UC Berkeley
Department of Electrical Engineering and Computer Sciences

Electrical Engineering 126: Probability and Random Processes

Problem Set 9
Fall 2017

Self-Graded Scores Due: 5 PM, Monday, November 13, 2017
Submit your self-graded scores via the Google form:
https://goo.gl/forms/JsXC6moMXQhIDGi13.
Make sure you use your **SORTABLE NAME** on CalCentral.

---

1. **Rotationally Invariant Random Variables**

   You have 2 independent and identically distributed continuous random variables, such that the joint density is rotation invariant (isotropic). Show that the random variables have the normal distribution.

   [*Hint*: For simplicity, you can take all random variables to be centered, i.e., with zero mean.]

   Solution:

   Let X and Y be such random variables. Since they are rotation invariant, they should be symmetric, and hence without loss of generality, we can assume that the density looks like:

$$f_X(x) = a\exp\bigl(-h(x^2)\bigr), \qquad f_Y(y) = a\exp\bigl(-h(y^2)\bigr).$$

   This is because $f_X(x)$ must be a function of x^2 only, so if $f_X(x) = k(x^2)$, then we may write $f_X(x) = \exp(-\ln(k(x^2)^{-1}))$ and take $h = \ln \circ k^{-1}$. This requires the density f_X to be strictly positive almost everywhere; this follows from rotational invariance and independence. The joint density is also isotropic, so we have:

$$f_{X,Y}(x,y) = A\exp\bigl(-g(r^2)\bigr)$$

   where, $r^2 = x^2+y^2$. Now, we know that the random variables are independent, and so $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. We get $h(x^2)+h(y^2) = g(x^2+y^2)$. Plugging $y = 0$, and assuming $h(0) = 0$, $h(x^2) = g(x^2)$ which implies $h(x) = g(x)$ for all $x > 0$. Also,

$$\begin{aligned}
h(x^2) + h(y^2) &= g(x^2+y^2) = h(x^2+y^2), \\
h(x) + h(y) &= h(x+y) && \text{for all } x > 0,\ y > 0, \\
h(x) &= bx, && \text{where } b \text{ is constant.}
\end{aligned}$$

   Now, plugging in, $f_X(x) = a\exp(-bx^2)$ which is Gaussian.

   As an alternative proof, for $t \in \mathbb{R}$, $tX + tY$ has the same distribution as $\sqrt{2}tX$, so $\varphi(t)^2 = \varphi(\sqrt{2}t)$, where φ is the characteristic function of X. By iterating,

$\varphi(t)^n = \varphi(\sqrt{n}t)$ for all positive integers n. So, letting $\exp c := \varphi(1)$, for $t^2 = a/b$, where a, b are positive integers, then $\varphi(t) = \varphi(\sqrt{a/b}) = \varphi(1/\sqrt{b})^a$ and $\varphi(1) = \exp c = \varphi(1/\sqrt{b})^b$, so we have $\varphi(t) = \exp(ca/b) = \exp(ct^2)$. We have found that the equation $\varphi(t) = \exp(ct^2)$ holds for all t such that $t^2 \in \mathbb{Q}$, and then by continuity of φ the equation must be true everywhere, and we have the characteristic function of a Gaussian.

2. **BSC: MLE & MAP**

You are testing a digital link that corresponds to a BSC with some error probability $\epsilon \in [0, 0.5]$.

(a) Assume you observe the input and the output of the link. How do you find the MLE of ϵ?

(b) You are told that the inputs are i.i.d. bits that are equal to 1 with probability 0.6 and to 0 with probability 0.4. You observe n outputs (n is a positive integer). How do you calculate the MLE of ϵ?

(c) The situation is as in the previous case, but you are told that ϵ has PDF $4 - 8x$ on $[0, 0.5)$. How do you calculate the MAP of ϵ given n outputs?

