Final Exam 2020 Solutions

1. Suppose the cdf of (X_1, X_2) is given by

$$F(x_1, x_2) = 1 - e^{-x_1} - e^{-x_2} + e^{-x_1 - x_2} \text{ for } x_1 \geq 0, x_1 \geq 0$$

and is 0 otherwise.
(a) (5 marks) Determine the joint probability density function of (X_1, X_2).

$$f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}$$

$$= \begin{cases} \frac{\partial}{\partial x_1}\left(e^{-x_2} - e^{-x_1 - x_2}\right) & x_1, x_2 \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} e^{-x_1 - x_2} & x_1, x_2 \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) (5 marks) Are X_1 and X_2 statistically independent? Justify your answer.

Yes they are statistically independent

as $f(x_1, x_2) = f_1(x_1) f_2(x_2)$

where $f_i(x_i) = \begin{cases} e^{-x_i} & x_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$

which equals $\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = f_1(x_1)$

when $i=1$ and equals $\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$

$= f_2(x_2)$ when $i = 2$.

(c) (5 marks) Determne the mean vector and variance matrix of (X_1, X_2).

$E(X_1) = \int_0^\infty x_1 e^{-x_1} dx_1$ $\quad u = x_1 \quad dv = e^{-x_1}$

$du = 1 \quad v = -e^{-x_1}$

integration by parts $-x_1 e^{-x_1} \Big|_0^\infty + \int_0^\infty e^{-x_1} dx_1$

2

$= 0 + 1 = 1$

and similarly $E(X_2) = 1$. So the mean vector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Also $E(X_1^2) =$

$\int_0^\infty x_1^2 e^{-x_1} dx_1$ $\quad u = x_1^2 \quad dv = e^{-x_1}$

$du = 2x_1 \quad v = -e^{-x_1}$

3

$= -x_1^2 e^{-x_1} \Big|_0^\infty + 2 \int_0^\infty x_1 e^{-x_1} dx_1 = 0 + 2 = 2$

So $Var(X_1) = Var(X_2) = 2 - 1^2 = 1$. Finally X_1 and X_2 are stat. ind. with finite variances so $cov(X_1, X_2) = 0$. Therefore the variance matrix is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(d) (5 marks) Determine the mean function and variance matrix of (Y_1, Y_2) where

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

$$E\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$\operatorname{Var}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} I \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 5 \\ 5 & 10 \end{pmatrix}$$

(e) (5 marks) Are Y_1 and Y_2 in (d) statistically independent? Justify your answer.

No they are not statistically independent because $cov(Y_1, Y_2) = 5 \neq 0$ and $cov(Y_1, Y_2)$ would be 0 if they are stat. ind.

(f) (5 marks) Determine the joint density function of (Y_1, Y_2).

Let $T(x_1, x_2) = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ where

$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. Then $\underset{\sim}{Y} = T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ has

density

1 $f_{\underset{\sim}{Y}}(y_1, y_2) = f_{\underset{\sim}{X}}(T^{-1}(y_1, y_2))\, J_T(T^{-1}(y_1, y_2))$

where $J_T(x_1, x_2) = \left| \det \begin{pmatrix} \frac{\partial T_1(x_1,x_2)}{\partial x_1} & \frac{\partial T_1(x_1,x_2)}{\partial x_2} \\ \frac{\partial T_2(x_1,x_2)}{\partial x_1} & \frac{\partial T_2(x_1,x_2)}{\partial x_2} \end{pmatrix} \right|$

1 $= \left| \det \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \right|^{-1} = |3+2| = 5$ and

1 $T^{-1}(y_1, y_2) = A^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$= \begin{pmatrix} (3y_1 - 2y_2)/5 \\ \frac{y_1 + y_2}{5} \end{pmatrix}$ and so

$f_{\underset{\sim}{Y}}(y_1, y_2) = \begin{cases} \frac{1}{5}\exp\left\{ -\frac{(3y_1 - 2y_2)}{5} - \frac{y_1 + y_2}{5} \right\}, & 3y_1 - 2y_2 \geq 0, \ y_1 + y_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$ 1

$= \begin{cases} \frac{1}{5}\exp\left\{ -\frac{4y_1}{5} + \frac{y_2}{5} \right\} & 3y_1 - 2y_2 \geq 0, \ y_1 + y_2 \geq 0 \\ 0 & \text{otherwise.} \end{cases}$

