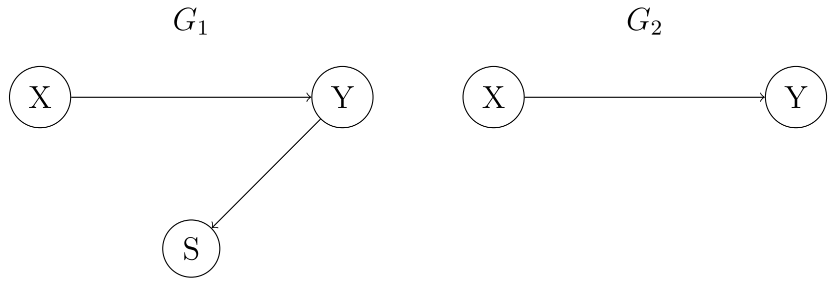
Lemma 1 Proof/Example

Canyon Foot

October 18, 2019

For binary random variables X, Y, S, we wish to construct two causal graphs, G_1, $G_2 = G_1 \backslash \{Y \to S\}$ with corresponding probability functions P_1, P_2 such that $P_2(y = 1|x = 1) = P_1(y = 1|x = 1, s = 1)$ but $P_2(y = 1|x = 1) \neq P_1(y = 1|x = 1)$. This demonstrates that the 'selection biased' conditional distribution $P_1(y = 1|x = 1, s = 1)$ is consistent with a conditional $P_2(y = 1|x = 1)$ that is different from the 'true' conditional $P_1(y = 1|x = 1)$. This is a violation of the definition of s-recoverability, therefore giving the result that G_1 is unrecoverable.



We begin by using the definition of conditional probability to expand $P_1(y = 1|x = 1, s = 1)$ as:

$$P_1(y = 1|x = 1, s = 1) = \frac{P_1(y = 1, x = 1, s = 1)}{P_1(x = 1, s = 1)}$$

Now, the definition of Markov factorization property for an SCM gives that $P_1(y = 1, x = 1, s = 1) = P_1(s = 1|y = 1)p(y = 1|x = 1)$. Then, expanding the denominator using the law of total probability, we have:

$$P_1(y = 1|x = 1, s = 1) = \frac{P_1(s = 1|y = 1)p(y = 1|x = 1)}{P_1(s = 1|y = 1)P(y = 0|x = 1) + P_1(s = 1|y = 0)P(y = x|yx = 1)}.$$

Given this formulation, we can assign explicit values to our probabilities. Let $P_1(y = 1|x = 1) = \frac{1}{2}$, therefore giving $P_1(y = 0|x = 1) = \frac{1}{2}$. Additionally, let $P_1(s = 1|y = 1) = \frac{3}{4}$ and $P_1(s = 1|y = 0) = \frac{1}{4}$. This means that we are more likely to 'sample' pairs (x, y) when $y = 1$ than when $y = 0$. Then

$$P_1(y = 1|x = 1, s = 1) = \frac{\frac{3}{4}\frac{1}{2}}{\frac{3}{4}\frac{1}{2} + \frac{1}{4}\frac{1}{2}} = \frac{3}{4}.$$

By our earlier condition, we make $P_2(y = 1|x = 1) = P_1(y = 1|x = 1, s = 1) = \frac{3}{4}$, and here we get the result since $P_2(y = 1|x = 1) \neq P_1(y = 1|x = 1) = \frac{1}{2}$.