Sandipan Dey 17.03.05

Probability

1. Basic theory: prob. spaces, random variables, expectation, convergence concepts, independence, conditioning, limit theorems.

2. Markov chains: classification of states, stationary distributions.

3. Martingales: inequalities, convergence theorems.

3×35 tests. Feller

Basic: Elementary prob. theory ..., Chung +
Hoel, Port, Stone (1st & 3rd)
Pitman

Next level: First Course — Chung.
Probability — Bauer
An intro. to prob. & measure — Parthasarathy

Breiman — Probability ✓
Williams — ~~[illegible]~~ Weighing the odds
Probability in martingales

P is a probability measure (a countably [illegible] non-ve measure with total mass 1) on the measurable space $(\Omega, \mathcal{F})$

$\mathcal{F}$: a σ-field subsets of Ω containing itself, elements of $\mathcal{F}$ are called events

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

Toss a coin ⟶ Head

⟶ Tail

⟶ Stand on edge.

Consider all possible ~~[illegible]~~ set of parameters that determine the coin's motion → this is called the sample space Ω.

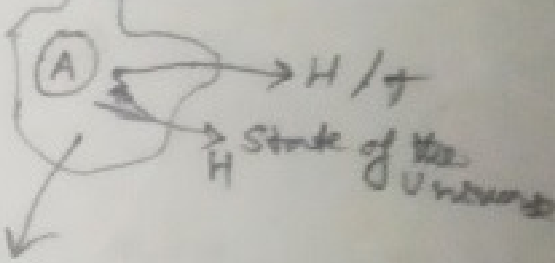
initial position

velocity

surface details

air friction

⋮



Ω : a set

elements of Ω, $\omega \in \Omega$ sample point

Ω is called 'sample space'

$A \subset \Omega$ to which we want to assign probability ~~[illegible] between~~ (number between 0 and 1)

the act of tossing the coin at a particular time in a particular surrounding in a particular manner picks (samples) a point from this space Ω [illegible]

$\mathcal{F}$ is a set of all $A \subset \Omega$ to which we want to assign a probability $P(A)$

~~[illegible]~~ desirable ($\mathcal{F}$ is collection of the subsets of Ω, called ~~[illegible]~~ events)

$A \in \mathcal{F}$, 'event' (subset of the sample space)

If an event A is of the type $A = \{\omega \in \Omega \mid R(\omega)\}$ for some property $R(\cdot)$, we may write $P(R)$ for $P(A)$.

$= \{$points of Ω that lead to head$\}$ or $A^c = \{$points of Ω that lead to tail$\}$

desired properties of $\mathcal{F}$:

(i) $\Omega \in \mathcal{F}$

$\phi \in \mathcal{F}$

(ii) $A, B \in \mathcal{F}$ 'A or B'

'A and B'.

$A \cup B \in \mathcal{F}$, $A \cap B \in \mathcal{F}$.

More generally, $A_1, A_2 \cdots \in \mathcal{F} \Rightarrow \bigcap_i A_i$, $\bigcup_i A_i \in \mathcal{F}$

(iii) $A \in \mathcal{F} \Rightarrow$ 'not A' $\in \mathcal{F}$, i.e., $A^c \in \mathcal{F}$

Remarks: (i) $\bigcap_i A_i \in \mathcal{F}$,

$A^c \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$.

(ii) can consider weaker requirement

$A_1, \cdots A_n \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}$.

$\bigcap_{i=1}^{n} A_i \in \mathcal{F}$.

closed under union & intersection

$\mathcal{F}$ is ~~**~~ said to be a σ-algebra

$= \sigma$-field

usually denotes some countable operation

$\mathcal{F}$ contains ϕ, Ω & is closed under complementation & countable unions and intersections)

we often need not consider all subsets.

P: 'probability of ...' $\quad P : \mathcal{F} \to [0,1]$.

$P(A)$: probability of A.

Desired properties of P:

(i) $P(\Omega) = 1$, $P(\phi) = 0$.

