Homework 1 HDP

Bizzaro Davide, Schiavo Leonardo

October 2021

Exercise 1. We want to prove the inequalities

$$\left(\frac{n}{m}\right)^m \leq \binom{n}{m} \leq \sum_{k=0}^{m} \binom{n}{k} \leq \left(\frac{en}{m}\right)^m \tag{0.1}$$

for all integers $m \in [1, n]$.

The first inequality is equivalent to $\frac{n^m (n-m)!}{n!} \leq \frac{m^m}{m!}$, which expanded is

$$\frac{n \cdots n \cdot (n-m) \cdots 2 \cdot 1}{n \cdot (n-1) \cdots 2 \cdot 1} \leq \frac{m \cdots m}{m \cdot (m-1) \cdots 2 \cdot 1}$$
$$\frac{n}{n} \cdot \frac{n}{n-1} \cdots \frac{n}{n-m+1} \cdot \frac{n-m}{n-m} \cdot \frac{n-m-1}{n-m-1} \cdots \frac{2}{2} \cdot \frac{1}{1} \leq \frac{m}{m} \cdot \frac{m}{m-1} \cdots \frac{m}{2} \cdot \frac{m}{1}$$
$$\frac{n}{n-1} \cdots \frac{n}{n-m+1} \leq \frac{m}{m-1} \cdots \frac{m}{2} \cdot m$$

Now we can notice that every factor on the left-hand side is less or equal than the respective factor on the right-hand side of the inequality[1], hence we obtain the thesis.

The second inequality is quite obvious because

$$\binom{n}{m} \leq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m-1} + \binom{n}{m} = \sum_{k=0}^{m} \binom{n}{k} \tag{0.2}$$

For the third inequality we multiply both sides by $(m/n)^m$, obtaining the equivalent thesis

$$\left(\frac{m}{n}\right)^m \sum_{k=0}^{m} \binom{n}{k} \leq e^m \tag{0.3}$$

Now we have that

$$\sum_{k=0}^{m} \binom{n}{k} \left(\frac{m}{n}\right)^m \leq \sum_{k=0}^{m} \binom{n}{k} \left(\frac{m}{n}\right)^k = \left(1 + \frac{m}{n}\right)^m \leq \left(1 + \frac{m}{n}\right)^n \nearrow e^m$$

where we used the fact that $\frac{m}{n} \leq 1$, the binomial theorem and a known limit. Therefore,

$$\left(\frac{m}{n}\right)^m \sum_{k=0}^{m} \binom{n}{k} \leq e^m \tag{0.4}$$

which was equivalent to the last inequality.

Exercise 2. We still consider the same set of centers

$$\mathcal{N} := \left\{ \frac{1}{k} \sum_{j=1}^{k} x_j \mid x_j \text{ are vertices of } P \right\} \tag{0.5}$$

used in the proof of Corollary 0.0.4, but now we want to give a better estimate of its cardinality. Recall that the number of ways to choose k elements with repetition (and this time without considering the order of choice, which

[1] Because, for each $i \in \mathbb{R}$, the real function $x \mapsto \frac{x}{x-i}$ is decreasing for $x \in]i, +\infty[$ (and because $m \leq n$).

is indeed irrelevant in our case) from a set with N elements is $\binom{N+k-1}{k}$. This is therefore greater or equal than the cardinality of $\mathcal{N}$, and by the previous exercise we get that, assuming WLOG $k \geq 1$,

$$|\mathcal{N}| \leq \binom{N+k-1}{k} \leq \left(\frac{e(N+k-1)}{k}\right)^k = \left(e\frac{k-1}{k} + \frac{eN}{k}\right)^k \leq \left(e + \frac{eN}{k}\right)^k. \tag{0.6}$$

In our case $k \coloneqq \lceil 1/\epsilon^2 \rceil$, so $1/k \leq \epsilon^2$ and

$$|\mathcal{N}| \leq \left(e + \frac{eN}{k}\right)^k \leq [e + e\epsilon^2 N]^{\lceil 1/\epsilon^2 \rceil} \tag{0.7}$$

which is what we wanted (with $C \coloneqq e$).