Solution:

(a) We observe the input X and the output Y. Thus, if $\mathbb{P}_\epsilon$ denotes the probability distribution when the error probability of the BSC is ϵ, then for $(x, y) \in \{0, 1\}^2$,

$$\epsilon_{\text{MLE}} = \arg\max_{\epsilon\in[0,0.5]} \mathbb{P}_\epsilon(X = x, Y = y) = \arg\max_{\epsilon\in[0,0.5]} \epsilon^{\mathbb{1}\{y\neq x\}}(1-\epsilon)^{\mathbb{1}\{y=x\}}.$$

Now if $x \neq y$, the expression is clearly maximized on the largest possible value of ϵ which is $\epsilon = 0.5$. If $x = y$, the expression is maximized for smallest value of ϵ which is 0.

(b) Suppose that we observe the outputs $y_1, \ldots, y_n$. Thus,

$$\epsilon_{\text{MLE}} = \arg\max_{\epsilon\in[0,0.5]} \mathbb{P}_\epsilon(Y_1 = y_1, \ldots, Y_n = y_n).$$

Since every use of the channel is independent we have,

$$\begin{aligned}
&\mathbb{P}_\epsilon(Y_1 = y_1, \ldots, Y_n = y_n) \\
&\quad= \prod_{i=1}^n \mathbb{P}_\epsilon(Y_i = y_i) \\
&\quad= \prod_{i=1}^n \big[\big(0.6(1-\epsilon) + 0.4\epsilon\big)\mathbb{1}\{y_i = 1\} + \big(0.4(1-\epsilon) + 0.6\epsilon\big)\mathbb{1}\{y_i = 0\}\big] \\
&\quad= \prod_{i=1}^n (0.6 - 0.2\epsilon)^{y_i}(0.4 + 0.2\epsilon)^{1-y_i} \\
&\quad= (0.6 - 0.2\epsilon)^{\sum_{i=1}^n y_i}(0.4 + 0.2\epsilon)^{n-\sum_{i=1}^n y_i}.
\end{aligned}$$

Let $t = \sum_{i=1}^{n} y_i$. As we can see, what matters for estimating ϵ is t. To find the maximizer of the expression, we first take the log and then set the derivative to 0. Thus,

$$\frac{-0.2t}{0.6 - 0.2\epsilon} + \frac{0.2(n-t)}{0.4 + 0.2\epsilon} = 0.$$

Solving the equation, we get

$$\epsilon_{\text{MLE}} = 3 - \frac{5t}{n}.$$

Of course, since we know that $0 \le \epsilon \le 0.5$, if ϵ_{MLE} is not in the interval $[0, 0.5]$ we should pick the closest point to it which will be either 0 or 0.5.

(c) This time we want to maximize $\mathbb{P}(Y_1 = y_1, \ldots, Y_n = y_n \mid \epsilon = \cdot) f_\epsilon(\cdot)$. Similar to the calculations of previous part, we want to maximize,

$$(4 - 8\epsilon)(0.6 - 0.2\epsilon)^t (0.4 + 0.2\epsilon)^{(n-t)}.$$

Taking the log and setting the derivative equal to 0 we have

$$\frac{-8}{4 - 8\epsilon} + \frac{-0.2t}{0.6 - 0.2\epsilon} + \frac{0.2(n-t)}{0.4 + 0.2\epsilon} = 0.$$

Then, we get the following quadratic equation.

$$\begin{aligned} 0 = &-8(0.6 - 0.2\epsilon)(0.4 + 0.2\epsilon) - 0.2t(4 - 8\epsilon)(0.4 + 0.2\epsilon) \\ &+ 0.2(n-t)(4 - 8\epsilon)(0.6 - 0.2\epsilon). \end{aligned}$$

One can solve the long quadratic equation analytically, and find ϵ_{MAP}. We skip the painful algebra here. (You also get full credit, if you find the quadratic equation.)