2. Suppose $Z_0, Z_1, \ldots$ are *i.i.d.* $N(0,1)$. With $T = \{1, 2, \ldots\}$ define the stochastic process $\{(t, X_t) : t \in T\}$ by $X_t = Z_t Z_{t-1}$.
(a) (5 marks) Determne the mean and autocovariance functions of this process.

$$\mu(t) = E(X_t) = E(Z_t Z_{t-1})$$

2

$$= E(Z_t)\,E(Z_{t-1}) \quad \text{since the } Z_i \text{ are mut. stat. ind.}$$

$$= 0 \quad \text{since } E(Z_t) = 0 \ \forall t.$$

$$\sigma(s,t) = \operatorname{cov}(X_s, X_t) = E(X_s X_t) \quad \text{since means} = 0$$

$$= E(Z_s Z_{s-1} Z_t Z_{t-1})$$

3

$$= \begin{cases} E(Z_s)E(Z_{s-1})E(Z_t)E(Z_{t-1}) & s < t-1 \\ E(Z_{t-2})E(Z_{t-1}^2)E(Z_t) & s = t-1 \\ E(Z_t^2)E(Z_{t-1}^2) & s = t \\ E(Z_{t+1})E(Z_t^2)E(Z_{t-1}) & s = t+1 \\ E(Z_s)E(Z_{s-1})E(Z_t)E(Z_{t-1}) & s > t+1 \end{cases}$$

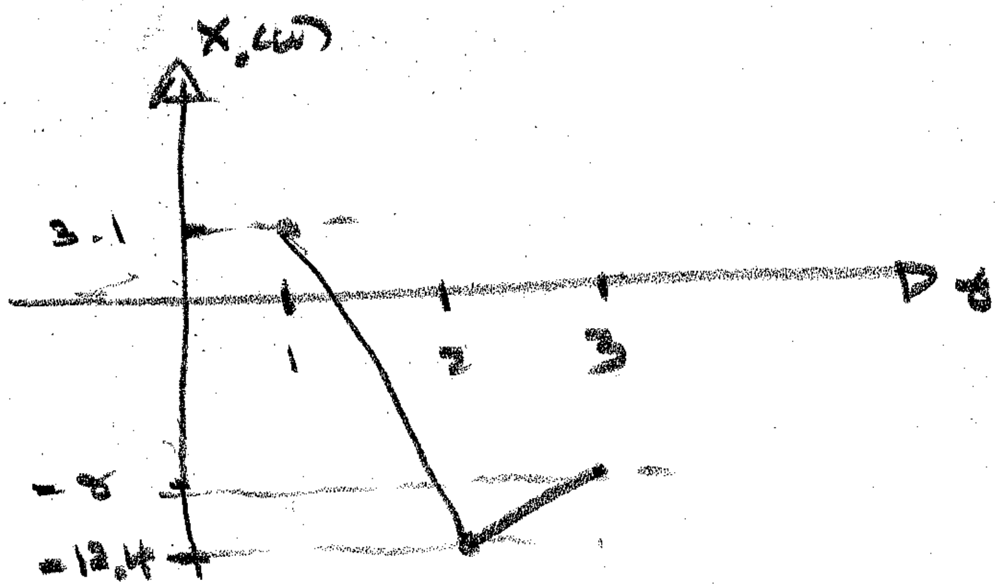
$$= \begin{cases} 0 & s < t-1 \\ 0 & s = t-1 \\ 1 \quad (\text{since } E(Z_t^2) = 1) & s = t \\ 0 & s = t+1 \\ 0 & s > t+1 \end{cases}$$

(b) (5 marks) If $Z_0 = 1, Z_1 = 3.1, Z_2 = -4, Z_3 = 2$, then plot the first three values of the sample function of the X_t process.

$X_1(\omega) = Z_0(\omega) Z_1(\omega) = 3.1$

$X_2(\omega) = Z_1(\omega) Z_2(\omega) = (3.1)(-4) = -12.4$

$X_3(\omega) = Z_2(\omega) Z_3(\omega) = (-4)(2) = -8$



(c) (10 marks) Determine the moment generating function $m_{X_t}(s)$ of X_t. (Hint: use the theorem of total expectation.) Does the mgf exist for all $s \in R^1$?

