Lecture 3
Random Variables

Nils Detering

January 15th, Online lecture

Random variable motivation

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ probability space.
- Instead of the actual events $\omega \in \Omega$, interested in related quantity.
    - Say you want to play monopoly. Use $\Omega = \{\omega : \omega = (\omega_1, \omega_2), \omega_i \in \{1,2,3,4,5,6\}, i \in \{1,2\}\}$.
    - Define probability measure by $\mathbb{P}(\{(\omega_1, \omega_2) = (k, j)\}) = \frac{1}{6^2}$ for $1 \leq k, j \leq 6$.
    - But actually you are interested in the sum $\omega_1 + \omega_2$.
    - Define function $X : \Omega \to \mathbb{N} \subset \mathbb{R}$ by

$$\omega \mapsto X(\omega) = \omega_1 + \omega_2.$$

Random variable informal

Definition

A function $X : \Omega \to \mathbb{R}$ is called a random variable.

Problem: We would like to say something about

$$\mathbb{P}(X \leq \alpha) = \mathbb{P}(\{\omega : X(\omega) \leq \alpha\}) = \mathbb{P}(X^{-1}((-\infty, \alpha])).$$

But: We know $\mathbb{P}(X^{-1}((-\infty, \alpha]))$ only if $X^{-1}((-\infty, \alpha]) \in \mathcal{F}$.
This depends on $\mathcal{F}$ and the function X.

Random variable more precise

We call a function $X : \Omega \to \mathbb{R}$ measurable if for all $\alpha \in \mathbb{R}$

$$\{\omega : X(\omega) \leq \alpha\} = X^{-1}((-\infty, \alpha]) \in \mathcal{F}.$$

Definition

A measurable function $X : \Omega \to \mathbb{R}$ is called a Random Variable (RV).

Obs: Every random variable induces a probability measure $\mathbb{P}_X$ on $\mathbb{R}$ by

$$\mathbb{P}_X(R) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in R\}) = \mathbb{P}(X^{-1}(R)).$$

for $R \in \mathbb{R}$.
We are often sloppy and write things like $\mathbb{P}(X \leq a)$ when we actually mean $\mathbb{P}(\{\omega : X(\omega) \leq a\})$!
By the definition of $\mathbb{P}_X$ it holds $\mathbb{P}(X \leq a) = \mathbb{P}_X((-\infty, a])$.

Two types of RVs: **Discrete** and **Continuous**, plus others which are rare.

- **Discrete**: It attains only finitely or countably many values, i.e. $X(\Omega) = \{x_1, x_2, \dots\}$, often $X(\Omega) = \mathbb{N}$ or $\mathbb{Z}$.
  - Examples: Bernoulli, Binomial, Geometric, Poisson,...
  - Define probability mass function (p.m.f.) of X, $p_X : X(\Omega) \to [0, 1]$ by $x \mapsto p_X(x) = \mathbb{P}(X = x)$
  - Obs. that $X(\Omega) \subset \mathbb{R}$ and p_X can be extended to $\mathbb{R}$ by setting $p_X(x) = 0$ for $x \in X(\Omega)^c$
  - Often combinatorial methods required for analysis.

- **Continuous**: It attains uncountably many values (for example $X(\Omega) = \mathbb{R}$ or $(0, \infty)$))
  - Examples: Normal, Exponential, Uniform, ...
  - Has a density function $f_X : \mathbb{R} \to \mathbb{R}_+$, s.t.

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx$$

  - Obs. $\mathbb{P}(a \leq X \leq a + \delta) \approx \delta f_X(a)$
  - Always $\mathbb{P}(X = x) = 0$ for any value $x \in \mathbb{R}$.
  - Use tools of calculus to analyze them.

Instead of $\mathbb{P}_X, f_X, p_X$ we often write $\mathbb{P}, f, p$ when it becomes clear from the context.

Summary of Properties

	discrete	continuous
Distribution $\mathbb{P}_X(R),\ R \subset \mathbb{R}$	$\sum_{x_i \in R} p_X(x_i)$	$\int_R f_X(x)dx$
Expectation $\mathbb{E}[g(X)]$, $g : X(\Omega) \to \mathbb{R}$	$\sum_{x_i \in X(\Omega)} g(x_i)p_X(x_i)$	$\int_{-\infty}^{\infty} g(x)f_X(x)dx$
Cumulative Distrib. $F_X(x) = \mathbb{P}(X \leq x)$	$\sum_{x_i \leq x} p_X(x_i)$	$\int_{-\infty}^{x} f_X(y)dy$
Inv. Cum. Distrib. $F_X^{-1}(u)$	$\inf\{x \in \mathbb{R} : \sum_{x_i \leq x} p_X(x_i) \geq u\}$	$\inf\{x \in \mathbb{R} : \int_{-\infty}^{x} f_X(y)dy \geq u\}$

Example

...

Inverse Cumulative Distribution $F_X^{-1} : (0,1) \to \mathbb{R}$ defined by

$$F_X^{-1}(u) = \inf\{x \in \mathbb{R} : F_X(x) \geq u\}$$

Inverse Cumulative Distribution

We will use F_X^{-1} to simulate a Random Variable with distribution F_X. For this let $U \sim \text{Unif}((0,1))$ (easy to simulate), then $F_X^{-1}(U) \stackrel{d}{=} X$. Why?....

Example

Joint probability/density

Say we have two random variables X, Y defined, both on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Both X and Y discrete:

- Define joint probability mass function by $p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$. Then
  - $\sum_{x,y} p_{X,Y}(x,y) = 1$
  - $p_X(x) = \sum_y p_{X,Y}(x,y)$
- and the conditional pmf of Y given $X = x$ by $p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$
- independent if $p_{X,Y}(x,y) = p_X(x)p_Y(y)$

Relation to conditional probability?....

Joint probability/density

Both X and Y continuous:

- There might exist a joint density function $f_{X,Y} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ s.t.
  - $\mathbb{P}((X, Y) \in R) = \iint_R f_{X,Y}(x, y) dxdy$ for $R \subset \mathbb{R} \times \mathbb{R}$
  - $\iint_{\mathbb{R}\times\mathbb{R}} f_{X,Y}(x, y) dxdy = 1$
  - $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
- and the conditional density of Y given $X = x$ by $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$
- independent if $f_{X,Y}(x, y) = f_X(x) f_Y(y)$

Example

...