

PS1: Introduction to Probability and Statistics

Deepthi Gorthi
AY250: Stellar Populations

September 4, 2016

Problem 1.

(a) In 1995, they introduced blue M&M's. Before then, the color mix in a bag of plain M&M's was 30% brown, 20% yellow, 20% red, 10% green, 10% orange, and 10% tan. Afterward, it was 24% blue, 20% green, 16% orange, 14% yellow, 13% red, 13% brown.

Suppose there are two bags of M&M's, one from 1994 and one from 1996 and you are randomly given one M&M from each bag. One is yellow, one is green. Using Bayes's theorem and a probability table to determine the relative probability that the yellow M&M came from the 1994 bag.

(b) Evaluate the "Evidence" and determine the normalized probability that the yellow M&M came from the 1994 bag.

Hint: This is similar to how we wrote out the Monty Hall problem in class.

Solution:

(a) The given distribution of M&M's can be summarized as follows:

Colour	1994	1996
Brown	0.3	0.13
Yellow	0.2	0.14
Red	0.2	0.13
Green	0.1	0.2
Orange	0.1	0.16
Other	0.1	0.24

Given data: One M&M was drawn from each sample and one of them is yellow and another green. To evaluate the relative probability that the yellow M&M was drawn from the 1994 bag (and hence the green from the 1996 bag), define the two hypothesis as:

H_1 : The yellow M&M is drawn from the 1994 bag and the green M&M from the 1996 bag.

H_2 : The yellow M&M is drawn from the 1996 bag and the green M&M from the 1994 bag.

Since we have no prior knowledge about either hypothesis, we should begin with flat priors- either hypothesis is equally likely. For probability of the data given the hypothesis we can multiply the probabilities of drawing the green and yellow M&Ms from the respective bag, since they are independent events.

Hypothesis	P(H)	P(D H)	P(D H)*P(H)
H_1	1/2	(0.2)*(0.2)	0.02
H_2	1/2	(0.14)*(0.1)	0.007

The relative probability that the yellow M&M was drawn from the 1994 bag is 0.02.

(b) The ‘evidence’ is also the probability of the data given all the possible hypotheses. This can be obtained by summing over the likelihoods of all the hypotheses.

$$P(D) = 0.02 + 0.007 = 0.027 \quad (1)$$

Hence the normalized probability that the yellow M&M came from the 1994 bag is:

$$P(H_1|D) = \frac{P(D|H_1) * P(H_1)}{P(D)} = \frac{0.02}{0.027} \sim 74\% \quad (2)$$

The normalized probability of our hypothesis that the yellow M&M was drawn from the 1994 bag is 74%.