Solutions to Practice Final Examination (Winter 2017)

There are 6 problems, each problem with multiple parts. Your answer should be as clear and readable as possible. Please justify any claim that you make.


(a) Find the joint pdf $f_{X,Z}(x,z)$ of $X$ and $Z$.
(b) Find the joint pdf $f_{Z,W}(z,w)$ of $Z$ and $W$.
(c) Find $E[Z|X]$.
(d) Find $E[X|Z]$.

Solution:

(a) For $z < x$, we have

$$F_{Z|X}(z|x) = P\{Z \leq z \mid X = x\} = 0.$$

For $0 \leq x \leq z$,

$$F_{Z|X}(z|x) = P\{Z \leq z \mid X = x\} = P\{X + Y \leq z \mid X = x\} = P\{Y \leq z - x \mid X = x\} \overset{(a)}{=} P\{Y \leq z - x\} = 1 - e^{-(z-x)},$$

where (a) follows from the independence of $X$ and $Y$. We therefore have

$$f_{Z|X}(z|x) = \begin{cases} e^{-(z-x)}, & \text{if } 0 \leq x \leq z \\
0, & \text{otherwise}. \end{cases}$$

Therefore,

$$f_{X,Z}(x,z) = \begin{cases} e^{-z}, & \text{if } 0 \leq x \leq z \\
0, & \text{otherwise}. \end{cases}$$
(b) From the previous part, we have, for \( 0 \leq x \leq z \),

\[
\begin{align*}
    f_{X|Z}(x|z) &= \frac{f_{X,Z}(x,z)}{f_Z(z)} \\
             &= \frac{f_{X,Z}(x,z)}{\int_{0}^{z} f_{X,Z}(x,z)dx} \\
             &= \frac{1}{z}.
\end{align*}
\]

Thus for \( z \geq 0 \), \( X \mid \{Z = z\} \sim \text{Unif}[0, z] \). We have \( W = X - Y = 2X - Z \). Therefore,

\[
F_{W|Z}(w|z) = P\{W \leq w \mid Z = z\} = P\{2X - Z \leq w \mid Z = z\} = P\{X \leq \frac{z + w}{2} \mid Z = z\} = \begin{cases} 
    0, & \text{if } w < -z \\
    \frac{z+w}{2z}, & \text{if } -z \leq w \leq z \\
    1, & \text{if } w > z.
\end{cases}
\]

Thus,

\[
f_{W|Z}(w|z) = \begin{cases} 
    \frac{1}{2z}, & \text{if } |w| \leq z \\
    0, & \text{otherwise},
\end{cases}
\]

which leads us to conclude that

\[
f_{Z,W}(z,w) = f_{W|Z}(w|z)f_Z(z) = \begin{cases} 
    \frac{1}{2}e^{-z}, & \text{if } |w| \leq z \\
    0, & \text{otherwise}.
\end{cases}
\]

(c) We have

\[
E[Z|X] = E[X + Y \mid X] = X + E[Y|X] = X + E[Y] = X + 1,
\]

where \( E[Y|X] = E[Y] \) since \( X \) and \( Y \) are independent.
(d) From part (b), we have $X \mid \{Z = z\} \sim \text{Unif}[0, z]$. Therefore,

$$E[X|Z] = \frac{Z}{2}.$$ 

2. **MMSE estimation (30 pts).** Let $X \sim \text{Exp}(1)$ and $Y = \min\{X, 1\}$.

(a) Find $E[Y]$.

(b) Find the estimate $\hat{X} = g(Y)$ of $X$ given $Y$ that minimizes the mean square error $E[(X - \hat{X})^2] = E[(X - g(Y))^2]$, and plot $g(y)$ as a function of $y$.

(c) Find the mean square error of the estimate found in part (b).

**Solution:**

(a) We have

$$E[Y] = E[\min\{X, 1\}]$$

$$= \int_0^\infty \min\{x, 1\} e^{-x} dx$$

$$= \int_0^1 xe^{-x} dx + \int_1^\infty e^{-x} dx$$

$$= -xe^{-x} - e^{-x} \bigg|_0^1 + e^{-1}$$

$$= 1 - e^{-1}.$$ 

(b) We have $g(y) = E[X \mid Y = y]$. For $y < 1$,

$$E[X \mid Y = y] = E[X \mid X = y] = y.$$ 

For $y = 1$, we have

$$E[X \mid Y = y] = E[X \mid X \geq 1]$$

$$\overset{(a)}{=} E[X] + 1$$

$$= 2,$$

where $(a)$ follows from the memorylessness property of the exponential distribution. Thus,

$$g(y) = \begin{cases} 
  y, & 0 \leq y < 1 \\
  2, & y = 1.
\end{cases}$$
The plot of $g(y)$ vs $y$ is shown in Fig. 1.

(c) For $0 \leq y < 1$, $\text{Var}(X \mid Y = y) = 0$. For $y = 1$,

$$\text{Var}(X \mid Y = y) = \text{Var}(X \mid X \geq 1)$$

$$(a) = \text{Var}(X)$$

$= 1,$

where the step $(a)$ follows from the memoryless property. We therefore have

$$\text{MSE} = E[\text{Var}(X \mid Y)]$$

$$= \text{Var}(X \mid Y = 1)P\{Y = 1\}$$

$$= e^{-1}.$$ 

3. Is the grass always greener on the other side? (30 pts). Let $X$ and $Y$ be two i.i.d. continuous nonnegative random variables with invertible common cdf $F$, i.e.,

$$P\{X \leq x\} = P\{Y \leq x\} = F(x).$$

(a) Find $P\{X > Y\}$ and $P\{X < Y\}$.

Suppose now that we observe the value of $X$ and make a decision on whether $X$ is larger or smaller than $Y$.

(b) Find the optimal decision rule $d(x)$ that minimizes the error probability. Your answer should be in terms of the common cdf $F$. 


(c) Find the probability of error for the decision rule found in part (b).

Solution:

(a) By symmetry, $P\{X > Y\} = P\{X < Y\} = 1/2$. Alternatively, let $f$ be the common pdf of $X$ and $Y$. Then

$$P\{X > Y\} = \int_0^{\infty} P\{X > Y \mid Y = y\} f(y) dy$$

$$= (a) \int_0^{\infty} P\{X > y\} f(y) dy$$

$$= \int_0^{\infty} (1 - F(y)) f(y) dy$$

$$= 1 - \int_0^{\infty} f(y) F(y) dy.$$

Here, $(a)$ follows from the independence of $X$ and $Y$. We now have, integrating by parts,

$$I := \int_0^{\infty} f(y) F(y) dy$$

$$= F(y)^2\bigg|_0^{\infty} - \int_0^{\infty} F(y) f(y) dy$$

$$= \lim_{y \to \infty} F(y)^2 - I$$

$$= 1 - I,$$

whence $I = 1/2$. Thus,

$$P\{X > Y\} = \frac{1}{2}.$$

By interchanging the roles of $X$ and $Y$, we conclude that

$$P\{X < Y\} = \frac{1}{2}.$$

Note: We can also compute $I$ by noting that

$$f(y) F(y) = \frac{1}{2} \frac{d}{dy} F(y)^2.$$
(b) Let us define a random variable $Z$ as

$$Z = \begin{cases} 
1, & \text{if } X > Y \\
0, & \text{if } X \leq Y. 
\end{cases}$$

Then, we have to find a decision rule $d(\cdot)$, such that $P\{d(X) \neq Z\}$ is minimized. We know that this should be the MAP decision rule. We have

$$p_{Z|X}(1|x) = P\{Z = 1 \mid X = x\} = P\{Y < X \mid X = x\} = P\{Y < x \mid X = x\} = F(x).$$

Therefore, $p_{Z|X}(0|x) = 1 - F(x)$, i.e., we should choose $d(x) = 1$ if $F(x) > 1 - F(x)$, i.e., if $x > F^{-1}(1/2)$ (which is the median of $X$ and is unique since $F$ is invertible). Thus the optimal decision rule is given by

$$d(x) = \begin{cases} 
1, & \text{if } x > F^{-1}(1/2) \\
0, & \text{if } x \leq F^{-1}(1/2). 
\end{cases}$$

In other words, we predict that $X$ is larger than $Y$ if the observed value of $X$ is larger than the median.

