Tutorial Problems: Concentration Inequalities

Sumohana Chennappayya and G V V Sharma*

1) Convergence of random variables: Define the following types of convergence of a sequence of random variables $(X_n : n \geq 1)$ to a random variable X: (5)
   a) Almost sure: $X_n \xrightarrow{a.s.} X$.
   b) Probability: $X_n \xrightarrow{p.} X$.
   c) Mean squared: $X_n \xrightarrow{m.s.} X$.
   d) Distribution: $X_n \xrightarrow{d.} X$.
2) Let Θ be uniformly distributed on the interval $[0, 2\pi]$. In which of the four senses (defined in the previous question) do each of the following two sequences converge. Identify the limits, if they exist, and justify your answers. (10)
   a) $(X_n : n \geq 1)$ defined by $X_n = \cos(n\Theta)$.
   b) $(Y_n : n \geq 1)$ defined by $Y_n = |1 - \frac{\Theta}{\pi}|^n$.
3) Find the moment generation function of $X \sim \mathcal{N}(0, \sigma^2)$. (5)
4) Use the Markov inequality to prove the following (with appropriate assumptions):
   a) The Chebyshev inequality. (5)
   b) The Chernoff bound. (5)
5) a) Give an example where the Markov inequality is tight. (5)
   b) Give an example where the Chernoff bound is tighter than the Chebyshev inequality. (5)
   c) Suppose we have 100 books with the number of pages uniformly distributed in [5, 50] so that the average number of pages is 27.50 and the variance is 168.75 pages. Upper bound the probability that the number of pages exceeds 3000 pages using both Chebyshev's inequality and CLT. What can you say about the two estimates? (5)
   d) Illustrate with a Python program the CLT for at least four different types of distributions. (5)
6) a) Use the basic inequalities to prove the weak law of large numbers (LLN). (5)
   b) Illustrate the efficacy of the LLN with a Python program using at least four different types of distributions. (5)
7) Definition: A real valued random variable X is said to be σ^2-sub-Gaussian if there exists a positive number σ such that $\mathbb{E}[e^{\lambda X)}] \leq e^{\frac{\sigma^2\lambda^2}{2}}$ for every $\lambda \in \mathbb{R}$. Show that the following random variables are σ^2-sub-Gaussian and find σ for each of them.
   a) A Rademacher random variable. (5)
   b) A Gaussian random variable with zero mean and variance σ^2. (5)
   c) A random variable that is zero mean and bounded in the interval $[a, b]$. (5)
8) Prove that for a σ^2-sub-Gaussian random variable with mean μ, $P[|X - \mu| \geq t] \leq \exp(\frac{-t^2}{2\sigma^2})$ for all $t > 0$. (5)
9) If X_i are independent, mean-zero, σ_i^2-sub-Gaussian random variables, show that $\sum_{i=1}^{n} X_i$ is $\sum_{i=1}^{n} \sigma_i^2$-sub-Gaussian. (5)
10) If X_i are bounded random variable (in $[a_i, b_i]$), show that $P(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t) \leq \exp\Big(-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\Big)$ and $P(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \leq -t) \leq \exp\Big(-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\Big)$ for all $t > 0$. (10)
11) Definition: A random variable X with mean μ is sub-exponential if there are non-negative parameters (ν, b) such that $\mathbb{E}[e^{\lambda(X-\mu)}] \leq e^{\frac{\nu^2\lambda^2}{2}}$ for all $|\lambda| < \frac{1}{b}$. Show that:
    a) The exponential random variable with parameter λ is sub-exponential. (5)
    b) The χ^2-random variable is sub-exponential. (5)
12) Suppose that X is a (ν, b)-sub-exponential random variable with mean μ, derive the tail bound $P[X \geq \mu + t]$ for all $t > 0$. (5)
13) If X_i are independent, (ν, b)-sub-exponential random variables, then $\sum_{i=1}^{n} X_i$ is $(\sum_{i=1}^{n} \nu_i, b_*)$-sub-

*The author is with the Department of Electrical Engineering, IIT, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

exponential where $b_* = \max_i b_i$. (5)

