Quiz – STA 5446 – 2005

1. Let $P : \Omega \to [0,1]$ be a probability law. Let $A, B \subseteq \Omega$. Show that
$P(A) = P(A|B) \cdot P(B) + P(A|B^c) P(B^c)$, where B^c denotes the complement of B.

2. Let A be an arbitrary set. Define $\inf A$, where $\inf A$ denotes the infimum of the set A.

3. If $\sum_{K=0}^{\infty} a_K$ is convergent, what do we know about the convergence of $\{a_K\}_K$, as $K \to \infty$? Can you prove this?

4. Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$, where I is an interval. Let $x_0 \in I$. Give two equivalent definitions for "f is continuous at x_0".

5. Let $f_n : I \subseteq \mathbb{R} \to \mathbb{R}$ be a sequence of functions, $n \in \mathbb{N}$.

a) What does "pointwise convergence" of $\{f_n\}_n$ mean?

b) What does "uniform convergence" of $\{f_n\}_n$ mean?

6. Can you state and <u>prove</u> Chebyshev's inequality?

7. Can you state and suggest a way of proving the central limit theorem (CLT)?

STA 5446- Fall '05

<u>Some results and definitions from real anlaysis, which are required for this course:</u>

We shall keep all our discussion confined to R^1 so it would suffice to know the below mentioned definitions/results only for functions/sets on a real line.

- Definitions
  - What do we mean by an infimum, supremum, least upper bound (lub) greatest lower bound (glb) of a set and Limsup and Liminf for a sequence of real numbers.
  - What are countable and uncountable sets?
  - The $\epsilon - \delta$ definition of (a) convergence of a series/ sequence and (b) continuity of a real valued function.
  - What is uniform continuity? What is a monotonic function?
  - In R^1 what are closed, open, compact and bounded sets?
- Results
  - How are the infimum, supremum, least upper bound (lub) greatest lower bound (glb) of a set related?
  - How are the Limsup and Liminf for a sequence of real numbers related?
  - Is the set of rationals countable?
  - Continuous fxns on a compact set is uniformly continuous
  - A subset of R^1 is closed and bounded $\iff$ it is compact (also known as Heine Borel theorem).
  - We should be able to work out elementary results like composition of 2 continuous functions is also continuous.
  - It would be good if we knew stuff from Calculus I, II like the Fundamental Theorem of Calculus and stuff, however we do not need things from Calculus III (e.g. multiple integrals Green's theorem etc).

Chapter 1

Motivation

SECTION 1 offers some reasons for why anyone who uses probability should know about the measure theoretic approach.
SECTION 2 describes some of the added complications, and some of the compensating benefits that come with the rigorous treatment of probabilities as measures.
SECTION 3 argues that there are advantages in approaching the study of probability theory via expectations, interpreted as linear functionals, as the basic concept.
SECTION 4 describes the de Finetti convention of identifying a set with its indicator function, and of using the same symbol for a probability measure and its corresponding expectation.
*SECTION *5 presents a fair-price interpretation of probability, which emphasizes the linearity properties of expectations. The interpretation is sometimes a useful guide to intuition.*

1. Why bother with measure theory?

Following the appearance of the little book by Kolmogorov (1933), which set forth a measure theoretic foundation for probability theory, it has been widely accepted that probabilities should be studied as special sorts of measures. (More or less true—see the Notes to the Chapter.) Anyone who wants to understand modern probability theory will have to learn something about measures and integrals, but it takes surprisingly little to get started.

For a rigorous treatment of probability, the measure theoretic approach is a vast improvement over the arguments usually presented in undergraduate courses. Let me remind you of some difficulties with the typical introduction to probability.

Independence

There are various elementary definitions of independence for random variables. For example, one can require factorization of distribution functions,

$$\mathbb{P}\{X \leq x, Y \leq y\} = \mathbb{P}\{X \leq x\}\,\mathbb{P}\{Y \leq y\} \qquad \text{for all real } x, y.$$

The problem with this definition is that one needs to be able to calculate distribution functions, which can make it impossible to establish rigorously some desirable properties of independence. For example, suppose $X_1, \ldots, X_4$ are independent random variables. How would you show that

$$Y = X_1 X_2 \left[\log\left(\frac{X_1^2 + X_2^2}{|X_1| + |X_2|} \right) + \frac{|X_1|^3 + X_2^3}{X_1^4 + X_2^4} \right]$$

is independent of

$$Z = \sin\left[X_3 + X_3^2 + X_3 X_4 + X_4^2 + \sqrt{X_3^4 + X_4^4} \right],$$

by means of distribution functions? Somehow you would need to express events $\{Y \le y, Z \le z\}$ in terms of the events $\{X_i \le x_i\}$, which is not an easy task. (If you did figure out how to do it, I could easily make up more taxing examples.)

You might also try to define independence via factorization of joint density functions, but I could invent further examples to make your life miserable, such as problems where the joint distribution of the random variables are not even given by densities. And if you could grind out the joint densities, probably by means of horrible calculations with Jacobians, you might end up with the mistaken impression that independence had something to do with the smoothness of the transformations.

The difficulty disappears in a measure theoretic treatment, as you will see in Chapter 4. Facts about independence correspond to facts about product measures.

Discrete versus continuous

Most introductory texts offer proofs of the Tchebychev inequality,

$$\mathbb{P}\{|X - \mu| \ge \epsilon\} \le \mathrm{var}(X)/\epsilon^2,$$

where μ denotes the expected value of X. Many texts even offer two proofs, one for the discrete case and another for the continuous case. Indeed, introductory courses tend to split into at least two segments. First one establishes all manner of results for discrete random variables and then one reproves almost the same results for random variables with densities.

Unnecessary distinctions between discrete and continuous distributions disappear in a measure theoretic treatment, as you will see in Chapter 3.

Univariate versus multivariate

The unnecessary repetition does not stop with the discrete/continuous dichotomy. After one masters formulae for functions of a single random variable, the whole process starts over for several random variables. The univariate definitions acquire a prefix *joint*, leading to a whole host of new exercises in multivariate calculus: joint densities, Jacobians, multiple integrals, joint moment generating functions, and so on.

Again the distinctions largely disappear in a measure theoretic treatment. Distributions are just image measures; joint distributions are just image measures for maps into product spaces; the same definitions and theorems apply in both cases. One saves a huge amount of unnecessary repetition by recognizing the role of image measures (described in Chapter 2) and recognizing joint distributions as measures on product spaces (described in Chapter 4).

Approximation of distributions

Roughly speaking, the central limit theorem asserts:

If $\xi_1, \ldots, \xi_n$ are independent random variables with zero expected values and variances summing to one, and if none of the ξ_i makes too large a contribution to their sum, then $\xi_1 + \ldots + \xi_n$ is approximately $N(0, 1)$ distributed.

What exactly does that mean? How can something with a discrete distribution, such as a standardized Binomial, be approximated by a smooth normal distribution? The traditional answer (which is sometimes presented explicitly in introductory texts) involves pointwise convergence of distribution functions of random variables; but the central limit theorem is seldom established (even in introductory texts) by checking convergence of distribution functions. Instead, when proofs are given, they typically involve checking of pointwise convergence for some sort of generating function. The proof of the equivalence between convergence in distribution and pointwise convergence of generating functions is usually omitted. The treatment of convergence in distribution for random vectors is even murkier.

As you will see in Chapter 7, it is far cleaner to start from a definition involving convergence of expectations of "smooth functions" of the random variables, an approach that covers convergence in distribution for random variables, random vectors, and even random elements of metric spaces, all within a single framework.

***

In the long run the measure theoretic approach will save you much work and help you avoid wasted effort with unnecessary distinctions.

2. The cost and benefit of rigor

In traditional terminology, probabilities are numbers in the range $[0, 1]$ attached to events, that is, to subsets of a sample space Ω. They satisfy the rules

(i) $\mathbb{P}\emptyset = 0$ *and* $\mathbb{P}\Omega = 1$

(ii) *for disjoint events $A_1, A_2, \ldots$, the probability of their union, $\mathbb{P}(\cup_i A_i)$, is equal to $\sum_i \mathbb{P}A_i$, the sum of the probabilities of the individual events.*

When teaching introductory courses, I find that it pays to be a little vague about the meaning of the dots in (ii), explaining only that it lets us calculate the probability of an event by breaking it into disjoint pieces whose probabilities are summed. Probabilities add up in the same way as lengths, areas, volumes, and masses. The fact that we sometimes need a countable infinity of pieces (as in calculations involving potentially infinite sequences of coin tosses, for example) is best passed off as an obvious extension of the method for an arbitrarily large, finite number of pieces.

In fact the extension is not at all obvious, mathematically speaking. As explained by Hawkins (1979), the possibility of having the additivity property (ii) hold for countable collections of disjoint events, a property known officially as ***countable additivity***, is one of the great discoveries of modern mathematics. In his 1902 doctoral dissertation, Henri Lebesgue invented a method for defining lengths of complicated subsets of the real line, in a countably additive way. The definition has the subtle feature that not every subset has a length. Indeed, under the usual axioms of set theory, it is impossible to extend the concept of length to *all* subsets of the real line while preserving countable additivity.

The same subtlety carries over to probability theory. In general, the collection of events to which countably additive probabilities are assigned cannot include all subsets of the sample space. The domain of the set function $\mathbb{P}$ (the ***probability measure***) is usually just a ***sigma-field***, a collection of subsets of Ω with properties that will be defined in Chapter 2.

Many probabilistic ideas are greatly simplified by reformulation as properties of sigma-fields. For example, the unhelpful multitude of possible definitions for independence coalesce nicely into a single concept of independence for sigma-fields.

The sigma-field limitation turns out to be less of a disadvantage than might be feared. In fact, it has positive advantages when we wish to prove some probabilistic fact about all events in some sigma-field, $\mathcal{A}$. The obvious line of attack—first find an explicit representation for the typical member of $\mathcal{A}$, then check the desired property directly—usually fails. Instead, as you will see in Chapter 2, an indirect approach often succeeds.

(a) Show directly that the desired property holds for all events in some subclass $\mathcal{E}$ of "simpler sets" from $\mathcal{A}$.

(b) Show that $\mathcal{A}$ is the smallest sigma-field for which $\mathcal{A} \supseteq \mathcal{E}$.

(c) Show that the desired property is preserved under various set theoretic operations. For example, it might be possible to show that if two events have the property then so does their union.

(d) Deduce from (c) that the collection $\mathcal{B}$ of all events with the property forms a sigma-field of subsets of Ω. That is, $\mathcal{B}$ is a sigma-field, which, by (a), has the property $\mathcal{B} \supseteq \mathcal{E}$.

(e) Conclude from (b) and (d) that $\mathcal{B} \supseteq \mathcal{A}$. That is, the property holds for all members of $\mathcal{A}$.

REMARK. Don't worry about the details for the moment. I include the outline in this Chapter just to give the flavor of a typical measure theoretic proof. I have found that some students have trouble adapting to this style of argument.

The indirect argument might seem complicated, but, with the help of a few key theorems, it actually becomes routine. In the literature, it is not unusual to see applications abbreviated to a remark like "a simple generating class argument shows ...," with the reader left to fill in the routine details.

Lebesgue applied his definition of length (now known as Lebesgue measure) to the construction of an integral, extending and improving on the Riemann integral. Subsequent generalizations of Lebesgue's concept of measure (as in the 1913 paper of Radon and other developments described in the Epilogue to

Hawkins 1979) eventually opened the way for Kolmogorov to identify probabilities with measures on sigma-fields of events on general sample spaces. From the Preface to Kolmogorov (1933), in the 1950 translation by Morrison:

> The purpose of this monograph is to give an axiomatic foundation for the theory of probability. The author set himself the task of putting in their natural place, among the general notions of modern mathematics, the basic concepts of probability theory—concepts which until recently were considered to be quite peculiar.
>
> This task would have been a rather hopeless one before the introduction of Lebesgue's theories of measure and integration. However, after Lebesgue's publication of his investigations, the analogies between measure of a set and probability of an event, and between integral of a function and mathematical expectation of a random variable, became apparent. These analogies allowed of further extensions; thus, for example, various properties of independent random variables were seen to be in complete analogy with the corresponding properties of orthogonal functions. But if probability theory was to be based on the above analogies, it still was necessary to make the theories of measure and integration independent of the geometric elements which were in the foreground with Lebesgue. This has been done by Fréchet.
>
> While a conception of probability theory based on the above general viewpoints has been current for some time among certain mathematicians, there was lacking a complete exposition of the whole system, free of extraneous complications. (Cf., however, the book by Fréchet ...)

Kolmogorov identified random variables with a class of real-valued functions (the ***measurable functions***) possessing properties allowing them to coexist comfortably with the sigma-field. Thereby he was also able to identify the expectation operation as a special case of integration with respect to a measure. For the newly restricted class of random variables, in addition to the traditional properties

(i) $\mathbb{E}(c_1X_1 + c_2X_2) = c_1\mathbb{E}(X_1) + c_2\mathbb{E}(X_2)$, for constants c_1 and c_2,

(ii) $\mathbb{E}(X) \geq \mathbb{E}(Y)$ if $X \geq Y$,

he could benefit from further properties implied by the countable additivity of the probability measure.

As with the sigma-field requirement for events, the measurability restriction on the random variables came with benefits. In modern terminology, no longer was $\mathbb{E}$ just an ***increasing linear functional*** on the space of real random variables (with some restrictions to avoid problems with infinities), but also it had acquired some continuity properties, making possible a rigorous treatment of limiting operations in probability theory.

[redacted]

STA 5446 : Probability and Measure

Lecture 1

Motivation

Reading: • Chapter 1 from Pollard's book.
• Motivation 1.1 and 1.2 of Chapter 1, Shorack's book.

1) Riemman / Lebesgue integral.

a) Riemman

For f $\boxed{\text{continuous}}$ and positive, defined on $[a,b]$, $\int_a^b f(x)\,dx$ represents a type of limiting sum of $f(x)$ values weighted by small lengths of intervals.

