Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
6.437 Inference and Information
Spring 2015

Problem Set 1

Issued: Tuesday, February 3, 2015 **Due:** Thursday, February 12, 2015

Problem 1.1

(a) Let x and y be independent identically distributed geometric random variables with parameter p. That is, x and y have common probability mass function

$$p(n) = p(1-p)^n, \quad n = 0, 1, \ldots .$$

Let $\mathsf{s} = \mathsf{x} + \mathsf{y}$.

(i) Find and sketch $p_{\mathsf{s}}(s)$.

(ii) Find and sketch $p_{\mathsf{x}|\mathsf{s}}(x|s)$ vs. x, with s viewed as a known integer parameter.

(iii) Find the conditional mean of x given $\mathsf{s} = s$, which is defined as

$$\mathbb{E}[\mathsf{x} \mid \mathsf{s} = s] = \sum_{x=-\infty}^{\infty} x\, p_{\mathsf{x}|\mathsf{s}}(x|s).$$

(iv) The conditional mean of x given s (s viewed as a random variable) is

$$\mu_{\mathsf{x}|\mathsf{s}} = \mathbb{E}[\mathsf{x}|\mathsf{s}] = \sum_{x=-\infty}^{\infty} x\, p_{\mathsf{x}|\mathsf{s}}(x|\mathsf{s}).$$

Since $\mu_{\mathsf{x}|\mathsf{s}}$ is a function of the random variable s, it too is a random variable. Find the probability mass function for $\mu_{\mathsf{x}|\mathsf{s}}$.

(b) Let x and y be independent identically distributed random variables with common density function

$$p(\alpha) = \begin{cases} 1 & 0 \le \alpha \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

Let $\mathsf{s} = \mathsf{x} + \mathsf{y}$.

(i) Find and sketch $p_{\mathsf{s}}(s)$.

(ii) Find and sketch $p_{\mathsf{x}|\mathsf{s}}(x|s)$ vs. x, with s viewed as a known real parameter.

(iii) The conditional mean of x given $\mathsf{s} = s$ is

$$\mathbb{E}\left[\mathsf{x} \mid \mathsf{s} = s\right] = \int_{-\infty}^{+\infty} x\, p_{\mathsf{x}|\mathsf{s}}(x|s)\, dx.$$

Find $\mathbb{E}\left[\mathsf{x} \mid \mathsf{s} = 0.5\right]$.

(iv) The conditional mean of x given s (s viewed as a random variable) is

$$\mu_{\mathsf{x}|\mathsf{s}} = \mathbb{E}\left[\mathsf{x}|\mathsf{s}\right] = \int_{-\infty}^{+\infty} x\, p_{\mathsf{x}|\mathsf{s}}(x|\mathsf{s})\, dx.$$

Since $\mu_{\mathsf{x}|\mathsf{s}}$ is a function of the random variable s, it too is a random variable. Find the probability density function for $\mu_{\mathsf{x}|\mathsf{s}}$.

Problem 1.2

A dart is thrown at random at a wall. Let (x, y) denote the Cartesian coordinates of the point in the wall pierced by the dart. Suppose that x and y are statistically independent Gaussian random variables, each with mean zero and variance σ^2, i.e.,

$$p_{\mathsf{x}}(\alpha) = p_{\mathsf{y}}(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\alpha^2}{2\sigma^2}\right).$$

(a) Find the probability that the dart will fall within the σ-radius circle centered at the point $(0, 0)$.

(b) Find the probability that the dart will hit in the first quadrant ($\mathsf{x} \geq 0$, $\mathsf{y} \geq 0$).

(c) Find the conditional probability that the dart will fall within the σ-radius circle centered at $(0, 0)$ given that the dart hits in the first quadrant.

(d) Let $\mathsf{r} = (\mathsf{x}^2 + \mathsf{y}^2)^{1/2}$, and $\mathsf{\Theta} = \tan^{-1}(\mathsf{y}/\mathsf{x})$ be the polar coordinates associated with (x, y). Find $\mathbb{P}\left(0 \leq \mathsf{r} \leq r,\ 0 \leq \mathsf{\Theta} \leq \Theta\right)$ and obtain $p_{\mathsf{r},\mathsf{\Theta}}(r, \Theta)$.
This observation leads to a widely used algorithm for generating Gaussian random variables.

