ISLE 因果推断学习小组第七讲

Causal Inference

Chapter 12&13 Part2

What if the exposure is a continuous variable?

当暴露是连续型变量时，如何进行因果推断

lyx 2022.8.19

Review: Real Workflow of Causal Inference

因果待估量（causal estimand）

$$\mathbb{E}(Y^a)$$

↑ **识别**

统计待估量（statistical estimand）

$$= \sum_l \mathbb{E}(Y|A=a, L=l)\,\mathbb{P}(L=l)$$

（标准化方法）

$$= \mathbb{E}\left[\frac{I(A=a)}{\mathbb{P}(A|L)}Y\right]$$

（逆概率加权方法）

↑ **估计**

估计量（estimator）

$$\sum_l \widehat{\mathbb{E}}(Y|A=a, L=l)\,\widehat{\mathbb{P}}(L=l)$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)}{\widehat{\mathbb{P}}(a|l_i)}y_i$$

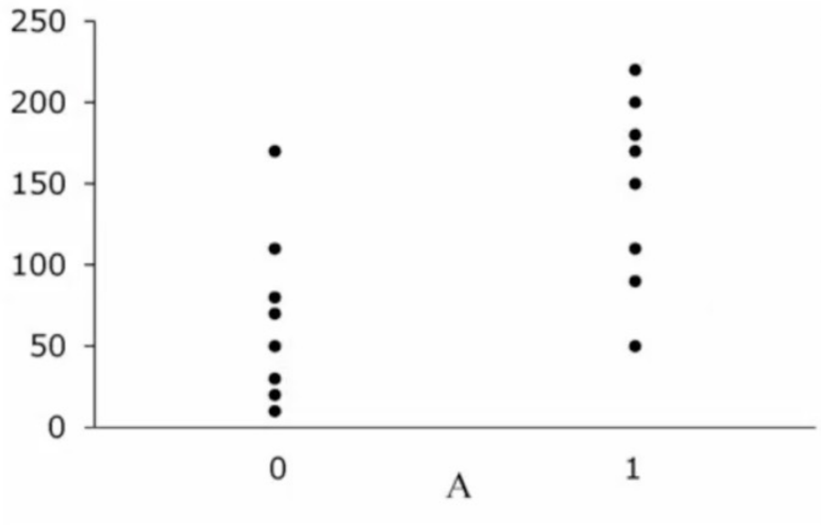
⚠️**正确因果推断的五大条件：**

- -3个识别条件：可交换性、一致性、正性
- -无测量误差
- -模型设置正确

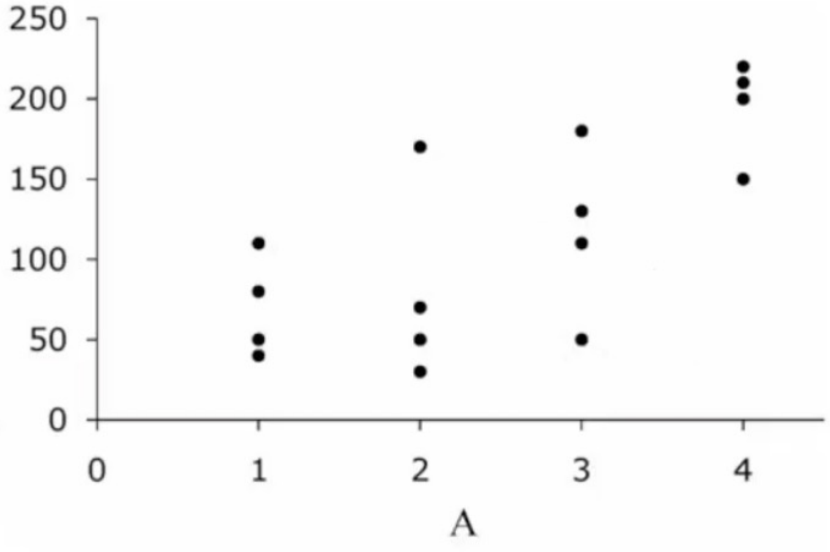
确保实验满足这些条件是最重要的事情，满足这些条件后，计算is trivial（作者原话，凡尔赛发言）

Review: Why model?

