UPPSALA UNIVERSITET

Föreläsningsanteckningar

Sannolikhet & Martingaler

Rami Abou Zahra

Inlämningsdatum
September 16, 2024

Contents

1. σ-algebras & Measure spaces

1.1. σ-algebras.

Definition 1.1 σ-algebra

A collection of subsets Σ of a set S is called a *σ-algebra* if:

- $\varnothing \in \Sigma$
- Is an algebra:
  - Closed under complements such that for $A \in \Sigma \Rightarrow A^c = S \backslash A \in \Sigma$
  - Closed under unions such that $A, B \in \Sigma \Rightarrow A \cup B \in \Sigma$
- Closed under countably infinite unions $A_i \in \Sigma$ for $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} A_i \in \Sigma$

Example:
$\Sigma = \{\varnothing, S\}$ is a σ-algebra on any set S.
Another example is $\mathcal{P}(S)$, which denotes the powerset.
Another example is $S = \mathbb{N}$, then $\Sigma = \{\varnothing, \mathbb{N}, \{2k : k \in \mathbb{N}\}, \{2k+1 : k \in \mathbb{N}\}\}$

Remark:
There exists many equivalent definitions of a σ-algebra. For example, instead of the first axiom of $\varnothing \in \Sigma$, an equivalent definition could be "Σ is non-empty", since then $\exists A \in \Sigma \Rightarrow A^c \in \Sigma \Rightarrow A \cup A^c = S \in \Sigma \Rightarrow (A \cup A^c)^c = \varnothing \in \Sigma$

Remark:
Closed under unions $\Rightarrow$ closed under finite unions since $A_1, \cdots, A_n \in \Sigma \Rightarrow A_1 \cup A_2 \in \Sigma, A_1 \cup A_2 \cup A_3 = \underbrace{(A_1 \cup A_2)}_{\in \Sigma} \cup A_3$, thus by induction $A_1 \cup \cdots \cup A_n \in \Sigma$

This does *not* imply Σ is closed under countable unions.

Counter-example:
Consider $S = [0, 1) \subseteq \mathbb{R}$. Let Σ be all finite unions of disjoint sets on the form $[a, b)$ such that $0 \leq a \leq b < 1$ (if $a = b \Rightarrow \varnothing$).
First and all algebra axioms are fulfilled, but the last one is not since we cvan consider $A_n = \left[\frac{1}{n}, 1\right)$.
Then $\bigcup_{i=2}^{\infty} = (0, 1) \notin \Sigma$

An algebra Σ is an algebra in an algebraic sense.
The symmetric difference $A \triangle B = (A \backslash B) \cup (B \backslash A)$. This behaves like "+" on Σ and intersections behave like multiplication.
Just like one would expect from an algebra, the multiplication is distributive over addition, eg. $C \cap (A \triangle B) = (C \cap A) \triangle (C \cap B)$

1.2. Measures.

Let Σ be a σ-algebra on S, and let μ_0 be a function from Σ_0 to $[0, \infty] = [0, \infty) \cup \{\infty\}$, essentially a function that assigns some value to subsets of Σ.
Intuitively, a measure should increase if we measure something bigger.

Definition 1.2 Additive and σ-additive measures

A measure μ_0 is called *additive* if $\mu_0(A \cup B) = \mu_0(A) + \mu_0(B)$ where A, B are disjoint sets.

A measure μ_0 is called *σ-additive* if this holds for ocuntable unions, i.e if A_n are pairwise disjoint, then $\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_0(A_n)$

Remark:
We say that μ_0 is a measure if μ_0 is σ-additive and $\mu_0(\varnothing) = 0$

Example:
$S = \{1, 2, \cdots, 6\}$, $\Sigma = \mathcal{P}(S)$ and set $\mu_0(A) = \frac{1}{6}|A|$. Note here that $\mu_0(S) = 1$

