Probability-2 Lecture-9

06 February 2024 11:18

$(\Omega, \mathcal{A}, P)$ - probability space.

Ω: real set; $\mathcal{A}$: σ-field; P: probability on $\mathcal{A}$.

$X: \Omega \to \mathbb{R}$ is called a <u>real r.v.</u> if (equivalent statements).

$$(*)\begin{cases} \{w : X(w) \le c\} \in \mathcal{A} & \forall\, c \in \mathbb{R}. \\ \{w : X(w) < c\} \in \mathcal{A} & \forall\, c \in \mathbb{R}. \\ \{w : X(w) > c\} \in \mathcal{A} & \forall\, c \in \mathbb{R}. \\ \{w : X(w) \ge c\} \in \mathcal{A} & \forall\, c \in \mathbb{R} \end{cases}$$

$\iff \{w : X(w) \in B\} \in \mathbb{R}$ $\forall\, B \in \mathcal{B}$

$\mathcal{B}$: Borel σ-field on $\mathbb{R}$.

$X : \Omega \to [-\infty, \infty]$ $(= \mathbb{R} \cup \{\infty\} \cup \{-\infty\})$ is called an extended real r.v. if $(*)$ holds. $\iff$

$$\{w : X(w) \in B\} \in \mathcal{A} \quad \forall\, B \in \mathcal{B}$$

and, $\{w : X(w) = \infty\} \in \mathcal{A}$
$\{w, X(w) = -\infty\} \in \mathcal{A}$.

Let X be a real random variable on $(\Omega, \mathcal{A}, P)$.

then, $P_X(B) := P(X^{-1}(B))$, $B \in \mathcal{B}$ defined a prob. on B and is called the probability dist^n of X.

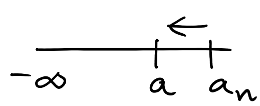
The function $F_X : \mathbb{R} \to \mathbb{R}$ defined by $F_X(a) = P_X((-\infty, a])$
$= P(X \le a)$ $\forall\, a \in \mathbb{R}$.

is called "Cumulative Distribution Function" (CDF).

<u>Properties of F_X:</u>

(i) F_X is non-decreasing.

(ii) F_X is right-continuous, non-negative.



$a \in \mathbb{R}$. $\quad (-\infty, a_n] \searrow (-\infty, a]$

$a_n \searrow a$. $\quad \Rightarrow P_X((-\infty, a_n]) \to P_X((-\infty, a])$

$a_n \searrow a. \quad \Rightarrow P_X((-\infty, a_n]) \to P_X((-\infty, a])$

$\Rightarrow F_X(a_n) \longrightarrow F_X(a).$

(iii) $F_X(a^-) = P(X<a)$

So, F_X is continuous at $a \Leftrightarrow$

$P(X<a) = P(X \leq a).$

$\Leftrightarrow P(X=a) = 0$

F_X has "jump discontinuity" at $a \leftrightarrow P(X=a) > 0$

and $\Delta F_X(a) = F_X(a) - F_X(a^-)$

$= P(X=a).$

(iv) $\lim_{a\to\infty} F_X(a) = 1.$ i.e, $a_n \nearrow \infty.$ $\therefore F_X(a_n) \to P_X(\mathbb{R}) = 1$

$(-\infty, a_n] \nearrow \mathbb{R}$

$\lim_{a\to-\infty} F_X(a) = 0.$ i.e, $a_n \searrow -\infty$

$(-\infty, a_n] \searrow \phi.$ $\therefore F_X(a_n) \longrightarrow P_X(\phi) = 0$

<u>Definition:</u> <u>(Discrete r.v.)</u>

A real random variable on $(\Omega, \mathcal{A}, P)$ is called <u>discrete</u> if X takes only countably many values, say, $D_X = \{x_1, x_2, \dots\}.$

In this case,

$X = \sum x_n \cdot 1_{A_n}$, where $\{A_n\}_{n\geq 1}$ is a partition of Ω by <u>sets</u> in $\mathcal{A}$. (non-empty)

