Ch 4 Abstract Theory of the Integral II

Fubini – Tonelli

(4.1) As a preparation for defining measures on product spaces, we discuss outer measures on an arbitrary space X

Defn. Let X be a set and 2^X the set of all subsets of X.

A set function $\mu^*: 2^X \to \mathbb{R}_+$ is called an outer measure on X if

(i) $\mu^*(\emptyset) = 0$

(ii) μ^* is monotone (i.e. $A \subseteq B \Rightarrow \mu^*(A) \le \mu^*(B)$)

(iii) μ^* is countably subadditive i.e. if $\{A_i\}$ is a sequence of sets

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} \mu^*(A_i)$$

Given an outer measure μ^* and $A \subseteq X$, we say that A is μ^*-measurable if

$$\forall E \subseteq X \quad \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Proposition Let μ^* be an outer measure on X and $\mathfrak{A}$ the class of μ^*-msble sets.

Then $\mathfrak{A}$ is a σ-algebra

The proof is entirely similar to our procedure for Lebesgue integration.

Left to the enthusiastic student !!

<u>Defn</u>. Let $\mathcal{C}$ be a collection of subsets of X s.t. $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$ and $A \in \mathcal{C} \Rightarrow A^c$ is a finite disjoint union of sets in $\mathcal{C}$. We call $\mathcal{C}$ a <u>semi-algebra</u>.

[N.B an algebra $\mathcal{O}\!\mathcal{L}$ of sets is defined by $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{O}\!\mathcal{L}$ and $A \cup B \in \mathcal{O}\!\mathcal{L}$ and $A^c \in \mathcal{O}\!\mathcal{L}$]

<u>Proposition</u> Let $\mathcal{C}$ be a semi-algebra. Then

(i) $\exists$ a smallest algebra $\mathcal{O}\!\mathcal{L}$ containing $\mathcal{C}$ ~~such that $\mu(\phi)=0$ if $\phi \in \mathcal{C}$~~

(This is called the algebra <u>generated</u> by $\mathcal{C}$)

(ii) S'pose μ is a non-negative set function monotone on $\mathcal{C}$ such that $\mu(\phi) = 0$ if $\phi \in \mathcal{C}$

The following conditions are equivalent:

(α) μ has a unique extension to a countably additive set function $\bar{\mu}$ on $\mathcal{O}\!\mathcal{L}$

(β) If $C \in \mathcal{C}$ is the union of a <u>finite</u> disjoint collection $\{C_i\}_{i=1}^n \subseteq \mathcal{C}$ then $\mu(C) = \sum_{i=1}^n \mu(C_i)$

(γ) If $C \in \mathcal{C}$ is the union of a <u>countable</u> disjoint collection $\{C_i\}_{i=1}^\infty \subseteq \mathcal{C}$ then $\mu(C) = \sum_{i=1}^\infty \mu(C_i)$

Proof. (i) The algebra $\mathcal{Q}$ is defined by adding to $\mathcal{C}$ all finite disjoint unions of sets in $\mathcal{C}$.
It requires a little work to show that this defines an algebra of sets.

(ii) $(\alpha) \Rightarrow (\gamma) \Rightarrow (\beta)$ is clear from the definition

$(\beta) \Rightarrow (\alpha)$ Now, if $A = \bigcup_{i=1}^{n} E_i$ where $E_i \in \mathcal{C}$ are disjoint, we define

$$\bar{\mu}(A) = \sum_{i=1}^{n} \mu(E_i)$$. We need to check that this defines $\bar{\mu}$ uniquely.

Suppose that $\{C_i\}_{i=1}^{n}$, $\{D_j\}_{j=1}^{m}$ are two finite disjoint collections of elements of $\mathcal{C}$ such that $A = \bigcup_i C_i = \bigcup_j D_j$

Note that $C_i = \bigcup_{j=1}^{m} C_i \cap D_j$ is a disjoint union of elements of $\mathcal{C}$, so by (β)

$$\mu(C_i) = \sum_{j=1}^{m} \mu(C_i \cap D_j)$$

Thus, by defin $\sum_{i=1}^{n} \mu(C_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \mu(C_i \cap D_j)$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} \mu(C_i \cap D_j) = \sum_{j=1}^{m} \mu(D_j)$$

Thus $\bar{\mu}$ is well-defined on the algebra $\mathcal{Q}$

Furthermore $\bar{\mu}$ is clearly monotone since μ is.
We now must show that $\bar{\mu}$ is countably additive

This follows from the monotone condition

If $\{C_i\}_{i=1}^{\infty}$ is a countable disjoint collection and $C = \bigcup_{i=1}^{\infty} C_i \in \mathscr{C}$, then for all n

$$\bar{\mu}(C) = \bar{\mu}\left(\bigcup_{i=1}^{\infty} C_i\right) \geq \bar{\mu}\left(\bigcup_{i=1}^{n} C_i\right) = \sum_{i=1}^{n} \mu(C_i)$$

Thus $\bar{\mu}(C) \geq \sum_{i=1}^{\infty} \mu(C_i)$

But since $\bar{\mu}$ is montone $\bar{\mu}(C) \leq \sum_{i=1}^{\infty} \mu(C_i)$

Thus we have equality and $\bar{\mu}$ is countably additive on $\mathcal{O}\!\mathcal{L}$

<u>Definition</u> Let $\mathcal{O}\!\mathcal{L}$ be an algebra of sets in X

A monotone set function μ on $\mathcal{O}\!\mathcal{L}$ s.t. $\mu(\phi) = 0$ is called a <u>premeasure</u> if $\forall$ disjoint sequence $\{A_i\}$ in $\mathcal{O}\!\mathcal{L}$ s.t. $\bigcup_{i=1}^{\infty} A_i \in \mathcal{O}\!\mathcal{L}$

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

Now suppose we have some algebra of sets in X equipped with a premeasure μ.

