Homework 1: Stat 510

Moulinath Banerjee

September 13, 2020

(1) (1) Problems from CB: 1.36, 1.39, 1.46. 1.51, 1.19. 1.21, 1.24, 1.31 (a) and (c).

(2) Consider a multinomial experiment where at any given trial, m distinct mutually exclusive and exhaustive outcomes can occur (e.g. a die is thrown and one of the six faces has to come up). Let p_j be the probability that the j'th outcome turns up in the trial. Then $\sum_{j=1}^{m} p_j = 1$. Consider repeating the multinomial experiment independently (i.e. you keep throwing the die repeatedly) till $m_1, m_2, \ldots, m_k$ outcomes of types 1 through k (where $k < m$) are obtained. What is the probability that you need N trials of the experiment for this to happen?

(3) **Simpson's paradox:** Consider M different departments at a university. The chance of a Type 1 student applying to a specific department k is p_k and that of a Type 2 student applying to the same department is q_k. We assume that $\sum p_k = \sum q_k = 1$. The probability of a Type 1 student being admitted to the k'th department once they apply is $a_{k,1}$ and that of a Type 2 student being admitted once they apply is $a_{k,2}$. Find the probability that a Type 1 student is admitted to the university, and also the probability of a Type 2 student being admitted.

Assume that each department k is either neutral to type or actually slightly favors Type 2 students, so that $a_{k,1} \leq a_{k,2}$ for each k (but not by much). Argue that despite this, the admissions data for the whole university taken together may demonstrate a huge bias towards Type 1 students. Something of this type did actually happen in Berkeley. See, for example, https://www.refsmmat.com/posts/2016-05-08-simpsons-paradox-berkeley.html. In other words, global 'biases' can be often be manufactured despite the non-existence of bias (or a reverse bias) simply because of the tension between conditional and unconditional probabilities.