Indian Institute of Technology, Madras
Abhishek Sinha November 26, 2019

Warm up Problems

1. **(PhD Candidate Selection Cut-Off)** The EE Department of IIT Madras is trying to fill one vacant PhD scholar position. There are n candidates whom the department interviews *sequentially*, and assigns a score in the range $[0,1]$ upon interviewing. Before the interviews, it is known that the scores of the candidates are independently and uniformly distributed in $[0,1]$. For $i \geq 1$, department interviews the $i+1^{\text{th}}$ candidate only if it rejects all previous candidates numbered 1 to i, based on *optimally pre-set threshold scores*. If the department selects the i^{th} candidate, it cannot interview candidates numbered $i+1$ to n and the selection process stops immediately. The goal of the department is to maximize the expected score of the selected candidate. Let M_n^* be the minimum score of the *first* candidate for which the department selects him under the optimal policy. Show that

$$\lim_{n\to\infty} n(1-M_n^*) = 2,$$

by the following series of steps or otherwise.

   (a) Let the expected score of the selected candidate under the optimal policy be M_k when there are k candidates yet to interview. By conditioning on the score of the next candidate and optimizing over the threshold score for him, show that

$$M_k = \frac{1}{2}(1 + M_{k-1}^2).$$

   (b) Define $\epsilon_n \equiv 1 - M_n$ and, from the above, derive a recursion in terms of ϵ_n. Show that

$$\lim_{n\to\infty} \frac{\epsilon_n}{\epsilon_{n-1}} = 1.$$

   (c) Show that

$$\frac{1}{n} - \frac{1}{n\epsilon_n} = -\frac{1}{2}\left(\frac{1}{n}\sum_{k=2}^{n} \frac{\epsilon_{k-1}}{\epsilon_k}\right).$$

   (d) Conclude the result by using (b), (c) in conjunction with the Cesaro mean theorem.

2. **(WLLN for Locally Correlated Random Variables)** Let $X_1, X_2, \ldots$ be i.i.d. sequence of random variables with

$$\mathbb{P}(X_1 = 1) = p, \;\; \mathbb{P}(X_1 = -1) = 1 - p, \;\; 0 < p < 1.$$

Define another (correlated) sequence of random variables $\{Y_i, \; i \geq 1\}$ as follows:

$$\begin{aligned} Y_n &= 1 \quad \text{if } X_n = 1 \text{ and } X_{n+1} = 1, \\ &= 0 \quad \text{o.w.} \end{aligned}$$

Show that

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \to p^2, \qquad \text{in probability.}$$

3. **(Expected Time to Cross a Threshold)** You are given a random number (uniformly distributed) between 0 and 1. To this, you add a second such random number. Keep adding numbers until the sum exceeds 1, and then stop. How many numbers N, in expectation, will you need? Can you find the variance of N?

Optional: Write a computer script and attach relevant plots to numerically verify your answers.

4. **(Lower Bound on Tail)** Assume that $x_1, x_2, \ldots, x_n$ are non-negative real numbers such that $\sum_{i=1}^{n} x_i^2 = n$ and $\sum_{i=1}^{n} x_i \geq s$. Prove that for any $0 \leq \lambda \leq 1$, at least $\lceil \frac{s^2(1-\lambda)^2}{n} \rceil$ of these numbers are larger than $\frac{\lambda s}{n}$. Can you turn this (deterministic) inequality to a probabilistic statement?

Hint: Cauchy-Schwartz inequality is your friend!

5. **(Probability that a random message passes two-dimensional parity check)** Consider a binary string of length mn arranged as a rectangular array of m rows and n columns. The set of all binary $m \times n$ array satisfying the condition

$$\sum_i X_{ij} = 0 \mod (2), \forall j$$
$$\sum_j X_{ij} = 0 \mod (2), \forall i$$

is called two-dimensional parity check code $\mathcal{C}_{m,n}$. Find the probability that a random binary string of length mn, each of whose bits are Bernoulli (1/2) is a codeword in $\mathcal{C}_{m,n}$.

For a string with 10000 bits, find the optimal m and n to minimize this probability.

6. **(Local Maximas in Permutations)** A permutation π of $\{1, 2, \dots, n\}$ (with $n \geq 3$) has a local maximum at a position k if the two neighbouring numbers (or, in case $k = 1$ or $k = n$, the one neighbouring number) are both smaller than the number in position k.

   Example: If $n = 5$, then the permutation $\{2, 1, 4, 5, 3\}$ has local maxima(s) in position(s) 1 and 4 (the numbers 2 and 5 respectively).

   What is the average number of local maxima of a permutation of $\{1, 2, \dots, n\}$, averaging over all such permutations?
   HINT: Use linearity of expectation.

7. **(Application of the CLT)** Evaluate the following limit

$$\lim_{x \to \infty} e^{-x} \sum_{k=0}^{x/2} \frac{x^k}{k!}.$$

   HINT: This problem, although looks quite formidable, becomes elementary if done in the right way. Think of how you can apply the celebrated Central Limit Theorem to evaluate this limit.