(ii) $P(A^c) = 1 - P(A)$

(iii) $A_1, A_2, \cdots A_n$ disjoint,

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

Countable additivity.

to each event $A \in \mathcal{F}$, is assigned a number $\in [0,1]$, called its probability denoted by $P(A)$. we expect that $\mathcal{F}$ and $P(\cdot)$ as a map from $\mathcal{F}$ to $[0,1]$ should satisfy certain requirements - must be σ-field

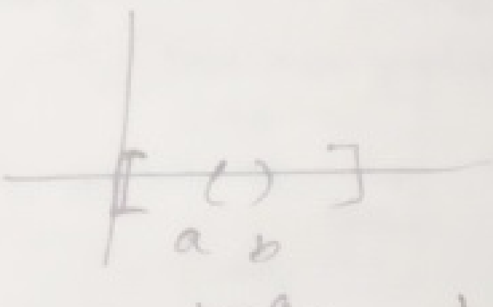
Remark: If we require only that

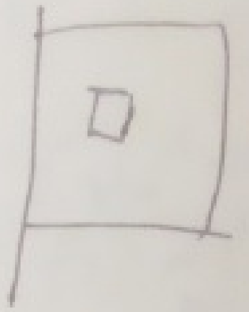
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_i P(A_i), \quad \{A_i\} \text{ disjoint}, \text{ finite additivity}$$

$(\Omega, \mathcal{F}, P)$ probability space

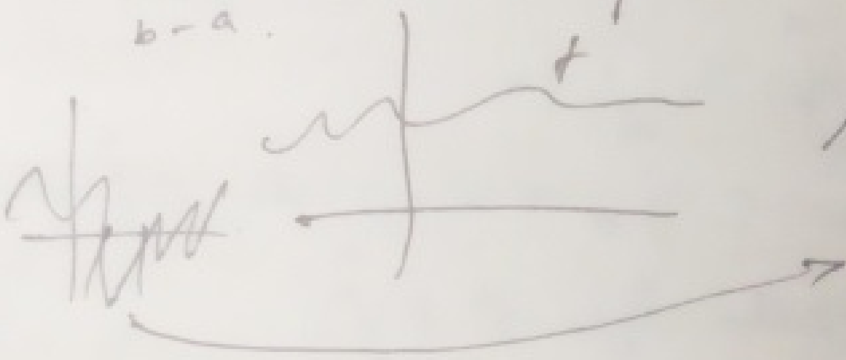
(i) – (iii) defines P as a positive measure on $(\Omega, \mathcal{F})$ with total mass 1.

'measure' length, area, volume.





examples of measures.



$$\mu(A) = \int_A f(x)\,dx.$$

negative measure

signed measure.

Ω, $\mathcal{F}$, P.

$\Omega = [0,1]$ $\mathcal{F} = ?$

Require: $\mathcal{F}$ containing all intervals.

Given $\mathcal{F}_0$, a collections of subsets of Ω, we can define $\sigma(\mathcal{F}_0)$ = the smallest σ-algebra containing $\mathcal{F}_0$

= the intersection of σ- algebras containing f_0

Need (i) $\exists$ a σ algebra containing f_0 —

take all subsets of Ω

(ii) observe: Arbitrary intersections of σ- algebras are σ- algebras (Unions are not).

Let f_α, $\alpha \in I$, be σ- algebras on Ω

$$\hat{f} = \bigcap_\alpha f_\alpha .$$

(i) $\Omega \in f_\alpha \; \forall \alpha \Rightarrow \Omega \in \hat{f}$,

similarly, $\phi \in \hat{f}$

(ii) $A \in \hat{f} \Rightarrow A \in f_\alpha \; \forall \alpha \Rightarrow A^c \in f_\alpha \; \forall \alpha$

$\Rightarrow A^c \in \hat{f}$

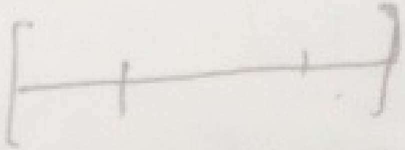
(iii) $A_1, A_2 \cdots \in \hat{f} \Rightarrow A_i \in f_\alpha \; \forall \alpha$,

$\Rightarrow \bigcup_i A_i \in f_\alpha \; \forall \alpha \Rightarrow$

$\bigcup_i A_i \in \hat{f}$.