• The m'th Riemman sum is formed by subdividing the domain of f:

$$RS_m = \sum_{i=1}^{m} f(x^*_{mi})\,(x_{mi} - x_{m,i-1}), \quad \text{where}$$

$a \equiv x_{m0} < x_{m1} < \dots < x_{mm} \equiv b$, and $x^*_{mi} \in [x_{m,i-1}, x_{m,i}]$.

• If the mesh of the partition $\to 0$ with m, $\int_a^b f(x)dx = \lim_m RS_m$.

- What if f is not continuous?
  Does the definition still work? No!

"How many" discontinuity points can f have such that integrability (in this sense) still holds?

How to MEASURE the set of "bad" points?

- Also, in general, we have statements of the type $f_n(x)$ "converges to" f and we would like to infer that the:

  "Integral of $f_n(x)$" "converges "to" Integral of $f(x)$

- If "convergence" is point-wise convergence and if "Integral" is Riemann integral we can find ourselves in the following situation:

Example Let $f_n : [0,1] \to \mathbb{R}$

$$f_n(x) = x^{n-1}, \quad n = 1, 2, \ldots$$

Then

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & \text{if } 0 \leq x < 1 \\ 1, & \text{if } x = 1 \end{cases} \equiv f(x)$$

Thus, we have a sequence of positive, continuous functions the limit of which is <u>NOT</u> continuous, hence the Riemann integral of $f(x)$ is not even defined, so we can not hope to get the desired result.

<u>b) Lebesgue</u> If f is a non-negative function (we'll see that the general case reduces to this), the Lebesgue integral can be also regarded as a limit of weighted sums, where now the weights are given by the "<u>measure</u>" of a set of points.

• The m-th Lebesgue sum (in analogy with the m-th Riemann sum) is:

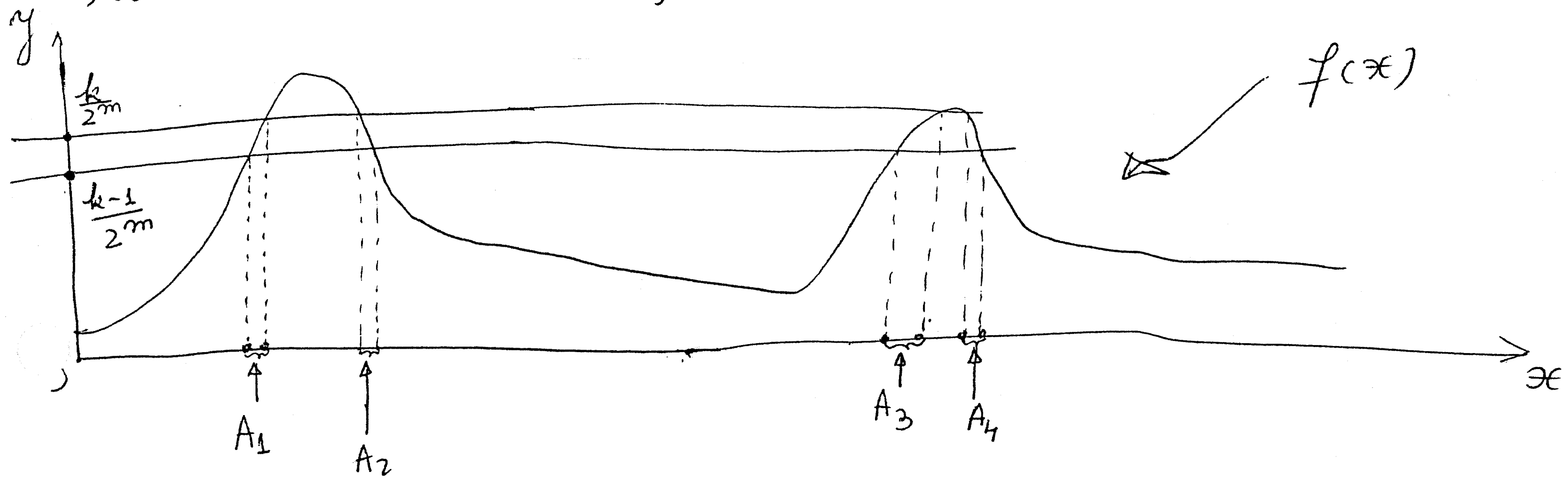
$$\boxed{LS_m = \sum_{k=1}^{m2^m} \frac{k-1}{2^m} \times \text{measure}\left(\left\{x : \frac{k-1}{2^m} \leq f(x) < \frac{k}{2^m}\right\}\right)}$$

and the integral is defined as $\lim_{m \to \infty} LS_m$.

This definition is of course not more general than the previous one <u>unless</u> it can be used for a richer class of functions than before.

- How should we define the class of functions $\mathcal{M}$ for which such a definition works?

- The characterization of $\mathcal{M}$ is intrinsically related to the definition of "measure".



Notice that, for this particular f,

$$\text{measure}\left(\left\{x : \frac{k-1}{2^m} \le f(x) < \frac{k}{2^m}\right\}\right) =$$

$$= \text{measure}(A_1 \cup A_2 \cup A_3 \cup A_4) \equiv \text{measure}(A).$$

- Connections with statistics:
- If we have $X \sim f(x)$, compute $E(g(X))$.
- If we have $X \sim \mu$, compute $E(g(X))$.

$\int g(x) f(x) dx$, $\int g(x) d\mu(x)$, $\int g(x) dF(x)$.

- Thus, we can start hoping that we can define the integral in this sense if we can assign a number to the "measure". As in the example, we need to measure a <u>set</u> A that, in turn, was obtained through an operation with sets. This will require the definition of a "set of sets", closed under some operations <u>the σ-field</u>. Then, we need to define as "measure" a set function (that is, a function whose argument is a set) that has some desirable properties (for example, it would allow us to compute the "measure" of A as a sum of the "measures" of $A_1, \ldots, A_4$).
- After all this is defined, we can then define the larger class $\mathcal{M}$ as the set of functions for which we can "measure" pre-images

as in the example. This will lead to the concept of <u>measurable functions</u>.

2) <u>Probability</u>

Given the sample space Ω, the traditional definition of probability, P, is:

1) $P(\emptyset) = 0, \quad P(\Omega) = 1$
2) For disjoint and countably many subsets of Ω, $A_1, A_2, \ldots$ we have
$P(\bigcup_i A_i) = \sum_i P(A_i)$ (countable additivity).

So, clearly $P : \mathcal{A} \subseteq \Omega \to [0, 1]$, where $\mathcal{A}$ is a collection of sets.

How big can $\mathcal{A}$ be such that 2 holds?

Can we take $\mathcal{A} = \mathcal{P}(\Omega) =$ all subsets of Ω?

<u>NO</u>: We cannot have the domain of P to be the entire $\mathcal{P}(\Omega)$ and still have countable additivity!

why

Question 1: How can we define the domain of P such that 2) holds.

Answer $\to$ σ- fields.

Question 2 How far can we go?
How can we construct measures on "complicated" spaces based on measures on "familiar" spaces? $\to$ Outer measures
Carathéodory extension theorem.

Everything

Lecture 2

Set Theory – notation

- A^c, $A \cap B = AB$, $A \cup B$,
- $A \cup B = A + B$
- $\bigcup_{i=1}^{n} A_i = \sum_{i=1}^{n} A_i$ } for disjoint sets.

- Ω = sample space.
- $\mathcal{P}(\Omega)$, 2^{Ω} = set of all possible subsets.
- $1_A(x)$ = the indicator of a set $= \begin{cases} 1, & \text{if } x \in A \\ 0, & x \notin A \end{cases}$.

- $A \setminus B \equiv A \cap B^c$
- $A \Delta B = A^c B \cup A B^c$ (the symmetric difference).
- A sequence of sets A_n is called <u>increasing</u> if $A_n \subset A_{n+1}$ $\forall n \geq 1$. $A_n \uparrow$
- A collection (class) of sets will be usually denoted by "curly" letters, i.e. $\mathcal{A}$.

Definition 1

A class $\mathcal{A}$ (nonvoid) of subsets A of a nonvoid set Ω is called a <u>σ- field</u> if:

a) If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$.

b) If $A_1, \dots, A_n, \dots$ is a countable collection of sets in $\mathcal{A}$, then $\bigcup_n A_n \in \mathcal{A}$.

<u>Remarks</u>: If $\mathcal{A}$ is a σ- field, then necessarily $\emptyset$ and Ω are in $\mathcal{A}$. Why? Done

<u>Remark 2</u>: If $\mathcal{A}$ is closed under complements and <u>finite</u> unions, $\mathcal{A}$ is called a <u>field</u>.

<u>Definition 2</u> $\mathcal{A}$ will be called a monotone class provided it contains $\bigcup_{n=1}^{\infty} A_n$ for all $\uparrow$ seq. A_n in $\mathcal{A}$ and it contains $\bigcap_{n=1}^{\infty} A_n$ for all $\downarrow$ A_n.

<u>Definition 3</u> $(\Omega, \mathcal{A})$ will be called a measurable space if $\mathcal{A}$ is a σ- field of subsets of Ω.

<u>Remark</u>: In the literature you may encounter "algebra" for field.

Remark If $\mathcal{A}$ is a field (or a σ-field), it is also closed under intersections. Why? Done

Prop 1 Arbitrary int. of fields, σ-fields, or monotone classes are fields, σ-fields or monotone classes, respectively.

Check!

Remark: Note that countable unions of σ-algebras need not be σ-algebras. See Ex. 1, Hwk 1.

Let $\mathcal{C}$ be a class of subsets of Ω.
Then, the minimal σ-field generated by $\mathcal{C}$ (that is, containing $\mathcal{C}$) is:

$$\sigma[\mathcal{C}] \equiv \bigcap \{ \mathcal{F}_\alpha : \mathcal{F}_\alpha \text{ is a } \sigma\text{-field of subsets of } \Omega \text{ for which } \mathcal{C} \subset \mathcal{F}_\alpha \}.$$

See Ex. 2, Hwk 1. intersection of all σ-fields containing $\mathcal{C}$

Example. Let $\Omega \equiv \{a, b, c, d, e\}$.

Let $\mathcal{C} = \{E_1, E_2\}$; $E_1 = \{a, b, c\}$, $E_2 = \{c, d,$

Find $\sigma[\mathcal{C}]$.

Solution (Notice that, of course, $\mathcal{C}$ itself is not a σ-field).

For such a simple example, the construction of $\sigma[\mathcal{C}]$ can be done by "completing" the class $\mathcal{C}$ to a σ-field, by adding to it subsets of Ω until Definition 1 is satisfied. Applying this process we get the following additional sets:

$$F_1 = \{a, b\} = E_1 \cap E_2^c, \quad F_2 = \{d, e\} = E_1^c \cap E_2$$

$$F_3 = \{c\} = E_1 \cap E_2, \quad F_4 = \{a, b, d, e\} = F_1 \cup F_3,$$

$\emptyset$ and Ω.

Thus $\sigma[\mathcal{C}] = \{\emptyset, F_1, F_2, F_3, F_1 \cup F_3, F_1 \cup F_2 = E_1, F_2 \cup F_3 = E_2, \Omega\}$.

NB For infinite σ-fields, we cannot list all the members. A more formal approach is needed, based on understanding the def. of $\sigma[\mathcal{C}]$. See also Ex. 3 Hwk1

Prop. 2

$\mathcal{A} \subset \Omega$ is a σ-field $\iff$ $\mathcal{A}$ is a field and a monotone class.

Solution: "$\Leftarrow$" We only need to show that $\mathcal{A}$ is closed under countable unions.

Notice that $A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots = \bigcup_{n=1}^{\infty} A_n =$

$$= \bigcup_{n=1}^{\infty} \left(\bigcup_{k=1}^{n} A_k \right) \equiv \bigcup_{n=1}^{\infty} B_n ,$$

where $B_n = \bigcup_{k=1}^{n} A_k$. Now, for each n, $B_n \in \mathcal{A}$ ($\mathcal{A}$ is a field). Also, $B_n = \bigcup_{k=1}^{n} A_k \subset B_{n+1} = \bigcup_{k=1}^{n} A_k \cup A_{n+1}$, thus $\{B_n\}_n \uparrow$, and the conclusion follows since $\mathcal{A}$ is a monotone class.

"$\Rightarrow$" check!

$\mathcal{A}$ is a σ-field

$\therefore$ $\mathcal{A}$ is closed under complements

$\mathcal{A}$ is closed under countable unions $\Rightarrow$ $\mathcal{A}$ is closed under finite unions

$\therefore$ $\mathcal{A}$ is a field

Also for any $B_n \in \mathcal{A}$ st $\{B_n\}_n \uparrow \Rightarrow \bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$ [$\mathcal{A}$ closed under countable unions]

$\therefore$ by def $\Rightarrow$ $\mathcal{A}$ is also a monotone class

<u>Definition 4</u> (Measures). — Definition 1.2 page 4.

<u>Proposition 1.2</u> (The monotone property of measures.. page 6.

Proof Needed

Proposition 3 (No proof)

(Continuity of measures: Gives conditions under which a finitely additive measure is also a countably additive measure).

If a finitely additive measure μ on either a field or a σ-field is either:

a) Continuous from below (i.e. $\mu(\lim A_n) = \lim(\mu A_n)$, for any $A_n \uparrow$)

or

b) Has $\mu(\Omega) < \infty$ and is continuous from above at $\emptyset$ (i.e. for any $A_n \downarrow \emptyset$, we have $\mu(\lim A_n) = \underset{\substack{\| \\ 0}}{\mu(\emptyset)} = \lim \mu(A_n)$),

then it is a countably additive measure.

with at least one $\mu(A_n) < \infty$.

does this imply $\mu(\Omega) < \infty$?

