Probability & Real Analysis

Wednesday, October 05, 2011
2:57 PM

Sponsored by the National Grid Foundation nationalgrid The power of action.

① Probability Space $(\Omega, \mathcal{F}, P)$

- algebra, σ-algebra; basic properties
- measurable space; example $(\mathbb{R}, \mathcal{B})$
- probability measure
- properties equivalent to countable additivity

② Construction of a probability measure on $\mathbb{R}$

③ Random variables and their distributions

① Ω - sample space (set of all possible outcomes of a random experiment)

just a set

Def: Ω is said to be discrete if it has either finite or countably infinite nr. of elements

(we discussed this case in the Prob. Refresher Course)

- the goal now is to develop a "language" for dealing with much larger sample spaces

ex: For Brownian Motion on $[0,T]$ the sample space is $\Omega = C([0,T])$: the set of all continuous functions on $[0,T]$ $\rightsquigarrow$ $w \in \Omega$ means "the whole path" on $[0,T]$

Def 1: A collection $\mathcal{A}$ of subsets of Ω is called an **ALGEBRA** if it has three of the following properties:

(1) $\Omega \in \mathcal{A}$

(2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ $\quad [A^c = \Omega \setminus A]$

(3) $A_1, A_2, \dots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

Lemma 1: Let $\mathcal{A}$ be an algebra of subsets of Ω, Then

(1) $\phi \in \mathcal{A}$

(2) $A_1, A_2, \dots, A_n \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{n} A_i \in \mathcal{A}$

(3) $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \setminus A_2 \in \mathcal{A}$

Lemma 2: If $\mathcal{A}$ is a finite algebra, then there exits non empty sets $E_1, E_2, \dots, E_m \in \mathcal{A}$ such that

$$\left.\begin{array}{l} (1)\ E_i \cap E_j = \phi \quad \text{for } i \neq j \\ (2)\ \Omega = \bigcup_{i=1}^{m} E_i \end{array}\right\} \text{we say } E_1, E_2, \dots, E_m \text{ form a partition of } \Omega$$

(3) For every set $A \in \mathcal{A}$ there exists a set $I \subseteq \{1, 2, \dots, m\}$ such that $A = \bigcup_{i \in I} E_i$

Proof: [Koralov, Sinai : Th. of Probability and random proc. p.5]

Remark: Lemma 2 says that every finite algebra is generated by a finite partition of Ω.

Definition 3: A collection of subsets of Ω, $\mathcal{F}$, is called a **σ-ALGEBRA** if $\mathcal{F}$ is an algebra that satifies

(4) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{F}$

(we say that $\mathcal{F}$ is closed under taking countable unions)

Lemma 4: If $\mathcal{F}$ is a σ-algebra, then

$A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcap_{n\geq 1} A_n \in \mathcal{F}$

Def 4: Let $\mathcal{E}$ a collection of snbsets of Ω

The **σ-algebra generated by $\mathcal{E}$**, $\sigma(\mathcal{E})$ is the smallest σ-algebra that contains $\mathcal{E}$.

notation : $$\sigma(\mathcal{E}) = \bigcap_{\substack{\mathcal{E} \subseteq \Sigma \\ \Sigma \text{ is } \sigma \text{ algebra.}}} \Sigma$$

Remark: intersection of any number of σ-algebcas is still a σ-algebra.

Definition 5: $(\Omega, \mathcal{F})$ is called a **MEASURABLE SPACE** (where $\mathcal{F}$ is a σ-algebra of Ω)

· elements of $\mathcal{F}$ are called **MEASURABLE SETS** or **EVENTS**

example : $(\mathbb{R}, \mathcal{B})$

$\mathbb{R}$ = set of all real numbers

$\mathcal{B}$ = Borel σ-algebra

$\mathcal{B}$ = Borel σ-algebra

Definition 6: Borel σ-algebra on $\mathbb{R}$ is the σ-algebra generated by all open sets in $\mathbb{R}$.

Lemma 5 $\mathcal{E} = \{(-\infty; a] \mid a \in \mathbb{Q}\}$ Then $\mathcal{B} = \sigma(\mathcal{E})$

$\rightsquigarrow$ $\mathcal{B}$ = very large set !!!

