Homework #3 CS 596 Yan Cui yj369

Problem 1:

Since we have:

$f(x)\,dx = P(X=x \le X+dx) = P(B)$, $(A_i \cap A_j) = \phi$, $\bigcup_{i=1}^{k} A_i = (H)$

Consider total probability:

$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_k)P(A_k)$

$\Rightarrow f(x)\,dx = f_1(x)dx \cdot p_1 + \cdots + f_k(x)dx \cdot p_k$

$\Rightarrow f(x) = p_1 f_1(x) + \cdots + p_k f_k(x)$

a). $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$, $p_1 = 99\%$

$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(0, \delta^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$, $p_2 = 1\%$

for $x_1, x_2 \sim N(0,1)$ we have

$f_1(x_1) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_1^2}{2}}$, $f_2(x_2) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_2^2}{2}}$

$\Rightarrow f_1(x_1, x_2) = \frac{1}{(\sqrt{2\pi})^2} \cdot e^{-\frac{1}{2}(x_1^2+x_2^2)}$

The same way:

$$f_2(x_1, x_2) = \frac{1}{\delta(\delta\cdot\sqrt{2\pi})^2} \cdot e^{-\frac{1}{2\delta^2}(x_1^2+x_2^2)}$$

Hence

$$f_{x_1x_2}(x_1, x_2) = \frac{0.99 \cdot \exp\left(-\frac{x_1^2+x_2^2}{2}\right)}{2\pi} + \frac{0.01 \cdot \exp\left(-\frac{x_1^2+x_2^2}{2\delta^2}\right)}{\delta^2 \cdot 2\pi}$$

b): $f(x_1) = \int f_{x_1x_2}(x_1, x_2)\,dx_2 = \frac{0.99\exp\left(-\frac{x_1^2}{2}\right)}{2\pi} \int e^{-\frac{1}{2}x_2^2} dx_2 + \frac{0.01\exp\left(-\frac{x_1^2}{2\delta^2}\right)}{\delta^2 2\pi} \int e^{-\frac{x_2^2}{2\delta^2}} dx_2$

(See next Page)

Since we have:

$$\int_{-\infty}^{+\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}$$

So: $f(X_1) = \frac{0.99\cdot\exp\left(-\frac{X_1^2}{2}\right)}{2\pi}\cdot\sqrt{2\pi} + \frac{0.01\cdot\exp\left(-\frac{X_1^2}{2\delta^2}\right)}{\delta^2 2\pi}\cdot\sqrt{\delta^2\cdot 2\pi}$

$$= \frac{0.99\cdot\exp\left(-\frac{X_1^2}{2}\right)}{\sqrt{2\pi}} + \frac{0.01\cdot\exp\left(-\frac{X_1^2}{2\delta^2}\right)}{\sqrt{\delta^2\cdot 2\pi}}$$

the same way we have

$$f(X_2) = \frac{0.99\cdot\exp\left(-\frac{X_2^2}{2}\right)}{\sqrt{2\pi}} + \frac{0.01\cdot\exp\left(-\frac{X_2^2}{2\delta^2}\right)}{\sqrt{\delta^2\cdot 2\pi}}$$

Hence $f(X_1)\cdot f(X_2) \neq f_{X_1X_2}(X_1, X_2)$

So X_1 and X_2 are not independent.

c). let $X \sim U(-1,1)$

let $Y = X^2$

So it is obviously that X and Y are not independent.

And $Cov(X,Y) = 0$ So they are uncorrelated

Problem 2:

a). $S = |X+s|$ and s is deterministic

$$P(S \le z) = P(|x+s| \le z) = P(-z \le X+s \le z) = P(-z-s \le X \le z-s)$$

$$= F_X(z-s) - F_X(-z-s)$$

Since $z \ge 0$,

$$f_S(z) = \frac{dP(S \le z)}{dz} = f_X(z-s) - (-1)\cdot f_X(-z-s)$$

$$= f_X(z-s) + f_X(-z-s)$$

b). Use total probability, we have

$$f_S(z) = f_S(z \mid s=0) \cdot P(s=0) + f_S(z \mid s=1) \cdot P(s=1)$$

from (a) we can have:

$$f_S(z) = \left[f_X(z) + f_X(-z) \right] \cdot 0.2 + \left[f_X(z-1) + f_X(-z-1) \right] \cdot 0.8$$

