(§P) Answer only 2 questions out of P1-P3

P1.

Let $X$ be an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{G}$ be a $\sigma$-subalgebra of $\mathcal{F}$.

(a) Carefully define (and justify the existence of) the conditional expectation $\mathbb{E}(X|\mathcal{G})$. 3 MARKS

(b) Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be $\sigma$-subalgebras of $\mathcal{F}$ with $\mathcal{G}_1 \subset \mathcal{G}_2$. Prove the smoothing property:

$$\mathbb{E}[\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1] = \mathbb{E}(X|\mathcal{G}_1) \ a.s.$$ 7 MARKS

(c) Suppose $X$ and $Y$ are jointly continuous with joint density function $f(x, y)$ and $\mathbb{E}(|Y|) < \infty$. Let $g(x)$ be the marginal density function of $X$, and define

$$h(x) = I_{\{g > 0\}}(x) \int_{-\infty}^{+\infty} y f(x, y) dy.$$ 10 MARKS

Show that $h$ is measurable and that $h(X)$ is a version of $\mathbb{E}[Y|\sigma(X)]$.

P2.

Let $\{X_n, n \geq 1\}$ be a sequence of finite-valued random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

(a) Show that if $X_n \to X$ in probability, there exists a subsequence $\{X_{n_i}, i \geq 1\}$ such that $X_{n_i} \to X$ a.s. 7 MARKS

(b) Show that if $X_n \to 0$ a.s., and if $\{X_n, n \geq 1\}$ are uniformly integrable, then $\mathbb{E}|X_n| \to 0$ as $n \to \infty$. 6 MARKS

(c) Show that it is always possible to find a sequence $\{a_n, n \geq 1\}$ of constants such that $\frac{X_n}{a_n} \to 0$ a.s. [Hint: Consider $\sum_{n=1}^{\infty} P(|\frac{X_n}{a_n}| > \frac{1}{n})$ for suitable $a_n$.] 7 MARKS

P3.

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\mathcal{C} \subset \mathcal{F}$. Let $\sigma(\mathcal{C})$ be the $\sigma$-algebra generated by $\mathcal{C}$.

(a) Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be subclasses of $\mathcal{F}$, each closed under finite intersection. Suppose that $P(A_1 \cap A_2) = P(A_1)P(A_2)$ for all $A_1 \in \mathcal{C}_1$ and $A_2 \in \mathcal{C}_2$. Let $\mathcal{F}_1 = \sigma(\mathcal{C}_1)$ and $\mathcal{F}_2 = \sigma(\mathcal{C}_2)$. Prove that $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent. 7 MARKS

(b) Suppose that $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$ are independent $\sigma$-subalgebras of $\mathcal{F}$. Show that $\sigma(\mathcal{F}_1 \cup \mathcal{F}_2)$ and $\mathcal{F}_3$ are independent. 7 MARKS

(c) Suppose $X$ and $Y$ are independent non-negative random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Define what it means to say $X$ and $Y$ are independent, and prove that $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. 6 MARKS
(§S.) Answer only 4 questions out of S1-S6

S1.

(a) Let the joint density for random variables \( X \) and \( Y \) be defined by
\[
f_{X,Y}(x, y) = 6(1 - y) \quad 0 < x < y < 1
\]
and zero otherwise.

(i) Evaluate \( E[f_X[X]] \) and \( E[f_X[X]^2] \), and hence evaluate \( V[f_X[X]] \).

5 MARKS

(ii) Derive the conditional density, \( f_{Y|X}(y|x) \), and the conditional expectation
\[
E[f_{Y|X}[(1 - Y)|X = x]],
\]
for general \( x, 0 < x < 1 \). Hence or otherwise, evaluate \( E[f_Y[Y]] \) and \( C_{f_{X,Y}}[X, Y] \), the covariance of \( X \) and \( Y \).

5 MARKS

(b) Let the joint density for random variables \( U \) and \( V \) be defined by
\[
f_{U,V}(u, v) = \frac{3}{2} u^2 (1 - |v|) \quad -1 < u < 1, -1 < v < 1
\]
and zero otherwise. Let \( A \equiv \{(u, v) : 0 < u < 1, 0 < v < u\} \). Compute \( P[(U, V) \in A] \).