(c) We have

$$P\{d(X) \neq Z\} = P\{X > Y, X \leq F^{-1}(1/2)\} + P\{X < Y, X > F^{-1}(1/2)\}$$

$$= P\{Y < X \leq F^{-1}(1/2)\} + P\{F^{-1}(1/2) < X < Y\}$$

$$= \int_0^{F^{-1}(1/2)} \int_0^x f(x)f(y)dydx + \int_{F^{-1}(1/2)}^\infty \int_x^\infty f(x)f(y)dydx$$

$$= \int_0^{F^{-1}(1/2)} f(x)F(x)dx + \int_{F^{-1}(1/2)}^\infty f(x)(1 - F(x))dx$$

$$= \int_0^{F^{-1}(1/2)} f(x)F(x)dx - \int_{F^{-1}(1/2)}^\infty f(x)F(x)dx$$

$$= \frac{1}{2} + \int_0^{F^{-1}(1/2)} F(F^{-1}(1/2)) - \frac{1}{2} \left(1 - (F(F^{-1}(1/2)))^2\right)$$

$$= \frac{1}{2} + \frac{1}{2} \left(1 - \frac{3}{8}\right)$$

$$= \frac{1}{4}.$$
Here, (a) follows from the observation made at the end of part (a).

4. **Sampled Wiener process (60 pts).** Let \( \{W(t), t \geq 0\} \) be the standard Brownian motion. For \( n = 1, 2, \ldots, \) let

\[
X_n = n \cdot W \left( \frac{1}{n} \right).
\]

(a) Find the mean and autocorrelation functions of \( \{X_n\} \).
(b) Is \( \{X_n\} \) WSS? Justify your answer.
(c) Is \( \{X_n\} \) Markov? Justify your answer.
(d) Is \( \{X_n\} \) independent increment? Justify your answer.
(e) Is \( \{X_n\} \) Gaussian? Justify your answer.
(f) For \( n = 1, 2, \ldots, \) let \( S_n = X_n / n \). Find the limit

\[
\lim_{n \to \infty} S_n
\]

in probability.

**Solution:**

(a) We have

\[
E[X_n] = nE[W(1/n)] = 0.
\]

For \( m, n \in \mathbb{N} \) and \( m \geq n \), we have

\[
E[X_m X_n] = mnE[W(1/m)W(1/n)]
\]

\[
= mn \cdot \min \{1/m, 1/n\}
\]

\[
= mn \cdot \frac{1}{m}
\]

\[
= n.
\]

Thus in general,

\[
E[X_m X_n] = \min \{m, n\}.
\]

(b) No. Since the autocorrelation function is not time-invariant, \( \{X_n\} \) is not WSS.
(c) Yes. Clearly, \( \{X_n\} \) is a Gaussian process (see the solution to part (e)) with mean and autocorrelation functions as found in part (a). Therefore, for integers \( m_1 < m_2 \leq m_3 < m_4 \), we have

\[
\mathbb{E}[(X_{m_2} - X_{m_1})(X_{m_4} - X_{m_3})] = \mathbb{E}[X_{m_2}X_{m_4}] + \mathbb{E}[X_{m_1}X_{m_3}] - \mathbb{E}[X_{m_2}X_{m_3}] - \mathbb{E}[X_{m_1}X_{m_4}]
\]

\[
= \min\{m_2, m_4\} + \min\{m_1, m_3\} - \min\{m_2, m_3\} - \min\{m_1, m_4\}
\]

\[
= m_2 + m_1 - m_2 - m_1
\]

\[
= 0
\]

\[
= \mathbb{E}[X_{m_2} - X_{m_1}]\mathbb{E}[X_{m_4} - X_{m_3}].
\]

Therefore, since \( (X_{m_2} - X_{m_1}) \) and \( (X_{m_4} - X_{m_3}) \) are jointly Gaussian and uncorrelated, they are independent. Now, for positive integers \( n_1 < n_2 < \cdots < n_k \) for some \( k \), \( (X_{n_1}, X_{n_2} - X_{n_1, \ldots, X_{n_k} - X_{n_{k-1}}} \) being a linear transformation of a Gaussian random vector, is itself Gaussian. Moreover, from what we just showed, \( (X_{n_1}, X_{n_2} - X_{n_1}, \ldots, X_{n_k} - X_{n_{k-1}}) \) are pairwise independent. Therefore, they are all independent, which implies that \( \{X_n\} \) is independent-increment. This implies Markovity.

(d) Yes. See the solution to part (c).

(e) Yes. For integers \( n_1, n_2, \ldots, n_k \) for any \( k \), we have

\[
\begin{bmatrix}
X_{n_1} \\
X_{n_2} \\
\vdots \\
X_{n_k}
\end{bmatrix} =
\begin{bmatrix}
n_1 & 0 & \cdots & 0 \\
0 & n_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & n_k
\end{bmatrix}
\begin{bmatrix}
W(1/n_1) \\
W(1/n_2) \\
\vdots \\
W(1/n_k)
\end{bmatrix}.
\]

Thus, \( [X_{n_1} \cdots X_{n_k}]^T \), being a linear transformation of a Gaussian random vector, is itself a Gaussian random vector. Therefore, \( \{X_n\} \) is Gaussian.

(f) Recall that \( X_n \sim N(0, n) \), which implies that \( X_n/\sqrt{n} \sim N(0, 1) \). Therefore, for any fixed \( \epsilon > 0 \), we have

\[
P\{|S_n| > \epsilon\} = P\{|X_n| > n\epsilon\}
\]

\[
= P\left\{\left|\frac{X_n}{\sqrt{n}}\right| > \epsilon\sqrt{n}\right\}
\]

\[
= 2Q(\epsilon\sqrt{n})
\]

\[
\rightarrow 0,
\]
as \( n \to \infty \). Therefore, \( \lim_{n \to \infty} S_n = 0 \) in probability. Alternatively, note that \( W(0) = 0 \) and \( W(t) \) is continuous with probability 1. Therefore

\[
\lim_{n \to \infty} S_n = \lim_{n \to \infty} W\left(\frac{1}{n}\right) = W(0) = 0.
\]

5. Poisson process (40 pts). Let \( \{N(t), t \geq 0\} \) be a Poisson process with arrival rate \( \lambda > 0 \). Let \( s \leq t \).

(a) Find the conditional pmf of \( N(t) \) given \( N(s) \).

(b) Find \( E[N(t)|N(s)] \) and its pmf.

(c) Find the conditional pmf of \( N(s) \) given \( N(t) \).

(d) Find \( E[N(s)|N(t)] \) and its pmf.

