

# High-Dimensional Statistics

Vitaliy Pozdnyakov

March 2020

Solutions to exercises from the book "High-Dimensional Probability", Roman Vershynin, 2020  
<https://www.math.uci.edu/~rvershyn/>

Notation:

- cc – coffee cup
- $\sum := \sum_{i=1}^N$
- $\prod := \prod_{i=1}^N$

## 1 Exercise 2.1.4 (1 cc)

Let  $g \sim N(0, 1)$ . Show that for all  $t \geq 1$ , we have

$$\mathbb{E} g^2 \mathbf{1}\{g > t\} = t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\{g > t\} \leq \left(t + \frac{1}{t}\right) \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

### 1.1 Solution

$$\mathbb{E} g^2 \mathbf{1}\{g > t\} = \int_{-\infty}^{+\infty} x^2 \mathbf{1}\{x > t\} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} x^2 e^{-x^2/2} dx$$

Make the substitution  $u = x$  and  $dv = x e^{-x^2/2} dx$ , then  $du = dx$  and

$$v = \int dv = \int x e^{-x^2/2} dx = \left[ z = -\frac{x^2}{2}; dz = -x dx \right] = - \int e^z dz = -e^z = -e^{-x^2/2}$$

Next, integrate by parts

$$\int_a^b u dv = uv \Big|_a^b - \int_a^b v du$$

$$\frac{1}{\sqrt{2\pi}} \int_t^{+\infty} x^2 e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \left( -x e^{-x^2/2} \Big|_t^{+\infty} \right) + \int_t^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\{g > t\}$$

Next, apply the theorem "Tails of the normal distribution":

$$\mathbb{P}\{g > t\} \leq \frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} e^{t^2/2}$$

Thereby,

$$t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\{g > t\} \leq t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} e^{t^2/2} = \left(t + \frac{1}{t}\right) \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

## 2 Exercise 2.2.3 (1 cc)

(Bounding the hyperbolic cosine)

Show that

$$\cosh(x) \leq \exp(x^2/2) \text{ for all } x \in \mathbb{R}.$$

### 2.1 Solution

Apply Taylor's expansion for the left side of inequation

$$\cosh(x)' = \sinh(x); \cosh(x)'' = \cosh(x); \cosh(x)''' = \sinh(x); \dots$$

$$\cosh(0) = 1; \sinh(0) = 0$$

Then

$$\begin{aligned} \cosh(x) &= \cosh(0) + \frac{\sinh(0)}{1!}x^1 + \frac{\cosh(0)}{2!}x^2 + \frac{\sinh(0)}{3!}x^3 + \dots \\ &= 1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + O(x^8) \end{aligned}$$

For the right side

$$\exp(x^2/2)' = \exp(x^2/2)x; \exp(x^2/2)'' = \exp(x^2/2)(x^2 + 1); \exp(x^2/2)''' = \exp(x^2/2)x(x^2 + 3); \dots$$

$$\exp(0^2/2) = 1; \exp(0^2/2)' = 0; \exp(0^2/2)'' = 1; \exp(0^2/2)''' = 0; \dots$$

Then

$$\begin{aligned} \exp(x^2/2) &= \exp(0^2/2) + \frac{\exp(0^2/2)'0}{1!}x^1 + \frac{\exp(0^2/2)''(0^2 + 1)}{2!}x^2 + \frac{\exp(0^2/2)'''(0^2 + 3)}{3!}x^3 + \dots \\ &= 1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{384} + O(x^8) \end{aligned}$$

Obviously,

$$1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + O(x^8) \leq 1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{384} + O(x^8)$$

Thus,

$$\cosh(x) \leq \exp(x^2/2) \text{ for all } x \in \mathbb{R}.$$

## 3 Exercise 2.2.7. (2 cc)

Let  $X_1, \dots, X_N$  be independent random variables. Assume that  $X_i \in [m_i, M_i]$  for every  $i$ . Prove that for any  $t > 0$

$$\mathbb{P} \left\{ \sum (X_i - \mathbb{E}X_i) \geq t \right\} \leq \exp \left( - \frac{2t^2}{\sum (M_i - m_i)^2} \right)$$

### 3.1 Solution

Let  $\xi_i = X_i - \mathbb{E}X_i$ . Apply Markov's inequality and a property of the exp function

$$\mathbb{P}\left\{\sum \xi_i \geq t\right\} = \mathbb{P}\left\{e^{\lambda \sum \xi_i} \geq e^{\lambda t}\right\} \leq e^{-\lambda t} \mathbb{E}\left[e^{\lambda \sum \xi_i}\right] \text{ for any } \lambda > 0$$