Problem 1.3

(a) Consider the random variables x, y whose joint density function is given by (see Fig. 1.3-1)

$$p_{\mathsf{x},\mathsf{y}}(x, y) = \begin{cases} 2 & \text{if } x, y \geq 0 \text{ and } x + y \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

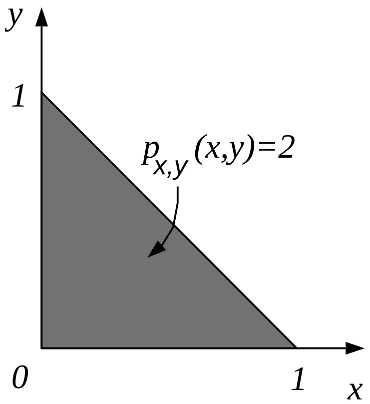

Figure 1.3-1

(i) Compute the covariance matrix

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_x & \lambda_{xy} \\ \lambda_{xy} & \lambda_y \end{bmatrix}.$$

(ii) Knowledge of y generally gives us information about the random variable x (and vice versa). We want to estimate x based on knowledge of y. In particular, we want to estimate x as an affine function of y, i.e.,

$$\hat{x} = \hat{x}(y) = ay + b,$$

where a and b are constants. Select a and b so that the expected mean-square error between x and its estimate $\hat{x}(y)$, i.e.,

$$\mathbb{E}\left[(\hat{x} - x)^2\right],$$

is minimized.

(b) Consider the random variables x, y whose joint density function is given by (see Fig. 1.3-2)

$$p_{x,y}(x,y) = \begin{cases} 1 & 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

Repeat steps (i) and (ii) of part (a), and compare your results to those on part (a).

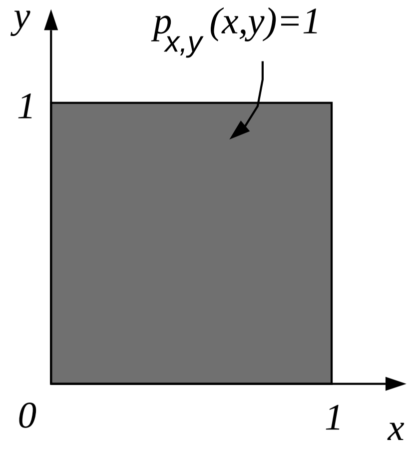

Figure 1.3-2

Problem 1.4 (practice)
Suppose x and y are random variables. Their joint density, depicted below in Figure 1.4-1, is constant in the shaded area and 0 elsewhere.