Solution:

(a) Assume \( 0 \leq n_s \leq n_t \). By the independent increment property of the Poisson process, we would get

\[
P\{N(t) = n_t|N(s) = n_s\} = P\{N(t) - N(s) = n_t - n_s|N(s) = n_s\}
= P\{N(t) - N(s) = n_t - n_s\}
= e^{-\lambda(t-s)} \frac{\lambda(t-s)^{n_t-n_s}}{(n_t-n_s)!}
\]

for \( n_s = 0, 1, \ldots \) and \( n_t = n_s, n_s + 1, \ldots \). Thus,

\[
N(t)|\{N(s) = n_s\} \sim n_s + \text{Poisson}(\lambda(t-s)).
\]

(b) From part (a), it immediately follows that

\[
E[N(t)|N(s)] = N(s) + \lambda(t - s).
\]

Therefore, the pmf of \( E[N(t)|N(s)] \) is

\[
p_{E[N(t)|N(s)]}(x) = \begin{cases} 
  e^{-\lambda s} \frac{(\lambda s)^k}{k!} & \text{if } x = k + \lambda(t - s), \quad k = 0, 1, \ldots \\
  0 & \text{otherwise}
\end{cases}
\]
(c) From part (a), the joint pmf of $(N(t), N(s))$ for $0 \leq n_s \leq n_t$, is
\[
\begin{align*}
P\{N(t) = n_t, N(s) = n_s\} &= P\{N(s) = n_s\}P\{N(t) = n_t | N(s) = n_s\} \\
&= e^{-\lambda_s} \left(\frac{(\lambda s)^{n_s}}{n_s!}\right) e^{-\lambda(t-s)} \left(\frac{(\lambda(t-s))^{n_t-n_s}}{(n_t-n_s)!}\right) \\
&= e^{-\lambda t} \frac{n_t^{n_s}(t-s)^{n_t-n_s}}{n_s!(n_t-n_s)!}.
\end{align*}
\]
Therefore, the conditional pmf of $N(s)|\{N(t) = n_t\}$ is for $n_t \geq n_s \geq 0$
\[
\begin{align*}
P\{N(s) = n_s | N(t) = n_t\} &= \frac{P\{N(s) = n_s, N(t) = n_t\}}{P\{N(t) = n_t\}} \\
&= \left(\frac{e^{-\lambda t} n_t^{n_s}(t-s)^{n_t-n_s}}{n_s!(n_t-n_s)!}\right) \left(\frac{e^{-\lambda t} (\lambda t)^{n_t}}{n_t!}\right)^{-1} \\
&= \binom{n_t}{n_s} \left(\frac{s}{t}\right)^{n_s} \left(1 - \frac{s}{t}\right)^{n_t-n_s}.
\end{align*}
\]
Hence,
\[
N(s)|\{N(t) = n_t\} \sim \text{Binom}\left(n_t, \frac{s}{t}\right).
\]

(d) From part (c), it immediately follows that
\[
E[N(s)|N(t)] = \frac{s}{t} N(t),
\]
and its pmf is
\[
p_{E[N(s)|N(t)]}(x) = \begin{cases} 
  e^{-\lambda t} \left(\frac{(\lambda t)^k}{k!}\right) & \text{if } x = \frac{s}{t} k, \quad k = 0, 1, \ldots \\
  0 & \text{otherwise}
\end{cases}
\]

6. Hidden Markov process (60 pts). Let $X_0 \sim N(0, \sigma^2)$ and $X_n = \frac{1}{2} X_{n-1} + Z_n$ for $n \geq 1$, where $Z_1, Z_2, \ldots$ are i.i.d. $N(0, 1)$, independent of $X_0$. Let $Y_n = X_n + V_n$, where $V_n$ are i.i.d. $\sim N(0, 1)$, independent of $\{X_n\}$.

(a) Find the variance $\sigma^2$ such that $\{X_n\}$ and $\{Y_n\}$ are jointly WSS.

Under the value of $\sigma^2$ found in part (a), answer the following.

(b) Find $R_{Y}(n)$. 

10
(c) Find $R_{XY}(n)$.
(d) Find the MMSE estimate of $X_n$ given $Y_n$.
(e) Find the MMSE estimate of $X_n$ given $(Y_n, Y_{n-1})$.
(f) Find the MMSE estimate of $X_n$ given $(Y_n, Y_{n+1})$.

Solution:

(a) If $\{X_n\}$ is WSS, then $\text{Var}(X_n) = \text{Var}(X_0) = \sigma^2$ for all $n \geq 0$. From the recursive relation, we would get

$$\text{Var}(X_n) = \frac{1}{4}\text{Var}(X_{n-1}) + \text{Var}(Z_n),$$

which implies $\sigma^2 = \frac{4}{3}$.

(b) First, note that for $n \geq 0$,

$$X_{m+n} = \frac{1}{2}X_{m+n-1} + Z_{m+n}$$
$$= \frac{1}{4}X_{m+n-2} + \frac{1}{2}Z_{m+n-1} + Z_{m+n}$$
$$= \ldots$$
$$= \frac{1}{2^n}X_m + \frac{1}{2^{n-1}}Z_{m+1} + \cdots + \frac{1}{2}Z_{m+n-1} + Z_{m+n}.$$

Hence, it follows that

$$R_X(n) = E[X_{m+n}X_n] = 2^{-n}E[X_m^2] = \frac{4}{3}2^{-|n|}.$$

Now we can find the autocorrelation function of $\{Y_n\}$ easily.

$$R_Y(n) = E[Y_{m+n}Y_m]$$
$$= E[(X_{m+n} + V_{m+n})(X_m + V_m)]$$
$$= E[X_{m+n}X_m + X_{m+n}V_m + V_{m+n}X_m + V_{m+n}V_m]$$
$$= R_X(n) + \delta(n)$$
$$= \frac{4}{3}2^{-|n|} + \delta(n)$$

Here $\delta(n)$ denotes the Kronecker delta function, that is,

$$\delta(n) = \begin{cases} 
1 & \text{if } n = 0 \\
0 & \text{otherwise.}
\end{cases}$$
(c) The cross correlation function $R_{XY}(n)$ is
\[
R_{XY}(n) = E[X_{m+n}Y_m] \\
= E[X_{m+n}X_m + X_{m+n}V_m] \\
= R_X(n) = \frac{4}{3}2^{-|n|}.
\]

(d) Since $X_n$ and $Y_n$ are jointly Gaussian, we can find the conditional expectation $E[X_n|Y_n]$, which is the MMSE estimate of $X_n$ given $Y_n$, as follows:
\[
E[X_n|Y_n] = E[X_n] + \frac{\text{Cov}(X_n,Y_n)}{\text{Var}(Y_n)}(Y_n - E[Y_n]) \\
= \frac{R_{XY}(0)}{R_{Y}(0)}Y_n \\
= \frac{4}{7}Y_n.
\]

(e) As in part (d), the MMSE estimate of $X_n$ given $(Y_n, Y_{n-1})$ is
\[
E[X_n|Y_n, Y_{n-1}] = E[X_n] + \Sigma_{X_n, (Y_n, Y_{n-1})} \Sigma_{(Y_n, Y_{n-1})}^{-1} \left( \begin{bmatrix} Y_n \\ Y_{n-1} \end{bmatrix} - E \begin{bmatrix} Y_n \\ Y_{n-1} \end{bmatrix} \right) \\
= \begin{bmatrix} R_{XY}(0) & R_{XY}(1) \\ R_{Y}(0) & R_{Y}(1) \end{bmatrix}^{-1} \begin{bmatrix} R_{Y}(0) & R_{Y}(1) \\ R_{Y}(0) & R_{Y}(1) \end{bmatrix} \begin{bmatrix} Y_n \\ Y_{n-1} \end{bmatrix} \\
= \begin{bmatrix} 4/3 & 2/3 \\ 7/3 & 2/3 \\ 2/3 & 7/3 \end{bmatrix}^{-1} \begin{bmatrix} Y_n \\ Y_{n-1} \end{bmatrix} \\
= \begin{bmatrix} 8/15 & 2/15 \end{bmatrix} \begin{bmatrix} Y_n \\ Y_{n-1} \end{bmatrix} \\
= \frac{8}{15}Y_n + \frac{2}{15}Y_{n-1}.
\]

(f) Since $(X_n, Y_n)$ are jointly WSS, from part (e) it immediately follows that the conditional expectation $E[X_n|Y_n, Y_{n+1}]$ has the same form with $E[X_n|Y_n, Y_{n-1}]$:
\[
E[X_n|Y_n, Y_{n+1}] = \frac{8}{15}Y_n + \frac{2}{15}Y_{n+1}.
\]
1. Additive exponential noise channel (60 pts). A device has two equally likely states $S = 0$ and $S = 1$. When it is inactive ($S = 0$), it transmits $X = 0$. When it is active ($S = 1$), it transmits $X \sim \text{Exp}(1)$. Now suppose the signal is observed through the additive exponential noise channel with output

$$Y = X + Z,$$

where $Z \sim \text{Exp}(2)$ is independent of $(X, S)$. One wishes to decide whether the device is active or not.