$\Rightarrow$ the two definations are equivalent.

$\sigma(f_0)$ = smallest σ-algebra containing f_0 = intersection of σ-algebras containing f_0



$$f_0 = \{A_1, A_2 \ldots A_6\}$$

$$f_0 = \{A_1, A_2, \ldots, A_6\}.$$

$$\sigma(f_0) = \{\phi, \Omega, A_1, \ldots A_6, A_1 \cup A_2 \ldots, A_1 \cup A_2 \cup A_3 \ldots\}.$$

S, $d(\cdot\cdot, \cdot\cdot)$ metric. is.

$$d(x,y) \geq 0.$$

$$d(x,z) \leq d(x,y) + d(z,y)$$

$$d(x,y) = 0 \text{ iff } x = y$$

$$\{x \mid d(x, x_0) < r\}.$$

Ball of radius r, centre x_0

Borel σ-algebras. (for open sets)

$\mathbb{R}$ | $M((a,b)) = b-a$

Ω

Example of a non-measurable set

$\mathcal{F}_0$ algebra $\longrightarrow$ $\sigma(\mathcal{F}_0)$ σ algebra

P on $(\Omega, \mathcal{F}_0) \rightarrow P$ on $(\Omega, \mathcal{F})$

need extension theorems.

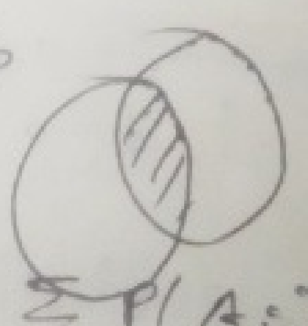
'Completion' $(\Omega, \mathcal{F}, P)$ ~~let~~ $\cancel{\hat{\mathcal{F}}}$

Let $\hat{\mathcal{F}} = \mathcal{F} \cup \{$arbitrary subsets of sets $A \in \mathcal{F}$ s.t. $P(A) = 0\}$

additional properties of P

(iii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

~~additional prop~~



(iv) $P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i)$

(iv) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Example of completion R.

$m((a,b)) = b - a$

Lebesgue σ-algebra

$P(\phi) = 0$

So, A is null event if $P(A) = 0$

sure event if $P(A) = 1$.

Statement S is said to be true almost surely if $P(S\text{ true}) = 1$

a.s.

$X = Y$

$r_1, r_2, \ldots$

$r_n \quad \frac{\epsilon}{2^n}$

Sum up ~~to~~ $< \epsilon$

if X is a set, a nonempty class $\mathcal{R}$ of subsets of X is called a ring if it is closed under the formation of set theoretic differences & finite unions. If moreover $X \in \mathcal{R}$, then $\mathcal{R}$ is called an algebra. A nonempty subclass of $\mathcal{P}(X)$ is a ring in the earlier sense iff it is a subring of $\mathcal{P}(X)$ in algebric sense

The class $\mathcal{P}(X)$ of X (power set) is a ring in usual algebric notation if. $\left.\begin{array}{l} + \equiv \Delta \\ \cdot \equiv \cap \end{array}\right\}$ defined requirement is that it is a σ-algebra

$(\Omega, \mathcal{F}, P)$ It's assumed that σ-algebra is Borel algebra

Borel algebra requires the notion of neighbourhood.

$(\Omega, \mathcal{F})$ — measurable space

$(\Omega, \mathcal{F}, P) \xrightarrow{\quad X \quad} E = \{H, T\}$

= a set

Desired properties of X:
want to assign prob. to sets

$\{\omega \mid X(\omega) \in A\}, A \subseteq E$

$\{\omega \in \Omega : X(\omega) \in A\}$

Range, not domain

Desired properties of ξ_s = < subsets $A \subset E$ under consideration >

(i) $E \in \xi_s$, $\phi \in \xi_s$

(ii) $A \in \xi_s \Rightarrow A^c \in \xi_s$

(iii) $A_i \in \xi_s \; \forall i \Rightarrow \bigcup_i A_i, \bigcap_i A_i \in \xi_s$

Let $(E, \mathcal{E})$ be another measurable space and $X: (\Omega, \mathcal{F}, P) \to (E, \mathcal{E})$ a r.v. (i.e. a measurable map)

$\mathcal{E}$ should be a σ-algebra

Want $\{w : X(w) \in A\}$ has a prob.

i.e., $\{w : X(w) \in A\} \in \mathcal{F}$.

