Probability Theroy based on Measure Theory

STAT 517

Dr. John Smith

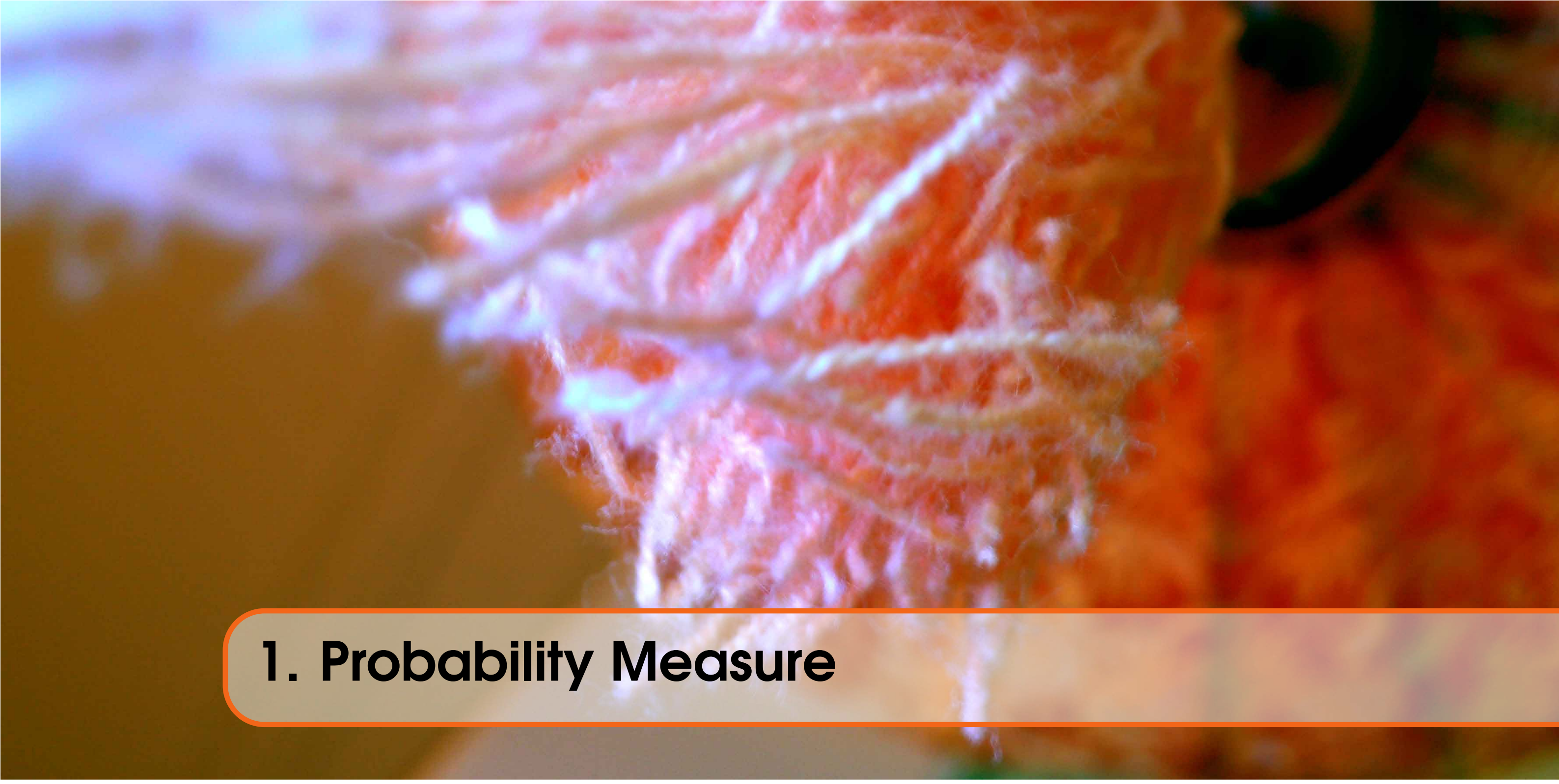
PUBLISHED BY PUBLISHER

BOOK-WEBSITE.COM

First printing, March 2013

Contents

Part One



1. Probability Measure

1.1 Probability on a Field

Definition 1.1.1 — Ω. Non emtpy set.

