UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Problem Set 6
Spring 2018

Issued: February 21, 2018 **Due:** 11:59 PM, Wednesday, February 28, 2018

---

1. **The Weak Law of Large Numbers**

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with common mean μ and associated transform M_X. We assume that $M_X(s)$ is finite when $s \in (-d, d)$, where d is some positive number. Let

$$\bar{X}_n = \frac{X_1 + \ldots + X_n}{n}.$$

(a) Show that the transform associated with $\bar{X}_n$ satisfies

$$M_{\bar{X}_n}(s) = M_X(s/n)^n.$$

(b) Suppose that the transform $M_X(s)$ has a first order Taylor series expansion around $s = 0$, of the form

$$M_X(s) = a + bs + o(s),$$

where $o(s)$ is a function that satisfies $\lim_{s\to 0} o(s)/s = 0$. Find a and b in terms of μ.

(c) Show that

$$\lim_{n\to\infty} M_{\bar{X}_n}(s) = \mathrm{e}^{\mu s}, \quad \text{for all } s \in (-d, d).$$

Hint: If $\{a_n\}_{n\in\mathbb{N}}$ is a real sequence with $\lim_{n\to\infty} a_n = a \in \mathbb{R}$, then $\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = \mathrm{e}^a$.

(d) Deduce that $\bar{X}_n \xrightarrow{\mathsf{d}} \mu$.

Note: From this together with Problem 6 we can conclude that $\bar{X}_n \xrightarrow{\mathbb{P}} \mu$.

2. **Huffman Questions**

Consider a set of n objects. Let $X_i = 1$ or 0 accordingly as the i-th object is good or defective. Let $X_1, X_2, \ldots, X_n$ be independent with $\mathbb{P}(X_i = 1) = p_i$; and $p_1 > p_2 > \cdots > p_n > 1/2$. We are asked to determine the set of all defective objects. Any yes-no question you can think of is admissible.

(a) Propose an algorithm based on Huffman coding in order to identify all defective objects.

(b) If the longest sequence of questions is required by nature's answers to our questions which are based on Huffman coding, then what (in words) is the last question we should ask? And what two sets are we distinguishing with this question?

Note: This problem is related to the 'Entropy and Information Content' section of Lab 4.

3. Channel Capacity of the Binary Symmetric Channel

Random Code

Each word $w \in \{0,1\}^k$, independently is encoded by a codeword $X_w = (X_w(1), \ldots, X_w(n))$ which consists of n i.i.d. Bernoulli(1/2) random variables. Define the encoding function $f_n : \{0,1\}^k \to \{0,1\}^n$ as

$$f_n(w) = \big(X_w(1), \ldots, X_w(n)\big).$$

Noise

During transmission, a codeword $X_w = (X_w(1), \ldots, X_w(n))$ is corrupted by noise $N = (N(1), \ldots, N(n))$ which is assumed to consist of n i.i.d. Bernoulli(p) random variables. Then the received message $Y_w = (Y_w(1), \ldots, Y_w(n))$ can be represented as $Y_w(i) = X_w(i) \oplus N(i)$, for $i = 1, \ldots, n$, where $\oplus$ denotes the XOR operation.

Decoding

We decode the received message Y_w, using the following decoding function $g_n : \{0,1\}^n \to \{0,1\}^k$

$$g_n(Y_w) = \begin{cases} u, & \text{if there is a } \textbf{unique}\ u \in \{0,1\}^k \text{ s.t. } Y_w \in \text{DecodeBox}(X_u) \\ \text{fail}, & \text{otherwise,} \end{cases}$$

where for any $\epsilon > 0$ we define

$$\text{DecodeBox}(x) = \left\{ y \in \{0,1\}^n : \left| \sum_{i=1}^{n} y(i) \oplus x(i) - pn \right| \le n\epsilon \right\}.$$

(a) Using the Chernoff bound argue that

$$\mathbb{P}\big(Y_w \notin \text{DecodeBox}(X_w)\big) \le 2\mathrm{e}^{-2n\epsilon^2}.$$

(b) Argue that $|\text{DecodeBox}(0)| = |\text{DecodeBox}(x)|$, for all $x \in \{0,1\}^n$.

(c) Argue that each component $Y_w(i)$ is marginally a Bernoulli(1/2) random variable.

(d) For $w \neq u$, show that

$$\mathbb{P}\big(Y_w \in \text{DecodeBox}(X_u)\big) = \frac{|\text{DecodeBox}(0)|}{2^n}.$$

(e) Fix a word $w \in \{0,1\}^k$. Argue that

$$\mathbb{P}\big(\exists u \neq w : Y_w \in \text{DecodeBox}(X_u)\big) \le 2^k \frac{|\text{DecodeBox}(0)|}{2^n}.$$

(f) Use the A.E.P. from Homework 5 Problem 5 (c) to deduce that

$$|\text{DecodeBox}(0)| \le 2^{n(H(p)+\epsilon')},$$

where ϵ' is some linear function of ϵ.

(g) How large can the channel capacity $C = k/n$ be in order to ensure that asymptotically both error probabilities go to zero?

4. Number of Parameters

(a) Let $Y_0, Y_1, \ldots, Y_n$ be binary random variables. How many parameters are reuqired to parametrize the joint distribution $\mathbb{P}(Y_0 = y_0, Y_1 = y_1, \ldots, Y_n = y_n)$?

(b) Let $Z_0, Z_1, \ldots, Z_n$ be binary, independent random variables. How many parameters are required to parametrize the joint distribution $\mathbb{P}(Z_0 = z_0, Z_1 = z_1, \ldots, Z_n = z_n)$?

(c) Let $X_0, X_1, \ldots, X_n, \ldots$ be a Markov chain with state space $S = \{0, 1\}$, initial distribution π_0 and transition probability matrix P. How many parameters are required to parametrize the joint distribution $\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n)$, where $X_0, X_1, \ldots, X_n$ are the first $n + 1$ random variables of the Markov chain given above?

(d) Say we want to construct a countably infinite sequence $Z_0, Z_1, \ldots, Z_n, \ldots$ of independent random variables, which is not a Markov chain. Which defining property of Markov chains must this sequence violate? Give a concrete example of such sequence of independent random variables.

5. Backwards Markov Property

Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain with state space $\mathcal{S}$. Show that for every $m, k \in \mathbb{N}$, with $m \geq 1$, we have

$$\mathbb{P}(X_k = i_0 \mid X_{k+1} = i_1, \ldots, X_{k+m} = i_m) = \mathbb{P}(X_k = i_0 \mid X_{k+1} = i_1),$$

for all states $i_0, i_1, \ldots, i_m \in \mathcal{S}$.

6. [Bonus] The CLT Implies the WLLN

The bonus question is just for fun. You are not required to submit the bonus question, but do give it a try and write down your progress.

(a) Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables. Show that if $X_n \xrightarrow{\mathsf{d}} c$, where c is a constant, then $X_n \xrightarrow{\mathbb{P}} c$.

(b) Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables, with mean μ and finite variance σ^2. Show that the CLT implies the WLLN.