Definition 1.3 Probability measures

All measures that sum up to 1 are called *probability measures*

Example:
$S = \mathbb{N}$, $\Sigma = \mathcal{P}(S)$ and set $\mu_0(A \in \Sigma) = |A|$. Here $\mu_0(S) = \infty$

Example:
$S = \mathbb{N}$, $\Sigma = \mathcal{P}(S)$ and set $\mu_0(A \in \Sigma) = \begin{cases} 0 & \text{if} \quad |A| < \infty \\ \infty & \text{if} \quad |A| = \infty \end{cases}$
This is an example of an additive but not σ-additive measure, since if $A_n = \{n\}$, then $\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) = \infty$, but $\sum_{n=1}^{\infty} \mu_0(A_n) = -1$

1.3. **Measure spaces.**

Definition 1.4 Measure space triplet

A *measure space* is a triplet (S, Σ, μ) where S is some set, Σ is a σ-algebra over S, and μ is a σ-additive function $\mu : \Sigma \to [0, \infty]$ such that $\mu(\varnothing) = 0$

Definition 1.5 Probability space

If $\mu(S) = 1$, then the triplet is called a *probability space.*

Example: (finite measure space)
Let $S = \{s_1, \cdots, s_k\}$ where $k \in \mathbb{N}$ be a set of outcomes. We also associate probabilities $p_1, \cdots, p_k$ to each $s_1, \cdots, s_k$ such that $\sum_i p_i = 1$. Let $\mu(A) = \sum_{s_i \in A} p_i \ \forall A \subseteq S$. If we let $\Sigma = \mathcal{P}(S)$, then (S, Σ, μ) is a measure and a probability space.

Example: (Lebesgue measure)
Let $S = \mathbb{R}$, $\Sigma = \mathcal{B}(\mathbb{R})$ be the Borel σ-algebra (smallest σ-algebra that makes open sets measureable, note that $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$) and let μ be something measuring length on finite unions of disjoint open intervals $A = (a_1, b_1) \cup \cdots \cup (a_n, b_n)$ such that $\mu(A) = |b_1 - a_1| + \cdots + |b_n - a_n|$

This μ is called the Lebesgue measure ($\mathcal{L}$)
Restricting S to $[0, 1]$, then we have a probability measure

$$\mu = \mathcal{L}\,|_{[0,1]}\,(A) = \mathcal{L}(A \cap [0, 1]) \Rightarrow ([0, 1], \mathcal{B}([0, 1], \mathcal{L}\,|_{[0,1]})) \quad \text{is a probability measure}$$

This is a formulation of uniform random numbers in $[0, 1]$

1.4. **Properties of measures.**
For a measure space, we have the following properties:

(1) $\mu(A \cup B) \leq \mu(A) + \mu(B)$
(2) $\mu\left(\bigcup A_i\right) \leq \sum \mu(A_i)$
(3) $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) - \mu(A_1 \cap A_2) - \cdots - \mu(A_{n-1} \cap A_n) + \mu(A_1 \cap A_2 \cap A_3) \cdots + (-1)^{n+1}\mu(A_1 \cap A_2 \cdots \cap A_n)$

Note that for the first two points, we have previously assumed that A, B were disjoint. This would be the case for "joint" sets.

Bevis 1.1

Consider $\mu(A) = \mu(A \backslash B \cup (A \cap B)) = \mu(A \backslash B) + \mu(A \cap B)$ and proceed. □

Remark:
For point 4, check Math Stackexchange

The idea is if we can consider some set that is measurable, we want to be able to say something about the compositions of those measurable sets so the idea is we include their subsets in the σ-algebra (in the space we set up) as well as keeping it closed in an algebraic sense.

