# Probabily Theory
# 1. Introduction

Peter Pfaffelhuber

April 17, 2024

# Basics

- $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space, i.e. $\Omega$ a set, $\mathcal{F}$ a $\sigma$-algebra, $\mathbf{P} : \mathcal{F} \to [0,1]$ a measure with $\mathbf{P}(\Omega) = 1$.
- Let $(E, r)$ be a metric space and $X : \Omega \to E$ measurable, i.e. for $\mathcal{F}' = \mathcal{B}(E)$ (Borel $\sigma$-algebra), we have $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{F}'$. Is $E = \overline{\mathbb{R}}$, then $X$ is called *real-valued*.
- $X_*\mathbf{P}(B) := \mathbf{P}(X \in B) = \mathbf{P}(X^{-1}(B))$ is called *distribution of* $X$.
- If $X_*\mathbf{P} = Y_*\mathbf{P}$, then $X, Y$ are called identically distributed and we write $X \stackrel{d}{=} Y$ or $X \sim Y$.

# Discrete distributions

- $\mu = \sum_{i \in \mathbb{N}_0} \delta_i$ is the counting measure on $\mathbb{N}_0$ and $f : \mathbb{N}_0 \to \mathbb{R}_+$. We denote by $f \cdot \mu$ the measure with

$$f \cdot \mu(A) := \sum_{i \in A}^{\infty} f(i) = \int_A f(x) \mu(dx).$$

- $X \sim B(n, p)$ means $X_* \mathbf{P} = f \cdot \mu$ with

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

- $X \sim \text{Poi}(\lambda)$ means $X_* \mathbf{P} = f \cdot \mu$ with

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!.}.$$

- $X \sim \text{geo}(p)$ means $X_* \mathbf{P} = f \cdot \mu$ with

$$f(x) = (1-p)^x p.$$

# Continuous distributions

- Let $\lambda$ be Lebesgue measure on $\mathcal{B}(\mathbb{R})$ and $f : \mathbb{R} \to \mathbb{R}_+$ measurable. We denote by $f \cdot \lambda$ the measure with
$$f \cdot \lambda(A) := \int_A f(x)\lambda(dx).$$
- $X \sim \exp(\lambda)$ means $X_*\mathbf{P} = f_\lambda \cdot \lambda$ with
$$f_\lambda(x) = \lambda e^{-\lambda x} 1_{x \geq 0}.$$
- $X \sim N(\mu, \sigma^2)$ means $X_*\mathbf{P} = f_{\mu,\sigma^2} \cdot \lambda$ with
$$f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big( - \frac{(x-\mu)^2}{2\sigma^2}\Big).$$

# expected values

- $X$ real-valued random variable (RV). Then
$$\mathbb{E}[X] := \int x X_* \mathbf{P}(dx)$$
is the *expected value* of $X$. It exists provided $\mathbb{E}[|X|] < \infty$.

- If $X_*\mathbf{P} = f \cdot \lambda$ and $g$ is measurable, then, if it exists,
$$\mathbf{E}[g(X)] = \int f(x)g(x)\lambda(dx)$$

- We set $\mathcal{L}^1 := \mathcal{L}^1(\mathbf{P}) := \{X : \mathbf{E}[|X|] < \infty\}$. For $X, Y, \in \mathcal{L}^1$,
$$X \leq Y \text{ almost certainly } \implies \mathbf{E}[X] \leq \mathbf{E}[Y],$$
$$\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y].$$

- $\mathbf{E}[X] = \mathbf{E}[X^+] - \mathbf{E}[X^-]$, if at least one term is finite.
$$\mathbf{E}[X] < \infty \implies \mathbf{P}(X < \infty) = 1.$$

# Measurability with respect to $\sigma(X)$

- The $\sigma$-algebra $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{F}'\}$ is *$\sigma$-algebra generated by* $X$.
- Lemma 6.2: Let $X, Z$ be RVs. Then, $Z$ is $\sigma(X)$-measurable iff there exists $\varphi$ measurable with $\varphi \circ X = Z$.

