Topics in Random Processes and Concentrations: EE 6112
Indian Institute of Technology Madras — March 25, 2020
Instructor: Abhishek Sinha

Problem Set 5

- This problem set is due on **April 10, 2020** in the class.
- Each problem carries 10 points.
- You may work on the problems in groups of size at most **two**. However, **each student must write their own solution**. If you collaborate on the problems, clearly mention the name of your collaborator.

---

1. **(Large deviation for sum of products)** Suppose X_i is an i.i.d. zero mean sequence of random variables with a finite everywhere moment generating function $M(\theta) = \mathbb{E}(\exp(\theta X_1))$. Argue the existence and express the following large deviations limit

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\bigg(\sum_{1\leq i\neq j\leq n)} X_i X_j > n^2 z) \bigg)$$

in terms of $M(\theta)$ and z.
HINT: Express the event in a manageable form.

2. **(Number of Power-Constrained Ternary Codewords)** Consider a ternary codeword of length n, consisting of n symbols, each taken from the set $\{0, 1, 2\}$. This codeword is to be transmitted over a power-constrained wireless link. Assume that the symbols $0, 1, 2$ takes $0, 1, 2$ unit of energy respectively, for transmission. Due to the power-constraint, only those sequences are eligible for transmission whose required average transmission energy is at most $\frac{1}{2}$.

In other words, for a sequence $(s_1, s_2, \ldots, s_n)$, if the i^{th} transmitted symbol s_i consumes w_i amount of energy, then the codeword is eligible for transmission if and only if

$$\frac{1}{n}\sum_{i=1}^{n} w_i \leq \frac{1}{2}.$$

Let $N(n)$ denote the number of eligible codewords of length n. Define

$$l = \lim_{n\to\infty} \frac{1}{n} \ln(N(n)).$$

Show that the above limit exists and compute l.

HINT: Which natural probability distribution on the codewords will turn this apparently combinatorial counting problem into a familiar problem on large deviation?

3. **(Link with time-varying capacity)** Consider calls accessing a link with time-varying capacity. The call holding times are independent and exponentially distributed with mean $\frac{1}{\mu}$, and each call requires one unit of bandwidth from the link to maintain QoS. The link has an available capacity of nc_0 in the time interval $[0, \tau_1)$, an available capacity of nc_1 in $[\tau_1, \tau_2)$, and an available capacity of nc_2 in $[\tau_2, \infty]$, where $\tau_2 > \tau_1 > 0$. We assume that $nc_0 > nc_1 > nc_2$ and $n > 0$. Suppose that there are $\lfloor n\alpha \rfloor$ calls in progress at time $t = 0$ (where $\alpha \in (0, c_0)$), and that no further calls ever arrive at the system. Let $N(t)$ denote the number of calls in progress at time t. Thus, the bandwidth requested by the calls at time t is $N(t)$. We are interested in estimating the probability that the required bandwidth ever exceeds the available capacity, up to a logarithmic equivalence. To this end, prove the following fact for appropriate functions $I_i(x)$:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\exists t \in [0, \infty) \text{ such that } N(t) \geq c(t)) = -\min_{1,2} I_i(c_i/\alpha),$$

where $c(t)$ is the available capacity at time t. Explicitly compute the functions $I_i(x)$.
Hint: Express the number of calls in progress at a time t as sum of appropriate indicator random variables.

4. **(Doubly-Constrained Sequences)** Let $\mathcal{Z}_n$ be the set of all length-n sequences of $0, 1$ such that:

   (a) 1 is always followed by 0 (so, for example, 001101 is not a legitimate sequence, but 001010 is)

   (b) the fraction of 0's in the sequence is at least α. (Namely, if $X_n(z)$ is the number of zeros in a sequence $z \in \mathcal{Z}_n$, then $X_n(z)/n \geq \alpha$ for every $z \in Z_n$).

   Let $Z_n = |\mathcal{Z}_n|$. Assuming that the limit $\lim_{n\to\infty} \frac{1}{n} \log Z_n$ exists, compute its limit for $\alpha = 0.7$. You may also assume that the limit is strictly monotonic in α in the neighbourhood of $\alpha = 0.7$.

   Hint: Use **Gartner-Ellis** Theorem with an *appropriately defined* Markov Chain on the sequences. Note that your Markov Chain needs to incorporate the constraint (a). The constraint (b) may be incorporated by the Large deviation theorem (similar to Problem 2).

Optional: Write an efficient computer program to compute Z_n for $n = 10^5$ and compute $\frac{\log Z_n}{n}$. Does this value corroborate your theoretical result?
Hint: Use Dynamic Programming with suitably chosen states. You may want to use IITM's excellent VIRGO Super Cluster facility for your computations (`https://cc.iitm.ac.in/node/189`).