For any $E \in 2^X$, define

$$\mu^*(E) = \inf_{E \subseteq \bigcup_{i=1}^{\infty} A_i} \sum_{i=1}^{\infty} \mu(A_i)$$

<u>Proposition</u> μ^* is an outer measure on 2^X.

<u>Proof</u>

~~[illegible]~~ Ex

<u>Lemma</u> If μ is a premeasure on $\mathcal{Q}$ and $A \in \mathcal{Q}$ and A_i any sequence of sets in $\mathcal{Q}$ s.t. $A \subseteq \bigcup A_i$, then $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$

<u>Proof</u> Let $B_n = A \cap A_n \cap A_{n-1}^c \cap \dots \cap A_1^c$

Then $B_n \in \mathcal{Q}$ and $B_n \subseteq A_n$ and A is a disjoint union of $\{B_n\}_{n=1}^{\infty}$. Thus

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

<u>Proof of the Prop^n</u>

Clearly μ^* is non-negative and monotone and $\mu^*(\emptyset) = 0$

We have to show that μ^* is countably sub-additive

Let $E \subseteq \bigcup_{i=1}^{\infty} E_i$. If $\mu^*(E_i) = \infty$ for some i, then

$$\sum_{i=1}^{\infty} \mu^*(E_i) = \infty \geq \mu^*(E) \geq \mu^*(E_i) = \infty$$ so we are finished.

Thus we may suppose $\mu^*(E_i) < \infty$ for all i.

Let $\varepsilon > 0$ and, for each i, choose a sequence $\{A_{ij}\}_{j=1}^{\infty}$ of sets in $\mathcal{Q}$ s.t. $E_i \subseteq \bigcup_{j=1}^{\infty} A_{ij}$ and

$$\sum_{j=1}^{\infty} \mu(A_{ij}) \leq \mu^*(E_i) + \frac{\varepsilon}{2^i}.$$

Then $$\mu^*(E) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu(A_{ij}) \leq \sum_i \left(\mu^*(E_i) + \frac{\varepsilon}{2^i}\right)$$

$$= \sum_i \mu^*(E_i) + \varepsilon.$$

This holds for any ε, so $\mu^*(E) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$

This shows that μ^* is countably sub-additive and hence is an outer measure $\square$

Properties of this extension

If $A \in \mathcal{A}$ and A is measurable for μ^*, then

$$\mu^*(A) = \mu(A)$$

Proof. We want to show that for $E \in 2^X$

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

This is clear if $\mu^*(E) = \infty$, so let's assume $\mu^*(E) < \infty$

Then we have $E \subseteq \bigcup_{i=1}^{\infty} A_i$ for some sequence $\{A_i\}_{i=1}^{\infty}$ of disjoint sets from $\mathcal{A}$ s.t. $\sum \mu(A_i) \leq \mu^*(E) + \varepsilon$

Then $E \cap A \subseteq \bigcup_{i=1}^{\infty}(A_i \cap A)$ and $E \cap A^c \subseteq \bigcup_{i=1}^{\infty}(A_i \cap A^c)$ are both coverings by sets in $\mathcal{A}$. Thus

$$\begin{aligned}
\mu^*(E) &\leq \mu^*(E \cap A) + \mu^*(E \cap A^c) && \text{(subadditivity)} \\
&\leq \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \cap A^c) && \text{(defn of } \mu^*) \\
&= \sum_{i=1}^{\infty} \mu(A_i \cap A) + \mu(A_i \cap A^c) \\
&= \sum_{i=1}^{\infty} \mu(A_i) && \text{(finite additivity of } \mu) \\
&\leq \mu^*(E) + \varepsilon.
\end{aligned}$$

Since this holds for any $\varepsilon > 0$, it is true for $\varepsilon = 0$ $\square$

Definition If $\mathcal{A}$ is an algebra of sets, let

$\mathcal{A}_\sigma$ = the sets formed by taking countable unions of elements of $\mathcal{A}$

$\mathcal{A}_{\sigma\delta}$ = the sets formed by taking countable intersections of elements of $\mathcal{A}_\sigma$.

Proposition A Let μ be a premeasure on an algebra $\mathcal{O}\!\mathcal{l}$ and let μ^* be the induced outer measure.

For every $E \in 2^X$ and for every $\varepsilon > 0$ $\exists A \in \mathcal{O}\!\mathcal{l}_\sigma$ such that $E \subseteq A$ and $\mu^*(A) = \mu^*(E) + \varepsilon$.

Proof. By def^n of μ^* there is a sequence $\{A_i\}_{i=1}^{\infty}$ of sets in $\mathcal{O}\!\mathcal{l}$ s.t. $E \subset \bigcup_{i=1}^{\infty} A_i$ and

$$\sum_{i=1}^{\infty} \mu(A_i) \leq \mu^*(E) + \epsilon.$$

Clearly $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{O}\!\mathcal{l}_\sigma$ and $\mu^*(A) \leq \sum^{\infty} \mu(A_i)$
$\leq \mu^*(E) + \epsilon$

Proposition B Let μ be a premeasure on an algebra $\mathcal{O}\!\mathcal{l}$ and μ^* the induced outer measure on 2^X. For every set E with $\mu^*(E) < \infty$ there is $A \in \mathcal{O}\!\mathcal{l}_{\sigma\delta}$ with $E \subseteq A$ and $\mu^*(E) = \mu^*(A)$

Proof For all n, use Proposition A to get $A_n \in \mathcal{O}\!\mathcal{l}_\sigma$ st. $E \subseteq A_n$ and $\mu^*(E) \leq \mu^*(A_n) \leq \mu^*(E) + \frac{1}{n}$

Now let $A = \bigcap_{n=1}^{\infty} A_n$. Then $A \in \mathcal{O}\!\mathcal{l}_{\sigma\delta}$ and

$\mu^*(E) = \mu^*(A)$

$\square$

Theorem (Carathéodory Extension Theorem)

Let μ be a premeasure on the algebra $\mathcal{A}$ and μ^* the induced outer measure on 2^X.

