Average causal effect

$(Z_i, Y_i(0), Y_i(1))$ are iid

$$\tau = E[Y_i(1) - Y_i(0)]$$
$$= E[Y_i(1)] - E[Y_i(0)]$$

randomization comes into picture

$$Z_i \perp\!\!\!\perp (Y_i(0), Y_i(1))$$

$$\tau = E(Y_i(1)) - E(Y_i(0))$$
$$= E(Y_i(1) \mid Z_i = 1) - E(Y_i(0) \mid Z_i = 0)$$
$$= E(Y_i \mid Z_i = 1) - E(Y_i \mid Z_i = 0)$$

$$\Rightarrow \hat{\tau} = \frac{1}{n_1}\sum_{i=1}^{n} Z_i Y_i - \frac{1}{n_0}\sum_{i=1}^{n}(1 - Z_i) Y_i$$

Thm $E(\hat{\tau}) = \tau$ unbiased estimator of ACE

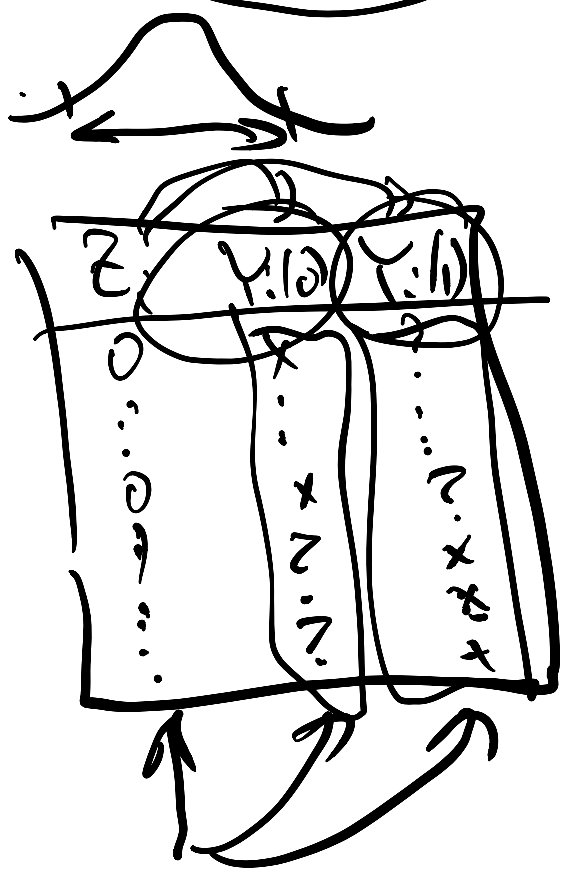
$$\mathrm{Var}(\hat{\tau}) = \frac{\mathrm{Var}(Y_i(1))}{n_1} + \frac{\mathrm{Var}(Y_i(0))}{n_0}$$

$$\widehat{\mathrm{Var}}(\hat{\tau}) = \frac{s_1^2}{n_1} + \frac{s_0^2}{2}$$

$$s_1^2 = \frac{1}{n_1 - 1}\sum Z_i (Y_i - \bar{Y}_1)^2$$

$$s_0^2 = \frac{1}{n_0 - 1} \sum_i (1 - Z_i)(Y_i - \bar{Y}_0)^2$$

$$\hat{\tau} \pm 1.96 \sqrt{\widehat{Var}(\hat{\tau})} \quad \text{(CLT)}$$



Propensity score

$$e(X_i) \triangleq \mathbb{P}(Z_i = 1 \mid X_i)$$

Thm $Z_i \perp\!\!\!\perp X_i \mid e(X_i)$

Pf $P(Z_i = 1 \mid \underline{X_i}, \underline{e(X_i)})$
$$= P(Z_i = 1 \mid e(X_i))$$

$$P(Z_i = 1 \mid X_i, e(X_i))$$
$$= P(Z_i = 1 \mid X_i)$$
$$= e(X_i)$$

$Z_i \in \{0, 1\}$

$$P(Z_i = 1 \mid e(X_i))$$
$$= E(Z_i \mid e(X_i))$$
$$= E(\underline{E(Z_i \mid X_i, e(X_i))} \mid e(X_i))$$
$$= E(\underline{E(Z_i \mid X_i)} \mid e(X_i))$$
$$= E(e(X_i) \mid e(X_i))$$
$$= e(X_i)$$

