Concentration Inequalities : Exercises solutions

Ludovic Schwartz

Contents

Chapter 2

Basic Inequalities

Exercise 2.1

Let $\mathbf{M}Z$ be a median of the square integrable variable Z (i.e $\mathbb{P}[Z \geq \mathbf{M}Z] \geq \frac{1}{2}$ and $\mathbb{P}[Z \leq \mathbf{M}Z] \geq \frac{1}{2}$). Show that :

$$|\mathbb{E}[Z] - \mathbf{M}Z| \leq \sqrt{Var(Z)}$$

Solution : We have the following inequalities :

$$\begin{aligned}
|\mathbb{E}[Z] - \mathbf{M}Z| &= |\mathbb{E}[Z - \mathbf{M}Z]| \\
&\leq \mathbb{E}[|Z - \mathbf{M}Z|] \\
&\leq \mathbb{E}[|Z - \mathbb{E}[Z]|] \\
&\leq \sqrt{\mathbb{E}[(Z - \mathbb{E}[Z])^2]}
\end{aligned}$$

Where the second inequality comes from the optimality of the median for the mean absolute error($\mathbf{M}Z \in \arg\min_{c \in \boldsymbol{R}} \mathbb{E}[|Z - c|]$) and the third inequality is Cauchy-Schwarz. □

Exercise 2.2

Let X be a random variable with median $\mathbf{M}X$, such that there exist positive constants a and b such that for all $t \geq 0$,

$$\mathbb{P}[|X - \mathbf{M}X| > t] \leq a \exp -\frac{t^2}{b}$$

Show that $|\mathbf{M}X - \mathbb{E}[X]| \leq \min\left(\sqrt{ab}, a\sqrt{b\pi}/2\right)$

Solution : We have the following :

$$\begin{aligned}
|\mathbb{E}[X] - \mathbf{M}X| &= |\mathbb{E}[X - \mathbf{M}X]| \\
&\leq \mathbb{E}[|X - \mathbf{M}X|] \\
&= \int_0^\infty \mathbb{P}[|X - \mathbf{M}X| > t]\, dt \\
&\leq \int_0^\infty a \exp -\frac{t^2}{b} \\
&\leq a \cdot \sqrt{b\pi}/2
\end{aligned}$$

For the second inequality, we will bound the variance and use the first exercise. We know that

$$Var(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \min_{c \in \boldsymbol{R}} \mathbb{E}\left[(X - c)^2\right]$$

In particular, plugging in the value of the median as c, we get :

$$\begin{aligned}
Var(X) &\leq \mathbb{E}\left[(X - \mathbf{M}X)^2\right] \\
&= \int_0^\infty \mathbb{P}\left[(X - \mathbf{M}X)^2 > t\right] dt \\
&= \int_0^\infty \mathbb{P}\left[|X - \mathbf{M}X| > \sqrt{t}\right] dt \\
&\leq \int_0^\infty a \cdot \exp -\frac{t}{b} dt \\
&\leq ab
\end{aligned}$$

Then using the result of the first exercise we have :

$$|\mathbb{E}[X] - \mathbf{M}X| \leq \sqrt{Var(X)} \leq \sqrt{ab}$$

□

Exercise 2.3 (Chebyshev-Cantelli Inequality)

Prove the following one-sided improvement of Chebyshev's inequality : for any real-valued random variable Y and $t > 0$,

$$\mathbb{P}[Y - \mathbb{E}[Y] \geq t] \leq \frac{Var(Y)}{Var(Y) + t^2}$$

Solution : We start by defining $Z := Y - \mathbb{E}[Y]$ and $v := \mathbb{E}\left[Z^2\right] = Var(Y)$

