11/12/2024

Lecture 20

Probabilities II / Τρεβέζας

Identically Distributed random variables (Ισόνομες τμ)

Definition: Let $X : \Omega_1 \to E$ and $Y : \Omega_2 \to E$ two random variables defined on $(\Omega_1, \mathcal{A}_1, P_1)$ and $(\Omega_2, \mathcal{A}_2, P_2)$ respectively. X and Y are called identically distributed (ισόνομες) if $P_X = P_Y$ on $(E, \mathcal{E})$ and we write $X \overset{d}{=} Y$ (d: distribution) or $X \overset{\mathcal{L}}{=} Y$ (in law).

Remarks (1) $X \overset{d}{=} Y \not\Rightarrow$ that they are defined on the same probability space. Μπορούμε να τις κάνουμε να είναι ορισμένες στον ίδιο χώρο πιθ, ∃ δηλαδή τέτοια κατασκευή, αλλά δε θα μας απασχολήσει

(2) $X \overset{d}{=} Y \Rightarrow E_{P_1}[h(X)] = E_{P_2}[h(Y)]$, $\forall h$ measurable if they are well defined.

(3) If X and Y are defined on the same probability space, then $X \overset{a.s.}{=} Y \Rightarrow X \overset{d}{=} Y$

(4) $X \overset{d}{=} Y \not\Rightarrow X \overset{a.s.}{=} Y$ (even if they are defined on the same probability space).

Proofs

(2) $E_{P_1}[h(X)] \overset{CVF}{=} E_{P_X}(h) \overset{X \overset{d}{=} Y}{=\!=} E_{P_Y}(h) = E_{P_2}[h(Y)]$

(3) Let $B \in \mathcal{E}$. Then

$\to P_1$ but we can use P in order to simplify

$$P_X(B) = P(X \in B) = P(X \in B, \underbrace{X = Y}_{\text{prob } 1}) = P(Y \in B, \underbrace{X = Y}_{\text{prob } 1}) = P(Y \in B) =$$

$$= P_Y(B)$$

$$\to P_X = P_Y \Rightarrow X \overset{d}{=} Y.$$

(4) $X = Z \sim N(0,1)$, $Y = -Z \sim N(0,1)$.

So, $X \overset{d}{=} Y$. But

$$P(X = Y) = P(Z = -Z) = P(Z = 0) \overset{Z \text{ cont.}}{=} 0 \Rightarrow X \overset{a.s.}{\neq} Y.$$

X τ.μ. συνεχής $\Leftrightarrow$ f συνεχής

$\Leftrightarrow$ $P(X = \text{σταθ}) = 0$

Distributions on $\mathbb{R}$ with density (κατανομές με πυκνότητα)

Definition: Let P be a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, λ the Lebesgue measure and $f : \mathbb{R} \to [0, +\infty]$ a Borel-measurable function. The function f is called a density of the measure P, if (→ not unique)

$$P(B) = \int_B f \, d\lambda = \int_B f(x)\lambda(dx)$$ if probability assigned to a Borel set can be written as a Lebesgue integral

$$\forall B \in \mathcal{B}(\mathbb{R}).$$

Remarks:

(1) $\int_{\mathbb{R}} f \, d\lambda = 1 \quad (= P(\mathbb{R}))$ necessary condition for a density.

Since $\int_{\mathbb{R}} f \, d\lambda = 1 < +\infty \Rightarrow f < +\infty, \ \lambda\text{-a.e.}$

not the derivative of f.

(2) If $f' = f, \lambda$-a.e, and f' is Borel-measurable, ≥ 0, then f' will also be a density of P. (a class of functions that are densities of P)

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad \int_B f' \, d\lambda \overset{\text{we should}}{=\!=} P(B).$$

Take $B \in \mathcal{B}(\mathbb{R})$

$$f = f', \lambda\text{-a.e} \Rightarrow f \cdot \mathbb{1}_B = f' \mathbb{1}_B, \lambda\text{-a.e.} \rightarrow \int f' \mathbb{1}_B \, d\lambda = \int f \mathbb{1}_B \, d\lambda \Rightarrow$$

$$\Rightarrow \int_B f' \, d\lambda = \int_B f \, d\lambda = P(B) \Rightarrow f' \text{ is also a density}$$

