NAME: Galen Byrd
DATE: 2/24/19
Homework: 04
STAT/CS 387

Problem 1

$\theta = E[C_1] - E[C_0] = 0.5 - 625 = -0.125$

$\alpha = E[Y \mid X = 1] - E[Y \mid X = 0] = 1 - .25 = .75$

In this example we see a positive association and a negative average causal effect. The intuition of my example is that the "treatment" X, has no effect on the outcome variable Y for almost all responders, but we assume we have an anti-responder. This means we observe a subject who goes untreated and survives, but we report that they would have died if they were treated.

X	**Y**	**C_0**	**C_1**
0	0	0	0*
0	0	0	0*
0	0	0	0*
0	1	1	0*
1	1	1*	1
1	1	1*	1
1	1	1*	1
1	1	1*	1

Problem 2

1)

Z=0	**Y=0**	**Y=1**	
X=0	0.81	0.09	P(X=0) = .9
X=1	0.09	0.01	P(X=1) = .1
	P(Y=0) = .9	P(Y=1) = .1	

Z=1	**Y=0**	**Y=1**	
X=0	0.25	0.25	P(X=0) = .5
X=1	0.25	0.25	P(X=1) = .5
	P(Y=0) = .5	P(Y=1) = .5	

2)

To show that X $\perp Y \mid Z$, we show that Pr(X=x & Y=y | Z=z) = Pr(X=x | Z=z) * Pr(Y=y | Z=z) is true for all cases.

Pr(X=1 & Y=1 | Z=0) = Pr(X=1 | Z=0) * Pr(Y=1 | Z=0), where 0.01 = 0.1 * 0.1

Pr(X=1 & Y=0 | Z=0) = Pr(X=1 | Z=0) * Pr(Y=0 | Z=0), where 0.09 = 0.9 * 0.1

Pr(X=0 & Y=1 | Z=0) = Pr(X=0 | Z=0) * Pr(Y=1 | Z=0), where 0.09 = 0.1 * 0.9.

Pr(X=0 & Y=0 | Z=0) = Pr(X=0 | Z=0) * Pr(Y=0 | Z=0), where 0.81 = 0.9 * 0.9.

Pr(X=0 & Y=0 | Z=1) = Pr(X=0 | Z=1) * Pr(Y=0 | Z=1), where 0.25 = 0.5 * 0.5.

Pr(X=1 & Y=0 | Z=1) = Pr(X=1 | Z=1) * Pr(Y=0 | Z=1), where 0.25 = 0.5 * 0.5.

Pr(X=0 & Y=1 | Z=1) = Pr(X=0 | Z=1) * Pr(Y=1 | Z=1), where 0.25 = 0.5 * 0.5.

Pr(X=1 & Y=1 | Z=1) = Pr(X=1 | Z=1) * Pr(Y=1 | Z=1), where 0.25 = 0.5 * 0.5.

3)

MARGINAL	**Y=0**	**Y=1**	
X=0	0.53	0.17	P(X=0) = .7
X=1	0.17	0.13	P(X=1) = .3
	P(Y=0) = .7	P(Y=1) = .3	

4)

To show that X and Y are not marginally independent we must show that $\Pr(X=x \mid Y=y) \neq \Pr(X)$ for at least one case. I continue with the case $\Pr(X=0 \mid Y=0) \neq \Pr(X=0)$, where $0.13/0.3 \neq 0.3$.

Problem 3

To show that $X \perp Z$ we must show that $f(X, Z) = f(X) * f(Z)$. We see that the probability function consistent with our DAG is $f(X, Y, Z) = f(X) * f(Z) * f(Y \mid X, Z)$. We proceed knowing $f(X, Z) = \sum_y f(X, Y, Z)$ and plug in the probability function consistent with our DAG.

$$= f(X) * f(Z) * \sum_y f(Y \mid X, Z)\ \}\text{Term sums to 1 as it is a sum over probability function}$$

$f(X, Z) = \Pr(X) * \Pr(Z)$

Next, we show $f(X \,\&\, Z \mid Y) \neq f(X \mid Y) * f(Z \mid Y)$ with a proof by contradiction.

