# CS281A/STAT241A - Homework II

September 11, 2014

**This assignment is due at the beginning of class on September 18.**

**1**. Suppose $x$ and $y$ are scalar random variables. Their joint density, depicted below in Figure 2-1, is constant in the shaded region and $0$ elsewhere.

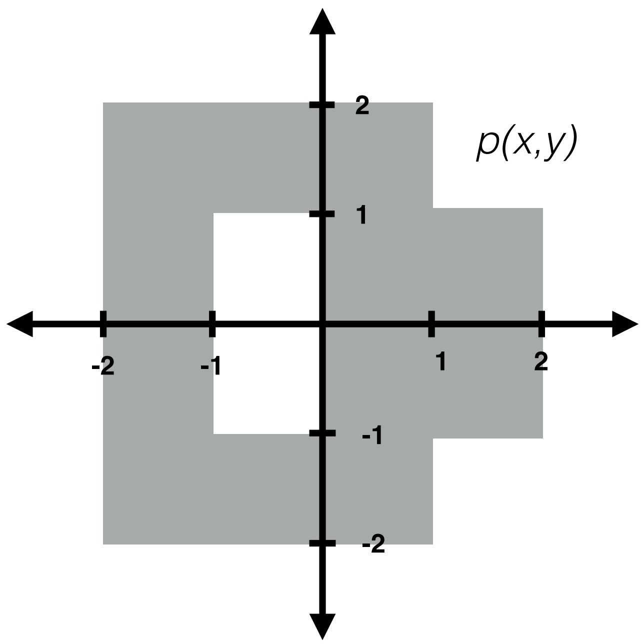


**Figure 1:** Figure 2-1

We want to decide if $x$ is less than or equal to zero after observing $y$.

(a) Determine the probabilities $\Pr[H_0] := \Pr[x \leq 0]$ and $\Pr[H_1] := \Pr[x > 0]$.

(b) Make fully labelled sketches of $p(y|H_0)$ and $p(y|H_1)$.

(c) Construct a rule $\hat{H}(y)$ deciding between $H_0$ and $H_1$ given an observation $y$ that minimizes the probability of error. Specify for which values of $y$ your rule chooses $H_1$ and for which it chooses $H_0$.

(d) What is the resulting probability of error?

(e) In the $(P_D, P_F)$ plane, sketch the operating characteristic of the likelihood ratio test for this problem.

(f) Is the point $(\frac{2}{3}, \frac{5}{6})$ achievable by *some* decision rule? If so, describe a test that achieves this value. If not, explain.

**2**. Suppose we are trying to determine if a professional athlete is using a performance enhancing drug. Assume that the maximum allowable concentration is $L$ parts per million in a blood sample, and the true concentration over all professional athletes is a uniform random number in the interval $[0, C]$ (in parts per million). Moreover, assume that our lab test is noisy, returning the true concentration plus a uniform error $\delta \in [-\Delta, \Delta]$. That is, when we perform a drug test, we measure

$$y = c + \delta$$

where $c \sim \text{unif}([0, C])$ and $\delta \sim \text{unif}([-\Delta, \Delta])$. We want to decide if $c$ is larger than $L$.

(a) Find the minimum probability of error detector and compute the associated probability of error.

(b) Suppose that we don't know the *a priori* distribution of $c$ and choose to use a maximum likelihood detector. Find the ML detector and the associated probability of error.

(c) Suppose we take two samples from the same player. That is we observe $y_1 = c + \delta_1$ and $y_2 = c + \delta_2$ and $\delta_1$ and $\delta_2$ are independent random variables distributed as $\text{unif}([-\Delta, \Delta])$. How do your answers to (a) and (b) change?

**3**. In a binary hypothesis testing problem, let $p_j$ denote that the probability that hypothesis $H_j$ is true. Recall that the minimax detection problem consists of solving the optimization problem

$$\text{minimize}_f \max_{p_0, p_1} \mathbb{E}[\ell(f(y), H)] \,.$$

That is, we seek to find the best decision rule for the least favorable prior probabilities $(p_0, p_1)$. Let $C_{ij}$ denote the loss $\ell(i, j)$.

(a) Show that the minimax detection problem is equivalent to the optimization problem

$$\begin{array}{ll} \text{minimize}_{f,t} & t \\ \text{subject to} & C_{00}(1 - P_F) + C_{10} P_F \leq t \\ & C_{01}(1 - P_D) + C_{11} P_D \leq t \end{array} \,.$$

(b) Using a Lagrange multiplier argument, show that we can lower bound the minimax risk by

$$\max_{\lambda, \mu \geq 0} \min_{f,t} t + \lambda(C_{00}(1 - P_F) + C_{10} P_F - t) + \mu(C_{01}(1 - P_D) + C_{11} P_D - t) \,.$$

(c) For fixed $\lambda$ and $\mu$, show that the optimum assignment $f$ is given by a Likelihood Ratio Test. What is the threshold for choosing between $H_0$ and $H_1$?

