ECE 561: Detection and Estimation Theory
Spring 2009
Midterm Exam
Issued: March 15th, 2009

March 15, 2009

GENERAL COMMENTS:

You are NOT allowed to collaborate on the midterm problem. Nevertheless, you are allowed to read/check the Internet for theoretical (mathematical) results that may help you address the problem. But please refrain from reporting other people's work and from trying to track down solutions to the problem online. You are also allowed to use the text and all the handouts for the class for your work.

The exam will test your knowledge about the material covered by March 11th, 2009. Some questions ask you to generalize or extend some of the analytical tools we studied in class, and hence may not have one unique solution, and there may be many different approaches that one can take to address them.

- 1) Uniformly most powerful tests

  Consider the hypothesis testing problem

  $H_0 : Y_k = \mathcal{N}_k, \quad k = 1, 2,$
  $H_1 : Y_k = \theta\, s_k + \mathcal{N}_k, k = 1, 2,$

  where $\mathcal{N}_{1,2}$ are independent standard Gaussian random variables, and $s_1 = 1$, $s_2 = -1$. The parameter θ is deterministic, but unknown parameter that takes on only two values, -1 and $+1$.

  - Is there a UMP test for this problem? If so, find its level α. If not, explain why such a test does not exist.
  - Show that an α-level GLRT for this problem is given by:

    $$d_{GLRT} = 1, \ \ \text{if} \, |y_1 - y_2| \geq \eta_\alpha,$$

    and zero otherwise.

– Show that the probability of detection for the GLRT test in the previous bullet is actually independent on θ. Express the probability of detection in terms of the parameter η_α.

- 2) Chernoff information and Kulback-Leibler Distance

  The Chernoff information for two distributions f_0 and f_1 is given by

$$C(f_0, f_1) = -\min_{0 \le s \le 1} \ln \left[\int f_1(y)^s \, f_0^{1-s} dy \right]$$

  For $s \in [0, 1]$, define the distribution:

$$f_s(y) = \frac{f_1(y)^s \, f_0(y)^{1-s} \, dy}{\int f_1(x)^s \, f_0(x)^{1-s} \, dx}.$$

  Show that the optimizing value of s, s^*, in the definition of $C(f_0, f_1)$ satisfies the equation:

$$D(f_{s^*}||f_0) = D(f_{s^*}||f_1).$$

  Furthermore, show that the optimal value of the Chernoff information equals to the two KL divergences above.

- 3) Azuma's Inequality for Martingales

  a) Prove the following claim.
  Let $\{X_k\}$, $k = 0, 1, 2, \ldots$ be a martingale sequence, such that almost surely, one has $|X_k - X_{k-1}| < c_k$, where c_k's are constants. In other words, the martingale sequence has bounded increments. Then

$$P\{X_n - X_0 \ge t\} \le \exp\left(\frac{-t^2}{2 \sum_{k=1}^n c_k^2} \right).$$

  b) Give the Azuma inequality for $X_k = \sum_{i=1}^k Y_i$, where the Y_i's are i.i.d. symmetric (i.e., variables that have the same probability for m and $-m$), integer-valued, bounded random variables. What does this bound assert about the growth rate of the sum?

  Can you use the above described bounds in the analysis of SPRTs?

- 4) (Difficult; Make out of it as much as you can) KL divergence between Binomial and Poisson distributions

  You learned in elementary probability classes that under appropriate conditions on the parameters of the distributions, the Poisson distribution is a good approximation for the Binomial distribution. Try to show that this can be mathematically expressed as follows: Given a Binomial distribution Q with parameters p and n, and a Poisson distribution P with parameter $\lambda = np$, it holds that

$$D(Q||P) \leq \left(\frac{1}{4} + \frac{np^2}{3} + \frac{p}{2} + \frac{1}{4n} \right) p^2.$$

Based on this result, what can you say about the asymptotic performance of simple detection schemes that try to discriminate between a Poisson and Binomial distribution? Focus on the Neyman-Pearson test, and the analysis from Cover+Thomas we did in class.

- 5) Composite hypothesis testing and SPRTs

In class, we discussed the composite hypothesis test for one simple and one composite hypothesis. We also finished the discussion on sequential hypothesis testing via Wald's zero-overshoot theory. How would you generalize the SPRT for the composite hypothesis scenario described above?