期中笔试

1.(1). 对 $Cov(X,Y)=0$.

y \ x	-1	0	1
-1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

对 $Cov(X,Y)<0$.

y \ x	-1	0	1
-1	0	0	$\frac{2}{3}$
0	0	0	0
1	$\frac{1}{3}$	0	0

对 $Cov(X,Y)>0$.

y \ x	-1	0	1
-1	$\frac{1}{3}$	0	0
0	0	$\frac{1}{3}$	0
1	0	0	$\frac{1}{3}$

(2). $E(Y|X) = \frac{E(XY)}{E(X)} = \frac{E(X)E(Y)+Cov(X,Y)}{E(X)}$

$= E(Y) + \frac{Cov(X,Y)}{E(X)} = E(Y) + \rho_{XY}\sigma_X\sigma_Y\frac{1}{E(X)}$

$\frac{\sigma(X)}{E(X)} = \frac{\sigma^2_{(X)}}{\sigma(X)E(X)}$ $\qquad D(X) = E(X^2) - E^2(X)$

$= \frac{E(X^2) - E^2(X)}{\sigma(X)\,E(X)}$ $\qquad$ 又 $E(X^2) = \sum_1^1 X^2 P(X)$ $\quad E(X) = \sum X P_X$ 又因X固定

$\therefore$ $\qquad \frac{E(X^2)}{E(X)} = X$

$\therefore E(Y|X) = E(Y) + \rho_{XY}\sigma_Y\frac{X-E(X)}{\sigma(X)}$

(3). $E(XY) = \int_y\int_x XY\,df_{XY}(x,y) \le \int_y\int_x \left(\frac{|x|^p}{p} + \frac{|y|^q}{q}\right) df(x,y)$

(4).

y \ x	-1	0	1
-1	$\frac{4}{27}$	$\frac{1}{27}$	$\frac{4}{27}$
0	$\frac{3}{27}$	$\frac{1}{27}$	$\frac{3}{27}$
1	$\frac{4}{27}$	$\frac{1}{27}$	$\frac{4}{27}$

有 $E(XY)=0$. $E(X)=0$ $E(Y)=0$.

$\therefore Cov(X,Y)=0$

但有 $P(X=-1,Y=-1) \neq P(X=-1)\cdot P(Y=-1)$. $\therefore$ 不独立

2， (1)，设边界坐标为 X_0, Y_0，即矩阵长度为 $2X_0 \times 2Y_0$

则 $F(x,y) = \frac{X+X_0}{2X_0} \times \frac{Y+Y_0}{2Y_0}$

$\therefore f(x,y) = \frac{1}{2X_0 \cdot 2Y_0}$

$E(XY) = \iint XY \cdot f(x,y)\,dxdy = \iint \frac{XY}{4X_0Y_0}\,dxdy$

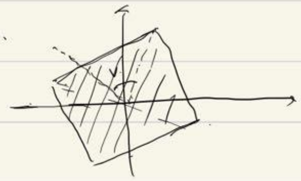
$= 0$

(2)，(a)，显然 $P(x,y) = \frac{1}{2X_0 \cdot 2Y_0}$　$P(x) = \frac{1}{2X_0}$　$P(Y) = \frac{1}{2Y_0}$

$\therefore P(x,y) = P(x)P(Y)$

(b)，显然不是独立的，所以不满足。

(3)，



如图为所需概率密度分布。

(4)　$\theta = 45°, 90°, 135°, 180°, 225°, 270°, 315°, 360°$ 时，满足至少一个。

(5)，关联：表示 X与Y间有某种函数关系，与线性相关、二次相关等。一般通过 P_{XY} 的值判断有无线性相关。

不关联：若 $P=0$ 则线性不相关，但可能有其他相关，因此不一定独立。

独立：即 X与Y无任何相关关系，通过判断 $P(x,y) = P(x)P(Y)$ 满足与否判断。

3, (1) 此式表示，在 X_i 发生情况下有 Y_i 的概率均值，和 X_i 不发生情况下有 Y_i 的概率均值的差如果显著则表示 X_i 对 Y_i 有因果影响。

(2) counterfactual 是指通过对已有结果假设再推理，估计其中一项影响因素的发生概率。

潜在结果是指一种可能的输出，与 counterfactual 不同在与上面是对因的假设，潜在结果是对果的假设。

$X \to Y$ 映射函数则是将所有映射情况都表现出，是一种理想情况。

(3) 必须在控制 X 而个体本身特征相同。

(4) 左图 $Y \perp X \mid Z$ 表示在 Z 发生时，Y 与 X 是独立的。

右图表示，在 Z 发生时，Y 在 X 的作用后也是关于 X 独立的。

4. (1). a, b, c, d 表示路径系数. f, g, h 表示相关系数.

