Mathematical foundations of Econometrics

G.Gioldasis, UniFe & prof. A.Musolesi, UniFe

April 7, 2016

Random variables

example: Lets consider the game of coin tossing. The sample space is $\Omega = \{H, T\}$, the involved *σ-algebra* is $\mathscr{F} = \{\emptyset, \Omega, \{H\}, \{T\}\}$ and the corresponding probability measure is $P(\{H\}) = P(\{T\}) = 1/2$ for a fair coin. Then, define a function X such that $X(\omega) = 1$ if $\omega = H$ and $X(\omega) = 0$ if $\omega = T$. X is a random variable with probabilities:

$$\begin{aligned} P(\{\omega \in \Omega : X(\omega) = 1\}) &\overset{s.n.}{=} P(X = 1) = P(\{H\}) = 1/2 \text{ and} \\ P(\{\omega \in \Omega : X(\omega) = 0\}) &\overset{s.n.}{=} P(X = 0) = P(\{T\}) = 1/2 \end{aligned} \tag{1}$$

For an arbitrary Borel set B we have:

$$P(\{\omega \in \Omega : X(\omega) \in B\}) \begin{cases} = & P(\{H\}) & = 1/2 \text{ if } & 1 \in B \text{ and } 0 \notin B \\ = & P(\{T\}) & = 1/2 \text{ if } & 1 \notin B \text{ and } 0 \in B \\ = & P(\{H, T\}) & = 1 \text{ if } & 1 \in B \text{ and } 0 \in B \\ = & P(\{\emptyset\}) & = 0 \text{ if } & 1 \notin B \text{ and } 0 \notin B \end{cases}$$

We need to confine the mapping $X : \Omega \to \mathbb{R}$ to those ω's for which we can make probability statements, i.e. $\{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F}$, where B is an arbitrary Borel set.

Random variables

Definition

Let $\{\Omega, \mathscr{F}, P\}$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is called a random variable defined on $\{\Omega, \mathscr{F}, P\}$ if for every Borel set B, $\{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F}$. So X is a random variable if it is $\mathscr{F}$ **measurable**.

The **space** or **range** of X is the set of real numbers: $\mathbb{D} = \{x : x = X(\omega), \omega \in \Omega\}$.

The set $\{\omega \in \Omega : X(\omega) \in B\}$ is called **inverse image** of B and is denoted by $X^{-1}(B)$, i.e.

$$X^{-1}(B) \stackrel{def}{=} \{\omega \in \Omega : X(\omega) \in B\}$$

Theorem

A function $X : \Omega \to \mathbb{R}$ is measurable $\mathscr{F}$ if and only if for all $x \in \mathbb{R}$ the sets $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F}$

Therefore, to verify that a real function $X : \Omega \to \mathbb{R}$ is $\mathscr{F}$ measurable, it is not necessary to verify that for all Borel sets B, but only for Borel sets of the type $(-\infty, x]$

Exercise

Consider the health insurance example. Recall that $\Omega = \{YH, YS, OH, OS\}$. The insurance premium X has to take into account only public information, i.e. only the age of the client. Consequently, X has to be measurable with respect to

$$\mathscr{F} = \{\emptyset, \Omega, \{YH, YS\}, \{OH, OS\}\}$$

So, for every $B \in \mathscr{B}$, $\{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F}$.

Define a random variable X of an experiment. For all Borel sets B, find the corresponding inverse images $X^{-1}(B)$.

Solution

Answer

Suppose that the random variable $X : \Omega \to \mathbb{R}$ equals $X = 50$ for a young client and $X = 100$ for an old one. According to this X, there are only four types of borel sets B. Therefore:

$$X^{-1}(B) = \begin{cases} \emptyset & \text{if} \quad 50 \notin B \text{ and } 100 \notin B \\ \{OH, OS\} & \text{if} \quad 50 \notin B \text{ and } 100 \in B \\ \{YH, YS\} & \text{if} \quad 50 \in B \text{ and } 100 \notin B \\ \Omega & \text{if} \quad 50 \in B \text{ and } 100 \in B \end{cases}$$

σ-algebra $\mathscr{F}_X$

Definition

Let X be a random variable. The *σ-algebra* $\sigma(X) = \mathscr{F}_X = \{X^{-1}(B), \forall B \in \mathscr{B}\}$ is called the *σ-algebra generated by* X.