$f(X, Y, Z) / f(Y) = f(X \mid Y) * f(Z \mid Y)$

$f(X) * f(Z) * f(Y \mid X, Z) / f(Y) = f(X \mid Y) * f(Z \mid Y)$

Because $X \perp Z$ we know $f(X, Z) = f(X) * f(Z)$,

$f(X, Z) * f(Y \mid X, Z) / f(Y) = f(X \mid Y) * f(Z \mid Y)$

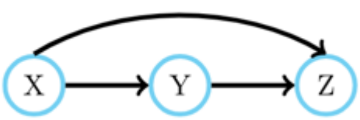
If X and Z were independent given Y we would see $f(X, Z) = f(X, Z \mid Y) = f(X \mid Y) * f(Z \mid Y)$, but we see

$f(X, Z \mid Y) * f(Y \mid X, Z) / f(Y) \neq f(X, Z \mid Y)$

Gives us a contradiction as the term on the right cannot equal itself divided by $f(Y)$ unless $f(Y) = f(Y \mid X, Z)$ which would imply that Y is independent of X and Z but it is not as it is a collider, therefore X and Z are dependent given Y. I understand a collider to be a variable that can be influenced by two independent variables. An example would be if waking up early (X) and aliens invading (Z) influence if it will be a sunny day (Y), where X and Z are independent.

Problem 4

1)



2) $Pr(Z=z \mid Y=y) = (Pr(Y=y) * Pr(Z = z)) / Pr(Y=y)$

$= Pr (Y=y \ \& \ Z=z) / Pr(Y=y)$

$= \sum_x Pr(X=x, Y=y, Z=z) / Pr(Y=y)$, then inputing our DAG,

$= \sum_x Pr(X=x)* Pr(Y=y \mid X=x)* Pr(Z=z \mid X=x, Y=y) / Pr(Y=y)$

$Pr(Z=1 \mid Y=1) = (Pr(X=1)* Pr(Y=1 \mid X=1)* Pr(Z=1 \mid X=1, Y=1)+ Pr(X=0)* Pr(Y=1 \mid X=0)* Pr(Z=1 \mid X=0 , Y=1)) / Pr(Y=1)$

$= (0.5*0.88*0.88+ 0.5*0.119*0.5)/ 0.5 = 0.83$

3)

