Probability Section

1. Probability: Problem 1

Let \( c \in \mathbb{R} \) be a constant, and consider a random vector \((X, Y)\) taking values in \(\mathbb{R}^2\) with probability density function:
\[
    f(x, y) = \frac{1}{2\pi} \exp \left\{ \frac{2cxy - (1 + c^2)x^2 - y^2}{2} \right\}.
\]

(a) Determine the distribution of \( X \).

(b) Determine the conditional distribution of \( Y \) given \( X \).

(c) Finally, determine a necessary and sufficient condition on \( c \) so that \( X \) and \( Y \) are independent.

2. Probability: Problem 2

Fix \( N \in \mathbb{N} \). Let \( Q \) be a \( N \times N \) matrix satisfying \( q_{ij} \geq 0 \) for all \( i \neq j \) and \( q_{ii} = -\sum_{j \neq i} q_{ij} \) for all \( i \). Let \( X \) be a continuous-time Markov chain with rate matrix \( Q \). Consider \( \tau := \inf\{t \geq 0 : X_t \neq X_0\} \), the first time \( X \) changes its state. Recall that
\[
    \tau \sim \text{Exponential}(-q_{ii}) \quad \text{if} \ X_0 = i, \quad \text{and} \quad \mathbb{P}(X_\tau = j \mid X_0 = i) = -q_{ij}/q_{ii}.
\]

Suppose \( X_0 = i \). Given a function \( f : \{1, 2, ..., N\} \to \mathbb{R} \), this question aims to find a second-order expansion of \( \mathbb{E} \left[ \int_0^\varepsilon f(X_t) \, dt \right] \) in \( \varepsilon \), as \( \varepsilon \downarrow 0 \).

For each \( \varepsilon > 0 \), consider the events \( A, B, \) and \( C \) that on the interval \([0, \varepsilon]\), the state of \( X \) does not change, changes exactly once, and changes twice or more, respectively.
3. Probability: Problem 3

Let \( \{Z_n\}_{n \in \mathbb{N}} \) be i.i.d. random variables with \( Z_n \sim \text{Exponential}(1) \), and \( \Lambda(t) \) be a nonnegative function that is integrable on \([0, T]\) for all \( T > 0 \). Consider the sequence of random times:

\[
\tau_0 := 0, \quad \tau_n := \inf \left\{ t \geq \tau_{n-1} : \int_{\tau_{n-1}}^{t} \Lambda(s) ds \geq Z_n \right\}, \quad \forall n \in \mathbb{N}.
\]

Define the counting process \( N(t) \) by \( N(t) := n \) for \( t \in [\tau_n, \tau_{n+1}) \).

(a) Show that for any \( t > 0 \) and \( n \in \mathbb{N} \),

\[
\mathbb{P}(N(t) = n \mid \tau_n) = \exp \left( -\int_{\tau_n}^{t} \Lambda(s) ds \right) 1_{\{\tau_{n} < t\}}.
\]

For (b), (c), and (d) below, assume that \( \Lambda(t) \equiv \lambda \) for some constant \( \lambda > 0 \).

(b) What can you say about the sequence of random variables \( \{\tau_n - \tau_{n-1}\}_{n \in \mathbb{N}} \)? Justify!

(c) Based on part (b), what can you conclude about the process \( N(t) \)?

(d) In particular, what’s the distribution of \( \tau_n \) for \( n \geq \mathbb{N} \)? Use the distribution of \( \tau_n \) and part (a) to derive \( \mathbb{P}[N(t) = n] \) for all \( n \in \mathbb{N} \) and \( t > 0 \).
4. Statistics: Problem 4

Let $X_1, X_2, \ldots, X_n$ be a random sample from the Normal distribution with mean $\theta$ and variance 1, with $\theta \in \mathbb{R}$.

(a) Show that the best unbiased estimator of $\theta^2$ is given as $\bar{X}^2 - 1/n$.

(b) Calculate the variance of the estimator $\bar{X}^2 - 1/n$. (Recall that for $Y \sim \chi^2(k)$, $E(Y) = k$, and $V(Y) = 2k$.)

(c) Is the estimator $\bar{X}^2 - 1/n$ efficient? Explain.

(d) Find the maximum likelihood estimator (MLE) for $\theta^2$, and find its bias and variance.

(e) Which estimator, $\bar{X}^2 - 1/n$ or the MLE estimator, has a lower mean square error (MSE)? Recall that the MSE of an estimator is defined as $\text{MSE}(\hat{\theta}) = B(\hat{\theta})^2 + V(\hat{\theta})$ (bias squared plus variance).

5. Statistics: Problem 5

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Uniform($\theta, 2\theta$) distribution, where $\theta > 0$.

(a) Find the method of moments (MOM) estimator of $\theta$, $\hat{\theta}_{MOM}$. (Recall that MOM estimators are obtained by equating the sample moments with theoretical moments, and solving for $\theta$).

(b) Find the MLE of $\theta$, $\hat{\theta}_{MLE}$, and find a constant $k$ such that $E_{\theta}(k\hat{\theta}_{MLE}) = \theta$

(c) Which of these two estimators can be improved using sufficiency, and how?


Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, each with probability density function $f(x; \theta) = \theta xe^{-\theta x^2/2}$ for $x \geq 0$.

(a) Determine the distribution of $\theta \sum_{i=1}^n X_i^2 / 2$.

For the next two parts, leave your final answer in terms of the critical values of a suitable distribution.

(b) Using part (a), derive a 100(1 - $\alpha$)% lower-confidence bound for $\theta$.

(c) Using part (a), derive the uniformly most powerful (UMP) test of size $\alpha$ for testing $H_0 : \theta = 1$ versus $H_1 : \theta > 1$. 