Proof: In either case we need to show that

$$\mu\left(\sum_1^\infty A_n\right) = \sum_1^\infty \mu(A_n).$$

a) $$\mu\left(\sum_1^\infty A_n\right) = \mu\left(\lim_{n\to\infty} \sum_{k=1}^n A_k\right) \underset{\boxed{\sum_{k=1}^n A_k \uparrow}}{=} \lim \mu\left(\sum_{k=1}^n A_k\right)$$

$$\underset{\uparrow}{=} \lim_n \sum_{k=1}^n \mu(A_k) = \sum_{k=1}^\infty \mu(A_k) \quad //$$

finite additivity

b) $$\mu\left(\sum_1^\infty A_n\right) \underset{\uparrow}{=} \mu\left(\sum_{k=1}^n A_k\right) + \mu\left(\sum_{n+1}^\infty A_k\right) =$$

finite a.

$$\underset{\uparrow}{=} \sum_{k=1}^n \mu(A_k) + \mu\left(\sum_{n+1}^\infty A_k\right)$$

finite a

Now: Note that $\mu\left(\sum_1^\infty A_n\right) \leq \mu(\Omega) < \infty$.

Also $\sum_{n+1}^\infty A_k \searrow \emptyset$ (since it is the "rest" of a convergent series)

$\Downarrow$

By cont. from above at $\emptyset \Rightarrow \lim_{n\to\infty} \mu\left(\sum_{n+1}^\infty A_k\right) = \mu(\emptyset) = 0$

q.ed.

Everything

Lecture 3' – The Lebesgue measure on $\mathbb{R}$

- <u>Definition</u> (It will be formally given in Lecture 4')
For $A \subset \mathbb{R}$ define the <u>Lebesgue measure</u> of A as $\lambda(A)$ = length of A.

- Let $I = \{ (a,b], (-\infty, b], (a,\infty) : \text{any } a, b \in \mathbb{R} \}$.

- <u>NOTE</u>: I is <u>NOT</u> a field (Why?) [Unions not included in I]

- Let $\mathcal{B}_I$ be the collection of sets consisting in <u>all</u> <u>finite disjoint</u> unions of elements in I.

- <u>NOTE</u>: $\mathcal{B}_I$ <u>is</u> a field.

- <u>Definition</u> $\mathcal{B} \equiv \sigma[\mathcal{B}_I]$ is called "the Borel sets of $\mathbb{R}$".

- NOTE : The following are Borel sets (that is, belong to $\mathcal{B}$).

a) (a,b) b) (a,∞) c) $(-\infty,a)$

d) $[a,b]$ e) $\{a\}$ f) Any finite set

g) any countable set h) The set $\mathbb{N}$ of natural #'s.

i) The set $\mathbb{Q}$ of rational #'s j) The set $\mathbb{R}\setminus\mathbb{Q}$ of irrational numbers.

k) $[a,\infty)$ l) $[a,b)$ m) $(a,b]$

n) The set $\mathbb{Z}$ of integer numbers.

- Define now a measure on $\mathcal{B}_I$; call it λ.

For each $A \in \mathcal{B}_I$, define

$$\lambda(A) = \sum_{j=1}^{J_0} \lambda(A_j) \; ; \quad \text{where } A = \bigcup_{j=1}^{J_0} A_j \text{ (disjoint union)},$$

and λ is the length. (We'll show in the next lecture that this is indeed a measure on $\mathcal{B}_I$). We continue to call this λ the Lebesgue measure on $\mathcal{B}_I$.

- The Carathéodory extension theorem (Fundamental

A (probability) measure defined on a field $\mathcal{C}$ has an unique extension to the generated σ-field $\mathcal{F} = \sigma[\mathcal{C}]$.

- A measure on $\mathcal{B}$, the Borel sets of $\mathbb{R}$.

By the above extension theorem, we can thus extend the Lebesgue measure from $\mathcal{B}_I$ to $\mathcal{B}$; we still call this extension the Lebesgue measure.

- Technical remark: The unique extension exists when the measure is σ-finite. We will only use the theorem for probability measures, which are σ-finite by definition.

- So far we have created the following triple $(\mathbb{R}, \mathcal{B}, \lambda)$, where $\mathcal{B} \subset 2^{\mathbb{R}}$.

NOTE: Alternative characterizations of $\mathcal{B}$.

a) If $\mathcal{A}$ is the family of all open intervals, then $\sigma[\mathcal{A}] = \mathcal{B}$.

b) If $\mathcal{C}$ is the family of all closed intervals, then $\sigma[\mathcal{C}] = \mathcal{B}$.

$*$

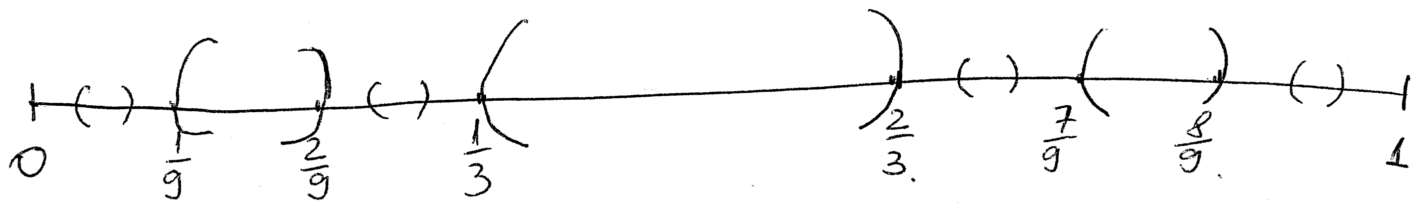
You can generate the same σ-field using diff generators.

- Question: Can we extend λ to all possible subsets of $\mathbb{R}$ (that is, to $2^{\mathbb{R}}$)?

Answer: NO (See Lecture notes 3, 4 and "Summary" for a very rigorous proof of this fact).

An example: The Cantor set.

- We begin with the the interval $[0,1]$. We then remove the open interval consisting in the middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$.



We then remove the open middle thirds of each of the two pieces, i.e. we remove $\left(\frac{1}{9}, \frac{2}{9}\right)$ and $\left(\frac{7}{9}, \frac{8}{9}\right)$. We then remove the four open middle thirds $\left(\frac{1}{27}, \frac{2}{27}\right)$, $\left(\frac{7}{27}, \frac{8}{27}\right)$, $\left(\frac{19}{27}, \frac{20}{27}\right)$ and $\left(\frac{25}{27}, \frac{26}{27}\right)$ of the remaining pieces.

- We continue inductively, at the n^{th} stage removing the 2^{n-1} middle thirds of all remaining sub-intervals, each of length $\frac{1}{3^n}$. The Cantor set is everything that is left over, after we have removed all these middle thirds.

1) We verify first that the Cantor set C is a Borel set (that is, it is in $\mathcal{B}$).

Let $C_0 = [0,1]$

$C_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$

$C_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{8}{9},1]$

. . . .

Then (check)

$C = \bigcap_{n=1}^{\infty} C_n$. Thus

$C^c = \bigcup_{n=1}^{\infty} C_n^c$; $\left.\begin{array}{l} C_n^c \in \mathcal{B} \text{ for all } n \\ \mathcal{B} \text{ is a } \sigma\text{-field} \end{array}\right\} \Rightarrow$

$C^c \in \mathcal{B} \Rightarrow (C^c)^c = C \in \mathcal{B}$, again since $\mathcal{B}$ is a (σ)-field.

2) We can compute $\lambda(C)$. It is zero (see Hwk 1).

Conclusion: There exists uncountable sets in $\mathcal{B}$, with Lebesgue measure 0. (We already know that the Lebesgue measure of a countable set is zero). How do you prove this!

NOT REQ

Lecture 3 Outer measures; The Caratheodory extension theorem.

Definition 1 (Outer measures).

Let $\mu^* : 2^\Omega \longrightarrow [0, \infty]$.

If μ^* satisfies the following prop. then μ^* is called an $\boxed{\text{outer}}$ measure :

Null: $\mu^*(\emptyset) = 0$

Monotone: $\mu^*(A) \leq \mu^*(B)$, for all $A \subset B$.

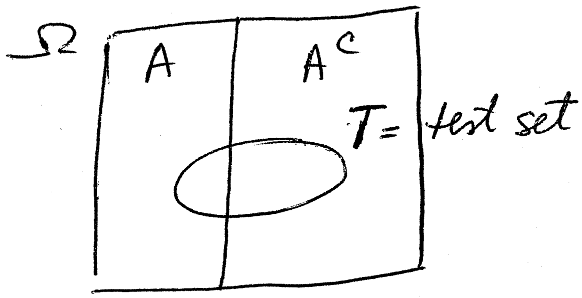
Countable subaditivity: $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

<u>Definition 2</u> $A \subset \Omega$ is called <u>μ^*-measurable</u>

if (1) $\mu^*(T) = \mu^*(TA) + \mu^*(TA^c)$, for all

subsets $T \subset \Omega$.

The sets T are called test sets.



<u>NB</u> Since μ^* is an outer measure, then by c.s-a.

we always have "$\leq$" in (1). $\left\{ \begin{array}{l} T = (TA) \cup (TA^c) \\ \mu(T) = \mu(TA \cup TA^c) \leq \mu(TA) + \mu(TA^c) \end{array} \right.$

<u>Definition 3</u> $\mathcal{A}^* \equiv \{ A \in 2^\Omega \mid A \text{ is } \mu^* \text{ measurable} \}$

<u>Note</u>: $A \in \mathcal{A}^* \iff \mu^*(T) \geq \mu^*(TA) + \mu^*(TA^c)$ for all $T \subset \Omega$.

Definition 4 (The outer extension of a measure μ).

Let Ω be arbitrary. Let μ be a measure on a field $\mathcal{C}$ of subsets of Ω. For each $A \in 2^{\Omega}$ we define:

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A \subset \bigcup_{n=1}^{\infty} A_n \text{ with } \underline{\text{all}}\ A_n \in \mathcal{C} \right\}.$$

- Then μ^* is called the outer extension of μ.
- $A_1, \dots, A_n, \dots$ are called Carathéodory coverings.

(We'll see later that μ^* is an outer measure on $(\Omega, 2^{\Omega})$ and that, moreover, it is a measure on $\mathcal{A}^*$).

Theorem (Carathéodory extension theorem).

- A measure μ on a field $\mathcal{C}$ can be extended to a measure on a σ-field $\sigma[\mathcal{C}]$, generated by $\mathcal{C}$, by defining:

$$\mu(A) = \mu^*(A) \quad \text{for each } A \in \mathcal{A} \equiv \sigma[C].$$

- Moreover, if μ is σ-finite on $\mathcal{C}$, then the extension is unique on $\mathcal{A}$ and is also σ-finite.

Comments: This theorem allows the construction of measures on "rich" classes as generalizations of measures on simpler classes.

Note We have denoted the measure on $\mathcal{C}$ and its extension on $\sigma(\mathcal{C})$ by the same letter, "μ".

Proof

Claim 1 μ^* is an outer measure on $(\Omega, 2$

Claim 2 a) $\mu^*|\mathcal{C} = \mu$ (That is, $\mu^*(C) = \mu(C)$ for all $C \in \mathcal{C}$).

b) $\mathcal{C} \subset \mathcal{A}^*$

Claim 3 $\mathcal{A}^*$ is a field.

Claim 4 μ^* is a f.a. measure on $\mathcal{A}^*$.

Claim 5 a) $\mathcal{A}^*$ is a σ-field.

b) $\mathcal{A}^* \supseteq \sigma(\mathcal{C})$

Claim 6 μ^* is c.a. on $\mathcal{A}^*$.

Claim 7 If μ is finite, then its extension is unique.

Claim 8 If μ is σ-finite, then the extension is unique.

Q.e.d.

Proof of Claim 1

<u>Null</u> $\mu^*(\phi) \equiv \inf\left\{ \sum_{n=1}^{\infty} \mu(A_n),\ \phi \subset \bigcup_{n=1}^{\infty} A_n,\ A_n \in \mathcal{C} \right\}$

Since $\{\phi, \phi, \dots\}$ is a covering of ϕ, $\phi \in \mathcal{C}$ and μ is a measure on $\mathcal{C}$, hence $\mu(\phi) = 0$, then $\mu^*(\phi) = 0$.

<u>Monotone</u> We need to show that for $A \subset B$ we have $\mu^*(A) \leq \mu^*(B)$.

Let $\{A_n\}_n$ be a covering for A and $\{B_n\}_n$ a covering for B.

Note that since $A \subset B \Rightarrow A \subset \bigcup_n B_n \Rightarrow$

$$\{\text{coverings for } A\} \supset \{\text{coverings for } B\}. \quad (1)$$

Now, $\mu^*(A) = \inf\limits_{\{\text{coverings of } A\}} \left\{ \sum_{n=1}^{\infty} \mu(\text{elements in the coverings for } A \right.$

$\mu^*(B) = \inf\limits_{\{\text{cov. of } B\}} \left\{ \sum_{n=1}^{\infty} \mu(\text{covering of } B) \right\}$.

Recall also that, in general, if $E \subset F \Rightarrow \inf E \geq \inf F$ (2)

Thus, by (1) and (2): $\mu^*(A) \leq \mu^*(B)$.

Countably subadditive

Recall the characterization of the infimum of a set

$m = \inf A \iff$ a) $m \leq x$, for all $x \in A$.

b) $\forall \varepsilon > 0 \quad \exists \; x_\varepsilon \in A$ s.t. $m \geq x - \varepsilon \iff x < m + \varepsilon$

We need to show that:

$$\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n), \text{ for all } A_n \in \Omega \text{ needed}$$

<u>Idea</u> We show that for an arbitrary $\varepsilon > 0$ we have

(*) $\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n) + \varepsilon$, and then take limit over $\varepsilon \to 0$ which yields the result.

<u>Show now</u> (*): Let $A_n \in \Omega$ be arbitrary.

