STA4001 Assign 1

杨景兰 Jinglan Yang 121090699

#1 **(Chapter 3, Exercise 3)** The joint probability mass function of X and Y, $p(x, y)$ is given by

$$\begin{array}{lll} p(1,1)=\frac{1}{9} & p(2,1)=\frac{1}{3} & p(3,1)=\frac{1}{9} \\ p(1,2)=\frac{1}{9} & p(2,2)=0 & p(3,2)=\frac{1}{18} \\ p(1,3)=0 & p(2,3)=\frac{1}{6} & p(3,3)=\frac{1}{9} \end{array}$$

Compute $E[X|Y=i]$ for $i=1,2,3$.

$E[X|Y=i] = \sum_x x \cdot P(X|Y=i)$

① $P(Y=1) = \frac{1}{9}+\frac{1}{3}+\frac{1}{9} = \frac{5}{9}$

$E[X|Y=1] = 1\times\frac{\frac{1}{9}}{\frac{5}{9}} + 2\times\frac{\frac{1}{3}}{\frac{5}{9}} + 3\times\frac{\frac{1}{9}}{\frac{5}{9}} = 1\times\frac{1}{5}+2\times\frac{3}{5}+3\times\frac{1}{5} = 2$

② $P(Y=2) = \frac{1}{9}+0+\frac{1}{18} = \frac{1}{6}$

$E[X|Y=2] = 1\times\frac{\frac{1}{9}}{\frac{1}{6}} + 2\times\frac{0}{\frac{1}{6}} + 3\times\frac{\frac{1}{18}}{\frac{1}{6}} = 1\times\frac{2}{3}+2\times 0+3\times\frac{1}{3} = \frac{5}{3}$

③ $P(Y=3) = 0+\frac{1}{6}+\frac{1}{9} = \frac{5}{18}$

$E[X|Y=3] = 1\times\frac{0}{\frac{5}{18}} + 2\times\frac{\frac{1}{6}}{\frac{5}{18}} + 3\times\frac{\frac{1}{9}}{\frac{5}{18}} = 1\times 0+2\times\frac{3}{5}+3\times\frac{2}{5} = \frac{12}{5}$

Hence, $E[X|Y=1]=2$, $E[X|Y=2]=\frac{5}{3}$, $E[X|Y=3]=\frac{12}{5}$

#2 **(Chapter 3, Exercise 18 (d))** Let $X_1, ..., X_n$ be independent random variables having a common distribution function that is specified up to an unknown parameter θ. Let $T = T(X_1, .., X_n)$ be a function of the data. If the conditional distribution $X_1, ..., X_n$ given $T(X_1, ..., X_n)$ does not depend on θ then T is said to be a sufficient statistic for θ. Show that, if X_i are Poisson random variables with mean θ, $T(X_1, ..., X_n) = \sum_{i=1}^n X_i$ is a sufficient statistic for θ.

By the Additive Property of poisson distribution, $T(X_1,\dots,X_n) = \sum_{i=1}^n X_i \sim \text{Poisson}(n\theta)$

if $\theta = \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}T$,

$P(X_1,\dots,X_n,T) = \prod_{i=1}^n e^{-\theta}\frac{\theta^{X_i}}{X_i!} = e^{-n\theta}\frac{\theta^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n X_i!} = e^{-T}\frac{\theta^T}{\prod_{i=1}^n X_i!}$

otherwise, $P(X_1,\dots,X_n,T)=0$

$P(X_1,\dots,X_n|T) = \frac{P(X_1,\dots,X_n,T)}{P(T)} = \frac{\frac{e^{-T}\theta^T}{\prod_{i=1}^n X_i!}}{\frac{e^{-n\theta}(n\theta)^T}{T!}} = \frac{T!}{n^T\prod_{i=1}^n X_i!}$, for $\theta=\frac{1}{n}T$

In all, $P(X_1,\dots,X_n|T) = \begin{cases} \frac{T!}{n^T\prod_{i=1}^n X_i!}, & \text{for } \theta=\frac{1}{n}T \\ 0, & \text{o.w.} \end{cases}$ $\Rightarrow$ independent of θ.

Therefore, $T(X_1,\dots,X_n) = \sum_{i=1}^n X_i$ is a sufficient statistic for θ.

#3 **(A Trading Model)** The fundamental value of the security is a random variable V taking two possible states H(high) and L(low) and $P(V = H) = p_0$. There are two types of traders in this market: informed traders and uninformed traders. The proportion of informed traders is $\alpha \in (0, 1)$. Once the market opens, the state of V is realized and informed traders get to know V. They will buy a share of the security if $V = H$ and sell it otherwise. Uninformed traders do not know the state of V. They will buy or sell the security randomly with equal probability. Compute the conditional probability that $V = H$ if you observe that a trader is trying to buy the security.