Exercise 3. We want to show that

$$\mathbb{E}|X|^p = \int_0^{+\infty} pt^{p-1}\mathbb{P}\{|X| > t\}dt \tag{0.8}$$

(whenever the right-hand side is finite), where X is a r.v. and $p \in (0, +\infty)$.

By the integral identity we have that

$$\mathbb{E}|X|^p = \int_0^{+\infty} \mathbb{P}\{|X|^p > s\}ds \tag{0.9}$$

and with a change of variables $s = t^p$ we get

$$\mathbb{E}|X|^p = \int_0^{+\infty} \mathbb{P}\{|X|^p > s\}ds = \int_0^{+\infty} \mathbb{P}\{|X|^p > t^p\}pt^{p-1}dt. \tag{0.10}$$

whenever the right-hand side is finite. To conclude, we just notice that $\mathbb{P}\{|X|^p > t^p\} = \mathbb{P}\{|X| > t\}$, since the two events $\{|X|^p > t^p\}$ and $\{|X| > t\}$ are clearly the same.

Exercise 4. We want to prove the inequality

$$\cosh(x) \leq \exp\left(\frac{x^2}{2}\right)$$

for all $x \in \mathbb{R}$.

Recall that

$$\cosh(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \quad \text{and} \quad \exp\left(\frac{x^2}{2}\right) = \sum_{k=0}^{\infty} \frac{2^{-k}}{k!}x^{2k}$$

for all $x \in \mathbb{R}$. Our thesis has now become

$$\sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \leq \sum_{k=0}^{\infty} \frac{2^{-k}}{k!}x^{2k}.$$

If it holds that $1/(2k)! \leq 2^{-k}/k!$ for all $k \in \mathbb{N}$, then we obtain the thesis. The fact that it holds is trivial for $k = 0$, so let us suppose $k \geq 1$. In this case the inequality we want to show is equivalent to

$$\begin{aligned} \frac{k(k-1)\cdots 2\cdot 1}{(2k)(2k-1)\cdots(2k-k+1)k(k-1)\cdots 2\cdot 1} &\leq \frac{1}{2^k} \\ \frac{1}{(2k)(2k-1)\cdots(k+1)} &\leq \frac{1}{2^k} \end{aligned}$$

or in one line

$$\begin{aligned} 2^k &\leq (2k)(2k-1)\cdots(k+1) \\ 2\cdots 2 &\leq (2k)\cdots(k+1) \end{aligned}$$

which is obviously true.

Exercise 5. Let $X_1, \dots, X_N$ be independent Bernoulli random variables with parameters p_i, $i = 1, \dots, N$. Let $S_N \coloneqq \sum_{i=1}^N X_i$ and $\mu \coloneqq \mathbb{E} S_N$. We want to show that

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq 2e^{-C\mu\delta^2} \tag{0.11}$$

for any $\delta \in (0, 1]$, where $C > 0$ is a constant.

First, let us show that for any $t < \mu$ we have

$$\mathbb{P}\{S_N \leq t\} \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t. \tag{0.12}$$

We redo the steps of the proof of Chernoff's inequality, but now considering $\lambda < 0$:

$$\mathbb{P}\{S_N \leq t\} = \mathbb{P}\{\lambda S_N \geq \lambda t\} = \mathbb{P}\{\exp(\lambda S_N) \geq \exp(\lambda t)\} \leq \tag{0.13}$$

$$\leq e^{-\lambda t}\mathbb{E}\exp(\lambda S_N) = e^{-\lambda t}\prod_{i=1}^N \mathbb{E}\exp(\lambda X_i) \leq \tag{0.14}$$

$$\leq e^{-\lambda t}\exp[(e^\lambda - 1)\mu] \tag{0.15}$$

and 0.12 follows by substituting $\lambda = \log(t/\mu)$ (which is indeed ≤ 0 because we are assuming $t \leq \mu$) into 0.15.