各方形框表示各变量. 两 ζ 表示误差.

(2). $E = dD + cC + \zeta_e$.

$D = aA + bB + \zeta_d$. $C = hA + gB$.

$\therefore E = adA + bdB + \zeta_d d + chA + \underbrace{cgB}_{\text{去掉}} + \zeta_e$.

又 $B = fA$.

$\therefore E = (ad + ch + bdf)A + \zeta_d d + \zeta_e$. 是 total effect.

其中 direct effect 是 0, indirect effect 是 $ad + ch + bdf + cgf$.

(3). no loops: 即不能有环路,如图是有向无环图.

only one curved arrow per path: 指一条通径只能有一次相关系数的参与,如上面将 cgB 项去掉.

no forward then back: 指计算 AD 关系时,不能将 $D-E-C-A$ 此条算入,因为这条在 D 处先前进后后退了。

(4).

$$y = By + \Gamma\xi + \zeta : \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \beta_{31} & \beta_{32} & 1 & 0 \\ 0 & 0 & \beta_{43} & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \begin{bmatrix} \Gamma_{11} & \Gamma_{21} + \phi_{21}\Gamma_{11} \\ \phi_{21}\Gamma_{22} & \Gamma_{22} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}$$

5, (1) $MI = \int P(x,y)\left(\ln \frac{P(x,y)}{P(x)} - \ln P(y)\right) dx\,dy$

$= \int P(x,y) \ln \frac{P(x,y)}{P(x)}\, dx\,dy - \int P(x,y) \ln P(y)\, dx\,dy$

$= L_{y|x} + E_y$

(2) 共同点在于. 都是对概率的对数值求和并使其最大,

E_y 是, 对 $\ln P(y)$ 的在一定概率下的求和. 的负值.

因此 MI 是反应 X 与 Y 的独立性的一项指标,

(3). 共同点, 内在机制上: 两者都是找出. 两变量的影响关系.

外在影响上: 两者都存对外在因素的考虑

不同点, 内在机制上: 因果推断要判断各中间因素, 机器学习则将中间因素视为黑箱.

外在影响: 因果推断是性地考虑外在因素. 机器学习将外在因素都纳入黑箱.

(4). ΔP 与 ΔL 间, 若每个个体间无差异, 则两者是等价的, ΔD 与 ΔU 同理

ΔP 与 ΔD 间. 若固定 X 与否与 Y 无关, 则. 两者等价, ΔL 与 ΔU 同理.

与2 共同点在于都是分析是否有因果关系,, 不同在于2没有用条件概率.

