Name: Weijun Zhu   Algorithmic Learning Theory HW#01

1.(a).

X_2 \ X_1	-1	0	1	$f_{X_2}()$
2	1/10	1/10	1/10	3/10 = 24/80
4	1/8	1/4	1/5	23/40 = 46/80
6	1/10	1/80	1/80	10/80 = 10/80
$f_{X_1}()$	26/80	29/80	25/80	1

(b).

$$E(X_1) = -1\times f_{X_1}(1) + 0\times f_{X_1}(0) + 1\times f_{X_1}(1)$$
$$= -1\times 26/80 + 0 + 1\times 25/80$$
$$= -1/80$$

$$E(X_2) = 2\times f_{X_2}(2) + 4\times f_{X_2}(4) + 6\times f_{X_2}(6)$$
$$= 2\times 24/80 + 4\times 46/80 + 6\times 10/80$$
$$= 292/80$$

(c).

$$Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2]$$
$$E[X_1X_2] = (-1)\cdot(2)\cdot(\tfrac{1}{10}) + (1)\cdot(2)\cdot(\tfrac{1}{10}) + (-1)\cdot(4)\cdot(\tfrac{1}{8}) + (1)\cdot(4)\cdot(\tfrac{1}{5})$$
$$+ (-1)\cdot(6)\cdot(\tfrac{1}{10}) + (1)\cdot(6)\cdot(\tfrac{1}{80})$$
$$= (-\tfrac{2}{10}) + (\tfrac{2}{10}) + (-\tfrac{4}{8}) + (\tfrac{4}{5}) + (-\tfrac{6}{10}) + (\tfrac{6}{80})$$
$$= -0.225$$
$$\therefore Cov(X_1, X_2) = -0.225 - (-\tfrac{1}{80})\cdot(\tfrac{292}{80})$$
$$= -0.179375$$
$$Var(X_1) = E(X_1^2) - E(X_1)^2$$
$$= [(-1)^2\times\tfrac{26}{80} + (0)^2\times\tfrac{29}{80} + (1)^2\times\tfrac{25}{80}] - (-\tfrac{1}{80})^2$$
$$= 0.635844371$$
$$Var(X_2) = E(X_2^2) - E(X_2)^2$$
$$= [(2)^2\times\tfrac{24}{80} + (4)^2\times\tfrac{46}{80} + (6)^2\times\tfrac{10}{80}] - (\tfrac{292}{80})^2$$
$$= 1.5775$$

(d). Let's compute correlation between X_1 & X_2.

$$CORR(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}} = \frac{-0.179375}{\sqrt{0.636}\cdot\sqrt{1.7557}} = -0.169749$$

So X_1 and X_2 are not independent, because correlation is not zero.

(e). $f_{X_1|X_2}(t_1|X_2=2) = \frac{f(t_1,2)}{f_{X_2}(2)} \rightarrow X_1=-1: \frac{f(-1,2)}{f_{X_2}(2)} = \frac{1/10}{3/10} = \frac{1}{3}$

$X_1=0: \frac{f(0,2)}{f_{X_2}(2)} = \frac{1/10}{3/10} = \frac{1}{3}$

$X_1=1: \frac{f(1,2)}{f_{X_2}(2)} = \frac{1/10}{3/10} = \frac{1}{3}$

$f_{X_2|X_1}(t|X_1=0) = \frac{f(t,0)}{f_{X_1}(0)} \rightarrow X_2=2: \frac{f(2,0)}{f_{X_1}(0)} = \frac{1/10}{29/80} = \frac{8}{29}$

$X_2=4: \frac{f(4,0)}{f_{X_2}(0)} = \frac{1/4}{29/80} = \frac{20}{29}$

$X_2=6: \frac{f(6,0)}{f_{X_2}(0)} = \frac{1/80}{29/80} = \frac{1}{29}$

(f). $E[X_1|X_2=2] = (X_1=-1) \times f_{X_1|X_2}(-1|X_2=2) + (X_1=0) \times f_{X_1|X_2}(0|X_2=2)$
$+ (X_1=1) \times f_{X_1|X_2}(1|X_2=2)$.
$= (-1) \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}$
$= 0$.

(g). $Var[X_2|X_1=0] = E(X_2^2|X_1=0) - [E(X_2|X_1=0)]^2$.

$E(X_2^2|X_1=0) = \sum X_2^2 \cdot f_{X_2|X_1}(X_2|X_1=0)$
$= (2)^2 \cdot (\frac{8}{29}) + (4)^2 \cdot (\frac{20}{29}) + (6)^2 \cdot (\frac{1}{29})$
$= 13.37931034$

$[E(X_2|X_1=0)]^2 = [\sum X_2 \cdot f_{X_2|X_1}(X_2|X_1=0)]^2$.
$= [2 \cdot \frac{8}{29} + 4 \cdot \frac{20}{29} + 6 \cdot \frac{1}{29}]^2$
$= 12.37098692$

$\therefore Var[X_2|X_1=0] = 13.38 - 12.37$
$= 1.01$

2.(a). We have, $\begin{cases} P(\text{Positive} \mid \text{cancer}) = 0.9 \\ P(\text{Positive} \mid \text{cancer}^c) = 0.09 \\ P(\text{cancer}) = 0.08 \end{cases}$

$$P(\text{cancer} \mid \text{Positive}) = \frac{P(\text{Positive} \mid \text{cancer}) \cdot P(\text{cancer})}{P(\text{Positive})}$$

$$P(\text{Positive}) = P(\text{Positive} \mid \text{cancer}) \cdot P(\text{cancer}) + P(\text{Positive} \mid \text{cancer}^c) \cdot P(\text{cancer}^c)$$
$$= 0.9 \times 0.08 + 0.09 \times 0.92$$
$$= 0.1548$$

$$P(\text{cancer} \mid \text{Positive}) = \frac{0.9 \times 0.08}{0.1548} = 0.465116279$$