1.5. **Monoticity of measure.**
Let (A_i) be a sequence of increasing sets in Σ such that $\varnothing \subseteq A_1 \subseteq \cdots \subseteq S$. Then:

$$\mu(A_i) = \mu(A_i \backslash A_{i-1} \cup (A_i \cap A_{i-1})) = \mu(A_i \backslash A_{i-1} \cup A_{i-1}) = \mu(A_i \backslash A_{i-1}) + \mu(A_{i-1}) \geq \mu(A_{i-1})$$

Thus, by induction, $\mu(A_1) \leq \mu(A_2) \leq \cdots$ and by monotone convergence the limit $\lim_{i\to\infty} \mu(A_i)$ exists in the extended positive real line.

Writing $A = \bigcup_{i=1}^{\infty} A_i$, we have $\mu(A) = \lim_{i\to\infty} \mu(A_i)$, this because:

$$A = A_1 \cup (A_2 \backslash A_1) \cup (A_3 \backslash A_2) \cup \cdots$$

$$\mu(A) = \mu(A_1) + \mu(A_2 \backslash A_1) + \mu(A_3 \backslash A_2) + \cdots = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_i \backslash A_{i-1})$$

where $A_0 = \varnothing = \lim_{n\to\infty} \mu(A_n)$

A similar result holds for decreasing sets, i.e $S \supseteq A_1 \supseteq A_2 \cdots \supseteq \varnothing$
We do the limit as $A = \bigcap_{i=1}^{\infty} A_i$ and by monotone convergence $\mu(A) = \lim_{i\to\infty} \mu(A_i)$ with similar proof.

Remark:
The last set in the decreasing sets does not necessarily have to be the empty set, recall that we are dealing with intersections instead of unions.

1.6. **Generated σ-algebras.**

Given any collection of subsets of $\mathfrak{A} \subseteq \mathcal{P}(S)$, the σ-algebra *generated* by $\mathfrak{A}$ is the smallest σ-algebra that cointains $\mathfrak{A}$ is denoted by $\sigma(\mathfrak{A}) = \bigcap_{\Sigma:\sigma\text{-alg}\&\mathfrak{A}\subseteq\Sigma}$
This is sometimes denoted by $< \mathfrak{A} >$

One can verify that this is indeed a σ-algbera:

(1) $\varnothing$ is contained in all σ-algebras, so $\varnothing$ is contained in all of the intersections
(2) If $A \in \sigma(\mathfrak{A})$, then $A \in \Sigma$ $\forall$ σ-algebras, but then $A^c \in \Sigma$ $\forall$ σ-algebras $\Rightarrow A^c \in \sigma(\mathfrak{A})$

The rest of the axioms for a σ-algebras are shown in an equivalent manner as in (2)

Example: (Borel σ-algebra)
Let $\mathcal{B}(S) = \sigma$(open subsets of S) (here we mean open in a topological sence since we need S to have a notion of opened-ness).
Since we mean open in a topological sense (which is defined as the complement of a closed set), we could have used the complement of a closed set to denote the open set, but since the complement is in the σ-algebra we may as well had the equivalent definition using the closed set all together.
This leads us to $\mathcal{B}(\mathbb{R}) = \sigma(\{(a, b) : a < b, a, b \in \mathbb{R}\})$. Instead of $\mathbb{R}$, any dense set could have worked as well (such as $\mathbb{Q}$)

Example:
Let $S = \{1, 2, 3 \cdots, 10\}$, and $\mathfrak{A} = \{\{1, 2\}, \{5\}\}$.
In order to generate a σ-algebra, we just need to recursively insert things that work with the axioms. For example, we need the empty set so we chuck in the empty set. We need the complement of the empty set so we chuck in the complement to the empty set. We need the complements to all the sets in $\mathfrak{A}$, so we add those as well, as well as their intersections.
We should then be left with just enough to call it a σ-algebra, and nothing more, hence the smallest σ-algebra:

$$\sigma(\mathfrak{A}) = \{\varnothing, S, \{1, 2\}, \{5\}, \{1, 2, 5\}, \{3, 4, 5, 6, 7, 8, 9, 10\}, \{1, 2, 3, 4, 6, 7, 8, 9, 10\}, \{3, 4, 6, 7, 8, 9, 10\}\}$$