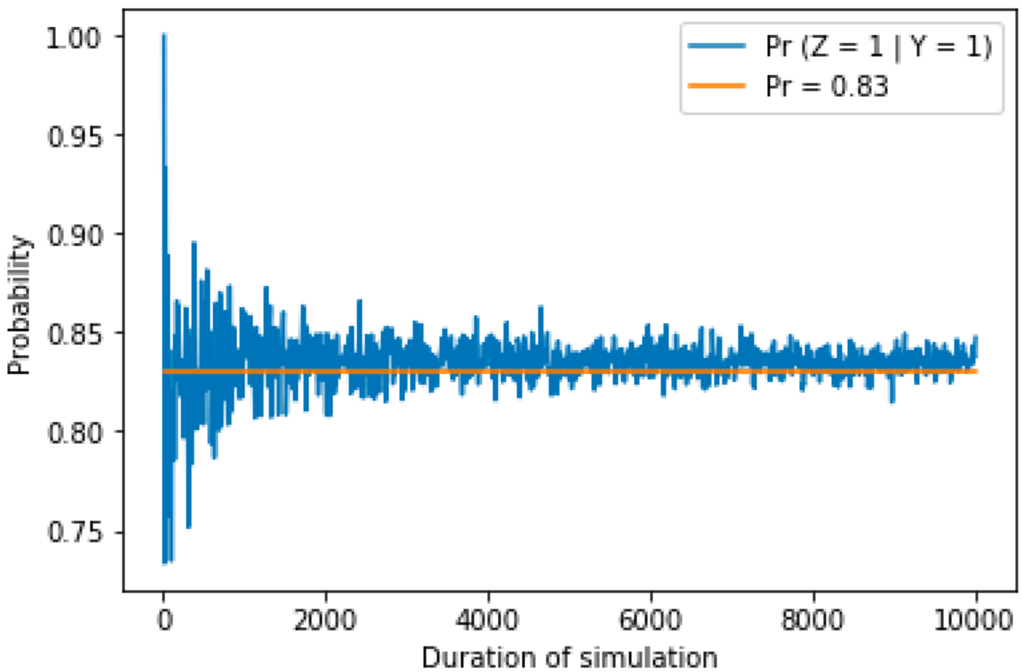


Fig 1: Pr(Z=1 | Y=1) for simulations of length up to 10,000

4) $Pr(Z=1 \mid Y:=y) = (Pr(Y:=y) * Pr(Z = 1)) / Pr(Y:=y)$

$= Pr (Y:=y \ \& \ Z=1) / Pr(Y:=y)$

$= \sum_x Pr(X=x, Y:=y, Z=1) / Pr(Y:=y)$, then inputing our DAG where Y was intervened,

$= \sum_x Pr(X=x) * Pr(Z=1 \mid X=x, Y:=y) / Pr(Y=y)$

$= \sum_x Pr(X=x) * Pr(Z=z \mid X=x, Y:=y)$

$Pr(Z=1 \mid Y:=1) = Pr(Z=1 \mid X=1, Y:=1) * Pr(X=1) + Pr(Z=1 \mid X=0, Y:=1) * Pr(X=0)$

$= 0.88*0.5 + 0.5*0.5 = 0.69$

5)

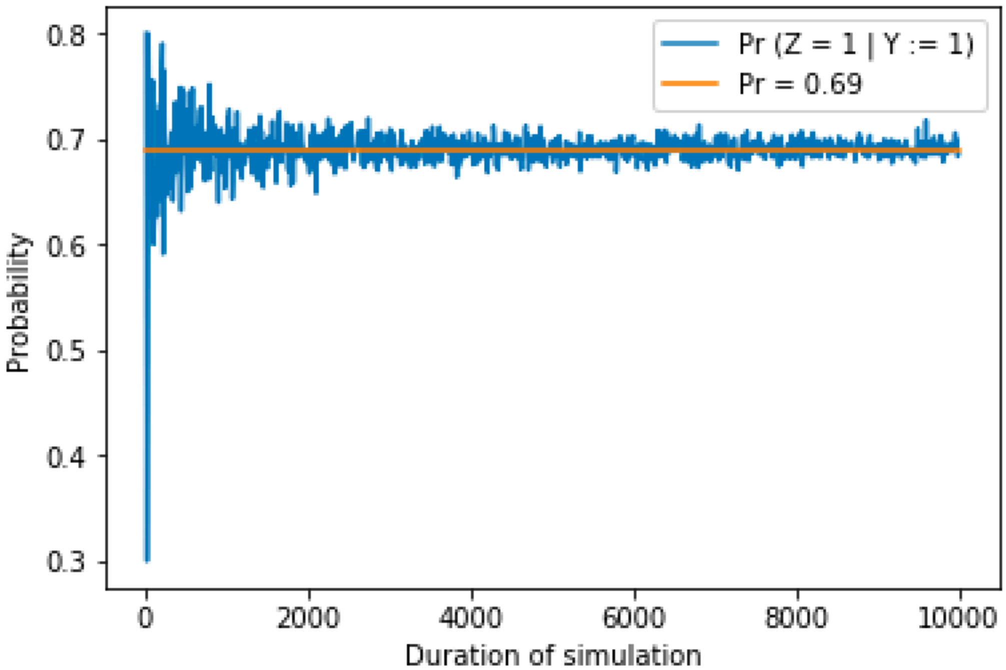


Fig 2: Pr(Z=1 | Y:=1) for simulations of length up to 10,000