3. **Poisson Process MAP**

Customers arrive to a store according to a Poisson process of rate 1. The store manager learns of a rumor that one of the employees is sending 1/2 of the customers to the rival store. Refer to hypothesis $X = 1$ as the rumor being true, that one of the employees is sending every other customer arrival to the rival store and hypothesis $X = 0$ as the rumor being false, where each hypothesis is equally likely. Assume that at time 0, there is a successful sale. After that, the manager observes $S_1, S_2, \ldots, S_n$ where n is a positive integer and S_i is the time of the ith subsequent sale for $i = 1, \ldots, n$. Derive the MAP rule to determine whether the rumor was true or not.

Solution:

Note that both hypotheses are a priori equally likely, so the MAP rule is equivalent to the ML rule. The interarrival times are independent conditioned on $X = 1$ and $X = 0$. The density of an interarrival interval given $X = 1$ is Erlang of order 2, so for $0 \le s_1 < \cdots < s_n$:

$$f_{S|X}(s_1, s_2, \ldots, s_n \mid 1) = \prod_{i=1}^{n} (s_i - s_{i-1}) \mathrm{e}^{-(s_i - s_{i-1})} = \mathrm{e}^{-s_n} \prod_{i=1}^{n} (s_i - s_{i-1})$$

The density of an interarrival interval given $X = 0$ is exponential, so:

$$f_{S|X}(s_1, s_2, \ldots, s_n \mid 0) = \mathrm{e}^{-s_n}$$

We can thus see, by taking the log of both expressions, we declare $X = 1$ if $\sum_{i=1}^{n} \log(S_i - S_{i-1}) \geq 0$, otherwise we declare $X = 0$.

4. **Gaussian LLSE**

The random variables X, Y, Z are i.i.d. $\mathcal{N}(0, 1)$.

(a) Find $L[X^2 + Y^2 \mid X + Y]$.

(b) Find $L[X + 2Y \mid X + 3Y + 4Z]$.

(c) Find $L[(X + Y)^2 \mid X - Y]$.

Solution:

(a) We note that

$$\mathbb{E}[(X^2 + Y^2)(X + Y)] = \mathbb{E}[X^3 + X^2Y + XY^2 + Y^3] = 0.$$

Hence,

$$\operatorname{cov}(X^2 + Y^2, X + Y) = 0,$$

so that

$$L[X^2 + Y^2 \mid X + Y] = \mathbb{E}[X^2 + Y^2] = 2.$$

(b) We find

$$\operatorname{cov}(X + 2Y, X + 3Y + 4Z) = \mathbb{E}[(X + 2Y)(X + 3Y + 4Z)] = 1 + 6 = 7$$

and

$$\operatorname{var}(X + 3Y + 4Z) = 1 + 9 + 16 = 26.$$

Hence,

$$L[X + 2Y \mid X + 3Y + 4Z] = \frac{7}{26}(X + 3Y + 4Z).$$

(c) We observe that $\operatorname{cov}(X + Y, X - Y) = 0$, so that $X + Y$ and $X - Y$ are independent. Hence,

$$L[(X + Y)^2 \mid X - Y] = \mathbb{E}[(X + Y)^2] = 2.$$

5. **Photodetector LLSE**

Consider a photodetector in an optical communications system that counts the number of photons arriving during a certain interval. A user conveys information by switching a photon transmitter on or off. Assume that the probability of the transmitter being on is p. If the transmitter is on, the number of photons transmitted over the interval of interest is a Poisson random variable Θ with mean λ, and if it is off, the number of photons transmitted is 0. Unfortunately, regardless of whether or not the transmitter is on or off, photons may be detected due to "shot noise". The number N of detected shot noise photons is a Poisson random variable N with mean μ, independent of the transmitted photons. Given the number of detected photons, find the LLSE of the number of transmitted photons.