We are now trying to prove $\mathbb{P}[Z \geq t] \leq \frac{v}{v+t^2}$ We have for any $u \geq 0$

$$\begin{aligned}
\mathbb{P}[Z \geq t] &= \mathbb{P}[Z + u \geq t + u] \\
&\leq \mathbb{P}\left[(Z+u)^2 \geq (t+u)^2\right] \\
&\overset{\text{(Markov)}}{\leq} \frac{\mathbb{E}\left[(Z+u)^2\right]}{(t+u)^2} \\
&\leq \frac{v+u^2}{(t+u)^2}
\end{aligned}$$

We're now going to optimize the value of this fraction over u. Let

$$\begin{aligned}
\varphi : \boldsymbol{R}^+ &\longrightarrow \boldsymbol{R} \\
u &\longmapsto \frac{v+u^2}{(t+u)^2}
\end{aligned}$$

We compute it's derivative :

$$\begin{aligned}
\varphi\prime(u) &= \frac{2u \cdot (t+u)^2 - (v+u^2) \cdot 2(t+u)}{(t+u)^4} \\
&= \frac{(t+u) \cdot (2ut - 2v)}{(t+u)^4}
\end{aligned}$$

In particular, φ reaches its minimum in $u^* := \frac{v}{t}$. We plug that value in the previous inequality :

$$\begin{aligned}
\mathbb{P}[Z \geq t] \leq \varphi(u^*) &= \frac{v + \frac{v}{t}^2}{(t+\frac{v}{t})^2} \\
&= \frac{v \cdot (t^2+v)}{(t^2+v)^2} \\
&= \frac{v}{v+t^2}
\end{aligned}$$

$\square$

Exercise 2.4 (PALEY-ZIGMUND INEQUALITY)

Show that if Y is a nonnegative random variable, then for any $a \in (0,1)$,

$$\mathbb{P}[Y - \mathbb{E}[Y]] \geq (1-a)^2 \frac{(\mathbb{E}[Y])^2}{\mathbb{E}[Y^2]}$$

Solution : We have :

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}\left[Y\mathbb{I}_{Y<a\mathbb{E}[Y]}\right] + \mathbb{E}\left[Y\mathbb{I}_{Y\geq a\mathbb{E}[Y]}\right] \\
&\leq \mathbb{E}\left[a\mathbb{E}[Y]\,\mathbb{I}_{Y<a\mathbb{E}[Y]}\right] + \sqrt{\mathbb{E}[Y^2]\cdot\mathbb{E}\left[\mathbb{I}^2_{Y\geq a\mathbb{E}[Y]}\right]} \\
&\leq a\mathbb{E}[Y] + \sqrt{\mathbb{E}[Y^2]\cdot\mathbb{E}\left[\mathbb{I}_{Y\geq a\mathbb{E}[Y]}\right]} \\
(1-a)\mathbb{E}[Y] &\leq \sqrt{\mathbb{E}[Y^2]\cdot\mathbb{P}[Y\geq a\mathbb{E}[Y]]}
\end{aligned}$$

Where the first inequality comes from Cauchy Schwarz on the right term. Taking the square and rearranging, we get the claimed result.

□

Exercise 2.5 (Moments VS. Chernoff Bounds)

Show that moment bounds for tail probabilities are always better than Cramér-Chernoff bounds. More precisely, let Y be a nonnegative random variable and let $t > 0$. The best moment bound for the tail probability $\mathbb{P}[Y \geq t]$ is $\min_q \mathbb{E}[Y^q]\, t^{-q}$ where the minimum is taken over all positive integers. The best Cramér-Chernoff bound is $\inf_{\lambda>0} \mathbb{E}[\exp \lambda(Y-t)]$. Prove that :

$$\min_q \mathbb{E}[Y^q]\, t^{-q} \leq \inf_{\lambda>0} \mathbb{E}[\exp \lambda(Y-t)]$$