Now,

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} = \mathbb{P}\{S_N \geq (1+\delta)\mu\} + \mathbb{P}\{S_N \leq (1-\delta)\mu\} \leq \tag{0.16}$$

$$\leq e^{-\mu}\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} + e^{-\mu}\left(\frac{e}{1-\delta}\right)^{(1-\delta)\mu} = \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}} + \frac{e^{-\delta\mu}}{(1-\delta)^{(1-\delta)\mu}} \tag{0.17}$$

by Chernoff's inequality and 0.12.

Hence, it is enough to prove that

$$\frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}} \leq e^{-C\mu\delta^2} \tag{0.18}$$

for all $\delta \in [-1, 0) \cup (0, 1]$ and a suitable constant $C > 0$.

We can rewrite it into the equivalent form[2]

$$e^{\delta + C\delta^2} \leq (1+\delta)^{(1+\delta)} \tag{0.19}$$

or better

$$\delta + C\delta^2 \leq (1+\delta)\log(1+\delta). \tag{0.20}$$

So, let us study when the function $f(x) = (1+x)\log(1+x) - x - Cx^2$ (extended by continuity in $x = -1$) is non-negative in the interval $[-1, 1]$. We notice that 0 is a root of $f(x)$ for every $C > 0$, and 1 is a root when $C = 2\log 2 - 1 = 0.386\dots$; let us show that this value of C is sufficient. We notice that $f'(x) = \log(1+x) - 2Cx$ and $f''(x) = \frac{1}{1+x} - 2C$, so that $f'(x) < 0$ for all $x \in (-1, 0)$, $f'(0) = 0$ and if

$$\bar{x} \coloneqq \frac{1}{2C} - 1 = \frac{1}{4\log 2 - 2} - 1 = 0.294\dots \in (0, 1),$$

then $f''(x) > 0$ (i.e. f is convex) for all $x \in (-1, \bar{x})$ and $f''(x) < 0$ (i.e. f is concave) for all $x \in (\bar{x}, 1)$. It is easy to see that this (together with the fact that $f(0) = f(1) = 0$) is enough to show that $f(x) \geq 0$ for all $x \in [-1, 1]$, and this concludes the proof.

Exercise 6. We want to show that for $p \geq 1$, the r.v. $X \sim N(0, 1)$ satisfies

$$||X||_{L^p}^p = \mathbb{E}|X|^p = 2^{p/2}\frac{\Gamma((1+p)/2)}{\Gamma(1/2)} \tag{0.21}$$

and that therefore $||X||_{L^p} \in O(\sqrt{p})$ as $p \to \infty$.

Recall that by definition $\Gamma(z) = \int_0^{+\infty} t^{z-1}e^{-t}dt$. Now,

$$||X||_{L^p}^p = \mathbb{E}|X|^p = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|^p e^{-x^2/2}dx = \frac{2}{\sqrt{2\pi}}\int_0^{+\infty}x^p e^{-x^2/2}dx. \tag{0.22}$$

[2] In case $\mu = 0$ the thesis is trivially verified, so WLOG $\mu \neq 0$.

We perform a change of varibles $t = x^2/2$ (meaning that $x = (2t)^{1/2}$ and $dx = (2t)^{-1/2}dt$):

$$||X||_{L^p}^p = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x^p e^{-x^2/2} dx = \tag{0.23}$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} (2t)^{p/2} e^{-t} (2t)^{-1/2} dt = \tag{0.24}$$

$$= \frac{2^{p/2}}{\sqrt{\pi}} \int_0^{+\infty} t^{(p-1)/2} e^{-t} dt = \tag{0.25}$$

$$= \frac{2^{p/2}}{\Gamma(1/2)} \Gamma((p+1)/2), \tag{0.26}$$

which is what we wanted. In the last passage we used the fact that (with the same change of varible as before)

$$\Gamma(1/2) = \int_0^{+\infty} t^{-1/2} e^{-t} dt = \sqrt{2} \int_0^{+\infty} e^{-x^2/2} dx = \sqrt{\pi}. \tag{0.27}$$