14) Bernstein bound: If X is a random variable with mean μ and variance σ^2 and satisfies the condition $|\mathbb{E}[(X-\mu)^k]| \leq \frac{1}{2}k!\sigma^2 b^{k-2}$, it is said to satisfy the Bernstein condition with parameter b. For such a random variable show that $\mathbb{E}[e^{\lambda(X-\mu)}] \leq e^{\frac{\lambda^2\sigma^2/2}{1-b|\lambda|}}]$ for all $|\lambda| < \frac{1}{b}$. (5)
15) The $\chi^2(n)$ random variable is defined as $\chi^2(n) = \sum_{i=1}^{n} X_i^2$ where X_i are independent standard normal random variables. Empirically check for sub-Gaussianity of $\chi^2(n)$ as a function of n. (5)
16) We saw a few examples of sub-Gaussian random variables in the class. Demonstrate the following with a Python script:
    a) Sub-Gaussianity of a Gaussian random variable with zero mean and variance σ^2. (5)
    b) Sub-Gaussianity of a Uniform random variable with range $[-a, a]$. (5)
    c) Non sub-Gaussianity of a Laplacian random variable with zero mean and variance $2b^2$. (5)
    d) Non sub-Gaussianity of a centered heavy tailed random variable of your choice. (5)
    e) Sub-Gaussianity of a sum of bounded random variables with zero mean. (5)

    As discussed in class, choose the variance of the "reference" Gaussian appropriately.
17) Recall the definition of the Cramer's transform from the quiz. Find the Cramer's transform of a centered Bernoulli random variable with parameter p. (5)
18) Bennett's inequality: we proved in class that for centered random variables $X_1, X_2, \ldots, X_n$ that satisfy $X_i \leq b$ for some $b > 0$ almost surely for all $i \leq n$, and have finite variance with $v = \sum_{i=1}^{n} X_i^2$ and $S = \sum_{i=1}^{n}(X_i - E[X_i])$, for all $\lambda > 0$, $\log\mathbb{E}e^{\lambda S} \leq n\log(1 + \frac{v}{nb^2}\phi(\lambda b)) \leq \frac{v}{b^2}\phi(\lambda b)$. Here $\phi(u) = e^u - u - 1$. Now show the following tail bound for any $t > 0$: $P(S \geq t) \leq \exp(-\frac{v}{b^2}h(\frac{bt}{v}))$ where $h(u) = (1+u)\log(1+u) - u$ for $u > 0$. (5)
19) Empirically compare (using a Python script) the sharpness/tightness of the tail bound due to Bennett's inequality with the Hoeffding's inequality and the Chernoff's inequality. Show appropriate plots to demonstrate your comparisons. (10)
20) Assume that the random variables $X_1, \ldots, X_n$ are independent and binary {-1, 1}-valued with $P\{X_i = 1\} = p_i$ and that $f : \{-1,1\}^n \longrightarrow \mathbb{R}$ has the bounded differences property with constants $c_1, \ldots, c_n$. Show that if $Z = f(X_1, \ldots, X_n)$, $Var(Z) \leq \sum_{i=1}^{n} c_i^2 p_i(1-p_i)$. (10)
21) Efron-Stein inequality: Recall the notation and formulation from class. $X_1, \ldots, X_n$ are independent random variables that take values from the set $\mathcal{X}$, $f : \mathcal{X}^n \longrightarrow \mathbb{R}$ is a square integrable function, $Z = f(X_1, \ldots, X_n), E_i(Z) = E[Z|X_1, \ldots, X_i], E^i(Z) = \int_{x_i\in\mathcal{X}} f(X_1, \ldots, x_i, \ldots X_n)dP(x_i), \Delta_i = E_i(Z) - E_{i-1}(Z)$. Prove the following equalities and inequalities that were claimed to be true in class:
    a) $Var(Z) = \sum_{i=1}^{n} \Delta_i^2$. (5)
    b) $E_i E^i(Z) = E_{i-1}(Z)$. (5)
    c) $\Delta_i^2 \leq E_i((Z - E^i(Z))^2)$. (5)
    d) $Var(Z) \leq \sum_{i=1}^{n} E(Z - E^i(Z))^2$. (5)