(a) Find $f_{Y|S}(y|0)$.
(b) Find $f_{Y|S}(y|1)$.
(c) Find $f_Y(y)$.
(d) Find $p_{S|Y}(0|y)$ and $p_{S|Y}(1|y)$.
(e) Find the decision rule $d(y)$ that minimizes the probability of error

$$P(S \neq d(Y)).$$

(f) Find the corresponding probability of error.

(Hint: Recall that $Z \sim \text{Exp}(\lambda)$ means that its pdf is $f_Z(z) = \lambda e^{-\lambda z}, \ z \geq 0.$)

Solution:

(a) Given $S = 0$, $X = 0$ and thus $Y = Z \sim \text{Exp}(2)$. Hence,

$$f_{Y|S}(y|0) = \begin{cases} 2e^{-2y}, & y \geq 0 \\ 0, & y < 0. \end{cases}$$

(b) Given $S = 1$, $X \sim \text{Exp}(1)$ and $Y$ is the sum of two independent exponential...
tial random variables. Hence,

\[
f_{Y|S}(y|1) = f_{X|S}(y) * f_Z(y) \\
= e^{-y} \mathbb{1}_{\{y \geq 0\}} * 2e^{-2y} \mathbb{1}_{\{y \geq 0\}} \\
= \int_{-\infty}^{\infty} 2e^{-2t} \mathbb{1}_{\{t \geq 0\}} e^{-(y-t)} \mathbb{1}_{\{y-t \geq 0\}} dt \\
= \int_{0}^{y} 2e^{-2t+y} dt \\
= \begin{cases} 
2e^{-y}(1-e^{-y}), & y \geq 0 \\
0, & \text{otherwise.}
\end{cases}
\]

(c) We have

\[
f_Y(y) = f_{Y|S}(y|0)P(S = 0) + f_{Y|S}(y|1)P(S = 1) \\
= \frac{1}{2} \left( 2e^{-2y} + 2e^{-y} - 2e^{-2y} \right) \\
= e^{-y}, \quad y \geq 0.
\]

Thus, \( Y \sim \text{Exp}(1) \).

(d) We have

\[
p_{S|Y}(0|y) = \frac{f_{Y|S}(y|0)P(S = 0)}{f_Y(y)} \\
= \frac{e^{-2y}}{e^{-y}} \\
= e^{-y}.
\]

We similarly have

\[
p_{S|Y}(1|y) = \frac{f_{Y|S}(y|1)P(S = 1)}{f_Y(y)} \\
= \frac{e^{-y}(1-e^{-y})}{e^{-y}} \\
= 1 - e^{-y}.
\]

(Alternatively, \( p_{S|Y}(1|y) = 1 - p_{S|Y}(0|y) \).)
(e) We have
\[ d(y) = \arg \max_{s \in \{0, 1\}} p_{S|Y}(s|y) \]
\[ = \begin{cases} 0, & e^{-y} > 1 - e^{-y} \\ 1, & \text{otherwise.} \end{cases} \]

The condition \( e^{-y} > 1 - e^{-y} \) is equivalent to \( y < \ln 2 \), and hence
\[ d(y) = \begin{cases} 0, & 0 \leq y < \ln 2 \\ 1, & y \geq \ln 2 \end{cases} \]

(f) We have
\[
P(d(Y) \neq S) = P(d(Y) \neq S, S = 0) + P(d(Y) \neq S, S = 1)
= P(Y \geq \ln 2|S = 0)P(S = 0) + P(Y < \ln 2|S = 1)P(S = 1)
= \frac{1}{2}\left( \int_{\ln 2}^{\infty} f_{Y|S}(y|0)dy + \int_{0}^{\ln 2} f_{Y|S}(y|1)dy \right)
= \frac{1}{2}\left( \int_{\ln 2}^{\infty} 2e^{-2y}dy + \int_{0}^{\ln 2} 2e^{-y}(1 - e^{-y})dy \right)
= \frac{1}{2}\left( e^{-2\ln 2} + 2(1 - e^{-\ln 2}) - (1 - e^{-2\ln 2}) \right)
= \frac{1}{2}\left( \frac{1}{4} + 2 - 1 - 1 + \frac{1}{4} \right)
= \frac{1}{4}\]  

2. Brownian bridge (40 pts). Let \( \{W(t)\}_{t=0}^{\infty} \) be the standard Brownian motion (Wiener process). Recall that the process is independent-increment with \( W(0) = 0 \) and
\[ W(t) - W(s) \sim N(0, t - s), \quad 0 \leq s < t. \]

In the following, we investigate several properties of the process conditioned on \( \{W(1) = 0\} \).

(a) Find the conditional distribution of \( W(1/2) \) given \( W(1) = 0 \).
(b) Find \( \mathbb{E}[W(t) | W(1) = 0] \) for \( t \in [0, 1] \).
(c) Find \( \mathbb{E}[(W(t))^2 | W(1) = 0] \) for \( t \in [0, 1] \).
(d) Find $E[W(t_1)W(t_2) \mid W(1) = 0]$ for $t_1, t_2 \in [0, 1]$.

**Solution:**

(a) By the property of a Brownian motion,

$$\begin{bmatrix} W(1/2) \\ W(1) \end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} 0 \\ 1/2 \\ 1/2 \\ 1 \end{bmatrix} \right).$$

Therefore,

$$E[W(1/2) \mid W(1)] = E[W(1/2)] + \frac{\text{Cov}(W(1/2), W(1))}{\text{Var}[W(1)]} \left( W(1) - E[W(1)] \right)$$

$$= \frac{1}{2} W(1).$$

Also,

$$\text{Var}[W(1/2) \mid W(1)] = \text{Var}[W(1/2)] - \frac{\text{Cov}(W(1/2), W(1))^2}{\text{Var}[W(1)]}$$

$$= \frac{1}{2} - \frac{1}{4}$$

$$= \frac{1}{4}.$$

Thus, $W(1/2) \mid \{W(1) = 0\} \sim \mathcal{N}(0, 1/4)$.

(b) For $t \in [0, 1]$,

$$\begin{bmatrix} W(t) \\ W(1) \end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} 0 \\ t \\ t \\ 1 \end{bmatrix} \right).$$

Therefore,

$$E[W(t) \mid W(1)] = E[W(t)] + \frac{\text{Cov}(W(t), W(1))}{\text{Var}[W(1)]} \left( W(1) - E[W(1)] \right)$$

$$= tW(1).$$

Thus, $E[W(t) \mid W(1) = 0] = 0$. 

16
(c) 

\[
\text{Var}[W(t)|W(1)] = \text{Var}[W(t)] - \frac{\text{Cov}(W(t), W(1))^2}{\text{Var}[W(1)]} = t - t^2.
\]

Thus, \( E[W(t)^2|W(1) = 0] = t(1 - t). \)

(d) For \( 0 \leq t_1 \leq t_2 \leq 1, \)

\[
\begin{bmatrix}
  W(t_1) \\
  W(t_2) \\
  W(1)
\end{bmatrix}
\sim \mathcal{N}
\begin{pmatrix}
  0, \\
  t_1 & t_1 & t_1 \\
  t_1 & t_2 & t_2 \\
  t_1 & t_2 & 1
\end{pmatrix}
\]

Hence,

\[
\text{Cov}
\begin{pmatrix}
  [W(t_1)] \\
  [W(t_2)]
\end{pmatrix}
| W(1)
= \text{Cov}
\begin{pmatrix}
  [W(t_1)] \\
  [W(t_2)]
\end{pmatrix}
- \text{Cov}
\begin{pmatrix}
  [W(t_1)], W(1)
\end{pmatrix}
\text{Var}[W(1)]^{-1}\text{Cov}
\begin{pmatrix}
  [W(t_1)], W(1)
\end{pmatrix}^T
= \begin{bmatrix}
  t_1 & t_1 \\
  t_1 & t_2 \\
  t_2 & t_2 \\
\end{bmatrix}
- \begin{bmatrix}
  t_1 \\
  t_1 \\
  t_2 \\
\end{bmatrix}
\begin{bmatrix}
  t_1 & t_2 \\
\end{bmatrix}
= \begin{bmatrix}
  t_1(1 - t_1) & t_1(1 - t_2) \\
  t_1(1 - t_2) & t_2(1 - t_2)
\end{bmatrix}.
\]

This shows that

\[
E[W(t_1)W(t_2)|W(1) = 0] = t_1(1 - t_2)
= \min(t_1, t_2) - t_1t_2.
\]

3. **Convergence of random processes (30 pts).** Let \( \{N(t)\}_{t=0}^\infty \) be a Poisson process with rate \( \lambda \). Recall that the process is independent increment and \( N(t) - N(s), \quad 0 \leq s < t, \) has the pmf

\[
p_{N(t) - N(s)}(n) = \frac{e^{-\lambda(t-s)}(\lambda(t-s))^n}{n!}, \quad n = 0, 1, \ldots.
\]

Define

\[
M(t) = \frac{N(t)}{t}, \quad t > 0.
\]
(a) Find the mean and autocorrelation function of \( \{M(t)\}_{t > 0} \).