Definition 1.1.2 — Paving. A collection of a subset of Ω is a paving.

Definition 1.1.3 — Field. A **field** $\mathscr{F}$ is a paving satisfying

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

(ii) $A \in \mathscr{F} \Rightarrow A^C \in \mathscr{F}$

(iii) $A, B, \in \mathscr{F}, \Rightarrow A \cup B \in \mathscr{F}$

Derived Properties about a Field

- $\emptyset \in \mathscr{F}$ (by (i) and (ii):

$$\begin{aligned} \Omega \in \mathscr{F} &\Rightarrow \Omega^C \in \mathscr{F} \\ &\Rightarrow \emptyset \in \mathscr{F}) \end{aligned}$$

- (i) can be replaced by "$\mathscr{R}$ is nonempty" because,
  Let $A \in \mathscr{F}$,

$$\begin{aligned} &\Rightarrow A^c \in \mathscr{F} \\ &\Rightarrow A^C \cup A \in \mathscr{F} \\ &\Rightarrow \Omega \in \mathscr{F} \end{aligned}$$

- $A \in \mathscr{F}, B \in \mathscr{F}, \Rightarrow, A \cap B \in \mathscr{F}$ because,

$$\begin{aligned} &(A \cap B)^C = A^C \cup B^C (DeMorgan's Law) \\ &A \cap B = (A^C \cup B^C)^C \end{aligned}$$

- $A_1, \ldots, A_m \in \mathscr{F} \Rightarrow A_1 \cup, \ldots, \cup A_m \in \mathscr{F}$ (mathematical induction)
- $A_1, \ldots, A_m \in \mathscr{F} \Rightarrow A_1 \cap, \ldots, \cap A_m \in \mathscr{F}$

Definition 1.1.4 — σ-Field. Similar to the definition of a field except for (iii). A paving satisfying

(i) $\Omega \in \mathscr{F}$
(ii) $A \in \mathscr{F} \Rightarrow A^C \in \mathscr{F}$
(iii) $A_1 \in \mathscr{F}, \ldots, A_m \in \mathscr{F}$
$\cup_{k=1}^{m} A_k \in \mathscr{F}$ (finite additvity)

If we replace (iii) from before by (iii') here:
For $A_1 \in \mathscr{F}, \ldots, A_m \in \mathscr{F}$

$$\cup_{k=1}^{\infty} A_k \in \mathscr{F}$$

then $\mathscr{F}$ is called a **σ-field**.

Derived Facts

- Again, (i) can be repalced by $\mathscr{F}$ no empty, (iii) can be replaced $A_1 \in \mathscr{F}, \ldots, A_m \in \mathscr{F}$

Example 1.1 $\Omega = (0, 1]$ (from now on all intervals are left open, right closed)

R Recall that σ-fields are generated by fields. Fancy scripts denote a σ-field. Fancy scripts with a zero subscript denote a field.

$\mathscr{B}_0$ is the collection of all finite union of disjoint intervals. Aside: something that we can consider easily, but computers cannot. (e.g. think of all stars, etc)

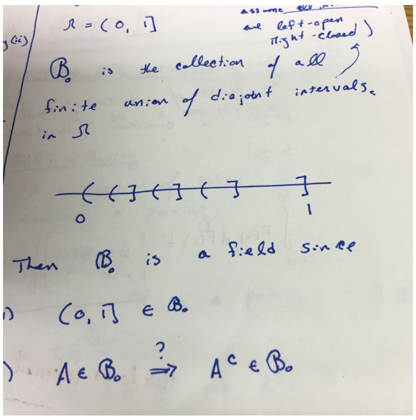


Figure 1.1: Finite unioin of three disjoint intervals.

