Integration and probability

Marco Biroli

October 19, 2020

Contents

Chapter 1

Introduction

1.1 Motivation

This courses aims at answering some basic question of measure theory. Examples are defining what is the length of a part of $\mathbb{R}$ or the surface of a part of $\mathbb{R}^2$. In a second time we will introduce integrals in the most general way in order to allow the integration of as many functions as possible. Then define and precise the mathematical structure allowing us to describe an infinite series of fair coins throws. Notice that the definition of an integral can follow from the definition of sections of $\mathbb{R}^2$ since we can characterize an integral by the area enclosed by the curve. Furthermore integrals and measures on intervals are also equivalent since integrating the identity over an interval should give the same value as the measure of that interval.

1.2 Initial definitions

Definition 1.2.1 (Countable Set)**.** A set S is said to be countable if and only if there exists a bijection in between S and $\mathbb{N}$.

Proposition 1.2.1. *All parts of a countable set are countable.*

Proof. We can simply start indexing by the smallest element. □

Proposition 1.2.2. *The image of a sequence is countable.*

Proof. By simply indexing the image by their antecedents we get a bijection.

□

Remark. $\mathbb{N}^2$ *is countable. The bijection is given by* $(n_1, n_2) \mapsto 2^{n_1}(2n_2 + 1) - 1$.

Proposition 1.2.3. *A countable union of countable (or finite) sets is countable (or finite).*

Proof. Let A_i be countable parts of a set X. For all i there exists a bijection $b_i : \mathbb{N} \to A_i$. And hence the bijection given by $(i, j) \mapsto b_i(j)$ maps $\mathbb{N}^2 \to \cup_i A_i$.

□

Proposition 1.2.4. *If X is countable the powerset $\mathcal{P}(X)$ is not. Or more generally no bijection exists in between X and $\mathcal{P}(X)$.*

Proof. Suppose by contradiction that there exists a bijective mapping $x \mapsto A_x$ of X to $\mathcal{P}(X)$. Now we consider the set $B := \{x : x \notin A_x\}$. Now suppose consider $A_y = B$ then this means that $y \in B \Leftrightarrow y \in A_y \Leftrightarrow y \notin B$. □

Definition 1.2.2 ($\limsup$)**.** Let $(x_n)_{n\in\mathbb{N}}$ be a sequence. Then introduce the sequence $s_n := \sup_{k\geq n} x_k$ then the sequence $(s_n)_{n\in\mathbb{N}}$ is naturally decreasing and therefore admits a limit in $[-\infty, \infty]$. Then the limit of $(s_n)_{n\in\mathbb{N}}$ is called the limsup and we denote:

$$\limsup_{n\to+\infty} x_n = \lim_{n\to+\infty} s_n = \inf_{n\in\mathbb{N}} s_n$$

Definition 1.2.3 ($\liminf$)**.** An equivalent definition gives:

$$\liminf_{n\to+\infty} x_n = \lim_{n\to+\infty} (\inf_{k\geq n} x_k) = \sup_{n\in\mathbb{N}} \inf_{k\geq n} x_k$$

Remark. *The advantage of the* $\limsup$ *and* $\liminf$ *with respect to the* $\lim$ *is that they are always defined even if they might diverge. It is furthermore trivial to see that* $\liminf x_n \leq \limsup x_n$.

Proposition 1.2.5. *The limit of* $(x_n)_{n\in\mathbb{N}}$ *exists if and only if* $\limsup x_n = \liminf x_n$.

Proof.
$$\forall n \in \mathbb{N}, i_n \leq x_n \leq s_n \Rightarrow \liminf x_n = \lim i_n \leq \lim x_n \leq \lim s_n = \limsup x_n$$
Hence if $\liminf x_n = \limsup x_n$ the $\lim x_n = \liminf x_n = \limsup x_n$. Inversely if $\lim x_n = \ell$ then from definition $\forall \varepsilon > 0, \exists N, s.t. \forall n \geq N, \ell - \varepsilon \leq x_n \leq l + \varepsilon$ and hence $s_n \to \ell$ and $i_n \to \ell$. □

Proposition 1.2.6. *If* $(y_n)_{n\in\mathbb{N}}$ *is a subsequence of* $(x_n)_{n\in\mathbb{N}}$ *then* $\liminf x_n \leq \liminf y_n \leq \limsup y_n \leq \limsup x_n$. *Hence any adherence point of* $(x_n)_{n\in\mathbb{N}}$ *is contained in the interval* $[\liminf x_n, \limsup x_n]$.

