Homework 5
HDP KNU+ Fall 2022

Hints are in the back of this homework set.

As in the previous homework sets, $C, C_1, C_2, \ldots$ and $c, c_1, c_2, \ldots$ denote positive absolute constants of your choice.

The subgaussian norm of a random variable X, introduced in Lecture 13, is like standard deviation on steroids. The standard deviation $\sigma(X) = \sqrt{\mathrm{Var}(X)}$ captures only the average spread of the distribution, but the subgaussian norm $\|X\|_{\psi_2}$ controls all moments, all tails, MGF, etc. (see Section 2.5.2 of the book). In the ideal world, the magnitude of the subgaussian norm would be comparable to the standard deviation. You will now check that (a) the subgaussian norm always dominates the standard deviation, (b) the converse holds in some natural cases, but (c) not always.

Problem 1 (Subgaussian norm as a variance proxy)

(a) Prove that the inequality

$$\sigma(X) \leq C_1 \|X - \mathbb{E}\, X\|_{\psi_2}$$

holds for any subgaussian random variable X.

(b) Prove that the reverse inequality

$$\|X - \mathbb{E}\, X\|_{\psi_2} \leq C_2\, \sigma(X), \tag{1}$$

holds for any normally distributed random variable $X \sim N(\mu, \sigma^2)$, and for the random variable X uniformly distributed on an interval $[a, b]$.

(c) Check that the random variable X that has Poisson distribution (with any parameter $\lambda > 0$) is not subgaussian. (Hence the bound (1) fails in this case, since the left hand side is infinite but the right hand side is finite.)

In Lecture 13 we proved the *integral identity*, which expresses the expectation of a nonnegative random variable X via the tails of X (see Lemma 1.2.1 in the book). You will now extend the integral identity to all random variables X, not necessarily nonnegative.

Problem 2 (General integral identity)

Prove that any random variable X satisfies

$$\mathbb{E}\, X = \int_0^\infty \mathbb{P}\left\{X > t\right\} dt - \int_{-\infty}^0 \mathbb{P}\left\{X < t\right\} dt.$$

In the Subgaussian Lemma we proved in Lecture 13, one of its equivalent parts requires a mysterious extra condition that the mean of X be zero. More precisely, this part says that if X is a subgaussian random variable and $\mathbb{E}\,X = 0$, then there exists $K > 0$ such that the MGF (moment generating function) of X is bounded as follows:

$$\mathbb{E}\exp(\lambda X) \le \exp(K^2\lambda^2) \quad \text{for all } \lambda \in \mathbb{R}. \tag{2}$$

(See Lecture 13 or Proposition 2.5.2 of the book.) You will now check that the mean zero assumption is essential, and in fact it can *never* be removed from this property.

Problem 3 (Subgaussian MGF bound implies mean zero)

Let X be a random variable that satisfies property (2). Prove that $\mathbb{E}\,X = 0$.

In data science, we often need to handle large number of random variables at the same time. Suppose we have N random points sampled from the standard normal distribution. Then, on average, the entire sample will lie within $O(\sqrt{\log n})$ from the origin. This is quite a good bound, since the logarithm grows slowly! You will now prove this bound in a general framework, for all subgaussian distributions:

Problem 4 (Maximum of subgaussians)

Let $X_1, \ldots, X_n$ be sub-gaussian random variables (not necessarily independent), and assume that $\|X_i\|_{\psi_2} \le K$ for all i. Show that

$$\mathbb{E}\left[\max_{i=1,\ldots,n} |X_i|\right] \le CK\sqrt{\log n}.$$

TURN OVER FOR HINTS

Hints

Hint for Problem 1. Part (a) should follow straight from one of the equivalent properties of a subgaussian distribution (see Lecture 13 or Section 2.5.2 of the book). Part (b) is also simple; just use one of the equivalent subgaussian properties (whichever is more convenient). For part (c), either use known expression for MGF of Poisson, or show that the tails are too heavy by first noting that $\mathbb{P}\{X \geq k\} \geq \mathbb{P}\{X = k\}$ and then using Strirling's approximation.

Hint for Problem 3. Bound $\exp(\lambda \mathbb{E} X)$ by $\mathbb{E} \exp(\lambda X)$ (how?), then use the assumption for $\lambda \to 0^+$ and for $\lambda \to 0^-$.

Hint for Problem 4. Bound $\mathbb{E} \max_i |X_i|$ by $\mathbb{E} \left(\sum_{i=1}^n |X_i|^p \right)^{1/p}$; move the expected value inside the sum (how?); use the subgaussian bound on the moments (Lecture 13 or Section 2.5.2); and finally optimize in p (or just guess what value of p would be good).