By the definition of μ^* and (10) we have:

$\forall \varepsilon' > 0 \quad \exists$ a covering $\{A_{nk} : k \geq 1\}$ of A_n

(that is $A_n \subset \bigcup_{k=1}^{\infty} A_{nk}$) such that
$\qquad\qquad\qquad\qquad\qquad \uparrow$
$\qquad\qquad\qquad\qquad\qquad \mathcal{C}$

$$\sum_{k=1}^{\infty} \mu(A_{nk}) \leq \mu^*(A_n) + \varepsilon'. \quad (11)$$

Since we can write this for any ε', we take

$\boxed{\varepsilon' = \frac{\varepsilon}{2^n}}$ for some $\varepsilon > 0$.

• Now, since $A_n \subset \bigcup_{k=1}^{\infty} A_{nk} \Rightarrow$

$\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{nk}$. Also, we have μ^* monotone (we have just proved it above). So:

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \mu^*\left(\bigcup_n \bigcup_k A_{nk}\right)$$

<u>Now</u>, note that we can read the definition of μ^* as follows:

$$\mu^*(B) = \inf_{\text{coverings}} \{ \text{sum}(\mu(\text{covering elements})),\ B \subset \text{covering} \}.$$

~~Hence~~, $\mu^*(B) \leq \text{sum}(\mu($~~a~~ elements of a <u>particular</u> covering ~~of~~ $B))$,

Since $\mu^*(B)$ is the infimum over such sums over covering elements.

• Noticing then that $\bigcup_n \bigcup_k A_{nk} \subseteq \bigcup_n \bigcup_k A_{nk}$, that is noting that the $\{A_{n,k}\}_{k,n}$'s form a covering for $\bigcup_n \bigcup_k A_{nk}$, we thus have

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_n \sum_k \mu(A_{nk})$$

$$\leq \sum_n \left(\mu^*(A_n) + \frac{\varepsilon}{2^n}\right)$$

By 11

$$= \sum_n \mu^*(A_n) + \varepsilon .$$

So, for $\varepsilon \to 0$ we have the desired result.

$\blacksquare$

Proof of Claim 2

a) $\mu^*|_{\mathcal{C}} = \mu$ //

We show that $\left\{ \begin{array}{l} \mu^*(C) \leq \mu(C) \\ \quad \text{and} \\ \mu^*(C) \geq \mu(C) \end{array} \right.$ for all $C \in \mathcal{C}$.

Let $C \in \mathcal{C}$.

Notice then that $C, \emptyset, \emptyset, \ldots, \emptyset, \ldots$ is always a covering for C; call it $C_1, C_2, \ldots, C_n, \ldots$

Notice also that, for this covering

$$\sum_{n=1}^{\infty} \mu(C_n) = \mu(C).$$

Recall that

$$\mu^*(C) = \inf \underbrace{\left\{ \sum_{n=1}^{\infty} \mu(A_n) ;\ A \subset \bigcup_n A_n,\ A_n \in \mathcal{C} \right\}}_{\equiv\, B}$$

Hence $\left. \begin{array}{l} \mu(C) \in B \\ \quad \text{and} \\ \mu^*(C) = \inf B \end{array} \right\} \Rightarrow \mu^*(C) \leq \mu(C).$

Show now $\mu(C) \le \mu^*(C)$.

Let $A_1, \ldots, A_n, \ldots$ be any covering of C.
Notice that $C = \bigcup_{n=1}^{\infty} (A_n \cap C) \in \mathcal{C}$.

Then:

$$\mu(C) = \mu\left(\bigcup_{n=1}^{\infty} (A_n \cap C)\right) \le \sum_{n=1}^{\infty} \mu(A_n \cap C)$$

Prop. 1.2 page 6, since the union is in $\mathcal{C}$

$$\le \sum_{n=1}^{\infty} \mu(A_n).$$

Thus, taking inf both sides, since the covering was arbitrary we have $\mu(C) \le \mu^*(C)$.

b) $\mathcal{C} \subset \mathcal{A}^*$

We show that any $C \in \mathcal{C}$ is in $\mathcal{A}^*$, that is we show that any $C \in \mathcal{C}$ is μ^*- measurable.

Let $C \in \mathcal{C}$.

Recall Definition 2 Page 2 (or Definition 1.3 page 4 PFS)

Then, we need to show that, for any $T \subset \Omega$, we have:
$$\mu^*(T) \geq \mu^*(TC) + \mu^*(TA^c). \quad //$$

Recall that μ^* is an infimum and recall the Th. of characterization (10) page 7.

. Let $\varepsilon > 0$ and T given.
By (10), $\exists$ a covering $\{A_n\}_{n=1}^{\infty} \subset \mathcal{C}$ of $T \in 2^{\Omega}$

Such that: $\mu^*(T) + \varepsilon \geq \sum \mu(A_n)$.

Write $A_n = CA_n + C^c A_n$ $\quad \Rightarrow$

Using the additivity of μ:

$$\mu^*(T) + \varepsilon \geq \sum \mu(CA_n) + \sum \mu(C^c A_n).$$
$$\geq \mu^*(CT) + \mu^*(C^c T)$$

using that $\{CA_n\}_n$ covers CT, $\{C^c A_n\}_n$ cover $C^c T$, and, again, that μ^* is an infimum.

Let now $\varepsilon \to 0$ and the proof is complete.

Proof of Claim 3

$\mathcal{A}^*$ is a field

(Recall that we have just proved, in Claim 2, that $\mathcal{A}^* \supset \mathcal{C}$).

We show:

1) $A \in \mathcal{A}^* \Rightarrow A^c \in \mathcal{A}^*$

2) $A, B \in \mathcal{A}^* \Rightarrow AB \in \mathcal{A}^*$

For 1): Since $A \in \mathcal{A}^* \Rightarrow \mu^*(T) = \mu^*(TA) + \mu^*(TA^c)$ (*)

for any T.

To show that $A^c \in \mathcal{A}^*$ we need to show that:

$$\mu^*(T) = \mu^*(TA^c) + \mu^*(T(A^c)^c) =$$

$$= \mu^*(TA^c) + \mu^*(TA)$$ which is identical to (*).

2) Again, a set G is in $\mathcal{A}^*$ if for all $T_1 \subset \Omega$, we have $\mu^*(T_1) = \mu^*(GT) + \mu^*(G^cT)$. (**)

Now, let $A, B \in \mathcal{A}^*$. We have:

$$\mu^*(T) = \mu^*(TA) + \mu^*(TA^c) \quad \text{By (**) applied to } A$$

$$= \mu^*(TAB) + \mu^*(TAB^c) + \mu^*(TA^cB) + \mu^*(TA^cB^c)$$

By (**) applied to B first with test set TA, then with test set TA^c

$$\geq \mu^*(TAB) + \mu^*(TAB^c + TA^cB + TA^cB^c)$$

Since, ~~by definit~~ we showed in Claim 1 that μ^* is countably subadditive

$$= \mu^*(TAB) + \mu^*(T(AB)^c). \text{ Thus,}$$

$AB \in \mathcal{A}^*$

q.e.d. this claim

NOT REG

Lecture 4

Carathéodory Theorem — cont

Proof of Claim 4 : See page 13, PfS. — μ^* is f.a. on $\mathcal{A}^*$

Proof of Claim 5 : $\mathcal{A}^*$ is a σ-field. — See page 13 pfS

b) $\mathcal{A}^* \supset \sigma[\mathcal{C}]$.

By Claim 3, $\mathcal{C} \subset \mathcal{A}^*$. Since $\mathcal{A}^*$ is a σ-field, then as in Exercise 1.1 page 4 PfS, $\sigma[\mathcal{C}] \subset \sigma[\mathcal{A}^*] = \mathcal{A}^*$

Proof of Claim 6 μ^* is c.a. on $\mathcal{A}^*$. — See page 14 PfS.

Claim 7 If μ is a finite measure, the extension is unique.

Recall that μ-finite measure means $\mu(\Omega) < \infty$.

Let μ_1 and μ_2 denote 2 extensions of μ to $\sigma[\mathcal{C}]$.

Recall that μ is a measure on the field $\mathcal{C}$.

- Let $\mathcal{M} \equiv \{ A \in \sigma[\mathcal{C}] : \mu_1(A) = \mu_2(A) \}$ (1)

We want to show that:

$$\mu_1(A) = \mu_2(A) \quad \text{for all } A \in \sigma[\mathcal{C}].$$

- We shall need

<u>Proposition 1.6 (page 10 PfS)</u>

Let $\mathcal{C}$ be a field. Then

$$\sigma[\mathcal{C}] = \mathcal{M}[\mathcal{C}]$$

The minimal σ-field generated by $\mathcal{C}$

The minimal monotone class generated by $\mathcal{C}$

- Notice now that:

$\mu_1(A) = \mu_2(A) = \mu(A)$, for any $C \in \mathcal{C}$, since both μ_1 and μ_2 are extensions of μ. $\Longrightarrow \mathcal{C} \subset \mathcal{M}$ (3)

- Now, by the above Proposition $\Longrightarrow$

$\sigma[\mathcal{C}] = \mathcal{M}[\mathcal{C}] =$ minimal monotone class containing $\mathcal{C}$. (4)

By (3) and (4) it is clear that;

if we prove that $\mathcal{M}$ is a monotone class,

then $\sigma[\mathcal{C}] \subset \mathcal{M}$, since $\sigma[\mathcal{C}]$ coincides with the individual monotone class containing $\mathcal{C}$.

And, since $\mathcal{M} \subset \sigma[\mathcal{C}]$ by definition, this would complete the proof.

$\mathcal{M}$ is a monotone class

Let $\{A_n\}_n$ be a monotone sequence of sets in $\mathcal{M}$ (say show that $\bigcup_n A_n = \lim A_n \in \mathcal{M}$.

Then, by Prop. 1.2 (and 1.4) $\Rightarrow$

$$\mu_1(\lim A_n) = \lim \mu_1(A_n)$$

$= \lim \mu_2(A_n)$ by the def. of $\mathcal{M}$ and $A_n \in \mathcal{M}$

$= \mu_2(\lim A_n)$ by Prop. 1.2 and 1.4.

Claim 8 Page 14 PfS.

$\boxed{\text{q.e.d. Theorem}}$

Conclusion – What have we achieved and how can we use it?

The Carathéodory extension theorem says that once we have defined a measure μ on a **field** $\mathcal{C}$ we can extend it to $\sigma[\mathcal{C}]$ by defining $\mu(A) = \mu^*(A)$, for all $A \in \sigma[\mathcal{C}]$.

- Also, notice that, in fact, μ^* is a c.a. measure on $\mathcal{A}^*$, which is a σ-field.
- When we extended μ from $\mathcal{C}$ to $\sigma[\mathcal{C}]$, can we actually go beyond $\sigma[\mathcal{C}]$?

Example We exemplify the previous ~~conts~~ construction and indicate further problems for $\Omega = \mathbb{R}$.

Definition 1 (It will be given formally in Lecture 5).
For $A \subset \Omega$ define the Lebesgue measure of A as $\lambda(A) =$ length of A.

- We give now an example of a σ- field $\subset 2^{\mathbb{R}}$ on which λ acts "well".

Example (Very important) Borel sets in $\mathbb{R}$.

- Let $I = \{ (a,b], (-\infty, b], (a, \infty) : a, b \in \mathbb{R} \}$.
  - Note that I IS NOT a field.
- Let B_I the collection of sets consisting in all finite disjoint unions of elements in I.
  - Note that B_I IS a field.

- Let $\mathcal{B} \equiv \sigma[B_I]$.

$\mathcal{B}$ is called "the Borel sets of $\mathbb{R}$".

- Define now a measure on B_I:

For each $A \in B_I$, define $\mu(A) = \sum_{j=1}^{J_0} \lambda(A_j)$, (15)

where $A = \bigcup_{j=1}^{J_0} A_j$ (disjoint union), and λ is the length. We'll show that this is indeed a measure on B_I in Lecture 5.

- How can I use μ to define a measure on $\mathcal{B}$.

Use the extension theorem:

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n),\ A \subset \bigcup_{n=1}^{\infty} A_n,\ A_n \in B_I \right\}.$$

Then, extend μ to $\mathcal{B}$ (and still define it μ)

$$\mu(A) = \mu^*(A), \text{ for all } A \in \sigma[\mathcal{C}].$$

NOTE (15) is NOT the only way we can define a measure on B_I. Since the definition of μ^* depends on the one of μ, then, of course, there are other ways to define measures on $\mathcal{B}$ $\longrightarrow$ See Lecture 5.

• We now return to the question:

Can we define ~~(extend)~~ μ beyond $\mathcal{B}$?

We need an extra definition and an additional result.

<u>Definition</u> (note that it works in general, not only for the particular μ above) (see page 15 Pf S).

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space.
If $\mu(A)=0$, then we call it a null set.
We call $(\Omega, \mathcal{A}, \mu)$ $\boxed{\text{complete}}$ if whenever
$B \subset A \in \mathcal{A}$ for some A with $\mu(A)=0$ then
$B \in \mathcal{A}$ (and, trivially, $\mu(B)=0$).

<u>Note</u>: One can always construct an $(\Omega, \mathcal{A}, \mu)$ ~~such that~~ and an $A \in \mathcal{A}$ s.t. $\left.\begin{array}{l}\mu(A)=0 \\ B \subset A\end{array}\right\}$ <u>but</u>

$B \notin \mathcal{A}$.

Example Let $\Omega = \{1, 2, 3, 4, 5\}$.

Let $\mathcal{C} = \{ \{1,2,3\}, \{3,4,5\} \}$.

Then $\sigma[\mathcal{C}] = \{ \emptyset, \{1,2\}, \{4,5\}, \{3\}, \{1,2,4,5\},$
$\{1,2,3\}, \{3,4,5\}, \{1,2,3,4,5\} \}$

Define $\underbrace{\mu(A) = \delta_4(A)}_{\text{The discrete measure}} = \begin{cases} 1, & \text{if } 4 \in A \\ 0, & \text{otherwise} \end{cases}$.