Proposition 1.2.7. *There exists a subsequence of* $(x_n)_{n\in\mathbb{N}}$ *which converges to the* $\limsup x_n$ *(or equivalently the* $\liminf x_n$*)*

Proof. We can choose a $k_n \geq n$ such that $s_n - \frac{1}{n} \leq x_{k_n} \leq s_n$ and $k_n > k_{n-1}$ and then the sequence $(x_{k_n})_{n\in\mathbb{N}}$ converges to $\limsup x_n$.

□

Definition 1.2.4 (Series)**.** Let $a_i, i \in \mathcal{I}$ be a not necessarily countable family of positive reals. Then we define the infinite sum as:
$$\sum_{i\in\mathcal{I}} a_i := \sup_{F\in\mathcal{I}, F \text{ finite}} \sum_{i\in F} a_i$$

Proposition 1.2.8. *If* $\sum_{i\in\mathcal{I}} a_i < +\infty$ *then the set* $\{i : a_i > 0\}$ *is finite or countable.*

Proof. Notice that:
$$\{i : a_i > 0\} \subset \bigcup_{k\in\mathbb{N}} \{i : a_i > \frac{1}{k}\}$$
Then each one of the sets in the union has the cardinality bounded by:
$$\#\{i : a_i > \frac{1}{k}\} \leq k \sum_{i\in\mathcal{L}} a_i$$

□

Proposition 1.2.9. *We now consider* $\mathcal{I}$ *countable. Then we have a bijection* $\sigma : \mathbb{N} \to \mathcal{I}$ *then:*
$$\sum_{i\in\mathcal{I}} a_i = \lim_{n\to\infty} \sum_{k=1}^{n} a_{\sigma(k)} =: \sum_{k=1}^{+\infty} a_{\sigma(k)}$$

Proof. We have that $\forall F \subset \mathcal{I}$ finite then $\sigma^{-1}(F)$ is finite, and therefore bounded for a certain N. Hence:
$$\sum_{i\in F} a_i = \sum_{k\in\sigma^{-1}(F)} a_{\sigma(k)} \leq \sum_{k=1}^{N} a_{\sigma(k)} \leq \sum_{k=1}^{+\infty} a_{\sigma(k)}$$
Now taking the sup for all F we get that:
$$\sum_{i\in\mathcal{I}} a_i \leq \sum_{k=1}^{+\infty} a_{\sigma(k)}$$
Then the other way around we simply have that:
$$\sum_{k=1}^{n} a_{\sigma(k)} = \sum_{i\in\sigma([\![1,n]\!])} a_i \leq \sum_{i\in\mathcal{I}} a_i$$
And finally be taking the limit this gives:
$$\sum_{k=1}^{+\infty} a_{\sigma(k)} \leq \sum_{i\in\mathcal{I}} a_i$$

□

Corollary 1.2.0.1. *Notice that the above is independent of* σ *and hence we trivially see that:*
$$\forall a_k \geq 0, \forall \sigma \in Aut(\mathbb{N}), \quad \sum_{k=1}^{+\infty} a_k = \sum_{k=1}^{+\infty} a_{\sigma(k)}$$

Proposition 1.2.10.

$$\sum_{(i,j)\in\mathcal{I}} a_{ij} = \sum_{i=1}^{+\infty}\left(\sum_{j=1}^{+\infty} a_{ij}\right) = \sum_{j=1}^{+\infty}\left(\sum_{i=1}^{+\infty} a_{ij}\right)$$

Proof. $F \subset \mathcal{I}$ finite implies that $\exists N$ such that $F \in [\![1, N]\!]^2$. Hence we have that:

$$\sum_{(i,j)\in F} a_{ij} \leq \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} \leq \sum_{i=1}^{N}\sum_{j=1}^{+\infty} a_{ij} \leq \sum_{i=1}^{+\infty}\left(\sum_{j=1}^{+\infty} a_{ij}\right)$$

Now taking the sup on all F we get that:

$$\sum_{(i,j)\in\mathbb{I}} \leq \sum_{i=1}^{+\infty}\left(\sum_{j=1}^{+\infty} a_{ij}\right)$$