Then $\mathscr{B}_0$ is a field.
(i) $(0, 1] \in \mathscr{B}_0$
(ii) $A \in \mathscr{B}_0 \Rightarrow A^c \in \mathscr{B}_0$
(iii) $A \in \mathscr{B}_o, B \in \mathscr{B}_o \Rightarrow A \cup B \in \mathscr{B}_o$

■

Wednesday August 24
$\mathscr{B}_0 =$ collection of finite unions of disjoin subintervals of (0, 1]. Is a field.

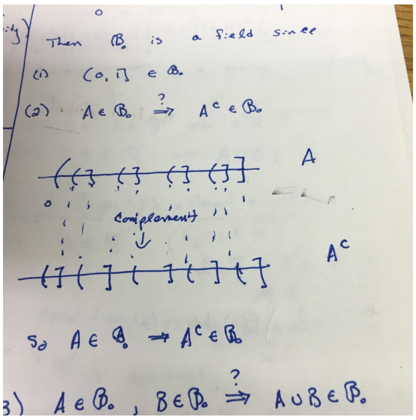


Figure 1.2: A and complement of A.

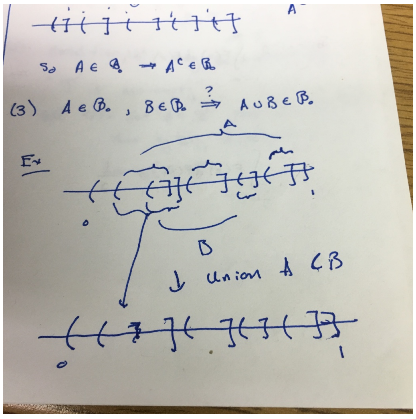


Figure 1.3: Union of A and B is still in $\mathscr{B}_o$

Definition 1.1.5 — Power Set. A σ-field is generated by a paving of power set. Let Ω be a set. The collection of all subsets of Ω is the power set written as 2^{Ω}.

R Where does this notation come from?
Consider ths case where Ω is finite

$$\Omega = \{\omega_1, \dots, \omega_n\}$$

Total number of subsets of Ω.
$\emptyset$, 1 element sets, 2-element sets, ..., n-element ests.

$$() + () + \cdots + = (1+1)^n$$

$\#(\mathscr{F}) = 2^{\#\Omega}$, so it seems reasonable to denote $\mathscr{F} = 2^{\Omega}$.
It is also easy to show that 2^{Ω} is a σ-field. (The largest, even. The smallest: $\{\emptyset, \Omega\}$ which is also a σ-field.)

$$\{\emptyset, \Omega\} \subseteq \sigma\text{-field} \subseteq 2^{\Omega}$$

It turns out we can extend notion of lenght from $\mathscr{B}_0$ to σ-field generated by $\mathscr{B}_o$.

Now, let $\mathscr{A}$ be a nonempty paving of Ω. We define

$$\sigma(\mathscr{A}) = \cap\{\mathscr{B} \subseteq 2^{\Omega} : \mathscr{B} \text{ is a } \sigma\text{-field}, \mathscr{A} \subseteq \mathscr{B}\}$$

OR rather, the *intersection* of all σ-fields that contains $\mathscr{A}$.

Let

$$\mathbb{F}(\mathscr{A}) = \{\mathscr{B} \subseteq 2^{\Omega} : \mathscr{B} \text{ is a } \sigma\text{-field}, \mathscr{B} \supseteq \mathscr{A}\}$$

Then,

$$\sigma(\mathscr{A}) = \cap\mathscr{B}$$

$$\mathscr{B} \in \mathbb{F}(\mathscr{A})$$

Derived Facts
$\mathbb{F}(\mathscr{A})$ is nonempty. For example, 2^{Ω} is a σ-field and $2^{\Omega} \supseteq \mathscr{A}$.
$\cap B$ is a σ-field. ($B \in \mathbb{F}(\mathscr{A})$)

R Get notes about notation/levels.

Proof. We will prove that indeed $\sigma(\mathscr{A})$ is a σ -field. Recall that we have three conditions above for σ-field.

(i)

$$\Omega \in \sigma(\mathscr{A})$$

$$\Omega \in \cap_{B \in \mathbb{F}(\mathscr{A})} B$$

Because: B is σ-field, $A \in B$, $\forall B \in \mathbb{F}(\mathscr{A})$.

