Name: ____________________

Question:	1	2	3	Total
Points:	8	12	15	35
Score:				

E1 244 - Detection & Estimation Theory (2019) Midterm Exam

Instructions

- The total time for this test is 1.5 hours.
- Write your name at the top of this cover sheet.
- Attach your solution sheets to this cover sheet and return everything including rough work.
- No class notes, calculators or electronic aids are permitted.
- Academic dishonesty will not be tolerated.

Useful formulas and definitions:

- **Exponential probability distribution.** The exponential probability distribution with mean $\beta > 0$ is defined by the probability density function $p(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$ if $x \geq 0$ and $p(x) = 0$ if $x < 0$.
- **Gaussian probability distribution.** The Gaussian probability distribution $\mathcal{N}(\mu, \sigma^2)$ is defined by the probability density function $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$, $x \in \mathbb{R}$.
- $\log 2 = 0.69$.

1. **(8 points) Traffic Intensity Testing**
Vehicles arrive at the Yeshwanthpur traffic junction in Bengaluru with independent inter-arrival times distributed according to the exponential distribution with mean either $\frac{1}{2}$s. or 1s. depending on whether the type of traffic is 'heavy' or 'light', respectively. Suppose you measure the first 10 vehicle inter-arrival times (in s.) to be 0.06, 0.24, 0.17, 0.07, 1.80, 0.25, 1.37, 0.42, 0.48, 0.69. Assuming that both the types of traffic are equally likely *a priori*, what would you guess about the traffic type to minimize the probability of incorrect guessing and why?

2. **Signal Detection**
Consider the signal model

$$Y_k = \sqrt{\theta}\,(s_k R_k + N_k), \quad 1 \le k \le n,$$

where $s_1, \ldots, s_n$ is a known signal sequence, $\theta \ge 0$ is a constant and $R_1, \ldots, R_n$, $N_1, \ldots, N_n$ are i.i.d. $\mathcal{N}(0,1)$ random variables.

(a) **(6 points)** Consider the hypothesis pair

$$\begin{aligned} H_0 &: \theta = \frac{a}{2} \\ &\text{vs.} \\ H_1 &: \theta = a, \end{aligned}$$

where $a > 0$ is a known constant. Describe the structure of the Neyman-Pearson optimal detector for given false-alarm level $\alpha \in (0,1)$. Derive its receiver operating characteristic (ROC) in terms of the standard normal cdf Φ for the signal sequence $s_1 = \cdots = s_n = 1$.

(b) **(6 points)** Consider the hypothesis pair

$$\begin{aligned} H_0 &: \theta = \frac{a}{2} \\ &\text{vs.} \\ H_1 &: \theta > \frac{a}{2}. \end{aligned}$$

Does a Uniformly Most Powerful (UMP) detector exist (why/why not)? What about a Locally Most Powerful (LMP) detector (why/why not)?

3. **Independence Testing**
Suppose you want to determine whether two random variables A and B are independent, by observing pairs $(A_1, B_1), (A_2, B_2), \ldots$ independently sampled from a joint distribution. More specifically, consider testing the hypotheses

$$H_0 : A_i \sim \mathcal{N}(0,1), B_i = A_i + Z_i, Z_i \sim \mathcal{N}(0,1), A_i \perp\!\!\!\perp Z_i, i = 1, 2, \ldots, n$$
(not independent)

vs.

$$H_1 : A_i \sim \mathcal{N}(0,1), B_i \sim \mathcal{N}(0,2), A_i \perp\!\!\!\perp B_i, i = 1, 2, \ldots, n$$
(independent),

where $X \perp\!\!\!\perp Y$ denotes that X, Y are independent random variables, and any random variables with different indices i are assumed independent. Note that the marginal distributions of A_i and B_i are the same under both the hypotheses.

(a) **(3 points)** Find the log-likelihood ratio for the i-th observed pair (A_i, B_i) as a function of A_i and B_i.

(b) **(6 points)** Under each hypothesis, find the expectation of the log-likelihood ratio $\log L_n$ for the whole sequence of observations $(A_1, B_1), (A_2, B_2), \ldots, (A_n, B_n)$.

(c) **(6 points)** Consider the hypothesis test that outputs H_1 if the log-likelihood ratio ($\log L_n$) of n observations exceeds a given threshold $\eta \in \mathbb{R}$, and H_0 otherwise. Suppose the number of pairs of observations (n) is extremely large (think 'almost ∞'). Argue, using the law of large numbers, why you would expect this test to detect the true hypothesis reliably, i.e., with very low error.