Definition 1.6 π-system

A π-*system* on a set S is a collection of subsets π such that $\varnothing \in \pi$ and if $A, B \in \pi$ then $A \cap B \in \pi$

Sats 1.1

Suppose $\mathfrak{A} \subseteq \mathcal{P}(S)$ is a π-system and suppose that μ_1, μ_2 are measures on $(S, \sigma(\mathfrak{A}))$ such that $\mu_1(A) = \mu_2(A)$ $\forall A \in \mathfrak{A}$

$\Rightarrow$ Then $\mu_1 = \mu_2$ on $(S, \sigma(\mathfrak{A}))$

In other words, π-systems uniquely determine a measure.

Example:
Let $S = \mathbb{R}$, $\mathfrak{A} = \{[-\infty, a) : a \in \mathbb{R}\}$, $\sigma(\mathfrak{A}) = \mathcal{B}(\mathbb{R})$
$\mathfrak{A}$ is a π-system and have any measure is uniquely defined on $\mathfrak{A}$.
note that $\mu([-\infty, a))$ is nothing but thte cumulative distribution function of the measure μ (in terms of a).
"Measure up to a point". The following gives justification to construct measures from small collections.

Sats 1.2: Caratheodorys extension theorem

If Σ_0 is an algebra and $\mu_0 : \Sigma \to [0, \infty]$ is a σ-additive, $\exists!$ μ on $\Sigma = \sigma(\Sigma_0)$ such that $\mu(A) = \mu_0(A)$ $\forall A \in \Sigma_0$

An important consequence is that the Lebesgue measure is unique (only one notion of length on $\mathcal{B}(\mathbb{R})$) defined through sets of the form $A = (a_1, b_1) \cup \cdots \cup (a_n, b_n)$ (disjoint union of open sets)

$$\mathcal{L}(A) = |b_1 - a_1| + \cdots + |b_n - a_n|$$

2. Probability Spaces

Probability spaces are normally denoted by $(\Omega, \mathcal{E}, \mathbb{P})$ where:

- Ω is the space of realisations
- $\mathcal{E}$ is the sets of events
- $\mathbb{P}$ is the probability measure

Example:
$\Omega = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$, $\mathbb{P}(A = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2})dx$, $A = (a, b)$
This models a normally distributed real number.

2.1. **Almost sure events.**

We say that an event occurs *almost surely* if $\mathbb{P}(\mathcal{E}) = 1$ (equivalently $\mathbb{P}(\mathcal{E}^c) = 0$)

Proposition:
Let $E_1, \cdots \in \mathcal{E}$ be such that $\mathbb{P}(E_i) = 1 \quad \forall i \in \mathbb{N}$
Then, $\mathbb{P}\left(\bigcap_{i=1}^{\infty} E_i\right) = 1$

Bevis 2.1

Note that since each of them have probability measure 1, their complement must have measure 0 so:

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} E_i^c\right) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(E_i^c)0$$

However, since:

$$0 \leq \mathbb{P}\left(\bigcup_{i \in \mathbb{N}} E_i^c\right) \leq 0 \Rightarrow \mathbb{P}\left(\bigcup_{i \in \mathbb{N}} E_i^c\right) = 0$$

$$\Rightarrow \mathbb{P}\left(\left(\bigcup_{i \in \mathbb{N}} E_i^c\right)^c\right) = 1$$

But we have de-Morgans law, i.e $\bigcup_{i \in \mathbb{N}} E_i^c = \left(\bigcap_{i \in \mathbb{N}} E_i\right)^c$, which yields:

$$\left(\left(\bigcap_{i \in \mathbb{N}} E_i\right)^c\right)^c = \bigcap_{i \in \mathbb{N}} E_i$$

$\square$

Remark:
This applies only to countable unions. If uncountable, we could consider

$$\Omega = [0, 1], \quad \Sigma = \mathcal{B}([0, 1]), \quad \mathbb{P} = \mathcal{L}\mid_{[0,1]}$$