Apply independence assumption

$$\mathbb{E}\left[e^{\lambda \sum \xi_i}\right] = \mathbb{E}\left[\prod e^{\lambda \xi_i}\right] = \prod \mathbb{E}\left[e^{\lambda \xi_i}\right]$$

Apply Hoeffding's Lemma

$$\mathbb{E}\left[e^{\lambda \xi_i}\right] \leq e^{\frac{\lambda^2 (M_i - m_i)^2}{8}}$$

Thus,

$$\mathbb{P}\left\{\sum \xi_i \geq t\right\} \leq e^{-\lambda t} \prod e^{\frac{\lambda^2 (M_i - m_i)^2}{8}} = e^{-\lambda t + \frac{\lambda^2}{8} \sum (M_i - m_i)^2}$$

Let  $\lambda = 4t / \sum (M_i - m_i)^2$ , then

$$\mathbb{P}\left\{\sum (X_i - \mathbb{E}X_i) \geq t\right\} \leq \exp\left(-\frac{2t^2}{\sum (M_i - m_i)^2}\right)$$

## 4 Exercise 2.2.8. (2 cc)

(Boosting randomized algorithms)

Imagine we have an algorithm for solving some decision problem (e.g. is a given number  $p$  a prime?). Suppose the algorithm makes a decision at random and returns the correct answer with probability  $1/2 + \delta$  with some  $\delta > 0$ , which is just a bit better than a random guess. To improve the performance, we run the algorithm  $N$  times and take the majority vote. Show that, for any  $\epsilon \in (0, 1)$ , the answer is correct with probability at least  $1 - \epsilon$ , as long as

$$N \leq \frac{1}{2\delta^2} \ln(1/\epsilon)$$

### 4.1 Solution

Let  $X_1, X_2, \dots, X_N$  be a random sequence with probabilities

$$\mathbb{P}\{X_i = 0\} = 1/2 - \delta$$

$$\mathbb{P}\{X_i = 1\} = 1/2 + \delta$$

Then the event {The answer is incorrect} can be written as

$$\begin{aligned} \left\{\sum X_i \geq N/2\right\} &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \sum (1/2 - \mathbb{E}X_i)\right\} \\ &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \sum (1/2 - 1/2 + \delta)\right\} \\ &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \delta N\right\} \end{aligned}$$

Apply Hoeffding inequality

$$\mathbb{P}\left\{\sum (X_i - \mathbb{E}X_i) \geq t\right\} \leq \exp\left(-\frac{2t^2}{\sum (M_i - m_i)^2}\right)$$

Make substitutions  $t = \delta N$ ,  $M_i = 1$ ,  $m_i = 0$

$$\mathbb{P}\left\{\sum(X_i - \mathbb{E}X_i) \geq \delta N\right\} \leq \exp(-2\delta^2 N)$$

We can observe that

$$\left\{N \leq \frac{1}{2\delta^2} \ln(1/\epsilon)\right\} \Rightarrow \{\exp(-2\delta^2 N) < \epsilon\}$$

Thereby,  $\mathbb{P}\{\text{The answer is incorrect}\} \leq \epsilon$  and  $\mathbb{P}\{\text{The answer is correct}\} \geq 1 - \epsilon$  under that constraint.

## 5 Exercise 2.2.10 (2 cc)

Let  $X_1, \dots, X_N$  be non-negative independent random variables with continuous distributions. Assume that the densities of  $X_i$  are uniformly bounded by 1.

- (a) Show that the MGF of  $X_i$  satisfies  $\mathbb{E}e^{-tX_i} \leq 1/t$  for all  $t > 0$ .  
(b) Deduce that, for any  $\epsilon > 0$ , we have  $\mathbb{P}\{\sum X_i \leq \epsilon N\} \leq (e\epsilon)^N$ .

### 5.1 Solution (a)

Let  $\phi_i(x)$  be a PDF of  $X_i$

$$\mathbb{E} \exp(-tX_i) = \int_0^\infty e^{-tx} \phi_i(x) dx$$

Since  $\phi_i(x) \leq 1$  for all  $i$  and  $x$ , then

$$\int_0^\infty e^{-tx} \phi_i(x) dx \leq \int_0^\infty e^{-tx} dx = \frac{1}{t} \text{ for all } t > 0$$

### 5.2 Solution (b)

Rewrite the probability as

$$\mathbb{P}\left\{\sum X_i \leq \epsilon N\right\} = \mathbb{P}\left\{\sum -X_i/\epsilon \geq -N\right\}$$