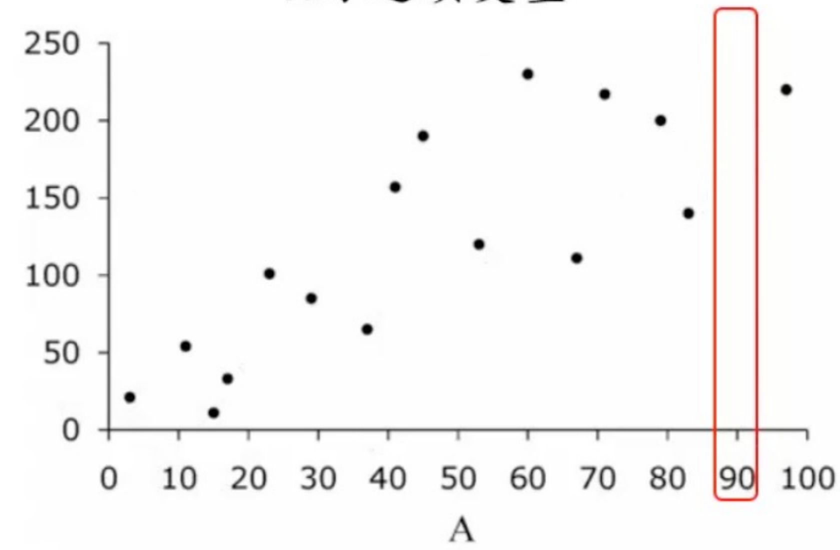
A为二分类

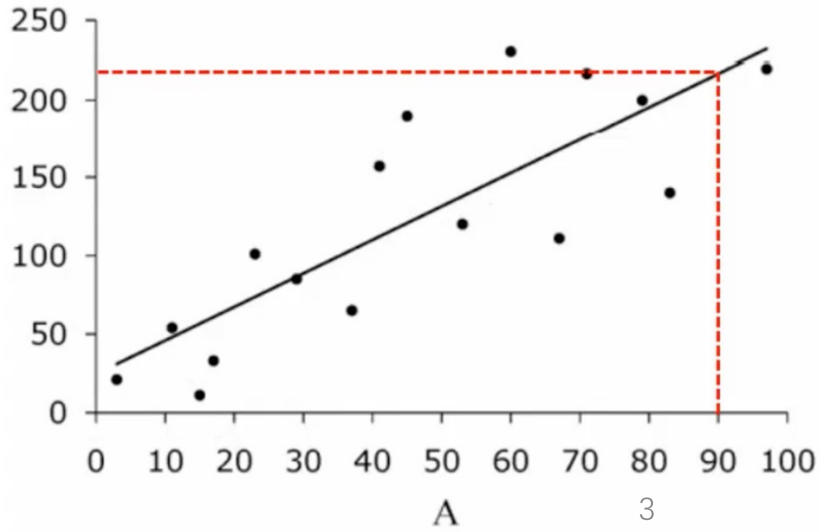


A为多分类



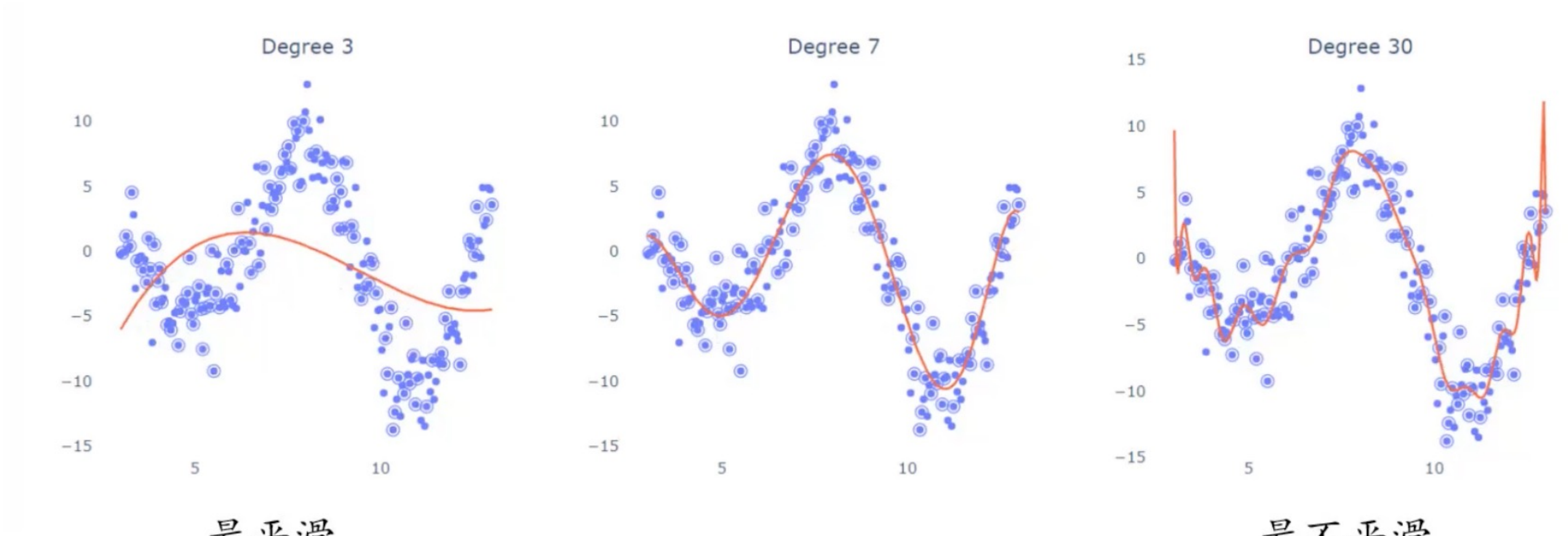
A为连续变量





模型：对数据联合分布的先验限制

复习：模型参数个数与平滑程度的关系



最平滑　　　　最不平滑

- Bias-Variance trade off

更多参数=更不平滑=更小的偏差=更大的方差（更宽的置信区间）

课堂学习思路

	暴露为二分类 混杂为二分类	→	暴露为二分类 混杂为连续型	→	暴露为连续型 混杂为二分类/连续型
➢ 数据	暴露为二分类 混杂为二分类	→	暴露为二分类 混杂为连续型	→	暴露为连续型 混杂为二分类/连续型
➢ 方法	非参数 IPW/standardization	→	参数化 IPW/standardization	→	IPW/standardization + Marginal structural models
	本书第 I 部分所讲		上周学习		本周学习

Outline

1. 边际结构模型
2. 连续型暴露、二分类混杂情形
3. 连续型暴露、连续型混杂情形
4. 在边际结构模型中考虑效应修饰