Proposition: If f_1 and f_2 are two Borel-measurable functions, ≥ 0 which are densities for the same P, then we should have $f_1 = f_2$ λ-a.e

Proof

Take $A = \{f_1 > f_2\}$ which is a Borel-measurable set, since f_1, f_2 are Borel-measurable.

$$P(A) = \int_A f_1 \, d\lambda = \int_A f_2 \, d\lambda \Rightarrow \int_A (\underbrace{f_1 - f_2}_{>0}) \, d\lambda = 0$$

We have $A = \{f_1 - f_2 > 0\}$

Lemma: If $g > 0$ on a set A, then:

$$\int_A g \, d\lambda = 0 \rightarrow \lambda(A) = 0 \quad \text{for any arbitrary measure}$$

We could also have $A \subset \{g > 0\}$ and have this Lemma hold.

<u>Lemma Proof</u>

Let $A_n = A \cap \{g \geq \frac{1}{n}\}$ $\forall n \geq 1$. Then $A_n \in \mathcal{B}(\mathbb{R})$ and

$$0 = \int_A g \, d\lambda \overset{g>0}{\underset{A_n \subset A}{\geq}} \int_{A_n} g \, d\lambda \geq \frac{1}{n} \lambda(A_n) \geq 0 \Rightarrow \lambda(A_n) = 0 \quad \forall n \geq 1.$$

Notice

$$\bigcup_{n \geq 1} (A \cap \{g \geq \tfrac{1}{n}\}) = A \cap \left(\bigcup_{n \geq 1} \{g \geq \tfrac{1}{n}\} \right) = A \cap \{g > 0\}.$$

So

$$\lambda(A) \overset{A \subset \{g>0\}}{=} \lambda(A \cap \{g > 0\}) = \lim_n \lambda(\underbrace{A \cap \{g \geq \tfrac{1}{n}\}}_{A_n}) = 0$$

because $(A_n) \uparrow$ and $\bigcup_n A_n = \lim_n A_n$. $\blacksquare$

Now, we have proved that $\lambda(\{f_1 > f_2\}) = 0 \Rightarrow \lambda(\{f_2 > f_1\}) = 0 \Rightarrow$

$$\Rightarrow \lambda(\{f_1 \neq f_2\}) = \lambda(\{f_1 > f_2\}) + \lambda(\{f_2 > f_1\}) = 0 \Rightarrow f_1 = f_2 \quad \lambda\text{-a.e.}$$

! Important Remark

If $(\Omega, \mathcal{A}, \mu)$ is a measure space and f is a Borel-measurable $f: \Omega \to [0, +\infty]$. The the set function:

$$\nu(A) = \int_A f \, d\mu$$

απο ολοκληρώματα παίρνουμε μέτρο.

is a measure on $(\Omega, \mathcal{A}, \mu)$.

Proof

$\nu(A) \geq 0, \ \forall A \in \mathcal{A}$

i) $\nu(\emptyset) = \int_{\emptyset} f \, d\mu = \int f \overset{0}{\overset{\shortparallel}{\mathbb{1}_{\emptyset}}} \, d\mu = 0$

ii) $(A_n)_+$ in $\mathcal{A}$, then

$$\nu\left(\bigcup_n A_n\right) = \int f \underbrace{\mathbb{1}_{\bigcup_n A_n}}_{\substack{\text{|| disjoint} \\ \sum_n \mathbb{1}_{A_n}}} d\mu = \int \left(\sum_n f \cdot \mathbb{1}_{A_n}\right) d\mu \overset{\text{Beppo-}}{\underset{\text{Levy}}{=}} \sum_n \int_{A_n} f \, d\mu = \sum_n \nu(A_n)$$

So when we define Lebesgue int in ≥ 0 function, we create measures

Corollary: If $f: \mathbb{R} \to [0, +\infty]$ is Borel measurable and $\int f \, d\lambda = 1$, then the set function P: $P(A) = \int_A f \, d\lambda$ is a probability measure and so f is a density of a probability measure.