(ii)

(iii) $A_1, \ldots, \in \cap_{B \in \mathbb{F}(\mathscr{A})} B, \forall B \in \mathbb{F}(\mathscr{A})$
$\Rightarrow \cup_{n=1}^{\infty} A_n \in B, \forall B \in \mathbb{F}(\mathscr{A})$

So, $\sigma(\mathscr{A}$ is a σ-field, we call it the σ-field, generated by $\mathscr{B}_o$. We know how tot assign lenth to members of $\mathscr{B}_o$, we now show the assignment can be extended to $\sigma(\mathscr{B}_o)$

■

■ **Example 1.2** Let $\mathscr{I}$ be the collection of *all* subintervals of (0,1].

Note that $\mathscr{I}$ is a smaller collection than $\mathscr{B}_0$ since $\mathscr{B}_0$ can have numerous different combinations of the sets.

Let

$$\mathscr{B} = \sigma(\mathscr{I})$$

.

This is a Borel-σ-field. (a member of $\mathscr{B}$ in Borel set.)
It turns out

$$\sigma(\mathscr{I}) = \sigma(\mathscr{B}_o)$$

This is because $\sigma(\mathscr{I})$ is a σ-field.

So,

$$\sigma(\mathscr{I}) \supseteq \mathscr{B}_o$$
$$\sigma(\mathscr{I}) \supseteq \sigma(\mathscr{B}_o)$$

Also,

$$\mathscr{I} \subseteq \mathscr{B}_o$$
$$\sigma(\mathscr{I}) \subseteq \sigma(\mathscr{B}_o)$$

Thus,

$$\sigma(\mathscr{I}) = \sigma(\mathscr{B}_o)$$

■

Definition 1.1.6 — Probability Measure. Probablity measures on field. Suppose $\mathscr{F}$ is a field on a nonempty set Ω. A probability measure is a function $P : \mathscr{F} \to \mathbb{R}$.

(i) $0 \le P(A) \le 1, \forall A \in \mathscr{F}$

(ii) $P(\emptyset) = 0$, $P(\Omega) = 1$

(iii) If $A_1, \ldots$ are disjoint emembers of $\mathscr{F}$ and $\cup A_n \in \mathscr{F}$ then we have countable additivity:

$$P(\cup A_n) = \sum_{n=1}^{\infty} P(A_N)$$

R Note that (iii) also implies finite additivity. Prove by adding infinite empty sets on end.

If Ω is nonempty set. And $\mathscr{F}$ is a σ-field on Ω. And P is a probability measure on $\mathscr{F}$.
Then $(\Omega, \mathscr{F}, P)$ is called a **probability space.**
And $(\Omega, \mathscr{F})$ is called a **measurable space.**

R If $A \subseteq B$, then $P(A) \le P(B)$. This is because we may write B as

$$B = A \cup (B \setminus A)$$

R

$$P(A) + P(B) = P(A \cup B) + P(A \cap B)$$

Friday August 26

Recall,
Probability measure on a field, $\mathscr{F}_0$.

- $P(A) + P(B) = P(A \cup B) + P(A \cap B)$
  - $P(A) = P(AB^C) + P(AB)$
  - $P(B) = P(BA^C) + P(AB)$
  - $P(A) + P(B) = P(AB^C) + P(BA^C) + 2P(AB)$
  - $P(A \cup B) = P(AB^C) + P(BA^C) + P(AB)$

- $P(A\cup B)=P(A)+P(B)-P(AB)$ By induction, we can prove if $A_1,\dots A_n$,

$$P(\cup_{k=1}^{n}A_k)=\sum_{k=1}^{n}P(A_k)-\sum_{i<j}P(A_iA_j)+\sum_{i<j<k}A_iA_j)+\cdots+(-1)^{n+1}P(A_1,\dots,A_n)$$

  Inclusion- Exclusion Formula
- If $A_1,\dots A_n\in\mathscr{F}$,

$$B_1=A_1$$

$$B_2=A_2\setminus A_1$$

$$\vdots$$

  Then,

$$\cup_{k=1}^{n}A_k=\cup_{k=1}^{n}B_k$$

  but the B_i are disjoint. Also $A_K\subseteq B_k\forall k=1,\dots,n$.

$$P(\cup_{k=1}^{n}A_k)=P(\cup_{k=1}^{n}B_k)=\sum_{k=1}^{n}B_k\leq\sum_{k=1}^{n}A_k$$

  Thus, $P(\cup_{k=1}^{n}A_k)\leq\sum_{k=1}^{n}A_k$. Finite subadditivity.