$m_X(s) = E(\exp\{sX_t\}) = E(\exp\{sZ_t Z_{t-1}\})$

3 $\overset{TTE}{=} E(E(\exp\{sZ_t Z_{t-1}\} | Z_{t-1}))$

2 $= E(\exp\{\frac{1}{2} s^2 Z_{t-1}^2\})$ since Z_t, Z_{t-1} are stat. ind so $Z_t | Z_{t-1} \sim N(0,1)$ and use formula for mgf of $N(0,1)$

$= \int_{-\infty}^{\infty} \exp\{\frac{1}{2} s^2 z^2\} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2} z^2\} dz$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(1-s^2) z^2\} dz$
which is finite iff $s \in (-1,1)$

3 $= \frac{1}{\sqrt{1-s^2}} \int_{-\infty}^{\infty} \frac{\sqrt{1-s^2}}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(1-s^2) z^2\} dz$

9 $\quad = 1$ since integral of $N(0, \frac{1}{1-s^2})$ density

2 So mgf exists only for $|s| < 1$.

3. Suppose that $h : R^4 \to R^1$ is given by $h(x_1, x_2, x_3, x_4) = x_1 + x_2 + x_3^2 + x_4^2$.
(a) (5 marks) Prove that h is convex.

$$\frac{\partial h}{\partial x_1} = 1 \quad \frac{\partial h}{\partial x_2} = 1 \quad \frac{\partial h}{\partial x_3} = 2x_3 \quad \frac{\partial h}{\partial x_4} = 2x_4$$

$$\frac{\partial^2 h}{\partial x_1^2} = 0 \quad \frac{\partial^2 h}{\partial x_1 \partial x_i} = 0 \quad i = 2, 3, 4$$

$$\frac{\partial^2 h}{\partial x_2^2} = 0 \quad \frac{\partial^2 h}{\partial x_2 \partial x_i} = 0 \quad i = 1, 3, 4$$

$$\frac{\partial^2 h}{\partial x_3^2} = 2 \quad \frac{\partial^2 h}{\partial x_3 \partial x_i} = 0 \quad i = 1, 2, 4$$

$$\frac{\partial^2 h}{\partial x_4^2} = 2 \quad \frac{\partial^2 h}{\partial x_4 \partial x_i} = 0 \quad i = 1, 2, 3$$

$\therefore$ the Hessian is

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \text{ and } \underline{x}' H \underline{x} = 2x_3^2 + 2x_4^2$$

$\geq 0 \quad \forall (x_1, x_2, x_3, x_4)$ so H is positive semidefinite

$\therefore h$ is convex

(b) (5 marks) Suppose that $\mathbf{X} = (X_1, X_2, X_3, X_4)'$ has a joint distribution with mean vector and variance matrix given by

$$\mu = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix}.$$

Determine a general lower bound on $E(h(\mathbf{X}))$.

By Jensen's inequality

$$E(h(\underset{\sim}{X})) \geq h(E(\underset{\sim}{X})) = h(1,-1,0,1)$$

$$= 1 + (-1) + 0^2 + 1^2 = 2.$$

(c) (5 marks) If $\mathbf{X} \sim N_4(\mu,\Sigma)$, then determine $E(h(\mathbf{X}))$ exactly.

$$E(h(\underset{\sim}{X})) = E(X_1) + E(X_2) + E(X_3^2) + E(X_4^2)$$

$$= 1 + (-1) + (2 - 0^2) + (3 - 1^2) = 2 + 2 = 4.$$

(x_1, x_2)

(d) (5 marks) What is the best affine predictor of $\mathbf{X}_2 = (X_3, X_4)$ when $\mathbf{X}_1 = (X_1, X_2) = (\cancel{1,1,2,2})$ is observed (no need to do all the arithmetic the formula is good enough)? Under what conditions is this also the best predictor and explain what "best" means.

The best affine predictor of $\underset{\sim}{X}_2$ is

$$h(x_1, x_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}' \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^{-1} \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$$

This is the best predictor when $\underset{\sim}{X} \sim N_4(\underset{\sim}{\mu}, \Sigma)$.