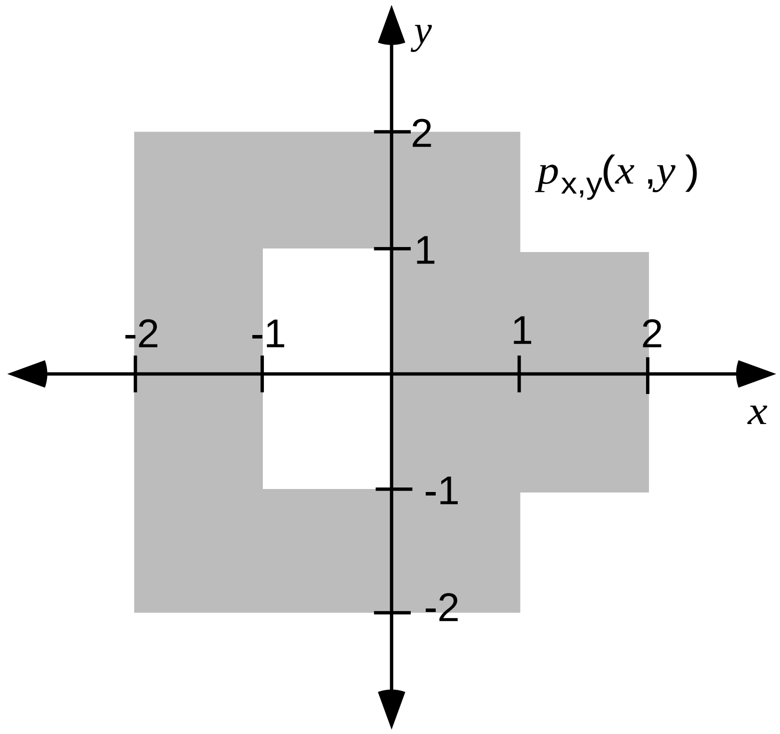

Figure 1.4-1

(a) Let $H = H_0$ when $x \leq 0$, and let $H = H_1$ when $x > 0$. Determine $P_0 = \mathbb{P}(H = H_0)$ and $P_1 = \mathbb{P}(H = H_1)$, and make fully labeled sketches of $p_{y|H}(y|H_0)$ and $p_{y|H}(y|H_1)$.

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

$$\mathcal{Y}_0 = \{y \mid \hat{H}(y) = H_0\}$$
$$\mathcal{Y}_1 = \{y \mid \hat{H}(y) = H_1\}.$$

What is the resulting probability of error?

Problem 1.5

In the binary communication system shown in Figure 1.5-1, messages $m = 0$ and $m = 1$ occur with a priori probabilities 1/4 and 3/4 respectively. Suppose that we observe r,

$$r = n + m,$$

where n is a random variable that is statistically independent of whether message $m = 0$ or $m = 1$ occurs.

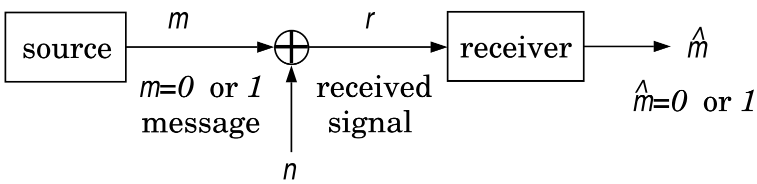

Figure 1.5-1

(a) Say n is a continuous valued random variable, and has the pdf shown in Figure 1.5-2.

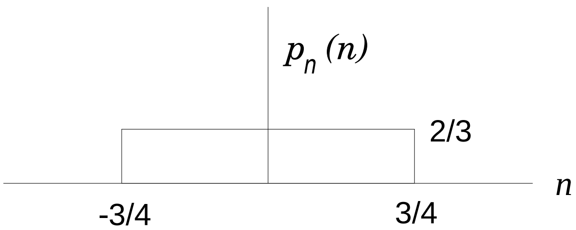

Figure 1.5-2

Find the minimum probability-of-error detector, and compute the associated probability of error, $\mathbb{P}(\hat{m} \neq m)$.

(b) Suppose that the receiver does not know the a priori probabilities, so it decides to use a maximum likelihood (ML) detector. Find the ML detector and the associated probability of error. Is the ML detector unique? Justify your answer. If your answer is no, find a different ML receiver and the associated probability of error.

(c) Now say n is discrete, and takes on the values -1, 0, 1 with probabilities 1/8, 3/4, 1/8 respectively. Find the minimum probability of error detector, and compute the associated probability of error, $\mathbb{P}(\hat{m} \neq m)$.

Problem 1.6

A disease has two varieties: the "0" strain and the "1" strain, with *a priori* probabilities p_0 and p_1 respectively.

(a) Initially, a rather noisy test was developed to determine which strain is present for patients who are known to have one of the two varieties. The output of the test is the sample value y_1 of a random variable y_1. Given that the strain is "0" ($\mathsf{H} = H_0$), $\mathsf{y}_1 = 5 + \mathsf{z}_1$, and given that the strain is "1" ($\mathsf{H} = H_1$), $\mathsf{y}_1 = 1 + \mathsf{z}_1$. The measurement noise z_1 is independent of H and is Gaussian with $\mathsf{z}_1 \sim N(0, \sigma^2)$. Give the MAP decision rule, i.e., determine the set of observations y_1 for which the decision is $\hat{\mathsf{H}} = H_1$. Let

$$\mathcal{E} = (\hat{\mathsf{H}} = H_0 \cap \mathsf{H} = H_1) \cup (\hat{\mathsf{H}} = H_1 \cap \mathsf{H} = H_0)$$

denote the event that there is a decision error. Give $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_0)$ and $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_1)$ in terms of $Q(\cdot)$.