c). $$P(s=0 \mid S=z) = \frac{P(S=z \mid s=0) \cdot P(s=0)}{P(S=z \mid s=0) \cdot P(s=0) + P(S=z \mid s=1) \cdot P(s=1)}$$

$$= \frac{P(S=z \mid s=0)}{P(S=z \mid s=0) + P(S=z \mid s=1)}$$

Problem 3:

a): For a specific sequence we have N_h heads and $N - N_h$ tails

So we know for a sequence S we have:

$$P(S) = p^{N_h} \cdot (1-p)^{N-N_h}$$

b): Assume for $N+1$ outcome, let the coin be a random variable X

So: $$p = P(X=H \mid S) = L(p \mid S) = p^{N_h} \cdot (1-p)^{N-N_h}$$

let $\hat{l} = \ln L$ and we have an indicator function

$$\mathbb{1}\{X=x\} = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

such that:

$\hat{p} = \underset{p}{\operatorname{argmax}}\, \hat{l}(p \mid s)$ assume $\nabla_p \hat{l} > 0$ we have $\frac{N_h}{p} - \frac{N-N_h}{1-p} = 0$

$$\Rightarrow \hat{p} = \frac{N_h}{N}$$

and $$\hat{p} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{x_i = x_0\}$$

c) Consider the Expectation. for some N, we have:

$$E[\hat{p}] = \frac{1}{N}\sum_{i=1}^{N} p = p$$

And from b). $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{X_i = x_0\}$

MSE:

$$MSE(\hat{p}) = E[(\hat{p}-p)^2] = E\left[\left(\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}\{X_i = x\} - p\right)^2\right]$$

$$= \frac{1}{N^2} E\left[\left(\sum_{i=1}^{N}\mathbb{1}\{X_i = x\}\right)^2\right] - 2\frac{2p}{N} E\left[\sum_{i=1}^{N}\mathbb{1}\{X_i = x\}\right] + p^2$$

Let variable $Y = \sum_{i=1}^{N}\mathbb{1}\{X_i = x\}$, $\hat{p} = \frac{Y}{N}$

For $Y = \sum_{i=1}^{N}\mathbb{1}\{X_i = x\}$ $\hat{p} = \frac{Y}{N}$

Y is Binomial distribution. such that, $Y \sim Bin(N, p)$

$E[Y] = N_b \cdot p$ and $E[Y^2] = N_b\, p\,((N_b-1)p+1)$

$$\Rightarrow MSE(\hat{p}) = \frac{1}{N^2} E[Y^2] - \frac{2p}{N} \cdot E[Y] + p^2$$

$$= \frac{Np((N-1)p+1)}{N^2} - \frac{2p}{N} \cdot Np + p^2$$

$$= \frac{Np^2 - p^2 + p}{N} - 2p^2 + p^2$$

$$= \frac{p - p^2}{N} = \frac{p(1-p)}{N}$$

such that when $N \to \infty$, we have $\frac{p(1-p)}{N} \to 0$

which is $MSE(\hat{p}) \to 0$, $\hat{p} \to p$

Problem 4:

a) By Induction and properties of Gaussian

$$X_2 = \alpha \cdot X_1 + W_1 \sim N\left(0, \frac{\alpha^2}{1-\alpha^2} + 1\right)$$

which is $X_2 \sim N\left(0, \frac{1}{1-\alpha^2}\right)$

suppose for $n = k$ we have:

$$X_k \sim N\left(0, \frac{1}{1-\alpha^2}\right) \quad \text{for } k \geq 2$$

So for $n = k+1$: we have

$$X_{k+1} = \alpha \cdot X_k + W_k \sim N\left(0, \frac{\alpha^2}{1-\alpha^2} + 1\right)$$

which is $X_{k+1} \sim N\left(0, \frac{1}{1-\alpha^2}\right)$

So X_n are Gaussian.

b) Since we know

$$f(X_1, X_2, \cdots, X_N) = f(X_1) \cdot f(X_2|X_1) \cdot \cdots f(X_N|X_{N-1})$$

