Probability and Mathematical Statistics: Homework #11

Due on September 18, 2022 at 11:59am

0

Name: **Zhu Zhelin**
Student ID: 2021533077

Problem 1 (mention the source of question, *e.g.*, BH CH0 #1)

(a) using the inverse function,we can get that $x = \frac{y-b}{a}$,predicting X from Y is also linear one,but with inverse function.

(b) $Cov(X, Y - cX) = 0, Cov(X, Y) - cCov(X, X) = \rho - cVar(X) = \rho - c = 0$, so c=$\rho$ then $V = Y - \rho X$ since the linear combination of bivariate Normal is also normal,it is a normal r.v.

(c) The same as b starting that $Cov(Y, X - dY) = 0, Cov(Y, X) - dCov(Y, Y) = 0$ $d = \rho$ W is $X - \rho Y$

(d) $E(Y|X) = E(cX + V|X) = \rho X + E(V|X) = \rho X + 0 = \rho X$ $E(X|Y) = E(dY + W|Y) = \rho Y$

(e) the linear prediction from X to Y is inverse from Y to X,while the Corr(X,Y) indicate the relationship between X and Y,by symmetry $Corr(X, Y) = Corr(Y, X) = \rho$,if it is close to 1,which means it has a positive relation between X and Y and to get information Y,if Corr is close to 1 ,we can put more weight on X to predict Y,the same is true from Y to predict X ,it makes sense when $\rho = 0$ which means the two are independent,we can't get information from X,if we use reverse,the prediction will be very large which makes no sense.

Problem 2 (BH CH0 #2)

(a) since for each j we randomly choose an r.v Xi,and Xi has random values,it is same as choose a random normal variable X_i,and so E(X_j*)=E(X_i)=μ $Var(X_j*) = \sigma^2$

(b) To calculate $E(\bar{X}*|X_1, X_2, \cdots, X_n) = \frac{X_1*+X_2*+*\cdots+X_n*}{n} = E(X_1*|X_1, X_2, X_3 \cdots X_n) = \frac{X_1+X_2+\cdots+X_n}{n}$, $Var(\bar{X^*}|X_1, X_2, \cdots X_n) = \frac{Var(X_1*|X_1,X_2,X_3\cdots X_n)}{n} = \sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n^2}$

(c) using $E(E(\bar{X}*|X_1, X_2, \cdots, X_n)) = E(\bar{X})$,we can know that $E(\bar{X}*) = E(\bar{X}) = \mu$,while $Var(\bar{X}*) = E(Var(\bar{X}*|X_1, X_2, \cdots, X_n)) + Var(E(\bar{X}*|X_1, X_2, \cdots, X_n)) = E(\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n^2}) + Var(\bar{X}) = \frac{(n-1)\sigma^2}{n^2} + \frac{\sigma^2}{n} = \frac{(2n-1)\sigma^2}{n^2}$

(d) we can know that X_j*'s are not independent ,furthermore,they are positively related,if for some j X_j* is large we can know that ,some sample of X_j is very large,since we can choose multiple times,the $\bar{X}*$ will tend to be large,if we can induce that X_j*'s are positively related,$Corr(X_i*, X_j*)$ is larger than zero so $Var(\frac{X_1*+X_2*+\cdots}{n}) = \frac{Var(X_1*)}{n} + \sum Cov(X_i*/n, X_j*/n)$ larger than $Var(\bar{X}) = \frac{Var(X_1)}{n}$ since the $Var(X_1*) = Var(X_1)$

Problem 3 (BH CH0 #3)

(a) (i) first to show $E(Y - L(Y|X)X) = 0$

$$\begin{aligned} &E(YX) - E(L(Y|X)X) = E(YX) - E(E(Y)X + \frac{Cov(X,Y)}{Var(X)}(X^2 - E(X)X))) \\ &E(YX) - E(X)E(Y) - Cov(X,Y) \end{aligned} \tag{1}$$

(ii) next to show $E(Y - L(Y|X)) = 0$,$E(Y) - E(E(Y) + \frac{Cov(X,Y)}{Var(X)}(X - E(X))) = E(Y) - E(E(Y)) = 0$

(iii) then to show $Y - L(Y|X)$ is uncorrelated to X,$Cov(Y - L(Y|X), X) = E((Y - L(Y|X))X) - E(Y - L(Y|X)X) = 0$ since $Y - L(Y|X)$ is linear combination of X and Y then add constant,it is also mvn, since the Cov is 0 they are independent.