(b) Does \( \{M(t)\}_{t > 0} \) converge in mean square as \( t \to \infty \), that is,
\[
\lim_{t \to \infty} E[(M(t) - M)^2] = 0
\]
for some random variable (or constant) \( M \)? If so, what is the limit?

Now consider
\[
L(t) = \frac{1}{t} \int_0^t \frac{N(s)}{s} \, ds, \quad t > 0.
\]

(c) Does \( \{L(t)\}_{t > 0} \) converge in mean square as \( t \to \infty \)? If so, what is the limit?

(Hint: \( \int 1/x \, dx = \ln x + C \), \( \int \ln x \, dx = x \ln x - x + C \), and \( \lim_{x \to 0} x \ln x = 0 \).)

**Solution:**

(a) We have
\[
E[M(t)] = \frac{E[N(t)]}{t} = \frac{\lambda t}{t} = \lambda.
\]

Also, for \( \tau \geq 0 \), we have
\[
E[M(t)M(t + \tau)] = \frac{E[N(t)N(t + \tau)]}{t(t + \tau)} = \frac{E[N(t)(N(t) + N(t + \tau) - N(t))]}{t(t + \tau)} = \frac{E[N(t)^2] + E[N(t + \tau) - N(t)]E[N(t)]}{t(t + \tau)} \tag{by independent-increment property}
\]
\[
= \frac{\lambda t + \lambda^2 t^2 + \lambda \tau \cdot \lambda t}{t(t + \tau)} = \frac{\lambda + \lambda^2 (t + \tau)}{t + \tau} = \lambda^2 + \frac{\lambda}{t + \tau}.
\]
Thus the autocorrelation function is given by

\[ R_M(s, t) = \lambda^2 + \frac{\lambda}{\max(s, t)}. \]

(b) We have

\[
\text{Var}[M(t)] = \mathbb{E}[M(t)^2] - \left( \mathbb{E}[M(t)] \right)^2
\]

\[ = R_M(t, t) - \lambda^2 \]

\[ = \frac{\lambda}{t}. \]

Thus if we let \( M = \lambda \), we have

\[
\lim_{t \to \infty} \mathbb{E}[(M(t) - \lambda)^2] = \lim_{t \to \infty} \text{Var}[M(t)]
\]

\[ = \lim_{t \to \infty} \frac{\lambda}{t} \]

\[ = 0. \]

This shows that \( M(t) \to M \) in mean square.

(c) We have

\[
\mathbb{E}[L(t)] = \frac{1}{t} \int_0^t \mathbb{E}[M(s)]ds
\]

\[ = \lambda. \]
Also,

\[
\text{Var}[L(t)] = \mathbb{E}[L(t)^2] - \left(\mathbb{E}[L(t)]\right)^2
\]

\[
= \frac{1}{t^2} \int_0^t \int_0^t \mathbb{E}[M(u)M(v)]dudv - \lambda^2
\]

\[
= \frac{1}{t^2} \int_0^t \int_0^t \left(\lambda^2 + \frac{\lambda}{\max(u, v)}\right)dudv - \lambda^2
\]

\[
= \frac{1}{t^2} \int_0^t \int_0^t \frac{\lambda}{\max(u, v)}dudv
\]

\[
= \frac{\lambda}{t^2} \left(\int_0^v \frac{1}{v}du + \int_v^t \frac{1}{u}du\right)dv
\]

\[
= \frac{\lambda}{t^2} \left(1 + \ln \left(\frac{t}{v}\right)\right)dv
\]

\[
= \frac{\lambda}{t^2} \left(t + t \ln t - t \ln t + t\right)
\]

\[
= \frac{2\lambda}{t}.
\]

Thus if we let \( L = \lambda \), we have

\[
\lim_{t \to \infty} \mathbb{E}[(L(t) - L)^2] = \lim_{t \to \infty} \text{Var}[L(t)]
\]

\[
= \lim_{t \to \infty} \frac{2\lambda}{t}
\]

\[
= 0.
\]

This shows that \( L(t) \to L \) in mean square.

4. Random binary waveform (40 pts). Let \( \{N(t)\}_{t=0}^\infty \) be a Poisson process with rate \( \lambda \), and \( Z \) be independent of \( \{N(t)\} \) with \( P(Z = 1) = P(Z = -1) = 1/2 \). Define

\[
X(t) = Z \cdot (-1)^{N(t)}, \quad t \geq 0.
\]

(a) Find the mean and autocorrelation function of \( \{X(t)\}_{t=0}^\infty \).

(b) Is \( \{X(t)\}_{t=0}^\infty \) wide-sense stationary?

(c) Find the first-order pmf \( p_{X(t)}(x) = P(X(t) = x) \).
(d) Find the second-order pmf \( p_{X(t_1),X(t_2)}(x_1, x_2) = P(X(t_1) = x_1, X(t_2) = x_2) \).

(Hint: \( \sum_{k \text{ even}} x^k/k! = (e^x + e^{-x})/2 \) and \( \sum_{k \text{ odd}} x^k/k! = (e^x - e^{-x})/2 \).)

Solution:

(a) Since \( Z \) is independent of the process \( N(t) \), we have

\[
E[X(t)] = E[Z] \cdot E[(-1)^{N(t)}] = 0,
\]

and, for \( \tau \geq 0 \),

\[
E[X(t)X(t + \tau)] = E[Z^2] \cdot E[(-1)^{N(t)+N(t+\tau)}] = E[(-1)^{N(t)+N(t+\tau)}] = \sum_{k=0}^{\infty} (-1)^k \frac{e^{-\lambda \tau} (\lambda \tau)^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda \tau} \frac{(-\lambda \tau)^k}{k!} = e^{-2\lambda \tau}.
\]

Thus the autocorrelation function is given by

\[
R_X(s, t) = e^{-2\lambda |s-t|}.
\]

(b) Since \( E[X(t)] \) is constant and \( R_X(s, t) \) depends only on \(|s-t|\), \( X(t) \) is wide-sense stationary.

(c) We have

\[
P(X(t) = 1) = P(X(t) = 1|Z = 1)P(Z = 1) + P(X(t) = 1|Z = -1)P(Z = -1)
= \frac{1}{2} \left( P((-1)^{N(t)} = 1) + P((-1)^{N(t)} = -1) \right)
= \frac{1}{2} \left( P(N(t) = \text{even}) + P(N(t) = \text{odd}) \right)
= \frac{1}{2}.
\]

Thus, \( p_{X(t)}(1) = p_{X(t)}(-1) = 1/2 \).
Let \( t_2 \geq t_1 \). Since the process \( N(t) \) is independent of \( Z \) and \((N(t_2) - N(t_1))\) is independent of \( N(t_1) \), we conclude that \( N(t_2) - N(t_1) \) is independent of \((N(t_1), Z)\).

Hence, \( N(t_2) - N(t_1) \) is also independent of \( X(t_1) = Z \cdot (-1)^{N(t_1)} \).

