Name: ______________________________

Question:	1	2	3	4	Total
Points:	10	10	5	10	35
Score:					

E1 244 - Detection & Estimation Theory - Mid Term exam

Instructions

- The total time for this test is 1.5 hours.
- Write your name on this question sheet.
- Attach your solution sheets to this question sheet and return everything.
- No calculators or electronic aids are permitted.
- Academic dishonesty will not be tolerated.

Useful formulas and definitions:

- **Gaussian probability distribution.** The Gaussian probability distribution $\mathcal{N}(\mu, \sigma^2)$ is defined by the probability density function $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}}$, $x \in \mathbb{R}$.
- **Bayesian composite hypothesis testing.** For the composite hypothesis test $H_0 : Y \sim \mathbb{P}_\theta$, $\theta \in \Theta_0$ vs. $H_1 : Y \sim \mathbb{P}_\theta$, $\theta \in \Theta_1$, with Θ_0 and Θ_1 disjoint, let a prior distribution be π on $\Theta \equiv \Theta_0 \cup \Theta_1$, and the costs be $C[i, \theta]$ for each hypothesis $i \in \{0, 1\}$ and parameter θ. The Bayes risk of a decision rule $\delta : \Gamma \to \{0, 1\}$ is defined to be the quantity

$$\int_\Theta \left(\sum_{i=0}^{1} \mathbb{P}_\theta[\delta(Y) = i]\, C[i, \theta] \right) \pi(\theta)\, d\theta.$$

- **Series sums.** $\sum_{i=1}^n i = \frac{n(n+1)}{2}$, $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$

1. **Testing for mixtures**
Given two probability distributions P_1 and P_2, here is how random variables $Y_1, \ldots, Y_n$ are defined to be generated by the mixture of P_1 and P_2: First, draw $Z \sim$ Bernoulli(1/2). If $Z = 0$, then generate $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} P_1$, and if $Z = 1$, then generate $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} P_2$.
Consider testing if $H_0 : Y_1, \ldots, Y_n$ are generated iid Bernoulli(1/2), or $H_1 : Y_1, \ldots, Y_n$ are generated by the mixture of Bernoulli(1/4) and Bernoulli(3/4).
   (a) **(5 points)** Write down a form for the Bayes-optimal test.

   Solution: For $y \equiv (y_1, \ldots, y_n)$, let $s(y) = \sum_{i=1}^n Y_i$. The Bayes-optimal test is the test that compares a threshold with the likelihood ratio

$$L(y) = \frac{p_1(y)}{p_0(y)} = \frac{\frac{1}{2}\left(\frac{1}{4}\right)^{s(y)}\left(\frac{3}{4}\right)^{n-s(y)} + \frac{1}{2}\left(\frac{3}{4}\right)^{s(y)}\left(\frac{1}{4}\right)^{n-s(y)}}{\left(\frac{1}{2}\right)^{s(y)}\left(\frac{1}{2}\right)^{n-s(y)}}$$
$$= \frac{3^{s(y)} + 3^{n-s(y)}}{2^{n+1}},$$

   or equivalently, the test that compares the statistic $3^{s(y)} + 3^{n-s(y)}$ to a threshold.

   (b) **(5 points)** Argue briefly (in words) what happens to the test statistic in your answer above, under each hypothesis, when n is large (i.e., why do you expect the test to work?).

   Solution: Under H_0, the statistic $3^{s(y)} + 3^{n-s(y)} \approx 2 \times 3^{n/2}$, whereas under H_1, $3^{s(y)} + 3^{n-s(y)} \approx 3^{n/4} + 3^{3n/4}$, which is much larger than $2 \times 3^{n/2}$.
   Alternative solution: Assuming uniform priors, the optimal Bayes test is to return H_1 if and only if the likelihood ratio

$$\frac{3^{s(y)} + 3^{n-s(y)}}{2^{n+1}} \geq 1.$$

   For large n, under H_0, the left hand side above is approximately

$$\frac{2 \cdot 3^{n/2}}{2^{n+1}} = \left(\frac{\sqrt{3}}{2}\right)^n \to 0.$$

   On the other hand, under H_1, the right hand side is approximately

$$\frac{3^{n/4} + 3^{3n/4}}{2^{n+1}} \approx \frac{3^{3n/4}}{2^{n+1}} \to \infty,$$

   so the test indeed works.