Now the other way around we have that:

$$\forall N, \forall M, \sum_{i=1}^{N}\sum_{j=1}^{M} a_{ij} \leq \sum_{(i,j)\in\mathbb{N}^2} a_{ij}$$

Taking the limit as $M \to +\infty$ with N fixed we get:

$$\forall N, \sum_{i=1}^{N}\sum_{j=1}^{+\infty} a_{ij} \leq \sum_{(i,j)\in\mathbb{N}^2} a_{ij}$$

Now taking the limit $N \to +\infty$ we get:

$$\sum_{i=1}^{+\infty}\left(\sum_{j=1}^{+\infty} a_{ij}\right) \leq \sum_{(i,j)\in\mathbb{N}^2} a_{ij}$$

Now by symmetry the third equality follows. □

Definition 1.2.5 (Absolute convergence)**.** Let $a_i, i \in \mathcal{I}$ be a not necessarily countable family of reals then the series is said to be absolutely convergent if and only if:

$$\sum_{i\in\mathcal{I}} |a_i| < +\infty$$

Proposition 1.2.11. *Now we define $a_i^+ = \max(a_i, 0)$ and $a_i^- = \max(-a_i, 0)$ then we have that $\forall i, a_i = a_i^+ - a_i^-$ and $|a_i| = a_i^+ + a_i^-$. If the family is absolutely convergent we have that:*

$$\sum_{i\in\mathcal{I}} a_i^+ - \sum_{i\in\mathcal{I}} a_i^- = \sum_{k=1}^{+\infty} a_{\sigma(k)}$$

Proof. We have that:

$$\sum_{k=1}^{n} a_{\sigma(k)} = \sum_{k=1}^{n} a_{\sigma(k)}^+ - a_{\sigma(k)}^- = \sum_{k=1}^{n} a_{\sigma(k)}^+ - \sum_{k=1}^{n} a_{\sigma(k)}^- \overset{n\to+\infty}{\longrightarrow} \sum_{k=1}^{+\infty} a_{\sigma(k)}^{+\infty} - \sum_{k=1}^{+\infty} a_{\sigma(k)}^-$$

□

Corollary 1.2.0.2. *From this we can deduce two corollaries when the terms are absolutely convergent we have that:*

- $\sum_{k=1}^{+\infty} a_k = \sum_{k=1}^{+\infty} a_{\sigma(k)}$
- $\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} a_{ij} = \sum_{j=1}^{+\infty}\sum_{i=1}^{+\infty} a_{ij}$

Definition 1.2.6 (Algebra of sets)**.** Let X be a set. We say that $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra of sets if it is stable by finite union and intersection and complement and $\emptyset, X \in \mathcal{A}$.

Definition 1.2.7 (Tribe or σ-algebra)**.** Let X be a set. We say that $\mathcal{A} \subset \mathcal{P}(X)$ is a tribe or σ-algebra if it is stable by countable union and intersection and complement and $\emptyset, X \in \mathcal{A}$.

Remark. *We give here a few examples. Trivially $\mathcal{P}(X)$ is a tribe. The if we have a finite partition of X:*

$$X = \bigcup_k X_k \quad \textit{where the} \quad X_k \quad \textit{are disjoint.}$$

Then any set $A \subset X$ of the form:

$$\forall \mathcal{I} \subset [\![1, l]\!], A = \bigcup_{i \in \mathcal{I}} X_i$$

Is a finite tribe.

Lemma 1.2.1. *All finite algebra is associated to a finite partition of the set.*

Proof. Let $\mathcal{A}$ be a finite algebra on X. Then we define:

$$\forall x \in X, A(x) := \bigcap_{A \in \mathcal{A}, x \in A} A$$

Now for x and y given we necessarily have that:

$$A(x) = A(y) \quad \text{or} \quad A(x) \cap A(y) = \emptyset$$

To show this formally let $x \in X$ and $B \in \mathcal{A}$. Then if $x \in B$ we must have that $A(x) \subset B$. Identically if $x \in B^c$ then we must have that $A(x) \subset B^c$ or in other words $A(x) \cap B = \emptyset$.

