University of Michigan Fall 2021

EECS501: Solution to Homework 8

1. **State TRUE or FALSE by giving reasons**

   (a) TRUE. Let us first find the joint PDF of Y_2 and Y_3: for $0 \le y_2 \le y_3$,

   $$f_{Y_2,Y_3}(y_2,y_3) = \int_{y_1=0}^{y_2} f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3)dy_1 = y_2 e^{-y_3}.$$

   We see that for $0 \le y_1 \le y_2$, we have

   $$f_{Y_1|Y_2,Y_3}(y_1|y_2,y_3) = \frac{f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3)}{f_{Y_2,Y_3}(y_2,y_3)} = \frac{1}{y_2}.$$

   Hence $E(Y_1|Y_2 = y_2, Y_3 = y_3) = y_2/2$.

   (b) FALSE. We have $E(X_i) = 0$, and

   $$\mathrm{Var}(X_i) = E(X_i^2) = \frac{1}{2}.$$

   CDF of Z_n converges to that of Gaussian with zero mean and variance $\frac{1}{2}$.

2. **Estimation of Gaussian Vector**

   $\underline{X}$ and $\underline{N}$ are independent. We need to show that they are jointly Gaussian by showing that $\underline{\alpha}^T\underline{X} + \underline{\beta}^T\underline{N}$ is Gaussian. Since each X and N are jointly Gaussian vectors and they are independent, we know $\underline{\alpha}^T\underline{X}$ and $\underline{\beta}^T\underline{N}$ are independent Gaussian random variables for any $\underline{\alpha}$ and $\underline{\beta}$. Therefore, their sum is also a Gaussian random variable. $\Rightarrow$ $\underline{X}$ and $\underline{N}$ are jointly Gaussian. Therefore, the MMSE of X from Y is equal to the LMMSE of X from Y which is given as follows. $\Rightarrow \begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix}$ is jointly Gaussian.

   $$R_Y = E(YY^T) = E[(HX+N)(HX+N)^T] = H\Sigma_X H^T + \Sigma_N$$

   $$R_{XY} = E(XY^T) = E[X(HX+N)^T] = R_X H^T + \underline{0} = R_X H^T$$

   $$E(\underline{X}|\underline{Y}) = R_{XY}R_Y^{-1}\underline{Y} = R_X H^T [H\Sigma_X H^T + \Sigma_N]^{-1}\underline{Y}$$

3. **Chernoff Inequality**

   (a) The Chernoff inequality is as follows.

   $$P(Z \ge x) \le e^{-n\sup_{\theta>0}(\theta x - \ln M_{X_i}(\theta))}$$

Since we have $e^{-n\sup_{\theta>0}(\theta x-\ln M_{X_i}(\theta))} \leq e^{-n(\theta x-\ln M_{X_i}(\theta))}$, $\forall\theta > 0$, we can write for all $\theta > 0$

$$P(Z \geq x) \leq e^{-n\theta x}e^{n\ln M_{X_i}(\theta)} = (\frac{M_{X_i}(\theta)}{e^{\theta x}})^n$$

We have $M_{X_i}(\theta) = 1 - p + pe^{\theta}$.

$$P(Z \geq p+\delta) \leq (\frac{1-p+pe^{\theta}}{e^{\theta(p+\delta)}})^n$$

Taking the derivative and setting it to zero:

$$\Rightarrow e^{\theta} = \frac{(1-p)(p+\delta)}{p(1-p-\delta)} \Rightarrow \frac{1-p+pe^{\theta}}{e^{\theta(p+\delta)}} = exp(-D(p+\delta||p))$$

$$\Rightarrow P(Z \geq p+\delta) \leq e^{-nD(p+\delta||p)}$$

(b) We have

$$P(Z \leq p-\delta) = P(-Z \geq \delta-p) \leq (\frac{M_{-X_i}(\theta)}{e^{\theta x}})^n = (\frac{1-p+pe^{-\theta}}{e^{\theta(\delta-p)}})^n$$

Taking the derivative and setting it to zero:

$$\Rightarrow e^{-\theta} = \frac{(1-p)(p-\delta)}{p(1-p+\delta)}$$

$$\Rightarrow P(Z \leq p-\delta) \leq e^{-nD(p-\delta||p)}$$

$$P(|Z-p| > \delta) \leq e^{-nD(p+\delta||p)} + e^{-nD(p-\delta||p)} \leq 2e^{-n\min\{D(p+\delta||p),D(p-\delta||p)\}}$$

So $P(|Z-p| > \delta)$ decays exponentially in n.