Remark:
So we obtain necessary and sufficient conditions for a function to be a density of the probability measure.

Definition: Let $(\Omega, \mathcal{A}, P)$ be a probability space and $X: \Omega \to \bar{\mathbb{R}}$ be a random variable. If $f: \mathbb{R} \to [0, +\infty]$ is a Borel-measurable, then we say that f is a density of X, if it is a density of its distribution P_X. (density function, probability density function)

Proposition: Let $(\Omega, \mathcal{A}, P)$ be a probability space and $X: \Omega \to \mathbb{R}$ be a random variable with density $f: \mathbb{R} \to [0, +\infty]$. If $h: \mathbb{R} \to \bar{\mathbb{R}}$ is measurable, then

$$E_P[h(X)] = \int_{\mathbb{R}} h(x) f(x)\, dx \quad (dx \equiv \lambda(dx)),$$

when some of these quantities is well defined.

Remark: This justifies the computation of the mean of a random variable which has density as defined in the course "Probability I".

Proof

Firstly, we have:

$$E_P[h(X)] \overset{CVF}{=} E_{P_X}(h).$$

So it suffices to show that:

$$E_{P_X}(h) = \int_{\mathbb{R}} h(x) f(x)\, dx$$

S1) Take $h = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$, $A_i \in \mathcal{B}(\mathbb{R})$, $1 \le i \le n$ (simple, ≥ 0, measurable)

$$E_{P_X}(h) = E_{P_X}\left[\sum_{i=1}^{n} a_i \mathbb{1}_{A_i}\right] = \sum_{i=1}^{n} a_i P_X(A_i) = \sum_{i=1}^{n} a_i \int_{A_i} f(x)\,dx =$$

$$= \int_{\mathbb{R}} \underbrace{\left(\sum_{i=1}^{n} a_i \mathbb{1}_{A_i}\right)}_{h(x)} f(x)\,dx = \int_{\mathbb{R}} h(x) f(x)\,dx.$$

S2) Take h to be measurable, ≥ 0. Then:

$\exists (h_n)\uparrow$ of simple, ≥ 0, measurable such that $h = \lim_n h_n$.

So

$$E_{P_X}(h) = E_{P_X}(\lim_n h_n) \overset{MCT}{=\!=} \lim_n E_{P_X}(h_n) \overset{S1}{=\!=} \lim_n \int_{\mathbb{R}} h_n(x) f(x)\,dx$$

$$= \int_{\mathbb{R}} \lim_n \left(h_n(x) f(x)\right) dx = \int_{\mathbb{R}} \lim_n h_n(x) f(x)\,dx =$$

$$= \int_{\mathbb{R}} h(x) f(x)\,dx$$

S3) Homework. $(h = h^+ - h^-)$ $\left(\int X(\omega)\,dP(\omega) = \int x\,dP_X(x)\right)$

Examples

Normal distribution

$X \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $x \in \mathbb{R}$

Exponential Distribution

$$X \sim Exp(\theta), \quad f(x) = \theta e^{-\theta x} \mathbb{1}_{(0,+\infty)}(x)$$
$\hookrightarrow \theta > 0$

Standard Cauchy

$$X \sim Cauchy(0,1), \quad f(x) = \frac{1}{\pi} \cdot \frac{1}{(1+x^2)}, \quad x \in \mathbb{R}$$

$(f(x) = f(-x))$

$X \overset{d}{=} -X.$

$E(X^+) = +\infty$

$E(X^-) = +\infty$

Discrete Distributions

Reminder: A probability measure P defined on a measurable space $(E, \mathcal{E})$ is said to be a <u>discrete probability measure</u> if

$\exists S \subset E$, countable with $P(S) = 1$.

Proposition: If P is a discrete probability measure on E, then:

$\forall h: E \to \bar{\mathbb{R}}$ we have

$$E[h(X)] = \int_{\mathbb{R}} h(x)\, dP(x) = \sum_{x \in S} h(x) P(\{x\})$$

(Homework)

Definition: A random variable: $X: \Omega \to E$ is called <u>discrete</u> if its distribution P_X is a discrete probability measure.

series of Dirac's weighted by probabilities.