Let $\mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A}$

(Note: Intersections of σ-algebras are σ-algebras!)

and μ' any extension of μ to $\mathcal{B}$

If $\mu^*(B) < \infty$ for a set $B \in \mathcal{B}$, then $\mu'(B) = \mu^*(B)$

If μ is a σ-finite measure, then the extension μ' to $\mathcal{B}$ is unique.

Proof. Let $A \in \mathcal{A}_\sigma$: it is a countable disjoint union of sets in $\mathcal{A}$ and so μ' is uniquely defined on $\mathcal{A}_\sigma$ and must coincide with μ^*

Let $B \in \mathcal{B}$ with $\mu^*(B) < \infty$. Choose $A \in \mathcal{A}_\sigma$ with $B \subseteq A$ and $\mu^*(A) \leq \mu^*(B) + \varepsilon$

Then $\mu'(B) \leq \mu'(A) = \mu^*(A) \leq \mu^*(B) + \varepsilon$

Thus it follows that $\mu'(B) \leq \mu^*(B)$ $\forall B \in \mathcal{B}$

Since the μ^* measurable sets form a σ-algebra containing $\mathcal{A}$, they must contain $\mathcal{B}$.

Let B be measurable, $A \in \mathcal{A}_\sigma$ s.t. $B \subseteq A$ with $\mu^*(A) \leq \mu^*(B) + \varepsilon$

Then $\mu'(A) = \mu^*(A) = \mu^*(B) + \mu^*(A \setminus B)$

Hence $\mu^*(A \setminus B) < \varepsilon$ if $\mu^*(B) < \infty$.

Thus $\mu^*(B) \le \mu^*(A) = \mu'(A) = \mu'(B) + \mu'(A \setminus B)$

$$\le \mu'(B) + \varepsilon$$

From this, it follows that $\mu^*(B) \le \mu'(B)$.

Now suppose μ is σ-finite. Let $\{X_i\}$ be a countable collection of disjoint sets in $\mathfrak{A}$ so that $X = \bigcup_{i=1}^{\infty} X_i$ and $\mu(X_i) < \infty$ $\forall i$.

Now any $B \in \mathcal{B}$ can be written $B = \bigcup_{i=1}^{\infty} (X_i \cap B)$

So we have $\mu'(B) = \sum_{i=1}^{\infty} \mu'(X_i \cap B)$

If $\bar{\mu}$ is a second measure ~~with these~~ on $\mathcal{B}$ which coincides with μ' on $\mathfrak{A}$, by the previous part

$\bar{\mu}(X_i \cap B) = \mu^*(X_i \cap B) = \mu'(X_i \cap B)$.

Thus $\bar{\mu}(B) = \sum \bar{\mu}(X_i \cap B) = \sum \mu'(X_i \cap B) = \mu'(B)$ $\square$

Summary of this section

Let $\mathcal{C}$ be a semi-algebra with a non-negative monotone set function μ on $\mathcal{C}$ such that

$$\mu(C) = \sum_{i=1}^{n} \mu(C_i) \quad \text{if} \quad C = \bigcup_{i=1}^{n} C_i \text{ is}$$

a finite ~~disjoint union of sets~~ union of disjoint sets in $\mathcal{C}$

THEN

- μ extends to a countably additive set function (also called μ) on $\mathcal{A}$ — the algebra generated by $\mathcal{C}$

- This extended μ is a premeasure on $\mathcal{A}$ which we use to define an ~~measure~~ outer measure μ^* on X

- The measurable sets of μ^* include $\mathcal{A}$ and $\mu^*|_{\mathcal{A}} = \mu$.

- Let $\mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A}$ (so it is contained in the measurable sets)

  Then there is a unique measure on $\mathcal{B}$ which coincides with μ on $\mathcal{A}$, viz μ^*.

Product Measures

Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be two measure space.

Consider the direct product $X \times Y$.

We want to make it into a ~~good~~ measure space.

Defn. If $A \subseteq X$ and $B \subseteq Y$ are msble sets then $A \times B \subseteq X \times Y$ is called a measurable rectangle. Let $\mathcal{R}$ denote the collection of all measurable rectangles

Define a set function λ on $\mathcal{R}$ by

$$\lambda(A \times B) = \mu(A)\nu B$$

Notice that $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$

and $(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c)$

so $\mathcal{R}$ is a semialgebra

Lemma Let $\{A_i \times B_i\}$ be a countable collection of disjoint measurable rectangles s.t.

$$A \times B = \bigcup_{i=1}^{\infty} A_i \times B_i \quad \text{is a msble rectangle}$$

Then $\lambda(A \times B) = \sum_{i=1}^{\infty} \lambda(A_i \times B_i)$

Proof Let $x \in A$ be fixed. Then for all $y \in B$

(x, y) is in exactly one rectangle $A_i \times B_i$

(as they are disjoint). Hence B is the disjoint union

of all the B_i such that $x \in A_i$ corresponds to B_i

Thus $\sum \nu(B_i) \chi_{A_i}(x) = \nu(B) \chi_A(x)$

so $\int \nu(B_i) \chi_{A_i}(x) d\mu(x) = \nu(B) \int \chi_A d\mu$

so $\sum \nu(B_i) \mu(A_i) = \nu(B) \mu(A) = \lambda(A \times B)$ $\square$

Thus we've shown that λ is a premeasure on the semi-algebra $\mathcal{R}$. Using the Carathéodory theorem we may extend λ to obtain a complete measure on $X \times Y$. We denote this by $\mu \times \nu$.

Note that $\mu \times \nu$ is defined on the smallest σ-algebra containing $\mathcal{R}$ (Called the product σ-algebra) and may be extended uniquely to the class of $\mu \times \nu$ measurable sets.