1. 边际结构模型（Marginal structural model)

$$E(Y^a) = \beta_0 + \beta_1 \mathrm{a}$$

- 对潜在结局Y^a建模（而不是观察结局）
- 暴露为二分类变量时，该模型为饱和模型

  待估量为$E(Y^{a=1})$、 $E(Y^{a=0})$

  模型参数为β_0、β_1
- 在这个例子中，β_1对应平均因果效应
- 当结局是二分类变量时，用边际结构logistic模型

连续型暴露、无混杂情形

? 当a是连续型变量，怎么估计因果效应$E(Y^a)-E(Y^{a'})$？ a和a’取任意值

个体编号	暴露（A）	观察结局（Y）	潜在结局										
			$Y_i(0)$	$Y_i(1)$	$Y_i(2)$	$Y_i(3)$	$Y_i(4.2)$	$Y_i(5)$	$Y_i(6.8)$	$Y_i(7)$	$Y_i(8)$	$Y_i(9)$	$Y_i(10)$
1	0	-10	-10	3		15		23		47	35		32
2	1	3	-10	3	?	15	?	23	?	47	35	?	32
3	3	7	-10	3		7		23		47	35		32
4	3	23	-10	3		23		23		47	35		32
5	5	23	-10	3		15		23		47	35		32
10	10	42	-10	3		15							