Example 1: In the previous exercise, the σ-algebra generated by the random variable X is $\mathscr{F}_X = \{\emptyset, \Omega, \{OH, OS\}, \{YH, YS\}\} = \mathscr{F}$.

Example 2: Roll a dice and let $X = 1$ if the outcome is even and $X = 0$ if the outcome is odd. Then

$$\mathscr{F}_X = \{\{1,2,3,4,5,6\}, \emptyset, \{2,4,6\}, \{1,3,5\}\} \subset \mathscr{F}$$

whereas $\mathscr{F}$ consists of all subsets of $\Omega = \{1,2,3,4,5,6\}$.

Note: Random variables "carry" information through $\sigma(X) = \mathscr{F}_X$. Therefore, they create a new measurable space $\{\Omega, \mathscr{F}_X\}$.

Example: Grades in an American high school

Suppose the performance of a student in an American high school exam. Let us assume then that the sample space of this "game" is $\Omega = [0, 100)$ (100 corresponds to excellence A'). A *σ-algebra* $\mathscr{F}$ could be the Borel subsets in $[0, 100)$, i.e.

$$\mathscr{F} = \mathscr{B} \cap [0, 100) \overset{def.}{=} \{B \cap [0, 100) : B \in \mathscr{B}\}$$

The grade of the student in the exam is a random variable $X : \Omega \to \mathbb{R}$ such that

$$X(\omega) = \begin{cases} 0 \text{ if} & \omega \in [0, 20) \ \ (grade\ F) \\ 1 \text{ if} & \omega \in [20, 40) \ \ (grade\ D) \\ 2 \text{ if} & \omega \in [40, 60) \ \ (grade\ C) \\ 3 \text{ if} & \omega \in [60, 80) \ \ (grade\ B) \\ 4 \text{ if} & \omega \in [80, 100) \ \ (grade\ A) \end{cases}$$

Grades A, B and C is a PASS while D and F means failure. The failure or pass of the examination can be described by another random variable $Y : \Omega \to \mathbb{R}$ such that:

$$Y(\omega) = \begin{cases} -1 \text{ if} & \omega \in [0, 40) \ \ (FAIL) \\ +1 \text{ if} & \omega \in [40, 100) \ \ (PASS) \end{cases}$$

Example: Grades in an American high school

Therefore, if one knows the value of X (grade of the examination) , he also knows the value of Y (pass or failure). Therefore, X conveys more information than Y. In other words, the *σ-algebra* $\sigma(X)$ generated by X contains $\sigma(Y)$. Indeed, $\sigma(X) \supset \sigma(Y)$ because:

$$\begin{aligned}\sigma(X) =& \{X^{-1}(A) : A \in \mathscr{B}\} = \\ & \{\Omega, \emptyset, [0,20), [20,40), [40,60), [60,80), [80,100), [0,40), [0,20) \cup [40,60), \\ & [0,20) \cup [60,80), [0,20) \cup [80,100), [20,60), [20,40) \cup [60,80), \\ & [20,40) \cup [80,100), [40,80), [40,60) \cup [80,100), [60,100), [0,60), \\ & [0,40) \cup [60,80), [0,40) \cup [80,100), [20,80), [20,60) \cup [80,100), [40,100), \\ & [0,20) \cup [60,100), [20,40) \cup [60,100), [0,80), [0,60) \cup [80,100), [20,100), \\ & [0,20) \cup [40,100), [0,40) \cup [60,100), [0,20) \cup [40,60) \cup [80,100), \\ & [0,20) \cup [40,80)\} \text{ (notice there are } 2^5 \text{ sets)}\end{aligned}$$

and

$$\sigma(Y) = \{Y^{-1}(A) : A \in \mathscr{B}\} = \{\Omega, \emptyset, [0,40), [40,100)\}$$

probability measure μ_X

Given a random variable X, we define for every Borel set $B \in \mathscr{B}$:

$$\mu_X(B) \overset{def.}{=} P(X^{-1}(B)) = P(\{\omega \in \Omega : X(\omega) \in B\}) \overset{s.n.}{=} P(X \in B)$$

It is proven that μ_X is a probability measure on $\{\mathbb{R}, \mathscr{B}\}$.

Definition

The probability measure μ_X defined above is called the *probability measure induced by* X.