So: $\mu(\{1,2\}) = 0$ BUT $\{1\} \notin \sigma[\mathcal{C}]$.

$\{1\} \subset \{1,2\}$

~~Definition 2~~

Proposition (See Hwk 2 and page 15 Pf S).

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

- Let $\hat{\mathcal{A}}_\mu \equiv \{ A \cup N \mid A \in \mathcal{A}$ and $N \subset (\text{some } B) \in \mathcal{A}$ with $\mu(B) = 0 \}$.
- Define $\hat{\mu}(A \cup N) = \mu(A)$ for all $A \in \mathcal{A}$ and for all $N \subset B \in \mathcal{A}$ with $\mu(B) = 0$.

Then $(\Omega, \hat{\mathcal{A}}_\mu, \hat{\mu})$ is a complete measure space and $\hat{\mu}|_{\mathcal{A}} = \mu$.

__Corollary 1__ $\hat{\mu}$ is the unique extension of $\mu \in \hat{\mathcal{A}}_\mu$.

__Corollary 2__ (Follows directly from the Carathéodory extension theorem and Corollary 1).

Let $\mathcal{C}$ be a field.

Let μ be a σ-finite measure on $\mathcal{C}$.

__Then__: a) The extension of μ from $\mathcal{C}$ to $\sigma[\mathcal{C}]$ is unique. (Carathéodory)

b) The further extension of μ from $\sigma[\mathcal{C}]$ to $\hat{\mathcal{A}}_\mu \equiv \hat{\sigma}[\mathcal{C}]_\mu$ is unique. (Corollary 1).

c) $\hat{\mathcal{A}}_\mu = \hat{\sigma}[\mathcal{C}]_\mu \subset \mathcal{A}^*$ = the collection of μ^* measurable sets.

__Proof__ Pages 15-16 PfS.

We prove here only Corollary 1.

- Recall that $\circledast \begin{cases} \hat{\mu}(A \cup N) = \mu(A) \text{ for all} \\ A \in \mathcal{A} \text{ and all } N \subset B \in \mathcal{A} \text{ with } \mu(B \end{cases}$

- Let ν be some other <u>extension</u> of μ.

We show that $\circledast$ holds with $\hat{\mu}$ replaced by ν, and thus $\hat{\mu} = \nu$.

<u>R. A.</u> Assume that $\exists\, A \in \mathcal{A}$ and $\exists\, N \subset B \in \mathcal{A}$ with $\mu(B) = 0$ such that

$$\nu(A \cup N) > \mu(A)$$

(Note that $A \subset A \cup N$ so we always have

$\nu(A \cup N) \geq \nu(A) = \mu(A)$, since ν is an extension)

<u>Thus</u>

$$\mu(A) < \nu(A \cup N) = \nu(A \cup (A^c N))$$

with $A^c N \subseteq \underbrace{A^c B}_{\text{a null set}}$

$$\underset{\substack{\uparrow \\ \nu \text{ measure on the completion}}}{=} \nu(A) + \nu(A^c N) \leq \nu(A) + \nu(B) =$$

$= \mu(A) + \mu(B)$, since ν extension and $A, B \in \mathcal{A}$.

Hence $\mu(B) > 0$, which contradicts $\mu(B) = 0$

q.e.d.

• Let us return now to our example.

Recall that we had μ defined on $\mathcal{B}$, at pages

Let us redenote μ by λ.

<u>Definition</u> The completion of the Lebesgue measure ~~on~~ $(\mathbb{R}, \mathcal{B}, \lambda)$ is still called (and denoted) Lebesgue measure (λ).

The "completed" family $\hat{\mathcal{B}}_\lambda$ is called the Lebesgue ~~measur~~ sets.

Note that $\hat{\mathcal{B}}_\lambda \subset 2^{\mathbb{R}}$. Can we go now beyond $\hat{\mathcal{B}}_\lambda$ in uniquely defining a measure on $2^{\mathbb{R}}$. No!

(See Prop. 2.1 page 16 in PfS).

<u>Example</u> There exists $D \in 2^{\mathbb{R}}$ that is not in $\hat{\mathcal{B}}_\lambda$. (Thus, not all subsets of $\mathbb{R}$ are measurable with respect to the Lebesgue measure (even if we complete it)).

<u>Proof</u> For any $x, y \in [0,1)$ define $x \sim y$ if $x - y \in \mathbb{Q}$. This defines a set of equivalence classes. Let D be the set that contains one (and only one) element of each equivalence set.

Write now: $\cancel{D_z}$ $$[0,1) = \sum_{z \in \mathbb{Q} \cap [0,1)} D_z, \quad (1)$$ where

$$D_z = \{ \underbrace{z + x \ (\text{modulo } 1)}_{\text{The rest of the division of } z+x \text{ by } 1} \mid x \in D \}$$

Notice that for different z's, D_z's are disjoint.

All D_z's have the same outer measure.

Note that $\lambda^*(D_z) = \lambda^*(z + D) = \lambda^*(D)$

λ^* here denotes the

outer extension of λ^*.
Denote $\lambda^*(D)$ by a.

Assume D is measurable, that is $\lambda^*(D) = \lambda(D) = a$.

Then, by ①:

$$\underline{1} = \lambda([0,1)) \underset{\substack{\uparrow \\ \text{c. additivity}}}{=} \sum_{z} \lambda(D_z) = \sum_{z} a = \begin{cases} 0, & \text{when } a = 0 \\ \infty, & \text{when } a > 0 \end{cases}$$

which leads to a contradiction.

Hence $D \notin \hat{B}_\lambda$.

Summary on the possible extensions of the Lebesgue measure on $(\mathbb{R}, \circ)$.

- We had B_I = field (the collection of sets consisting in all finite disjoint unions of elements in I defined at page 5, Lecture 4).

- We defined a measure on B_I. At page 6 we called it μ and it was
$$\mu(A) = \sum_{j=1}^{J_0} \lambda(A_j) \; ; \; A = \sum_{j=1}^{J_0} A_j .$$
λ was called the Lebesgue measure (to be rigorously defined in Lecture 5).

NB From now on, by abuse of notation, we will also denote μ by λ.

(1) Thus: we have λ, the Lebesgue measure, defined on B_I.

(2) By Carathéodory's Th.: We extend λ to $B = \sigma[B_I]$.
We continue to denote the extension by λ.
Thus: we now have a measure, λ, still called the Lebesgue measure, defined on B, the Borel sets of $\mathbb{R}$.

(3) We construct now $\widehat{B}_\lambda$ = the "completion of B" with respect to λ.
$\widehat{B}_\lambda$ is called "the Lebesgue sets of $\mathbb{R}$".

We further extend λ from ② above to $\widehat{\mathcal{B}}_\lambda$. This extension is called "the completion" of λ, and we still denote it by λ.

The picture is then:
$$(\mathbb{R}, \mathcal{B}_I, \lambda)$$
$$\Downarrow \text{extended}$$
$$(\mathbb{R}, \mathcal{B}, \lambda)$$
$$\Downarrow \text{extended}$$
$$(\mathbb{R}, \widehat{\mathcal{B}}_\lambda, \lambda)$$

<u>Question</u>: Is $\widehat{\mathcal{B}}_\lambda = 2^{\mathbb{R}}$?

<u>Proposition</u>: $\widehat{\mathcal{B}}_\lambda \subset 2^{\mathbb{R}}$. (That is: there exists a set $D \in 2^{\mathbb{R}}$ such that $D \notin \widehat{\mathcal{B}}_\lambda$).

<u>Proof</u>. ~~Assume~~ In what follows we will construct a set D and show that $D \notin \widehat{\mathcal{B}}_\lambda$. Assume we constructed D.

- We prove that $D \notin \widehat{\mathcal{B}}_\lambda$ by contradiction. The idea is as follows. Assume $D \in \widehat{\mathcal{B}}_\lambda$.
  We proved at page 9 that $\widehat{\mathcal{B}}_\lambda \subset \mathcal{B}^*$.
  Thus $D \in \widehat{\mathcal{B}}_\lambda \Rightarrow D \in \mathcal{B}^*$.
- Recall now that if we have a measure λ on $\mathcal{B}_I$ we can define the outer measure λ^* ~~of~~ any set $D \in 2^{\mathbb{R}}$. Thus, we can always calculate $\lambda^*(D)$.

But: pay attention here, an <u>outer measure</u> on $2^{\mathbb{R}}$ is <u>NOT</u> <u>a measure</u> on $2^{\mathbb{R}}$.

However, we also proved that the outer measure λ^* on $2^{\mathbb{R}}$ is a <u>measure</u> on $\mathcal{B}^*$, and we always have $\lambda^*(A) = \lambda(A)$, for any $A \in \mathcal{B}$.

<u>We construct now D.</u>

- Consider the interval $[0,1)$ and define the equivalence relationship: $x \sim y$ iff $x - y \in \mathbb{Q} \cap [0,1)$, ~~for all $x, y \in [0,$~~ where $x, y \in [0,1)$.
- Let $\hat{x}$ be the equivalence class of $x \in [0,1)$, that is $\hat{x} = \{ y \in [0,1) \mid x - y \in \mathbb{Q} \cap [0,1) \}$.
- Choose now one (<u>and only one</u>) element y of this set and call it the representative of the class; denote it by r_x.

  Obviously, $x - r_x \in \mathbb{Q} \cap [0,1)$.
- Let $D = \{ r_x \,;\, x \in [0,1) \}$; $D_z = \{ z + r_x, \ r_x \in D \}$, for all $z \in \mathbb{Q} \cap [0,1)$.
- Claim $[0,1) = \sum_{z \in \mathbb{Q} \cap [0,1)} D_z$.

  We only show $[0,1) \subseteq \sum_{z \in \mathbb{Q} \cap [0,1)} D_z$.

Let $x \in [0,1)$. Need to show that $\exists z \in [0,1) \cap \mathbb{Q}$ s.t. $x \in D_z$. That is, show that $\exists z$ s.t. $x = z + r_x$. By the definition of D, define $z = x - r_x$ and we are done.

Assume now that this $D \in \widehat{\mathcal{B}}_\lambda$. Then $D \in \mathcal{B}^*$.

Notice also that if $D \in \widehat{\mathcal{B}}_\lambda$, we also have that $D_z \in \widehat{\mathcal{B}}_\lambda$. Thus $D_z \in \mathcal{B}^*$, for all $z \in \mathbb{Q} \cap [0,1)$.

Recall that $\mathcal{B}^*$ is a σ-field, and so

$$\lambda^*\Big(\bigcup_{z \in \mathbb{Q}\cap[0,1)} D_z\Big) = \sum_{z \in \mathbb{Q}\cap[0,1)} \lambda^*(D_z).$$

<u>Now</u>, $\lambda^*(D_z) = \lambda^*(z + D) = \lambda^*(D) \overset{\text{Notation}}{=} a$ $\Rightarrow$

using the definition of λ^* (with inf), and the fact that $\lambda(\{z\}) = 0$

$$1 = \lambda([0,1)) = \lambda^*\Big(\bigcup_{z \in \mathbb{Q}\cap[0,1)} D_z\Big) = \sum_{z \in \mathbb{Q}\cap[0,1)} a.$$

since λ^* coincides with λ on $\mathcal{B}$

<u>So</u>:
- If $a = 0 \Rightarrow 1 = 0$ contradiction!
- If $a > 0 \Rightarrow 1 = \sum_z a = \infty$ contradiction!

q.e.d.

REQ

Lecture 5 The Correspondence Theorem

Definition 1

A measure μ on $\mathbb{R}$ assigning finite values to finite intervals is called a Lebesgue-Stieltjes (LS) measure.

Definition 2

Let F be a finite function on $\mathbb{R}$.

If
a) F is increasing $\nearrow$
b) F is right continuous

then F is called a generalized df (gdf).

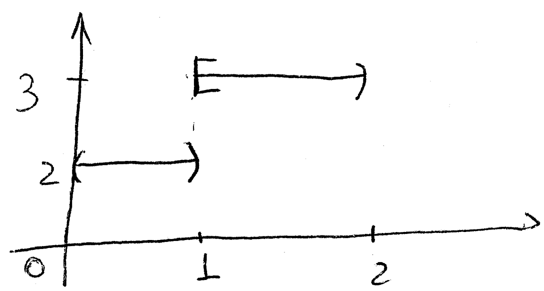
Remarks: • Recall that a function g is called right continuous at x_0 if $\lim_{\substack{x \to x_0 \\ x > x_0}} g(x) = g(x_0)$.

• We sometimes denote $\lim_{\substack{x \to x_0 \\ x > x_0}} g(x)$ by $\lim_{x \searrow x_0} g(x)$.

• Similarly, g is called left continuous if

$$\lim_{\substack{x \to x_0 \\ x < x_0}} g(x) \equiv \lim_{x \nearrow x_0} g(x) = g(x_0)$$

$$g(x) = \begin{cases} 2, & x \in (0,1) \\ 3, & x \in [1,2) \end{cases}$$



g is not continuous at 1, but it is right continuous.

- We denote $F_-(x_0) = \lim_{\substack{x \to x_0 \\ x < x_0}} F(x)$.

<u>Definition</u> 1) $\Delta F(\cdot) = F(\cdot) - F_-(\cdot)$ is called the mass function of F.

Notice that $\Delta F(x_0) = F(x_0) - F_-(x_0)$ is the "jump" of F at x_0.

2) $F(a,b] = F(b) - F(a)$ for all $a < b$ is called the increment function F.

<u>Note</u> $F_-(0) = 0$ for the representative gdf.

Theorem (Correspondence theorem).