Solution : We denote $m := \min_q \mathbb{E}[Y^q]\, t^{-q}$. In particular, we have that for any integer q, $\mathbb{E}[Y^q] \geq m \cdot t^q$. We will now fix a $\lambda > 0$. We have :

$$\begin{aligned}
\mathbb{E}[\exp \lambda(Y-t)] &= e^{-\lambda t}\mathbb{E}\left[\sum_{q=0}^{\infty} \frac{(\lambda Y)^q}{q!}\right] \\
&\overset{\text{(Fubini–Tonelli)}}{=} e^{-\lambda t} \sum_{q=0}^{\infty} \frac{(\lambda^q \mathbb{E}[Y]^q)}{q!} \\
&\geq e^{-\lambda t} \sum_{q=0}^{\infty} \frac{(\lambda^q \cdot m \cdot t^q)}{q!} \\
&= m \cdot e^{-\lambda t} \sum_{q=0}^{\infty} \frac{(\lambda^q t^q)}{q!} \\
&= m \cdot e^{-\lambda t} \cdot e^{\lambda t} \\
&= m
\end{aligned}$$

□

Exercise 2.6

Let Z be a real-valued random variable. Show that the set of positive numbers $S = \{\lambda > 0 : \mathbb{E}\left[e^{\lambda Z}\right] < \infty\}$ is either empty or an interval with left end point equal to 0. Let $b = \sup S$. Show that $\varphi_Z(\lambda) = \log \mathbb{E}\left[e^{\lambda Z}\right]$ is convex and infinitely many times differentiable on $I = (0, b)$ with $\varphi'_Z(0) = \varphi_Z(0) = 0$ and the Cramér transform of Z equals $\varphi^*_Z = \sup_{\lambda \in I}(\lambda t - \varphi_Z(\lambda))$

Solution : We start with the first claim, we are going to prove that:

$$\text{For any } 0 < \lambda < \mu, \quad \mu \in S \implies \lambda \in S$$

This will prove that S is either empty or an interval with left end point equal to 0. Let $0 < \lambda < \mu$ with $\mu \in S$. We have :

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda Z}\right] &= \mathbb{E}\left[e^{\lambda Z}\mathbb{I}_{Z\leq 0} + e^{\lambda Z}\mathbb{I}_{Z>0}\right] \\
&= \mathbb{E}\left[1 \cdot \mathbb{I}_{Z\leq 0}\right] + \mathbb{E}\left[e^{\mu Z}\right] \\
&< \infty
\end{aligned}$$

Now for the differentiability, let $0 < \alpha < \beta < b$, $\lambda \in (\alpha, \beta)$ and $k \in \mathbb{N}$. We have :

$$\frac{d^k}{d\lambda} e^{\lambda Z} = Z^k e^{\lambda Z}$$

To prove the differentiability of φ on (α, β), it is enough to provide a domination of $\frac{d^k}{d\lambda}$ on (α, β). We have :

$$\begin{aligned}
|Z^k e^{\lambda Z}| &\leq |Z|^k e^{\lambda Z}\mathbb{I}_{Z\leq 0} + |Z|^k e^{\lambda Z}\mathbb{I}_{Z>0} \\
&\leq |Z|^k e^{\alpha Z}\mathbb{I}_{Z\leq 0} + (|Z|^k \cdot e^{-\frac{b-\beta}{2}Z})(e^{\frac{b+\beta}{2}Z})\mathbb{I}_{Z>0}
\end{aligned}$$

The left term is integrable because it is bounded and the right term is integrable as the product of a bounded function and an integrable function. Now since $\mathbb{E}\left[e^{\lambda Z} > 0\right]$ on S, we have proven that φ is infinitely differentiable on I.

To prove convexity, we use the previous result and compute the 2nd order derivative of φ. For any $\lambda \in I$:

$$\begin{aligned}
\varphi'(\lambda) &= \frac{\mathbb{E}\left[Ze^{\lambda Z}\right]}{\mathbb{E}\left[e^{\lambda Z}\right]} \\
\varphi'' &= \frac{\mathbb{E}\left[Z^2 e^{\lambda Z}\right] \cdot \mathbb{E}\left[e^{\lambda Z}\right] - \mathbb{E}\left[Ze^{\lambda Z}\right]^2}{\mathbb{E}\left[e^{\lambda Z}\right]^2} \overset{\text{(C.S)}}{\geq} 0
\end{aligned}$$

Where the second inequality comes from Cauchy-Schwarz Inequality. This proves the convexity of φ.

