Juana Sanchez, Juana Sanchez, Juana Sanchez **Homework 1**

PART I. (a)"The Brexit debate and the prosecutor's fallacy" by Gianlua Baio. Significance, June 11, 2019

Population of interest: Voters that voted in the referendum.

Let L= "voted to Leave European Union"

RaS="The voter was racist and Anti-semite"

A journalist, Self, made a statement claiming that P(L |Ras) ≈ 1. Another journalist, Francois, interpreted this as P(Ras|L) ≈ 1. This is called the Prosecutor's fallacy in Law.

The author of the article claims that Francois interpretation of Self's statement would be true only if 50% of the voters in the Brexit referendum were Ras. This follows from Bayes theorem, the statement made by Self and what is known of the proportion of voters that voted Leave (P(L)= 0.5189)

$$\Pr(RaS \mid L) = \frac{\Pr(L \mid RaS)\Pr(RaS)}{\Pr(L)}$$

If we assume P(L|RaS) = 1 (Self's claim), P(RaS) =0.5, and P(L) = 0.5189, then P(RaS|L) = 0.9635768 ≈ 1.
But if P(RaS) was smaller than 0.5, say 0.2, then P(RaS|L) =0.3854307.

The point then is that Francois does not know Bayes theorem, and has failed to account for P(RaS), the prior probability in the numerator that a voter is racist and anti-semite. If that probability is small, the claim that Francois is making about Self's statement is wrong.

The message is: Do not ignore the prior probability and use Bayes theorem properly

(a) Maureen Grays' video"Judgments under Uncertainty. Applications of probabilistic reasoning"

Population: Bikini Bottom

Let Y= "An individual in Bikini Bottom is tough" and X= "sponge bob-type individuals in Bikini Bottom"

Bouncer wants to figure out P(Y|X) but because it is difficult, the bouncer uses P(X|Y) instead (which is incorrect). This is called the "representativeness heuristic" in Psychology. It is the same problem as in the Prosecution Fallacy.

This is something that is known from the membership of the club.

By Bayes Theorem,

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Say P(X|Y) is 0.05 (5% of Bikini Bottom tough individuals are like Sponge Bob)
P(X) =0.1 (10%of Bikini Bottom individuals are sponge Bob-type of individuals).

If P(Y) = 0.01 (only 1% of the population in Bikini Bottom is tough), then P(Y|X)= 0.005.

If P(Y) = 0.5 (half of the population in Bikini Bottom is tough), then P(Y|X)= 0.25.
If P(Y) =0.7 (70% of the population in Bikini Bottom is tough), then P(Y|X) = 0.35.
If P(Y)= 1 then P(Y|X) = 0.5.

Thus, even though in appearance (based on P(X|Y) =0.05, appearances based on the Bouncer experience among tough individuals, it seems that Sponge Bob is not tough), if the prior probability of an individual of the population being tough is taken into account, Sponge Bob could have a high probability of being tough, up to 50% chance.

P(Y) is, in Maureen Gray's video the base rate. The message is: do not ignore the prior probability and use Bayes theorem properly.

(c) "A case study in conditional probability" by Lakshya Malhotra.

Based on data for Jun, Jul, Year, Floods the author codes the data for June as follows:

Population: June of all years in Kerala.

F="flooding", J = "More than 500 mm rain in June"

Interested in P(F |J)=54/(39+54) = 0.5806452 Uses a basic table to calculate

JUN_GT_500	0	1
FLOODS		
0	19	39
1	6	54

If interested in P(J|F) we use Bayes theorem.

$$P(J|F) = \frac{P(F|J)P(J)}{P(F)}$$

Again from the table, P(J/F) = 54/(54+6) = 0.9

Malhotra, like Maureen Gray and Gianlua Baio, is also trying to bring up the distinction between two different conditional probabilities, in this case P(F|J) and P(J|F) and points out how the P(J|F) depends on the prior probability P(J). We should not be surprised to see rain larger than 500 mm in June of years when there was flooding because the prior probability P(J) is very high (June is one of the most rainy months of the year), the author says.

There is no mention of people confusing P(J|F) as P(F|J) or viceversa,as in Gray's video and Baio's article, but Malhotra points out that looking at P(F|J) is useful to understand that all floods are not caused by rain only, since P(F|J)= 0.58. On the other hand, the high P(J) helps explain the high P(J|F).

In contrast with Gray and Baio, Malhotra uses real data, not hypotheticals, to illustrate the points made.

(b) "Fairness in Artificial Intelligence" by Prof Chris Piech

The author is trying to bring up the point that AI can make mistakes for some populations. In order to illustrate that, use is made of several criteria for determining whether an algorithm is fair. All these criteria involve conditional probability. The example is about an algorithm deciding whether a person will repay a loan.

Population: the population represented in the training data used by the algorithm.
G= "algorithm's guess" (1=algorithm guesses that it will repay, 0 = guesses not repay)
T= "Truth" (1=person will repay, 0 =the person will not repay)
D=population (0, or 1, so there are two populations)

$D=0$

	$G=0$	$G=1$
$T=0$	0.21	0.32
$T=1$	0.07	0.28

$D=1$

	$G=0$	$G=1$
$T=0$	0.01	0.01
$T=1$	0.02	0.08

Parity: compares conditional probabilities P(G=1|D=0) and P(G=1|D=1) Is probability of a positive prediction the same for all segments of the population? If yes, the algorithm is fair.