- Let F be a finite function on $\mathbb{R}$, F right continuous, F increasing, $F_-(0) = 0$. (1)
- Let μ be a Lebesgue-Stieltjes measure on $(\mathbb{R}, \mathcal{B})$. (2)

Then

$$\mu((a,b]) = F(a,b] = F(b) - F(a), \quad (3)$$

for all $-\infty \le a < b \le +\infty$,

establishes a 1-1 corresp. between LS measures on $(\mathbb{R}, \mathcal{B})$ and the representative members of gdf's. $\bigstar$

Example: 1) If λ is the Lebesgue measure, $F(x) = x$

2) For the Dirac measure δ_{x_0}, we take $F(x) = 1_{[x_0,\infty)}(x)$ (Check!)

<u>Proof</u> $\rightarrow$ (Not required)

<u>Step 1</u> Let μ be a LS measure on $(\mathbb{R}, \mathcal{B})$.
We need to show that we can define, via (3), a function F that satisfies (1).

$$\text{Define } \quad F(x) = \begin{cases} \mu(\{0\}) - \mu((x,0]), & \text{for } x < 0 \\ \mu(\{0\}), & \text{for } x = 0 \\ \mu(\{0\}) + \mu((0,x]), & \text{for } x > 0 \end{cases} \qquad (4)$$

1) $F \uparrow$ Obvious, since $\mu(\{0\})$ and $\mu((x,0])$ are positive quantities.

2) <u>F is right continuous</u>
Only need to check at 0.
So, we need to check that

$$\lim_{\substack{x \to 0 \\ x > 0}} F(x) = F(0) = \mu(\{0\}).$$

~~Note that~~ $\cancel{\lim_{\substack{x \to 0 \\ x > 0}} F(x) \le \lim_{n \to \infty} F(\tfrac{1}{n})}$

Recall the characterization of the continuity of a function via sequences:

- f is continuous at $x_0 \iff \forall\, x_n \xrightarrow[n\to\infty]{} x_0$, then
$$\lim_{n\to\infty} f(x_n) = f(x_0).$$
- f is right continuous at $x_0 \iff \forall\, x_n \searrow x_0 \Rightarrow$
$$f(x_n) \to f(x_0).$$

<u>Thus</u>: Let $a_n \searrow 0$ $\;(a_1 \geq a_2 \geq \ldots \geq a_n \geq \ldots \geq 0)$.

We show that $\lim_{n\to\infty} F(a_n) = F(0) = \mu(\{0\})$

$(0, a_n]$

$$\lim_{n\to\infty} F(a_n) = \mu(\{0\}) + \underbrace{\lim_{n\to\infty} \mu((0, a_n])}$$

by propositio

$$\underset{\text{Prop. 1.1.2.}}{=} \mu(\{0\}) + \mu\Big(\bigcap_{n=1}^{\infty} (0, a_n]\Big) =$$

$$= \mu(\{0\}) + \mu(\emptyset) = \mu(\{0\}).$$

3) <u>Finite</u>: Note that $0 \leq F(a,b] < \infty$ for all finite a, b. Use again Prop. 1.1.2 and (4) to check that F is finite.

Since $F\_(0) = 0$, ~~and then~~ F is indeed the representative gdf. $\Big\}$ How?

<u>Step 2</u> Assume now that we have a representative gdf F satisfying (1) page 4.

- ~~Define~~ Let $\mathcal{J}$ be the collection of all finite intervals $(a, b]$.

- The idea is to show that ~~that~~ we can define, via (3), a measure μ on

$$\mathcal{C}_F \equiv \{\text{all finite disjoint unions of intervals in } \mathcal{C}_I\}, \text{ where}$$

$$\mathcal{C}_I \equiv \{\text{all intervals } (a, b], (-\infty, b], (a, \infty)\,;\, a, b \in \mathbb{R}\}.$$

Then, by Carathéodory Thm, once μ on $\mathcal{C}_F$ is defined we know that it has an unique extension to $\mathcal{B} \equiv \sigma[\mathcal{C}_F]$.

- Furthermore, we show that defining a c.a. measure on $\mathcal{C}_F$ reduces to defining a c.a. measure on $\mathcal{J}$. For this, notice that $(a, \infty) = \bigcup_{n=1}^{\infty} (a, n]$ and $(-\infty, b] = \bigcup_{n=1}^{\infty} (-n, b]$ and, moreover, $(a, \infty) = \sum_{n=1}^{\infty} I_n$, $(-\infty, b] = \sum_{n=1}^{\infty} I_n'$ using the construction of Prop 1.1.2, for $I_n, I_n' \in \mathcal{J}$.

Hence, let $\mu((a,b]) = F(b) - F(a)$.

We show that μ is c.a (countably additive) on $\mathcal{I}$.

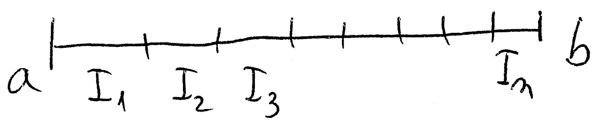
Assume $(a,b] = \sum_{n=1}^{\infty} (a_n, b_n] \equiv \sum_{n=1}^{\infty} I_n$.

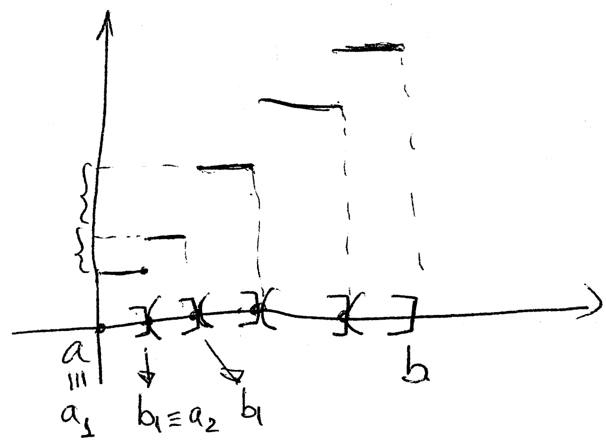
Show that $\mu\left(\sum_{n=1}^{\infty} I_n\right) = \sum_{n=1}^{\infty} \mu(I_n)$

---

a) $\sum_{n=1}^{\infty} \mu(I_n) \leq \mu\left(\sum_{n=1}^{\infty} I_n\right)$

Let n be fixed. Then, $\sum_{k=1}^{n} I_k \subset (a,b]$.





$$\sum_{k=1}^{n} \mu(I_k) = \sum_{k=1}^{n} F(b_k) - F(a_k) = F(b_1) - F(a_1) + F(b_2) - F(a_2) + \dots + F(b_n) - F(a_n) \leq$$

$$\leq F(b) - F(a) = \mu((a,b]).$$

Taking $n \to \infty$ gives a).

(b) $\mu\left(\sum_{n=1}^{\infty} I_n\right) \leq \sum_{n=1}^{\infty} \mu(I_n)$

Thus, we need to show that

$$\mu((a,b]) = F(a,b] \leq \sum_{n=1}^{\infty} \mu(I_n), \text{ for } b > a+\varepsilon, \ \varepsilon > 0.$$

We show that:

(c) $F(b) - F(a+\varepsilon) \leq \sum_{n=1}^{\infty} \mu(I_n) + \varepsilon$.

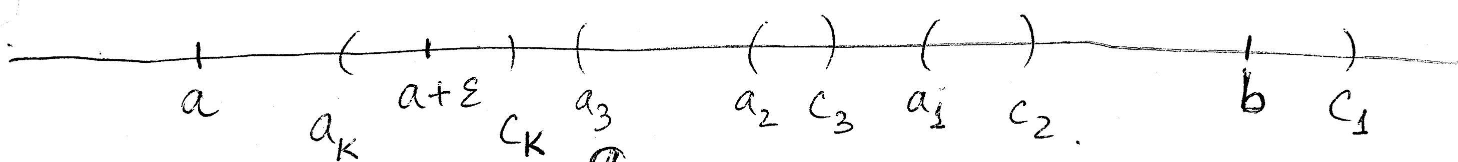
We then let $\varepsilon \downarrow 0$ and use the right continuity of F to obtain (b).

Prove now (c). Recall $F(b) - F(a+\varepsilon) = F(a+\varepsilon, b]$.

Also, note that $\underbrace{[a+\varepsilon, b]}_{\text{Compact set}} \subset (a,b] = \bigcup_{n=1}^{\infty} (a_n, b_n]$ disjoint.

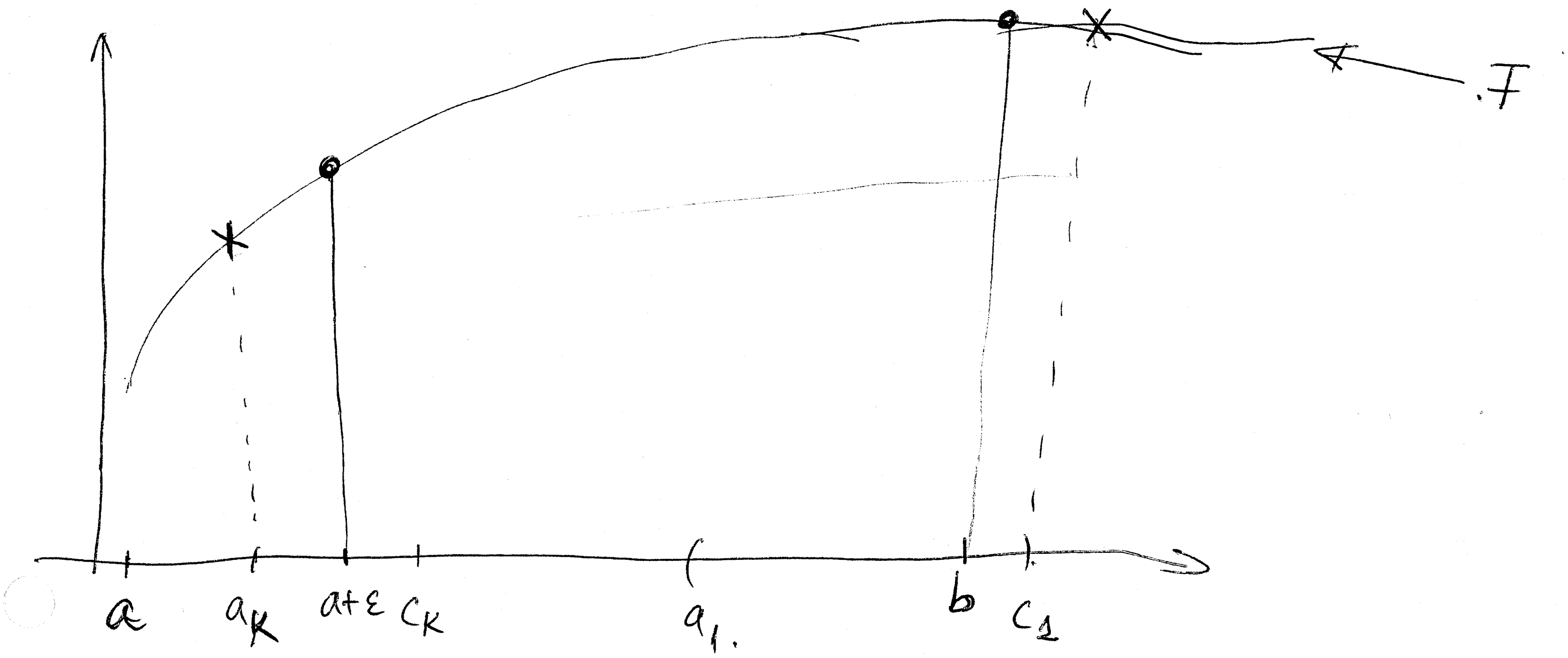
$$\subseteq \underbrace{\bigcup_{n=1}^{\infty} (a_n, \underbrace{b_n + \varepsilon_n}_{\shortparallel\ c_n})}; \text{ for any } \varepsilon_n > 0$$

Heine-Borel $\Longrightarrow$ $\exists$ a finite cover $\{(a_k, c_k)\}_{k=1}^{K}$ of open cover $[a+\varepsilon, b]$.



Recall that F is right continuous. Then for the construction above, choose ε_n small such that

$$F(b_n) - F(b_n+\varepsilon_n) \equiv F(b_n, b_n+\varepsilon_n] < \frac{\varepsilon}{2^n}, \text{ for each } n \geq 1.$$



Then:

$$F(a+\varepsilon, b] \leq F(a_K, c_1] \leq \sum_{k=1}^{K} F(a_k, c_k] =$$

$$= \sum_{k=1}^{K} F(c_k) - F(a_k) = \sum_{k=1}^{K} F(b_k+\varepsilon_k) - F(b_k) + F(b_k) - F(a_k)$$

$$= \sum_{k=1}^{K} F(b_k) - F(a_k) + \sum_{k=1}^{K} \frac{\varepsilon}{2^k} \leq$$

$$\leq \sum_{k=1}^{\infty} \mu(I_k) + \varepsilon.$$

Check that μ is non-negative and that $\mu(\emptyset) = 0$. ?

Thus μ is a c.a. measure on $\mathcal{C}_F$.

We need to show that it is well defined.

Let $\underset{\underset{\mathcal{C}_F}{\cap}}{A} = \sum_n I_n = \sum_m I'_m$ be two representations of A.

We need to show that $\mu\left(\sum_n I_n\right) = \mu\left(\sum_m I'_m\right)$.

Define $\mu(A) \equiv \sum_n \mu(I_n)$.

Thus, we show $\sum_n \mu(I_n) = \sum_m \mu(I'_m)$.

Note that
$$I'_m = A \cap I'_m = \sum_n I_n I'_m \Rightarrow$$
$$I_n = A \cap I_n = \sum_m I'_m I_n$$

$$\sum_m \mu(I'_m) = \sum_m \sum_n \mu(I_n I'_m) = \sum_n \sum_m \mu(I_n I'_m) =$$
$$= \sum_n \mu(I_n)$$

q.e.d.

