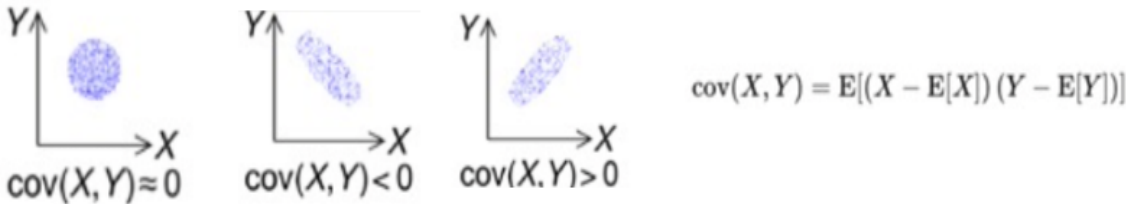


上 海 交 通 大 学 版权所有，不得翻印
(2022 至 2023 学 年 第 1 学 期 期 中 笔 试 + 平 时 作 业)

课程名称 CS7323 从数据学习因果结构

考试时间: _____	小时	_____	分钟
班级号 _____	学号 _____		
姓名 _____	成绩 _____		

1. 回答下面的相关的概率基础问题：
- (1) 对下面的每一个示意图，各举一个与示意图含义一致的概率分布 $p(x, y)$ ，计算相应的 $\text{cov}(x, y)$ 并说明与图示结果一致。



(2) 推导

$$E(Y | X) = E(Y) + \rho_{X,Y} \cdot \sigma_Y \frac{X - E(X)}{\sigma_X}, \quad \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

(3) 证明

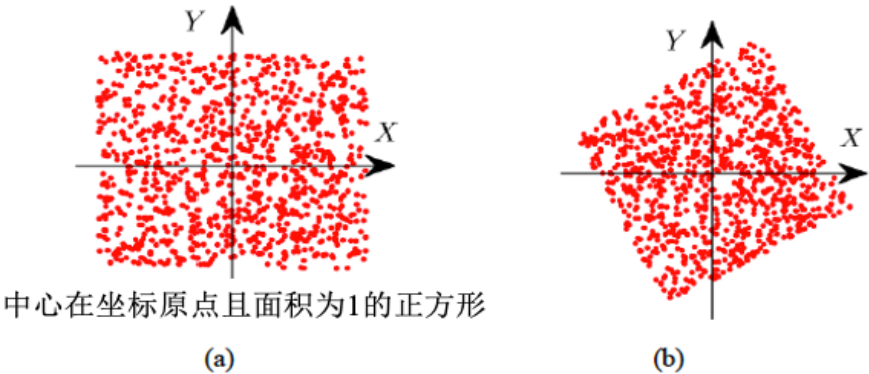
Hölder's inequality: Let p and q satisfy $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, and $1/p + 1/q = 1$. The Hölder's inequality states that $E |XY| \leq (E |X|^p)^{1/p} (E |Y|^q)^{1/q}$.

(4) 举例说明

If $X \perp\!\!\!\perp Y$, then $E[XY] = E[X]E[Y]$, and $\text{cov}[X, Y] = E[XY] - E[X]E[Y] = 0$.

The converse is not true: $\text{cov}[X, Y] = 0 \not\Rightarrow X \perp\!\!\!\perp Y$.

2. 图 (a) 和 (b) 所示概率分布 $p(x, y)$ 在 Support (即红色点域) 上是均匀的

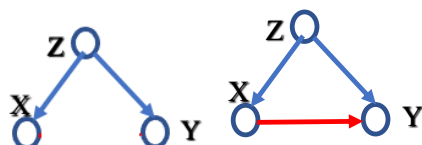
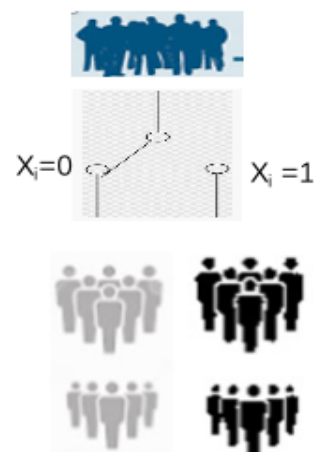