6. (1). 图1，配右边式子，

图2配左边式子，因为X由Y产生，固定X对Y概率不影响。

图3配右边式子，

图4配左边式子，因为X与Y间存在门Z，Y受Z影响。

(2). A，$P(Y|doX) = P(Y|X)$

B，$P(Y|doX) =$

(3) 显然，ABC和BCD在全部考虑时是C型，在考虑BC时是B型，

ACD是A型，

(4). C对A由于A是C后点所以无影响，对B，则考虑B型对B产生影响。

D对A由于C的截断所以无影响，对B则考虑C型对B产生影响。

7. (1). 由左上表看似得出，吃药的康复率比吃安慰剂高，但对男性女性分别研究，

则都得出吃安慰剂的康复率更高。

(2). 由贝叶斯网络，Y由X Z得到，X由Z得到。

因此：$P(X,Y,Z) = P(Z)P(X|Z)P(Y|X,Z)$

(3). $P(Y|do\,X) = \sum_z P(z)P(y|z,x)$

$P(Y|X) = \sum \frac{P(Y,X)}{P(X)}$

(4). $P(\text{康复}|\text{吃药}) = 50\%$

$P(\text{康复}|do\,\text{吃药}) = 0.5\times0.6 + 0.5\times0.2 = 40\%$

8. (1) 在医院中，有呼吸道疾病的人得骨病的概率比没呼吸道疾病的人高，然而在一般人中，则不存在这一现象，原因是医院和一般大众环境不同导致的。

(2) 不可以，此图是用于解释史密斯悖论的。

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

$P(Y \mid x) = \sum \frac{P(Y, x)}{P(x)}$

(4) 先设，1：车，2：羊，3：羊

若不换门，显然是 $\frac{1}{3}$ 概率选中。

若换门，则如表：

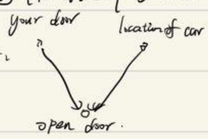
选择的门	1	1	2	3
开的门	2	3	3	2
换后的门	3	2	1	1

如表所示，只有当第一次选到车时，才会因换门而失去奖品，

即选到车的概率等于第一次选到羊的概率，即 $\frac{2}{3}$。

用因果图解释，不换门时，得到车的概率与车的位置无关

但换门时由于有关系，



开的门与车的位置有强相关，

引入了车位置的信息，此时换门，则多引入一个信息，概率会变大

换门产生概率变化就是因为引入了车辆位置这一潜在变量，这也可以解释为什么。

9. (1). $P(A,B,C,D,E) = P(E|C)P(C|AB)\,P(D|B)\,P(A)P(B)$.

原先C在A未激活时，有.

$$P(C) = 0.9\times0.8\times0.1 + 0.75\times0.8\times0.9 + 0.6\times0.2\times0.9 + 0.05\times0.2\times0.9$$

$$= 0.633.$$

A激活后. $P(C) = 0.9\times0.1 + 0.75\times0.9 = 0.765$.

E 在 A 非激活时.

$$P(E) = 0.15\times0.633 + 0.95\times0.367 = 0.443$$

激活后

$$P(E) = 0.15\times0.765 + 0.95\times0.235 = 0.337.$$

(2). $P(Y, do\,x, z_1, z_2, z_3) = P(Y|do\,x, z_3, z_2)\,\underbrace{P(do\,x|z_1,z_3)}_{1}\,P(z_3|z_1,z_2)\,P(z_1)P(z_2)$.

$\Rightarrow \sum P(Y, do\,x, z_1,z_2,z_3) = \sum P(Y|do\,x, z_1, z_2)\,P(z_3|z_1,z_2)\,P(z_2)P(z_1)$

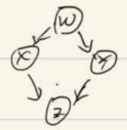
$\Rightarrow P(Y|do\,x)\,\underbrace{P(do\,x)}_{1} = \sum P(Y|do\,x, z_1, z_2)\,P(z_3|z_1,z_2)\,P(z_2)P(z_1)$

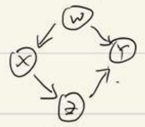
$\Rightarrow P(Y|do\,x) = \sum P(Y|do\,x, z_1, z_2)\,P(z_3|z_1,z_2)\,P(z_2)P(z_1)$

(3). ① $P(Y|Z)P(Z|X)P(X)$

② $P(Y|Z)P(X|Z)P(Z)$

③ $P(X|Z)P(Z|Y)P(Y)$.

(4) ① 与原图类似的.  共4个. $P(Z|X,Y)P(X|W)P(Y|W)P(W)$.

② 一个反向.  共6个. $P(Y|W,Z)P(Z|X)P(X|W)P(W)$.

观察到共分以上两类.

10, (1). 在高斯噪声下, 线性, $Y=aX+E_Y$ 与 $X=bY+E_X$ 图是相同的, 无法判别谁是谁的因.

在标准情况下, 显然左图. X的均值分布是正常的而 Y的分布不合理 因而可判断. X与Y的因果关系,

(2). $E(Y|do\,X) = \int Y f_Y(x_0, u_Y) f_Z(u_Z) . dY.$

(3). $A = \frac{I}{I-B}$. 应满足 $I-B \neq 0$.

(4). 工具变量是独立于混杂因子与Y无直接路径, 与X强关联用于求X,Y间因果关系而引入的

只有(a)是.

(b)与X无强关联. (c)与混杂因子不独立. (d)与Y有直接路径.