And for each $n = 2, \cdots, N$

$$X_n \Big| X_1 = x_1, \cdots, X_{N-1} = x_{n-1} = x_{n-1} \overset{d}{=} (\alpha X_{n-1} + W_n) \Big| X_1 = x_1, \cdots, X_{n-1} = x_{n-1}$$

$$\overset{d}{=} \alpha x_{n-1} + W_n \sim N(\alpha X_{n-1}, 1)$$

Hence

$$f(X_1, \cdots, X_N) = f(X_1) \cdot \left(\frac{1}{\sqrt{2\pi}}\right)^{n-1} \cdot \exp\left\{-\frac{1}{2} \cdot \sum_{i=2}^{N} (X_i - \alpha X_{i-1})^2\right\}$$

$$= \frac{(1-\alpha^2)^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \cdot \exp\left\{-\frac{1}{2}\left[(1-\alpha^2)X_1^2 + (X_2 - \alpha X_1)^2 + \cdots + (X_n - \alpha X_{n-1})^2\right]\right\}$$

c). $\hat{\theta}_{MLE}(X) = \arg\max_{\theta} f(X_1, X_2, \cdots, X_N)$

$$\sum_{i=1}^{N} \nabla_\theta \log f(X_1, X_2, \cdots, X_N) = 0$$

That is:

$$\ln f(X_1, X_2, \cdots, X_N) = \frac{1}{2}\ln(1-\alpha^2) - \frac{n}{2}\ln(2\pi) - \frac{(1-\alpha^2)X_1^2 + (X_2 - \alpha X_1)^2 + \cdots + (X_N - \alpha X_{N-1})^2}{2}$$

Hence

$$\frac{\partial \ln f}{\partial \alpha} = \frac{1}{2}\cdot\frac{1}{1-\alpha^2}\cdot(-2\alpha) - \frac{-2\alpha X_1^2 - 2X_1(X_2 - \alpha X_1) - \cdots - 2X_{N-1}(X_N - \alpha X_{N-1})}{2}$$

$$= -\frac{\alpha}{1-\alpha^2} + \sum_{i=1}^{N-1} X_i \cdot X_{i+1} - \alpha \cdot \sum_{i=2}^{N-1} X_i^2$$

And $\frac{\partial \ln f}{\partial \alpha} = 0$, let $m = \sum_{i=1}^{N-1} X_i \cdot X_{i+1}$ $\quad n = \sum_{i=2}^{N-1} X_i^2$

So:

$$-\frac{\alpha}{1-\alpha^2} + m - n\cdot\alpha = 0 \Rightarrow \frac{\alpha}{1-\alpha^2} = m - n\alpha$$

$$\Rightarrow n\cdot\alpha^3 - m\alpha^2 - (n+1)\alpha + m = 0 \Rightarrow (n\alpha - m)(\alpha^2 - 1) - \alpha = 0$$

assume $a = n$, $b = -m$, $c = -(n+1)$, $d = m$

$$\Rightarrow a\cdot\alpha^3 + b\alpha^2 + c\alpha + d = 0 \quad \text{assume } T(\alpha) = 0$$

Since $a \geq 0$ and Δ of the formula have $\Delta > 0$ (I have calculated)

And for $\alpha = -1$, $T(-1) = -(-1) = 1 > 0$

for $\alpha = 1$, $T(-1) = -1 < 0$

So we for $|\alpha| < 1$ we have a $\hat{\alpha} \in (-1, 1)$ and only one $\hat{\alpha}$ such that $T(\hat{\alpha}) = 0$ and $|\hat{\alpha}| < 1$

Go back to $\ln f(X_1, X_2, \cdots, X_N)$, because $\frac{1}{2}\ln(1-\alpha)^2$ and $-\frac{(1-\alpha^2)X_1^2}{2}$ is less important compared to the other terms.

So we can simply simplify $\ln f$ to be:

$$\ln f = -\frac{n}{2}\ln(2\pi) - \frac{(X_2 - \alpha X_1)^2 + \cdots + (X_N - \alpha X_{N-1})^2}{2}$$

(Continuing Page 6)

$$\frac{\partial \ln f}{\partial \alpha} = -\frac{-2x_1(x_2-\alpha x_1) - \cdots - 2x_{N-1}(x_N - \alpha x_{N-1})}{2} = 0$$

$$\Rightarrow \quad \alpha(x_1 + \cdots + x_{N-1}) = x_1 x_2 + \cdots + x_{N-1} x_N$$

$$\Rightarrow \quad \hat{\alpha} = \frac{\sum_{i=1}^{N-1} x_i x_{i+1}}{\sum_{i=1}^{N-1} x_i^2}$$