$E(Y|Z)$
$= E(E(Y|X,Z)|Z)$

$E(Y)$
$= E(E(Y|X))$

Thm Under unconfoundedness,

$$Z_i \perp\!\!\!\perp Y_i(0), Y_i(1) \mid \boxed{e(X_i)}$$

$e(X_i) = P(Z_i = 1 \mid X_i)$

often estimated from (X_i, Z_i) pairs

$\hat{e}(X_i)$ = logistic regression

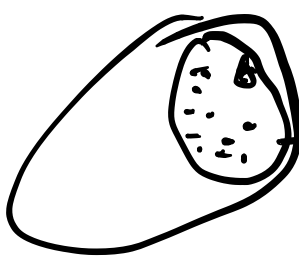
X_i $\quad e(X_i) = P(Z_i = 1 \mid X_i)$

$= .7$ say

for $X_i = x$

$= .7$

for $X_i = x'$



$e(X_i) = .7$

$X_i = x$ $\quad e(x) = .7$

$X_i = x'$ $\quad e(x') = .7$

.7 of the $Z_i = 1$

.3 of the $Z_i = 0$

for both x and x'

- Inverse propensity score weighted

$$\tau = E(Y_i(1)) - E(Y_i(0))$$

Thm $E(Y_i(1))$

$$\approx \hat{E}\left(\frac{Y_i Z_i}{\hat{e}(X_i)}\right)$$

$$\left(E(Y_i(0)) = E\left(\frac{Y_i(1-Z_i)}{1-e(X_i)}\right) \right)$$

Propensity score

$$e(X_i) = P(Z_i = 1 \mid X_i)$$

Thm $\underline{Z_i} \perp\!\!\!\perp \underline{X_i} \mid \underline{e(X_i)}$

build a model $\hat{e}(X_i)$ using (say) logistic and $(\underline{X_i}, \underline{Z_i})$

Tower property

$$E(X) = E(E(X|Y))$$

$$E(X|Z) = E(E(X|Y, Z)|Z)$$

- Inverse propensity score weighting

Thm $E(Y_i(1)) = E\left(\frac{Y_i Z_i}{e(X_i)}\right)$

$\left(E(Y_i(0)) = E\left(\frac{Y_i(1-Z_i)}{1-e(X_i)}\right) \right)$

empirical average

$\Rightarrow$ estimator:

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i Z_i}{\hat{e}(X_i)} - \frac{1}{n}\sum_{i=1}^{n} \frac{(1-Z_i) Y_i}{1-\hat{e}(X_i)}$$

(estimates $\tau = E(Y_i(1) - Y_i(0))$
$= E(Y_i(1)) - E(Y_i(0))$

$$E\left(\frac{Y_i Z_i}{e(X_i)}\right) = E(Y_i(1))$$

Pf

$$E\left(\frac{Y_i Z_i}{e(X_i)}\right) \quad Y_i = \begin{cases} Y_i(1) & Z_i = 1 \\ Y_i(0) & Z_i = 0 \end{cases}$$

$$= E\left(E\left(\frac{Y_i Z_i}{e(X_i)} \,\middle|\, X_i\right)\right)$$

$$= E\left(E\left(\frac{Y_i(1) Z_i}{e(X_i)} \,\middle|\, X_i\right)\right)$$

$$= E\left(\frac{E(Y_i(1) \mid X_i)\, E(Z_i \mid X_i)}{e(X_i)}\right)$$

$$= E\left(\frac{E(Y_i(1) \mid X_i)\, P(Z_i = 1 \mid X_i)}{e(X_i)}\right)$$

$$= E\left(\frac{E(Y_i(1) \mid X_i)\, \cancel{P(Z_i = 1 \mid X_i)}}{\cancel{P(Z_i = 1 \mid X_i)}}\right)$$