2. Consider the hypothesis test

$$H_0 : Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Binomial}(m, \theta_0), \quad \text{vs.}$$
$$H_1 : Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Binomial}(m, \theta), \quad \theta > \theta_0,$$

where $\theta_0 \in (0, 1)$ and m are assumed to be known constants.

(a) **(5 points)** For a <u>fixed</u> $\theta > \theta_0$, what is the general form of the Neyman-Pearson optimal test?

Solution: The likelihood ratio of observations $y \equiv (y_1, \ldots, y_n)$ is

$$\frac{p_1(y)}{p_0(y)} = \prod_{k=1}^{n} \frac{\theta^{y_k}(1-\theta)^{m-y_k}}{\theta_0^{y_k}(1-\theta_0)^{m-y_k}} = \left(\frac{\theta}{\theta_0}\right)^{s(y)} \left(\frac{1-\theta}{1-\theta_0}\right)^{mn-s(y)}$$
$$= \left(\frac{\theta(1-\theta_0)}{\theta_0(1-\theta)}\right)^{s(y)} \left(\frac{1-\theta}{1-\theta_0}\right)^{mn}$$

where $s(y) = \sum_{k=1}^{n} y_k$, and the N-P optimal test is to compare this to a suitable threshold. This is equivalent to comparing $s(y) = \sum_{k=1}^{n} y_k$ to a threshold.

(b) **(5 points)** Does there exist a Uniformly Most Powerful (UMP) test for θ_0 vs. $\theta > \theta_0$?

Solution: Yes. The N-P optimal test for level α is the same for testing θ_0 vs. θ for each $\theta > \theta_0$, so it follows that this common NP-optimal test is indeed a UMP test for the entire family (θ_0, ∞).

3. **(5 points)** Suppose you are looking to detect the linear signal $s_k = \beta k$, $k = 1, \ldots, n$, in additive, iid Gaussian noise of mean 0 and variance σ^2, where $\beta\ \sigma^2$ are known. Determine the Neyman-Pearson optimal detector and its detection-vs-false alarm probability performance (in terms of the standard normal cdf Φ).

Solution: This is the hypothesis test

$$H_0 : \underline{Y} \sim \mathcal{N}(0, \sigma^2 I)$$
$$\text{vs.}$$
$$H_1 : \underline{Y} \sim \mathcal{N}(\underline{s}, \sigma^2 I).$$

The N-P optimal detector for this problem is the matched filter $\sum_{k=1}^{n} s_k Y_k \geq \tau$ or equivalently, $\sum_{k=1}^{n} k Y_k \geq \tau'$.

If the desired false alarm probability is α, then we must have that the optimum detection probability is

$$1 - \Phi(\Phi^{-1}(1-\alpha) - d) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{\frac{\underline{s}^T \underline{s}}{\sigma^2}}\right)$$
$$= 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{\frac{\sum_{k=1}^{n} \beta^2 k^2}{\sigma^2}}\right)$$
$$= 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{\frac{\beta^2 n(n+1)(2n+1)}{6\sigma^2}}\right).$$

4. **Composite hypothesis testing**

Consider the hypothesis test

$$H_0 : X \sim \mathcal{N}(\theta, 1), \theta < 0$$
$$\text{vs.}$$
$$H_1 : X \sim \mathcal{N}(\theta, 1), \theta \geq 0,$$

where θ is assumed to have the $\mathcal{N}(1, 1)$ prior distribution. Let the costs be uniform for $\theta < 0$ and $\theta \geq 0$, i.e., for each $i \in \{0, 1\}$ and $\theta \in \mathbb{R}$,

$$C[i, \theta] = \begin{cases} 0, & \text{if } (i = 0, \theta < 0) \text{ or } (i = 1, \theta \geq 0) \\ 1, & \text{otherwise.} \end{cases}$$

(a) **(5 points)** What is the posterior probability distribution of θ given the observation $X = x$?