The function $f: E \to [0,1]$, where $f(x) = P(X = x)$ is called <u>probability function</u> or <u>probability mass function</u> or <u>mass function</u> of the random variable X.

Theorem :

$$\text{distribution on } \mathbb{R} \begin{cases} \to \text{discrete} \\ \to \text{continuous} \begin{cases} \to \text{absolutely continuous (απολύτως)} \\ \to \text{singularly continuous (ιδιάζουσα)} \end{cases} \end{cases}$$

discrete: $\exists S \subset \mathbb{R} : P(S) = 1$

continuous: F is continuous $\longleftrightarrow$ $P(\{x\}) = 0 \quad \forall x \in \mathbb{R}$

$\downarrow$ (the distribution function)

absolutely continuous : P has a density

singularly continuous: $\exists A \in \mathcal{B}(\mathbb{R}) : \lambda(A) = 0$, such that $P(A) = 1$ and P is continuous.

Absolute Continuity

Remark: P has a density $\longleftrightarrow$ $P(B) = \int_B \underbrace{f}_{\text{density.}} d\lambda$.

Then,

$\lambda(B) = 0 \Rightarrow P(B) = 0$ Lebesgue measure 0 gives 0 probability

Definition : Let μ, ν be two measures on a measure space. $(E, \mathcal{E})$.
We say that ν is absolutely continuous with respect to μ
(πάντα είναι αναφορικά με κάποιο μέτρο) if:

$\mu(A) = 0 \Rightarrow \nu(A)$

We write $\nu \ll \mu$.

Remark: We conclude that if P has a density, then:

$P \ll \lambda$ (density $\equiv$ Radon–Nikodym derivative)

Radon-Nikodym Theorem (II)

$\mu \ \lambda$

$P \ll \lambda$, λ σ-finite measure $\Rightarrow$

$$\Rightarrow P(A) = \int_A f \, d\lambda$$

$$\mu(A) = \int_A f \, d\lambda$$

f → Radon-Nikodym derivative. We can denote: $f = \frac{d\mu}{d\lambda}$

So,

$$\mu(A) = \int_A \frac{d\mu}{d\lambda} \, d\lambda.$$

We can write: $\mu(A) = \int_A d\mu$ Αν μπορείς να βρεις μέτρο απολ. συνεχ

λ $\mu \ll \lambda$

wrt άλλο μέτρο, τότε μπορείς να το γράψεις σαν μια $\mu(A) = \int_A \frac{d\mu}{d\lambda} \, d\lambda$

Μπορείς να ολοκληρώνεις ως προς άλλο την Radon-Nikodym παρ. του ως προς το άλλο.

So we can have these representations:
of measures as integrals of other measures.

Product measures (μέτρα γινόμενο)

Definition: Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces. We call measurable rectangle (μετρήσιμο ορθογώνιο) on $X \times Y$, every set of the form $A \times B$, where $A \in \mathcal{A}$, $B \in \mathcal{B}$.

Additionally, we define

$$\mathcal{A} \otimes \mathcal{B} := \sigma(\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\})$$

the σ-algebra generated from the measurable rectangles and we call it the product σ-algebra (σ-άλγεβρα γινόμενο) of $\mathcal{A}$ and $\mathcal{B}$.

Proposition: If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two spaces of σ-finite measure, then

$\exists$ a unique measure m on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$: $m(A \times B) = \mu(A) \cdot \nu(B)$ $\forall A \in \mathcal{A}$, $\forall B \in \mathcal{B}$.

Prob I:
independent X, Y: $P(X, Y) = P(X) \cdot P(Y)$.

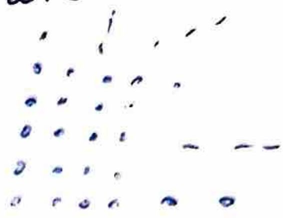
This measure is denoted $\mu \otimes \nu$ and is called product measure of μ and ν.

The space $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ is called product space (χώρος γινόμενο) of the initial ones.