Some conventions,
If $A_1,\dots$ is a sequence of sets, we say $A_n\uparrow A$ if

1. $A_1\subseteq A_2\subseteq\dots$
2. $\cup_{k=1}^{\infty}A_k=A$

If $A_1,\dots$ is a sequence of sets, we say $A_n\downarrow A$ if

1. $A_1\supseteq A_2\supseteq\dots$
2. $\cap_{k=1}^{\infty}A_k=A$

Theorem 1.1.1 If P is a probability measure on a field $\mathscr{F}$ Then,

1. Continuity from below.
   If $A_n\in\mathscr{F}\forall n, A\in\mathscr{F}$

$$A_n\uparrow A$$

   then

$$P(A_n)\uparrow P(A)$$

2. Continuity from above.
   If $A_n\in\mathscr{F}\forall n.A\in\mathscr{F}A_n\downarrow A$, then $P(A_n)\downarrow P(A)$
3. Countable subadditivity.
   If $A_n\in\mathscr{F}\forall n.\cup_{k=1}^{\infty}A_k\in\mathscr{F}$ then

$$P(\cup_{n=1}^{\infty}A_k)\leq\sum_{n=1}^{\infty}P(A_k)$$

Proof. 1. If $A_1,\dots A_n\in\mathscr{F}$,

$$B_1=A_1$$

$$B_2=A_2\setminus A_1$$

$$B_3 = A_3 \setminus A_2$$

$$\vdots$$

then, $B_1, \ldots$ are disjoint.

$$\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n$$

$P(A) = P(\cup_{n=1}^{\infty} A_n) = P(displaystyle \cup_{n=1}^{\infty} B_n) = displaystyle \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} displaystyle \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} P(A_n)$

2. $A_n \downarrow A \Leftrightarrow A_n^C \uparrow A^C$
   But by (1),
   $P(A_n^C) \uparrow P(A^C)$ $1 - P(A_n) \uparrow 1 - P(A)$ $P(A_n) \downarrow P(A)$
3. By finite subadditivity,

$$P(\cup^n k = 1 A_k) \leq displaystyle \sum^n k = 1 P(A_k) \leq displaystyle \sum_{n=1}^{\infty} P(A_n)$$

But since, by (1), because
$\cup_{k=1}^{n} A_k \uparrow \cup_{n=1}^{\infty} A_n$
$P(\cup_{k=1}^{n} A_k) \uparrow P(\cup_{n=1}^{\infty} A_n)$
S0,
$P(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$

■

R $A \in \mathscr{F} =$ "A is F-set".

1.2 Extention of Probability Measure to a σ-field

Let f be a function $f : D \to R$.
Let $\tilde{D}$ be another set such that

$$D \subseteq \tilde{D}$$

An extantion of f onto $\tilde{D}$ is

$$\tilde{f} : \tilde{D} \to R$$

Such that $f(x) = \tilde{f}(x) \forall x \in D$
$\tilde{f}$ is an extention of f on D.
We say f has unique extention, $\tilde{f}$ onto $\tilde{D}$ if

1. $\tilde{f}$ is an extension of f to $\tilde{D}$.
2. if g is another extension of f to $\tilde{D}$ then $\tilde{f} = g$ on D.

Theorem 1.2.1 A probability measure on a field has a unique extension on the σ-field generated by this field.
Means: $\mathscr{F}_0$ is a field
P is a probability measure on $\mathscr{F}_0$
Then there exists a probability measure, Q on $\sigma(\mathscr{F})$ such that $Q(A) = P(A) \forall A \in \mathscr{F}_0$

Moreover, if $\tilde{Q}$ is another probability measure on $\sigma(\mathscr{F}_0)$ such that $\tilde{Q} = P(A) \forall A \in \mathscr{F}$ then $\tilde{Q} = Q$.

