TOPICS IN PROBABILITY. PART I: CONCENTRATION

Exercise sheet 1: Basic concentration inequalities

Exercise 1 (Recap: Markov's inequality and direct corollaries)**.**

- *Suppose that X is a random variable, $a > 0$. Show that $\mathbb{P}[|X| \geq a] \leq \frac{\mathbb{E}[|X|]}{a}$.*
- *Prove extended version of Markov's inequality. Namely, if $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is monotonically increasing function, then $\mathbb{P}[|X| \geq c] \leq \frac{\mathbb{E}[\phi(|X|)]}{\phi(c)}$ for all $c > 0$ such that $\phi(c) \neq 0$.*
- *Conclude Chebyshev's inequality. More precisely, assume that X is an integrable random variable with mean μ and a finite non-zero variance σ^2, $a > 0$ and show that $\mathbb{P}[|X - \mu| \geq a] \leq \frac{\sigma^2}{a^2}$. Recall that the variance is given by $\sigma^2 = \mathbb{E}[(X - \mu)^2]$.*
- *(Exponential inequalities) Let X be a random variable, $b \in \mathbb{R}$. Prove (using Markov's inequality applied to suitable non-negative random variables) that $\mathbb{P}[X \geq b] \leq e^{-tb}M(t)$ for all $t \geq 0$ and $\mathbb{P}[X \leq b] \leq e^{-tb}M(t)$ for all $t \leq 0$, where $M(t) = \mathbb{E}[e^{tX}]$ is the moment generating function (MGF) of X.*
  Note that we can optimize the bounds by taking in the above inequalities infimum over $t \geq 0$, $t \leq 0$, respectively. The resulting inequalities are called Chernoff bounds in some literature (sometimes more specific results are called that, see Exercise 4).

Proof. 1)-2) Let us directly prove the more general second statement, which implies the first one by taking $\phi(x) = x$. Since ϕ is non-negative and monotonically increasing,

$$\phi(|X|) \geq \phi(|X|)\mathbf{1}_{\{|X| \geq c\}} \geq \phi(c)\mathbf{1}_{\{|X| \geq c\}}.$$

By taking expectation and dividing all sides of inequality by $\phi(c)$ we get the result.

3) Apply 2) with random variable $X - \mu$ and $\phi(x) = x^2$.

4) Let $t \geq 0, s \leq 0$, then by strict monotonicity of exponential function and 1),

$$\mathbb{P}[X \geq b] = \mathbb{P}[e^{tX} \geq e^{tb}] \leq e^{-tb}M(t),$$
$$\mathbb{P}[X \leq b] = \mathbb{P}[e^{sX} \geq e^{sb}] \leq e^{-sb}M(s).$$

□

Exercise 2 (p-moments via tails and applications)**.**

Let X be a random variable, $p \in (0, \infty)$. Show that

$$\mathbb{E}[|X|^p] = \int_0^\infty pt^{p-1}\mathbb{P}[|X| > t]\mathrm{d}t$$

whenever the rhs is finite.

Suppose now that X is a random variable satisfying $\mathbb{P}[|X| \geq t] \leq 2e^{-bt}$ for some $b > 0$. Show that $\mathbb{E}[|X|^n] \leq 2b^{-n}n!$ and conclude that $\|X\|_{\mathrm{L}^n} \leq \frac{cn}{b}$ for some uniform constant $c > 0$ (Stirling's formula might be helpful).

Proof. Let us first prove that for a non-negative random variable Y, $\mathbb{E}[Y] = \int_0^\infty \mathbb{P}[X > t]\mathrm{d}t$. This follows easily since for any $x \geq 0$, $x = \int_0^x 1\mathrm{d}t = \int_0^\infty \mathbf{1}_{t<x}\mathrm{d}t$. By substituting x with Y

and taking the expectation on both sides gives the claim.
Now by the claim and the change of variable $u = t^{1/p}$ we obtain

$$\mathbb{E}[|X|^p] = \int_0^\infty \mathbb{P}[|X|^p > t]\mathrm{d}t = \int_0^\infty pu^{p-1}\mathbb{P}[|X| > u]\mathrm{d}u.$$

Assume now that $\mathbb{P}[|X| \geq t] \leq 2e^{-bt}$ for some $b > 0$. Then,

$$\begin{aligned}\mathbb{E}[|X|^n] &= \int_0^\infty nu^{n-1}\mathbb{P}[|X| > u]\mathrm{d}u \leq 2n\int_0^\infty u^{n-1}e^{-bu}\mathrm{d}u \\ &= 2nb^{-n}\int_0^\infty y^{n-1}e^{-y}\mathrm{d}y = 2nb^{-n}\Gamma(n) = 2b^{-n}n!.\end{aligned}$$

By Stirling's approximation we have $n! = \sqrt{2\pi n}n^n e^{-n}e^{R_n}$, where R_n is some number satisfying $0 \leq R_n \leq 1/(12n)$. Combining with the above bound we get

$$\|X\|_{\mathrm{L}^n} \leq (2b^{-n}n!)^{1/n} = (2\sqrt{2\pi n}e^{R_n})\frac{n}{be} \leq \frac{cn}{b}.$$

The last inequality follows for some universal constant since R_n is bounded and $n^{1/(2n)}$ is bounded universally since as $n \to \infty$, it converges to 1. □

Exercise 3 (Tails of normal distribution)**.**
Let $X \sim \mathcal{N}(\mu, \sigma^2)$ be a Gaussian random variable.