▶ **面临挑战**：并非所有暴露剂量都有数据。对于没有数据的暴露剂量，无法得知其对应潜在结局。

上模型！用已有数据建模，例如 $E(Y^a)=\beta_0+\beta_1 a$
（也可以是其他形式，以此为例，下同）

非饱和边际结构模型

$$E(Y^a) = \beta_0 + \beta_1 \mathrm{a}$$

- 暴露为连续型变量时，该模型为非饱和模型

  待估量为$E(Y^{a1})$、 $E(Y^{a2})\cdots$

  模型参数为β_0、β_1
- 暴露剂量$a_1 a_2$之间的因果效应变化值为β_1（ $a_1 - a_2$)

Outline

1. 边际结构模型
2. 连续型暴露、二分类混杂情形
3. 连续型暴露、连续型混杂情形
4. 在边际结构模型中考虑效应修饰

2.1 连续型暴露、二分类混杂情形——标准化法

❓有混杂变量L，怎么估计因果效应$E(Y^a)-E(Y^{a'})$？ a和a'取任意值

性别（L）	个体编号	暴露（A）	观察结局（Y）	潜在结局				
				Y^0	Y^1	$Y^{3.5}$	Y^5	Y^8
0	1	0	3	5.06	7.33	13.00	16.40	23.21
	2	1	9	5.06	7.33	13.00	16.40	23.21
	3	3.5	13	5.06	7.33	13.00	16.40	23.21
	4	5	18	5.06	7.33	13.00	16.40	23.21
	5	8	22	5.06	7.33	13.00	16.40	23.21
1	6	0	2	3.39	5.67	11.39	14.82	21.67
	7	1	5	3.39	5.67	11.39	14.82	21.67
	8	1	6	3.39	5.67	11.39	14.82	21.67
	9	3.5	12	3.39	5.67	11.39	14.82	21.67
	10	3.5	13	3.39	5.67	11.39	14.82	21.67
	11	5	16	3.39	5.67	11.39	14.82	21.67
	12	8	20	3.39	5.67	11.39	14.82	21.67
$\widehat{\mathbb{E}}(Y^a)$				4.09	6.36	12.06	15.48	22.31

解决方案：

①按L分层

②对不同L取值分别建立模型，填补潜在结局

$$\mathbb{E}(Y^a|L=0) = 5.06+2.27a$$

$$\mathbb{E}(Y^a|L=1) = 3.38+2.29a$$

③计算各暴露下的结局均值（标准化）

④对$E(Y^a)$建立边际结构模型

$$\widehat{\mathbb{E}}(Y^a) = 4.08+2.27a$$

2.2 连续型暴露、二分类混杂情形——IPW法

性别 L	个体编号	暴露 A	观察结局 Y	$\widehat{\mathbb{P}}(A=a\|L)$	$\frac{I(a_i=a)y_i}{\widehat{\mathbb{P}}(A=a\|L)}$
	1	0	3		
	2	1	9		
0	3	3.5	13		
	4	5	18		
	5	8	22		
	6	0	2		
	7	1	5		
	8	1	6		
1	9	3.5	12		
	10	3.5	13		
	11	5	16		
	12	8	20		

$\widehat{\mathbb{E}}(Y^a)$

- 感兴趣的因果待估量： $\tau = \mathbb{E}(Y^{a_2}) - \mathbb{E}(Y^{a_1})$
- IPW公式：$\mathbb{E}(Y^a) = \mathbb{E}\left[\frac{I(A=a)}{\mathbb{P}(A|L)}Y\right] = \frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)}{\mathbb{P}(a|l_i)}y_i$
- 识别并估计平均因果效应：