$$= E\left(E(Y_i(1) \mid X_i)\right)$$

$$= E(Y_i(1))$$

$$E z = 0 \cdot \cancel{P(z=0)}$$

$$1 \cdot P(z=1)$$

$$= P(z=1)$$

$$E\left(\frac{Y_i z_i}{e(X_i)}\right)$$

$$E(aX) = aEX$$

$$= E\left(E\left(\frac{Y_i(1) z_i}{e(X_i)} \,\middle|\, X_i\right)\right)$$

$$= E\left(\boxed{\frac{1}{e(X_i)}} \boxed{E(Y_i(1) z_i \mid X_i)}\right)$$

$$= E\left(\frac{1}{e(X_i)} E(Y_i \mid X_i) \cdot E(z_i \mid X_i)\right)$$

Matching

X_i, Z_i, $Y_i(0)$, $Y_i(1)$

consider $i \rightarrow Z_i = 1$ (say)

find $i' \rightarrow Z_{i'} = 0$

s.t. $\boxed{X_i \approx X_{i'}}$

$$\hat{\tau}_i = \underline{Y_i} - \underline{Y_{i'}}$$

$$\hat{\tau} = \frac{1}{n} \sum_i (Y_i - Y_{i'})$$

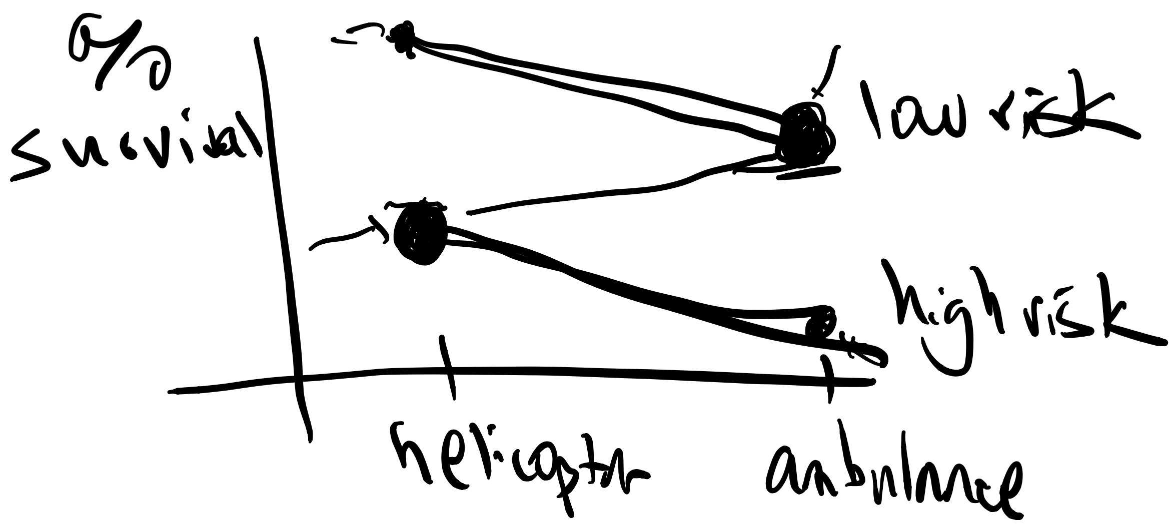
Simpson's paradox

arises as one compares conditionals and marginals

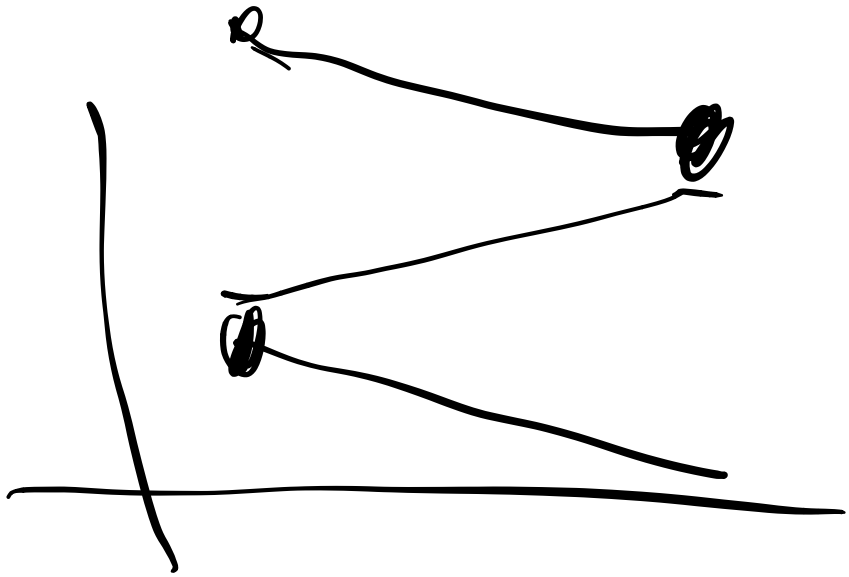
- paradox: a trend can occur at all levels of a certain variable, but when combines these levels, the trend can reverse