$\mathcal{B}$ is countably generated

(there is a countable collection of sets $\mathcal{E}$ such that $\sigma(\mathcal{E}) = \mathcal{B}$)

• for the proof see : Jacod & Protter : (p. 8)

Definition 6: A probability measure defined on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that

(1) $P(\Omega) = 1$

(2) for every sequence $\{A_n\}_{n \geq 0} \subseteq \mathcal{F}$, such that $A_i \cap A_j = \phi$ for $i \neq j$,

$$\boxed{P\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} P(A_n)}$$

" countable additivity "

Theorem 1: If P is a probability measure on $(\Omega, \mathcal{F})$ then

(i) $P(\phi) = 0$

(ii) P is (finitely) additive

(iii) If $A_1, A_2 \in \mathcal{F}$ and $A_1 \subseteq A_2 \Rightarrow P(A_1) \leq P(A_2)$

Theorem 2: Let $\mathcal{F}$ be a σ-algebra and $P: \mathcal{F} \to [0,1]$ is such that $P(\Omega) = 1$ and is additive. Then the following are equivalent

(i) P is countably additive

(ii) If $A_n \in \mathcal{F}$, $n \geq 1$ and $A_n \downarrow \emptyset$ then $P(A_n) \downarrow 0$

(iii) If $A_n \in \mathcal{F}$, $n \geq 1$ and $A_n \downarrow A$ then $P(A_n) \downarrow P(A)$

(iv) If $A_n \in \mathcal{F}$, $n \geq 1$ and $A_n \uparrow \Omega$ then $P(A_n) \uparrow 1$

(v) If $A_n \in \mathcal{F}$, $n \geq 1$ and $A_n \uparrow A$ then $P(A_n) \uparrow P(A)$

Notation:

- $A_n \downarrow A$ means : $A_{n+1} \subseteq A_n$ for all n and $\bigcap_{n=1}^{\infty} A_n = A$
- $A_n \uparrow A$ means : $A_n \subseteq A_{n+1}$ for all n and $\bigcup_{n=1}^{\infty} A_n = A$

Definition 7: P : probability measure on $(\Omega, \mathcal{F})$. $A \subseteq \Omega$ is said to be **NULL SET** if there is a $B \in \mathcal{F}$ such that $A \subseteq B$ and $P(B) = 0$

Definition 8: We say that a property holds **ALMOST SURELY** (a.s) (with respect to P) if the set of all ω for which the property does not hold forms a null set.

Theorem 3: Let P be a probability on $\mathcal{F}$ and

$\mathcal{N}$ be the collection of all null sets. Then
$\mathcal{F}' = \{ A \cup N : A \in \mathcal{F}, N \in \mathcal{N} \}$ is a σ-algebra called **the P-completion of $\mathcal{F}$**. This is the smallest σ-algebra containing $\mathcal{F}$ and $\mathcal{N}$, and P extends uniquely as a probability (still denoted P) on $\mathcal{F}'$ by setting

$$\boxed{P(A \cup N) = P(A) \quad \text{for} \quad A \in \mathcal{F}, N \in \mathcal{N}.}$$

Given an algebra $\mathcal{A}_0$ that generates $\mathcal{F} = \sigma(\mathcal{A}_0)$ we construct P on $\mathcal{A}_0$ so that

(1) $P(\Omega) = 1$

(2) If $A_1, A_2, \ldots \in \mathcal{A}_0$, $A_i \cap A_j \neq \emptyset$ if $i \neq j$, and $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_0$

then $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$

Theorem 4: Each $P : \mathcal{A}_0 \to [0,1]$ that satisfies (1) & (2) has a unique extension on $\mathcal{F}$.

② Special case : prob measures on $\mathbb{R}$

$(\mathbb{R}, \mathcal{B})$

$\mathcal{B} = \sigma(\mathcal{C}) \qquad \mathcal{C} = \{ (-\infty ; a] \mid a \in \mathbb{Q} \}$

Definition 9: The distribution function of P is given by

$$\boxed{P((-\infty;x]) \triangleq F(x)}$$

Theorem 5: The distribution function $F(\cdot)$ characterizes the prob. measure

Proof: Let $\mathcal{B}_0$ be the set of finite disjoint unions of intervals of the form: $(x;y]$ where $-\infty \le x \le y \le \infty$

ex:

- $\mathcal{B}_0$ is an algebra
- since every $(-\infty; a]$, $a \in \mathbb{Q}$ is an element of $\mathcal{B}_0$ then $\mathcal{B} \subseteq \sigma(\mathcal{B}_0)$
- $(a,b] = \bigcap_{n=1}^{\infty} (a, b+\frac{1}{n})$ so $\mathcal{B}_0 \subseteq \mathcal{B} \Rightarrow \sigma(\mathcal{B}_0) \subseteq \mathcal{B}$