Apply Markov's inequality and a property of the exp function

$$\mathbb{P}\left\{\sum -X_i/\epsilon \geq -N\right\} = \mathbb{P}\left\{\exp\left(\sum -X_i/\epsilon\right) \geq \exp(-N)\right\} \leq e^N \prod \mathbb{E}e^{-X_i/\epsilon}$$

Apply the result from the section (a) and then

$$\mathbb{P}\left\{\sum X_i \leq \epsilon N\right\} \leq e^N \prod \epsilon = (e\epsilon)^N$$

## 6 Exercise 2.3.2 (2 cc)

(Chernoff's inequality: lower tails)

Modify the proof of Theorem 2.3.1 to obtain the following bound on the lower tail. For any  $t < \mu$ , we have

$$\mathbb{P}\{S_N \leq t\} \leq e^{-\mu} \left(\frac{e\mu}{t}\right)^t$$

## 6.1 Solution

Since  $t < \mu$ ,  $t > 0$  and  $S_N > 0$  we can rewrite the probability as

$$\mathbb{P}\{S_N \leq t\} = \mathbb{P}\{(t/\mu)^{S_N} \geq (t/\mu)^t\}$$

Apply Markov's inequality

$$\mathbb{P}\{(t/\mu)^{S_N} \geq (t/\mu)^t\} \leq (\mu/t)^t \mathbb{E}(t/\mu)^{S_N}$$

Consider the expectation

$$\mathbb{E}(t/\mu)^{S_N} = \prod \mathbb{E}(t/\mu)^{X_i}$$

$$\mathbb{E}(t/\mu)^{X_i} = (t/\mu)^1 p_i + (t/\mu)^0 (1 - p_i) = 1 + [(t/\mu) - 1] p_i \leq \exp\left(\left[\frac{t}{\mu} - 1\right] p_i\right)$$

$$\prod \mathbb{E}(t/\mu)^{X_i} \leq \exp\left(\left[\frac{t}{\mu} - 1\right] \sum p_i\right) = \exp\left(\left[\frac{t}{\mu} - 1\right] \mu\right) = e^t e^{-\mu}$$

Thereby

$$\mathbb{P}\{S_N \leq t\} \leq e^{-\mu} \left(\frac{e\mu}{t}\right)^t$$

## 7 Exercise 2.3.3 (2 cc)

(Poisson tails)

Let  $X \sim \text{Pois}(\lambda)$ . Show that for any  $t > \lambda$ , we have

$$\mathbb{P}\{X \geq t\} \leq e^{-\lambda} \left(\frac{e\lambda}{t}\right)^t$$

## 7.1 Solution

Chernoff's inequality

$$\mathbb{P}\{S_N \geq t\} \leq e^{-\mu} \left(\frac{e\mu}{t}\right)^t$$

Assume that  $N \rightarrow \infty$ ,  $\max p_i \rightarrow 0$  and  $\mathbb{E}S_N \rightarrow \lambda < \infty$ . Then, using Poisson limit theorem we have

$$S_N \rightarrow \text{Pois}(\lambda)$$

Thereby

$$\mathbb{P}\{X \geq t\} \leq e^{-\lambda} \left(\frac{e\lambda}{t}\right)^t$$

## 8 Exercise 2.3.5 (3 cc)

(Chernoff's inequality: small deviations)

Show that, in the setting of Theorem 2.3.1, for  $\delta \in (0, 1]$  we have

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq 2e^{-c\mu\delta^2}$$

where  $c > 0$  is an absolute constant.

## 8.1 Solution

Consider Chernoff's inequality with  $t > \mu$  (upper tails). Let  $t = \mu + \delta\mu$ , so

$$\mathbb{P}\{S_N - \mu \geq \delta\mu\} \leq e^{-\mu} \left( \frac{e}{1+\delta} \right)^{\mu+\delta\mu}$$

In opposite, if  $t < \mu$  and  $t = \mu - \delta\mu$  (lower tails) we obtain

$$\mathbb{P}\{S_N - \mu \leq -\delta\mu\} \leq e^{-\mu} \left( \frac{e}{1-\delta} \right)^{\mu-\delta\mu}$$

Apply addition theorem of probability for lower and upper tails

$$\mathbb{P}(\{S_N - \mu \geq \delta\mu\} \cup \{S_N - \mu \leq -\delta\mu\}) = \mathbb{P}\{|S_N - \mu| \geq \delta\mu\}$$

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq e^{-\mu} \left( \frac{e}{1-\delta} \right)^{\mu-\delta\mu} + e^{-\mu} \left( \frac{e}{1+\delta} \right)^{\mu+\delta\mu} \leq 2e^{-\mu} \left( \frac{e}{1+\delta} \right)^{\mu+\delta\mu}$$