Examples: (1) Let ν_1 be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Then, since $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu_1)$ is a σ-finite measure space. We can define the product space

$$(\mathbb{N}^2, \mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N}), \nu_1 \otimes \nu_1). \quad (1)$$

We will show that $(1) = (\mathbb{N}^2, \mathcal{P}(\mathbb{N}^2), \nu_2)$, where ν_2 is the counting measure on $\mathbb{N}^2$.



i) $\underline{\mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N}) = \mathcal{P}(\mathbb{N}^2)}$.

We have that:

$$\mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) = \sigma(\{A \times B : A, B \subset \mathbb{N}\}) \subset \mathcal{P}(\mathbb{N}^2).$$

Now, if $\Delta \subseteq \mathbb{N}^2$, then Δ is countable. So

$$\Delta = \bigcup_{\substack{i \in I \\ \downarrow \\ \text{countable}}} \{(a_i, b_i)\} = \bigcup_{i \in I} \underbrace{\{a_i\}}_{A_i} \times \underbrace{\{b_i\}}_{B_i} \in \sigma(\{A \times B : A, B \subset \mathbb{N}\}) \Rightarrow$$

$$\Rightarrow \Delta \in \mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N})$$

Therefore:

$$\mathcal{P}(\mathbb{N}^2) = \mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N}).$$

ii) $\nu_2(A \times B) = |A \times B| = |A| \cdot |B| = \nu_1(A) \cdot \nu_1(B) = (\nu_1 \otimes \nu_1)(A \times B)$

By the uniqueness of the extension, since they coincide in the rectangles, and because ν_1 is σ-finite:

$$\nu_2 = \nu_1 \otimes \nu_1.$$

(2) Let λ_1 be the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Then, since $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ is a σ-finite measure space, we can define the product measure space:

$\mathbb{R} = \bigcup_n [-n,n]$

$\lambda_1([-n,n]) = 2n$ finite.

$$(\mathbb{R}^2, \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}), \lambda_1 \otimes \lambda_1).$$

It can be shown:

$$(\mathbb{R}^2, \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}), \lambda_1 \otimes \lambda_1) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda_2),$$

where λ_2 is the Lebesgue measure on $\mathbb{R}^2$ which assigns the <u>area</u> of $A \in \mathcal{B}(\mathbb{R}^2)$.
(εμβαδόν)

<u>Proposition</u>: If $(X_i, \mathcal{A}_i, \mu_i)$ $1 \le i \le n$ are n spaces of σ-finite measure, then

cartesian product

$\exists$ a unique measure m on the measurable space $\left(\prod_{i=1}^{n} X_i, \bigotimes_{i=1}^{n} \mathcal{A}_i\right)$, where

$$\bigotimes_{i=1}^{n} = \sigma\left(\{A_1 \times A_2 \times \dots \times A_n : \underset{1 \le i \le n}{A_i \in \mathcal{A}_i}\}\right),$$

such that:

$$m(A_1 \times A_2 \times \dots \times A_n) = \mu_1(A_1) \cdot \mu_2(A_2) \cdots \mu_n(A_n) \quad \forall A_i \in \mathcal{A}_i, \ 1 \le i \le n.$$

and the measure m is called product measure of $\mu_1, \mu_2, \dots, \mu_n$ and is denoted by:

$$\bigotimes_{i=1}^{n} \mu_i \text{ or } \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_n$$

Examples:

(1) $\left(\mathbb{N}^n, \bigotimes_{i=1}^{n} \mathcal{P}(\mathbb{N}), \bigotimes_{i=1}^{n} \nu_1\right) \overset{HW}{=\!=} \left(\mathbb{N}^n, \mathcal{P}(\mathbb{N}^n), \nu_n\right)$

↳ counting measure on n dimensions.

(2) $\left(\mathbb{R}^n, \bigotimes_{i=1}^{n} \mathcal{B}(\mathbb{R}), \bigotimes_{i=1}^{n} \lambda_1\right) \overset{HW}{=\!=} \left(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^n), \lambda_n\right)$

↳ n-dimensional Lebesgue measure.

are exactly the measures that are induced by independent r.v.