Conclusion: $\boxed{\sigma(\mathcal{B}_0) = \mathcal{B}}$

$\rightsquigarrow$ now we have $P((x,y]) = F(y) - F(x)$ and for every $A \in \mathcal{B}_0$: $A = \bigcup_{1 \le i \le n} (x_i, y_i]$, $y_i < x_{i+1}$ and

$$P(A) = \sum_{i=1}^{n} \left(F(y_i) - F(x_i)\right)$$

Theorem 6: A function F is the distribution function of a (unique) probability measure on $(\mathbb{R}, \mathcal{B})$ if and only if

(i) F is non-decreasing

(ii) F is right continuous

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

<u>Proof</u>: If for some P : $F(x) = P((-\infty; x])$ then (i), (ii), (iii) can be easily checked by using the properties of P

⇝ (i) follows from Th1, part (iii) ;

⇝ (ii) follows from the representation

$$(-\infty; x] = \bigcap_{n=1}^{\infty} (-\infty; x_n] \quad \text{when} \quad x_n \downarrow x$$

and part (iii) of theorem 2.

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P((-\infty; x_n]) = P((-\infty; x]) = F(x)$$

⇝ (iii) follows from Theorem 2, parts (ii) and (iv)

The difficult part is to prove that given an F that satisfies (i) - (iii) we can find a prob. measure P on $(\mathbb{R}, \mathcal{B})$ such that $F(x) = P((-\infty; x])$

⇝ the way to prove this is to construct such a measure

(1) define it by hand on intervals of the form $(x, y]$ by $P((x,y]) = F(y) - F(x)$

⇝ this will define P on $\mathcal{B}_0$

(2) check the countable additivity property

(3) apply the extension theorem

· the actual work is in proving (2) : Since (2) is equivalent

- the actual work is in proving (2) : Since (2) is equivalent to any of the 4 properties of Theorem 2 we can choose to prove that if $A_n \in \mathcal{B}_0$, $A_n \downarrow \emptyset$ then $P(A_n) \downarrow 0$

(the proof is omitted ; it is not too hard but rather technical)

$\rightsquigarrow$ it can be found on p.41 of Jacod & Protter

Corollary : $F(\cdot)$ distr. function of P on $(\mathbb{R}, \mathcal{B})$

$F(x-) = \lim_{y \uparrow x} F(y)$ $\left(\underset{x \to \infty}{\text{limit}} \text{ exists} \right)$

for all $x < y$

(i) $P\big((x,y]\big) = F(y) - F(x)$

(ii) $P\big([x,y]\big) = F(y) - F(x-)$

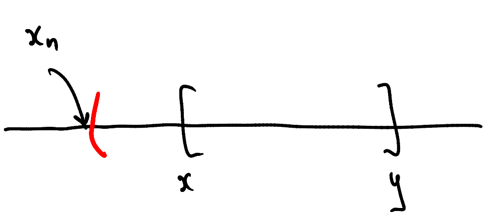
(iii) $P\big([x,y)\big) = F(y-) - F(x-)$

(iv) $P\big((x,y)\big) = F(y-) - F(x)$

(v) $P\big(\{x\}\big) = F(x) - F(x-)$

- $P(\{x\}) = 0$ iff F is continuous at x

<u>Proof</u> : (ii) $[x,y] = \bigcap_{n=1}^{\infty} (x_n, y]$ where $x_n \uparrow x$



$$P([x,y]) = P\left(\bigcap_{n=1}^{\infty} (x_n, y]\right) = \lim_{n\to\infty} P((x_n,y])$$

$$= \lim_{n\to\infty} \left[F(y) - F(x_n)\right] = F(y) - \underbrace{\lim_{n\to\infty} F(x_n)}_{\lim\limits_{x_n \uparrow x} F(x_n) = F(x-)} = F(y) - F(x-)$$

③ Random variables

Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces and $X : E \to F$ be a function

Definition 10: X is **MEASURABLE** (relative to $\mathcal{E}$ and $\mathcal{F}$) if for every $A \in \mathcal{F} \Rightarrow X^{-1}(A) \in \mathcal{E}$

Recall : $X^{-1}(A) := \{\omega \in E \mid X(\omega) \in A\}$

Shorthand notation : $X^{-1}(A) = \{X \in A\}$

- In probability theory, when we think of E as a sample space, we call every measurable function a random variable

- When $F = \mathbb{R}$, we shall always take $\mathcal{F}$ to be $\mathcal{B}$.