We have

\[
P(X(t_2) = 1|X(t_1) = 1) = P\left(\frac{X(t_2)}{X(t_1)} = 1|X(t_1) = 1\right)
\]
\[
= P((-1)^{N(t_2) - N(t_1)} = 1|X(t_1) = 1)
\]
\[
= P((-1)^{N(t_2) - N(t_1)} = 1)
\]
\[
= P(N(t_2) - N(t_1) = \text{ even })
\]
\[
= \sum_{k \text{ even}} e^{-\lambda(t_2 - t_1)} \frac{(\lambda(t_2 - t_1))^k}{k!}
\]
\[
= e^{-\lambda(t_2 - t_1)} \left( \frac{e^{\lambda(t_2 - t_1)} + e^{-\lambda(t_2 - t_1)}}{2} \right)
\]
\[
= 1 + e^{-2\lambda(t_2 - t_1)}/2.
\]

Similarly,

\[
P(X(t_2) = -1|X(t_1) = -1) = P(N(t_2) - N(t_1) = \text{ even })
\]
\[
= 1 + e^{-2\lambda(t_2 - t_1)}/2, \text{ and}
\]

\[
P(X(t_2) = -1|X(t_1) = 1) = P(X(t_2) = 1|X(t_1) = -1)
\]
\[
= P(N(t_2) - N(t_1) = \text{ odd })
\]
\[
= \sum_{k \text{ odd}} e^{-\lambda(t_2 - t_1)} \frac{(\lambda(t_2 - t_1))^k}{k!}
\]
\[
= e^{-\lambda(t_2 - t_1)} \left( \frac{e^{\lambda(t_2 - t_1)} - e^{-\lambda(t_2 - t_1)}}{2} \right)
\]
\[
= 1 - e^{-2\lambda(t_2 - t_1)}/2.
\]

Thus,

\[
p_{X(t_1), X(t_2)}(x_1, x_2) = \begin{cases} 
1 + e^{-2\lambda(t_2 - t_1)}/4, & (x_1, x_2) = (1, 1) \text{ or } (-1, -1) \\
1 - e^{-2\lambda(t_2 - t_1)}/4, & (x_1, x_2) = (1, -1) \text{ or } (-1, 1).
\end{cases}
\]
Finally, if we remove the restriction $t_2 \geq t_1$, the above becomes

$$p_{X(t_1), X(t_2)}(x_1, x_2) = \begin{cases} 
\frac{1 + e^{-2\lambda|t_2-t_1|}}{4}, & (x_1, x_2) = (1, 1) \text{ or } (-1, -1) \\
\frac{1 - e^{-2\lambda|t_2-t_1|}}{4}, & (x_1, x_2) = (1, -1) \text{ or } (-1, 1).
\end{cases}$$
1. **Nonlinear and linear MMSE estimation (30 pts).** Let $X$ and $Y$ be two random variables with joint pdf

$$f_{X,Y}(x, y) = \begin{cases} x + y, & 0 \leq x \leq 1, \ 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) (10 points) Find the linear MMSE estimator of $X$ given $Y$.

(b) (10 points) Find the corresponding MSE.

(c) (10 points) Find the MMSE estimator of $X$ given $Y$. Is it the same as the linear MMSE estimator?

**Solution:**

(a) The linear MMSE estimator of $X$ given $Y$ has the form

$$\hat{X} = \frac{\mathbb{Cov}(X, Y)}{\text{Var}(Y)}(Y - \mathbb{E}[Y]) + \mathbb{E}[X].$$

$$f_X(x) = \int_0^1 (x + y) dy = xy + \frac{y^2}{2} \bigg|_0^1 = x + \frac{1}{2}.$$

$$\mathbb{E}[X] = \int_0^1 xf_X(x) dx = \int_0^1 x^2 + \frac{x}{2} dx = \frac{x^3}{3} + \frac{x^2}{4} \bigg|_0^1 = \frac{7}{12}.$$ 

$$\mathbb{E}[X^2] = \int_0^1 x^2 f_X(x) dx = \int_0^1 x^3 + \frac{x^2}{2} dx = \frac{x^4}{4} + \frac{x^3}{6} \bigg|_0^1 = \frac{5}{12}.$$ 

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}.$$ 

By the symmetry of $f_{X,Y}(x, y)$, $\mathbb{E}[Y] = \mathbb{E}[X]$ and $\text{Var}(Y) = \text{Var}(X)$.

$$\mathbb{E}[XY] = \int_0^1 \int_0^1 xy(x + y) dx dy = \int_0^1 \int_0^1 x^2 y + xy^2 dx dy = \int_0^1 \frac{x^3 y}{3} + \frac{x^2 y^2}{2} \bigg|_0^1 dy$$

$$= \int_0^1 \frac{y^2}{3} + \frac{y^2}{2} dy = \frac{y^2}{6} + \frac{y^3}{6} \bigg|_0^1 = \frac{1}{3}.$$
\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \left( \frac{7}{12} \right)^2 = \frac{48}{144} - \frac{49}{144} = -\frac{1}{144}.
\]

So,
\[
\hat{X} = -\frac{1}{11}(Y - \frac{7}{12}) + \frac{7}{12} = -\frac{1}{11}Y + \frac{7}{11}.
\]

(b) The MSE is given by
\[
\text{MSE} = \text{Var}(X) - \frac{\text{Cov}(X, Y)^2}{\text{Var}(Y)}
\]
so
\[
\text{MSE} = \frac{11}{144} - \frac{(-1/144)^2}{11/144} = \frac{5}{66}.
\]

(c) The pdf of \(X\mid\{Y = y\}\) is
\[
f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{x + y}{y + \frac{1}{2}}.
\]
\[
\mathbb{E}(X|Y = y) = \int_0^1 x f_{X|Y}(x|y)dx
\]
\[
= \int_0^1 \frac{x + y}{y + \frac{1}{2}} dx
\]
\[
= \int_0^1 \frac{x^2 + xy}{y + \frac{1}{2}} dx
\]
\[
= \left. \frac{x^3 + \frac{x^2 y}{2}}{y + \frac{1}{2}} \right|_0^1
\]
\[
= \frac{1}{3} + \frac{y}{2}
\]
\[
= \frac{2 + 3y}{3 + 6y}.
\]

So, the MMSE estimator is
\[
\mathbb{E}[X|Y] = \frac{2 + 3Y}{3 + 6Y}.
\]

It is different from the linear MMSE estimator.
2. **Convergence (30 pts).** Consider the sequence of i.i.d. random variables $X_1, X_2, \ldots$ with

$$X_i = \begin{cases} 
0 & \text{w.p. } \frac{1}{2}, \\
2 & \text{w.p. } \frac{1}{2}, 
\end{cases}$$

for all $i \geq 1$.

Define the sequence

$$Y_n = \begin{cases} 
X_n, & \text{for all } n \text{ w.p. } \frac{1}{3}, \\
\frac{1}{2}X_n, & \text{for all } n \text{ w.p. } \frac{1}{3}, \\
0, & \text{for all } n \text{ w.p. } \frac{1}{3}. 
\end{cases}$$

Let

$$M_n = \frac{1}{n} \sum_{i=1}^{n} Y_i.$$

(a) (10 points) Determine the probability mass function (pmf) of $Y_n$.

(b) (10 points) Determine the random variable (or constant) that $M_n$ converges to (in probability) as $n$ approaches infinity. Justify your answer.

(c) (10 points) Use the central limit theorem to estimate the probability that the random variable $M_{84}$ exceeds $\frac{2}{3}$.

