**Chapter 3 – Analyzing a multicategory response**

The previous two chapters provided analysis methods for when there were binary responses. The purpose of this chapter is to generalize some of these previous methods to allow for more than two response categories. Examples include:

* Canadian political party affiliation – Conservative, New Democratic, Liberal, Bloc Quebecois, or Green
* Chemical compounds in drug discovery experiments – Positive, blocker, or neither
* Cereal shelf-placement in a grocery store – Bottom, middle, or top
* Beef grades – Prime, choice, select, standard, utility, and commercial
* Five-level Likert scale – Strongly disagree, disagree, neutral, agree, or strongly agree.

For these examples, some responses are ordinal (e.g., Likert scale) and some are not (e.g., chemical compounds). We will investigate both ordinal and nominal (unordered) multicategory responses within this chapter.

**Section 3.1 – Multinomial probability distribution**

The multinomial probability distribution is the extension of the binomial distribution to situations where there are more than two categories for a response.

Notation:

* Y denotes the response category with levels of j = 1, …, J
* Each category has a probability of πj = P(Y = j).
* n denotes the number of trials
* n1, …, nJ denote the response count for category j, where 

The probability mass function for observing particular values of n1, …, nJ is



The dmultinom() function evaluates this function, and the rmultinom() function simulates observations. When J = 2, the distribution simplifies to the binomial distribution.

For n trials, the likelihood function is simply the probability mass function. The maximum likelihood estimate of πj is .

Questions:

* What would the probability mass function look like if there was only one trial?
* What would the likelihood function be if 1 trial was observed, then another 1 trial was observed independently of the previous trial, … , so that there were m total one-trial sets observed in succession?
* What would the likelihood function be if n trials were observed, then another n trials were observed independently of the previous trial, … , so that there were m total n-trial sets observed in succession?
* Consider the same scenario as in the last question, but now with the possibility of different probabilities for each set. What would the likelihood function be?

Example: Multinomial simulated sample (Multinomial.R)

As a quick way to see what a sample looks like in a multinomial setting, consider the situation with n = 1,000 trials, π1 = 0.25, π2 = 0.35, π3 = 0.2, π4 = 0.1, and π5 = 0.1. Below is how we can simulate a sample:

> pi.j <- c(0.25, 0.35, 0.2, 0.1, 0.1)

> set.seed(2195) # Set seed to reproduce the sample

> n.j <- rmultinom(n = 1, size = 1000, prob = pi.j)

> data.frame(n.j, pihat.j = n.j/1000, pi.j)

 n.j pihat.j pi.j

1 242 0.242 0.25

2 333 0.333 0.35

3 188 0.188 0.20

4 122 0.122 0.10

5 115 0.115 0.10

Suppose there are m = 5 separate sets of n = 1000 trials.

> set.seed(9182)

> n.j <- rmultinom(n = 5, size = 1000, prob = pi.j)

> n.j

 [,1] [,2] [,3] [,4] [,5]

[1,] 259 259 237 264 247

[2,] 341 346 374 339 341

[3,] 200 188 198 191 210

[4,] 92 106 89 108 107

[5,] 108 101 102 98 95

> n.j/1000

 [,1] [,2] [,3] [,4] [,5]

[1,] 0.259 0.259 0.237 0.264 0.247

[2,] 0.341 0.346 0.374 0.339 0.341

[3,] 0.200 0.188 0.198 0.191 0.210

[4,] 0.092 0.106 0.089 0.108 0.107

[5,] 0.108 0.101 0.102 0.098 0.095

Notice the variability from one set to another.