□

Definition 1.2.8 (Additive measure)**.** Let $\mathcal{A} \in \mathcal{P}(X)$ be an algebra and $m : \mathcal{A} \to [0, +\infty]$ a function. Then we say that m is an additive measure if:

- $m(\emptyset) = 0$
- If $A \cap B = \emptyset$ then $m(A \cup B) = m(A) + m(B)$

Definition 1.2.9 (Measure or σ-additive measure)**.** Let $\mathcal{F} \subset P(X)$ be a tribe and $m : \mathcal{F} \to [0, +\infty]$ a function. We say that m is a measure if:

- $m(\emptyset) = 0$.
- If given a countable family of disjoint sets $(A_i)_{i \in \mathbb{N}}$ we have that $m(\bigcup_{i=1}^{+\infty} A_i) = \sum_{i=1}^{+\infty} m(A_i)$

Remark. *Notice that the notation is confusing: a measure is stronger than an additive measure. Which is why we sometimes refer to measures as σ-additive measures in order to underline the difference.*

Proposition 1.2.12. *When $m : \mathcal{A} \to [0, +\infty]$ is an additive measure on an algebra the following properties are equivalent:*

1. *If $A_i \in \mathcal{A}$ are a countable disjoint family and $\bigcup_i A_i \in \mathcal{A}$ then $m(\bigcup_i A_i) = \sum_i m(A_i)$*
2. *If $A, A_i \in \mathcal{A}$ and $A \subset \bigcup_i A_i$ then $m(A) \leq \sum_i m(A_i)$.*

In such a case we say that m is σ-additive on $\mathcal{A}$.

Proof. We start by proving (1) $\Rightarrow$ (2). Let $A_i \in \mathcal{A}$ we define $\tilde{A}_i$ by:

$$\tilde{A}_1 = A_1, \quad \tilde{A}_n = A_n - \tilde{A}_{n-1}$$

Then the family $\tilde{A}_i$ is disjoint by construction and $\bigcup_i A_i = \bigcup_i \tilde{A}_i$. Now for $A \subset \bigcup_i A_i$ we have that:

$$A \subset \bigcup_i \tilde{A}_i \Rightarrow A = \bigcup_i \left(\tilde{A}_i \cap A \right) \Rightarrow m(A) = \sum_i m(\tilde{A}_i \cap A) \leq \sum_i m(A_i)$$

Now the other implication (2) $\Rightarrow$ (1) follows from the opposite equality:

$$A = \bigcup A_i \Rightarrow m(A) \leq \sum_i m(A_i)$$

Which is true since:

$$\forall n, \quad \bigcup_{i=1}^{n} A_i \subset A$$

Hence $m(A) \geq \sum_{i=1}^{n} m(A_i)$ and hence at the limit $n \to +\infty$ we get:

$$m(A) \geq \sum_{i=1}^{+\infty} m(A_i)$$

□

Definition 1.2.10 (Image and pre-image of algebras of sets and tribes)**.** Let $f : \Omega \to X$ an application. Then if $\mathcal{A}$ is an algebra (or respectively tribe) on Ω, then we define the image algebra (or respectively tribe) by:

$$f_\star \mathcal{A} = \{A \subset X, f^{-1}(A) \in \mathcal{A}\}$$

Inversely if $\mathcal{A}$ is an algebra (or respectively tribe) on X then the pre-image algebra (or respectively tribe) is given by:

$$f^\star \mathcal{A} = \{f^{-1}(A), A \in \mathcal{A}\}$$

Proof. The fact that these are indeed algebras of sets or tribes follow directly from the basic properties of image and pre-image sets:

$$f^{-1}(A^c) = f^{-1}(A)^c \quad \text{and} \quad f^{-1}(\bigcap_{i \in \mathcal{I}} A_i) = \bigcap_{i \in \mathcal{I}} f^{-1}(A_i) \quad \text{and} \quad f^{-1}(\bigcup_{i \in \mathcal{I}} A_i) = \bigcup_{i \in \mathcal{I}} f^{-1}(A_i)$$

□

Definition 1.2.11 (Image measure)**.** Let $f : (\Omega, \mathcal{A}, m) \to X$ be an application. We define the image measure or law as the measure defined by:

$$(f_\star m)(Y) := m(f^{-1}(Y)) \quad \text{defined on} \quad f_\star \mathcal{A}$$

Definition 1.2.12 (Finite and probability measures)**.** A measure m on X is said to be finite if $m(X) < +\infty$ and it is said to be of probability if $m(X) = 1$.

