Name: Brandon Jaipersaud
UTORid: jaipers5
Student Number: 1002521130

CSC384 – A4 Q3

Variables:

A = sensitive attribute (i.e. gender) = {M, F}

C = classification (i.e Hyperlipidemia prediction) = {T, F}

Y = ground truth (i.e whether the patient actually has Hyperlipidemia) = {T, F}

Note that probabilities are rounded to 2 decimal places.

Case 1: Separation but no sufficiency

Joint Distribution Table for P(A, C, Y)

A	C	Y	P(A, C, Y)
M	T	T	0.20
M	T	F	0.25
M	F	T	0.10
M	F	F	0.10
F	T	T	0.10
F	T	F	0.08
F	F	T	0.05
F	F	F	0.12

Separation:

P(C=T | Y=T, A=M) = P(C=T | Y=T) = 0.67

No Sufficiency:

P(Y=T | C=T, A=M) = 0.44
P(Y=T | C=T) = 0.48

➔ P(Y=T | C=T, A=M) != P(Y=T | C=T)

Case 2: Sufficiency but no separation

Joint Distribution Table for P(A, C, Y)

A	C	Y	P(A,C,Y)
M	T	T	0.10
M	T	F	0.15
M	F	T	0.05
M	F	F	0.15
F	T	T	0.20
F	T	F	0.30
F	F	T	0.02
F	F	F	0.03

Sufficiency:

$P(Y=T \mid C=T, A=M) = P(Y=T \mid C=T) = 0.40$

No Separation

$P(C=T \mid Y=T, A=M) = 0.67$
$P(C=T \mid Y=T) = 0.81$

➔ $P(C=T \mid Y=T, A=M) \mathrel{!}= P(C=T \mid Y=T)$