(iv) thus $Y - L(Y|X)$ is independent of $H(X)$ given any H

(v) thus Y-$L(Y|X)$ is uncorrelated with $H(X)$

(vi) $Cov(Y - L(Y|X), H(X))$=$E((Y - L(Y|X))H(X)) - E(Y - L(Y|X))E(H(X)) = 0$ thus $E((Y - L(Y|X))(H(X))) = 0$,

(vii) thus $Y - L(Y|X)$ is orthgonal to H(X)

(viii) thus $L(Y|X)$ is a projection ,due to uniqueness of projection,it is equal to $E(Y|X)$

(b) $MMSE = E(\Theta|X)$ $\Theta|X = x \sim Beta(1 + x, n - x + 1)$,so E$(\Theta|X) = \frac{x+1}{n+2}$ (by posterior probability and the corresponding expectation)
then llse =$E(\Theta) + \frac{Cov(X,\Theta)(X - E(X))}{Var(X)}$ where E(Θ)=$\frac{1}{2}$,$Var(X) = E(Var(X|\Theta)) + Var(E(X|\Theta))$ since $X|\Theta$ is Bin(n,Θ),so it is equal to$E(n\Theta(1 - \Theta)) + Var(n\Theta) = \frac{n(n+2)}{12}$
E(X)=$E(E(X|\Theta)) = E(n\Theta) = \frac{n}{2}$,E$(\Theta X)$=$E(E(\Theta X|\Theta)) = E(n\Theta^2)$,Cov(X,$\Theta$)=E$(\Theta X)$-E$(\Theta)$E(X)=$\frac{n}{12}$
so llse is also $\frac{X+1}{n+2}$ which is same to mmse.

Problem 4 (BH CH0 #4)

(a) $p_k = q^k p, H(X) = \sum_{k=0}^{\infty} p_k \log_2(1/p_k) = \sum(-p\log_2 qkp^k - p\log_2(p)q^k) = -\sum(pq\frac{d(q^k)}{dq} + p\log_2 pq^k) = -\frac{q}{p}\log_2 q - \log_2 p$

(b) first lhs is$\sum P(X=Y|Y=x_k)P(Y=x_k) = \sum p_k^2$ the rhs is equal to $p_1^{p_1}p_2^{p_2}\cdots$
$log(X)$ is concave,using jesson we just need to prove $log(\sum p_k^2) \leq \sum p_k log p_k$,$log(\sum p_k^2) = log(\sum p_k \cdot p_k)$ where $\sum p_k = 1$ so using jesson ,we can easily get this true,since it is concave ,lhs$\geq$ rhs

Problem 5 (BH CH0 #5)

(a) let N_t be the t time all toys arrive with $\lambda = 1$ rate we can know that $N_t^j|N_t = n \sim Bin(n, p_j)$ it is similiar to chicken egg hatching problems,using this model we can easily know for each toys ,they are independent and $\sim Pois(p_j t)$ add them totally we can get $N_t \sim Pois(t)$

(b) we can easily get the T=max(Y$_1, \cdots$), since getting all the toys is the same to wait for the last toys,then we can get that the $Y_j \sim Expo(p_j)$ derivate from the $Pois(p_j t)$ processes,it is same to ask the first arrival time of each possion processes,to calculate $E(T)$,I want to first calculate the PDF of T$P(T > t) = 1 - P(T \leq t) = P(Y_1 \leq t, Y_2 \leq t, Y_3 \leq t, \cdots) = 1 - \prod_{i=1}^{n}(1 - e^{-p_i t})$ $E(T) = \int_0^\infty P(T > t)dt = \int_0^\infty (1 - \prod_{i=1}^{n}(1 - e^{-p_i t}))dt$

(c) since N represent the total toys' arrival time X_i represent the interval time between two toys,so add the interval time up until you collect each types of toys are equal to wait for the last type of toys be collected.$E(T) = E(E(X_1 + X_2 + \cdots X_N|N = n)) = E(NE(X_1|N)) = E(E(T|N)) = E(N)$

Problem 6 (BH CH0#6)

1. I will first solve general problem A=hu ,B=sheng ,C=wei notice that q and r is always same ,so $E(AABC) = E(AABC|1_{th}A)p + 2qE(AABC|1_{th}B) = p(E(ABC)+1) + 2q(E(AAC)+1)$ then continue condition,we can get the final Expectation is $p(\frac{(3+2q)p}{2q} + \frac{2p^2+2q^2+2p^2q+2q^2p+2pq}{p(1-q)} + 1) + 2q(\frac{p^2+q^2+p^2q+pq^2+pq}{q(1-q)^2} + \frac{q}{1-q} + \frac{(2+p)q}{p(1-q)} + 1)$ when p=q=r,$7\frac{1}{3}$ then B is the answer
2. to use the function we can get C min ≈ 6.9