5 MARKS

(c) Let the joint density for random variables \( R \) and \( S \) be defined by
\[
f_{R,S}(r, s) = 2r \quad 0 < r < 1, 0 < s < 1
\]
and zero otherwise. Find the probability
\[
P[R^2 < S < R] .
\]

5 MARKS

S2.

(a) The cumulant generating function for random variable \( X \), \( K_X(t) \), is defined by
\[
K_X(t) = \log M_X(t) = \log \mathbb{E}[e^{tX}]
\]
for \( t \in (-h, h) \), some \( h > 0 \). Show that
\[
K_X(\lambda t_1 + (1 - \lambda) t_2) \leq \lambda K_X(t_1) + (1 - \lambda) K_X(t_2)
\]
for \( t_1, t_2 \in (-h, h) \), and \( 0 \leq \lambda \leq 1 \).

6 MARKS
(b) Consider the probability density functions $f_0$ and $f_\beta$ for $\beta > 0$ such that

$$f_\beta(x) = \frac{f_0(x) \exp{-\beta h(x)}}{Z(\beta)}$$

for function $h(x) > 0$, where $Z(\beta)$ is the normalizing constant for density $f_\beta$.

(i) Show that

$$\frac{dZ(\beta)}{d\beta} = -Z(\beta) E_{f_0}[h(X)]$$

and find a similar expression for

$$\frac{d}{d\beta} \left\{ \frac{1}{Z(\beta)} \right\}.$$ 

6 MARKS

(ii) Let $g(\beta)$ be defined by

$$g(\beta) = E_{f_\beta}[h(X)].$$

Show that $g(\beta)$ is a non-increasing function of $\beta$. 

8 MARKS

S3.

(a) Suppose that

$$F_X(x) = \frac{(x + 1)^2 - 1}{(x + 1)^2} \quad 0 < x < \infty$$

with $F_X(x) = 0$ for $x \leq 0$. Let $Y_n$ be the maximum order statistic derived from a random sample $X_1, \ldots, X_n$ from this distribution.

(i) Show that in this case $Y_n$ has no limiting distribution. 

4 MARKS

(ii) Consider the transformed variable

$$Z_n = \frac{1}{\sqrt{n}} Y_n.$$ 

Show that $Z_n$ does have a limiting distribution, and identify it.

4 MARKS

(b) Now suppose that $X_1, \ldots, X_n$ are independent Exponential (1) random variables. Let $Y_n$ be the maximum order statistic derived from a random sample $X_1, \ldots, X_n$.

(i) Find the distribution function for $Z_n = Y_n - a$ for constant $a$. 

4 MARKS

(ii) Find a sequence of constants $\{a_n\}$ such that, for $z > 0$,

$$\lim_{n \to \infty} F_{Z_n}(z) = \exp{-\exp{-z}}.$$ 

3
and hence find an approximation to the probability

\[ P \left[ Y_n > y_0 \right] \]

for large \( n \), for fixed \( y_0 > 0 \).  

4 MARKS

(c) The times between earthquakes of a certain magnitude that occur in Southern California are assumed to be independent Exponential random variables with rate parameter \( \lambda = 0.01 \) (per day), so that the expected time between earthquakes is 100 days. Over a given recording period, \( n = 200 \) earthquakes were observed, and the longest time between any two consecutive earthquakes was 1021 days. Comment on the validity of the Exponential distribution assumption in the light of these data.

4 MARKS

S4.

Let \( \vec{X} = [X_1, X_2, \ldots, X_n]' \) be a random sample from an Exponential distribution with density

\[ f_\theta(x) = \theta e^{-\theta x} I[x \geq 0], \theta > 0. \]

(a) Find the maximum likelihood estimator (MLE) of \( \theta \), say \( \hat{\theta}_n \), and show that \( \hat{\theta}_n \) is consistent and asymptotically normal (without resorting to known results regarding MLEs). Invoke clearly any other theorem you need.