(b). K = the number of positive people test

$$P(\text{cancer} \mid K=2) = \frac{P(K=2 \mid \text{cancer}) \cdot P(\text{cancer})}{P(K=2)}$$

$$P(K=2 \mid \text{cancer}) = \binom{2}{2} 0.9^2 = 0.81$$

$$P(K=2 \mid \text{cancer}^c) = \binom{2}{2} 0.09^2 = 0.0081$$
$$P(K=2) = 0.81 \times 0.08 + 0.0081 \times 0.92 = 0.072252$$

$$\therefore P(\text{cancer} \mid K=2) = 0.896860987$$

(c). $P(\text{Negative}) = 1 - P(\text{Positive}) = 1 - 0.1548 = 0.8452$
$P(\text{cancer}^c) = 1 - P(\text{Cancer}) = 1 - 0.08 = 0.92$
$P(\text{Negative} \mid \text{cancer}^c) = 1 - P(\text{Positive} \mid \text{cancer}^c) = 1 - 0.09 = 0.91$

$$P(\text{cancer}^c \mid \text{Negative}) = \frac{P(\text{Negative} \mid \text{Cancer}^c) \cdot P(\text{cancer}^c)}{P(\text{Negative})}$$

$$= \frac{0.91 \times 0.92}{0.8452}$$

$$= 0.990534785$$

扫描全能王 创建

3.(a). If A & B are independent, then we have $P(A|B)=P(A)$ and $P(A\cap B)=P(A)\cdot P(B)$

① (not A, B): $P(A^c, B)=P(B)-P(A,B)$

$$=P(B)-P(A)\cdot P(B)=[1-P(A)]\cdot P(B)$$
$$=P(A^c)\cdot P(B).$$

② (A, not B): $P(A, B^c)=P(A)-P(A,B)=P(A)-P(A)\cdot P(B)=[1-P(B)]\cdot P(A)$
$$=P(B^c)\cdot P(A)=P(A)\cdot P(B^c).$$

③ (not A, not B): $P(A^c, B^c)=P(A^c)-P(A^c,B)=P(A^c)-P(A^c)\cdot P(B).$
$$=[1-P(B)]\cdot P(A^c)$$
$$=P(A^c)\cdot P(B^c).$$

(b). - The Probability of sure event is 1, and sure event means it always happen, do not affect by other events. Then we can get
$P(\text{sure event} \mid \text{other events})=1$.
So, we can say sure event is independent of all other events.

- The impossible event has probability "0", and it means the event never happen whatever other things happen. Then we get
$P(\text{impossible event} \mid \text{other events})=0$
So, we can say impossible event is independent of all othe events.

(c). Assume that A = the event with Probability "0" ; B = all other events
If A & B are not independent, we get
$$P(B|A)=\int_{(x,y)} f_{X|Y}(X=B|y=A)\,dx \ ; \ P(B|A)=\frac{P(A|B)\cdot P(B)}{P(A)}.$$
We can see if $P(A)=0$, $P(B|A)$ is undefined.
So, A & B should be independent.

(d). When A & B are independent, $P(A\cap B)=P(A)\cdot P(B)$, We can say that $A\cap B\neq\phi$, and $P(A|B)=P(A)$ & $P(B|A)=P(B)$

4(a). We can not get absolute answer, fair or fake.
By geometric distribution, $f(k)=(1-P)^{k-1}P$
If the coin is fair, we get $P=0.5$ and $1-P=0.5$

$$f(4)=(1-0.5)^3\times0.5=0.0625$$

This means you toss 4 times, and the 4th time you get head.
The probability of this suitation is 0.0625.

If the coin is fake, we get $P=0.2$ and $1-P=0.8$.

$$f(4)=(1-0.2)^3\times0.2=0.1024$$

This means you toss 4 times, and the 4th time you get head.
The Probability of this suitation is 0.1024.

So we can say the probability of fake coin is bigger than fair coin, but we can not get exactly answer.

(b). $(1-P_1)^x P_1 \geq (1-P_2)^x P_2$; $0\leq P_1 \leq P_2 \leq 1$

$(1-P_1)^x P_1 - (1-P_2)^x P_2 \geq 0$

$\ln(1-P_1)^x + \ln P_1 - \ln(1-P_2)^x - \ln P_2 \geq 0$

$\ln(1-P_1)^x + \ln P_1 \geq \ln(1-P_2)^x + \ln P_2$

$x\ln(1-P_1) + \ln P_1 \geq x\ln(1-P_2) + \ln P_2$

$$x \geq \frac{\ln P_2 - \ln P_1}{\ln(1-P_1) - \ln(1-P_2)}$$

$$x \geq \log_{\left(\frac{1-P_1}{1-P_2}\right)}\left(\frac{P_2}{P_1}\right)$$

$\therefore$ The minimum number of failure before success is $\log_{\left(\frac{1-P_1}{1-P_2}\right)}\left(\frac{P_2}{P_1}\right)$