example

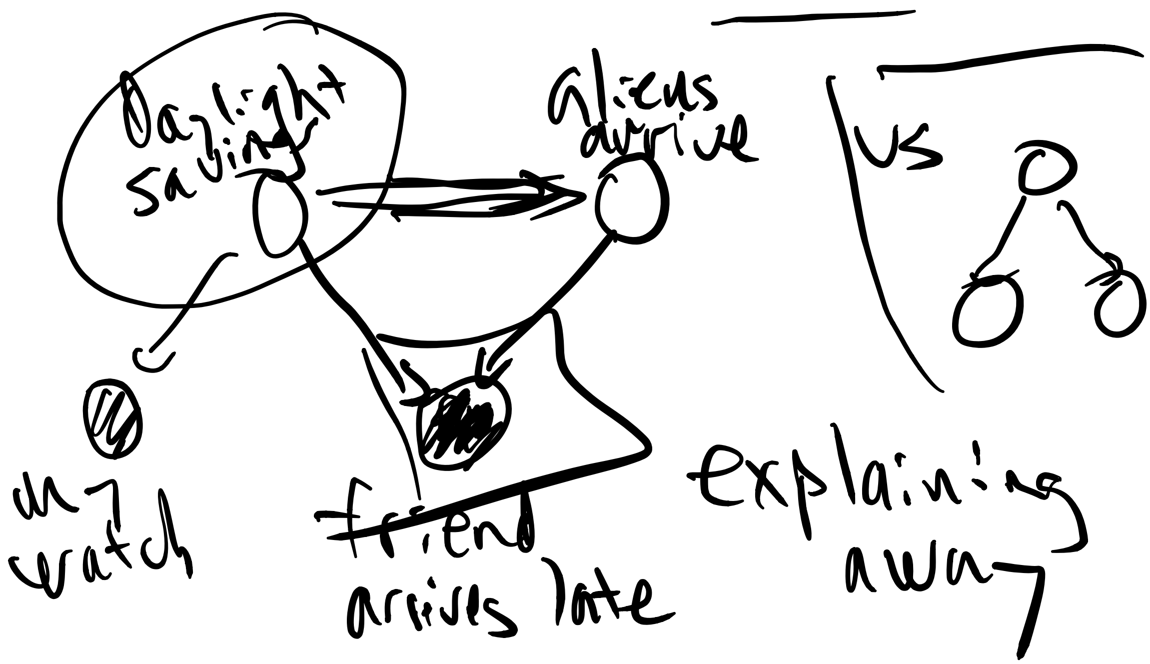
medical evacuation helicopters are worse at saving lives than traditional ambulances



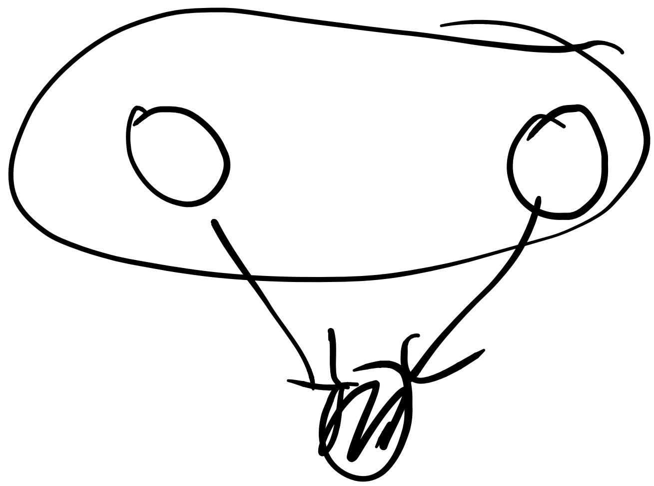
H/A depends on L/H



Selection bias



(induced dependency)



Imbens & Rubin

Causal Inference

Pearl

Causality

X & Pearl

The Book of Why