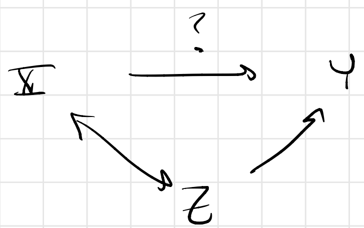
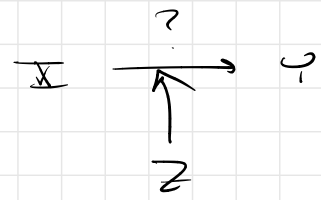
3/29/22

Confounders (Z)



Effect modifier (Z)



(interaction)

Hosmer-Lemeshow Test: (Sec. 5.2.3)

Goodness-of-fit test where we have ungrouped binary (0/1) observations + continuous predictor.

$$\left[\begin{array}{l} H_0: \text{our model} \\ H_a: \text{perfect model} \end{array}\right]$$

① Fit the model (logistic regression)

$$\log\left(\frac{\hat{\pi}_i}{1-\hat{\pi}_i}\right) = \beta_0 + \beta_1 x_i \qquad i = 1, 2, 3, \ldots, n$$

② Generate fitted probabilities $\hat{\pi}_i$

③ Order $\hat{\pi}_i$ from smallest to largest.

④ Group the data by quantiles of the fitted probabilities.
→ # of groups should allow ≥ 5 expected events in each group.

⑤ For each group, compute the observed & expected # of successes:

expected # = Sum of fitted probs.

(b) Compare observed counts to expected via a chi-squared test stat:

$$\mathbb{X}^2_{HL} = \sum_{j=1}^{g} \frac{(obs_j - exp_j)^2}{exp_j\left(1 - \frac{exp_j}{n_j}\right)} \overset{H_0}{\dot\sim} \chi^2_{(g-2)}$$

n_j = # of obs. in group j

g = # of groups

Model in matrix terms:

$$g(\underset{n\times 1}{\underline{\Psi}}) = \underset{n\times p}{\mathbb{X}}\ \underset{p\times 1}{\underline{\beta}}$$