Examples If $X = Y = \mathbb{R}$, then $\mu \times \nu$ is Lebesgue measure on $\mathbb{R}^2$.

<u>Definition</u> Given $x \in X$ and $E \subseteq X \times Y$, define

$$E_x = \{ y \in Y : (x,y) \in E \}$$

as the <u>section</u> of E at x.

Note that $\chi_{E_x}(y) = \chi_E(x,y)$

Consider the $\mu \times \nu$ msble sets in $X \times Y$

<u>Lemma</u> - Let $x \in X$ and $E \in \mathcal{R}_{\sigma\delta}$. Then E_x is a msble subset of Y.

<u>Proof</u>. This is true if $E \in \mathcal{R}$ by definition of $\mathcal{R}$

Let $E \in \mathcal{R}_\sigma$, so that $E = \bigcup_{i=1}^{\infty} E_i$, where $E_i \in \mathcal{R}$ for all i

Then $\chi_{E_x}(y) = \chi_E(x,y) = \sup_i \chi_{E_i}(x,y)$

$$= \sup_i \chi_{(E_i)_x}(y)$$

Since $E_i \in \mathcal{R}$, $(E_i)_x$ is msble, so $\chi_{(E_i)_x}$ is a msble function. Thus the sup. is measurable. It follows that χ_{E_x} is a msble fn, which implies that E_x is a msble set

Now suppose $E \in \mathcal{R}_{\sigma\delta}$ - we may suppose $E = \bigcap_{i=1}^{\infty} E_i$ where $E_i \in \mathcal{R}_\sigma$.

Now $\chi_{E_x} = \chi_E(x,y) = \inf_i \chi_E(x,y)$
$$= \inf_i \chi_{(E_i)_x}(y)$$

Since $(E_i)_x$ is a msble set, it follows as above that E_x is msble □

<u>Lemma</u> Let $E \in \mathcal{R}_{\sigma\delta}$ and s'pose that $\mu \times \nu(E) < \infty$. Define $g(x) = \nu(E_x)$. Then g is a msble function of x and
$$\int_X g\, d\mu = \mu \times \nu(E)$$

<u>Proof</u>. The lemma is clearly true if $E \in \mathcal{R}$. If $E \in \mathcal{R}_\sigma$ has the form $E = \bigcup_{i=1}^{\infty} E_i$ with $E_i \in \mathcal{R}$ pairwise disjoint, let $g_i(x) = \nu((E_i)_x)$. Then $g_i \geq 0$, g_i is msble and $g = \sum_{i=1}^{\infty} g_i$. Letting $g^n = \sum_{i=1}^{n} g_i$, we see that $g^n \nearrow g$. Then g is msble and so by the monotone convergence thm

$$\int g\, d\mu = \sum_{i=1}^{n} g_i\, d\mu = \sum_{i=1}^{n} \mu \times \nu(E_i) = \mu \times \nu(E)$$

Thus the lemma holds if $E \in \mathcal{R}_\sigma$

Now suppose $E \in \mathcal{R}_{\sigma\delta}$. Choose a sequence $\{E_i\}_{i=1}^{\infty}$ in $\mathcal{R}_\sigma$ with $E_{i+1} \subseteq E_i$ and $E = \bigcap_{i=1}^{\infty} E_i$ and $(\mu \times \nu)(E_i) < \infty$ for all i.

It follows that $\mu \times \nu(E) < \infty$, so there exists $A \in R_\sigma$ with

$$\mu\times\nu(A) \leq \mu\times\nu(E) + \varepsilon$$

Let $g_i(x) = \nu((E_i)_x)$. Then $g(x) = \lim_{i\to\infty} g_i(x)$, so g is msble. Furthermore $\nu((E_{i+1})_x) \leq \nu((E_i)_x)$. So again by the MCT we have

$$\int g(x)\,d\mu(x) = \lim_i \int g_i(x)\,d\mu(x) = \lim_i \mu\times\nu(E_i)$$

$$= \mu\times\nu(E).$$

We now come to the two big theorems on product measures.

<u>Theorem</u> (Fubini) Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be two measure spaces, and f a $\mu\times\nu$ integrable function on $X\times Y$. Then

(i) For μ a.e. x, f_x defined by $f_x(y) = f(x,y)$ is ν integrable on Y, and for ν a.e. y, f_y defined by $f_y(x) = f(x,y)$ is μ-integrable on X.

(ii) The function defined a.e. on X by

$$x \mapsto \int_Y f_x(y)\, d\nu(y)$$ is μ-integrable

and the function defined a.e. on Y by

$$y \mapsto \int_X f_y(x)\, d\mu(x)$$ is ν-integrable

(iii) $$\int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x) = \int_Y \left[\int_X f(x,y)\, d\mu(x)\right] d\mu(y)$$

$$= \int_{X\times Y} f(x,y)\, d\mu\times\nu(x,y)$$

<u>Proof</u> Clear in the case where $f(x,y) = \chi_C(x,y)$ where C is a measurable subset of $X\times X$.

Hence, it's clear for simple functions.

Since f is integrable on $X\times Y$ we can find a sequence $\{\varphi_n\}$ of simple functions s.t.

$\int |f-\varphi_n|\, d\mu\times\nu \to 0$ and $f-\varphi_n \to 0$ in measure

and $\varphi_n \le f$.

