UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Discussion 5
Fall 2021

1. **Curse of Dimensionality**

In this problem, we will use the law of large numbers to illustrate a statistical phenomenon. In particular, consider the hypercube $[-1,1]^n$ in $\mathbb{R}^n$, and let $X_1, \dots, X_n$ be iid Uniform($[-1,1]$).

(a) For $\epsilon > 0$ consider the set

$$A_{n,\epsilon} := \{x \in \mathbb{R}^n : (1-\epsilon)\sqrt{n/3} < \|x\|_2 < (1+\epsilon)\sqrt{n/3}\},$$

which is the ϵ-boundary of a ball with radius $\sqrt{n/3}$ centered at the origin. For low dimensions $n = 1, 2$ and $\epsilon = 1/10$, compute the fraction of volume of $[-1,1]^n$ which comes from $A_{n,\epsilon}$.

(b) Show that as n gets large, most of the volume of the hypercube comes from $A_{n,\epsilon}$. Comment on why this contradicts the intuition developed in part (a).

Solution:

(a) For $n = 1$: we can compute the fraction to be $2(2/10)\sqrt{1/3}/2 = \frac{1}{5\sqrt{3}}$.
For $n = 2$: the area of the ring is given by

$$\pi(1 + 1/10)^2(2/3) - \pi(1 - 1/10)^2(2/3) = \frac{4}{15}\pi.$$

Thus the fraction is $\pi/15$.

(b) Note that X_i^2 are iid since they are functions of iid variables. Since $\mathbb{E}[X_i^2] = 1/3$, we have by the weak law of large numbers that

$$\frac{X_1^2 + \cdots + X_n^2}{n} \xrightarrow{p} 1/3.$$

In particular this implies that

$$P\left(|(X_1^2 + \cdots + X_n^2) - n/3| < n\delta\right) \to 1.$$

But note that $R^2 = (X_1^2 + \cdots + X_n^2)$ is just the squared radius, so put another way this says

$$P((1-3\delta)n/3 < R^2 < (1+3\delta)n/3) \to 1.$$

Now picking δ sufficiently small (i.e. so that $(1-3\delta, 1+3\delta) \subset ((1-\epsilon)^2, (1+\epsilon)^2)$), this implies

$$P((1-\epsilon)\sqrt{n/3} < R < (1+\epsilon)\sqrt{n/3}) \to 1.$$

But as our $X_1, \dots, X_n$ are each uniform on $[-1,1]$, the probability of this region is precisely the ratio of its volume to the volume of the hypercube. Thus

$$\frac{\text{Vol}(A_{n,\epsilon} \cap [-1,1]^n)}{2^n} \to 1,$$

as desired. This result should be pretty bizarre! Note that it is true for every ϵ, no matter how small. So even if $\epsilon = 0.000001$, eventually for a high enough dimension n, most of the volume of the hypercube comes from the thin ϵ-boundary of a sphere. This phenomena arises frequently when dealing with large data sets, in which the dimension could indeed be quite large.

2. **Product of Rolls of a Die**

A fair die with labels (1 to 6) is rolled until the product of the last two rolls is 12. What is the expected number of rolls?

[Hint: You can model this process as a Markov chain with 3 states. Choose your states according to the outcome of last roll. For example, assign one state if it is outcome was 1 or 5 (which is useless if you want the product to be 12). If the outcome was 2,3,4 or 6, it's useful and can be assigned another state. Assign third state to the case when the product last two outcomes was 12.]

Solution:

According to the hint, we model this process as a Markov chain with 3 states. The states correspond to the outcome of the last roll. If the last outcome is 1 or 5, it is useless for getting a product of 12, and we say that the Markov chain is in state s_1. If the last outcome is one of 2, 3, 4, or 6, the outcome is useful, and we say that the Markov chain is in state s_2. If the product of the last two rolls is 12, we say that the Markov chain is in state s_3. Then the probability transition matrix is

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 \\ 1/3 & 1/2 & 1/6 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let T_i be the expected number of rolls that is needed to get to state s_3, starting from state s_i, $i = 1, 2$. Then we have

$$\begin{aligned} T_1 &= 1 + \frac{1}{3}T_1 + \frac{2}{3}T_2, \\ T_2 &= 1 + \frac{1}{3}T_1 + \frac{1}{2}T_2. \end{aligned}$$

Solving the equations, we get $T_1 = 10.5$ and $T_2 = 9$. Then the expected number of rolls is

$$T = 1 + \frac{1}{3}T_1 + \frac{2}{3}T_2 = 10.5.$$

3. **Concentration for Binomials & Gaussians**

For sums of bounded zero-mean i.i.d. random variables $S_n = X_1 + \ldots + X_n$, a Chernoff-type inequality tells us that

$$P(|S_n| \geq t\sqrt{n}) \leq C\exp(-ct^2),$$

for some constants $C, c > 0$.

(a) (Optional) Prove the inequality above.

(b) Let $Z \sim \mathcal{N}(0, 1)$. Show that

$$P(|Z| \geq t) \leq C \exp(-ct^2).$$

Solution:

(a) See Azuma's inequality.

(b) Let $Y_i \sim \text{Bernoulli}(1/2)$. Rescale these, so that $X_i = 2Y_i - 1$, where X_i is zero mean and unit variance. Let $S_n = X_1 + \cdots + X_n$. By CLT, we know that $S_n/\sqrt{n}$ converges in distribution to Z. That is, we have

$$P(|Z| \geq t) \leftarrow P(|S_n/\sqrt{n}| \geq t) \leq C \exp(-ct^2)$$

as $n \to \infty$. Since the LHS is fixed for all n, this implies that

$$P(|Z| \geq t) \leq C \exp(-ct^2).$$

Bounds of this form are known as sub-gaussian concentration.