Calibration: Compares conditional probabilities P(G=T|D=0) and P(G=T|D=1) Is the probability that the algorithm is correct the same in all populations? If yes, the algorithm is fair. This means: Is P(G=0,T=0 |D=0) the same as P(G=0,T=0|D=1)? Could also mean: Is P(G=1,T=1|D=0) the same as P(G=1,T=1 |D=1)? If yes, in all cases, the algorithm is fair.

Equality of odds: Compares conditional probabilities P(G=1|D=0, T=1) and P(G=1|D=1,T=1). If those two are the same, the algorithm is fair.

According to the data in the table above, the algorithm does not satisfy parity or calibration but it satisfies equality of odds.

The author, after introducing the criteria, shows also an example on facial recognition algorithms used by some major high tech companies and reports that those showed more accuracy for certain populations. One reason is that the majority demographics in their training data pool have higher weight on the algorithm. Why is that?

Parity

$$P(G=1,T=1|D=0)=\frac{P(D=0|G=1,T=1)P(G=1,T=1)}{P(D=0)}$$

$$P(G=1,T=1|D=1)=\frac{P(D=1|G=1,T=1)P(G=1,T=1)}{P(D=1)}$$

If all in those ratios is the same, except the P(D=0) and P(D=1), it is not hard to see that if P(D=0)> P(D=1), then $P(G=1,T=1|D=0)$> $P(G=1,T=1|D=1)$.

That is, if the training data has many more individuals of one particular population, D=0, then the probability of being right is larger for that population than for the population of which there are not that many in the training data. I other words, the algorithm makes the wrong decision more often for individuals in the population less represented in the training data.

Thus, although Piech uses conditional probability and Bayes theorem, like all other three authors do, the part of Bayes formula that plays a role in the accuracy of an algorithm and its fairness according to the calibration criterion is the denominator of the Bayes formula: the proportion of the population that is represented in the training data. In the other authors, it was the prior probability in the numerator.

PART II.

In (a) a journalist, Self, made a statement claiming that P(L |Ras) ≈ 1. Another journalist, Francois, interpreted this as P(Ras|L) ≈ 1. This is called the Prosecutor's fallacy in Law. The message is: Francois failed to account for the prior probability P(Ras) in the formula for Bayes theorem to calculate P(Ras|L). This is the same message as in (b), where indicates that the Bouncer wants to figure out P(Y|X) but because it is difficult, the bouncer uses P(X|Y) instead (which is incorrect), and ignores the prior probability P(Y). This is called the "representativeness heuristic" in Psychology. It is the same problem as in the Prosecution Fallacy in (a).
In (c), we see the same message as in (a) and (b) appearing indirectly in the discussion that high P(J) helps explain the high P(J|F). The author in (c) is not using the formulas for Bayes but illustrates that point with the table of real data. On the other hand, in (d) **Piech uses conditional probability and Bayes theorem, like all other three authors do, but the part of Bayes formula that plays a role in the accuracy of an algorithm and its fairness according to the calibration criterion is the denominator of the Bayes formula: the proportion of the population that is represented in the training data. In the other authors, it was the prior probability in the numerator.**

Specific Rubric:

file contains names of students (5 pts)
file neat; parts separated; sources labeled; no ambiguous language (10 pts)
file is pdf; name of three students in filename (5 pts)

PART I

Brexit article
I(a) Be specific about what probabilities are which (conditional, total, joint, etc.) [2 pts]
I(a) Key Idea: Confusion about specific probability being discussed [2 pts]
I(a) Key Idea: Prosecutor's fallacy [2 pts]
I(a) Key Idea: Priors Matter [2 pts]
I(a) Key Idea: Bayes Theorem [2 pts]

Maureen Grays' video
I(b) Be specific about what probabilities are which (conditional, total, joint, etc.) [2 pts]
I(b) Key Idea: representativeness heuristic / stereotypes / small samples [2 pts]
I(b) Key Idea: what's wrong with using representativeness [2 pts]
I(a) Key Idea: Priors / Base Rates Matter; Base Rate Neglect [2 pts]
I(a) Key Idea: Bayes Theorem [2 pts]

"A case study in conditional probability"
I(c) Be specific about what probabilities are which (conditional, total, joint, etc.) [2 pts]
I(c) Key Idea: difference between P(F|J) and P(J|F) [2 pts]
I(c) Key Idea: Priors Matter [2 pts]
I(c) Key Idea: Bayes Theorem [2 pts]

"Fairness in Artificial Intelligence"
I(d) Be specific about what probabilities are which (conditional, total, joint, etc.) [2 pts]
I(d) Key Idea: parity [2 pts]
I(d) Key Idea: calibration [2 pts]
I(d) Key Idea: equality of odds [2 pts]
I(d) Key Idea: the proportion of the population represented in the training data [2 pts]

PART II

II all articles / videos discuss Bayes theorem and difficulties with intuitive reasoning [10 pts]
II I(a,b,c) all focus on how the prior is a key element in correct reasoning [10 pts]
II I(b,d) focus on representativeness [10 pts]
II I(a,c) focus on the difficulty of interpreting which conditional probability is being discussed [10 pts]

And partial points given based on how well these items are discussed.