Consider the inequality

$$\begin{aligned} 2e^{-\mu} \left( \frac{e}{1+\delta} \right)^{\mu+\delta\mu} &\leq 2e^{-c\mu\delta^2} \\ \frac{e^{\delta\mu}}{(1+\delta)^{\mu+\delta\mu}} &\leq e^{-c\mu\delta^2} \\ e^{\delta+c\delta^2} &\leq (1+\delta)^{1+\delta} \end{aligned}$$

Take log

$$\begin{aligned} \delta + c\delta^2 &\leq (1+\delta) \ln(1+\delta) \\ c &\leq \frac{(1+\delta) \ln(1+\delta) - \delta}{\delta^2} \end{aligned}$$

Thereby,

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq 2e^{-c\mu\delta^2}$$

for all  $c \leq [(1+\delta) \ln(1+\delta) - \delta]/\delta^2$ .

## 9 Exercise 2.3.8 (2 cc)

(Normal approximation to Poisson)

Let  $X \sim Pois(\lambda)$ . Show that, as  $\lambda \rightarrow \infty$ , we have

$$\frac{X - \lambda}{\sqrt{\lambda}} \rightarrow \mathcal{N}(0, 1) \text{ in distribution.}$$

## 9.1 Solution

Rewrite the RV  $X$  as a sum  $Z_1 + \dots + Z_N$  where  $Z_i \sim Pois(\lambda/N)$  and apply Lindeberg-Levy central limit theorem

$$\frac{X - \mathbb{E}X}{\sqrt{\text{Var}(X)}} = \frac{X - N\mathbb{E}Z_i}{\sqrt{N\text{Var}(Z_i)}} = \frac{X - \lambda}{\sqrt{\lambda}} \rightarrow \mathcal{N}(0, 1) \text{ as } N \rightarrow \infty$$

## 10 Exercise 2.4.2 (1 cc)

(Bounding the degrees of sparse graphs)

Consider a random graph  $G \sim G(n, p)$  with expected degrees  $d = O(\log n)$ . Show that with high probability (say, 0.9), all vertices of  $G$  have degrees  $O(\log n)$ .

### 10.1 Solution

Since  $d_i > 0$  and  $n > 0$ , then for any  $i$  we have  $d_i = O(\log n) \equiv \exists C_i : d_i \leq C_i \log n$  and we need to prove the following statement

$$\mathbb{P}\{\forall i : d_i \leq C_i \log n\} \geq 0.9$$

where  $C_i$  is a constant. We can rewrite that probability as

$$\mathbb{P}\{\exists i : d_i \geq C_i \log n\} \leq \sum \mathbb{P}\{d_i \geq C_i \log n\} \leq 0.1$$

Apply Chernoff's inequality for upper bound (since  $d \leq C_i \log n$ )

$$\sum \mathbb{P}\{d_i \geq C_i \log n\} \leq ne^{-d} \left( \frac{ed}{C_i \log n} \right)^{C_i \log n} \leq 0.1$$

Note that there exists  $C_d$  such that  $d \leq C_d \log n$  and then

$$ne^{-C_d \log n} \left( \frac{eC_d}{C_i} \right)^{C_i \log n} = n^{C_i + C_d + 1} \left( \frac{C_d}{C_i} \right)^{C_i \log n} \leq 0.1$$

Hence, the statement holds for any  $C_i$  such that

$$C_i \geq \frac{C_d \log n + \log n + \log 10}{\log n} \left[ W \left( \frac{C_d \log n + \log n + \log 10}{C_d e \log n} \right) \right]^{-1}$$

where  $W(\cdot)$  is the product log function.

## 11 Exercise 2.4.3 (2 cc)

(Bounding the degrees of very sparse graphs)

Consider a random graph  $G \sim G(n, p)$  with expected degrees  $d = O(1)$ . Show that with high probability (say, 0.9), all vertices of  $G$  have degrees

$$O\left(\frac{\log n}{\log \log n}\right)$$

### 11.1 Solution

Apply the same approach as in the previous exercise and obtain

$$\sum \mathbb{P}\left\{d_i \geq C_i \frac{\log n}{\log \log n}\right\} \leq ne^{-C_d} \left( \frac{eC_d \log \log n}{C_i \log n} \right)^{(C_i \log n)/(\log \log n)} \leq 0.1$$

Hence, the statement holds for any  $C_i$  such that

$$C_i \geq \frac{(\log n - C_d + \log 10) \log \log n}{\log n} \left[ W \left( \frac{\log n - C_d + \log 10}{C_d e} \right) \right]^{-1}$$

where  $W(\cdot)$  is the product log function.