Put $p(x_n) = P(A_n) = P(X = x_n)$

$P_X(B) = \sum_{n: x_n \in B} p(x_n) \cdot 1_B(x_n).$

$\|$

$$P_X(B) = \sum_{n: x_n \in B} p(x_n)\, I_B(x_n)$$

$$\| $$

$$P(X^{-1}(B))$$

$$\| $$

$$P(X \in B)$$

$$F_X(a) = \sum_{n: x_n \leq a} p(x_n)$$

<u>Definition: (Continuous Random Variable):</u>

A real r.v. is called **Continuous** if F_X is continuous everywhere.

$$\boxed{X \text{ continuous} \iff P(X=a)=0 \quad \forall\ a \in \mathbb{R}.}$$

ie, X is continuous if it has <u>no point mass</u>.

<u>Special Case of continuous random variable:</u> (Absolutely continuous r.v.)

Let $f: \mathbb{R} \to \mathbb{R}$ be a non-negative function s.t.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1, \quad \text{and} \quad F_X(a) = \int_{-\infty}^{a} f(x)\,dx.$$

why special?

This integral may <u>not</u> be simply Riemann.

$\therefore$ there may exist r.v. s.t it satisfies the def^n of cont. r.v but <u>not</u> this property above.

<u>Result:</u> Let $F: \mathbb{R} \to \mathbb{R}$ be:

(i) Non-decreasing.

(ii) Right-continuous everywhere

(iii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$.

then, $\exists$ probability space $(\Omega, \mathcal{A}, P)$, & a real r.v X on $(\Omega, \mathcal{A}, P)$, s.t. $F \equiv F_X$.

<u>Proof.</u> clearly, $0 \leq F(a) \leq 1 \quad \forall\ a \in \mathbb{R}$.

Proof. clearly, $0 \leq F(a) \leq 1 \quad \forall\, a \in \mathbb{R}$.

for $-\infty \leq a \leq b \leq \infty$,

denote $I_{a,b} = \{x \in \mathbb{R} : a < x \leq b\}$

$\mathcal{S} = \{ I_{a,b} : -\infty \leq a \leq b \leq \infty \}$ is a semifield.

Define μ on $I_{a,b}$ by $\mu(I_{a,b}) = F(b) - F(a)$

$\therefore\ \mu \geqslant 0,\ \mu(\mathbb{R}) = 1 \qquad (F(+\infty) := 1,\ F(-\infty) := 0)$

To prove:

If $I_{a_n,b_n} \in \mathcal{S}_{n \geqslant 1}$ are disjoint & $\bigcup_n I_{a_n,b_n} = I_{a,b}$,

then $\mu(I_{a,b}) = \sum_n \mu(I_{a_n,b_n})$

Lemma 1: $\left.\begin{array}{l} I_{a_n,b_n},\ n \geqslant 1 \text{ disjoint} \\ \text{and } \bigcup_n I_{a_n,b_n} \subset I_{a,b} \end{array}\right] \Rightarrow \sum_n \mu(I_{a_n,b_n}) \leq \mu(I_{a,b})$

Lemma 2: $\left.\begin{array}{l} I_{a_n,b_n},\ n \geqslant 1 \text{ s.t.} \\ I_{a,b} \subset \bigcup_n I_{a_n,b_n} \end{array}\right] \Rightarrow \mu(I_{a,b}) \leq \sum_n \mu(I_{a_n,b_n})$.

Proof: (Lemma 1).

hypothesis implies that for any $n \geqslant 1$,

$$\bigcup_{k=1}^{n} I_{a_k,b_k} \subset I_{a,b}.$$

Order the a_i's, $1 \leqslant k \leq n$ in increasing order:

$$a \leqslant a_1 \leqslant b_1 \leq a_2 \leq b_2 \leqslant \cdots \leq a_n \leqslant b_n \leq b.$$

$$\therefore\ F(b) - F(a) \geqslant F(b_n) - F(a_1)$$

$$\overset{\shortparallel}{\mu(I_{a,b})} \geqslant \sum_{k=1}^{n} \left(F(b_k) - F(a_k)\right) = \sum_{k=1}^{n} \mu(I_{a_k,b_k})$$

$$\mu(I_{a,b}) \geq \sum_{k=1}^{n} \left(F(b_k) - F(a_k)\right) = \sum_{k=1}^{n} \mu(I_{a_k,b_k})$$

for countable, take $n \to \infty$.