P.S. Further Thoughts : For $a\alpha^3 + b\alpha^2 + c\alpha + d = 0$

$a = n$, $b = -m$, $c = -(n+1)$, $d = m$

$$m = \sum_{i=1}^{N-1} x_i \cdot x_{i+1} \qquad n = \sum_{i=2}^{N-1} x_i^2$$

we have:

$$\sum_{i=1}^{N-1} x_i x_{i+1} - \sum_{i=2}^{N-1} \left[x_i \cdot (\alpha x_i) \right] = \frac{\alpha}{1-\alpha^2}$$

Since x_i and w_i are independent for all $i = 1, 2, \cdots, N$

So I think it can be made as: $\alpha x_i = x_{i+1} - w_{i+1}$

$$\Rightarrow \sum_{i=1}^{N-1} x_i x_{i+1} - \sum_{i=2}^{N-1} x_i x_{i+1} + \alpha \sum_{i=2}^{N-1} x_i w_{i+1} = \frac{\alpha}{1-\alpha^2}$$

$$\Rightarrow \left(x_1 x_2 + \sum_{i=2}^{N-1} x_i w_{i+1} \right) \cdot (1-\alpha^2) = \alpha$$

Assume $t = x_1 x_2 + \sum_{i=2}^{N-1} x_i w_{i+1}$

$$\Rightarrow \quad t \cdot (1-\alpha^2) = \alpha \quad \Rightarrow \quad \alpha = \frac{-1 \pm \sqrt{1+4t^2}}{2t} \quad \text{And } |\alpha| < 1$$

$$\Rightarrow \quad \hat{\alpha} = \frac{-1 + \sqrt{1+4t^2}}{2t}$$

Problem 5

a) $H_0: X \sim N(\mu_0, \Sigma_0)$

$H_1: X \sim N(\mu_1, \Sigma_1)$

$\Rightarrow$

f_0:

$$f_0(x) = (2\pi)^{-N/2} \cdot |\Sigma_0|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0)\right\}$$

$$f_1(x) = (2\pi)^{-\frac{N}{2}} \cdot |\Sigma_1|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(x-\mu_1)^T \cdot \Sigma_1^{-1}(x-\mu_1)\right\}$$

By Bayesian Approach: $P(H_0) = P(H_1) = 0.5$ we have

$C_{01} = C_{10} = 1$, $C_{00} = C_{11} = 0$

$$\lambda(x) = \frac{f_1(x)}{f_0(x)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(C_{10}-C_{00})P(H_0)}{(C_{01}-C_{00})\cdot P(H_1)} = 1$$

$\Rightarrow$ LRT: $f_1(x) > f_0(x)$ Decide H_1

$f_1(x) < f_0(x)$ Decide H_0

$f_1(x) = f_0(x)$ Either is OK

Then:

$$\ln(\lambda(x)) = \ln\left(\frac{f_1}{f_0}\right) = -\frac{1}{2}\left(\ln|\Sigma_1| - \ln|\Sigma_0|\right) - \frac{1}{2}\left[(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) - (x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0)\right]$$

$$\underset{H_0}{\overset{H_1}{\gtrless}} \ln 1 = 0$$

b). if $\Sigma_0 = \Sigma_1$, $\Rightarrow |\Sigma_0| = |\Sigma_1|$ ~~So we have~~ and $\Sigma_0^T = \Sigma_0, (\Sigma_0^{-1})^T = \Sigma_0^{-1}$

Suppose $(x-\mu_0)^T = (a_1\ a_2\ a_3 \cdots a_N)$

$(x-\mu_1)^T = (b_1\ b_1\ b_3 \cdots b_N)$

~~$\Sigma_0^{-1} = [C_1\ C_2 \cdots C_N]$, C_i is column vector of column i in Σ_0^{-1}~~

So: we have:

$$(x-\mu_1)^T \Sigma_0^{-1}(x-\mu_1) - (x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0) \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

c) if $\mu_0 = \mu_1$ we have:

$$(x-\mu_1)^T \Sigma_0^{-1}(x-\mu_1) - (x-\mu_1)\Sigma_0^{-1}(x-\mu_1) = 0.$$

So we know either H_0 or H_1 is OK

It makes sense because $P(H_0) = P(H_1) = 0.5$

That is $P(H_0) + P(H_1) = 1$ which is we have definitely known that X is in a situation that:

$$X \sim N(\mu_0, \Sigma_0)$$