Since φ_n is increasing to f, $\varphi_n(x,y) \le f(x,y)$

So f_x is the pwise increasing limit of $(\varphi_n)_x$

Thus f_x is msble and by the MCT

$$\lim_{n\to\infty} \int (\varphi_n)_x \, d\nu = \int f_x \, d\nu$$ and this is an increasing sequence

Since the φ_n are simple functions, each of the functions $x \longmapsto \int (\varphi_n)_x \, d\nu$ is measurable and integrable (in x). It follows that $\int f_x \, d\nu$ is also measurable and integrable in x, and

$$\iint f_x \, d\nu \; d\mu(x) = \lim_{n\to\infty} \iint (\varphi_n)_x \, d\nu \, d\mu(x)$$

$$= \lim_{n\to\infty} \int \varphi_n \, d\mu \times \nu$$

It follows that

$$\iint f(x,y) \, d\nu(y) \, d\mu(x) = \int f \, d(\mu \times \nu)$$

The other ~~inequ~~ equality follows by symmetry $\square$

<u>Theorem</u> (Tonelli) Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be two σ-finite measure spaces, and let f be a nonnegative fn on $X \times Y$

Suppose that

(i) For a.e. x $f_x(y) = f(x,y)$ is an integrable fn on Y and for a.e. y, $f_y(x) = f(x,y)$ is an integ. fn. on X

(ii) The functions $x \mapsto \int f(x,y)\,d\nu(y)$ and $y \mapsto \int f(x,y)\,d\mu(x)$ are both integrable on X (resp. Y)

Then f is integrable and

$$\int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x) = \int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y)$$
$$= \int_{X \times Y} f(x,y)\,d\mu\times\nu(x,y)$$

<u>Proof</u> Since X, Y are σ-finite, $X \times Y$ is σ-finite

Write $X \times Y = \bigcup_{i=1}^{\infty} Z_i$ with $Z_i \cap Z_j = \phi$ for $i \neq j$ and $(\mu \times \nu)(Z_i) < \infty$ $\forall i$.

It will suffice to show that f is integrable on each of the Z_i's.

To see this, let $f_n = f \wedge n = \begin{cases} f(x) & \text{if } f(x) \leq n \\ n & \text{if } f(x) > n \end{cases}$

Then $f_n \nearrow f$. By the Bounded Convergence Theorem f_n is integrable, and by the MCT, the theorem follows. $\square$

Infinite product spaces

Let $\{(X_i, \mathcal{B}_i, \mu_i)\}_{i=1}^{\infty}$ be a sequence of measure spaces.

Define $X = \prod_{i=1}^{\infty} X_i = \{ x = (x_1, x_2 \dots) : x_i \in X_i \}$

Let $\mathcal{C}$ = set of all rectangles $\{B_1 \times \dots \times B_n : n \in \mathbb{N},\ B_i \in \mathcal{C}_i \ i=1 \dots n\}$

where $B_1 \times \dots \times B_n = \{ x \in X : x_1 \in B_1, x_2 \in B_2, \dots, x_n \in B_n \}$

Then $\mathcal{C}$ is a semi-algebra and

$\mu = \bigotimes_{i=1}^{\infty} \mu_i$ may be defined on $\mathcal{C}$ by

$$\mu(B_1 \times \dots \times B_n) = \mu_1(B_1)\mu_2(B_2) \dots \mu_n(B_n)$$

Using the Carathéodory theorem, we may define $\mathcal{B}$ = the smallest σ-algebra containing $\mathcal{C}$ and extend μ to a measure on $\mathcal{B}$

$(X, \mathcal{B}, \mu)$ is called the <u>infinite product space</u>

<u>Example</u>: We already saw an example when $X_i = \{0, 1\}$. Choose $\mu_i(0) + \mu_i(1) = 1$ for $i = 1, 2, \dots$

Then X = all sequences of 0's and 1's

$$\mu(\{0\} \times \{1\} \times \{0\} \times \{0,1\} \times \dots \times \{0,1\} \times \dots)$$

$$= \mu_1(0)\mu_2(1)\mu_3(0) \qquad \text{(finite product)}$$

Let $(X, \mathcal{A}, \mu)$ be a probability space, i.e. a measure space with $\mu(X)=1$.

Probabilists interpret the elements in X as <u>events</u> (X is the <u>sample space</u>) and elements of $\mathcal{A}$ as <u>collections of events</u>, which may be the outcome of some observations. If $A \in \mathcal{A}$, then $\mu(A)$ is thought of as the probability that this collection of events occurs

eg - infinite product space $\prod \{0,1\} \longleftrightarrow$ coin tossing

- unit interval $[0,1] \longleftrightarrow$ probability that an atom emits an α-particle within unit time

<u>Random variables</u> are measurable functions

— 1 —

Given a probability space, a sub σ-algebra $\mathcal{B} \subseteq \mathcal{A}$ corresponds heuristically to partial information (eg info. only up to n tosses of the coin, or up to time t)

Often the space X is denoted by Ω - and if a pt $\omega \in \Omega$ is chosen and the probability that it lies in $B \in \mathcal{B}$ is $\mu(B)$, which we denote by $P(B)$.

The observer, not knowing ω, might only know the information that $\left.\begin{array}{l}\omega \in B \\ \text{or } \omega \notin B\end{array}\right\}$ for a given $B \in \mathcal{B}$ - but not for all $A \in \mathcal{A}$.

How can we model this?

Of course, given two sets A, B, the probability that a w belongs to A w.r.t. a smaller set B

is $P(A|B) = \frac{P(A\cap B)}{P(B)}$ unless $P(B)=0$, in which case it isn't defined

So suppose we know $P(A|B_i)$ for $B_1, B_2, \ldots$

and suppose $\mathcal{B}$ is the σ-algebra generated by the disjoint sets $B_1, B_2, \ldots$

Let $f(w) = \frac{P(A\cap B_i)}{P(B_i)}$ if $w \in B_i$, $i=1,2,\ldots$

This random variable is denoted $P(A\|\mathcal{B})$ — the conditional probability of A given $\mathcal{B}$.

If $P(B_i)=0$ for $B_i \neq \emptyset$, this definition is not well-made. In that case, we take $P(A\|\mathcal{B})$ to be an arbitrary constant value, so actually $P(A\|\mathcal{B})$ is a whole family of measurable functions defined a.e. Choosing one of these is called choosing a <u>version</u> of $P(A\|\mathcal{B})$.