4. **Concentration Bounds** Define X_i be the amount that the gambler wins at game i. We have $P(X_i = 1) = 0.3$, $P(X_i = 0) = 0.2$, and $P(X_i = -1) = 0.5$. We also define $Z = \sum_{i=1}^{400} X_i$. We want to calculate $P(Z > 0)$. We have $E[X_i] = -0.2$ and $Var[X_i] = E[X_i^2] - E[X_i]^2 = 0.8 - 0.04 = 0.76$.

$$E[Z] = \sum_{i=1}^{400} E[X_i] = 400 \times (-0.2) = -80$$

$$Var[Z] = \sum_{i=1}^{400} Var[X_i] = 400 \times (0.76) = 304$$

Chebyshev:

$$P(Z > 0) = P(Z-(-80) > 80) \leq P(|Z-(-80)| > 80) \leq \frac{Var[Z]}{80^2} = 0.0475$$

Central Limit Theorem:

Define

$$\hat{Z} = \frac{1}{\sqrt{400}} \frac{\sum_{i=1}^{400}(X_i - E[X_i])}{\sqrt{Var[X_i]}} = \frac{Z - E[Z]}{\sqrt{Var[Z]}} \approx \mathcal{N}(0,1)$$

We write

$$P(Z > 0) = P(Z - (-80) > 80) = P(\frac{Z + 80}{\sqrt{304}} > \frac{80}{\sqrt{304}}) \approx 1 - \Phi(4.5883) = 2.2342 \times 10^{-6} \approx 0$$

Chernoff Bound:

$$P(Z > 0) \leq e^{-400 \sup_{\theta > 0} -\Lambda(\theta)}$$

We have $M_{X_i}(\theta) = E[e^{\theta X_i}] = 0.3e^{\theta} + 0.5e^{-\theta} + 0.2$, and therefore,

$$\Lambda(\theta) = \ln(0.3e^{\theta} + 0.5e^{-\theta} + 0.2)$$

To find the supremum, we take the derivative of $\Lambda(\theta)$ and set it to 0. We have

$$\Lambda'(\theta^*) = \frac{0.3e^{\theta^*} - 0.5e^{-\theta^*}}{0.3e^{\theta^*} + 0.5e^{-\theta^*} + 0.2} = 0 \Rightarrow e^{\theta^*} = \sqrt{5/3}$$

Therefore, we have

$$P(Z > 0) \leq e^{400\Lambda(\theta^*)} = e^{400\Lambda(\theta^*)} = 3.3882 \times 10^{-5}$$

5. **Wick's theorem (also Isserli's theorem)**

The number of all distinct ways of partitioning the $2n$ random variables into n pairs of (X_i, X_j) is $\frac{(2n)!}{n!2^n}$. The reason is that we first order $2n$ variables and the pairs would be determined based on the order (the first two together, the second two together, ...). The number of ways to order $2n$ items is $(2n)!$. Then we divide this number by the number of ways each partition is over counted. To change the order of items but have the same partition, the items in each pair can change their positions (2^n ways), and the position of pairs can change with each other ($n!$ ways)).

Using the above result, we have

$$E[X_1 X_3^2 X_4] = E[X_1 X_4]E[X_3^2] + 2E[X_1 X_3]E[X_3 X_4] = C_{14}C_{33} + 2C_{13}C_{34}$$

$$E[X_1^2 X_2^2] = E[X_1^2]E[X_2^2] + 2E[X_1 X_2]^2 = C_{11}C_{22} + 2C_{12}^2$$

$$E[X_1^6] = 15E[X_1^2]^3 = 15C_{11}^3$$

6. **LMMSE** We have

$$\hat{E}[X|Y] = E[X] + Cov(X,Y)/Var(Y)(Y - E(Y))$$

Since $X \sim exponential(\lambda)$, we have $E[X] = \frac{1}{\lambda}$. We also have $E[Y] = E[X] + E[N] = E[X]$. We can write

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[(X - E[X])(X + N - E[X])] = Var(X) = \frac{1}{\lambda^2}$$

$$Var(Y) = E[(Y - E[X])^2] = E[(X + N - E[X])(X + N - E[X])] = Var(X) + Var(N) = \frac{1}{\lambda^2} + \sigma^2$$

Therefore, we have

$$\hat{E}[X|Y] = \frac{1}{\lambda} + \frac{\frac{1}{\lambda^2}}{(\frac{1}{\lambda^2} + \sigma^2)}(Y - \frac{1}{\lambda}) = \frac{1}{\lambda} + \frac{\frac{1}{\lambda^2}}{(\frac{1}{\lambda^2} + \sigma^2)}(Y - \frac{1}{\lambda}) = \frac{1}{(1 + \sigma^2\lambda^2)}Y + \frac{\sigma^2\lambda}{(1 + \sigma^2\lambda^2)}$$

For LMMSE error, $E[(X - \hat{E}[X|Y])^2]$, we use orthogonality principle, which is $E[(X - \hat{E}[X|Y])\hat{E}[X|Y]] = 0$.

$$\begin{aligned}
E[(X - \hat{E}[X|Y])^2] &= E[(X - \hat{E}[X|Y])X] = E[X^2] - E[\hat{E}[X|Y]X] \\
&= \frac{2}{\lambda^2} - \frac{1}{(1 + \sigma^2\lambda^2)}E[XY] - \frac{\sigma^2\lambda}{(1 + \sigma^2\lambda^2)}E[X] \\
&= \frac{2}{\lambda^2} - \frac{1}{(1 + \sigma^2\lambda^2)}E[X(X + N)] - \frac{\sigma^2\lambda}{(1 + \sigma^2\lambda^2)}\frac{1}{\lambda} \\
&= \frac{2}{\lambda^2} - \frac{1}{(1 + \sigma^2\lambda^2)}\frac{2}{\lambda^2} - \frac{\sigma^2}{(1 + \sigma^2\lambda^2)} = \frac{\sigma^2}{1 + \lambda^2\sigma^2}
\end{aligned}$$