To finish the exercise, it remains to study the case in which the expectation of Z is 0. To prove that φ is continuously differentiable on $[0,b)$, we simply need to verify that its derivative has a right limit in 0. It is clear that $Ze^{\lambda Z} \underset{\lambda\to 0}{\to} Z$. Then :

$$\begin{aligned}\mathbb{E}\left[Ze^{\lambda Z}\right] &= \mathbb{E}\left[Ze^{\lambda Z}\mathbb{I}_{Z\leq 0}\right] + \mathbb{E}\left[Ze^{\lambda Z}\mathbb{I}_{Z>0}\right] \\ &\underset{\lambda\to 0}{\to} \mathbb{E}\left[Z\mathbb{I}_{Z\leq 0}\right] + \mathbb{E}\left[Z\mathbb{I}_{Z>0}\right] = \mathbb{E}\left[Z\right] = 0\end{aligned}$$

Where the left expectation converges because of dominated convergence and the right one because of monotonic convergence. We have proven continous differntiability of φ on $[0,b)$ and that $\varphi(0) = \varphi'(0) = 0$. ☐

Exercise 2.7

Prove that if Z is a centered normal random variable with variance σ^2, then

$$\sup_{t>0}\left(\mathbb{P}\left[Z\geq t\right]\exp\frac{t^2}{2\sigma^2}\right) = \frac{1}{2}$$

Solution : Without loss of generality, we assume that $\sigma = 1$. Let $t > 0$, we have

$$\begin{aligned}\mathbb{P}\left[Z\geq t\right] &= \int_t^{+\infty}\frac{1}{\sqrt{2\pi}}\exp -\frac{u^2}{2}\,du \\ &= \int_0^{+\infty}\frac{1}{\sqrt{2\pi}}\exp -\frac{(u+t)^2}{2}du \\ &= \exp -\frac{t^2}{2}\cdot\int_0^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-ut\right)\exp\left(-\frac{u^2}{2}\right)du \\ \mathbb{P}\left[Z\geq t\right]\cdot\exp\left(\frac{t^2}{2}\right) &= \int_0^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-ut\right)\exp\left(-\frac{u^2}{2}\right)du\end{aligned}$$

Then we notice that $\exp(-ut) \underset{t\to\infty}{\to} 1$ pointwise. More importantly it converges in an increasing manner and since everything inside the integral is nonnegative, we can use the monotone convergence theorem to conclude that :

$$
\begin{aligned}
\sup_{t>0} \mathbb{P}\left[Z \geq t\right] \cdot \exp\left(\frac{t^2}{2}\right) &= \sup_{t>0} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-ut\right) \exp\left(-\frac{u^2}{2}\right) du \\
&= \lim_{t\to\infty} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-ut\right) \exp\left(-\frac{u^2}{2}\right) du \\
&= \int_0^{+\infty} \lim_{t\to\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-ut\right) \exp\left(-\frac{u^2}{2}\right) du \\
&= \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \exp -\frac{u^2}{2} du = \frac{1}{2}
\end{aligned}
$$

□

Exercise 2.9 (Sub-Gaussian lower tail for nonnegative random variables)

Let X be a nonnegative random variable with finite second moment. Show that for any $\lambda > 0$, $\mathbb{E}\left[e^{-\lambda(X-\mathbb{E}[X])}\right] \leq e^{\lambda^2 \mathbb{E}\left[X^2\right]/2}$. In particular, if $X_1, \ldots, X_n$ are independent nonnegative random variables, then for any t¿0,

$$
\mathbb{P}\left[S \leq -t\right] \leq \exp\left(\frac{-t^2}{v}\right)
$$

where $S = \sum_{i=1}^n (X_i - \mathbb{E}\left[X_i\right])$ and $v = \sum_{i=1}^n \mathbb{E}\left[X_i^2\right]$