Proof: (lemma 2):

Claim: enough to prove this with $-\infty < a < b < \infty$.

Suppose $b = +\infty$. Replace $I_{a,b} = I_{a,b\wedge m}$

$\therefore$ Assume $-\infty < a < b < \infty$.

Step 1: $-\infty < c < d < \infty$ finite union of sets.

$$I_{c,d} \subset \bigcup_{k=1}^{n} I_{c_k,d_k} \Rightarrow \mu(I_{c,d}) \leq \sum_{k=1}^{n} \mu(I_{c_k,d_k})$$

true for $n=1$.

Assume: true for $(n-1)$

If $I_{c,d} \subset \bigcup_{k=1}^{n} I_{c_k,d_k}$, assume $d \in I_{c_n,d_n}$ (W.L.O.G). last interval

$$\therefore I_{c,c_n} \subset \bigcup_{k=1}^{n-1} I_{c_k,d_k}$$

$[\because \mu(I_{c_n,d}) \leq \mu(I_{c_n,d_n})]$

$$\mu(I_{c,d}) = \mu(I_{c,c_n}) + \mu(I_{c_n,d}) \leq \sum_{k=1}^{n-1} \mu(I_{c_k,d_k}) + \mu(I_{c_n,d_n})$$

added qty. added qty.

$$= \sum_{k=1}^{n} \mu(I_{c_k,d_k})$$

$$F(d) - F(c_n) + F(c_n) - F(c) \leq \sum_{k=1}^{n-1} \mu(I_{c_k,d_k}) + F(d_n) - F(c_n)$$

Fix $\varepsilon > 0$.

Use right continuity of F to get $a < \tilde{a} < b$ s.t.

$$F(\tilde{a}) < F(a) + \varepsilon/2$$

define $\tilde{b}_n = b_n$ if $b_n = \infty$.

choose $\tilde{b}_n > b_n$ such that

$$F(\tilde{b}_n) < F(b_n) + \frac{\varepsilon}{2^{n+1}}$$

2^{n+1}

$$I_{a,b} = (a,b] \in \bigcup_n I_{a_n, b_n}$$

$$\Rightarrow [\tilde{a}, b] \subset \bigcup_n (a_n, \tilde{b}_n)$$

$\exists\ n$ s.t.

$$I_{a,b} \subset [\tilde{a}, b] \subset \bigcup_{k=1}^{n} (a_k, \tilde{b}_k) \subset \bigcup_{k=1}^{n} I_{a_k, \tilde{b}_k}$$

$$\mu(I_{\tilde{a},b}) \leq \sum_{k=1}^{n} \mu(I_{a_k, \tilde{b}_k})$$

$$\Rightarrow F(b) - F(\tilde{a}) \leq \sum_{k=1}^{\infty} \left(F(\tilde{b}_k) - F(a_k)\right) + \varepsilon$$

$$F(\tilde{b}_k) < F(b) + \frac{\varepsilon}{2^{n+1}},$$

& there are n such terms.

Similarly for $F(a_k)$.

Take $X : x \mapsto x$ on $(\mathbb{R}, \mathcal{B}, \mu)$.

Ex: Show, $F \equiv F_X$

Example: Let $F : \mathbb{R} \to \mathbb{R}$ be

$$F(a) = \begin{cases} 0, & \text{if } a \leq 0 \\ a, & 0 < a \leq 1 \\ 1, & a > 1. \end{cases}$$

Here, note that

$\lim_{a \to -\infty} F(a) = 0$

$\lim_{a \to \infty} F(a) = 1.$

$F(a) \geq 0$, non-decreasing.

We can find a real r.v s.t. F is the distn fn of X.