- *Using Chernoff bounds (optimized exponential inequalities) estimate $\mathbb{P}[X - \mu \geq c]$ for $c \geq 0$.*
- *Now suppose that that $\mu = 0, \sigma^2 = 1$. Show that for all $c > 0$,*

$$\left(\frac{1}{c} - \frac{1}{c^3}\right)\frac{1}{\sqrt{2\pi}}e^{-c^2/2} \leq \mathbb{P}[X \geq c] \leq \frac{1}{c}\frac{1}{\sqrt{2\pi}}e^{-c^2/2}.$$

  Hint: use density function, for the lower bound $g(x) = (1 - 3x^{-4})\mathbf{1}_{x>c} \leq 1$ might be helpful.

Compare bounds of first and second bullet points in the case $\mu = 0, \sigma^2 = 1$.

Proof.

- By Chernoff bounds,

$$\begin{aligned}\mathbb{P}[X - \mu \geq c] &\leq \inf_{t\geq 0}(e^{-t(\mu+c)}M(t)) = \exp\left(-\sup_{t\geq 0}((\mu + c)t - \log M(t))\right) \\ &= \exp\left(-\sup_{t\geq 0}(ct - \sigma^2t^2/2)\right) = \exp\left(-\frac{c^2}{2\sigma^2}\right).\end{aligned}$$

  Verify the last inequality by finding the critical point of the function $t \mapsto ct - \sigma^2t^2/2$.
- Let us start by showing the upper bound,

$$\begin{aligned}\mathbb{P}[X \geq c] &= \frac{1}{\sqrt{2\pi}}\int_c^\infty e^{-x^2/2}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-c^2/2}e^{-cy}e^{-y^2/2}\mathrm{d}y \\ &\leq \frac{e^{-c^2/2}}{\sqrt{2\pi}}\int_0^\infty e^{-cy}\mathrm{d}y = \frac{1}{c}\frac{1}{\sqrt{2\pi}}e^{-c^2/2}.\end{aligned}$$

  The lower bound follows from the identity

$$\int_c^\infty (1 - 3x^{-4})e^{-x^2/2}\mathrm{d}x = \left(\frac{1}{c} - \frac{1}{c^3}\right)e^{-c^2/2}.$$

The second upper bound is sharper for large c's, but for very small c's it is weaker. $\square$

Exercise 4 (Large deviations from LLN, Chernoff)**.**
Let $X_i \in \mathrm{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ be iid random variables with expectation μ and MGF $M(t) = \mathbb{E}[e^{tX_1}]$. Let $S_n := \sum_{i=1}^n X_i$. Show that for all $n \in \mathbb{N}$ and $a \in \mathbb{R}$,

$$\begin{cases} \mathbb{P}\left[\frac{S_n}{n} \geq a\right] \leq e^{-nI(a)} & \text{if } a \geq \mu, \\ \mathbb{P}\left[\frac{S_n}{n} \leq a\right] \leq e^{-nI(a)} & \text{if } a \leq \mu, \end{cases}$$

where $I(a) := \sup_{t\in\mathbb{R}}(at - \log M(t))$. Note that the supremum is taken over all $\mathbb{R}$.
You can follow the following steps:

- *Reduce to the case $\mu = 0$;*
- *Exponential Markov inequality;*
- *Use Jensen's inequality to conclude that the supremum in $I(a)$ can be taken over all of $\mathbb{R}$.*

Proof. Let us prove the case $a \geq \mu$. The other case follows fully analogously. We proceed in three steps:

- We can assume wlog that $\mu = 0$, since otherwise we can consider $Y_i = X_i - \mu$, which are again iid random variables, but now centered. And the statement for Y_i with $a \geq 0$ rewritten in terms of X_i looks like $\mathbb{P}\left[\frac{S_n}{n} \geq a + \mu\right] \leq e^{-nI(a+\mu)}$, so for $a' = a + \mu \geq \mu$ we get the result.
- For all $t \geq 0, n \in \mathbb{N}$ by exponential Markov's inequality,
$$\mathbb{P}\left[\frac{S_n}{n} \geq a\right] = \mathbb{P}[S_n \geq na] \leq e^{-tna}\mathbb{E}[e^{tS_n}] = e^{-tna}M(t)^n = e^{-(at-\log M(t))n}.$$
- Now we want to optimize the bound over $t \geq 0$. Namely,
$$\mathbb{P}\left[\frac{S_n}{n} \geq a\right] \leq \inf_{t\geq 0} e^{-(at-\log M(t))n} = \exp\left(-n \sup_{t\geq 0}(at - \log M(t))\right).$$
It is left to check that $\sup_{t\geq 0}(at - \log M(t)) = \sup_{t\in\mathbb{R}}(at - \log M(t)) = I(a)$. The latter is true since for $t < 0$ and $a \geq \mu = 0$ and by Jensen's inequality we get
$$at - \log M(t) \leq -\log \mathbb{E}[e^{tX_1}] \leq \mathbb{E}[-\log(e^{tX_1})] = -t\mu = 0 = a \cdot 0 - \log M(0).$$
Thus, it suffices to consider the supremum over non-negative numbers.

$\square$