Note: The random variable X maps the probability space $\{\Omega, \mathscr{F}, P\}$ into a new probability space $\{\mathbb{R}, \mathscr{B}, \mu_X\}$. This approach is more accessible to the researcher because it describes events in the $\mathbb{R}$ space rather than in an Ω space. The random variable is mapped back into the probability space $\{\Omega, \mathscr{F}_X, P\}$.

Cumulative distribution functions

Definition

Let X be a random variable with induced probability measure μ_X. The function $F(x) = \mu_X((-\infty, x]) \stackrel{def.}{=} P(X \leq x)$, $x \in \mathbb{R}$ is called the cumulative distribution function of X.

Theorem

A cumulative distribution function of a random variable is always ***right continuous****, i.e.* $\forall x \in \mathbb{R}$, $\lim_{\delta \downarrow 0} F(x + \delta) = F(x)$ *and* ***monotonic non decreasing****, i.e.* $F(x_1) \leq F(x_2)$ *if* $x_1 < x_2$ *with* $\lim_{x \downarrow -\infty} F(x) = 0$, $\lim_{x \uparrow +\infty} F(x) = 1$.

Theorem

The set of discontinuity points of a distribution function of a random variable is ***countable****.*

Discrete, continuous and mixed random variables

Definition

A random variable X is **discrete** if its corresponding cumulative distribution F_X is a step function. X is **continuous** if F_X is absolutely continuous for all x in $\mathbb{R}$. X is **mixed** if it is neither discrete nor continuous.

Definition (Alternative)

A random variable X is **discrete** if the range $\mathbb{D}$ of X is finite or countably infinite.

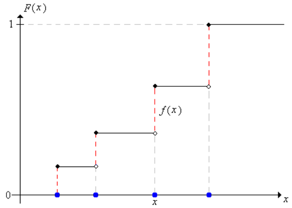
Therefore, if the range of a random variable X is uncountably infinite, then X is not discrete.

Examples:

- The number of everyday crashes in a city is a discrete random variable.
- An estimate of the surface of Mars is a continuous random variable.
- The point where a telescope focuses or the point where a missile hits the ground are continuous random variables.

Question: What type is the daily rainfall in a district, measured in mm?

Discrete, continuous and mixed random variables



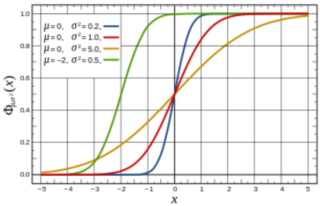


Figure 1: Cumulative distribution of discrete and continuous random variable

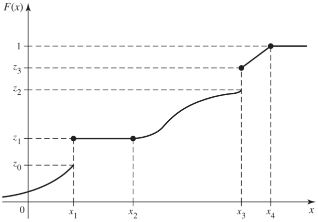


Figure 2: Cumulative distribution of mixed random variable

Independence of random variables

Definition

Let X_j be a sequence of random variables or vectors defined on a common probability space $\{\Omega, \mathscr{F}, P\}$. X_1 and X_2 are pairwise independent if for all Borel sets B_1, B_2, the sets $A_1 = \{\omega \in \Omega : X_1(\omega) \in B_1\}$ and $A_2 = \{\omega \in \Omega : X_2(\omega) \in B_2\}$ are independent. The sequence X_j is independent if for all Borel sets B_j the sets $A_j = \{\omega \in \Omega : X_j(\omega) \in B_j\}$ are independent.

Theorem

Let $X_1, \ldots, X_n$ *be random variables and denote, for* $x \in \mathbb{R}$ *and* $j = 1, \ldots n,$, $A_j(x) = X^{-1}((-\infty, x]) = \{\omega \in \Omega : X_j(\omega) \leq x\}$. *Then* $X_1, \ldots, X_n$ *are independent iff for arbitrary* $(x_1, x_2, \ldots, x_n)$ *the sets* $A_1(x_1), \ldots, A_n(x_n)$ *are independent.*

Theorem

The random variables $X_1, \ldots, X_n$ *are independent if and only if the joint distribution function* $F(x)$ *of* $X = (X_1, \ldots, X_n)^T$ *can be written as the product of the distribution functions* $F_j(x_j)$ *of the* X_j*'s., i.e.* $F(x) = \prod_{j=1}^{n} F_j(x_j)$, *where* $x = (x_1, \ldots, x_n)^T$.