CIS 399 – Homework 2

Sophia Trump
strump@brynmawr.edu

February 19, 2019

Partners: Kennedy Ellison.

Problem 1

Show that $(1 - FNR) = \frac{A}{A+C}$.

From the definition of equality of false negative rates, we know that $(1 - FNR)$ can be rewritten as

$$1 - \frac{C}{A+C} \tag{1}$$

Simplifying (1),

$$\begin{aligned}
1 - \frac{C}{A+C} &= \frac{1}{1} - \frac{C}{A+C} \\
&= \frac{1(A+C)}{1(A+C)} - \frac{C}{A+C} \\
&= \frac{A+C}{A+C} - \frac{C}{A+C} \\
&= \frac{A+C-C}{A+C} \\
&= \frac{A}{A+C}
\end{aligned}$$

Thus, from the above, $(1 - FNR) = \frac{A}{A+C}$. QED.

Problem 2

Show that $(1 - PPV) = \frac{B}{A+B}$.

From the definition of equality of positive predictive value, we know that $(1 - PPV)$ can be rewritten as

$$1 - \frac{A}{A+B} \tag{2}$$

Simplifying (2),

$$\begin{aligned} 1 - \frac{A}{A+B} &= \frac{1}{1} - \frac{A}{A+B} \\ &= \frac{1(A+B)}{1(A+B} - \frac{A}{A+B} \\ &= \frac{A+B}{A+B} - \frac{A}{A+B} \\ &= \frac{A+B-A}{A+B} \\ &= \frac{B}{A+B} \end{aligned}$$

Thus, from the above, $(1 - PPV) = \frac{B}{A+B}$. QED.

Problem 3

Show that $\frac{BR}{1-BR} = \frac{A+C}{B+D}$.

From the definition of the base rate of a population, BR can be rewritten as

$$\frac{A+C}{|P|} \tag{3}$$

Substituting (3), $\frac{BR}{1-BR} = \frac{A+C}{B+D}$ becomes

$$\frac{\frac{A+C}{|P|}}{1 - \frac{A+C}{|P|}} \tag{4}$$

From the definition of $|P|$, (4) can be further rewritten as

$$\frac{\frac{A+C}{A+B+C+D}}{1 - \frac{A+C}{A+B+C+D}} \tag{5}$$

Simplifying 5,

$$\begin{aligned} \frac{\frac{A+C}{A+B+C+D}}{1 - \frac{A+C}{A+B+C+D}} &= \frac{\frac{(A+C)(A+B+C+D)}{A+B+C+D}}{1 - \frac{(A+C)(A+B+C+D)}{A+B+C+D}} \\ &= \frac{A+C}{(A+B+C+D)-(A+C)} \\ &= \frac{A+C}{B+D} \end{aligned}$$

Thus, from the above, $\frac{BR}{1-BR} = \frac{A+C}{B+D}$. QED.

Problem 4

Using the results from Problems 1, 2, and 3, show that:

$$FPR = (\frac{BR}{1-BR})(\frac{1-PPV}{PPV})(1-FNR)$$

From Problem 3, we know $\frac{BR}{1-BR}$ can be rewritten as

$$\frac{A+C}{B+D} \tag{6}$$

From Problem 2, we know $(1 - PPV)$ can be rewritten as

$$\frac{B}{A+B} \tag{7}$$

From Problem 1, we know $(1 - FNR)$ can be rewritten as

$$\frac{A}{A+C} \tag{8}$$

Substituting (6), (7), and (8), $(\frac{BR}{1-BR})(\frac{1-PPV}{PPV})(1 - FNR)$ becomes

$$(\frac{A+C}{B+D})(\frac{\frac{B}{A+B}}{\frac{A}{A+B}})(\frac{A}{A+C}) \tag{9}$$

Simplifying (9),

$$\begin{aligned}
(\frac{A+C}{B+D})(\frac{\frac{B}{A+B}}{\frac{A}{A+B}})(\frac{A}{A+C}) &= (\frac{A+C}{B+D})(\frac{\frac{B(A+D)}{A+B}}{\frac{A(A+B)}{A+B}})(\frac{A}{A+C}) \\
&= (\frac{A+C}{B+D})(\frac{B}{A})(\frac{A}{A+C}) \\
&= \frac{(A+C)BA}{(B+D)A(A+C)} \\
&= \frac{BAA+BAC}{(AB+AD)(A+C)} \\
&= \frac{BAA+BAC}{AAB+AAD+CAB+CAD} \\
&= \frac{BA(A+C)}{A(AB+AD+CB+CD)} \\
&= \frac{B(A+C)}{AB+AD+CB+CD} \\
&= \frac{B(A+C)}{(A+C)(B+D)} \\
&= \frac{B}{B+D}
\end{aligned}$$

From the definition of equality of false positive rates, we know FPR can be rewritten as

$$\frac{B}{B+D}$$

This is equivalent to the simplified version of $(\frac{BR}{1-BR})(\frac{1-PPV}{PPV})(1 - FNR)$, as shown above. Thus, $FPR = (\frac{BR}{1-BR})(\frac{1-PPV}{PPV})(1 - FNR)$. QED.