Solution : We have the following elementary inequalities :

$$
\begin{aligned}
\forall x \geq 0,\ \exp -x \leq 1 - x + \frac{x^2}{2} \\
\forall x \in \mathbb{R},\ 1 + x \leq \exp x
\end{aligned}
$$

Using this we get :

$$
\begin{aligned}
\mathbb{E}\left[e^{-\lambda(X-\mathbb{E}[X])}\right] &= e^{\lambda\mathbb{E}[X]} \cdot \mathbb{E}\left[e^{-\lambda X}\right] \\
&\leq e^{\lambda\mathbb{E}[X]} \cdot \mathbb{E}\left[1 - \lambda X + \frac{\lambda^2 X^2}{2}\right] \\
&\leq e^{\lambda\mathbb{E}[X]} \cdot \left(1 + \left(-\lambda\mathbb{E}\left[X\right] + \frac{\lambda^2}{2}\mathbb{E}\left[X^2\right]\right)\right) \\
&\leq e^{\lambda\mathbb{E}[X]} \cdot e^{-\lambda X + \frac{\lambda^2}{2}\mathbb{E}\left[X^2\right]} \\
&= e^{\lambda^2\mathbb{E}\left[X^2\right]/2}
\end{aligned}
$$

□

Exercise 2.16

Prove that if X is a sub-Gaussian random variable with variance factor v then $Var(X) \leq v$.

Solution : We define $\phi(\lambda) := \log \mathbb{E}\left[e^{\lambda X}\right]$ the log moment generating function of X. The subgaussian condition can be written as

$$\forall \lambda \in \mathbb{R},\, \phi(\lambda) \leq \frac{\lambda^2 v}{2}$$

We also have that $\phi(0) = \phi'(0)$ and $\phi''(0) = Var(X)$ we can now use a Taylor expansion around 0 to get the following result :

$$\begin{aligned}
0 &\leq \frac{\lambda^2 v}{2} - \phi(\lambda) \\
&= \frac{\lambda^2}{2} \cdot (v - Var(X)) + o(\lambda^2) \\
&\sim \frac{\lambda^2 v}{2} \cdot (v - Var(X))
\end{aligned}$$

This is only possible if $Var(X) \leq v$

$\square$

Chapter 3

Bounding the Variance

Exercise 3.2

Assume that the random variables $X_1, \ldots, X_n$ are independent and binary {-1,1}-valued with $\mathbb{P}[X_i = 1] = p_i$ and that $f : \{-1,1\}^n \to \mathbb{R}$ has the bounded differences property with constants $c_1, \ldots, c_n$. Show that if $Z = f(X_1, \ldots, X_n)$,

$$Var(Z) \leq \sum_{i=1}^{n} c_i^2 p_i (1 - p_i)$$

Solution : We use Efron Stein's Inequality. Let $X_1', \ldots, X_n'$ be a independent copy of $X_1, \ldots, X_n$ and $Z_i = f(X_1, \ldots, X_{i-1}, X_i', X_{i+1}, \ldots, X_n)$. We also define $Z_i^{(\pm)} = f(X_1, \ldots, X_{i-1}, \pm 1, X_{i+1}, \ldots, X_n)$. We have that :

$$\begin{aligned}
Var(Z) &\leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[(Z - Z_i)^2\right] \\
&\leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}^{(i)}[(Z - Z_i)^2]\right] \\
&\leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[p_i^2 \cdot (Z_i^{(+)} - Z_i^{(+)})^2 + 2p_i(1 - p_i) \cdot (Z_i^{(+)} - Z_i^{(-)})^2 + (1 - p_i)^2 \cdot (Z_i^{(-)} - Z_i^{(-)})^2\right] \\
&\leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[2p_i(1 - p_i)c_i^2\right] \\
&\leq \sum_{i=1}^{n} p_i(1 - p_i)c_i^2
\end{aligned}$$

$\boxtimes$

Bibliography