— | —

In a more general setting, let $\mathcal{B}$ be any sub σ-algebra of $\mathcal{A}$. Imagine an observer who knows, for each $A\in\mathcal{A}$ whether $w\in A$ or not.

Think of $\mathcal{B}$ as a collection of experiments.

For $A\in\mathcal{A}$, fixed, we define a finite measure on $\mathcal{B}$ by $\nu(B) = \mu(A\cap B)$ $\forall B\in\mathcal{B}$

Then $\mu(B)=0 \Rightarrow \nu(B)=0$. By the Radon-Nikodym theorem applied to μ and ν, $\exists$ a $\mathcal{B}$-msble function $f = \frac{d\nu}{d\mu}$ s.t. $\forall B \in \mathcal{B}$

$$\mu(A \cap B) = \nu(B) = \int_B f \, d\mu$$

The function f is denoted $P(A \| \mathcal{B})$

It's a random variable with two properties

(i) $P(A \| \mathcal{B})$ is $\mathcal{B}$-msble and integrable

(ii) $P(A \| \mathcal{B})$ satisfies

$$\int_B P(A \| \mathcal{B}) \, dP = P(A \cap B) \quad \forall B \in \mathcal{B}.$$

There will in general be many such random variables but any two are equal with probability 1 (i.e. <u>almost surely</u>) A specific such random variable is called a <u>version</u> of the conditional probability

<u>Example</u>. Take $X = [0,1]$, $\mathcal{B}_n = \sigma$-alg. generated by

$$\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right) : j = 0, \ldots 2^n - 1$$

Then for a set A

$$P(A \| \mathcal{B})(\omega) = \frac{1}{2^n} \mu\left(A \cap \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right)\right) \text{ if } \frac{j}{2^n} \le \omega < \frac{j+1}{2^n}$$

Note that $\quad 0 \le P(A \| \mathcal{B}) \le 1 \quad$ a.s.

and if $\quad A = \bigcup_{n=1}^{\infty} A_n \quad$ is a disjoint union, the

$$P(A \| \mathcal{B}) = \sum_n P(A_n \| \mathcal{B}).$$

Conditional Expectations

Defn Let f be an integrable random variable on $(X, \mathcal{A}, \mu)$ and $\mathcal{B} \subseteq \mathcal{A}$ a σ-algebra.

There exists a random variable

$$E_\mu(f \,\|\, \mathcal{B}) \quad \text{such that}$$

(i) $E_\mu(f \,\|\, \mathcal{B})$ is $\mathcal{B}$-msble and integrable

(ii) $\forall B \in \mathcal{B}$ $\quad \int_B E_\mu(f \,\|\, \mathcal{B})\, d\mu = \int_B f\, d\mu$

Proof. Define a measure ν on $\mathcal{B}$ by $\nu(B) = \int_B f\, d\mu$. This is finite, as f is integrable and $\nu \ll \mu$. Let $h = \frac{d\nu}{d\mu}$. Then denote

$$h = E_\mu(f \,\|\, \mathcal{B}) \qquad \square$$

As usual, ~~this~~ $E_\mu(f \,\|\, \mathcal{B})$ is only defined a.e. Any such h is called a <u>version</u> of the conditional expectation.

$E_\mu(f \,\|\, \mathcal{B})$ can be interpreted as the expected value of f for someone who knows, for each $B \in \mathcal{B}$, whether $w \in B$ or $w \notin B$.

If $\mathcal{B} = \mathcal{A}$, then $E_\mu(f \,\|\, \mathcal{A}) = f$ a.e.

If $\mathcal{B} = \{\emptyset, \Omega\}$ then $E_\mu(f \,\|\, \mathcal{B}) = \int_\Omega f\, d\mu$

<u>Theorem</u> Suppose f, g and $\{f_n\} \in L^1(X, \mu)$

Then

(i) If $f = a$ (const) a.e. then $E_\mu(f \| \mathcal{B}) = a$

(ii) If a, b are constants then

$$E_\mu(af + bg \| \mathcal{B}) = a E_\mu(f \| \mathcal{B}) + b E_\mu(g \| \mathcal{B})$$

(iii) $|E_\mu(f \| \mathcal{B})| \leq E_\mu(|f| \| \mathcal{B})$

(iv) If $\lim_n f_n = f$ a.e. and $|f_n| \leq g$, then

$$\lim_n (E_\mu f_n \| \mathcal{B}) = E_\mu(f \| \mathcal{B}) \text{ a.e.}$$

(v) If f is $\mathcal{B}$-msble and if g and $f \cdot g \in L^1$, then

$$E_\mu(fg \| \mathcal{B}) = f E_\mu(g \| \mathcal{B})$$

(vi) (Jensen's Inequality)

$$|E_\mu(f \| \mathcal{B})|^p \leq E_\mu(|f|^p \| \mathcal{B}) \text{ for } 1 \leq p \leq \infty$$

— 1 —

<u>Martingales</u>

Let $f_1, f_2, \ldots$ be a sequence of random variables on $(X, \mathcal{A}, \mu)$ and let $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \ldots$ be a sequence of σ-algebras. Then $(f_n, \mathcal{A}_n)$ is called a (sub) martingale if

(i) $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ $\forall n$

(ii) f_n is $\mathcal{A}_n$ msble (f_n is <u>adapted</u> to $\mathcal{A}_n$)

(iii) $\sup_n \int |f_n| \, d\mu < \infty$

(iv) $E_\mu(f_{n+1} \| \mathcal{A}_n) = f_n$ a.e.

<u>Submartingale</u>

$(\leq f_n \text{ a.e.})$

Example $\Omega = [0,1]$

$$A_n = \sigma\text{-algebra gen by } \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right)$$

Take $f \in L^1([0,1])$.