① 按照 L 分层计算 $\widehat{\mathbb{P}}(A=a|L)$

② 计算不同暴露取值对应的 $\frac{I(a_i=a)y_i}{\widehat{\mathbb{P}}(A=a|L)}$

③ 计算不同暴露取值对应的 $\widehat{\mathbb{E}}(Y^a) = \frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)}{\widehat{\mathbb{P}}(a|l_i)}y_i$

④ 对 $\mathbb{E}(Y^a)$ 建立边际结构模型 $\mathbb{E}(Y^a) = \beta_0 + \beta_1 a$

2.2 连续型暴露、二分类混杂情形——IPW法

①计算 $\widehat{\mathbb{P}}(A=a|L)$

②计算不同暴露取值对应的 $\dfrac{I(a_i=a)y_i}{\widehat{\mathbb{P}}(A=a|L)}$

性别 L	个体编号	暴露 A	观察结局 Y	$\hat{P}(A=0\|L)$	$\hat{P}(A=1\|L)$	$\hat{P}(A=3.5\|L)$	$\hat{P}(A=5\|L)$	$\hat{P}(A=8\|L)$	$\frac{I(a_i=0)y_i}{\widehat{\mathbb{P}}(0\|L)}$	$\frac{I(a_i=1)y_i}{\widehat{\mathbb{P}}(1\|L)}$	$\frac{I(a_i=3.5)y_i}{\widehat{\mathbb{P}}(3.5\|L)}$	$\frac{I(a_i=5)y_i}{\widehat{\mathbb{P}}(5\|L)}$	$\frac{I(a_i=8)y_i}{\widehat{\mathbb{P}}(8\|L)}$
	1	0	3						$\frac{3}{1/5}$	0	0	0	0
	2	1	9						0	$\frac{9}{1/5}$	0	0	0
0	3	3.5	13	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	0	0	$\frac{13}{1/5}$	0	0
	4	5	18						0	0	0	$\frac{18}{1/5}$	0
	5	8	22						0	0	0	0	$\frac{22}{1/5}$
	6	0	2						$\frac{2}{1/7}$	0	0	0	0
	7	1	5						0	$\frac{5}{2/7}$	0	0	0
	8	1	6										
1	9	3.5	12	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{1}{7}$					
	10	3.5	13										
	11	5	16										
	12	8	20										

④对 $\mathbb{E}(Y^a)$ 建立边际结构模型

$$\widehat{\mathbb{E}}(Y^a)=1.61+2.48a$$

③计算不同暴露取值对应的 $\frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)}{\widehat{\mathbb{P}}(a|l_i)}y_i$ →

$\widehat{\mathbb{E}}(Y^a)$	2.42	6.96	12.71	16.81	20.83

2.42= $\frac{1}{12}$ (3×5 + 2×7)

Outline

1. 边际结构模型
2. 连续型暴露、二分类混杂情形
3. 连续型暴露、连续型混杂情形
4. 在边际结构模型中考虑效应修饰

3.1 连续型暴露、连续型混杂——IPW方法

年龄 L	个体编号	暴露 A	观察结局 Y	$\hat{f}(A=a\|L)$	$\frac{I(a_i=a)y_i}{\hat{f}(a\|L)}$
36	1	0	3		
42	2	0	-10		
	3	1	47		
47	4	0	7		
	5	2.2	4		
	6	1	10		
48	7	0	23		
	8	3	8		
51	9	0	0		
56	10	1	10		

$\widehat{\mathbb{E}}(Y^a)$

- 感兴趣的因果待估量： $\tau = \mathbb{E}(Y^{a_2}) - \mathbb{E}(Y^{a_1})$
- IPW公式： $\mathbb{E}(Y^a) = \mathbb{E}\left[\frac{I(A=a)}{\mathbb{P}(A|L)}Y\right] = \frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)}{\mathbb{P}(a|l_i)}y_i$

