Homework One

EMALCA 2025: High-dimensional Statistics

Due in class on July 2 (Submit in groups of 2 or 3)

1. Consider $X \sim \text{Exp}(1)$, i.e., an exponential random variable with rate 1.

   a) (**2.5/3 points**) Verify that X is not sub-Gaussian by showing that

   $$\frac{\|X\|_p}{\sqrt{p}} \to \infty, \text{ as } p \to \infty.$$

   Here, $\|X\|_p = (\mathbb{E}(|X|^p))^{1/p}$.

   b) (**2.5/3 points**) Show that X is sub-Exponential.

   c) (**2.5/3 points**) Show that $\sqrt{X}$ is sub-Gaussian.

2. (**2.5 points**) Show that exponential, Pareto and Cauchy distributions are not sub-gaussian.

3. (**2.5 points**) Show that $\|\cdot\|_{\psi_2}$ satisfies the triangle inequality. Hint: Use the convexity of the function $f(x) = e^{x^2}$. You can also assume that the infimum is attained in the definition.

4. (**2.5 points**) Let $X_1, \ldots, X_n$ be mean zero, independent sub-Exponential random variables with parameters $\nu, \alpha > 0$. For any vector $a = (a_1, \ldots, a_n)^\top \in \mathbb{R}^n$ define the weighted sum

   $$S(a) = \sum_{i=1}^{n} a_i X_i.$$

   Show that for any $t > 0$, we have that

   $$\mathbb{P}\left(|S(a)| > t\right) \leq 2\exp\left(-C\min\left\{\frac{t^2}{\nu^2\|a\|_2^2}, \frac{t}{\nu\|a\|_\infty}\right\}\right)$$

   for some positive constant C.

Additional information

Pareto distribution

Notación [editar]

Si X es una variable aleatoria continua con distribución Pareto con parámetros $\alpha > 0$ y $x_m > 0$ entonces escribimos $X \sim \text{Pareto}(\alpha, x_m)$.

Función de densidad [editar]

La función de densidad de una variable aleatoria $X \sim \text{Pareto}(\alpha, x_m)$ es

$$f_X(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$$

para $x_m \leq x$.

Momentos [editar]

El n-ésimo momento sólo está definido para $n < \alpha$ y en tal caso es

$$\mathrm{E}[X^n] = \frac{\alpha x_\mathrm{m}^n}{\alpha - n}$$

Cauchy distribution

Parámetros	x_0 (real) $\gamma > 0$ escala (real)
Función de densidad (pdf)	$\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$
Función de distribución (cdf)	$\frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right)+\frac{1}{2}$
Media	no definida
Mediana	x_0
Moda	x_0
Varianza	no definida

Exponential distribution

Definición [editar]

Función de Densidad [editar]

Se dice que una variable aleatoria continua X tiene una **distribución exponencial** con parámetro $\lambda > 0$ y escribimos $X \sim \text{Exp}(\lambda)$ si su función de densidad es

$$f_X(x) = \lambda e^{-\lambda x}$$

para $x \geq 0$.

El n-ésimo momento de la variable aleatoria X es

$$\mathrm{E}[X^n] = \frac{n!}{\lambda^n}$$