5 MARKS

Now suppose that we observe only \( Y_i = [X_i] + 1 \), where the square bracket means the largest integer smaller than or equal to \( X_i \), \( i = 1, \ldots, n \).

(b) Find the MLE of \( \theta \), say \( \tilde{\theta}_n \) based on the random sample \( \vec{Y} = [Y_1, \ldots, Y_n]' \).

5 MARKS

(c) Find the asymptotic distribution of \( \tilde{\theta}_n \).

5 MARKS

(d) Find the asymptotic relative efficiency of \( \hat{\theta}_n \) with respect to \( \tilde{\theta}_n \) and comment on the situations where \( \theta \) is small and where \( \theta \) tends to infinity.

5 MARKS

S5.

The class of power series distributions consists of families of probability mass functions (pmf) of the form

\[ f_\gamma(x) = \frac{a(x)}{K(\gamma)} \gamma^x I[x \in \{0, 1, 2, \ldots\}] \]

where \( \gamma > 0 \), and where \( a(x) \) satisfies \( \sum_{x=0}^{\infty} a(x) \gamma^x = K(\gamma) < \infty \). Suppose that \( \vec{X} = [X_1, \ldots, X_n]' \) is a random sample from a family of power series distributions given by \( \{f_\gamma : \gamma > 0\} \). Let \( \vec{Y} = \sum_{i=1}^{n} X_i \).

(a) Show that \( Y \) has a power series distribution.  

2 MARKS
(b) Show that $Y$ is $\gamma$-sufficient.

(c) Show that $Y$ is $\gamma$-complete if the set $\{\gamma > 0 : K(\gamma) < \infty\}$ contains an interval.

(d) Let $A(y) = \sum\sum_{i=1}^{n} x_i = y \prod_{i=1}^{n} a(x_i)$. Show that the uniformly minimum variance unbiased estimator (UMVUE) of $\gamma^r$ is

$$
\tau(Y) = \begin{cases} 
0, & \text{if } Y < r \\
A(Y - r)/A(Y), & \text{if } Y \geq r
\end{cases}
$$

for any $Y$ with positive probability under any $\gamma > 0$.

Optional hint: You may wish to consider an expression for $K(\gamma)^n \mathbb{E}_\gamma T$ for a candidate estimator $T$.

(e) Let $X = [X_1, \ldots, X_n]'$ be a random sample from a Geometric($p$) distribution with pmf $f_p(x) = p(1-p)^x I[x \in \{0, 1, \ldots\}]$. Obtain the UMVUE of $p$.

(f) Under the same setting as in (e), find the UMVUE of $\theta = (1-p)/p$.

S6.

(a) State the Neyman-Pearson Lemma. Make sure to clearly state the criteria and conclusions for the so-called necessity and sufficiency parts of the lemma.

For the rest of this question, let $X \sim U(\theta, \theta + 1)$, $\theta \in \{0, \theta_1\}$ for some $\theta_1 > 0$. We wish to test $H_0 : \theta = 0$ against $H_1 : \theta = \theta_1$.

(b) Let $\theta_1 = 1$. Find the most powerful level $\alpha$ test for all $1 \geq \alpha > 0$. Indicate which necessity condition of the Neyman-Pearson Lemma is not met.

(c) Let $\theta_1$ be a value in $(0.95, 1)$. Show that there exists a most powerful level 5% test that has size strictly smaller than 5%. Explain what condition in the Neyman-Pearson Lemma is not satisfied.

(d) Let $\theta_1$ be a value in $(0, 0.95]$.

i. Find a most powerful level 5% test.

ii. Show that this test is not unique.

(e) Sketch on one graph the maximum power attainable in the alternative for $H_1 : \theta = \theta_1$ as a function of $\theta_1 \in (0, 1]$. 

5 MARKS
INSTRUCTIONS

Answer only two questions out of Section P.
Answer only four questions out of Section S.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td></td>
</tr>
</tbody>
</table>

This exam comprises the cover and 5 pages of questions.