当混杂是二分类时，这些值的全部组合都有现成数据，该组合下的概率可直接计算。

① 根据观测到的A和L数据构建暴露模型并估计$\hat{f}(A=a|L)$

② 计算不同暴露取值对应的 $\frac{I(a_i=a)y_i}{\hat{f}(a|L)}$

③ 计算不同暴露取值对应的 $\widehat{\mathbb{E}}(Y^a) = \frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)y_i}{\hat{f}(a|L)}$

④ 对 $\mathbb{E}(Y^a)$ 建立边际结构模型 $\mathbb{E}(Y^a) = \beta_0 + \beta_1 a$

3.1 连续型暴露、连续型混杂——IPW方法

① 根据观测到的A和L数据构建暴露模型并估计$\hat{f}(A=a|L)$

①估计 $\hat{f}(A=a|L)$

年龄 L	个体编号	暴露 A	观察结局 Y	$\hat{f}(A=0\|L)$	$\hat{f}(A=1\|L)$	$\hat{f}(A=2.2\|L)$	$\hat{f}(A=3\|L)$
36	1	0	3	0.346	0.306	0.088	0.020
42	2	0	-10	0.310	0.344	0.129	0.034
	3	1	47				
47	4	0	7	0.270	0.361	0.170	0.053
	5	2.2	4				
	6	1	10				
48	7	0	23	0.262	0.363	0.178	0.057
	8	3	8				
51	9	0	0	0.235	0.366	0.205	0.072
56	10	1	10	0.190	0.355	0.250	0.101

f(A|L) 是条件概率密度函数
(conditional probability density function, PDF)

⚠Unfortunately, pdfs are generally hard to estimate correctly, which is why using IP weighting for continuous treatments will often be **dangerous.**

One should be careful when using IP weighting for continuous treatments because the effect estimates may be exquisitely **sensitive to the choice of the model for the conditional density** f(A|L)

3.1 连续型暴露、连续型混杂——IPW方法

①估计 $\hat{f}(A=a|L)$

②计算不同暴露取值对应的 $\frac{I(a_i=a)y_i}{\hat{f}(a|L)}$

年龄 L	个体编号	暴露 A	观察结局 Y	$\hat{f}(A=0\|L)$	$\hat{f}(A=1\|L)$	$\hat{f}(A=2.2\|L)$	$\hat{f}(A=3\|L)$	$\frac{I(a_i=0)y_i}{\hat{f}(0\|L)}$	$\frac{I(a_i=1)y_i}{\hat{f}(1\|L)}$	$\frac{I(a_i=2.2)y_i}{\hat{f}(2.2\|L)}$	$\frac{I(a_i=3)y_i}{\hat{f}(3\|L)}$
36	1	0	3	**0.346**	**0.306**	**0.088**	**0.020**	3/0.346	0	0	0
42	2	0	-10	0.310	0.344	**0.129**	**0.034**	-10/0.310	0	0	0
	3	1	47					0	47/0.344	0	0
47	4	0	7	0.270	0.361	0.170	**0.053**	7/0.270	0	0	0
	5	2.2	4					0	0	4/0.170	0
	6	1	10					0	10/0.361	0	0
48	7	0	23	0.262	**0.363**	**0.178**	0.057	23/0.262	0	0	0
	8	3	8					0	0	0	8/0.057
51	9	0	0	**0.235**	**0.366**	**0.205**	**0.072**	0	0	0	0
56	10	1	10	**0.190**	**0.355**	**0.250**	**0.101**	0	10/0.355	0	0
								9.01	19.25	2.35	14.04

③计算不同暴露取值对应的 $\hat{\mathbb{E}}(Y^a)=\frac{1}{n}\sum_{i=1}^{n}\frac{I(a_i=a)y_i}{\hat{f}(a|L)}$

④对 $\mathbb{E}(Y^a)$ 建立边际结构模型

4. 在边际结构模型中考虑效应修饰

假设戒烟L对体重Y的影响有性别差异，性别用V表示

边际结构模型：$E(Y^a|V) = \beta_0 + \beta_1 a + \beta_2 aV + \beta_3 V$

- 女性：V=1 $E(Y^a|V=1) = \beta_0 + \beta_1 a + \beta_2 a + \beta_3$
- 男性：V=0 $E(Y^a|V=0) = \beta_0 + \beta_1 a$

当$\beta_2 \neq 0$时，效应修饰作用存在

⚠️效应修饰 VS混杂

many students have difficulty understanding the distinction because the same statistical methods–stratification (Chapter 4) or regression (Chapter 15)–are often used both for confounder adjustment and detection of effect modification.