As for the limit with $p \to \infty$, what we need to prove is that there exists a constant $C > 0$ such that

$$\frac{||X||_{L^p}}{\sqrt{p}} = \frac{\sqrt{2}}{\sqrt{p}} \frac{\Gamma((1+p)/2)^{1/p}}{\Gamma(1/2)^{1/p}} \leq \frac{\sqrt{2}}{\sqrt{p}} \Gamma((1+p)/2)^{1/p} \leq C \tag{0.28}$$

for all p sufficiently large. Now, by the properties of the function Γ, we have that $\Gamma((1+p)/2) = \frac{p+1}{2} \frac{p-1}{2} \cdots \frac{3}{2} \Gamma(1/2)$ if p is even and $\Gamma((1+p)/2) = ((1+p)/2)!$ if p is odd. In both cases it is clear that $\Gamma((1+p)/2) \leq ((1+p)/2)^{(p+1)/2}$ for all p sufficiently large. Therefore, for such large p's,

$$\frac{\sqrt{2}}{\sqrt{p}} \Gamma((1+p)/2)^{1/p} \leq \frac{\sqrt{2}}{\sqrt{p}} ((1+p)/2)^{\frac{p+1}{2p}} \xrightarrow{p \to \infty} 1.$$

This is enough to conclude.

Exercise 7. We want to compute the sub-gaussian norm of a random variable $X \sim N(0, \sigma^2)$. By definition the sub-gaussian norm is

$$||X||_{\psi_2} := \inf \left\{ t > 0 : \mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] \leq 2 \right\}$$

We need to calculate the expected value explicitly:

$$\begin{aligned} \mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] &= \int_{\mathbb{R}} \exp\left(\frac{x^2}{t^2} \right) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left[-x^2 \left(\frac{1}{2\sigma^2} - \frac{1}{t^2} \right) \right] dx \end{aligned}$$

which is finite for $t > \sqrt{2}\sigma$. We can use the change of variables

$$-x^2 \left(\frac{t^2 - 2\sigma^2}{2t^2\sigma^2} \right) = -\frac{y^2}{2}$$

and obtain

$$\mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2} \right) \frac{t\sigma}{\sqrt{t^2 - 2\sigma^2}} dy = \frac{t}{\sqrt{t^2 - 2\sigma^2}}$$

Now, if we want

$$\frac{t}{\sqrt{t^2 - 2\sigma^2}} \leq 2$$

we need to take $t \geq \sqrt{\frac{8}{3}}\sigma$. Hence the sub-gaussian norm of X is $||X||_{\psi_2} = \sqrt{\frac{8}{3}}\sigma$.

Exercise 8. Similarly to the previous exercise we want to compute the sub-gaussian norm of $X \sim SymBer$. The Symmetric Bernoulli r.v. is characterized by $P(X = -1) = P(X = 1) = 1/2$. We compute the expected value as before:

$$\mathbb{E}\left[\frac{X^2}{t^2} \right] = \exp\left(\frac{1}{t^2} \right)$$

and $\exp(1/t^2) \leq 2$ when

$$t \geq \sqrt{\frac{1}{\ln(2)}}.$$

Taking the infimum we obtain the sub-gaussian norm $||X||_{\psi_2} = \sqrt{\frac{1}{\ln(2)}}$.

Exercise 9. We want to prove the inequality

$$||X||_{\psi_2} \leq \frac{||X||_\infty}{\sqrt{\ln(2)}}$$

Let's consider the following sets:

$$\begin{aligned} A =& \left\{ t > 0 : \mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] \leq 2 \right\} \\ B =& \left\{ t > 0 : \exp\left(\frac{||X||_\infty^2}{t^2} \right) \leq 2 \right\} \end{aligned}$$

We notice that $B \subseteq A$, hence $\inf A \leq \inf B$. We observe the following equivalence chain

$$\begin{aligned} \exp\left\{ \frac{||X||_\infty^2}{t^2} \right\} &\leq 2 \\ \frac{||X||_\infty^2}{t^2} &\leq \ln(2) \\ t &\geq \frac{||X||_\infty}{\sqrt{\ln(2)}} \end{aligned}$$

We deduce that $\inf B = ||X||_\infty / \sqrt{\ln(2)}$. In conclusion we have

$$||X||_{\psi_2} = \inf(A) \leq \inf(B) = \frac{||X||_\infty}{\sqrt{\ln(2)}}$$

and the thesis.

