Information Theory
Practice Set 6

Lakshmi Prasad Natarajan

<u>*Solutions are not to be returned*</u>

Practice Set

See the last page for hints for hints to selected questions.

1. Exercise Problems from Cover and Thomas: 8.1, 8.3(a), 8.4, 8.10
2. Is it possible that a continuous random variable with a well-defined probability density function $f(x)$ has $h(X) = +\infty$?
3. Suppose X is discrete with $P(X = +1) = P(X = -1) = 0.5$, and Z is Gaussian with mean 0 and variance σ^2. Let $Y = X + Z$. Note that Y is continuous and X is discrete. What is $I(X;Y)$ in this case (there is no closed form expression)?
4. Reading exercise from Cover and Thomas:
   (a) Section 8.3
   (b) (Not part of the course syllabus) From the paragraph immediately after equation (8.49) till the paragraph immediately before Example 8.5.1.
5. Let X be a real valued random variable, and $a \in \mathbb{R}$. Explain why the following steps are correct.

$$\begin{aligned}\mathbb{E}((X-a)^2) &= (a - \mathbb{E}X)^2 + \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 \\ &\geq \mathrm{Var}(X), \quad \text{with equality if and only if } a = \mathbb{E}X.\end{aligned}$$

   `Remark`: If we have to "estimate" or "approximate" a random variable X with a constant a, then the estimate $\hat{X} = a$ that minimizes the *mean square error* $\mathbb{E}(X - \hat{X})^2$ is the expected value of X. This estimate $\hat{X} = \mathbb{E}X$ is called the minimum mean square error (MMSE) estimate of X. We will discuss more about this in future lectures.

6. Suppose $X = (X_1, X_2)$ is Gaussian with zero mean and covariance

$$\boldsymbol{K} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

   Assume $-1 < \rho < +1$.

   (a) What is $\det \boldsymbol{K}$? Is $\boldsymbol{K}$ invertible?
   (b) Find $\boldsymbol{K}^{-1}$.
   (c) From the last step in Lecture 17 we know that

$$h(X) = \frac{1}{2}\log\left((2\pi)^2 \det \boldsymbol{K}\right) + \frac{1}{2}\log e \sum_{i=1}^{2}\sum_{j=1}^{2}\left(\boldsymbol{K}^{-1}\right)_{i,j}(\boldsymbol{K})_{i,j}.$$

   Find $h(X)$ as a function of ρ.
   (d) What are the values of $h(X_1)$ and $h(X_2)$?
   (e) Find $I(X_1;X_2)$ as a function of ρ. Plot this as a function of ρ, and try to interpret it.

7. *Definition of Gaussian random vector.* We say that a collection of random variables $X_1, \dots, X_n$ forms a Gaussian random vector if any linear combination of $X_1, \dots, X_n$ is Gaussian distributed, i.e., for any choice of constants $a_1, \dots, a_n$, the sum $\sum_i a_i X_i$ has a Gaussian distribution.

   Fact. If $X_1, \dots, X_n$ are independent Gaussian random variables, then their sum is Gaussian distributed.

   Now answer the following questions.

   (a) Suppose $X_1, \dots, X_n$ are independent with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Argue that $(X_1, \dots, X_n)$ is a Gaussian random vector, and find its covariance matrix and density function.

   (b) With X as in part (a), and assuming $\boldsymbol{A}$ is any $m \times n$ matrix (this is not necessarily square or invertible), argue that $Y = \boldsymbol{A}X$ is a Gaussian random vector. What is the mean and covariance matrix of Y ?

8. Suppose $X = (X_1, \dots, X_n)$ contains iid components $X_i \sim \mathcal{N}(0, \sigma^2)$.

   (a) Find the covariance matrix $\boldsymbol{K}$ for X.

   (b) What is the probability density function of X ?

   (c) Suppose $\boldsymbol{A}$ is an invertible $n \times n$ matrix, and $Y = \boldsymbol{A}X$. Then what is the mean vector and the covariance matrix of Y?

   (d) Y is also a Gaussian random vector. What is its density function ?

Hints for selected questions
in the next page

Hints

- Q.1

  8.4: The median of the exponential density $\lambda e^{-\lambda x}$ is $\frac{1}{\lambda \log e}$. For three digit accuracy we need $\Delta = 10^{-3}$. Find $h(X)$ and the approximate value of $H(X_\Delta)$.

- Q.2

  Yes, it is possible. We have seen a discrete random variable, call it Y, with $\mathcal{Y} = \{2, 3, 4, \dots\}$ and $H(X) = +\infty$ (see previous Practice Sets). Now define a continuous random variable X as follows:

  $$f_X(x) = P[Y = k] \text{ if } k \leq x < k + 1$$

  where $x \geq 2$, and k is an integer. Define $f_X(x) = 0$ for $x < 2$. Find $h(X)$ and relate this to $H(Y)$.

- Q.3

  Here, use the fact $I(X;Y) = P[X = -1]h(Y|X = -1) + P[X = +1]h(Y|X = +1)$. You will not be able to get a closed form solution for $h(Y)$. But you can plot $h(Y) - h(Y|X)$ in a computer algebra software as a function of σ^2.

- Q.7

  (b): mean is $\boldsymbol{A}\mu$ where $\mu = (\mu_1, \dots, \mu_n)$ is a column vector. Covariance is $\boldsymbol{A}\boldsymbol{A}^t$.