Outer Measure $P^* : 2^\Omega \to \mathbb{R}$
For any $A \in 2^\Omega$ $(A \subseteq \Omega)$

$$P^*(A) = \inf\{\sum_{n=1}^{\infty} P(A_n) : A_1, A_2, \ldots \text{is a sequence of } \mathscr{F}_0 \text{ sets, } A \subseteq \cup_{n=1}^{\infty} A_n\}$$

Inner Measure
$P_*(A) = 1 - P^*(A)$

Define the paving $\mathscr{M}$ as followes

$$\mathscr{M} = \{A \in 2^\Omega : E \in 2^\Omega, P^*(E) = P^*(E \cap A) + P^*(E \cap A^C)\}$$

Monday August 29

P^* satisfies the following probabilities:
(i) $P^*(\emptyset) = 0$
(ii) $P^*(A) \geq 0 \quad \forall A \in 2^\Omega$
(iii) $A \subseteq B \Rightarrow P^*(A) \subseteq P^*(B)$
(iv) $P^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P^*(A_n))$

Proof. (i) Take $\{\emptyset, \emptyset, \ldots\}$.

$$\emptyset \in \mathscr{F}_0, \quad \emptyset \cup_{n=1}^{\infty} \emptyset$$

So,

$$P^*(\emptyset) \leq \sum_{n=1}^{\infty} P(\emptyset) = 0$$

Note,

$$P(A) \geq 0 \quad \forall A$$

So,

$$P^*(\emptyset) \geq \emptyset$$

Thus,

$$P^*(\emptyset) = \emptyset$$

(ii) Already done as part of (i).
(iii) Let $A \subseteq B$

$$P^*(A) = \inf\{\sum_{n=1}^{\infty} P(A_n), A_n \in \mathscr{F}_0, A \subseteq \cup A_n\}$$

Now, if $B_1, \cdots \in \mathscr{F}_0 \subseteq \cup B_n$
Then,

$$A \subseteq B \subseteq \cup_n B_n$$

If $\{\{B_n\}_{n=1}^{\infty} : B_n \in \mathscr{F}_0, B \subseteq \cup_n B_n\} \subseteq \{\{A_n\}_{n=1}^{\infty} : A_n \in \mathscr{F}_0, A \subseteq \cup_n A_n\}$
Or in short, Collection 1 $\subseteq$ Collection 2.

So the inf of a larger set is smaller than (or equal to) the inf of a smaller set.

So,
$P^*(A) = \inf\{\sum_{n=1}^{\infty} P(A_n), A_n \in \text{ collection \#1}\} \leq P^*(B) = \inf\{\sum_{n=1}^{\infty} P(B_n), A_n \in \text{ collection \#2}\} = P^*(B)$

(iv) Want

$$P^*(\cup_n A_n) \leq \sum_n P^*(A_n)$$

$P^*(A_n) = \inf\{\sum_{n=1}^{\infty} P(A_n) : A_{nk} \in \mathscr{F}_0, A \subseteq \cup_k A_{nk}\}$

Let $\varepsilon > 0$, by defnition of there exists,

$$\{B_n\}_{n=1}^{\infty}$$

such that

$$\sum_{k=1}^{\infty} P(B_{nk}) \leq P^*(A_n) + \frac{\varepsilon}{2^n}$$

So,

$$\cup_n A_n \subseteq \cup_{n,k} B_{nk}$$

and,

$$\begin{aligned} P^*(\cup_n A_n) &\leq \sum_{n,k} P(B_{nk}) \\ &< \sum_n P^*(A_n) + \sum_n (\varepsilon 2^{-n}) \\ P^*(\cup A_n) &< \sum_n P^*(A_n) + \varepsilon \quad \forall \varepsilon > 0 \end{aligned}$$

Simply put,
b
So,

$$P^*(\cup_n A_n) \leq \sum_n P^*(A_n)$$

■

By definition, $A \in \mathscr{M}$ if and only if $P^*(EA) + P^*(EA^C) = P^*(E)$.