- (1) 写出图 (a) 上概率密度的数学表达式, 并给出 $E(xy) = ?$
- (2) 判断 “两个分布都不满足 $p(x, y) = p(x)p(y)$ ” 的对错, 并给出理由。
- (3) 将图 (a) 上概率密度分布以中心为圆心旋转一个角度 v , 画出所得的概率密度分布 $p(x, y|v)$ 的示意图,
- (4) 考虑上述所得的 $p(x, y|v)$ 且 v 只取值一周, 指出在哪些 v 值上, 有
 $p(x, y|v) = p(x|v)p(y|v)$ 和 $E(xy|v) = E(x|v)E(y|v)$, 两者之一或同时成立。
- (5) 以上为例, 讨论关联 dependence、不相关、与独立在概念上的异同。

3. 在随机控制实验 RCT 的因果分析中, 基本做法是计算和比较下面的差值是否显著:

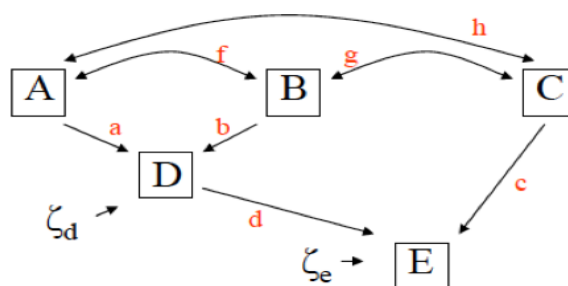
$$\bar{\Delta} = \frac{1}{m} \sum_{i=1}^n x_i Y_i - \frac{1}{n-m} \sum_{i=1}^n (1-x_i) Y_i.$$

- (1) 借助右图, 说明上式的含义。
- (2) 指出 counterfactual、潜在输出、与 $X \rightarrow Y$ 映射函数在概念上的异同。
- (3) 在满足什么条件下, 比较上面差值 $\bar{\Delta}$ 的做法可以称为是在分析因果效应, 解释你的理解。
- (4) 借助下图, 解释 $Y \perp X | Z$ 与 $Y_x \perp X | Z$ 的异同。

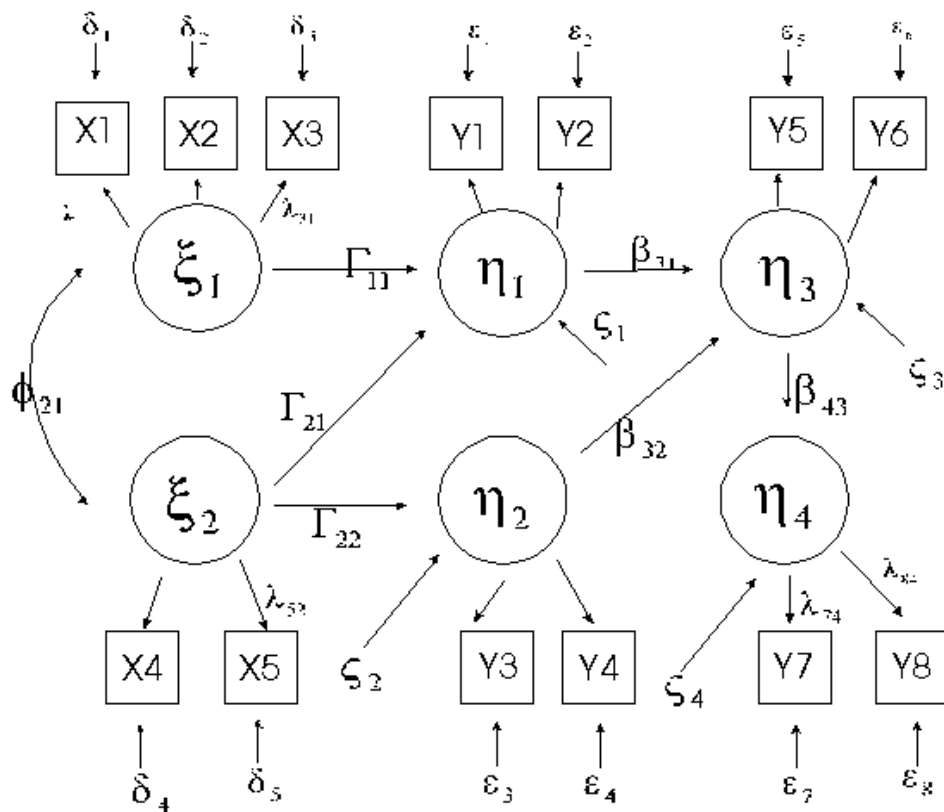


4. 回答下述有关路径分析的问题

- (1) 解释右图中各种边、方框、与字母符号的含义。
- (2) 以 A 与 E 方框为例, 计算对应的 direct effect、indirect effect、与 total effect。
- (3) 用右图解释正确路径应满足的三规则:
 No loops, only one curved arrow per path, no forward then back.