Then $f_n(x) = \frac{1}{2^n} \int_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} f(t)\,dt$ if $x \in \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right)$

is a martingale

eg $X = \prod_1^\infty \{0,1\}$ $\quad \mu_n = \left(\frac{1}{2}, \frac{1}{2}\right)$ $\quad \mu = \bigotimes \mu_n$

$A_n = \mathcal{C}_n = \sigma$-alg generated by rectangles $B_1 \times \dots \times B_n$ for n fixed

$f \in L^1(X)$

$$f_n(x) = \int \cdots \int f(x_1, x_2, \dots, x_n, \xi_{n+1}, \xi_{n+2} \dots)\, d\mu_{n+1}(\xi_{n+1})\, d\mu_{n+2}(\xi_{n+2}) \dots$$

is a martingale

— | —

Indeed, let f be any integrable fn. in $L^1(X, A, \mu)$ and let $A_1 \subseteq A_2 \subseteq A_3 \dots \subseteq A_n \subseteq \dots \subseteq A$ be a sequence of σ-algebras. Then $f_n = E_\mu(f \| A_n)$ is a martingale

<u>Theorem</u> (The martingale convergence theorem)

Let $f_1, f_2, \dots, f_n, \dots$ be a (sub) martingale s.t.

$$K = \sup_n \int |f_n| \, d\mu < \infty$$

Then $\exists$ a random vble f on X s.t. $f_n \to f$ a.e.

and $\int |f| \, d\mu \le K$. [If unif. integ $f_n \to f$ in L^1]

<u>Proof</u>. In fact, we want to define f by

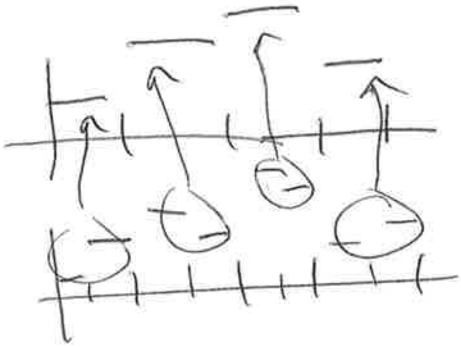
$$E_\mu(f, A_n) = f_n$$

To do this, let $f = \limsup f_n$ and then show that $\limsup_n f_n = \liminf_n f_n$ by analysing the number of upcrossings in the martingale.

<u>Examples</u> On $[0,1]$

f_2

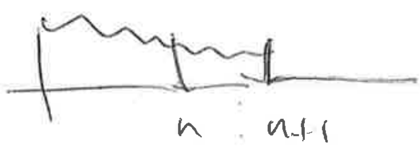
f_3



average value

Then $f_i \to f$ a.e.

On $\mathbb{R}$ $\quad f_n =$ σ-alg gen by $\mathcal{B} \cap [0, n]$



$f_n \to f$ in L^1.

To prove the martingale convergence theorem requires two basic inequalities

<u>Lemma A</u> (Kolmogorov's Inequality)

If $X_1, X_2, \ldots, X_n$ is a submartingale, then for $\alpha > 0$

$$\mu\{\max_{i \le n} X_i \geq \alpha\} \le \frac{1}{\alpha}\int |X_n|\, d\mu$$

<u>Proof</u> Let $A_1 = \{\omega : X_1(\omega) \geq \alpha\}$

$$A_k = \{\omega : \max_{i<k} X_i(\omega) < \alpha \le X_k(\omega)\}$$

$$A = \bigcup_{k=1}^{n} A_k = \{\omega : \max_{i \le n} X_i(\omega) \geq \alpha\}$$

Since $A_k \in \mathcal{F}_k = \sigma(X_1, \ldots X_k)$

$$\int_A X_n\, d\mu = \sum_{k=1}^{n} \int_{A_k} X_n\, d\mu = \sum_{k=1}^{n} \int_{A_k} E\{X_n \| \mathcal{F}_k\}\, d\mu$$

$$\geq \sum_{k=1}^{n} \int_{A_k} X_k\, d\mu \geq \alpha \sum \mu(A_k) = \alpha\, \mu(A)$$

Thus $\alpha\, \mu\{\omega : \max_{i \le n} X_i \geq \alpha\} \le \int_{\{\omega : \max_{i \le n} X_i(\omega) \geq \alpha\}} X_n\, d\mu \le \int X_n^+ \le \int |X_n|\, d\mu$ $\Box$

<u>Note</u> if $\{X_i\}$ is a mgale, $|X_i|$ is a submartingale

so $\mu\{\max_{i \le n} |X_i| \geq \alpha\} \le \frac{1}{\alpha}\int |X_n|\, d\mu$

<u>Definition</u> If $[\alpha, \beta]$, $\alpha < \beta$, is an interval and $X_1 \ldots X_n$ are random variables, the <u>number of upcrossings</u> of $[\alpha, \beta]$ by $X_1(\omega) \ldots X_n(\omega)$ is the number of times the sequence $X_1(\omega) \ldots X_n(\omega)$ passes from below α to above β.

Specifically, it is the number of integers u, v

$1 \le u < v \le n$ s.t. $X_u \le \alpha$

$\alpha < X_i < \beta$ for $u < i < v$ (vacuous condition if $u+1=v$)

and $X_v \geq \beta$

Theorem (The upcrossings lemma)

If $X_1, X_2, \dots, X_n$ is a submartingale and

$U(\omega) = \#$ upcrossings of $[\alpha, \beta]$ by $X_1(\omega), \dots, X_n(\omega)$

then

$$\int U(\omega)\, d\mu(\omega) \le \frac{1}{\beta-\alpha}\left(\int |X_n|\, d\mu + \alpha\right)$$

Proof. We require a lemma

Lemma B. If $X_1, \dots, X_n$ is a submartingale and τ_1, τ_2 are stopping times (functions $X \to \mathbb{N}$ s.t. $\{\tau_i(\omega) \le k\} \in \mathcal{F}_k$) such that $1 \le \tau_1(\omega) \le \tau_2(\omega) \le n$ a.e. ω

then $$\int X_{\tau_1(\omega)}(\omega)\, d\mu(\omega) \le \int X_{\tau_2(\omega)}(\omega)\, d\mu(\omega)$$