Definition 1.2.13. An application $f : (\Omega, \tau) \to (X, \mathcal{T})$ is said to be measurable if $\forall Y \in \mathcal{T}, f^{-1}(Y) \in \tau$ or in other words: $\mathcal{T} \subset f_\star \tau$ or $f^\star \mathcal{T} \subset \tau$.

Chapter 2

[...] Note Gap [...]

Chapter 3

Central limit theorem.

3.1 [...] Note Gap [...]

3.2

Lemma 3.2.1. *We start with a Lemma which will be useful for proving the Hoeffding inequality. Let $\tilde{f}$ be a centered random variable (taking only a finite amount of values) then, for $\theta > 0$, we have:*

$$\mathbb{E}[e^{\theta \tilde{f}}] \leq e^{\frac{C\theta^2}{8}} \quad \textit{where} \quad C = (\max \tilde{f} - \min \tilde{f})^2$$

Proof. We take:

$$g(\theta) := \ln\Big(\mathbb{E}[e^{\theta \tilde{f}}]\Big)$$

Then g is a C^∞ function. Notice the following properties:

$$g(0) = 0 \quad \text{and} \quad g'(\theta) = \frac{\mathbb{E}[\tilde{f} e^{\theta \tilde{f}}]}{\mathbb{E}[e^{\theta \tilde{f}}]} \quad \text{hence} \quad g'(0) = E[\tilde{f}] = 0$$

Hence we know that close to $\theta = 0$ we will have $g(\theta) \leq c\theta^2$ for some c. In order to get the value of the constant we need to compute the second derivative:

$$g''(\theta) = \frac{\mathbb{E}[\tilde{f}^2 e^{\theta \tilde{f}}]\mathbb{E}[e^{\theta \tilde{f}}] - \mathbb{E}[\tilde{f} e^{\theta \tilde{f}}]^2}{\mathbb{E}[e^{\theta \tilde{f}}]^2} = \mathbb{E}\left[\tilde{f}^2 \frac{e^{\theta \tilde{f}}}{\mathbb{E}[e^{\theta \tilde{f}}]}\right] - \mathbb{E}\left[\tilde{f} \frac{e^{\theta \tilde{f}}}{\mathbb{E}[e^{\theta \tilde{f}}]}\right]^2$$

Now the trick is to notice that this is the variance of a probability law on $\tilde{Y}$ given by:

$$P_\theta(\tilde{y}) = \frac{P(\tilde{y}) e^{\theta \tilde{y}}}{\mathbb{E}[e^{\theta \tilde{f}}]}$$

Which sums indeed to one and hence is a well defined probability. Now notice the trivial remark: $\mathrm{Var}(X) = (\max X - \min X)^2/4$ for any random variable X taking a finite amount of values. This follows from the fact that we can change the value of X by a constant without changing the variance in order to get that $\max X = -\min X$ then we have that:

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \leq \mathbb{E}[X]^2 \leq (\max X)^2 = \left(\frac{\max X - \min X}{2}\right)^2$$

Hence we get that $g''(\theta) \leq \frac{C}{4}$ which gives $g(\theta) \leq \frac{c}{8}\theta^2$. ☐

Theorem 3.2.2 (Inequality of Hoeffding)**.** *The inequality is stated as:*

$$P(|\tilde{S_n}| \geq a) \leq 2\exp\left(-\frac{2a^2}{Cn}\right) \quad \textit{where} \quad C = (\max f - \min f)^2$$

Proof. To prove the Hoeffding inequality we use the Lemma stated above to evaluate (let $\theta > 0$):

$$P(\tilde{S_n} \geq a) = P(e^{\theta \tilde{S_n}} \geq e^{\theta a}) \leq e^{-\theta a}\mathbb{E}[e^{\theta \tilde{S_n}}] = e^{-\theta a}\prod_{i=1}^{n} \mathbb{E}[e^{\theta \tilde{f}_i}] = e^{-\theta a}\mathbb{E}[e^{\theta \tilde{f}}]^n \leq \exp\left(\frac{nC\theta^2}{8} - \theta a\right)$$

