplied by 2 because you could choose the *Q. alba* first or the *Q. rubra* first = 0.218.

Question 9: Similar to question 7, choosing a *Q. alba* first followed by a *Q. rubra* would represent a single permutation within our sample space of 9 permutation options. The probability of this event occurring then would be $1/9 = 0.109$.

Question 10: Sample space of Poisson distributions is infinite.

Question 11: The sample space of a Binomial distribution is the number of trials. In this case, it is 10.

Question 12: Both the Poisson and Binomial distributions are based on outcomes of trials having 2 values- either a 1 (presence/success) or a 0 (absence/failure). This type of predicted outcome (a binary response) is perfectly suited for count data as counts represent the presence or absence of an event occurring. Both of these distributions are also well-suited for counts because they take each event as an independent occurrence with an equal probability of occurring. With counts, most events are independent of each other and are equally probable, making both the Binomial and Poisson distributions great fits for count data.

Depending on the nature of the count, and more specifically the number of trials, you can choose which of these two distributions is most appropriate. For studies with a fixed number of trials (such as a study surveying toad presence/absence on 5 different plots), the Binomial distribution is best suited. For studies with an infinite number of trials (such as counting hawks at a watch site during migration), a Poisson distribution would be best.

Question 13: A binomial distribution is used for data that has a binary outcome (presence/absence in the case of counts) and is bounded within a fixed number of trials (unlike Poisson, where trials are infinite). Binomial distributions are therefore better count models than Poisson distributions when the number of trials is set. An example would be counting the presence/absence of chickadees at 15 bird feeders, each of which is only surveyed once. Since we are limiting the sample space to 15 (the number of trials conducted), the Binomial distribution is the best fit for this set of data.