Problem 5

The statement from Problem 4 holds for the entire population as well as for each group individually. Suppose that all three fairness notions are satisfied by our hypothesis, i.e. $FPR_1 = FPR_2$, $FNR_1 = FNR_2$, and $PPV_1 = PPV_2$. Further, assume that all of these values, as well as the base rates, are neither 0 nor 1. Show that this implies that the base rates of the groups must be equal.

From the given, we know that $FPR_1 = FPR_2$. From Problem 4, we know this can be rewritten as

$$FPR = (\frac{BR}{1-BR})(\frac{1-PPV}{PPV})(1-FNR) \tag{10}$$

From the given, we know that (10) holds for the entire population, as well as for each group individually. Thus, it holds that for groups 1 and 2,

$$(\frac{BR_1}{1-BR_1})(\frac{1-PPV_1}{PPV_1})(1-FNR_1) = (\frac{BR_2}{1-BR_2})(\frac{1-PPV_2}{PPV_2})(1-FNR_2) \tag{11}$$

From the given, we know that $FNR_1 = FNR_2$ and that $PPV_1 = PPV_2$. Let the variable s represent FNR_1 and FNR_2, since they are equal. Similarly, let the variable g represent PPV_1 and PPV_2 since they are equal. Thus, (11) can be rewritten as

$$(\frac{BR_1}{1-BR_1})(\frac{1-g}{g})(1-s) = (\frac{BR_2}{1-BR_2})(\frac{1-g}{g})(1-s) \tag{12}$$

Simplifying (12), (and keeping in mind that none of these values are 0),

$$\begin{aligned}
(\frac{BR_1}{1-BR_1})(\frac{1-g}{g})(1-s) &= (\frac{BR_2}{1-BR_2})(\frac{1-g}{g})(1-s) \\
(\frac{BR_1}{1-BR_1})(1-s) &= (\frac{BR_2}{1-BR_2})(1-s) \\
\frac{BR_1}{1-BR_1} &= \frac{BR_2}{1-BR_2} \\
BR_1(1-BR_2) &= BR_2(1-BR_1) \\
BR_1 - BR_1BR_2 &= BR_2 - BR_1BR_2 \\
BR_1 &= BR_2
\end{aligned}$$

Thus, the base rates of the groups must be equal, since $BR_1 = BR_2$, as shown above. QED.

Problem 6

Show that if our hypothesis makes no mistakes (i.e. $B_1 = B_2 = 0$ and $C_1 = C_2 = 0$), all three fairness notions will be satisfied, regardless of the base rates for each group.

We must show that $FPR_1 = FPR_2$, $FNR_1 = FNR_2$, and $PPV_1 = PPV_2$, regardless of base rates.

1. Proof that $FPR_1 = FPR_2$: From the definition of FPR, we know that FPR for groups 1 and 2 respectively can be rewritten as

$$\frac{B_1}{B_1 + D_1} \tag{13}$$

$$\frac{B_2}{B_2 + D_2} \tag{14}$$

From the given, we know that $B_1 = B_2 = 0$. Hence, (13) and (14) can be rewritten respectively as

$$\frac{0}{0 + D_1} \tag{15}$$

$$\frac{0}{0 + D_2} \tag{16}$$

However, both (15) and (16) simplify to 0. Thus, $FPR_1 = 0$ and $FPR_2 = 0$, and $FPR_1 = FPR_2$. QED.

2. Proof that $FNR_1 = FNR_2$: From the definition of FNR, we know that FNR for groups 1 and 2 respectively can be rewritten as

$$\frac{C_1}{A_1 + C_1} \tag{17}$$

$$\frac{C_2}{A_2 + C_2} \tag{18}$$

From the given, we know that $C_1 = C_2 = 0$. Hence, (17) and (18) can be rewritten respectively as

$$\frac{0}{A_1 + 0} \tag{19}$$

$$\frac{0}{A_2 + 0} \tag{20}$$

However, both (19) and (20) simplify to 0. Thus, $FNR_1 = 0$ and $FNR_2 = 0$, and $FNR_1 = FNR_2$. QED.

3. Proof that $PPV_1 = PPV_2$: From the definition of PPV, we know that PPV for groups 1 and 2 respectively can be rewritten as

$$\frac{A_1}{A_1 + B_1} \tag{21}$$

$$\frac{A_2}{A_2 + B_2} \tag{22}$$

From the given, we know that $B_1 = B_2 = 0$. Hence, (21) and (22) can be rewritten respectively as

$$\frac{A_1}{A_1 + 0} \tag{23}$$

$$\frac{A_2}{A_2 + 0} \tag{24}$$

However, both (23) and (24) simplify to 1. Thus, $PPV_1 = 1$ and $PPV_2 = 1$, and $PPV_1 = PPV_2$. QED.

From parts 1, 2, and 3, it is shown that all three fairness notions are satisfied, regardless of base rates. QED.