Then $\mathbb{P}(X = x) = 0$ (where X is some randomly chosen number and x is some fixed number). Taking the complement of this event yields $\mathbb{P}(X \neq x) = 1$ so $\mathbb{P}(X \neq x : x \in \mathbb{Q}) = 1$

2.2. **Liminf and limsup.**

Recall from real analysis:

$$\left.\begin{array}{l}\lim_{n\to\infty}\sup x_n = \lim_{n\to\infty}\sup_{m\geq n} x_n \\ \lim_{n\to\infty}\inf x_n = \lim_{n\to\infty}\inf_{m\geq n} x_n\end{array}\right\}\text{Limits exists in the extended reals and the limit exists iff limsup} = \text{liminf}$$

Recall that if $\lim_{n\to\infty}\sup x_n \geq x \Leftrightarrow \exists$ a subsequence $(x_n)_k$ with limit $\geq x$ and the opposite for liminf.

There exists a similar notion for sets.
Let $E_1, \cdots$ be events (sets)

$$\left.\begin{array}{l}\lim_{n\to\infty}\inf E_n = \bigcup_{n\geq 1}\bigcap_{m\geq n} E_n \\ \lim_{n\to\infty}\sup E_n = \bigcap_{n\geq 1}\bigcup_{m\geq n} E_n\end{array}\right\}$$

Some intuition here is definitely necessary.
For the first one, we are taking intersections of less and less sets (increasing sequence of sets), then finally unions. Think of this as events that eventually will appear

For the second one, it is decreasing (because of the intersection outside), all points will occur infinitely often.

Lemma 2.1: Fatous Lemma

Let $E_1, \cdots$ be events, then:

$$\mathbb{P}\left(\liminf_n E_n\right) \leq \liminf_n \mathbb{P}(E_n)$$

Bevis 2.2: Fatous Lemma

Let $F_n = \bigcap_{m\geq n} E_m$, i.e $E_n = \bigcup_{n\in\mathbb{N}} F_n$.
Here F_n is an increasing sequence of sets, which implies $F_n \in E_m$ $\forall m \geq n$, so $\mathbb{P}(F_n) \leq \mathbb{P}(E_m)$ $\forall m \geq n$
However, this also implies that $\mathbb{P}(F_n) \leq \inf_{m\geq n}\mathbb{P}(E_m)$

F_n is increasing $\Rightarrow$ probabilities are incfreasing $\Rightarrow \lim_{n\to\infty}\mathbb{P}(F_n)$ exists

$$\Rightarrow P\left(\bigcup_n^{\infty} F_n\right) = P\left(\liminf_n E_n\right)$$

$$\Rightarrow \lim_{n\to\infty}\mathbb{P}(F_n) \leq \lim_{n\to\infty}\inf \mathbb{P}(E_n) \qquad \text{by } \mathbb{P}(F_n) \leq \inf_{m\geq n}\mathbb{P}(E_m)$$

This yields finally $\mathbb{P}(\lim_n \inf E_n) \leq \lim_n \inf \mathbb{P}(E_n)$, which is what we wanted to prove. □

Note:
The reverse Fatous lemma can be proved by flipping everything (signs, inequalities, infimum to supremum etc.)

Lemma 2.2: Borel-Cantelli Lemma

Let $E_1, \cdots$ be a sequence of events such that $\sum_{n=1}^{\infty}\mathbb{P}(E_n) < \infty$

Then $\mathbb{P}(\lim_n \sup E_n) = 0 = \mathbb{P}(\text{"infinitely many } E_n \text{ occur"})$

Bevis 2.3: Borel-Cantelli Lemma

Recall what the limsup is, i.e $\lim_n \sup E_n = \bigcap_{n\in\mathbb{N}} \underbrace{\bigcup_{m\geq n} E_m}_{G_n}$

