Inverse Probability Weighting

In backdoor criterion part, we learned that PA(X) always satisfies backdoor criterion. Hence, we can represent postintervention causal effect as such:

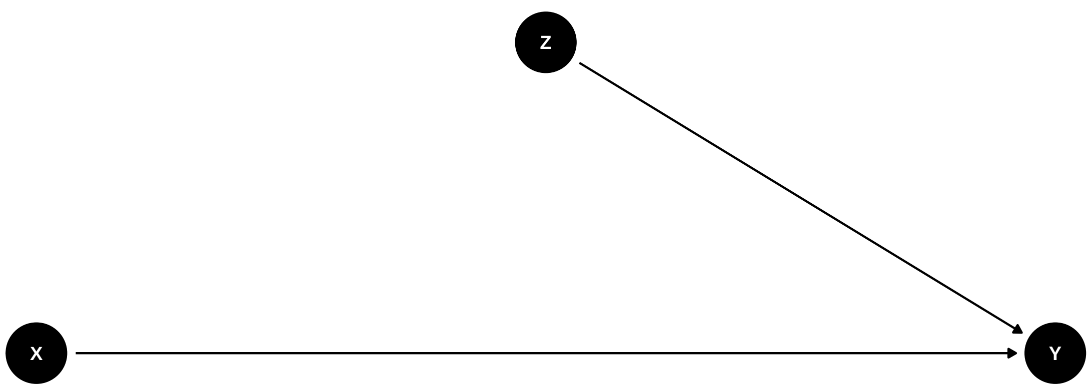
$P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, PA = z)P(PA = z)$

Modifying this equation, we can obtain the following:

$P(Y = y \mid do(X = x)) = \sum_z \frac{P(Y=y, X=x, PA=z)}{P(X=x|PA=z)}$

The factor $P(X = x \mid PA = z)$ in the denominator is known as the "propensity score."

Now, assume we have a graphical model with an intervention, do(X=x), on the model.



$P(Y = y \mid do(X = x))$

$= \ P_m(Y = y \mid X = x))$ (by definition)

$= \ \sum_z P_m(Y = y \mid X = x, Z = z)P(Z = z \mid X = x)$ (Bayes' rule)

$= \ \sum_z P_m(Y = y \mid X = x, Z = z)P(Z = z)$ $(X \perp\!\!\!\perp Z)$

$= \ \sum_z P(Y = y \mid X = x, Z = z)P(Z = z)$ $(X \perp\!\!\!\perp Z)$ (invariance relations)

$= \ \sum_z \frac{P(Y=y|X=x,Z=z)P(X=x|Z=z)P(Z=z)}{P(X=x|Z=z)}$

$= \ \sum_z \frac{P(Y=y,X=x,Z=z)}{P(X=x|Z=z)}$

From this result, we can see that postintervention causal effect can be computed by multiplying the pre-treatment distribution of (X,Y,Z) by a factor $1/P(X = x \mid Z = z)$, propensity score. Namely, each case $(Y = y, X = x, Z = z)$in the population should boost its probability by the inverse conditional probability of assignment to a treatment condition given a set of observed covariates. This is the reason why this method is called "inverse probability weighting."

Reference

Pearl, J., Glymour, M., & Jewell, N. P. (2019). Causal inference in statistics a primer. Wiley.

Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.