(4) 根据下图将式子 $\eta = B\eta + \Gamma\xi + \zeta$ 中的每一符号都用相应的矩阵和向量表示出来。



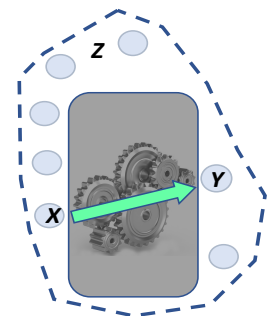
5. 探讨下面与因果分析相关的问题：

(1) 证明 $MI = \int P(X,Y) \ln \frac{P(X,Y)}{P(X)P(Y)} dXdY = L_{Y|X} + E_Y$

$$L_{Y|X} = \int P(X,Y) \ln P(Y|X) dXdY, \quad E_Y = - \int P(Y) \ln P(Y) dY.$$

(2) 讨论 最大似然监督学习与最大化 $L_{Y|X}$ 的共同点，解释 E_Y 的含义，进而解释 MI 的含义、以及它与最大似然监督学习的关系。

(3) 进一步借助右图，分别从内在机制与外在影响，讨论因果推断与机器学习的共同点和不同点。



(4) 讨论下面四个对因果效应的度量，在什么情形下是一样的？在什么情形下是不同的，为什么？

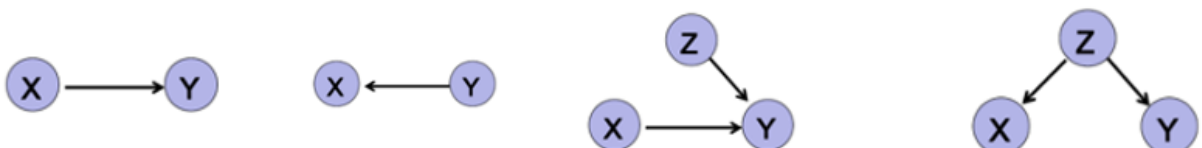
$$\Delta P = P(Y|X=1) - P(Y|X=0), \quad \Delta\mu = E(Y|X=1) - E(Y|X=0)$$

$$\Delta D = P(Y|do X=1) - P(Y|do X=0), \quad \bar{\Delta}\mu = E(Y|do X=1) - E(Y|do X=0)$$

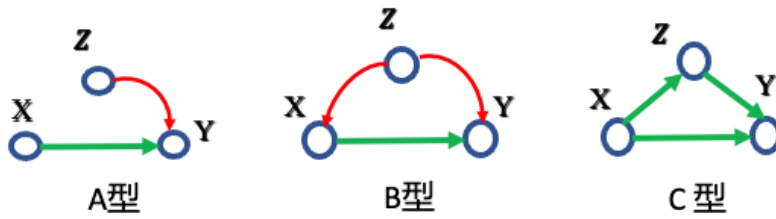
特别地，在题 2 的 RCT 分析场景下，指出它们与 $\bar{\Delta}$ 的共同点和不同点。

6. 强制干预 $X = x$ 记为 $do X = x$ 并简化为 $do x$

(1) 将每个式子 $p(Y|x) = p(Y|do x)$ 和 $p(Y|x) = p(Y) \neq p(Y|x)$ 正确配对到合适的图表示，并解释理由。

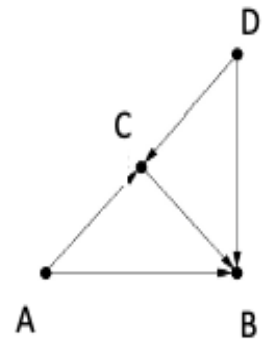


- (2) 分别就这三种 DAG，给出 $P(Y|do\ x)$ 的计算公式，并比较三者的共同点和不同点，进而给出对 $P(Y|X)$ 和 $P(Y|do\ X)$ 之异同的理解。