Now the choice of θ is still free hence to make the bound as tight as possible we simply need to minimize the argument of the exponential. The minimum is obtained for:

$$\theta = \theta_\star = \frac{4a}{nC}$$

Then replacing up top we obtain the bound:

$$P(\tilde{S_n} \geq a) \leq \exp\left(\frac{2a^2}{nC} - \frac{4a^2}{nC}\right) = \exp\left(-\frac{2a^2}{nC}\right)$$

The same reasoning applies for $P(\tilde{S_n} \leq -a)$ which gives the prefactor 2 required. □

Proposition 3.2.1. *The Hoeffding inequality gives another proof of the Central Limit Theorem.*

Proof. We write:

$$P(|\tilde{S_n}| \geq \varepsilon n^\alpha) \leq 2\exp\left(-\frac{2\varepsilon^2}{C}n^{2\alpha-1}\right)$$

Then for ε fixed the r.h.s. is the term of a convergent series in n. Hence from the first Borel-Cantorelli Lemma we have that almost surely all the l.h.s. is satisfied in other words:

$$P(\tilde{S_n} \geq \varepsilon n^\alpha \quad \text{for an infinite amount of} \quad n) = 0 \Rightarrow \frac{\tilde{S_n}}{n^\alpha} \xrightarrow{a.s} 0$$

□

3.3 Lower bounds.

Theorem 3.3.1. *We have that $P(S_n \quad bounded\ ) = 0$, supposing that the law of f is non-trivial.*

Proof. We fix $a \neq 0$ such that $P(a) > 0$. Then we want to show that $\forall k, \exists n, s.t. f_{n+1} = a, \cdots, f_{n+k} = a$. Denote by $B_n = \{f_{n+1} = a, \cdots, f_{n+k} = a\}$ then we have from independence that $P(B_n) = P(f = a)^k$. The events B_n themselves though are not independent however the events $A_n = B_{kn}$ are. Hence from the second Borel Cantelli lemma we have that $P(A_n \quad \text{i.o.}) = 1$ which tells us that $P(\bigcup_n A_n) = 1$. □

Lemma 3.3.2. *For $a \in Imf$ such that $P(f = a) > 0$ and T the first time for which $f_T = a$ then $\mathbb{E}[T] = \frac{1}{P(f=a)}$.*

Proof. This is simply the expectancy of a geometric law. □

Proposition 3.3.1. *We have that $\forall A > 0, \exists \delta > 0, s.t. \forall n, P(\tilde{S_n} \geq A\sqrt{n}) \geq \delta$.*

Proof. We prove the proposition only in the simplest case where $Y = \{0,1\}$ and $P(f = 1) = \rho$ however the reasoning can be generalized. Then we start by proving the following result:

$$P_\rho(S_n = k) \geq \frac{1}{C\sqrt{n}}\exp\left(-nC|\frac{k}{n} - \rho|^2\right)$$

If we look at:

$$\frac{P_\rho(S_n = k)}{P_{\frac{k}{n}}(S_n = k)} = \frac{\rho^k(1-\rho)^{n-k}}{(\frac{k}{n})^k(1-\frac{k}{n})^{n-k}} = \exp\left(-nh(\frac{k}{n},\rho)\right) \quad \text{where} \quad h(s,\rho) = s\ln\frac{s}{\rho} + (1-s)\ln\frac{1-s}{1-\rho}$$

Then one can notice that:

$$h(s,\rho) \leq C|s-\rho|^2 \quad \text{where} \quad C \quad \text{can depend on} \quad \rho$$

Then plugging this back on top we get:

$$P_\rho(S_n = k) \geq P_{\frac{k}{n}}(S_n = k)\exp\left(-nC|\frac{k}{n} - \rho|^2\right)$$

Now using the Stirling formula we can bound the prefactor as desired and it gives the desired result. Then using this result we get that:

$$P(\tilde{S}_n \geq A\sqrt{n}) = \sum_{k \geq n\rho - A\sqrt{n}} P(S_n = k) \geq \sum_{n\rho + A\sqrt{n} \leq k \leq n\rho + (A+\frac{1}{2})\sqrt{n}} P(S_n = k) \quad \text{where} \quad \tilde{S}_n = S_n - n\rho$$

Then in the sum we have $\sqrt{n}$ terms and we can minorate each term with the previous result which gives:

$$P(\tilde{S}_n \geq A\sqrt{n}) \geq \sqrt{n}\left(\frac{1}{C\sqrt{n}}\exp\left(-nC\frac{(A+\frac{1}{2})^2}{n}\right)\right) \geq \frac{1}{C}\exp\left(-C(A+\frac{1}{2})^2\right)$$

□

Theorem 3.3.3. *We have that* $P(\limsup_{n\to+\infty} \frac{\tilde{S_n}}{\sqrt{n}} = +\infty) = 1$.

