Comments on
Poisson Distribution Presentation Notes
(second draft)
By Jeremy Mahoney

John McCuan

October 5, 2023

1 Probability and Frequency

Your second draft looks great, and your presentation was great as well.

As I've thought about your project, there is one other question that occurs to me could be worth considering.[1] In some sense, I would say this is a kind of pedagogical and theoretical weakness of your presentation if there is one. The question is:

> Why should one use the probability
>
> $$p = \frac{\lambda}{n}$$
>
> for the probability in the binomial distribution?

You simply say: "Assuming the mole is equally likely to emerge at any time, the probability the mole appears during an interval is approximately λ/n." You also

[1] I'm not suggesting this project presentation is "incomplete" or consideration of this topic is something to which you (Jeremy) should give specific attention this semester. But it may be something you wish to follow up on in the future. On the other hand, it may be worth considering if you wish to think deeply about something related while this topic is on your mind. And on the third hand you may not even appreciate the question or think it is a valid one. Obviously, a lot of the questions I have about probability (and statistics...and combinatorics) are not considered to be of much interest to the experts in the field. These are not topics addressed in textbooks.

seem to go somewhat in the direction of this question in your Exercise 2, but that seems to be a somewhat different question. This seems something worth taking on directly: Why the ratio λ/n?

I suggested one possible answer: the probability $p = \lambda/n$ is the probability for which the average value, i.e., expectation of the binomial distribution with n trials is λ. Thus, the probability $p = \lambda/n$ is chosen so that the average value matches the observed average number of appearances of the mole over the entire time interval t. Still, while this is **an** explanation, it does not strike me as the most insightful one.

Here is something in the direction of an alternative: Imagine first that over some particular time interval of length t, there are exactly the average number λ appearances of the mole. Now I'm going to suggest something a little strange: Think of each appearance of the mole as a "thrown dart." This anology is suggested by Orloff and Booth. The board is the time interval $[0, t)$, and we imagine the interval is divided into n intervals of equal length, each of which as you suggest is "equally likely" to be the location the dart hits. If there are 100 subintervals, then the probability the dart hits any one location (a success on that trial) is 1/100. If λ darts are thrown, or λ appearances of the mole take place, then by some principle of adding probabilities the probability that a particlar unterval $[kt/100, (k+1)t/100)$ is hit is $\lambda/100$. Hence, the natural probability for success in each of the n trials is λ/n. At least this is some kind of explanation.

More generally, the (or a) relationship between a probability and **ratios of outcomes** is a topic which seems, on the one hand, to be somewhat intuitively "clear." On the other hand, there should be some well-defined principles of this relation which do not seem to be anywhere spelled out. This same kind of question was lurking behind the very first discussion of the course concerning the base rate fallacy. You may recall that I translated the given probabilities, e.g., the base rate or the rate of false negatives, into ratios of outcomes for a specific population. Why are such translations valid, and how are they made?

Like with many other topics in counting, probability, combinatorics, etc., it seems the standard practice is to just kind of spit out an answer and hope whoever is listening agrees without giving the matter too much critical thought. At least some of these "fuzzy areas" are susceptible to further explanation and at least the formulation of some kind of clearly stated principles. It is interesting that a few such principles, e.g., Orloff and Booth's *rule of product* (Class 1, installment 2 notes) are often attempted but are usually too vague or ill formulated to actually be correct.

2 Additional comments and suggestions

1. I guess I did your Exercise 3 in class. (Sorry about that.)

2. If I have time I may type up solutions to your other exercises...or maybe some other student would like to take that on.

3. From my point of view, I think some of the text and the mathematical calculations in your project could be "tightened up," but I'm not sure this would be a real improvement for most of your readers.

4. You mention at the beginning "guessing the number of days you would see a certain number of appearances." I guess this comes back to some kind of question of the interpretation of frequency. At any rate, it seems to me a more direct interpretation of the Poisson distribution is to contemplate, for a given number of appearances, the ratio of days that number of appearances is observed to the number of all days of observation: If I go out and count for 100 days, on how many of those days will I see the mole pop up 4 times?

5. Some explanation of "big O" notation might have been good to include. I did try to offer something in the final assignment.