Note here that G_n is a decreasing sequence of sets, so $\lim_{n\to\infty} \sup E_n \subseteq G_n \quad \forall m \in \mathbb{N}$ and $\mathbb{P}(\lim_{n\to\infty \sup E_n}) \leq \mathbb{P}(G_m) \quad \forall m \in \mathbb{N}$

In particular, this is bounded above by:

$$\sum_{k=m}^{\infty} \mathbb{P}(E_k) \leq \mathbb{P}(\lim_{n\to\infty} \sup E_n)$$

But $\sum_{k=m}^{\infty} \mathbb{P}(E_k) \to 0$ as $m \to \infty$ since $\sum \mathbb{P}(E_n) < \infty$, so $\mathbb{P}(\lim_n \sup E_n) = 0$ $\square$

Example: (Coin toss)
Let E_n be the event that the first n coin toss in a sequence of tosses is heads. We have $\mathbb{P}(E_n) = 2^{-n}$ (assuming a fair coin) and $\sum_{n=1}^{\infty} \mathbb{P}(E_n) = 1 < \infty$ (since $\sum_{n=1}^{\infty} 2^{-n} \to 1$)
Thus, by the Borel-Cantelli lemma, $\mathbb{P}(\lim_n \sup E_n) = 0$ (finintely many values in which E_n occurs $\Rightarrow$ the run of heads will end almost surely)

3. Random Variables

Definition 3.7 Measurable functions

Let (S, Σ, μ) be a measure space. We say that $f : S \to \mathbb{R}$ is *measurable* if all pre-images of all Borel sets are in Σ:

$$f^{-1}(A) \in \Sigma \Rightarrow A \in \mathcal{B}(\mathbb{R})$$

Note:

$$f^{-1}(A) = \{s \in S : f(s) \in A\} \in \Sigma \quad \forall A \in \mathcal{B}(\mathbb{R})$$

- $m\Sigma$ are all measurable functions with respect to Σ
- $(m\Sigma)^+$ are all non-negative measurable functions with respect to Σ
- $b\Sigma$ are all bounded measurable functions with respect to Σ

Remark:
This can be generalized as functions $f : S \to T$ where (T, Σ', ν) is a measure space.

Lemma 3.1

We have:
(1) $f^{-1}(A^c) = (f^{-1}(A))^c$
(2) $f^{-1}\left(\bigcup_i A_i\right) = \bigcup_i f^{-1}(A_i)$
(3) $f^{-1}\left(\bigcap_i A_i\right) = \bigcap_i f^{-1}(A_i)$

Bevis 3.1

We shall only prove number 2, but the rest is proved in a similar manner:

$$\text{If } x \in f^{-1}\left(\bigcup_i A_i\right) \Leftrightarrow f(x) \in \bigcup_i A_i \Leftrightarrow \exists i \text{ s.t } f(x) \in A_i$$
$$\Leftrightarrow x \in f^{-1}(A_i) \Leftrightarrow x \in \bigcup_i f^{-1}(A_i)$$

□

Proposition:
If $f : S \to \mathbb{R}$ is continuous then it must also be measurable with respect to the Borel σ-algebra $\mathcal{B}(\mathbb{R})$

Bevis 3.2

Follows from topology, since pre-images of any open set is open as well as some help using the following: □

Proposition:

If $C \subseteq \mathcal{P}(\mathbb{R})$ is a collection such that $\sigma(C) = \overbrace{\mathcal{B}(\mathbb{R})}^{\text{codomain}}$, then $f : S \to \mathbb{R}$ is measurable with respect to $\underbrace{\mathcal{B}(S)}_{\text{domain}}$ if and only if $f^{-1}(A) \in \mathcal{B}(S) \quad \forall A \in C$

Examples:
In order to check whether $f : S \to \mathbb{R}$ is measurable, it suffices to check one of these:

- $f^{-1}(A) \in \Sigma \quad \forall A$ where A is an open set
- $f^{-1}((a, b)) \in \Sigma \quad \forall a < b \in \mathbb{R}$
- $f^{-1}([-\infty, a)) \in \Sigma \quad \forall a \in \mathbb{R}$
- Anything that generates $\mathcal{B}(\mathbb{R})$

Lemma 3.2

If $f_1, f_2 : S \to \mathbb{R}$ are measurable functions, then $f_1 + f_2$ is measurable

Bevis 3.3

We want to show that addition is measurable, we can do this by considering $(f_1+f_2)^{-1}((x,\infty)) \in \Sigma$ given that f_1, f_2 are measurable of course.