(b) A budding medical researcher determines that the test is making too many errors. A new measurement procedure is devised with two observation random variables y_1 and y_2, y_1 being the same as in part (a). Under hypothesis H_0, $\mathsf{y}_2 = 5 + \mathsf{z}_1 + \mathsf{z}_2$, and under hypothesis H_1, $\mathsf{y}_2 = 1 + \mathsf{z}_1 + \mathsf{z}_2$. Assume that z_2 is independent of both z_1 and H and that $\mathsf{z}_2 \sim N(0, \sigma^2)$. Find the MAP decision rule for $\hat{\mathsf{H}}$ in terms of the joint observation (y_1, y_2), and find $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_0)$ and $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_1)$. (Hint: Find $p_{\mathsf{y}_2|\mathsf{y}_1,\mathsf{H}}(y_2|y_1, H_0)$ and $p_{\mathsf{y}_2|\mathsf{y}_1,\mathsf{H}}(y_2|y_1, H_1)$.)

(c) Explain in layman's terms why the medical researcher should learn more about probability.

(d) Now, suppose that z_2 in part (b) is uniformly distributed between 0 and 1 rather than being Gaussian. We are still given that z_2 is independent of both z_1 and H. Find the MAP decision rule for $\hat{\mathsf{H}}$ in terms of the joint observation (y_1, y_2) and find $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_0)$ and $\mathbb{P}(\mathcal{E}|\mathsf{H} = H_1)$.

(e) Finally, suppose that z_1 is also uniformly distributed between 0 and 1. Again, find the MAP decision rule and error probabilities.

Problem 1.7 (practice)

You are presented with two sealed envelopes, each with money inside. You don't know how much money each envelope holds. You get to keep one of the envelopes, and thus you would like to select the one with the most money in it. To help you make your decision, you choose an envelope by flipping a fair coin, and look inside. You may either keep that envelope, or switch and keep the other one.

(a) Suppose that the amounts of money in each envelope are independent identically distributed random variables according to some distribution $p_\$(\cdot)$ over the positive real line. What strategy should you follow to maximize your probability of selecting the envelope with the most money inside? Characterize your solution as completely as possible.

(b) Suppose that the amounts of money in each envelope are completely unknown, except that the amounts are different from one another. This corresponds to what is referred to as a *composite* hypothesis test. In this scenario, symmetry might suggest that one could do no better than to just pick either of the envelopes to keep entirely at random, thereby achieving a probability of 1/2 of selecting the envelope with the most money.

Show that this intuition is wrong. In particular, consider the following randomized decision strategy: you generate a realization z, independent of the initial coin flip, from any continuous distribution $p_z(\cdot)$ over the positive real line. If the amount of money in the envelope you choose is less than z, you switch to the other envelope. Otherwise, you keep the original envelope. Show that with only very mild conditions on $p_z(\cdot)$ this strategy achieves a probability strictly greater than 1/2 of selecting the envelope with the most money.

Problem 1.8

The traditional shell game is an example of a confidence trick, but in this problem we consider a shell game that is legitimate and for which probabilistic analysis is the important skill. Suppose there are three shells. Your opponent hides a pea under one of the three shells at random and remembers which one; the other two shells have nothing under them. A priori, you do not know which shell has the pea. If you ultimately guess the correct shell, you win \$1. Otherwise you lose \$1.

(a) If the rules of the game are simply that you pick a shell and turn it over to reveal what is under it, what is your strategy and your expected winnings?

(b) Suppose after you pick your shell, you don't turn it over right away. Instead, your opponent turns over one of the remaining shells that has nothing under it. Given what your opponent turned over, what is the probability that you will win, i.e., that your pick has the pea under it?

(c) Suppose further that after following the process of part (b) you are allowed to switch your selection to the remaining shell your opponent chose not to turn over. How does your probability of winning change, if at all, by switching your selection. Should you switch your selection? Explain.