$P(V=H|\text{buy}) = \frac{P(\text{buy}|V=H)P(V=H)}{P(\text{buy}|V=H)P(V=H)+P(\text{buy}|V=L)P(V=L)}$

$= \frac{P(V=H)[P(\text{buy}|\text{informed},V=H)P(\text{informed}|V=H) + P(\text{buy}|\text{uninformed},V=H)P(\text{uninformed},V=H)]}{P(\text{buy}|V=H)P(V=H) + P(V=L)[P(\text{buy}|\text{informed},V=L)P(\text{informed}|V=L) + P(\text{buy}|\text{uninformed},V=L)P(\text{uninformed}|V=L)]}$

$= \frac{p_0\cdot[\alpha+\frac{1}{2}(1-\alpha)]}{p_0[\alpha+\frac{1}{2}(1-\alpha)]+(1-p_0)[0+\frac{1}{2}(1-\alpha)]}$

$= \frac{p_0(1+\alpha)}{p_0(1+\alpha)+(1-p_0)(1-\alpha)}$

$= \frac{p_0+\alpha p_0}{1+2\alpha p_0-\alpha}$

Hence, $P(V=H|\text{buy}) = \frac{p_0+\alpha p_0}{1+2\alpha p_0-\alpha}$

#4 **(Chapter 3, Exercise 26)** You have two opponents with whom you alternate play. Whenever you play A, you win with probability p_A; whenever you play B, you win with probability p_B, where $p_B > p_A$. If your objective is to minimize the expected number of games you need to play to win two in a row, should you start with A or with B?
Hint: Let $E[N_i]$ denote the mean number of games needed if you initially play i. Derive an expression for $E[N_A]$ that involves $E[N_B]$; write down the equivalent expression for $E[N_B]$ and then subtract.

1°. Suppose starting with A.

If losing with first attempt, then after the first attempt, restart with B.

If winning with first attempt, we further condition on the second round.

Then, if winning with second attempt, stop; o.w., restart with A.

$$E[N_A] = (1-P_A)\,E[N_A \mid X_A^1 = 0] + P_A\,E[N_A \mid X_A^1 = 1]$$

$$= (1-P_A)(1+E[N_B]) + P_A(1+P_B+(1-P_B)(1+E[N_A]))$$

$$= 1-\cancel{P_A} + E[N_B] - P_A E[N_B] + \cancel{P_A} + \cancel{P_AP_B} + P_A - \cancel{P_AP_B} + P_A E[N_A] - P_A P_B E[N_A]$$

$$= 1+P_A+(1-P_A)E[N_B] + P_A(1-P_B)E[N_A] \qquad ①$$

2°. Suppose starting with B.

$$E[N_B] = (1-P_B)\,E[N_B \mid X_B^1 = 0] + P_B\,E[N_B \mid X_B^1 = 1]$$

$$= 1+P_B+(1-P_B)E[N_A] + P_B(1-P_A)E[N_B] \qquad ②$$

According to ①②,

$$E[N_A] = \frac{1+P_A}{1-P_A+P_AP_B} + \frac{1-P_A}{1-P_A+P_AP_B}\cdot\frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)}, \quad E[N_B] = \frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)}$$

subtract $E[N_A]$, $E[N_B]$:

$$E[N_A] - E[N_B] = \frac{1+P_A}{1-P_A+P_AP_B} + \frac{1-P_A}{1-P_A+P_AP_B}\cdot\frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)} - \frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)}$$

$$= \frac{1+P_A}{1-P_A+P_AP_B} + \left(\frac{1-P_A}{1-P_A+P_AP_B} - 1\right)\cdot\frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)}$$

$$= \frac{1+P_A}{1-P_A+P_AP_B} - \frac{P_AP_B}{1-P_A+P_AP_B}\cdot\frac{2-P_AP_B+P_A(P_B)^2}{P_AP_B(2-P_A-P_B+P_AP_B)}$$

$$= \frac{1}{1-P_A+P_AP_B}\left(\frac{(1+P_A)P_AP_B(2-P_A-P_B+P_AP_B) - P_AP_B(2-P_AP_B+P_A(P_B)^2)}{P_AP_B(2-P_A-P_B+P_AP_B)}\right)$$

$$= \frac{1}{1-P_A+P_AP_B}\cdot\frac{P_AP_B(P_A-P_B)(1-P_A+P_AP_B)}{P_AP_B(2-P_A-P_B+P_AP_B)}$$

since $P_A, P_B \in [0,1]$ and $P_B > P_A$,

$1-P_A+P_AP_B > 0$, $P_AP_B(2-P_A-P_B+P_AP_B) > 0$, but $P_AP_B(P_A-P_B)(1-P_A+P_AP_B) < 0$.

Hence, $E[N_A] - E[N_B] < 0 \Rightarrow E[N_A] < E[N_B]$.

Therefore, to minimize # of game you need to play to win two in a row, you should start with A.