Name: ______________________________

Question:	1	2	3	4	5	Total
Points:	6	9	10	6	9	40
Score:						

E1 244 - Detection & Estimation Theory 2018 - Final exam

Instructions

- Write your name on this question sheet.
- Attach your solution sheets to this question sheet and return everything.
- The total time for this test is 3 hours.
- No calculators or electronic aids are permitted.
- Academic dishonesty will not be tolerated.

1. Consider sequentially testing the hypotheses

$$\begin{aligned} H_0 &: Y_k = N_k,\ \ k = 1, 2, \ldots, \quad \textit{vs.} \\ H_1 &: Y_k = S_k + N_k,\ \ k = 1, 2, \ldots, \end{aligned}$$

where $N_1, N_2, \ldots$, are iid $\mathcal{N}(0, 1)$ random variables and $S_1, S_2, \ldots$, are iid $\mathcal{N}(0, 2)$ random variables, with the N's being independent of the S's.

(a) **(3 points)** Design a sequential probability ratio test, SPRT(a, b), such that the test has maximum error probability $\max(P_F, P_M) \approx 0.01$, using Wald's approximations.

(b) **(3 points)** Find (approximately) the expected stopping times of the test you designed, under both hypotheses. You may use the fact that the Kullback-Leibler divergence between a $\mathcal{N}(\mu_1, \sigma_1^2)$ distribution and a $\mathcal{N}(\mu_2, \sigma_2^2)$ distribution is

$$\log(\sigma_2/\sigma_1) + [(\mu_1 - \mu_2)^2 + \sigma_1^2 - \sigma_2^2]/(2\sigma_2^2).$$

2. Suppose $X_1, \ldots, X_n$ are iid samples from the distribution with probability mass function $p_\theta(x) = e^{-\theta}\left(1 - e^{-\theta}\right)^{x-1}$, $x = 1, 2, 3, \ldots$, with $\theta \in (0, \infty)$.

(a) **(4 points)** Find a scalar sufficient statistic for θ. Can you show that it is minimal sufficient?

(b) **(2 points)** Find a method of moments estimator for θ. (You can use $\sum_{x=1}^{\infty} x\lambda^x = \lambda/(1-\lambda)^2$ for $\lambda \in (0, 1)$.)

(c) **(3 points)** Now assume that θ has the prior probability density $\pi(\theta) = \beta e^{-\beta\theta}$ for $\theta \in (0, \infty)$ and $\pi(\theta) = 0$ otherwise, for some known positive constant β. Find the mode of the posterior distribution of θ (the maximum a posteriori estimator) given $X_1, \ldots, X_n$.

3. You are in charge of estimating the ambient temperature in the IISc campus. To do this, you can read (i.e., make a single measurement) from a temperature sensor on campus (#1). You also have access to two additional temperature sensors, at Malleshwaram (#2) and Hebbal (#3). But due to travel constraints, you can only choose to read from one of them, and output a linear function of the two sensors (IISc and one other) as your final estimate.

   Suppose that each location's temperature—X_1 (IISc), X_2 (Malleshwaram) and X_3 (Hebbal)—is a random variable with mean 0 (in a suitable scale), and that the covariance matrix of $[X_1, X_2, X_3]$ has the rows $[2, -1, 2]$, $[-1, 6, 3]$ and $[2, 3, 5]$ in order. Moreover, each sensor adds independent noise of mean 0 and variance 1 to the true temperature before outputting its reading.

   (a) **(4 points)** Compute the best (in mean-square error) linear estimator of X_1 using sensors #1 and #2. What is its mean-square error?

   (b) **(4 points)** Compute the best (in mean-square error) linear estimator of X_1 using sensors #1 and #3. What is its mean-square error?

   (c) **(2 points)** Which of the above combinations would you prefer?

4. Consider testing the following hypotheses using independent observations $X_1, \ldots, X_n, Y_1, \ldots, Y_n$.

$$
\begin{aligned}
H_0 : \quad & X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(0, 1) \\
& Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \mathcal{N}(0, 1), \qquad \textit{vs.} \\
H_1 : \quad & X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(0, 1) \\
& Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1), \quad \mu > 0.
\end{aligned}
$$

   (a) **(3 points)** What is an optimal level-α test of H_0 versus H_1 when μ equals a <u>fixed</u> $\mu_0 \in (0, \infty)$?

   (b) **(3 points)** Does there exist a Uniformly Most Powerful (UMP) test of H_0 versus H_1? Why/why not?

5. Consider the following model for a medical patient's treatment outcome, as a function of the dosage of an administered drug. If the dosage level is x, then the probability that the patient is cured is p, where p and x follow the relationship[1]

$$\log \frac{p}{1-p} = \theta_0 + \theta_1 x,$$

   with $\theta_0, \theta_1 \in \mathbb{R}$ being unknown parameters.

   Suppose a population of patients is divided into m groups of size n each, and each patient in group i is treated with (known) dosage level x_i, for $1 \le i \le m$. At the end of the treatment period, Y_i out of the n patients from group i are observed to be successfully cured, in every group $1 \le i \le m$. Assuming that treatment outcomes for each patient are independent,

   (a) **(3 points)** Write down the probability of seeing the data $(Y_1, \ldots, Y_m) = (y_1, \ldots, y_m)$ under a pair of model parameters (θ_0, θ_1).

[1] also called the *linear logistic* model in statistics/machine learning

(b) **(3 points)** Can you find a (non-trivial) sufficient statistic for (θ_0, θ_1)?

(c) **(3 points)** Describe, clearly but briefly, how you would find the maximum likelihood estimate of (θ_0, θ_1) given the observations $Y_1, \ldots, Y_n$ (a closed-form solution is not expected).
(This is called Logistic regression.)