Theorem 7: Let $\mathcal{C}$ be a collection of subsets of $\mathcal{F}$ such that $\sigma(\mathcal{C}) = \mathcal{F}$. The function $X : E \to F$ is measurable (with respect to $\mathcal{E}$ and $\mathcal{F}$) iff $X^{-1}(\mathcal{C}) \subset \mathcal{E}$

<u>Proof</u> : "$\Rightarrow$" holds by definition

"$\Leftarrow$" we assume that $X^{-1}(\mathcal{C}) \subset \mathcal{E}$, we need to prove that $X^{-1}(A) \in \mathcal{E}$ for all $A \in \mathcal{F}$.

- $X^{-1}\left(\bigcup_{n \geq 1} A_n\right) = \bigcup_{n \geq 1} X^{-1}(A_n)$

- $X^{-1}\left(\bigcap_{n \geq 1} A_n\right) = \bigcap_{n \geq 1} X^{-1}(A_n)$

- $X^{-1}(A^c) = \left[X^{-1}(A)\right]^c$

Define $\mathcal{A} = \{ A \in \mathcal{F} : X^{-1}(A) \in \mathcal{E} \}$. Then $\mathcal{C} \subseteq \mathcal{A}$

$\rightsquigarrow$ by the 3 above properties and the fact that $\mathcal{E}$ is a σ-algebra we conclude that $\mathcal{A}$ is a σ-algebra.

$\rightsquigarrow$ therefore $\mathcal{A} \supseteq \sigma(\mathcal{C}) = \mathcal{F}$, but $\mathcal{A} \subseteq \mathcal{F}$ by definition so we have to have $\mathcal{A} = \mathcal{F}$

<u>Corollary</u> : Let $(F, \mathcal{F}) = (\mathbb{R}, \mathcal{B})$ and X_n, X be real valued functions on E

(i) X is measurable iff $\{X < a\} \in \mathcal{E}$ for all a

(ii) If $\{X_n\}_{n \geq 1}$ are measurable $\Rightarrow$

$$\inf_{n \geq 1} X_n , \quad \sup_{n \geq 1} X_n , \quad \limsup_{n \geq 1} X_n , \quad \liminf_{n \geq 1} X_n$$

are measurable as well.

(iii) If $\{X_n\}$ are measurable and $X_n \to X$ pointwise then

X is measurable.

<u>Proof</u>: (i) follows from Lemma 5 and Lemma 7.

(ii) $\{X_n \le a\} \in \mathcal{E}$; $\{\sup X_n \le a\} = \bigcap_{n=1}^{\infty} \{X_n \le a\} \in \mathcal{E}$

($\inf X_n$ is similar, take $\{X_n < a\}$ etc)

$$\limsup_{n\to\infty} X_n = \underbrace{\inf_n \underbrace{\sup_{m\geq n} X_m}_{\text{measurable}}}_{\text{measurable}}$$

(iii) If $\lim_{n\to\infty} X_n = X$ then $X = \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n$

<u>Theorem 8</u> Let X be measurable from $(E, \mathcal{E})$ to $(F, \mathcal{F})$ and Y be measurable from $(F, \mathcal{F})$ to $(G, \mathcal{G})$; then $Y \circ X$ is measurable from $(E, \mathcal{E})$ to $(G, \mathcal{G})$

<u>Theorem 9</u> : Every continuous function X from $\mathbb{R}^n$ to $\mathbb{R}$ is measurable ($(\mathbb{R}^n, \mathcal{B}^n)$ to $(\mathbb{R}, \mathcal{B})$)

<u>Theorem 10</u> : Let $(F, \mathcal{F}) = (\mathbb{R}, \mathcal{B})$ and $(E, \mathcal{E})$ be any measurable space. Then :

(i) For $A \subset E$, 1_A is measurable iff A is measurable

(ii) If $X_1, X_2, \dots, X_n$ are real measurable functions on $(E, \mathcal{E})$ and f is a Borel function on $\mathbb{R}^n$ (i.e. such that $f^{-1}(B) \in \mathcal{B}^n$ for all $B \in \mathcal{B}$) then

$f(X_1, X_2, \dots, X_n)$ is measurable

(iii) If X, Y are measurable then so are

$X+Y$, $X \cdot Y$, $X \vee Y$, $X \wedge Y$, X/Y (if $Y \neq 0$)

Definition 11: Let X be a real valued random variable on some on some probability space $(\Omega, \mathcal{F}, P)$. The distribution of X is the probability measure on $(\mathbb{R}, \mathcal{B})$ defined by

$$\boxed{P_X(B) = P(X^{-1}(B)) = P(X \in B)}$$

Remark: It is an easy check that P_X is a probability measure

$$F_X(x) := P_X((-\infty; x]) = P(X \leq x)$$

is called the cumulative distr. function of X.

When $F_X(x) = \int_{-\infty}^{x} f(y)\,dy$ for some f we say that X has a density f.