STAT 2857A – Lecture 4 Examples and Exercises

Solutions – Revised 22/09/24

Example 4.1

a) The conditional probabilities defined in the question are:

   i) The probability that a person assigned to be female at birth is colour blind:

$$P(B|F) = 1/200.$$

   ii) The probability that a person assigned to be male at birth is colour blind:

$$P(B|F') = 1/12.$$

b) The probability that a randomly selected newborn is female and colour blind is:

$$P(F \cap B) = P(B|F)P(F) = .4878 \times 1/200 = .002439.$$

The probability that a randomly selected newborn is male and colour blind is:

$$P(F' \cap B) = P(B|F')P(F') = (1 - .4878) \times 1/12 = .04268.$$

Therefore, the probability that a newborn assigned to be female at birth is colour blind is .00244. The probability that a newborn assigned to be male at birth is colour blind is .0427.

c) Note that F and F' partition the sample space – i.e., they are mutually exclusive and exhaustive. Everyone is assigned to be either male or female at birth. Applying the law of total probability we have

$$P(B) = P(F \cap B) + P(F' \cap B) = .04512.$$

Therefore, the overall probability of being born colour blind is about .0451 or 4.51%.

d) The event we are interested in is that a baby is assigned to be male given that it is colour blind, $F'|B$. To compute this probability, we need to apply Bayes' rule so we can switch the direction of conditioning:

$$P(F'|B) = \frac{P(B|F')P(F')}{P(B)} = \frac{1/12 \times (1 - .4878)}{.04512} = .9460.$$

Hence, the probability that a newborn that is colour blind is assigned to be male at birth is .946.

e) Let C be the event that a randomly selected person is born on an odd day. The event we are interested in is $B \cup C$. Ignoring leap years and assuming that the outcomes are equally likely, the probability of C is

$$P(C) = 187/365 = .5123.$$

Applying the rule for the probability of a union we have

$$P(B \cup C) = P(B) + P(C) - P(B \cap C).$$

We know that

$$P(B) = .04512 \quad \text{and} \quad P(C) = .5123.$$

We then need to compute

$$P(B \cap C) = P(B|C)P(C).$$

it seems odd to believe that whether or not someone is colour blind depends on the day of the year in which they are born. In this case, we might assume that $P(B|C) = P(B|C')$. The law of total probability then tells us that

$$P(B) = P(B|C)P(C) + P(B|C')P(C') = P(B|C)[P(C) + P(C')] = P(B|C)$$

so that

$$P(B \cap C) = P(B)P(C).$$

This implies that B and C are independent events, which we will explore next day. Putting this together we find that

$$P(B \cap C) = .04512 + .5123 - .04512(.5123) = .5343.$$

The probability that someone is colour blind or born on an odd day of the month is about .534.

Exercise 4.1

a) Let H be the event that an individual is planning to vote for Kamala Harris, U the event that the individual comes from Utah, and V the event that they come from Vermont. The conditional probabilities defined are

$$P(H|U) = .28$$

and

$$P(H|V) = .70.$$

b) By the multiplication rule

$$P(U \cap H) = P(H|U)P(U).$$

Assuming that the number of voters in each state is proportional to the population size, we find

$$P(U) = 3,417,734/334,235,923 = .0102.$$

Then

$$P(U \cap H) = .28(.0102) = .002856.$$

The probability that a randomly selected US voter is from Utah and supports Kamala Harris is .00286.

c) Note that U and U' form a partition of the US population (the events are mutually exclusive and exhaustive). The law of total probability then implies that

$$P(H) = P(H|U)P(U) + P(H|U')P(U').$$

Solving for $P(H|U')$ gives

$$P(H|U') = \frac{P(H) - P(H|U)P(U)}{P(U')}.$$

Assuming that the number of voters in each state is proportional to the population size, we find

$$P(U) = 3,417,734/334,235,923 = .0102.$$

and

$$P(U') = 1 - P(U) = .9898.$$

Then

$$P(H|U') = \frac{.48 - .28(.0102)}{.9898} = .482.$$

Given the information, the probability that a voter not from Utah plans to vote for Kamala Harris is .482. I.e., 48.2% of voters from outside of Utah plan to vote for Harris.

d) Applying Bayes' rule

$$P(U|H) = \frac{P(H|U)P(U)}{P(H)}.$$

We know that $P(H|U) = .28$ and $P(H) = .48$. Assuming that the proportion of voters in each state is proportionate to the population we have

$$P(U) = 3,417,734/334,235,923 = .0102.$$

Hence

$$P(U|H) = \frac{.28(.0102)}{.48} = .005965.$$

Similarly

$$P(V) = 647,464/334,235,923 = .00194$$

and

$$P(V|H) = \frac{.70(.00194)}{.48} = .002825.$$

Hence, the probability that a randomly selected voter who supports Kamala Harris comes from Utah (.00597) is higher than the probability that they come from Vermont (.00283). This may seem counterintuitive at first. How can it be more likely that a Harris supporter comes from Utah when she has half the support in that state? The answer is simply that there are many more people in Utah. Despite her low support percent-wise, there are approximately 947 thousand Harris supporters in Utah and only 452 thousand in Vermont.

e) The key assumption is that the number of voters in a state is proportionate to the population size. Nate Silver, and American statistician, economist, and writer, famously predicted the outcome of the election in all 50 states in 2012, and then got things completely wrong in 2016. The problem – voter turnout. Silver predicted confidently (with probability .71) that Hillary Clinton would win the 2016 election. In the end, she was forced to concede to Donald Trump. The problem for Silver was that supporters of Donald Trump were much more likely to vote meaning that the polls did not accurately predict the election even if they did accurately reflect the overall support within each state.