Since they are individually measurable, this means that the pre-images of this open set is in Σ, i.e

$$f_1^{-1}((x,\infty)), f_2^{-1}((x,\infty)) \in \Sigma$$

We use the fact that $x < f_1(s) + f_2(s) \Leftrightarrow \exists q \in \mathbb{Q}$ such that $q < f_1(s)$ and $x - q < f_2(s)$ (this reminds of the construction of the Dedekind sets, which is justified since there must be a rational number between x and $f_1 + f_2$ since we can just decrease the denominator to make a DIY ε)

$$\Rightarrow (f_1+f_2)^{-1}((x,\infty)) = \underbrace{\bigcup_{q\in\mathbb{Q}} \underbrace{\left(\underbrace{\overbrace{f_1^{-1}((q,\infty))}^{\in\Sigma}}_{s \text{ s.t } q < f_s(s)} \cap \underbrace{\overbrace{f_2^{-1}((q,\infty))}^{\in\Sigma}}_{x-q<f_2(s)} \right)}_{\in\Sigma}}_{\in\Sigma}$$

$\square$

Remark:
$\underbrace{f_1 \circ f_2}_{\text{"multiplication"}}$ is measurable by a similar proof. In fact, any infinite linear combination is measurable.

Lemma 3.3

Compositions of measurable functions is measurable

Bevis 3.4

$$(f_1 \circ f_2)^{-1}(A) = \underbrace{f_2^{-1} \circ \underbrace{f_1^{-1}(A)}_{\substack{\text{measurable}\\ \in\Sigma}}}_{\substack{\text{measurable}\\ \in\Sigma}}$$

$\square$

Lemma 3.4

If $f_n : S \to \mathbb{R}$ is a sequence of measurable functions $\forall n \in \mathbb{N}$, then

- $\inf_n f_n$
- $\sup f_n$
- $\lim_n \inf f_n$
- $\lim_n \sup f_n$

are measurable. Moreover, the event that it exists is measurable, i.e

$$\left\{ s \in S : \lim_{n\to\infty} f_n(s) \text{ exists and is finite} \right\} \in \Sigma$$

Bevis 3.5

Note that $(\inf_n f_n)^{-1}([x,\infty)) = \left\{ s \in S : \underbrace{\inf_n f_n(s) \in [x,\infty)}_{\Leftrightarrow \inf_n f_n(s) \geq x} \right\}$

Then all events have to be $\geq x$, i.e intersection:

$$\bigcap_{n\in\mathbb{N}} \underbrace{\{s \in S : f_n(s) \geq x\}}_{=f_n^{-1}([x,\infty))\in\Sigma} \in \Sigma$$

This can be concluded naturally since f_n is measurable, and hence $\inf_n f_n$ is measurable. Similar reasoning shows that $\sup_n f_n$ is measurable.

Note that $\lim_{n\to\infty} \inf f_n(s) = \sup_{n\in\mathbb{N}} \inf_{m\geq n} f_n(s)$ which is just a composition of measurable functions which we have shown is measurable $\Rightarrow \lim_n \inf f_n$ is measurable. Similar reasoning shows that $\lim_n \sup f_n$ is measurable.

