Date: Monday August 12th, 2019
Time: 1pm-5pm

Instructions

- Answer only two questions from Section P. If you answer more than two questions, then only the FIRST TWO questions will be marked.

- Answer only four questions from Section S. If you answer more than four questions, then only the FIRST FOUR questions will be marked.

<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 page and six pages of questions.
• You may use any result that is known to you, but you must state the name of the result (law/theorem/lemma/formula/inequality) that you are using, and show the work of verifying the condition(s) for that result to apply.

• For the problems with multiple parts, you are allowed to assume the conclusion from the previous part in order to solve the next part, whether or not you have completed the previous part.

P1. Let \( \{X_n : n \geq 1\} \) be a sequence of independent and identically distributed \( \mathbb{R} \)-valued random variables on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Assume that \( \mathbb{E}[|X_1|] < \infty \). Set \( S_n := \sum_{j=1}^{n} X_j \) for every \( n \geq 1 \). Show that:

a) \( \{\frac{S_n}{n} : n \geq 1\} \) is uniformly integrable. \(\text{(10 marks)}\)

b) \( \frac{S_n}{n} \rightarrow \mathbb{E}[X_1] \) almost surely as well as in \( L^1 \), as \( n \rightarrow \infty \). \(\text{(10 marks)}\)

P2. (a) Show that a monotone function \( f : \mathbb{R} \rightarrow \mathbb{R} \) is measurable. \(\text{(10 marks)}\)

(b) Suppose that \( \mu \) is a measure on \((\mathbb{R}, \mathcal{B}(\mathbb{R}))\) which is (a) translation invariant (i.e. \( \mu(A + x) = \mu(A) \) for all \( A \in \mathcal{B}(\mathbb{R}) \)) and (b) finite on bounded sets (i.e. if \( A \in \mathcal{B}(\mathbb{R}) \) satisfies \( A \subset [-x, x] \) for some \( 0 < x < \infty \), then \( \mu(A) < \infty \)). Show that \( \mu \) is a multiple of Lebesgue measure. \(\text{(10 marks)}\)
P3. Given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a constant \(p \in (0, 1)\), let \(\{X_n : n \geq 1\}\) be a sequence of i.i.d. random variables with the common distribution

\[
\mathbb{P}(X_1 = 1) = p \quad \text{and} \quad \mathbb{P}(X_1 = 0) = 1 - p.
\]

Define a mapping \(D : \Omega \to [0, 1]\) by

\[
D = \sum_{n=1}^{\infty} \frac{X_n}{2^n}.
\]

Obviously \(D\) is well-defined as a random variable because the series above is convergent almost surely. Denote by \(\mu_p\) the law of \(D\) (therefore, \(\mu_p\) is a probability measure on \(([0, 1], \mathcal{B}([0, 1]))\)) and \(F_p\) the distribution function of \(D\).

a) Prove that \(F_p\) is a continuous function. \(\text{(10 marks)}\)

b) For each \(x \in [0, 1]\) and \(n \geq 1\), define

\[
E_n(x) := \begin{cases} 
1, & \text{if } 2^{n-1}x - \left\lfloor 2^{n-1}x \right\rfloor \geq \frac{1}{2}, \\
0, & \text{if } 2^{n-1}x - \left\lfloor 2^{n-1}x \right\rfloor < \frac{1}{2},
\end{cases}
\]

where \(\lfloor s \rfloor\) denotes the integer part of \(s \geq 0\). Prove that

\[
\mu_p \left( \left\{ x \in [0, 1] : \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} E_n(x) = p \right\} \right) = 1,
\]

from where argue that if \(p_1 \neq p_2\), then \(\mu_{p_1} \perp \mu_{p_2}\), and in particular, if \(p \neq \frac{1}{2}\), \(\mu_p \perp \lambda\) where \(\lambda\) is the Lebesgue measure on \(\mathbb{R}\). \(\text{(10 marks)}\)
S1. The type of medical care a patient receives may vary with the age of the patient. Here is the joint probability mass function associated with data obtained in a large study of women who had a breast lump investigated, whether or not each woman received a mammogram and a biopsy when the lump was discovered. Define

\[ Y_1 = \begin{cases} 
0, & \text{if age < 65,} \\
1, & \text{if age } \geq 65.
\end{cases} \quad \text{and} \quad Y_2 = \begin{cases} 
0, & \text{if the test was not done,} \\
1, & \text{if the test was done.}
\end{cases} \]

Then the following table provides \( P_{ij} = P(Y_1 = i, Y_2 = j) \):