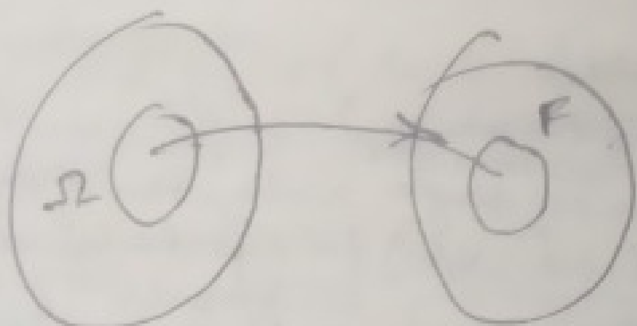
$\Omega = [0, 1]$ $\quad A \notin \mathcal{F}$

$X(w) = 1$ on A

$\quad\quad\;\; 0$ on A^c

$\{w : X(w) \in (-.5, .5)\} = A$, we can't assign

$X^{-1}(A) = \{w : X(w) \in A\}$.



Require: $\forall A \in \mathcal{E}$, $X^{-1}(A) \in \mathcal{F}$.

$X^{-1}(A) \in \mathcal{F} \;\; \forall A \in \mathcal{E}$.

$\rightarrow$ f is 'measurable'.

$X : (\Omega, \mathcal{F}, P) \longrightarrow (E, \mathcal{E})$

X is said to be an E-valued random variable if $X^{-1}(A) \triangleq \{\omega : X(\omega) \in A\} \in \mathcal{F}$ $\forall A \in \Xi\}$

Can define a prob. measure μ on (E, Ξ) by: $\mu(A)$

$$= P(\langle \omega : X(\omega) \in A \rangle)$$

$$= P(X \in A).$$

μ = the law of X

(i) [Principle of ~~...~~ (equally likely outcomes) insufficient reason.

(ii) Physical reasoning.

(iii) worst case - max. entropy

since one observes a single realization of a r.v. X, i.e. $X(\omega)$ for a particular $\omega \in \Omega$ the probability space $(\Omega, \mathcal{F}, P)$ is a hypothetical entity and its choice is by no means unique.

$$\Omega = \langle a_1 \cdots a_n \rangle$$

$$\int P(x) \log(P(x)) \, dx$$

E. T. Jaynes (Non-Baysian)

max. ~~entrop interpretation~~ entropy method

The image μ of P under X is a probability measure on $(E, \mathcal{E})$, called the law of X and denoted by $\mathcal{L}(X)$.

The events $\{\omega \mid X(\omega) \in A\}$ for $A \in \mathcal{E}$ form a sub-σ-field of $\mathcal{F}$ called the σ-field generated by X and denoted by $\sigma(X)$. More generally, given a family $X_\alpha, \alpha \in I$, of random variables on $(\Omega, \mathcal{F}, P)$ taking values in measurable spaces $(E_\alpha, \mathcal{E}_\alpha)$, $\alpha \in I$, respectively, the σ-field generated by $X_\alpha, \alpha \in I$, denoted by $\sigma(X_\alpha, \alpha \in I)$, is the smallest sub-$\sigma$ field w.r.t. which they are all measurable.

Now an experiment such as tossing a coin or rolling a die picks a point from the sample space and maps it to another space E. ($E = \{$head, tail$\}$ and $\{1, 2, \ldots, 6\}$ resp.) Thus it's a map $X : \Omega \to E$. Since our idea is to have sets of the type $\{\omega \mid X(\omega) \in B\}$ to be events (i.e. elements of $\mathcal{F}$) for a suitable collection of subsets $B \subset E$, considerations analogous to the above for $\mathcal{F}$ suggest that we equip E with a σ-field $\mathcal{E}$ and require a map $X : (\Omega, F) \to (E, \mathcal{E})$ to be measurable. Such a map is called E-valued r.v.