The last statement in Lemma 3.4 can be decomposed into the following:

$$\left\{ s \in S : \lim_{n\to\infty} f_n(s) \text{ exists and is finite} \right\} \in \Sigma =$$

$$\left\{ s \in S : \lim_n \inf f_n(s) > -\infty \right\} \cap \left\{ s \in S : \lim_n \sup f_n(s) < \infty \right\} \cap \left\{ s \in S : \lim_n \inf f_n(s) = \lim_n \sup f_n(s) \right\}$$

This is measurable since all of the 3 sets are measurable (pre-images of open sets). Think of it in the following way:

- $> -\infty \Rightarrow (-\infty, \infty]$ which is an open set
- $< \infty \Rightarrow [-\infty, \infty)$ which is an open set
- $= \Rightarrow \{0\}$ which is an open set

Since compositions of intersections are measurable, the proof is complete. □

Definition 3.8 Random Variable

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
A measurable funcftion $X : \mathcal{F} \to \mathbb{R}$ is called a *random variable.*

Example:
Let $\Omega = \{1, \cdots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P} = \frac{1}{6}|A|$ (rolling a die)
Define $X(\omega) = \begin{cases} 1 & \omega \in \{1,3,5\} \quad \text{(odd)} \\ 0 & \omega \in \{2,4,6\} \quad \text{(even)} \end{cases}$
This is a random variable. One can verify this by checking pre-images of open sets of the range of the random variable $\{\varnothing, \{1,0\}, \{1\}, \{0\}\}$
By taking the discrete topology (collection of all subsets of S) $\mathcal{P}(S)$, this sneaky random variable is actually a random variable.
$Y(\omega) = \omega$ is also a random variable here since $\forall A \in \mathcal{F} = \mathcal{P}(\Omega) \subseteq \mathbb{R}$
Note that we have 2 distinct spaces, Ω could have not been a subset of $\mathbb{R}$, so Y would not have been a random variable since then $\mathcal{P}(\Omega) \not\subseteq \mathbb{R}$

Random variables "collapse" the space S due to their inherent injectivity. One way to measure this collapse is thinking of Borel σ-algebras in terms of this random variable. I.e, the smallest σ-algebra such that X is measurable

$$\overbrace{\sigma}^{\text{ensures }\sigma\text{-alg}} \left(\left\{ \underbrace{X^{-1}(A)}_{\text{ensure measurability}} : A \in \mathcal{B}(\mathbb{R}) \right\} \right) = \sigma(X)$$

In particular, $(\Omega, \sigma(X), \mathbb{P})$ is sufficient for X to be measurable (with respect to this space).

Example:

$$X(\omega) = \begin{cases} 1 & \omega \in \{1,3,5\} \quad \text{(odd)} \\ 0 & \omega \in \{2,4,6\} \quad \text{(even)} \end{cases} \qquad\qquad Y(\omega) = \omega$$

In this example,

$$\sigma(Y) = \sigma(\{Y^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}) = \sigma\overbrace{(\{\{1\}, \cdots, \{6\}\})}^{\mathcal{P}(\Omega)=\mathcal{F}}$$

$$\sigma(X) = \{\text{pre-images of neighborhoods of } 1 \ \& \ 0\} = \{\varnothing, \Omega, \{1,3,5\}, \{2,4,6\}\}$$

The last one may be difficult to grasp, but think of it like constructing the following set

$$\{\{\text{neither } 0 \vee 1 = \varnothing\}, \{\text{both } 0 \vee 1\}, \{\text{pre-image of } 1\}, \{\text{pre-image of } 0\}\}$$

This yields the smallest σ-algebra that contains these but still is $\neq \mathcal{F}$

Knowing nothing about a measurable/probability space is not possible, we always know things such as $\mu(\Omega) = 1$ and $\mu(\varnothing) = 0$. We could say that "if we know nothing, then we know those 2 things"
Conversely, if we know that $\mathcal{P}(\Omega)$ or $\mathcal{F}$ is a σ-algebra, then we know everything (we know the probability of every event happening).
For a constant random variable X, we know nothing (pre-images is $\{\varnothing, \Omega\}$). We can encode information using this.

Definition 4.9 Law

Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. A *law* $\mathcal{L}_X(A)$