Thus, there may be some advantage to teaching these concepts using *marginal structural models*, because then methods for confounder adjustment (IP weighting) are distinct from methods for detection of effect modification (adding treatment-covariate product terms to a marginal structural model).

4. 在边际结构模型中考虑效应修饰

- 参数估计方法模型： $\mathbb{E}[Y|A,V] = \theta_0 + \theta_1 A + \theta_2 VA + \theta_3 V$

- 标准化： $\mathbb{E}[Y^a|V] = \sum_l \mathbb{E}[Y|A=a, L=l|V]\mathbb{P}[L=l|V]$

- IPW: $\mathbb{E}[Y^a|V] = \mathbb{E}\left[\frac{I(A=a)}{\mathbb{P}(A|V,l)}\right]$ （根据IPW公式计算在不同性别的因果效应）

- 模型：引入协变量V，即 $\mathbb{E}[Y^a|V] = \beta_0 + \beta_1 a + \beta_2 Va + \beta_3 V$

➢ 数据

暴露为二分类 混杂为二分类	→	暴露为二分类 混杂为连续型	→	暴露为连续型 混杂为二分类/连续型

➢ 方法

非参数 IPW/standardization → 参数化 IPW/standardization → IPW/standardization + Marginal structural models

---

➢ 因果待估量

$E(Y^{a=1}) - E(Y^{a=0})$　　$E(Y^{a=1}) - E(Y^{a=0})$　　$E(Y^{a=a_1}) - E(Y^{a=a_2})$

- 借助边际结构模型$E(Y^a) = f(a)$

➢ 关键步骤

各值可直接计算得到，
不需要建模，故称非参数

标准化

利用可观测到的$Y^{a=1}$数据建模：

$\mathbb{E}(Y^{a=1} \mid L) = \beta_0 + \beta_1 l$

利用可观测到的$Y_i(0)$数据建模：

$\mathbb{E}(Y^{a=0} \mid L) = \beta_0 + \beta_1 l$

IPW

拟合logistic回归方程，估计 $\mathbb{P}(A = a|L)$：

$\mathbb{P}(A = 1|L) = \text{logit}^{-1}(\beta_0 + \beta_1 l)$

$\mathbb{P}(A = 1|L) = \text{logit}^{-1}(-6.22 + 0.13l)$

$\mathbb{P}(A = 0|L) = 1 - \mathbb{P}(A = 1|L)$

构建暴露模型并估计$\hat{f}(A = a|L)$

对 $\mathbb{E}(Y^a)$ 建立边际结构模型

Wrap up

1、混杂为二分类时，标准化法：分层构建模型$\mathrm{E}(\mathrm{Y}^{a}|L=0)=f_1(a)$、$\mathrm{E}(\mathrm{Y}^{a}|L=1)=f_2(a)$

IPW：分层计算不同暴露水平的概率$\mathrm{P}(A=a|L=0)$、$\mathrm{P}(A=a|L=1)$

2、混杂为连续型时，IPW：构建条件概率密度函数f(A|L) 来估算不同暴露水平、不同混杂组合的概率

$\mathrm{P}(A=a|L=l)$

3、当暴露、混杂中的任意一个或两个同时变成连续型变量时，就需要用观察到的数据构建相关模型来估算因果推断中需要的一些值。除此之外，基本步骤与非参数的IPW或标准化法是一致的。

Thank you

lyx 2022.8.19