We know that P^* is subadditive.

So, by subadditivity we know,

$$P^*(E) \leq P^*(AE) + P^*(A^C E)$$

Therefore, to show $A \in \mathscr{M}$ we only need to show

$$P^*(E) \geq P^*(AE) + P^*(A^C E)$$

$\mathscr{M}$ is defined by P^* and P^* is defined using $\mathscr{F}_0$ so $\mathscr{M}$ is indireclty tied to $\mathscr{F}_0$.

Lemma 1. $\mathscr{M}$ is a field.

Proof. (i) $\Omega \in \mathscr{M}$

$$\begin{aligned} A &= \Omega \\ P^*(\emptyset) &= 0 \\ P^*(E) + P^*(\emptyset) &= P^*(E) \end{aligned}$$

(ii) $A \in \mathscr{M} = A^C \in \mathscr{M}$

$$\begin{aligned} P^*(E) &= P^*(EA) + P^*(A^C E) \\ &= P^*(EA^C) + P^*(AE) \\ &= P^*(EA^C) + P^*((A^C)^C E) \end{aligned}$$

(iii) $A, B \in \mathscr{M} \to A \cap B \in \mathscr{M}$

$$\begin{aligned} &B \in \mathscr{M} \Rightarrow P^*(E) = P^*(Eb) + P^*(B^C E) \quad \forall E \\ &A \in \mathscr{M} \Rightarrow P^*(BE) = P^*((BE)A) + P^*(A^C(BE)) \\ &A \in \mathscr{M} \Rightarrow P^*(B^C E) = P^*((B^C E)A) + P^*(A^C(B^C E)) \end{aligned}$$

Hence,

$$P^*(BE) + P^*(B^C E) = P^*((BE)A) + P^*(A^C(BE)) + P^*((B^C E)A) + P^*(A^C(B^C E))$$

$$\begin{aligned} P^*(A^C(BE)) + P^*((B^C E)A) + P^*(A^C(B^C E)) &\geq P^*((A^C BE) \cup (AB^C E) \cup (A^C BE)) \\ &= P^*(E \cap [A^C B \cup AB^C \cup A^C B^C]) \\ &= P^*(E \cap (AB)^C) \end{aligned}$$

$$\begin{aligned} P^*(E) &= P^*(BE) + P^*(B^C E) \\ &= P^*((BE)A) + (P^*(A^C(BE)) + P^*((B^C E)A) + P^*(A^C(B^C E))) \\ &\geq P^*(ABE) + P^*(E(AB)^C) \end{aligned}$$

So, $A, B \in \mathscr{M}$

∎

Lemma 2. If $A_1, A_2, \ldots$ is a sequence of disjoint $\mathscr{M}$-sets then for each $E \subseteq \Omega$,

$$P^*(E \cap (\cup_k A_k)) = \sum_k P^*(E \cap A_k)$$

Proof. First, prove this statement for finite sequence.

$$A_1, \ldots, A_n$$

by mathematical induction.

If $n = 1$ this is 'trivial',

$$P^*(E \cap A_1) = P^*(E \cap A_1)$$

If $n = 2$ we need to show,

$$P^*(E(A_1 \cup A_n)) = P^*(EA_1) + P^*(EA_2)$$

Because $A_1 \in \mathscr{M}$,

$$P^*(E(A_1 \cup A_2)) = P^*(E(A_1 \cup A_2))A_1 + P^*(E(A_1 \cup A_2)A_1^2)$$

$$E(A_1 \cup A_2) = E(A_1A_2 \cup A_1A_2 = EA_1$$

$$E(A_1 \cup A_2)A_1^C = E(A_1A_1^C \cup A_2A_2^C)$$

So,
$P^*(E(A_1 \cup A_2)) = P^*(EA_1) + P^*(EA_2)$

■

2. General Measure

3. Integration with Respect to a Measure

4. Random Variable

5. Convergence in Probability/Limit Theorem

6. Radon-Nikodym Derivative Theorem

7. Special Topics

Index