**Solution:**

(a) $Y_n \sim Y$, where $Y$ has the pmf

$$p_Y(y) = \mathbb{P}(Y=y|Y=X)\mathbb{P}(Y=X) + \mathbb{P}(Y=y|Y=X)\mathbb{P}(Y=X) + \mathbb{P}(Y=y|Y=0)\mathbb{P}(Y=0)$$

$$= p_X(y)\frac{1}{3} + p_X(2y)\frac{1}{3} + p_0(y)\frac{1}{3}$$

$$= \begin{cases} 
0 & \text{w.p. } \frac{2}{3}, \\
1 & \text{w.p. } \frac{1}{6}, \\
2 & \text{w.p. } \frac{1}{6}.
\end{cases}$$
(b) \( M_n \) is the sample mean of \( Y_1, Y_2, \ldots, Y_n \). By the weak law of large numbers (WLLN), it converges in probability to \( E[Y] \), where

\[
E[Y] = \sum_{y \in Y} yp_Y(y)
\]

\[
= 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6}
\]

\[
= \frac{1}{2}
\]

(c) By the Central Limit Theorem (CLT)

\[
\frac{M_n - E[Y]}{\sigma_Y / \sqrt{n}} \to Z \sim N(0, 1).
\]

\[
E[Y^2] = \sum_{y^2 \in Y} y^2 p_Y(y)
\]

\[
= 0^2 \cdot \frac{2}{3} + 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6}
\]

\[
= \frac{5}{6}.
\]

So,

\[
\text{Var}(Y) = E[Y^2] - (E[Y])^2 = \frac{5}{6} - (\frac{1}{2})^2 = \frac{7}{12}
\]

and

\[
\sigma_Y = \sqrt{\frac{7}{12}}.
\]

\[
P(M_n > x) = P \left( \frac{M_n - E[Y]}{\sigma_Y / \sqrt{n}} > \frac{x - \frac{1}{2}}{\sqrt{7/12n}} \right).
\]

Setting \( n = 84 \) and \( x = \frac{2}{3} \), we get

\[
P(M_n > \frac{2}{3}) \approx P(Z > 2) = Q(2).
\]

3. Poisson Process (40 pts). Let \( \{N(t), t \geq 0\} \) be a Poisson process with arrival rate \( \lambda > 0 \).
(a) (10 points) Let \( T_M \) be the time of the \( M \)-th arrival. Find \( \mathbb{E}[T_M] \) and \( \text{Var}(T_M) \).

(b) (10 points) Let \( s \leq t \). Assume \( k \) arrivals occur in \( t \) seconds, that is \( N(t) = k \). Show that the conditional distribution of \( N(s) \) given \( N(t) = k \) satisfies \( N(s) \mid \{N(t) = k\} \sim \text{Binom}(k, \frac{s}{t}) \).

(c) (10 points) Let \( s \leq t \). Determine the conditional expectation \( \mathbb{E}[N(s) \mid N(t)] \) and give its probability mass function (pmf).

(d) (10 points) Assume \( N(t) = k \). Determine the probability that all \( k \) arrivals occur in the first \( \frac{t}{2} \) seconds.

**Solution:**

(a) The interarrival times \( X_1, X_2, \ldots X_M \) are i.i.d. \( \text{Exp}(\lambda) \) random variables, with mean \( 1/\lambda \) and variance \( 1/\lambda^2 \). Now,

\[
T_M = \sum_{i=1}^{M} X_i.
\]

Therefore, by independence, \( \mathbb{E}[T_M] = M/\lambda \) and \( \text{Var}(T_M) = M/\lambda^2 \).

(b) We can write

\[
P(N(s) = j \mid N(t) = k) = \frac{P(N(s) = j, N(t) = k)}{P(N(t) = k)}
= \frac{P(N(s) = j, N(t) - N(s) = k - j)}{P(N(t) = k)}
\]

\[\overset{(a)}{=} \frac{P(N(s) = j)P(N(t) - N(s) = k - j)}{P(N(t) = k)}
= \frac{(\lambda s)^j}{j!} e^{-\lambda s} \frac{(\lambda(t-s))^{k-j}}{(k-j)!} e^{-\lambda(t-s)} / \frac{\lambda^k}{k!}
\]

\[= \frac{k!}{j!(k-j)!} \frac{s^j}{t^k} \left( 1 - \frac{s}{t} \right)^{k-j}
= \binom{k}{j} \left( \frac{s}{t} \right)^j \left( 1 - \frac{s}{t} \right)^{k-j}
\]

where \((a)\) follows from the independent increment property.

So, \( N(s) \mid \{N(t) = k\} \sim \text{Binom}(k, \frac{s}{t}) \).
(c) Since \( N(s) \mid \{N(t) = k\} \sim \text{Binom}(k, \frac{s}{t}) \),

\[
E[N(s) \mid N(t) = k] = k \left( \frac{s}{t} \right).
\]

Therefore,

\[
E[N(s) \mid N(t)] = N(t) \left( \frac{s}{t} \right).
\]

The pmf is given by

\[
p_{E[N(s) \mid N(t)]}(x) = \begin{cases} \frac{(\lambda t)^k}{k!} e^{-\lambda t} & \text{if } x = k \left( \frac{s}{t} \right), k \geq 0, \\ 0 & \text{otherwise.} \end{cases}
\]

(d) We have that

\[
P(N(s) = k \mid N(t) = k) = \binom{k}{k} \left( \frac{s}{t} \right)^k = \left( \frac{s}{t} \right)^k.
\]

So,

\[
P(N \left( \frac{t}{2} \right) = k \mid N(t) = k) = \left( \frac{1}{2} \right)^k.
\]

4. Moving Average Process (40 pts). Let \( Z_0, Z_1, Z_2, \ldots \) be i.i.d. \( \sim \mathcal{N}(0, 1) \). Let \( Y_n = Z_{n-1} + Z_n \) for \( n \geq 1 \).

(a) (10 points) Find the mean function and autocorrelation function of \( \{Y_n\} \).

(b) (5 points) Is \( \{Y_n\} \) wide-sense stationary? Justify your answer.

(c) (10 points) Is \( \{Y_n\} \) Gaussian? Justify your answer.

(d) (5 points) Is \( \{Y_n\} \) strict-sense stationary? Justify your answer.

(e) (10 points) Is \( \{Y_n\} \) Markov? Justify your answer. [Hint: Compare \( E(Y_3 \mid Y_1, Y_2) \) to \( E(Y_3 \mid Y_2) \).]

Solution:

(a) We have

\[
E[Y_n] = E[Z_{n-1}] + E[Z_n] = 0, \ \forall n.
\]
From the properties of \( \{Z_i\} \), we get

\[
R_Y(m, n) = E[Y_m Y_n] \\
= E[(Z_{m-1} + Z_m)(Z_{n-1} + Z_n)] \\
= E[Z_{m-1}Z_{n-1}] + E[Z_{m-1}Z_n] + E[ZmZ_{n-1}] + E[ZmZ_n] \\
= \begin{cases} 
2, & m = n \\
1, & |m - n| = 1 \\
0, & \text{otherwise}. 
\end{cases}
\]

(b) \( E[Y_n] \) and \( R_Y(m, n) \) are both time-invariant, in the sense that \( E[Y_n] \) does not depend on \( n \) and \( R_Y(m, n) \) depends only on \( |m - n| \).

Therefore \( \{Y_n\} \) is wide-sense stationary (WSS).

(c) \((Y_1, \ldots, Y_n)\) is a linear transform of the Gaussian random vector \((Z_1, \ldots, Z_n)\).

Therefore \( \{Y_n\} \) is a Gaussian random process.

(d) Since \( \{Y_n\} \) is Gaussian and WSS, it is strict-sense stationary (SSS).

