Choose the Right Probability Distribution (Education and Science Communication)

Dr. F.J. Rodenburg, L. Outhuis & L. Bikker

2022-08-08

These exercises assume you are familiar with the basics of **probability distributions**. That means that you should know when the following distributions are good approximations or not:

- The normal distribution
- The Poisson distribution
- The binomial distribution

For these questions, you have to **identify the outcome** from a short description of a study, and **choose the appropriate probability distributions** that could be used in the analysis.

Q1 — Genetically Modified Food

To explore the societal acceptance of genetically modified food, a survey is conducted. The first question asks whether the respondent has ever knowingly eaten genetically modified food before, or not. What is the outcome and what probability distribution can be associated with it?

Answer:

Q2 — Vaccine Acceptance

Research suggests older individuals were more willing to take vaccines during the COVID-19 pandemic.[1] To study the relationship between age and vaccine acceptance, a survey could be conducted. What would be the outcome and what probability distribution can be used to approximate it?

Answer:

Q3 — Writing an Article

Scientific writing is an important skills when writing for publication. When published, the number of times that an article is cited by other works is a way to measure the impact of the article and the author. What probability distribution can be used to approximate the number of citations articles receive?

Answer:

Q4 — Effectiveness of Course Material

Imagine that you are working on new biology course material to introduce to students. To test the effectiveness of this new course material, you designed a test about the content of the course material and divide the participants of the test in two groups. One group is making the test after studying the old course material, the other group is making the test after studying the new course material. You note down the average scores on this test of both groups. What will be the outcome and what probability distribution can be used to approximate this?

Answer:

Q5 — Likert-Scale

When analyzing the results of a survey, there are 10 questions about the ethics of laboratory animals and 10 questions about the ethics of genetically modified organisms (GMOs). The respondents answered all these questions on a Likert scale (e.g. a scale from 1–5, with 1 = strongly disagree; to 5 = strongly agree).

1. Can you approximate Likert-scale data with any of the three types of probability distributions covered so far? Explain.
2. If you were to average the responses of similar questions, what probability distribution could be used to approximate this outcome?

Answer:

Q6 — Gender Preference

In a small study, 5 females and 5 males of comparable age distribute flyers for a new exhibition in Naturalis. If researchers want to assess whether people are more inclined to accept flyers from males or females, what would be the outcome and what probability distribution could be used to approximate it?

Answer:

References

1. Truong, J., Bakshi, S., Wasim, A., Ahmad, M. & Majid, U. What factors promote vaccine hesitancy or acceptance during pandemics? A systematic review and thematic analysis. *Health Promotion International* **37**, (2021).