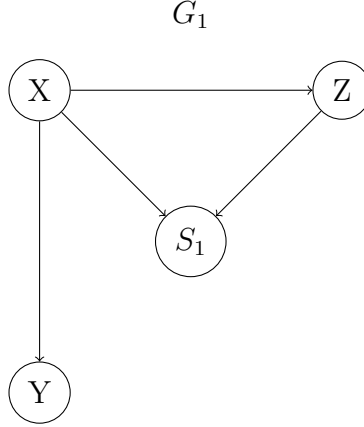


Recovering Joint from two biased distributions

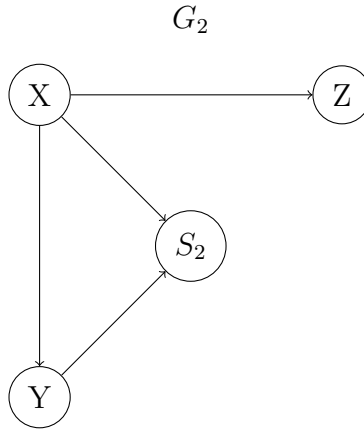
Canyon Foot

November 26, 2019

Here I give a simple example of two causal graphs measuring the same variables but with different selection mechanisms. In this case, the full joint distribution is recoverable only if we have access to the distributions for both graphs. We assume external information $P(x)$.



As always, we have access to $P(x, y, z|S_1 = 1)$, and since $Y \perp\!\!\!\perp S|X$, we also have $P(y|x)$. Additionally, since we have assumed access to $P(x)$, we could obtain $P(x, y)$ if we wanted. Notice the because of their is an edge for Z to S , we cannot obtain $P(z)$.



The story here is similar. We can get $P(x, y, z|S_2 = 1)$, $P(z|x)$, and $P(x, z)$ (using the external data). Then, using the chain rule for probability,

$$\begin{aligned} P(X, Y, Z) &= P(Y|X, Z)P(X, Z) \\ &= P(Y|X)P(Z|X)P(X) \end{aligned}$$

Since we have all of these quantities we can find $P(X, Y, Z)$ without issue. However, to do so we need to use $P(Z|X)$ and $P(Y|X)$, which means we need both biased distributions.