Solutions to Final Exam

(Total: 200 points)

There are 4 problems. The first 3 problems have 4 parts each, while the last problem has 8 parts. Each part is uniformly worth 10 points.

Your answer should be as clear and readable as possible. In particular, if the answer involves a pmf or pdf, make sure to identify the values or intervals for which the pmf or pdf is nonzero.

1. *Order statistics.*
   Let X_1, X_2, X_3 be independent and uniformly drawn from the interval $[0,1]$. Let Y_1 be the smallest of X_1, X_2, X_3, let Y_2 be the median (second smallest) of X_1, X_2, X_3, and let Y_3 be the largest of X_1, X_2, X_3. For example, if $X_1 = .3, X_2 = .1, X_3 = .7$, then $Y_1 = .1, Y_2 = .3, Y_3 = .7$. The random variables Y_1, Y_2, Y_3 are called the *order statistics* of X_1, X_2, X_3.

   (a) What is the probability $\mathrm{P}\{X_1 \le X_2 \le X_3\}$?

   (b) Find the pdf of Y_1.

   (c) Find the pdf of Y_3.

   (d) (Difficult.) Find the pdf of Y_2.
   (Hint: $Y_2 \le y$ if and only if at least two among X_1, X_2, X_3 are $\le y$.)

   Solution:

   (a) By symmetry, $\mathrm{P}\{X_i \le X_j \le X_k\}$ should be identical for all $i \ne j \ne k$. Since there are $3! = 6$ such (i,j,k), the probability should be $1/6$.

   (b) We have $\mathrm{P}\{Y_1 > y\} = \mathrm{P}\{X_1, X_2, X_3 > y\} = \mathrm{P}\{X_1 > y\}\mathrm{P}\{X_2 > y\}\mathrm{P}\{X_3 > y\} = (1-y)^3$. Hence, $f_{Y_1}(y) = \frac{d}{dy}(1-(1-y)^3) = 3(1-y)^2$ for $0 \le y \le 1$.

   (c) We can use the similar steps to part (b) to find $f_{Y_3}(y) = 3y^2,\ 0 \le y \le 1$. Alternatively, we can see that by symmetry $1 - Y_3$ and Y_1 should have the same pdf, which gives the same answer.

   (d) The event $\{Y_2 \le y\}$ can be expressed as the union of following mutually exclusive events

$$\begin{aligned}\{Y_2 \le y\} = \{X_1, X_2 \le y, X_3 > y\} &\cup \{X_2, X_3 \le y, X_1 > y\} \\ &\cup \{X_3, X_1 \le y, X_2 > y\} \cup \{X_1, X_2, X_3 \le y\}.\end{aligned}$$

By symmetry $P\{X_1, X_2 \le y, X_3 > y\} = P\{X_2, X_3 \le y, X_1 > y\} = P\{X_3, X_1 \le y, X_2 > y\} = y^2(1-y)$. Hence, $\mathrm{P}\{Y_2 \le y\} = 3y^2(1-y) + y^3 = 3y^2 - 2y^3$. By taking the derivative, we have $f_{Y_2}(y) = 6y(1-y)$ for $0 \le y \le 1$. This distribution is known as the Beta$(2,2)$ distribution.

2. *Fair coins.*
We are given two coins: Coin 1 with bias (=probability of heads) 1/2 and Coin 2 with random bias $P \sim \mathrm{Unif}[0,1]$. We pick one at random and flip it three times independently. The value of the bias does not change during the sequence of tosses. Let X be the number of heads.

(a) Find the conditional pmf of X given that Coin 1 is selected.

(b) Find the conditional pmf of X given that Coin 2 is selected.

(c) Find the optimal decision rule $D(x) \in \{1,2\}$ for deciding which coin is flipped such that the probability of decision error is minimized.

(d) Find the associated probability of error.

Solution: Let $\Theta \in \{1,2\}$ denote the coin we select.

(a) Note that when we select Coin 1, the number of heads follow the binomial distribution: $X|\{\Theta = 1\} \sim \mathrm{B}(3, 1/2)$. Hence, $p_{X|\Theta}(0|1) = p_{X|\Theta}(3|1) = 1/8$, $p_{X|\Theta}(1|1) = p_{X|\Theta}(2|1) = 3/8$.

(b) With Coin 2, $p_{X|\Theta}(0|2) = \mathrm{P}\{X = 0|\Theta = 2\} = \int_0^1 (1-p)^3 dp = 1/4$. Similarly, $p_{X|\Theta}(1|2) = \int_0^1 3p(1-p)^2 dp = 1/4$, $p_{X|\Theta}(2|2) = p_{X|\Theta}(3|2) = 1/4$.

(c) Since $p_\Theta(1) = p_\Theta(2) = 1/2$, the optimal decision rule compares $p_{X|\Theta}(x|1)$ vs. $p_{X|\Theta}(x|2)$. Therefore, $D(x) = 2$ if $x = 0, 3$, and $D(x) = 1$ if $x = 1, 2$.

(d) We have

$$\begin{aligned}\mathrm{P}\{\Theta \neq \hat{\Theta}\} &= \frac{1}{2}\mathrm{P}\{X = 0 \text{ or } 3 \,|\, \Theta = 1\} + \frac{1}{2}\mathrm{P}\{X = 1 \text{ or } 2 \,|\, \Theta = 2\} \\ &= \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{2} = \frac{3}{8}.\end{aligned}$$

3. *Estimation.*
Let $X \sim N(0, P)$ and $Z \sim N(0, N)$ are independent. Let $Y = X + Z$.