Lecture 6 Measurable functions

Everything

<u>Definition 1</u> Let $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ two measure spaces. Let $X : \Omega \longrightarrow \Omega'$.

We call X <u>measurable</u> if

$$X^{-1}(\mathcal{A}') \subset \mathcal{A}.$$

<u>Notation</u> 1) For $A \in \mathcal{A}'$, the <u>inverse image of A</u> is

$$X^{-1}(A) \equiv \{\omega \in \Omega \mid X(\omega) \in A)$$
$$\equiv [X \in A]$$

2) $X^{-1}(\mathcal{A}') \equiv \{X^{-1}(A) \mid A \in \mathcal{A}'\}$.

<u>Comments</u> 1) In definition 1, we in fact call X $\mathcal{A}'$-$\mathcal{A}$-measurable.

2) Let $\bar{\mathcal{B}} = \sigma[\mathcal{B}, \{+\infty\}, \{-\infty\}]$.

In the special case

$$X : (\Omega, \mathcal{A}) \longrightarrow (\bar{\mathbb{R}}, \bar{\mathcal{B}}), \text{ if}$$

X is $\bar{\mathcal{B}}$-Ω measurable, we call it simply <u>measurable</u>.

<u>Proposition 1</u> (Measurability criteria).

Let $X : (\Omega, \mathcal{A}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$.

Suppose $\sigma[\mathcal{C}] = \bar{\mathcal{B}}$. <u>Then</u>:

1) X is measurable if and only if $X^{-1}(\mathcal{C}) \subset \mathcal{A}$. does it hold for any σ-field?

2) X is measurable if and only if $X^{-1}([-\infty, x]) \in \mathcal{A}$, for all $x \in \mathbb{R}$.

<u>Proof</u> 1) „$\Leftarrow$" Assume $X^{-1}(\mathcal{C}) \subset \mathcal{A}$ and show that X is measurable, that is, show that $X^{-1}(\bar{\mathcal{B}}) \subset \mathcal{A}$.

We use Proposition 2.1.2 page 22:

$X^{-1}(\sigma[\mathcal{C}]) = \sigma[X^{-1}(\mathcal{C})]$, for any collection $\mathcal{C}$ of subsets of $\overline{\mathcal{B}}$.

Thus $X^{-1}(\overline{\mathcal{B}}) = X^{-1}(\sigma[\mathcal{C}]) =$
$= \sigma[X^{-1}(\mathcal{C})]$

We have $\left.\begin{array}{l} X^{-1}(\mathcal{C}) \subset \mathcal{A} \\ \mathcal{A}\ \sigma\text{-field} \end{array}\right\} \Rightarrow \sigma[X^{-1}(\mathcal{C})] \subset \underset{\mathcal{A}}{\underset{\parallel}{\sigma[\mathcal{A}]}}$

So: $X^{-1}(\overline{\mathcal{B}}) \subset \mathcal{A}$.

$\Rightarrow$" Check!

2) Based on Prop 2.1.2 page 22 and on 1), we only need to check that
$$\overline{\mathcal{B}} = \sigma[\{[-\infty, x] : x \in \mathbb{R}\}].$$

Since $\mathcal{B} = \sigma[\mathcal{C}_I]$, $\mathcal{C}_I = \{(a,b], (-\infty, b], (a, \infty), -\infty < a < b < \infty\}$

it enough to check that each interval in $\mathcal{C}_I$ can be obtained through various operations from intervals of the type $[-\infty, x]$.

Since, indeed,

$$(a,b] = [-\infty, b] \cap [-\infty, a]^c,$$

$$[-\infty, b) = \bigcup_1^{\infty} [-\infty, b - \tfrac{1}{n}]$$

$$\{-\infty\} = \bigcap_n [-\infty, -n], \quad \{\infty\} = \bigcap_n [-\infty, n]^c,$$

the proof is complete.

Proposition 2.2 page 25

"Common" functions are measurable. See proof page.

Example 1: If X is measurable, cX, $c>0$, is measurable.

Cf. Proposition 2.1, it is enough to show that for $Z \equiv cX$, we have

$$Z^{-1}([-\infty, x]) \in \mathcal{A}, \quad \text{for all } x \in \mathbb{R}$$

$$Z^{-1}([-\infty, x]) = \{\omega \in \Omega \mid Z(\omega) \leq x\} =$$

$$= \{\omega \in \Omega \mid X(\omega) \leq \tfrac{x}{c}\} \in \mathcal{A},$$

for all $x \in \mathbb{R}$, since is measurable.

$$\begin{tikzcd} (\Omega, \mathcal{A}) \arrow[r, "X"] \arrow[dr, "g\circ X"'] & (\mathbb{R}, \mathcal{B}) \arrow[d, "g"] \\ & (\mathbb{R}, \mathcal{B}) \end{tikzcd}$$

__Example 2__ If X is measurable and g is continuous, then the composite function $g(X)$ is measurable.

__Proof.__ If g is measurable then, noting that $g(X) = g \circ X$, we have

$$(g\circ X)^{-1}(\bar{\mathcal{B}}) = X^{-1}(g^{-1}(\bar{\mathcal{B}})) \subset X^{-1}(\bar{\mathcal{B}}) \subset \mathcal{A}$$

See page 22, Prop. 1.1.

$g : (\bar{\mathbb{R}}, \bar{\mathcal{B}}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$

Since X is measurable

- To prove the result for continuous functions, notice that <u>any continuous function</u> is measurable ;

from $(\bar{\mathbb{R}}, \bar{\mathcal{B}}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$

Since:

$$g^{-1}(\mathcal{B}) = g^{-1}(\sigma[\text{open sets}]) = \sigma[g^{-1}(\text{open sets})] \subset$$

Prop 1.2 page 22

$$\subset \sigma[\text{open sets}] \subset \bar{\mathcal{B}}$$

g continuous $\Leftrightarrow$ $g^{-1}(\text{open set}) \subset \text{open set}$.

Also $g^{-1}(\{+\infty\}) = \emptyset \in \bar{\mathcal{B}}$ and $g^{-1}(\{-\infty\}) = \emptyset \in \bar{\mathcal{B}}$.

Definition 2 Let $(\Omega, \mathcal{A})$ be a measure space.

a) For $A \in \Omega$, $1_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise.} \end{cases}$ $\rightarrow$

the indicator function of A.

b) A <u>simple</u> function is of the form

$$X(\omega) \equiv \sum_{i=1}^{n} x_i 1_{A_i}(\omega), \text{ for } \sum_{i=1}^{n} A_i = \Omega, \text{ with}$$

all $A_i \in \mathcal{A}$ and $x_i \in \mathbb{R}$.

*** Proof Req. also

Proposition 2.3 (Measurability via simple functions).

**

a) Simple functions are measurable.

** Proof required

b) $X : \Omega \rightarrow \bar{\mathbb{R}}$ is measurable $\Leftrightarrow$ X is the limit of a sequence of simple functions.

($\mathbb{R}$ line including $-\infty$ & $+\infty$)

No proof

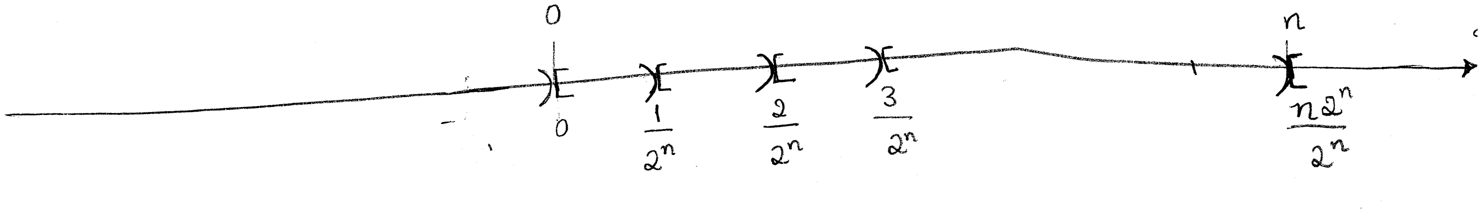
c) If $X \geq 0$ is measurable, then X is the limit of simple functions that are ≥ 0 and $\uparrow$.

Proof: a) See Hwk 3

b).
- The idea is to construct a sequence X_n such that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, for all $\omega \in \Omega$.
- If $\{X_n\}_n$ are measurable, then by Prop. 2.2, page 25, X is also measurable.
- Define

$$X_n \equiv \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \left\{ 1_{\left[\frac{k-1}{2^n} \le X < \frac{k}{2^n}\right]} - 1_{\left[\frac{k-1}{2^n} \le -X < \frac{k}{2^n}\right]} \right\}$$

$$+ \, n \cdot \left\{ 1_{[X \ge n]} - 1_{[-X \ge n]} \right\}.$$



Notice now that

$$\left\{ \left\{ \left[\frac{k-1}{2^n} \le X < \frac{k}{2^n} \right], \left[\frac{k-1}{2^n} \le -X < \frac{k}{2^n} \right] \right\}_{k=1}^{n \cdot 2^n}, \; [X \ge n], \; [-X \ge n] \right\}$$

form a partition of Ω and all the elements of this partition are in $\mathcal{A}$, since X is measurable (and we use Prop. 2.1 page 25).

Hence, by a), $\{X_n\}_n$ is measurable.

It remains to show that its limit is X.

- Let ω be arbitrary fixed.
- Notice that if $|X(\omega)| \geq n \iff X(\omega) \geq n$ or $X(\omega) \leq -n$

then $X_n(\omega) = \begin{cases} n, & \text{if } X(\omega) \geq n \\ -n, & \text{if } X(\omega) \leq -n \end{cases}$

so $\lim\limits_{n \to \infty} X_n(\omega) = \infty = X(\omega)$

• Thus, it is enough to show that

$$\lim_{n\to\infty} X_n(\omega) = X(\omega), \text{ if } |X(\omega)| < n.$$

$$\Updownarrow$$

$$-n < X(\omega) < n$$

Enough to consider the case $X(\omega) \in [0, n)$, since $X(\omega) \in (-n, 0]$ is identical.

Now, for $X(\omega) \in [0, n)$ write

$$X(\omega) = X(\omega) \cdot 1 = X(\omega) \sum_{k=1}^{n2^n} \underbrace{1_{\left[\frac{k-1}{2^n} \le X < \frac{k}{2^n}\right]}(\omega)},$$

since ~~these sets form a partition of~~ $X(\omega)$ belongs to one of these intervals.

Then:

$$|X_n(\omega) - X(\omega)| = \left| \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n} - X(\omega)\right) 1_{\left[\frac{k-1}{2^n} \le X(\omega) < \frac{k}{2^n}\right]}(\omega) \right|$$

$$\le \sum_{k=1}^{n2^n} \left| \frac{k-1}{2^n} - X(\omega) \right| 1_{\left\{\omega \,\middle|\, \frac{k-1}{2^n} \le X(\omega) < \frac{k}{2^n}\right\}}(\omega) \le$$

$\le \frac{1}{2^n}$, since only term in the sum is nonzero, namely the one corresponding to the indicator of the set of ω's ~~to which~~ for which $X(\omega)$ is in an interval of length $\frac{1}{2^n}$.

Since $|X_n(\omega) - X(\omega)| \le \frac{1}{2^n} \Rightarrow \lim_{n\to\infty} X_n(\omega) = X(\omega)$, for any $\omega \in \Omega$.

c) Proof at page 26. (Not required)

<u>Proposition 2.4</u> The discontinuity set is measurable.
Page 26.

APPENDIX C

The Problem of Measure

Assuming the axiom of choice, it is known, as shown in §3.4, that there exist non-measurable sets for Lebesgue measure. There is still the question whether Lebesgue measure could be extended to all subsets of $\mathbb{R}$ as a countably additive measure. More generally, the "problem of measure" asks whether there is a measure μ on all subsets of an uncountable set X, with $\mu(X) = 1$ and $\mu\{x\} = 0$ for each point $x \in X$. This appendix will prove a partial answer, which will be used in Appendix E.

***C.1 Theorem* (Banach and Kuratowski).** Assuming the continuum hypothesis, there is no measure μ defined on all subsets of $I := [0, 1]$ with $\mu(I) = 1$ and $\mu\{x\} = 0$ for each $x \in I$.

PROOF. The proof will be based on the following:

C.2 Lemma. Assuming the continuum hypothesis, there exist subsets A_{ij} of I for i and $j = 1, 2, \ldots$, with the following properties:

(a) For each i, the sets A_{ij} for different j are disjoint and their union is I.
(b) For any sequence $k(i)$ of positive integers,

$$\bigcap_i \bigcup_{1 \le j \le k(i)} A_{ij} \text{ is at most countable.}$$

PROOF. For any two sequences $\{n_i\}$ and $\{k_i\}$ of positive integers, $\{n_i\} \le \{k_i\}$ will mean $n_i \le k_i$ for all i. The following will be proved first:

C.3 Lemma. Assuming the continuum hypothesis, there is a set $\mathscr{F}$ of sequences of integers, where $\mathscr{F}$ has cardinality c (the cardinality of I), such that for every sequence $\{m_j\}$ (in $\mathscr{F}$ or not), the set of all sequences $\{n_j\}$ in $\mathscr{F}$ with $\{n_j\} \le \{m_j\}$ is at most countable.