Solution:

Let T be the number transmitted photons and D be the number of detected photons. We are looking for:

$$L[T \mid D] = \mathbb{E}[T] + \frac{\operatorname{cov}(T, D)}{\operatorname{var} D}(D - \mathbb{E}[D])$$

We find each of these terms separately. We can see by the tower property that $\mathbb{E}[T] = p\lambda$. Now, we have:

$$\begin{aligned}
\operatorname{cov}(T, D) &= \mathbb{E}[(T - \mathbb{E}[T])(D - \mathbb{E}[D])] \\
&= \mathbb{E}[(T - \mathbb{E}[T])(T - \mathbb{E}[T] + N - \mathbb{E}[N])] \\
&= \mathbb{E}[(T - \mathbb{E}[T])^2] + \mathbb{E}[(T - \mathbb{E}[T])(N - \mathbb{E}[N])] = \operatorname{var} T \\
&= p(\lambda^2 + \lambda) - (p\lambda)^2
\end{aligned}$$

where the second to last equality follows since T and N are independent. We now find:

$$\operatorname{var} D = \operatorname{var}(T + N) = \operatorname{var} T + \operatorname{var} N = p(\lambda^2 + \lambda) - (p\lambda)^2 + \mu$$

Finally, we have $\mathbb{E}[D] = \mathbb{E}[T] + \mathbb{E}[N] = p\lambda + \mu$. Putting these together, we have the LLSE (no need to simplify the equation).

6. **Exam Difficulties**

The difficulty of an EE 126 exam, Θ, is uniformly distributed on $[0, 100]$, and Alice gets a score X that is uniformly distributed on $[0, \Theta]$. Alice gets her score back and wants to estimate the difficulty of the exam.

(a) What is the LLSE for Θ?

(b) What is the MAP of Θ?

Solution:

(a) We need $\mathbb{E}[\Theta]$, σ_X^2, and $\operatorname{cov}(X, \Theta)$. First, $\mathbb{E}[\Theta] = 50$. The variance of X is

$$\operatorname{var} X = \mathbb{E}[\operatorname{var}(X \mid \Theta)] + \operatorname{var} \mathbb{E}[X \mid \Theta].$$

The first term can be found as follows.

$$\operatorname{var}(X \mid \Theta) = \frac{\Theta^2}{12} \implies \mathbb{E}[\operatorname{var}(X \mid \Theta)] = \int_0^{100} \frac{\theta^2}{12} \cdot \frac{1}{100} \, \mathrm{d}\theta = \frac{10000}{36}$$

By noting $\mathbb{E}[X \mid \Theta] = \Theta/2$, the second term is

$$\frac{1}{4}\frac{10000}{12} = \frac{10000}{48}.$$

Thus,

$$\operatorname{var} X = \frac{70000}{144}.$$

Now, the covariance can be found as follows.

$$\operatorname{cov}(X, \Theta) = \mathbb{E}[\Theta X] - \mathbb{E}[\Theta]\, \mathbb{E}[X]$$

We found $\mathbb{E}[\Theta]$ above, and $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \Theta]] = \mathbb{E}[\Theta/2] = 25$. Also,

$$\mathbb{E}[\Theta X] = \mathbb{E}\big[\mathbb{E}[\Theta X \mid \Theta]\big] = \mathbb{E}\Big[\frac{\Theta^2}{2}\Big] = \frac{\operatorname{var}\Theta + \mathbb{E}[\Theta]^2}{2} = \frac{10000/12 + 2500}{2}$$
$$= \frac{5000}{3}.$$

Thus,

$$\operatorname{cov}(X, \Theta) = \frac{1250}{3}.$$

Then, the LLSE of Θ is

$$L[\Theta \mid X] = \mathbb{E}[\Theta] + \frac{\operatorname{cov}(X, \Theta)}{\sigma_X^2}(X - \mathbb{E}[X]) = 50 + \frac{6}{7}(X - 25).$$

(b) Given X, $X \le \Theta \le 100$. In order to maximize

$$f_{X|\Theta}(x \mid \theta) = \frac{1}{\theta},$$

one should choose $\hat{\Theta} = X$.