(a) Find the minimum mean squared error (MSE) estimator of X given Y.

(b) What is the associated MSE?

(c) Find the minimum MSE estimator of X^2 given Y.

(d) Find the minimum MSE *linear* estimator of X^2 given Y.
(Hint: You can use symmetry to find $E(X^3)$ rather easily.)

Solution:

(a) First note that X and Y are jointly Gaussian (why?). In particular, we have

$$X|\{Y=y\} \sim N\left(\frac{Py}{P+N}, \frac{PN}{P+N}\right).$$

(To get this, we can directly calculate the joint pdf or use the fact that the minimum MSE estimator is linear for jointly Gaussian random variables.) Hence, the best MSE estimator is $E(X|Y) = PY/(P+N)$.

(b) The resulting MSE is $E\mathrm{Var}(X|Y) = \mathrm{Var}(X|Y) = PN/(P+N)$.

(c) Since

$$\mathrm{Var}(X|Y) = E(X^2|Y) - \big[E(X|Y)\big]^2,$$

we have

$$E(X^2|Y) = \mathrm{Var}(X|Y) + \big[E(X|Y)\big]^2 = \frac{PN}{P+N} + \left(\frac{PY}{P+N}\right)^2,$$

which is the minimum MSE estimator of X^2 given Y.

(d) First note that $E(X^3) = 0$ because the Gaussian pdf is an even function, while x^3 is an odd function. Since $E(X^2Y) = E(X^3 + XZ) = E(X^3) = 0$, X^2 and Y are uncorellated. Therefore, the best linear estimator is $E(X^2) = P$.

4. *Autoregressive process.*
Let $X_0 \sim N(0,1)$ and $X_n = \alpha X_{n-1} + Z_n$ for $n \geq 1$, where $Z_1, Z_2, \ldots$ are i.i.d. $\sim N(0, 1-\alpha^2)$ and independent of X_0. Assume $-1 < \alpha < 1$.

(a) Find the mean and autocorrelation function of X_n.

(b) Is X_n a wide-sense stationary process?

(c) Is X_n a Gaussian process?

(d) Is X_n an independent-increment process?

(e) Given the observation $X_1, X_2, \ldots, X_k$, find the minimum MSE *linear* estimator of X_{k+1}. (Hint: You might first consider the best *nonlinear* estimator.)

(f) What is the associated MSE?

(g) Given the observation $X_1, X_2, \ldots, X_k$, find the minimum MSE *linear* estimator of X_{k+2}.

(h) What is the associated MSE?

Solution:

(a) We have $E(X_n) = \alpha E(X_{n-1}) + E(Z_n) = \alpha E(X_{n-1})$. Since $E(X_0) = 0$, we can recursively check that $E(X_n) = 0$ for all n.

Since X_{n-1} and Z_n are independent for each n, $R_X(n, n) = E(X_n^2) = \alpha^2 + (1 - \alpha^2) = 1$ for all n. Now note that we can represent X_{n+k} as

$$X_{n+k} = \alpha^k X_n + \alpha^{k-1} Z_{n+1} + \alpha^{k-2} Z_{n+2} + \cdots + \alpha Z_{n+k-1} + Z_{n+k}.$$

Hence,

$$E(X_n X_{n+k}) = \alpha^k E(X_n^2) = \alpha^k$$

for all n and $k \geq 0$. Therefore, $R_X(n, m) = \alpha^{|n-m|}$.

(b) The answer is yes. Since the mean is constant and the autocorrelation function $R_X(n, m)$ is a function only of $|n - m|$, X_n is wide-sense stationary. (In fact, it is strict-sense stationary.)

(c) The answer is yes. For every n, the random vector $(X_0, \ldots, X_n)$ is a linear transformation of the random vector $(X_0, Z_1, \ldots, Z_n)$, which is Gaussian. Hence, $(X_0, \ldots, X_n)$ is also Gaussian for every n. Therefore the random process X_n is Gaussian.

(d) The answer is no. Consider X_0 and $\Delta = X_1 - X_0 = (\alpha - 1)X_0 + Z_1$. Since we can easily check that X_0 and Δ are correlated, they are dependent as well.

(e) Since X_n is a Gaussian process, the best (nonlinear) estimator is linear. Put in another way, $E(X_{k+1}|X_1, \ldots, X_k)$ is the best linear estimator of X_{k+1} given $X_1, \ldots, X_k$. But

$$\begin{aligned} E(X_{k+1}|X_1, \ldots, X_k) &= E(\alpha X_k + Z_{k+1}|X_1, \ldots, X_k) \\ &= \alpha X_k + E(Z_{k+1}) = \alpha X_k. \end{aligned}$$

Hence the best linear estimator is αX_k.

(f) The associated MSE is $E(X_{k+1} - \alpha X_k)^2 = E(Z_{k+1}^2) = 1 - \alpha^2$.

(g) Since $X_{k+2} = \alpha^2 X_k + \alpha Z_{k+1} + Z_{k+2}$, following the same steps in (e), we have $E(X_{k+2}|X_1, \ldots, X_k) = \alpha^2 X_k$, which is the best (linear) estimator.

(h) The associated MSE is $E(X_{k+2} - \alpha^2 X_k)^2 = \alpha^2 E(Z_{k+1}^2) + E(Z_{k+2}^2) = (1 + \alpha^2)(1 - \alpha^2) = 1 - \alpha^4$.