

- a way to identify causal effects
- conditioning + marginalizing = standardization
- to get from  $E(Y^a|X=x)$  to  $E(Y^a)$  we need to marginalize (average over) the dist. of  $X$ : means: avg./sum over all levels of  $X$

$$E_x(Y^a) = \sum_x E(Y^a|X=x) P(X=x) \quad (\text{for } X \text{ categorical})$$

(1) stratify      (2) averaging

$$= \sum_x E(Y|A=a, X=x) P(X=x) \quad \left. \begin{array}{l} \text{subpopulation &} \\ \text{observable} \end{array} \right\} \text{known as "standardization"} \quad \left. \begin{array}{l} \text{marginal dist. of } X \end{array} \right.$$

**Standardized mean**  $\hat{Y} = \text{avg. potential outcome}$

**Standardization:**

- involves stratifying followed by averaging
  - Step 1: obtain treatment effect within each stratum (i.e. for each  $x \in X$ )
  - Step 2: pool obtained treatment effect across stratum, weighting by the probability of each stratum

↳ Illustration of applying standardization method:

$A = \{\text{saxagliptin, sitagliptin}\}$  for diabetes patients

$Y = \text{MACE}$  (major adverse cardiac event)

Challenge:

- users of saxagliptin more likely to have used other oral antidiabetic (OAD) drug in the past
- patients with past use of OAD drugs are at higher risk for MACE

MAIN IDEA: compute MACE-rate for  $A=1$  and  $A=0$  users for two subpopulations:

(1) patients w/o prior OAD use (subpop. 1)  $\left. \begin{array}{c} x_1 \\ Y \end{array} \right\} X$  (in practice  $X$  would consist of more than 1 variable)  
 (2) patients w/ prior OAD use (subpop. 2)  $\left. \begin{array}{c} x_2 \\ Y \end{array} \right\} X$

⇒ average MACE-rate by share of population for each subpopulation

⇒ is CE if within levels of prior OAD use variable ( $X$ ), treatment can be thought of as randomized ⇒ ignorability given prior OAD use.

↳ Example with data:

		$Y$		Total
		$MACE = y_{ls} = 1$	$MACE = y_o = 0$	
$A$	$Saxa = y_{ls} = 1$	350	3650	4000
	$Saxa = y_o = 0$	500	6500	7000
Total		850	10150	11000

$$- P(Y=1 | A=1) = \frac{350}{4000} = 0.088 = 8.8\%$$

$$- P(Y=1 | A=0) = \frac{500}{7000} = 0.071 = 7.1\%$$

NOW STRATIFY BY  $X$  (prior OAD use) variable:

in each subpopulation, the risk of MACE is the same regardless treatment A

$$\begin{aligned} X = \text{no OAD use} = 0 & \quad P(Y|A=1) = \frac{50}{11000} \\ P(X=0) = \frac{5000}{11000} & \quad P(Y|A=0) = \frac{200}{4000} \\ & \quad P(Y|A=1) = \frac{300}{3000} \quad X = \text{OAD use} = 1 \\ & \quad P(Y|A=0) = \frac{300}{3000} \quad P(X=1) = 1 - P(X=0) = \frac{6000}{11000} \quad (\Rightarrow P(X=0)) \end{aligned}$$

		MACE		Total
		$y=1$	$y=0$	
$A$	$A=1$	50	950	1000
	$A=0$	200	3800	4000
Total		250	4750	5000

		MACE		Total
		$y=1$	$y=0$	
$A$	$A=1$	300	2700	3000
	$A=0$	300	2700	3000
Total		600	5400	6000

$$E(Y^1) = P(X=0) \cdot [E(Y|A=1, X=0)] + P(X=1) \cdot [E(Y|A=0, X=1)] \Leftrightarrow$$

based on  $P(Y|A=1)$  in subpopulation  $X=0$ .

based on  $P(Y|A=0)$  in subpopulation  $X=1$ .

$$\Leftrightarrow \frac{5}{11} \cdot \frac{50}{11000} + \frac{6}{11} \cdot \frac{300}{3000} = \frac{5}{11} \cdot \frac{5}{100} + \frac{6}{11} \cdot \frac{3}{30} = \frac{25}{1100} + \frac{18}{330} = 0.0772 \approx 7.72\%$$

⇒ result:  $E(Y^1) = E(Y^0)$  therefore  $ATE = E(Y^1) - E(Y^0) = 0$

Problem with standardization:

- in practice  $X$  is high-dimensional → stratifying would lead to many not only 1 variable empty cells/subpopulations w/o units (as in example above with  $X$  being OAD usage) ↳ e.g. stratification with age, blood pressure

↳ better/scalable alternatives: matching, inverse probability of treatment weighting propensity score methods } observational studies

instrumental variable methods } natural experiments