L9

Bayesian Network.

ChainRule: $P(X_j \mid X_1 \cdots X_{j-1}) = P(X_j \mid pa_j) = \prod_{i=1} P(x_i \mid pa_i)$

根据ChainRule可以拆分联合分布，再根据独立性可以绘出DAG.

d-separation: $X \to Y \to Z$ "V-structure"

$X \leftarrow Y \leftarrow Z$ ✓ $X \to Y \leftarrow Z$ X

$X \leftarrow Y \to Z$ $(X \perp Z \mid Y)$ $(X \not\perp Z \mid Y)$

Observation Equivalence:

same skeletons, same V-structure.

Functional Causal Models:

SEM: $x_i = f_i(pa_i, u_i)$ "f" usually linear.

Abduction: 根据观察求U → Action: 更改网络.

→ Prediction: 用M_x和U计算新的值.

L10

后门准则: Z不是X的后继且阻断X Y之间指向X的路径，则称Z是后门.

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

和propensity score的关系

$$P(y \mid do(X=x)) = \sum_z \frac{P(X=x, Y=y, \overset{PA}{Z}=z)}{\boxed{P(X=x \mid \underset{PA}{Z}=z)}}$$ Propensity score

前门准则:

条件① Z中断 X→Y有向边

② X,Z无后门.

③ Z,Y后门被X阻断.

则有 $P(Y=y \mid do(X=x)) = \sum_z \sum_{x'} P(Y=y \mid Z=z, X=\boxed{x'})$

$$P(X=\boxed{x'}) P(Z=z \mid X=x)$$

adjustment: if $P(x,z) > 0$, then

$$P(y \mid do(x)) = \sum_z P(z \mid x) \sum_{x'} P(y \mid z, x') P(x')$$

Rules of do-Calculus.

$\hat{x}$: do(X=x) $G_{\bar{X}}$: deleting all edges go into X

$G_{\underline{X}}$: deleting all edges go out from X

R_1: $P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w)$ if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\bar{X}}}$

R_2: $P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w)$ if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\bar{X}\underline{Z}}}$

R_3: $P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w)$ if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\bar{X}\overline{Z(W)}}}$

Z(W) means the set of Z-nodes that are not ~~ancestor~~ ancestors of any W-node in $G_{\bar{X}}$.

L11:

Parameter learning:

$$\log P(D \mid \theta) = \log \prod_m \prod_i p(x_i^m \mid x_{\pi_i}, \theta) = \sum_i \sum_m \log p(x_i^m \mid x_{\pi_i}, \theta)$$

$$\theta_{ijk} \overset{def}{=} P(x_i = j \mid x_{\pi_i} = k)$$

$$N_{ijk} \overset{def}{=} \sum_m I(x_i^m = j, x_{\pi_i}^m = k)$$

$$\ell = \log \prod_m \prod_{i,j,k} \theta_{ijk}^{N_{ijk}} = \sum_m \sum_{ijk} N_{ijk} \log \theta_{ijk} \quad \hat{\theta}_{ijk}^{ML} = \frac{N_{ijk}}{\sum_{j'} N_{ij'k}}.$$

Structure Learning: Constraint-based method.

首先从统计上得到条件独立性，确定骨架

然后基于条件独立性确定 V-structure (cond不能有V-stru

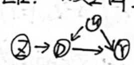
然后根据 no-v-structure 和 no-circle 确定其它边.

Search & score method:

Score: e.g. log-likelihood

Optimization: e.g. hill climbing algorithm.

L12. 工具变量法.



4类人: $G_i = g(W_i(0), W_i(1)) = \begin{cases} nt & \text{if } W_i(0)=0 \ W_i(1)=0 \\ co & \text{if } W_i(0)=0 \ W_i(1)=1 \\ df & \text{if } W_i(0)=1 \ W_i(1)=0 \\ at & \text{if } W_i(0)=1 \ W_i(1)=1 \end{cases}$

假设不存在第三类和第四类人 (One side noncompliance)

$$ITT_{W,co} = \frac{1}{N_{co}} \sum_{i:G_i=co} (W_i(1) - W_i(0))$$

$$ITT_{W,nc} = \frac{1}{N_{nc}} \sum_{i:G_i=nc} (W_i(1) - W_i(0)) = 0$$

$$ITT_W = \underset{N_{nc}/N}{\pi_{nc}} ITT_{W,nc} + \underset{N_{co}/N}{\pi_{co}} ITT_{W,co}$$

$$ITT_Y = ITT_{Y,co} \pi_{co} + ITT_{Y,nc} \cdot \pi_{nc}$$

$$ITT_{Y,co} = \frac{1}{N_{co}} \sum_{i:G_i=co} Y_i(1, W_i(1)) - Y_i(0, W_i(0))$$

$$ITT_{Y,nc} = \frac{1}{N_{nc}} \sum_{i:G_i=nc} (Y_i(1, W_i(1)) - Y_i(0, W_i(0))) = 0$$

$$\tau_{cace} = ITT_{Y,co} = \frac{ITT_Y}{ITT_W}.$$ (nc项均为0)

$$\hat{ITT}_Y = ACE(Z \to Y) = E(Y \mid Z=1) - E(Y \mid Z=0)$$

$$\hat{ITT}_W = ACE(Z \to W) = E(W \mid Z=1) - E(W \mid Z=0)$$

Traditional IV method

$$\beta = \frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{E(D \mid Z=1) - E(D \mid Z=0)}$$

Assume $Z \perp U$, $Z \not\perp D$, $Z \perp Y \mid (D, U)$ and $Y = \beta D + \gamma U + \varepsilon_Y$.

Traditional IV:
线性假设. ↔ Potential Outcome Framework.
unit-level
目标群体为所有人.

Potential Outcome Framework.
Z_i 无混杂.
Exclusive restriction on potential outcome.
$Y_i(z,w) = Y_i(1-z,w)$
目标群体是 compliers
One-side ~~condition~~. noncompliance.

证明 back-criterion 和 front-criterion 一般方法.

$P_m(*) = P(* \mid do(X=x))$

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

$P_m(Z=z) = P(Z=z)$ (因 X 到 Z 无有向边

$P_m(Z=z \mid X=x) = P_m(Z=z)$ 当 X 与 Z d-seprated 时.

Rule1: The Causal Effect Rule

PA 为 X 父结点. $P(Y=y \mid do(X=x)) = \sum_z P(Y=y \mid X=x, P_A=z) P(P_A=z)$

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

↳ Propensity score.

前后门准则证明:

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

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

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

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

前门:

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

后门

$= \sum_z P(Z=z \mid X=x) \sum_{x'} P(Y=y \mid X=x', Z=z) P(X=x')$

Rubain's Framework for RAM (obs study

1. Assess balance

2. if not improve balance key: propensity score

Design ① trimming
② Blocking
③ Matching

3. Estimation
① weighting
② Blocking + regression
③ matching + regression

Analysis

$$\hat{\tau}^{ht} = \frac{1}{N}\sum_{i=1}^{N}\frac{W_i Y_i^{obs}}{e(X_i)} - \frac{1}{N}\sum_{i=1}^{N}\frac{(1-W_i)\cdot Y_i^{obs}}{1-e(X_i)}$$