- (3) 这三种 DAG 是否存在于右图中为部件？找出你认为的那些，并讨论它们在右图整个 DAG 中如何协同起作用。

- (4) 在 C 节点 do 干预后，说明 C、D 将分别如何影响 A 和 B。



7. 回答下述问题：

- (1) 借助下表解释什么是 Yule-Simpson 悖论。

	康复	未康复	康复率
吃药	20	20	50%
安慰剂	16	24	40%

性别	康复	未康复	康复率
男性	18	12	60%
女性	2	8	20%

性别	康复	未康复	康复率
男性	18	12	60%
女性	7	3	70%

性别	康复	未康复	康复率
女性	2	8	20%
安慰剂	9	21	30%

- (2) 按照这个 DAG，证明

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

- (3) 给出 $P(Y|X)$ 和 $P(Y|do\ X)$ ，并比较。

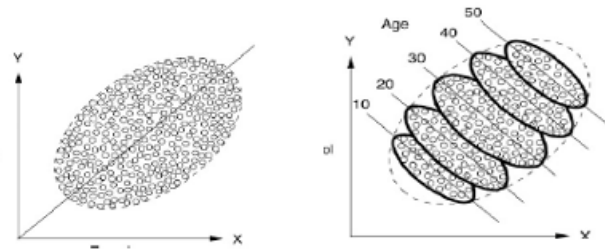
- (4) 根据上表，分别计算 $P(\text{康复}|\text{吃药})$ 和 $P(\text{康复}|do\ \text{吃药})$ 。

8. 回答下述问题：

- (1) 借助下表解释什么是 Berkson 悖论？

	General Population			Hospitalized in Last Six Months		
Respiratory disease? ↓	Bone disease? ↓			Bone disease? ↓		
	Yes	No	% Yes	Yes	No	% Yes
Yes	17	207	7.6	5	15	25.0
No (control)	184	2,376	7.2	18	219	7.6

(2) 可以用下图解释 Berkson 悖论吗？为什么？

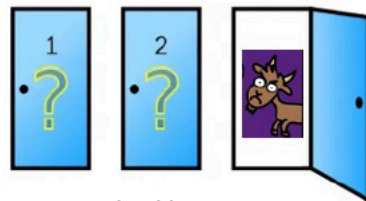


(3) 给出 $P(Y|X)$ 和 $P(Y|do X)$ 。

(4) 解释 MONTY HALL 问题

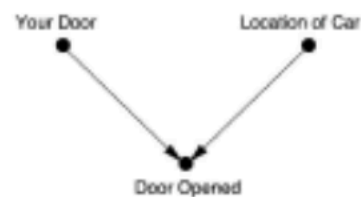
MONTY HALL PROBLEM

in September 1990, a writer named Marilyn vos Savant wrote a column in Parade magazine: "you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say #1, and the host, whoknows what's behind the doors, opens another door, say #3, which has a goat. He says to you, 'Do you want to pick door #2?' Is it to your advantage to switch your choice of doors?"



vos Savant argued that contestants should switch doors.

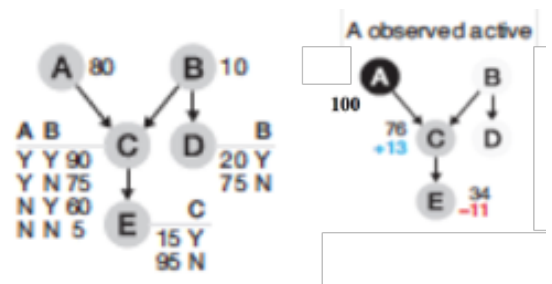
Door 1	Door 2	Door 3	Outcome If You Switch	Outcome If You Stay
Auto	Goat	Goat	Lose	Win
Goat	Auto	Goat	Win	Lose
Goat	Goat	Auto	Win	Lose



Causal diagram for *Let's Make a Deal*.

9. 考虑概率图模型的下述问题：

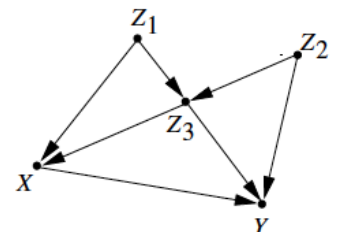
(1) 写出分布 $p(A,B,C,D,E)$ 对应于左图的乘积分解式，图中的数字是用百分比表示的概率，进而推导说明右图中因 A 节点激活而引起的数字变化。

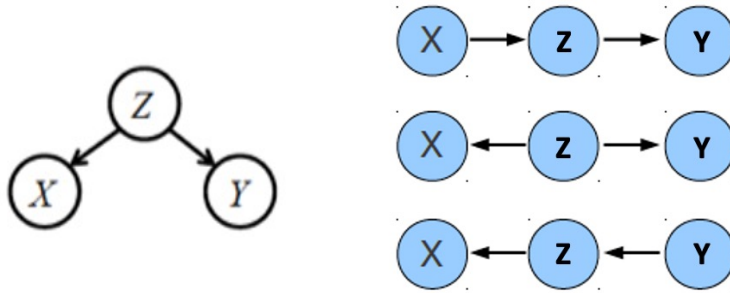


(2) 对于右图，推导出计算公式：

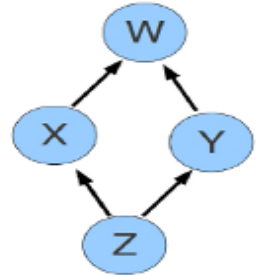
$$P(y|do(x_0)) = \sum_{z_1, z_2, z_3} P(z_1)P(z_2)P(z_3|z_1, z_2)P(y|z_2, z_3, x_0)$$