PROOF. Let $\mathscr{S}$ be the set of all sequences of positive integers. Then $\mathscr{S}$ has cardinality c (§13.1). By the continuum hypothesis, there is a well-ordering $\leqq$ on $\mathscr{S}$ such that for each $y \in \mathscr{S}$, $\{x: x \leqq y\}$ is countable. For each $\alpha \in \mathscr{S}$, let f_α be a function from $\mathbb{N}$ onto $\{x: x \leq \alpha\}$. Define a sequence $g_\alpha := \{g_\alpha(n)\}_{n\geq 0}$ of positive integers by $g_\alpha(n) := f_\alpha(n)(n) + 1$ for $n = 0, 1, \ldots$. (Then g_α is called a "diagonal" sequence.) It is not true that $g_\alpha \leq x$ for any of the countably many $x \leqq \alpha$. Let $\mathscr{F}$ be the set of all sequences g_α, $\alpha \in \mathscr{S}$. Then if $g_\alpha \leq x$, it must not be the case that $x \leqq \alpha$, so $\alpha \leqq x$, and the set of such α is countable. Now $\mathscr{F}$ is uncountable, since each g_α is the sequence y for some $y \in \mathscr{S}$, and if the set of such y were countable, they would have a supremum $\beta \in \mathscr{S}$ for $\leqq$, but the sequence g_β is different from all the sequences $y \leqq \beta$, a contradiction. So by the continuum hypothesis, $\mathscr{F}$ has cardinality c, proving Lemma C.3. □

Now to prove Lemma C.2, let h be a 1–1 function from $[0, 1]$ onto $\mathscr{F}$. Thus each $h(x)$ is a sequence $\{h(x)_n\}_{n\geq 0}$. Define sets A_{ij} by $x \in A_{ij}$ iff $j = h(x)_i$. Then for a fixed value of i, the sets A_{ij} for different j are clearly disjoint. Their union over all j gives the whole interval $[0, 1]$.

Let $\{k_i\} := \{k(i)\}$ be any sequence of positive integers. Let $x \in \bigcap_i \bigcup_{j\leq k(i)} A_{ij}$. Then by definition of A_{ij}, we have $h(x)_i \leq k(i)$ for all i, so $h(x) \leq \{k_i\}$. By Lemma C.2, there are only countably many sequences in $\mathscr{F}$ which are $\leq \{k_i\}$, and since h is 1–1, there are only countably many such $x \in [0, 1]$, finishing the proof of Lemma C.2. □

Now to prove Theorem C.1, choose $k(i)$ for each $i \geq 1$ such that $\mu(B_i) < 1/2^{i+1}$ where $B_i := \bigcup_{j>k(i)} A_{ij}$. By Lemma C.2, the intersection of the complements of all the B_i is countable, so it has μ measure 0. Thus $1 = \mu(\bigcup_i B_i) < \sum_{i\geq 1} 1/2^{i+1} = 1/2$, a contradiction. □

Notes

This appendix is based on the paper of Banach and Kuratowski (1929). For more information on the problem of measure ("measurable cardinals," "real-valued measurable cardinals," etc.), see, for example, Jech (1978).

References

Banach, Stefan, and Casimir [Kazimierz] Kuratowski (1929). Sur une généralisation du problème de la mesure. *Fund. Math. 14*: 127–131.

Jech, Thomas (1978). *Set Theory*. Academic Press, New York.

Definitions only

Lecture 7 Convergence

I. Convergence almost everywhere (a.e.)

Definition Let $X_1, \dots, X_n, \dots$ be a sequence of measurable functions defined on $(\mathcal{A}, \Omega, \mu)$ with values in $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$.

$$X_n \xrightarrow[n \to \infty]{a.e.} X \quad \text{if}$$

$X_n(\omega) \xrightarrow[n \to \infty]{} X(\omega)$ for all $\omega \notin N$, for some $N \in \mathcal{A}$ with $\mu(N) = 0$.

So: $X_n \xrightarrow[n \to \infty]{a.e} X$ if we have pointwise convergence except for a set of measure 0.

Note If $X_n \xrightarrow{a.e} X$, then $\exists\ \tilde{X}$ measurable such that $X = \tilde{X}$ a.e.

(See Hwk 3).

Thus, we can always redefine the limit X such that, except for a set of measure 0, it is measurable.

Proposition 1 (mimics the result for real numbers) (Page 29)

Let $(X_n)_n$ measurable a.e. finite.

Let X measurable a.e. finite.

Then:

$$X_n \xrightarrow[n \to \infty]{a.e} X \iff X_n - X_m \xrightarrow{a.e.} 0 \text{ as } m \wedge n \to \infty$$

(Contrast with: Any sequence of real numbers is convergent $\iff$ it is a Cauchy sequence).

Terminology used in the book:

If $X_n(\omega)$ is a Cauchy sequence a.e, it will be called mutually convergent a.e. ☆☆

Proof of Proposition 1 — Exercise.

Proposition 2 (Page 29)

Let $X, X_1, X_2, \ldots$ be finite measurable functions. We denote the:

a) convergence set by $[X_n \to X]$

b) mutual convergence set by $[X_n - X_m \to 0]$.

Then:

(1) $[X_n \to X] = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} [|X_m - X| < \frac{1}{k}]$

(2) $[X_m - X_n \to 0] = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} [|X_m - X_n| < \frac{1}{k}]$

Hence, the mutual convergence and the convergence sets are measurable. (Why?).

Proof We show that (1) holds.
(2) is identical.

Recall that $[X_n \to X] = \{\omega \in \Omega \mid X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)\}$

- For each $\omega \in \Omega$, $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega) \iff$

$\forall \varepsilon > 0 \quad \exists n_\varepsilon$ s.t. $\forall n \geq n_\varepsilon \quad |X_n(\omega) - X(\omega)| < \varepsilon$

$\iff \boxed{\forall \varepsilon = \frac{1}{K} > 0 \quad \exists n_K \geq 1 \text{ s.t. } \forall n \geq n_K \ |X_n(\omega) - X(\omega)| < \frac{1}{K}}$

Note now that $\forall \longleftrightarrow \bigcap$

$\exists \longleftrightarrow \bigcup$

Thus, reading this line from right to left, we have: $\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} [|X_m - X| < \frac{1}{k}]$

Notation:

$\forall \frac{1}{K} > 0$ (that is, $\forall k$), $\exists n \geq 1$ s.t. $\forall m \geq n$

$|X_m(\omega) - X(\omega)| < \frac{1}{K}$.

Important set : The divergence set :

$$[\omega \mid X_n(\omega) \not\to X(\omega)] = [X_n \to X]^c =$$

$$= \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left[|X_m - X| \geq \tfrac{1}{k} \right]$$

Notice that :

a) $D_{kn} \equiv \bigcup_{m=n}^{\infty} \left[|X_m - X| \geq \tfrac{1}{k} \right] \downarrow$ in n

b) $A_k \equiv \bigcap_{n=1}^{\infty} D_{kn} \uparrow$ in k

Hence $[X_n \to X]^c = \bigcup_{k=1}^{\infty} A_k, \quad A_k \uparrow .$

Proposition 3 (Page 30).

Consider finite measurable X_n's and X.

1) $$X_n \xrightarrow[n\to\infty]{a.e.} X \iff \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}[|X_m - X_n| > \varepsilon]\right) = 0,$$

for all $\varepsilon > 0$.

(Convergence to a finite limit $\iff$ Cauchy).

2) If $\mu(\Omega) < \infty$ then:

$$X_n \xrightarrow[n\to\infty]{a.e.} X \iff \mu\left(\bigcup_{m=n}^{\infty}[|X_m - X_n| \geq \varepsilon]\right) \xrightarrow[n\to\infty]{} 0,$$

for all $\varepsilon > 0$ Also holds if we have X

$$\iff \mu\left(\left[\max_{n \leq m \leq N} |X_m - X_n| \geq \varepsilon\right]\right) \leq \varepsilon,$$

for all $N \geq n \geq n_\varepsilon$, for all $\varepsilon > 0$.

Remarks

1. Showing that $(X_n)_n$ is convergent ~~us~~ using the Cauchy criterion is very useful, as it does **NOT** require any knowledge of X.

2. Note that in 1) the measure is required to be exactly zero, and the event is "complicated".

3. We can use Proposition 1.1.2 to say that $\mu(\bigcap_n F_n) = \lim_n \mu(F_n)$ only if we impose $\mu(\Omega) < \infty$. This will be very useful for us, as we will only work (in the near future!) with $\mu = P$, hence finite.

Proof Exercise! Use Prop 1.1.2. and a reasoning as at Page 4.

$$X_n \xrightarrow[n\to\infty]{a.e.} X \iff X_m - X_n \xrightarrow[m\wedge n\to\infty]{a.e} 0.$$

- $X_m - X_n \xrightarrow{a.e} 0 \underset{\text{Definition}}{\overset{}{\iff}} X_m(\omega) - X_n(\omega) \xrightarrow[m\wedge n\,\infty]{} 0$ for all $\omega \in N^c$, for some N s.t. $\mu(N) = 0$. (That is $\mu(N^c) = 1$)

- Write out $X_m(\omega) - X_n(\omega) \xrightarrow[n\wedge m'\to\infty]{} 0$,

$$\forall \varepsilon > 0 \quad \underbrace{\exists\ n_\varepsilon \geq 1 \quad \text{s.t.} \quad \forall\, m, n \geq n_\varepsilon \geq 1 : |X_m(\omega) - X_n(\omega)| \leq \varepsilon}_{\circledast}$$

Let $\varepsilon > 0$ be arbitrary, fixed.

Then the set $\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega \in \Omega \,|\, |X_n(\omega) - X_m(\omega)| \leq \varepsilon\}$ $\Rightarrow$ contains all the ω's s.t. $\circledast$ holds, for that ε.

$$\mu\left(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} [|X_n - X_m| \leq \varepsilon]\right) = 1 \Longrightarrow$$

$$X_m - X_n \xrightarrow[m\wedge n\to\infty]{} 0 \iff \forall \varepsilon > 0\ \mu\Big(\bigcap_{n=1}^{\infty} \underbrace{\bigcup_{m=n}^{\infty} [|X_m - X_n| > \varepsilon]}\Big) = 0$$

- ~~Now~~ $\iff \forall \varepsilon > 0 \ \lim_{n\to\infty} \mu\left(\bigcup_{m=n}^{\infty} [|X_m - X_n| > \varepsilon]\right) = 0$

$\Longleftrightarrow \forall \varepsilon > 0 \quad \exists\, \underline{\underline{n_\varepsilon}} \quad s.t. \quad \forall\, n \geq n_\varepsilon$

$$\mu\Big(\bigcup_{m=n_\varepsilon}^{\infty} [\,|X_m - X_n| > \varepsilon\,] \Big) \leq \varepsilon \qquad (1)$$

Notice now that (since n_ε is $\underline{\underline{\text{fixed}}}$)

$$\bigcup_{m=n_\varepsilon}^{\infty} [\,|X_m - X_n| > \varepsilon\,] = \bigcup_{N=n_\varepsilon}^{\infty} \underbrace{\bigcup_{m=n_\varepsilon}^{N} [\,|X_m - X_n| > \varepsilon\,]}_{A_N}$$

$A_{n_\varepsilon} = [\,|X_{n_\varepsilon} - X_n| > \varepsilon\,]$: For $N = n_\varepsilon$

: For $N = n_\varepsilon + 1$

$A_{n_\varepsilon + 1} = [\,|X_{n_\varepsilon} - X_n| > \varepsilon\,] \cup [\,|X_{n_\varepsilon+1} - X_n| > \varepsilon\,]$.

- - - - - - - - .

So $\{A_N\} \nearrow$

$\underline{\underline{\text{Thus}}}$ (1) $\Longleftrightarrow \forall \varepsilon > 0 \quad \exists\, n_\varepsilon \quad s.t. \quad \forall\, n \geq n_\varepsilon$

$$\mu\Big(\bigcup_{N=n_\varepsilon}^{\infty} \bigcup_{m=n_\varepsilon}^{N} [\,|X_m - X_n| > \varepsilon \Big) \leq \varepsilon$$

$\Longleftrightarrow \forall \varepsilon > 0 \quad \exists\, n_\varepsilon \quad s.t. \quad \forall\, n \geq n_\varepsilon$

$$\lim_{\substack{N \to \infty \\ (N \geq n_\varepsilon)}} \mu\Big(\bigcup_{m=n_\varepsilon}^{N} |X_m - X_n| > \varepsilon \Big) \leq \varepsilon.$$

$\Longleftrightarrow \forall\, \varepsilon > 0 \quad \exists\, n_\varepsilon \quad s.t. \quad \forall\, n, N \geq n_\varepsilon$

$$\mu\Big(\bigcup_{m=n_\varepsilon}^{N} |X_m - X_n| > \varepsilon \Big) \leq \varepsilon.$$

Notice further that

$$\bigcup_{j=1}^{J} [|Y_j| > a] = [\max_{1\le j\le J} |Y_j| > a].$$

Equivalently $A \equiv \bigcap_{j=1}^{J} [|Y_j| \le a] = [\max_{1\le j\le J} |Y_j| \le a] \equiv B$

"$\subseteq$" $\omega \in \bigcap_{j=1}^{J} [|Y_j| \le a] \Rightarrow |Y_j(\omega)| \le a \quad \forall j = 1, .., J \Rightarrow$

$\Rightarrow \max |Y_j(\omega)| \le a \Rightarrow \omega \in B.$

"$\supseteq$" Let $\omega \in B$.

We always have: $|Y_j(\omega)| \le \max |Y_j(\omega)|$, for all j

Since $\omega \in B \Rightarrow |Y_j(\omega)| \le \max |Y_j(\omega)| \le a. \Rightarrow \omega \in A$]

Thus

$X_m - X_n \xrightarrow{a.e} 0 \iff \forall \varepsilon > 0, \exists n_\varepsilon \ge 1$ s.t. $\forall N, n \ge n_\varepsilon$

we have $\mu(\max_{n\le m\le N} |X_m - X_n| \ge \varepsilon) \le \varepsilon.$

q.e.d.

Remark By the def of $\overline{\lim}$ notice that

$$\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} [|X_m - X| > \varepsilon] = \overline{\lim} A_n, \quad \text{where}$$

$$A_n \equiv [|X_n - X| > \varepsilon].$$

Thus $X_n \xrightarrow{a.e} X$ iff $\mu(\overline{\lim} A_n) = 0, \forall \varepsilon > 0.$