Exercise 10. We want to show that the subgaussian norm in indeed a norm. Just a reminder of the definition

$$||X||_{\psi_2} = \inf\left\{ t > 0 : \mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] \leq 2 \right\}$$

1. The positivity is obvious since the infimum is taken over the real positive numbers.

2. If $X = 0$ we have $\exp(0/t^2) = 1$ for all $t > 0$, hence $||X||_{\psi_2} = 0$. Viceversa, if $||X||_{\psi_2} = 0$ we want to show that $X = 0$ almost everywhere. By hypothesis we know that $\lim_{t \to 0} \mathbb{E}(\exp(X^2/t^2)) \leq 2$. Ab absurdum there exists a set $M \subset \Omega$ such that $\mathbb{P}(M) > 0$ in which $|X| \neq 0$. Since

$$M = \bigcup_{n=1}^{\infty} \left\{ |X| > \frac{1}{n} \right\}$$

and

$$\left\{ |X| > \frac{1}{n} \right\} \subseteq \left\{ |X| > \frac{1}{n+1} \right\}$$

we have that

$$\lim_{n \to \infty} \mathbb{P}\left(|X| > \frac{1}{n} \right) = \mathbb{P}(M)$$

Thus there must exist an $\bar{n}$ such that $\mathbb{P}(|X| > 1/\bar{n}) > \mathbb{P}(M) - \epsilon > 0$. Now letting $A = \{|X| > 1/\bar{n}\}$,

$$\lim_{t \to 0} \mathbb{E}\left[\exp\left(\frac{X^2}{t^2} \right) \right] \geq \lim_{t \to 0} \int_A \exp(X^2/t^2) d\mathbb{P} \geq \lim_{t \to 0} \exp \frac{(1/\bar{n})^2}{t^2} \mathbb{P}(A) = +\infty.$$

This is in contradiction with our hypothesis, so we conclude that $X = 0$ almost everywhere.

3. Homogeneity: for any $\lambda \neq 0$,

$$
\begin{aligned}
||\lambda X||_{\psi_2} &= \inf\left\{t > 0 : \mathbb{E}\left[\exp\left(\frac{\lambda^2 X^2}{t^2}\right)\right] \leq 2\right\} \\
&= |\lambda| \inf\left\{\frac{t}{|\lambda|} > 0 : \mathbb{E}\left[\exp\left(\frac{X^2}{(t/\lambda)^2}\right)\right] \leq 2\right\} \\
&= |\lambda| ||X||_{\psi_2}
\end{aligned}
$$

4. Last the triangular inequality $||X+Y||_{\psi_2} \leq ||X||_{\psi_2} + ||Y||_{\psi_2}$.

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\left(\frac{X+Y}{||X||_{\psi_2} + ||Y||_{\psi_2}}\right)^2\right)\right] &= \mathbb{E}\left[\exp\left(\left(\frac{||X||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}\frac{X}{||X||_{\psi_2}} + \frac{||Y||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}\frac{Y}{||Y||_{\psi_2}}\right)^2\right)\right] \\
&\leq \frac{||X||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}\mathbb{E}\left[\exp\left(\left(\frac{X}{||X||_{\psi_2}}\right)^2\right)\right] + \frac{||Y||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}\mathbb{E}\left[\exp\left(\left(\frac{Y}{||Y||_{\psi_2}}\right)^2\right)\right] \\
&\leq \frac{||X||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}2 + \frac{||Y||_{\psi_2}}{||X||_{\psi_2} + ||Y||_{\psi_2}}2 \leq 2
\end{aligned}
$$

where we used the convexity of the function $x \mapsto \exp(x^2)$ in the first inequality and the definition of sub-Gaussian norm in the second. Again by the definition of sub-Gaussian norm, we get that $||X||_{\psi_2} + ||Y||_{\psi_2} \geq ||X+Y||_{\psi_2}$.