(d) Show that we can choose $\lambda$ and $\mu$ so that the lower bound is matched by a feasible upper bound. That is, find an assignment of $f$ and $t$ such that they are feasible for the original problem and achieve the minimum in the lower bound.

4. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Prove that $g$ is convex if and only if

$$g(z) \geq g(x) + \nabla g(x)^T (z - x) \, .$$

for all $x$ an $z$.

5. A density $p(x)$ is said to be *log-concave* if $\log p(x)$ is a concave function. Show the following popular probability distributions are log concave.

   (a) The *multivariate Gaussian distribution*, $\mathcal{N}(\mu, \Lambda)$, for any mean parameter $\mu$ and covariance $\Lambda$.

   (b) The *gamma density*, defined by

$$p(x) = \frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} \, .$$

   where $\Gamma$ is the ordinary Gamma function, $\lambda \geq 1$, and $\alpha > 0$.

   (c) The *Dirichlet density* on the unit simplex:

$$p(x) = \frac{\Gamma\left(\sum_{i=1}^{n+1} \lambda_i\right)}{\Gamma(\lambda_1) \cdots \Gamma(\lambda_{n+1})} x_1^{\lambda_1 - 1} \ldots x_n^{\lambda_{n-1}} \left(1 - \sum_{i=1}^{n} x_i\right)^{\lambda_{n+1}-1} \, .$$

   where $x$ is restricted to be nonnegative and have sum less than $1$. Here, the parameter $\lambda$ has all components greater than or equal to $1$.

*It may be useful to consult Section 3.5 in Boyd and Vandenberghe for ideas on how to solve this problem.*

6. Suppose $x$ and $y$ are scalar random variables. Their joint density, depicted below in Figure 2-2, is constant in the shaded region and $0$ elsewhere.

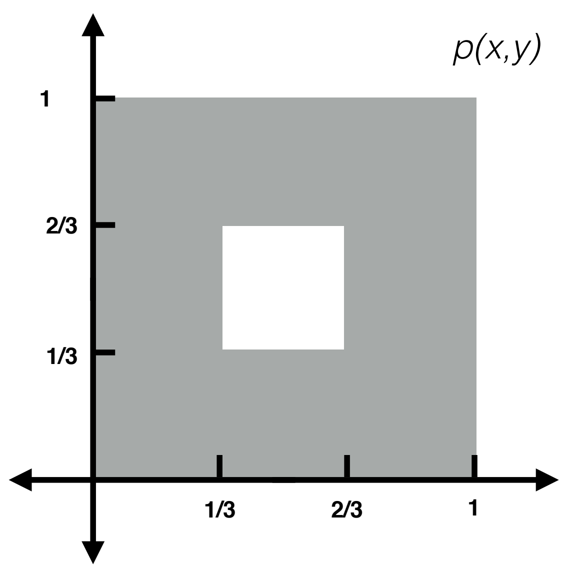


**Figure 2:** Figure 2-2

(a) Determine $\hat{x}_{\text{BLS}}(y)$, the Bayes least-squares estimate of $x$ given the observation $y$.

(b) Are $x$ and $y$ uncorrelated? Are $x$ and $y$ statistically independent? Explain your reasoning.

**7**. The number of times you check your Facebook page in a particular hour in the day is a Poisson random variable with mean $\alpha b$ where $\alpha > 0$ is a universal constant and $b$ quantifies how bored you are. Let $y_k$ denote the number of times you check Facebook between $k$ and $k+1$ o'clock. Conditioned on your boredom variable $b$, the $y_k$ are statistically independent random variables. For each $9 \leq i \leq 17$,

$$\Pr[y_i = k|b] = \frac{(\alpha b)^k e^{-\alpha b}}{k!} \qquad i = 9, 10, 11, \ldots$$

(a) Suppose $p(b) = B^{-1}e^{-b/B}$ for $b \geq 0$ and $p(b) = 0$ for $b < 0$. Determine $\hat{b}_{\text{BLS}}$, the Bayes least-squares estimate of $b$, and the resulting mean-square estimation error. You may find the following identity useful:

$$\int_0^\infty x^k e^{-ax} dx = \frac{k!}{a^{k+1}}\,.$$

(b) As a check on your answer to part (a), verify that when $B$ tends to infinity, the estimate in part (a) reduces to

$$\hat{b}_{\text{BLS}} \to \frac{1}{9\alpha}\left(1 + \sum_{i=9}^{17} y_i\right)\,.$$