(3) 写出下面每个 DAG 图对应的 $p(x,y,z)$, 并说出你的观察。



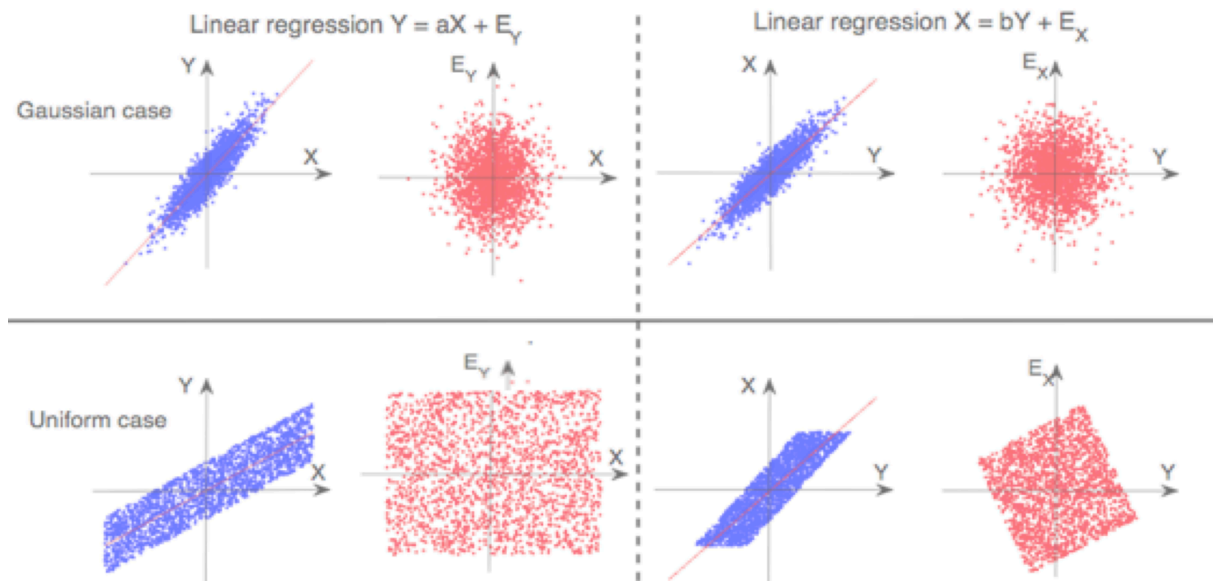


- (4) 骨架指 DAG 去掉方向后得到的无向图，写出与右图 DAG 骨架一样且 v 节点一样的所有 DAG，并写出每个 DAG 图对应的 $p(x,y,z,w)$ ，说出你的观察。



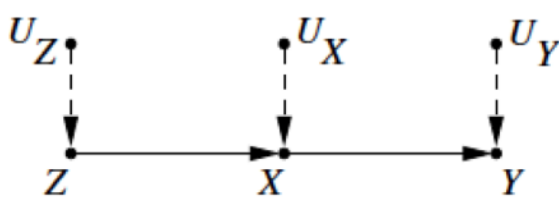
10. 考虑下述问题：

- (1) 说明因果方向 $Y=aX+E$ 在上排情况不可以辨识，而在下排情况可辨识。



- (2) 按照下图求出

$$E(Y|do(x_0)) = ?$$



$$\begin{aligned} z &= f_Z(u_Z) \\ x &= f_X(z, u_X) \\ y &= f_Y(x, u_Y) \end{aligned}$$

(3) 将 $X=BX+e$ 写成 $X=Ae$, $A=?$ 且应满足什么条件? 如果有一个矩阵 W 使 $WA=PD$, 这里 P 是置换矩阵而 D 是对角矩阵, 说明 W 与 A 是什么关系, 若进一步要求 A 是下三角矩阵, W 与 A 又是什么关系?

(4) 什么是工具变量 IV? 下面情况中, 哪些是有效的 IV, 哪些不是, 并给出理由。

