Saarland University
Faculty of Mathematics and Computer Science
Department of Mathematics
Prof. Dr. Roland Speicher
Dr. Johannes Hoffmann

Exercises for the lecture
High Dimensional Analysis: Random Matrices and Machine Learning
Summer term 2023
Sheet 3
Hand-in: Friday, 26.05.2023, 22:00 Uhr via CMS

Exercise 1 (2+5+3+5 points). Fix $p \in [0, 1]$. Let $y_1, \ldots, y_n$ be independent Bernoulli random variables with

$$\mathrm{P}\{y_i = 1\} = p, \qquad \mathrm{P}\{y_i = 0\} = 1 - p$$

and consider $y := y_1 + \ldots + y_n$. Let $\delta > 0$.

(a) Show that $E[\exp(\lambda y_i)] \leq \exp(p(\exp(\lambda) - 1))$ holds for every $\lambda > 0$.

(b) Conclude the following classic Chernoff bound:

$$\mathrm{P}\{y \geq (1+\delta)np\} \leq \left(\frac{\exp(\delta)}{(1+\delta)^{1+\delta}} \right)^{np}.$$

Hint: we know from class that

$$\mathrm{P}\{y \geq \alpha\} \leq \exp(-\lambda\alpha) \prod_{i=1}^{n} E[\exp(\lambda y_i)] \quad \textit{for any} \quad \lambda > 0.$$

(c) Assume you are rolling a fair six-sided dice n times. Apply (b) to estimate the probability to roll a six at least 70% of the experiments.

(d) Compare the estimate of (b) with the estimates from the Markov and the Chebyshev Inequalities. Run a simulation of the experiment in (c) to test how tight the predictions of the three bounds are for $n \in \{1, 5, 25, 100\}$ (use 1,000 repetitions of each experiment to get sensible data).

please turn over

Exercise 2 (6 + 6 points).

(a) Let x be a sub-exponential centred random variable, i.e. a one-dimensional real random variable with mean zero and such that there exists a constant $c > 0$ satisfying

$$E[\exp(\lambda x)] \leq \exp(c^2\lambda^2) \quad \text{for all} \quad |\lambda| \leq \frac{1}{c}.$$

Prove that we then have

$$\mathrm{P}\left\{x \geq \alpha\right\} \leq \begin{cases} \exp\left(-\dfrac{\alpha^2}{4c^2}\right), & \text{if } \alpha \leq 2c, \\ \exp\left(-\dfrac{\alpha}{2c}\right), & \text{if } \alpha > 2c. \end{cases}$$

(b) In the proof of Theorem 2.2. we have shown that for a standard Gaussian random vector $x \sim N(0, I_p)$ we have the concentration

$$\mathrm{P}\left\{\left|\|x\|^2 - p\right| \geq \varepsilon\sqrt{p}\right\} \leq 2\exp\left(-\frac{\varepsilon^2}{16}\right).$$

However, this was only for the case where $\varepsilon\sqrt{p} \leq p$, but the proof actually works for all $\varepsilon\sqrt{p} \leq 2p$. Complement this now by a corresponding estimate also for the case of large deviations $\varepsilon\sqrt{p} > 2p$.

Exercise 3 (7 points). Show that every bounded random variable is sub-Gaussian: let x be a real random variable that is bounded, i.e., for some $a, b \in \mathbb{R}$ we have

$$\mathrm{P}\{a \leq x \leq b\} = 1.$$

Assume also that x is centred, i.e., $E[x] = 0$. Then there exists a $c \in \mathbb{R}$ such that we have for all λ

$$E[\exp(\lambda x)] \leq \exp(c\lambda^2).$$

The best constant is actually given by $c = \frac{(b-a)^2}{8}$, but here we are satisfied with any bound.

Hint: for symmetric distributions the situation is easy; in the non-symmetric case one might try to symmetrize the situation by going over, as in our proof of Theorem 3.2., from $E[\exp(\lambda x)]$ *to* $E[\exp(\lambda(x - y))]$, *where y is an independent copy of x.*

Exercise 4 (2 + 4 points). Consider the following statement: if $h := f \circ g$ is the composition of two convex functions $f, g : \mathbb{R} \to \mathbb{R}$, then h is also convex.

(a) Give a counterexample to show that the statement is not true in general.

(b) Repair the statement by introducing an additional assumption on f and g and prove the statement under this assumption.