(e) The process is Gaussian, so the conditional expectation is the linear MMSE estimator.

\[
E[Y_3|Y_1, Y_2] = \Sigma_{XY}^\top \Sigma_{Y_1Y_2}^{-1} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \\
= \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \\
= \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \\
= \frac{1}{3}(2Y_2 - Y_1).
\]

On the other hand,

\[
E[Y_3|Y_2] = \frac{\text{Cov}(Y_3, Y_2)}{\text{Var}(Y_2)} Y_2 = \frac{1}{2} Y_2 \neq E[Y_3|Y_1, Y_2].
\]

This means that \( f_{Y_3|Y_1,Y_2}(y_3|y_1, y_2) \neq f_{Y_3|Y_2}(y_3|y_2) \), so the process is not Markov.
5. **WSS process through linear filter (40 pts).** Let \( Y(t) \) be a short-term integration of a WSS process \( X(t) \):

\[
Y(t) = \frac{1}{T} \int_{t-T}^{t} X(u) du.
\]

The frequency response \( H(f) \) of this linear integration system is

\[
H(f) = e^{-j\pi fT} \frac{\sin(\pi fT)}{\pi fT}.
\]

Suppose the input \( X(t) \) has mean \( \mathbb{E}[X(t)] \) and autocorrelation function

\[
R_X(\tau) = \begin{cases} 
1 - \frac{|	au|}{T}, & |\tau| \leq T \\
0, & \text{otherwise}
\end{cases}
\]

(a) (10 points) Determine the constant \( a \) such that \( \mathbb{E}[Y(t)] = a \mathbb{E}[X(t)] \).

(b) (10 points) Find \( S_Y(f) \).

(c) (10 points) Find \( R_Y(\tau) \). (You can leave your answer in the form of a convolution.)

(d) (10 points) Determine explicitly the average power of the output \( \mathbb{E}[Y^2(t)] \).

Hint: You may use the transform pair \( R_X(\tau) \leftrightarrow T \left( \frac{\sin(\pi fT)}{\pi fT} \right)^2 \) and Fourier Transform relationships from the tables provided.

**Solution:**

(a) We have

\[
\mathbb{E}[Y(t)] = H(0) \mathbb{E}[X(t)].
\]

\[
H(0) = e^{-j\pi fT} \frac{\sin(\pi fT)}{\pi fT} \bigg|_{f=0} = 1.
\]

So,

\[
\mathbb{E}[Y(t)] = \mathbb{E}[X(t)].
\]

(b) Since

\[
S_Y(f) = |H(f)|^2 S_X(f),
\]

we have

\[
S_X(f) = \mathcal{F}\{R_X(\tau)\} = T \left( \frac{\sin(\pi fT)}{\pi fT} \right)^2.
\]
where $\mathcal{F}(\cdot)$ denotes Fourier Transform. Also,

$$|H(f)|^2 = \left(\frac{\sin(\pi f T)}{\pi f T}\right)^2.$$ 

Therefore,

$$S_Y(f) = T \left(\frac{\sin \pi f T}{\pi f T}\right)^4.$$ 

(c) We have that

$$R_Y(\tau) = \mathcal{F}^{-1}\{S_Y(f)\}$$

$$= \mathcal{F}^{-1}\left\{ T \left(\frac{\sin \pi f T}{\pi f T}\right)^4 \right\}$$

$$= \frac{1}{T} \mathcal{F}^{-1}\left\{ T \left(\frac{\sin \pi f T}{\pi f T}\right)^2 \right\}^2$$

$$= \frac{1}{T} R_X(\tau) * R_X(\tau)$$

$$= \frac{1}{T} \int_{-\infty}^{\infty} R_X(s) R_X(\tau - s) ds$$

where (a) follows from the Time Convolution property of the Fourier Transform.

Alternatively, another approach is to write

$$R_Y(\tau) = h(\tau) * R_X(\tau) * h(-\tau)$$

where $h(\tau)$ is the impulse response of the linear integrator.

In particular,

$$h(\tau) = \frac{1}{T} u(\tau - T)$$

where $u(\tau)$ is the unit step function.

Now,

$$h(\tau) * h(-\tau) = \frac{1}{T} R_X(\tau)$$

so, again,

$$R_Y(\tau) = \frac{1}{T} R_X(\tau) * R_X(\tau).$$
(d) We have
\[ E[Y^2(t)] = R_Y(0) \]
and
\[ R_Y(0) = \frac{1}{T} R_X(\tau) * R_X(\tau) \]
\[ = \frac{1}{T} \int_{-\infty}^{\infty} R_X(s) R_X(-s) ds \]
\[ = \frac{1}{T} \int_{-\infty}^{\infty} \left(1 - \frac{|s|}{T}\right)^2 ds \]
\[ = \frac{2}{T} \int_{0}^{\infty} \left(1 - \frac{s}{T}\right)^2 ds \]
\[ = \frac{2}{T} \int_{0}^{\infty} \left(1 - \frac{2s}{T} + \frac{s^2}{T^2}\right) ds \]
\[ = \frac{2}{T} \left(s - \frac{s^2}{T} + \frac{s^3}{3T^2}\right) \bigg|_{0}^{T} \]
\[ = \frac{2}{T} \left(T - \frac{T^2}{T} + \frac{T^3}{3T^2}\right) \]
\[ = \frac{2}{3} \]

6. Optimal linear estimation (20 pts). Let \( X(t) \) be a zero-mean WSS process with autocorrelation function
\[ R_X(\tau) = e^{-|\tau|}. \]

(a) (10 points) Find the MMSE estimator for \( X(t) \) of the form
\[ \hat{X}(t) = aX(t - t_1) + bX(t - t_2), \]
where \( t_1 = t_0 \) and \( t_2 = 2t_0 \), where \( t_0 > 0 \).

(b) (10 points) Find the MSE of this estimator.

Solution:

(a) We use the orthogonality principle to determine the coefficients \( a \) and \( b \).

Once could also note that, because the process is zero-mean, the linear MMSE of \( X(t) \) given the random vector \( X(t - t_0), X(t - 2t_0) \) will be of
this form. Then one could find the linear MMSE estimator using the
techniques developed for that.

By the orthogonality principle,

\[ E[(X(t) - \hat{X}(t))X(t - t_1)] = 0. \]
\[ E[(X(t) - \hat{X}(t))X(t - t_2)] = 0. \]

Therefore,

\[ E[(X(t) - aX(t - t_0) - bX(t - 2t_0))X(t - t_0)] = 0. \]
\[ E[(X(t) - aX(t - t_0) - bX(t - 2t_0))X(t - 2t_0)] = 0. \]

This means

\[ E[X(t)X(t - t_0)] - aE[X(t - t_0)X(t - t_0)] - bE[X(t - 2t_0)X(t - t_0)] = 0 \]

or \( R_X(t_0) - aR_X(0) - bR_X(-t_0) = 0. \)

Similarly,

\[ E[X(t)X(t - 2t_0)] - aE[X(t - t_0)X(t - 2t_0)] - bE[X(t - 2t_0)X(t - 2t_0)] = 0 \]

or \( R_X(2t_0) - aR_X(t_0) - bR_X(0) = 0. \)

This gives a system of 2 linear equations for \( a \) and \( b \):

\[
\begin{bmatrix}
R_X(0) & R_X(-t_0) \\
R_X(t_0) & R_X(0)
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
R_X(t_0) \\
R_X(2t_0)
\end{bmatrix}.
\]

From the definition of \( R_X(\tau) \), this becomes

\[
\begin{bmatrix}
1 & e^{-|t_0|} \\
e^{-|t_0|} & 1
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
e^{-|t_0|} \\
e^{-2|t_0|}
\end{bmatrix}.
\]

Solving, either by direct observation from the first equation or by inverting
the \( 2 \times 2 \) matrix, we get

\[ a = e^{-|t_0|}, \ b = 0. \]

Therefore, noting that \( t_0 > 0 \), we have

\[ \hat{X}(t) = e^{-t_0}X(t - t_0). \]
(b) The MSE is given by

$$E[(X(t) - \hat{X}(t))^2] = E[(X(t) - \hat{X}(t))X(t)] - E[(X(t) - \hat{X}(t))\hat{X}(t)]$$

$$\equiv E[(X(t) - \hat{X}(t))X(t)] + 0$$

$$= E[X^2(t)] - E[\hat{X}(t)X(t)]$$

$$= R_X(0) - E[e^{-|t_0|}X(t - t_0)X(t)]$$

$$= R_X(0) - e^{-|t_0|}R_X(-t_0)$$

$$= 1 - e^{-2|t_0|}$$

where (a) uses the orthogonality of the estimation error and the estimate. Since $t_0 > 0$, we have

$$MSE = 1 - e^{-2t_0}$$