Solution: We have $X \sim \mathcal{N}(\theta, 1)$ where $\theta \sim \mathcal{N}(1, 1)$, so the posterior probability (density) of θ given $X = x$ is[1], by Bayes' rule,

$$\begin{aligned}
\mathbb{P}\left[\theta \mid X = x\right] &= \frac{\mathbb{P}\left[X = x \mid \theta\right] \, \mathbb{P}\left[\theta\right]}{\mathbb{P}\left[X = x\right]} \\
&= \frac{e^{-(x-\theta)^2/2} \cdot e^{-(\theta-1)^2/2}}{2\pi \cdot \mathbb{P}\left[X = x\right]} \\
&= \frac{1}{2\pi \cdot \mathbb{P}\left[X = x\right]} \, e^{-\left[\theta^2 - \theta(x+1) + \frac{x^2+1}{2}\right]} \\
&= \frac{1}{2\pi \cdot \mathbb{P}\left[X = x\right]} \, e^{-\left[\left(\theta - \frac{x+1}{2}\right)^2 - \frac{(x+1)^2}{4} + \frac{x^2+1}{2}\right]} \\
&= \frac{1}{2\pi \cdot \mathbb{P}\left[X = x\right]} \, g(x) \, e^{-\left(\theta - \frac{x+1}{2}\right)^2},
\end{aligned}$$

for some quantity $g(x)$ not dependent on θ. But this means that *the posterior distribution of θ, given $X = x$, is Gaussian with mean $\frac{x+1}{2}$ and variance $\frac{1}{2}$.*

(b) **(5 points)** What is the Bayes optimal decision rule for the observation $X = x$?

Solution: We know, from the theory of Bayes composite hypothesis testing for uniform costs, that the Bayes optimal decision rule can be viewed as (a) first update the prior on θ to the posterior distribution of θ given $X = x$, and then (b) declare the hypothesis that has the least expected cost under the posterior distribution. By part (a) above, given that $X = x$, θ follows a $\mathcal{N}((x + 1)/2, 1/2)$ distribution. Under this distribution for θ, declaring H_0 would incur expected cost

$$\int_0^\infty \frac{1}{\frac{1}{\sqrt{2}}\sqrt{2\pi}} \, e^{-\left(\theta - \frac{x+1}{2}\right)^2} d\theta,$$

whereas declaring H_1 would incur expected cost

$$\int_{-\infty}^0 \frac{1}{\frac{1}{\sqrt{2}}\sqrt{2\pi}} \, e^{-\left(\theta - \frac{x+1}{2}\right)^2} d\theta.$$

Since the Gaussian pdf is symmetric about its mean, it follows that the Bayes optimal rule is:

Declare H_0 if $(1+x)/2 < 0$ and H_1 if $(1+x)/2 \geq 0$, i.e., Declare H_0 if and only if $x < -1$.

Alternative solution (without using part (a)). A Bayes' optimal composite hypothesis test is to declare H_1 if and only if

$$\begin{aligned}
& \frac{\mathbb{P}\left[X = x \mid \theta \geq 0\right]}{\mathbb{P}\left[X = x \mid \theta < 0\right]} \geq \frac{\pi_0}{\pi_1} := \frac{\mathbb{P}\left[\theta < 0\right]}{\mathbb{P}\left[\theta \geq 0\right]} \\
\Leftrightarrow \quad & \frac{\mathbb{P}\left[X = x, \theta \geq 0\right]}{\mathbb{P}\left[X = x, \theta < 0\right]} \geq 1 \\
\Leftrightarrow \quad & \int_0^\infty e^{-(x-\theta)^2/2} \cdot e^{-(\theta-1)^2/2} d\theta \geq \int_{-\infty}^0 e^{-(x-\theta)^2/2} \cdot e^{-(\theta-1)^2/2} d\theta \\
\Leftrightarrow \quad & \int_0^\infty e^{-\left[\theta^2 - \theta(x+1) + \frac{x^2+1}{2}\right]} d\theta \geq \int_{-\infty}^0 e^{-\left[\theta^2 - \theta(x+1) + \frac{x^2+1}{2}\right]} d\theta \\
\Leftrightarrow \quad & \int_0^\infty e^{-\left(\theta - \frac{x+1}{2}\right)^2} d\theta \geq \int_{-\infty}^0 e^{-\left(\theta - \frac{x+1}{2}\right)^2} d\theta \\
\Leftrightarrow \quad & \frac{x+1}{2} \geq 0 \quad \text{(since the pdf of a Gaussian is symmetric about its mean)} \\
\Leftrightarrow \quad & x \geq -1.
\end{aligned}$$