[N.B. Since $\int |X_{\tau(\omega)}(\omega)|\, d\mu(\omega) \le \sum_{k=1}^{n} \int |X_n|\, d\mu$ the functions are integrable]

Proof Let $\Delta_k = X_k - X_{k-1}$. Then

$$\int (X_{\tau_1} - X_{\tau_2})\, d\mu = \int \left(\sum_{k=1}^{n} \Delta_k \chi_{\tau_1 < k < \tau_2}\right) d\mu$$

$$= \sum_{k=1}^{n} \int_{\tau_1 < k < \tau_2} \Delta_k\, d\mu$$

Since $\{w : \tau_1(w) < k < \tau_2(w)\} = \{w : \tau_1(w) \le k-1\} \cap \{w : \tau_2(w) \le k-1\}^c \in \mathcal{F}_{k-1}$

the integrals are non-negative $\square$

<u>Proof of the upcrossings theorem</u>

Let $Y_k = \max\{0, Y_k - \alpha\}$

$\theta = \beta - \alpha$

Then U is also the # of upcrossings of $[0, \theta]$ by the submartingale $Y_1 \dots Y_n$.

Define inductively $\tau_0, \tau_1, \dots$

$\tau_0 = 1$

τ_k for <u>even</u> k is the smallest j s.t. $\tau_{k-1} < j \le n$ and $Y_j \ge \theta$, and n if there is no such j

τ_k for <u>odd</u> k is the smallest j s.t. $\tau_{k-1} < j \le n$ and $Y_j = 0$ and is n if there is no such j.

Note that $\tau_k \le n$ and $\tau_{k-1} = n \Rightarrow \tau_k = n$

Also $\tau_{k-1} < n \Rightarrow \tau_{k-1} < \tau_k$ and hence $\tau_k = \tau_{k+1} = \dots = n$

— 1 —

If k is odd and τ_{k-1} is a stopping time, then for $j < n$

$\{w : \tau_k(w) = j\} = \bigcup_{i=1}^{j-1} \{w : \tau_{k-1}(w) = i, \ Y_{i+1}(w) \ge 0 \dots Y_{j-1}(w) \ge 0$ and $Y_j(w) = 0\}$

this is in $\mathcal{F}_j$, and $(\tau_k = n) = [\tau_k \le n-1]^c$ this is in $\mathcal{F}_n$

so τ_k is a stopping time

A similar argument for k even shows that the τ_k are stopping times, as τ_0 is.

Since $Y_n = Y_{\tau_n} \geq Y_{\tau_n} - Y_{\tau_0}$

$$Y_n \geq \sum_e (Y_{\tau_k} - Y_{\tau_{k-1}}) + \sum_o (Y_{\tau_k} - Y_{\tau_{k-1}})$$

even odd

By Lemma B, the last sum has non-negative integral

so $\int Y_n \, d\mu \geq \int \sum_e (Y_{\tau_k} - Y_{\tau_{k-1}}) \, d\mu$

Suppose now k is even

If $\tau_{k-1} = n$ then $Y_{\tau_{k-1}} = Y_n = Y_{\tau_k}$

while if $\tau_{k-1} < n$ then $Y_{\tau_{k-1}} = 0$. In either case $Y_{\tau_k} - Y_{\tau_{k-1}} \geq 0$. On the other hand, if $1 \leq u < v \leq n$ $Y_u = 0$ and $0 < Y_i < \theta$ for $u < i < v$, $Y_v \geq \theta$ (that is u, v contribute to an upcrossing of $[0, \theta]$ by $Y_1 \ldots Y_n$). Then $\exists$ even k $1 \leq k \leq n$ s.t. $\tau_{k-1} \leq u < v = \tau_k$.

Thus $Y_{\tau_{k-1}} = 0$ and $Y_{\tau_k} \geq \theta$ and so $Y_{\tau_k} - Y_{\tau_{k-1}} \geq \theta$

Since different upcrossings give different values of k

$\sum_e Y_{\tau_k} - Y_{\tau_{k-1}} \geq U\theta$ so $\int Y_n \, d\mu \geq \theta \int U \, d\mu$

In terms of the original random variables

$$(\beta - \alpha) \int U \, d\mu \leq \int_{X_n > \alpha} (X_n - \alpha) \, d\mu = \int |X_n| \, d\mu + |\alpha| \quad \square$$

Proof of the martingale convergence theorem

Fix α, β and let U_n = number of upcrossings of $[\alpha, \beta]$ by $X_1 \cdots X_n$. Then

$$\int U_n \, d\mu \leq \frac{\int |X_n| + |\alpha|}{\beta - \alpha} \leq \frac{K + |\alpha|}{\beta - \alpha}$$

Since U_n is non-decreasing and $\int U_n \, d\mu$ is bounded, the monotone convergence theorem implies $\sup_n U_n$ is integrable, hence finite valued a.e.

Let
$$\left.\begin{aligned} X^* &= \limsup_{i \to \infty} X_i \\ X_* &= \liminf_{i \to \infty} X_i \end{aligned}\right\} \text{ possibly } \infty$$

If $X_* < \alpha < \beta < X^*$, then U_n must go to ∞.

Since U_n is a.e. bounded

$$\mu\{\omega : X_*(\omega) < \alpha < \beta < X^*(\omega)\} = 0$$

Now $\{\omega : X_* < X^*\} = \bigcup_{\substack{\alpha, \beta \\ \text{rational}}} \{X_* < \alpha < \beta < X^*\}$

Therefore $X_* = X^*$ a.e. and X_k converges to their common value X — which may be ∞.

By Fatou's lemma

$$\int |X| \leq \liminf \int |X_n| \leq K$$

Since X is integrable, it is finite a.e. $\square$