<table>
<thead>
<tr>
<th>( Y_1 )</th>
<th>( Y_2 )</th>
<th>( P_{ij} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.124</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.190</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.321</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.365</td>
</tr>
</tbody>
</table>

a) Find the marginal probability functions for \( Y_1 \) and \( Y_2 \). (6 marks)
b) Find the conditional probability function for \( Y_2 \) given \( Y_1 = 1 \). Are \( Y_1 \) and \( Y_2 \) independent? Why? (8 marks)
c) Were the tests omitted on older patients more or less frequently than would be the case if testing were independent of age? (6 marks)

S2. The velocities \( V \) of gas particles can be modeled by the Maxwell distribution, whose probability density function is given by

\[ f(v) = 4\pi \left( \frac{m}{2\pi KT} \right)^{3/2} v^2 e^{-v^2/(2mKT)}, \quad v > 0, \]

where \( m \) is the mass of the particle, \( K \) is Boltzmann’s constant, and \( T \) is the absolute temperature.

a) Find the mean velocity of these particles. (10 marks)
b) The kinetic energy of a particle is given by \((1/2)mV^2\). Find the mean kinetic energy for a particle. (10 marks)
S3. Suppose the probability that a person will suffer an adverse reaction from a medication is 0.001.

a) What is the probability that 2 or more will suffer an adverse reaction if the medication is administered on 1000 individuals. Use \( \lim_{n \to \infty} (1 - \frac{1}{n})^n = e^{-1} \) to provide an approximate answer. (6 marks)

b) Use a Poisson approximation to answer part (a). (6 marks)

c) Use a Normal approximation to answer part (a). Does a Normal approximation produce a reasonable result? Why? (8 marks)
S4. Suppose $X$ is a random variable with a probability distribution $P(X \in B)$, where $B \in \mathcal{B}$ is any Borel set. Let $X_1, X_2, \ldots, X_n$ be iid random variables from this distribution.

(a) For a fixed Borel set $A$, denote $\theta_A = P(X \in A)$. Using the above iid sample, find a non-trivial sufficient statistic for $\theta_A$. \hfill (6 marks)

(b) Based on the above iid sample, construct an approximate $100(1 - \alpha)\%$ confidence interval for $\theta_A$. [Use your choice of approximation, for large $n$.] \hfill (6 marks)

(c) Using your choice of approximation, for large $n$: explain in details how to test $H_0 : \theta_A = \theta_0$ versus $H_1 : \theta_A \neq \theta_0$, at a significance level $\alpha$. Note that $\theta_0$ is a pre-specified value. \hfill (8 marks)

S5. Let $X_1, X_2, \ldots, X_n$ be iid random variables with the probability density function

$$f(x; \theta) = \begin{cases} \frac{r}{\theta} x^{r-1} e^{-x^r/\theta} & ; y > 0 \\ 0 & ; \text{elsewhere} \end{cases}$$

where $\theta > 0$ is an unknown parameter, and $r$ is a known positive integer. This is called a **Weibull distribution**, which belongs to the exponential family. Note that $E(X_i^r) = \theta$.

(a) Find both the moment and maximum likelihood estimators of $\theta$. \hfill (8 marks)

(b) Find the UMVUE of $\theta$. \hfill (4 marks)

(c) Derive the likelihood ratio statistic, and construct an approximate $100(1 - \alpha)\%$ confidence interval for $\theta$. Draw a plot to explain how this technique is used to obtain the confidence interval. \hfill (8 marks)
S6. Let $X_1, X_2, \ldots, X_m$ be iid random variables from a Poisson distribution with parameter $\lambda_1$, and let $Y_1, Y_2, \ldots, Y_n$ be iid random variables from a Poisson distribution with parameter $\lambda_2$. All $m + n$ variables are independent, and $(\lambda_1, \lambda_2)$ are both unknown. Let $X = (X_1, X_2, \ldots, X_m)$ and $Y = (Y_1, Y_2, \ldots, Y_n)$.

(a) Find the UMVUE and the MLE of $\eta = (\lambda_1 - \lambda_2)^2$. (8 marks)

(b) Find the Cramér-Rao lower bound (CRLB) for the variance of any unbiased estimator of $\eta$. (6 marks)

(c) Consider using a likelihood ratio statistic for testing

$$H_0 : \eta = 0$$
$$H_1 : \eta \neq 0.$$ 

Design an approximate test of size $\alpha$ for testing these hypotheses. (6 marks)