By best is meant that $h(x_1, x_2)$ minimizes $E\left(\| \underset{\sim}{X}_2 - g(\underset{\sim}{X}_1) \|^2 \right)$ among all functions g of $\underset{\sim}{X}_1$.

4. Suppose that $X_1, \ldots, X_n$ is a sample (*i.i.d.*) from a distribution with mean μ and variance $\sigma^2 > 0$.
(a) (5 marks) Determine the limiting value of the sequence of random variables $W_n = \frac{1}{n}\sum_{i=1}^n X_i^2$ as $n \to \infty$. Explain clearly what this convergence means and justify your answer.

$W_n \xrightarrow{wp1} E(X_1^2) = \sigma^2 + \mu^2$

by the Strong Law of Large Numbers

(okay to use $W_n \xrightarrow{P} \sigma^2 + \mu^2$ by Weak Law too)

This means $P(\{\omega : \lim_{n\to\infty} W_n(\omega) = \sigma^2 + \mu^2\}) = 1$

(This means for any $\delta > 0$ and $\varepsilon > 0$ $\exists N_{\delta,\varepsilon}$ s.t. $\forall n > N_{\varepsilon,\delta}$ $P(|W_n - (\sigma^2+\mu^2)| > \delta) < \varepsilon$)

(b) (5 marks) Repeat part (a) but for the sequence of random variables $Y_n = \bar{X}/s$ where $\bar{X} = \frac{1}{n}\sum_{i-1}^{n} X_i$, $s^2 = \frac{1}{n}\sum_{i-1}^{n}(X_i - \bar{X})^2$.

$\bar{X} \xrightarrow{wp1} \mu$ by the SLLN

$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$ and $\frac{1}{n}\sum_{i=1}^{n} x_i^2 \xrightarrow{wp1} \mu^2 + \sigma^2$

$\bar{x}^2 = \bar{x}\,\bar{x} \xrightarrow{wp1} \mu^2$ and so

$s^2 \xrightarrow{wp1} \mu^2 + \sigma^2 - \mu^2 = \sigma^2$

which implies $s = \sqrt{s^2} \xrightarrow{wp1} \sigma$ by the continuity of $\sqrt{\ }$ and so by the continuity of $h(x,y) = x/y$ when $y \neq 0$ we have $\bar{x}/s \xrightarrow{wp1} \mu/\sigma$

Also could use $\bar{x} \xrightarrow{d \text{ or } P} \mu$, $s^2 \xrightarrow{d \text{ or } P} \sigma^2$ and then by Slutsky's Theorem $\bar{x}/s \rightarrow \mu/\sigma$

(c) (10 marks) Now consider the sequence of random variables $Z_n = n^{1/2}(\bar{X} - \mu)/s$. ~~Compare~~ Determine the limiting distributions of Y_n and Z_n and justify your answer.

5 $Z_n = n^{1/2}(\bar{X}-\mu)/s \xrightarrow{d} N(0,1)$ by
3 the CLT, and Slutsky's Theorem while
2 $Y_n \xrightarrow{d} \mu/\sigma$ a constant since
convergence wp1 implies convergence in
distribution.

5. (10 marks) Suppose that $Z_0, Z_1, \ldots$ are *i.i.d.* $N(0,1)$ and $X_n = 1/n + \alpha Z_n + \beta Z_{n-1}$. Determine whether or not $\{(n, X_n) : n \in \mathbb{N}\}$ is a stationary Gasussian process.

$$\underline{X}_n = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \\ \vdots \\ 1/n \end{pmatrix} + \begin{pmatrix} \beta & \alpha & 0 & \cdots & 0 & 0 \\ 0 & \beta & \alpha & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & \cdots & & & \beta & \alpha \end{pmatrix} \begin{pmatrix} Z_0 \\ Z_1 \\ \vdots \\ Z_n \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1/2 \\ \vdots \\ 1/n \end{pmatrix} + A\underline{Z}, \text{ where } \underline{Z} \sim N_{n+1}(\underline{0}, I)$$

and $A\underline{Z} \sim N_n(\underline{0}, AA')$ and

so $\underline{X}_n \sim N_n\left(\begin{pmatrix} 1 \\ 1/2 \\ \vdots \\ 1/n \end{pmatrix}, AA'\right)$. This

implies that X_n process is a Gaussian process but it is not stationary because the mean function is not constant.