  $\Leftarrow$: clear

  $\Rightarrow$ for $Z = 1_A$: Here, $A = X^{-1}(A')$ for some suitable $A'$.

  Thus $Z = 1_A = 1_{X^{-1}(A')} = 1_{A'} \circ X$.

# Convergence results

- $X, X_1, X_2, \ldots$ real-valued RVs. Then,

$$A := \{X_n \xrightarrow{n\to\infty} X\} := \{\omega : X_n(\omega) \xrightarrow{n\to\infty} X(\omega)\} \in \mathcal{F}$$

  If $\mathbf{P}(A) = 1$, we say $X_n \xrightarrow{n\to\infty} X$ almost surely.

- Proposition 6.3:
  1. *Lemma of Fatou:* $\liminf_{n\to\infty} \mathbf{E}[X_n] \geq \mathbf{E}[\liminf_{n\to\infty} X_n]$.
  2. *Theorem of monotonic convergence:* If $X_1, X_2, \cdots \in \mathcal{L}^1$ and $X_n \uparrow X$ almost surely, then
  $$\mathbf{E}[X_n] \xrightarrow{n\to\infty} \mathbf{E}[X].$$
  3. *Theorem of dominated convergence:* If $X_n \xrightarrow{n\to\infty} X$ is almost surely and $|X_1|, |X_2|, \cdots \leq Y$ almost surely with $\mathbf{E}[Y] < \infty$. Then,
  $$\mathbf{E}[X_n] \xrightarrow{n\to\infty} \mathbf{E}[X].$$

# Markov and Chebyshev inequality

- Proposition 6.4: Let $X \geq 0$ and $x \geq 0$. Then the Markov inequality

$$\mathbf{P}(X \geq x) \leq \frac{\mathbf{E}[X]}{x}$$

holds. If $X$ is a real-valued RV and $p, x \geq 0$. Then the Chebyshev inequality holds, i.e.

$$\mathbf{P}(|X| \geq x) \leq \frac{\mathbf{E}[|X|^p]}{x^p}.$$

Proof: Since $x \cdot 1_{X \geq x} \leq X$, we can write

$$x \cdot \mathbf{P}(X \geq x) = \mathbf{E}[x \cdot 1_{X \geq x}] \leq \mathbf{E}[X].$$

# Minkowski and Hölder inequality

- Proposition 6.5 $X, Y$ RVs with values in $\mathbb{R}$.
  1. If $0 < p, q, r \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Then,

     $$\mathbf{E}[|XY|^r]^{1/r} \leq \mathbf{E}[|X|^p]^{1/p} \cdot \mathbf{E}[|Y|^q]^{1/q} \qquad \text{(Hölder inequality)}$$

     Specifically for $p = q = 2$,

     $$\mathbf{E}[|XY|] \leq \mathbf{E}[|X|^2]^{1/2} \cdot \mathbf{E}[|Y|^2]^{1/2}. \qquad \text{(Cauchy-Schwarz inequality)}$$

  2. The Minkowski inequality is

     $$\mathbf{E}[|X+Y|^p]^{1/p} \leq \mathbf{E}[|X|^p]^{1/p} + \mathbf{E}[|Y|^p]^{1/p}, \qquad 1 \leq p \leq \infty$$

     $$\mathbf{E}[|X+Y|^p] \leq \mathbf{E}[|X|^p] + \mathbf{E}[|Y|^p], \qquad 0 < p < 1$$

# Jensen's inequality

- Proposition 6.6: Let $X \in \mathcal{L}^1$ and $\varphi$ be convex. Then,

$$\mathbf{E}[\varphi(X)] \geq \varphi(\mathbf{E}[X]).$$

- Lemma 6.7: Let $q > 0$ and $X \in \mathcal{L}^q$ real-valued random variable. Then, for $p \leq q$

$$\mathbf{E}[|X|^q] = \mathbf{E}[(|X|^p)^{q/p}] \geq \mathbf{E}[|X|^p]^{q/p}.$$

In particular, $\mathcal{L}^q \subseteq \mathcal{L}^p$.