UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Problem Set 12
Spring 2020

1. **Statistical Estimation**

   Given $X \in \{0, 1\}$, the random variable Y is exponentially distributed with rate $3X + 1$.

   (a) Assume $\mathbb{P}(X = 1) = p \in (0, 1)$ and $\mathbb{P}(X = 0) = 1 - p$. Find the MAP estimate of X given Y.

   (b) Find the MLE of X given Y.

2. **Laplace Prior & ℓ^1-Regularization**

   Suppose you draw n i.i.d. data points $(x_1, y_1), \dots, (x_n, y_n)$, where n is a positive integer and the true relationship is $Y = WX + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. (That is, Y has a linear dependence on X, with additive Gaussian noise.) Further suppose that W has a prior distribution with density

   $$f_W(w) = \frac{1}{2\beta} \mathrm{e}^{-|w|/\beta}, \qquad \beta > 0.$$

   (This is known as the **Laplace distribution**.) Show that finding the MAP estimate of W given the data points $\{(x_i, y_i) : i = 1, \dots, n\}$ is equivalent to minimizing the cost function

   $$J(w) = \sum_{i=1}^{n} (y_i - wx_i)^2 + \lambda |w|$$

   (you should determine what λ is). This is interpreted as a one-dimensional ℓ^1-regularized least-squares criterion, also known as LASSO.

3. **Hypothesis Testing for Uniform Distribution**

   Assume that

   - If $X = 0$, then $Y \sim \text{Uniform}[-1, 1]$.
   - If $X = 1$, then $Y \sim \text{Uniform}[0, 2]$.

Using the Neyman-Pearson formulation of hypothesis testing, find the optimal randomized *decision rule* $r : [-1, 2] \to \{0, 1\}$ with respect to the criterion

$$\min_{\text{randomized } r:[-1,2]\to\{0,1\}} \mathbb{P}\big(r(Y) = 0 \bigm| X = 1\big)$$
$$\text{s.t. } \mathbb{P}\big(r(Y) = 1 \bigm| X = 0\big) \le \beta,$$

where $\beta \in [0, 1]$ is a given upper bound on the false positive probability.

4. **Gaussian Hypothesis Testing**

   Consider a hypothesis testing problem that if $X = 0$, you observe a sample of $\mathcal{N}(\mu_0, \sigma^2)$, and if $X = 1$, you observe a sample of $\mathcal{N}(\mu_1, \sigma^2)$, where $\mu_0, \mu_1 \in \mathbb{R}$, $\sigma^2 > 0$. Find the Neyman-Pearson test for false alarm $\alpha \in (0, 1)$, that is, $\mathbb{P}(\hat{X} = 1 \mid X = 0) \le \alpha$.

5. **BSC Hypothesis Testing**

   Consider a BSC with some error probability $\epsilon \in [0.1, 0.5)$. Given n inputs and outputs (x_i, y_i) of the BSC, solve a hypothesis problem to detect that $\epsilon > 0.1$ with a probability of false alarm at most equal to 0.05. Assume that n is very large and use the CLT.

   Hint: The null hypothesis is $\epsilon = 0.1$. The alternate hypothesis is $\epsilon > 0.1$, which is a **composite hypothesis** (this means that under the alternate hypothesis, the probability distribution of the observation is not completely determined; compare this to a **simple hypothesis** such as $\epsilon = 0.3$, which *does* completely determine the probability distribution of the observation). The Neyman-Pearson Lemma we learned in class applies for the case of a simple null hypothesis and a simple alternate hypothesis, so it does not directly apply here.

   To fix this, fix some specific $\epsilon' > 0.1$ and use the Neyman-Pearson Lemma to find the optimal hypothesis test for the hypotheses $\epsilon = 0.1$ vs. $\epsilon = \epsilon'$. Then, argue that the optimal decision rule does not depend on the specific choice of ϵ'; thus, the decision rule you derive will be *simultaneously* optimal for testing $\epsilon = 0.1$ vs. $\epsilon = \epsilon'$ for all $\epsilon' > 0.1$.

6. **Voltage MAP**

   You are trying to detect whether voltage V_1 or voltage V_2 was sent over a channel with independent Gaussian noise $Z \sim N(V_3, \sigma^2)$. Assume that both voltages are equally likely to be sent.

   (a) Derive the MAP detector for this channel.

(b) Using the Gaussian Q-function, determine the average error probability for the MAP detector.

(c) Suppose that the average transmit energy is $(V_1^2 + V_2^2)/2$ and that the average transmit energy is constrained such that it cannot be more than $E > 0$. What voltage levels V_1, V_2 should you choose to meet this energy constraint but still minimize the average error probability?

7. [Bonus] p-Value

The bonus question is just for fun. You are not required to submit the bonus question, but do give it a try and write down your progress.

Let us define what the p-value of a hypothesis test is. Given an observation Y and a constraint of β on the PFA, the Neyman-Pearson rule will either declare that the alternate hypothesis is true or not. The constraint on the PFA controls the trade-off between declaring the alternate hypothesis to be true when it is not (false alarm), and declaring the alternate hypothesis to be true when it is (correct detection). Therefore, for very high values of β, the hypothesis test will declare that the alternate hypothesis is true, and for very low values of β, the hypothesis test will declare that the null hypothesis is true.

(Intuitively, the smaller the value of β, the more conservative the resulting hypothesis test is, i.e., it will be more reluctant to declare that the alternate hypothesis is true.)

The **p-value** of the observation is the smallest value of β such that the alternate hypothesis is declared true.

Think about this carefully, and explain why the p-value is *not* the probability that the alternate hypothesis is true.