Proof. We take $A \geq -\min f$. We consider the subsequence n_k such that $n_1 = 1$ and $n_{k+1} = n_k + 4n_k^2$. Then we have that:

$$P(|\tilde{S}_{n_{k+1}} - \tilde{S}_{n_k}| \geq a) = P(|\tilde{S}_{n_{k+1}-n_k}| \geq a)$$

Then we consider the events:

$$B_k = \{\tilde{S}_{n_{k+1}} - \tilde{S}_{n_k} \geq A\sqrt{n_{k+1} - n_k}\} = \{\tilde{S}_{4n_k} \geq 2An_k\}$$

Now notice that the B_k are independent and from the proposition above we have that $P(B_k) \geq \delta$. Hence the second Borel Cantorelli Lemma tells us that $P(B_k \text{ i.o.}) = 1$. Therefore the exists an infinite amount of k such that:

$$\tilde{S}_{n_{k+1}} \geq \tilde{S}_{n_k} + 2An_k \geq n_k(2A + \min f) \geq n_k A \quad \text{since} \quad A \geq -\min f$$

Furthermore we have that $n_{k+1} \leq 16n_k^4$ hence $n_k \leq \frac{\sqrt{n_{k+1}}}{4}$. Therefore we get that:

$$\tilde{S}_{n_{k+1}} \geq \frac{A}{4}\sqrt{n_{k+1}} \quad \text{and so} \quad \limsup_{n\to+\infty} \forall A > 0, a.s., \frac{\tilde{S}_n}{\sqrt{n}} \geq \frac{A}{4}$$

Hence taking the intersection on discretized sequence of A we get that:

$$a.s. \limsup_{n\to+\infty} \frac{\tilde{S}_n}{\sqrt{n}} = +\infty$$

□

Theorem 3.3.4. *For general culture know that:*

$$\limsup \frac{\tilde{S}_n}{\sqrt{2\,Var[f]n\ln(\ln(n))}} \xrightarrow{a.s.} 1$$

3.4 Remarks on σ-algebras.

Definition 3.4.1 (Generated tribe)**.** Let $B \in \mathcal{P}(X)$ we call the generated tribe by B the smallest tribe containing B.

Lemma 3.4.1. *Such a tribe always exists since an intersection of tribes is a tribe.*

Proof. Let $\mathcal{A}_i$ for $i \in \mathcal{I}$ be tribes on X. Then $\mathcal{A} = \cap_{i\in\mathcal{I}}\mathcal{A}_i$ is a tribe since: $\forall A_n \in \mathcal{A}, \forall i \in \mathcal{I}, A_n \in A_i \Rightarrow \cup_n A_n \in \mathcal{A}_i$ and hence $\cup_n A_n \in \cap_i \mathcal{A}_i$. □

Theorem 3.4.2. *Let $\mathcal{A}$ be an algebra of parts of X and m an additive probability measure on $\mathcal{A}$, which is also a σ-additive algebra. Then there exists a unique probability measure m on τ the generated tribe from $\mathcal{A}$.*

Proposition 3.4.1. *Let $f : (\Omega, \tau) \to (X, \mathcal{T})$ an application and let $\mathcal{B} \subset \mathcal{T}$ generating $\mathcal{T}$. If $f^{-1}(B) \in \tau, \forall B \in \mathcal{B}$, then f is measurable.*

Proof. We consider the set $\mathcal{E}$ of parts of X such that the preimage of f is measurable: $\mathcal{E} = \{Y \in \mathcal{P}(X), f^{-1}(Y) \in \tau\}$. We have already shown that is a tribe (denoted by $f_\star\tau$). Then this tribe contains B hence it contains $\mathcal{T}$. □

Proposition 3.4.2. *The following part of $\mathbb{R}$ generate the same tribe (the Borel tribe):*

- *The intervals*
- *The closed intervals*
- *The open intervals*
- *The open sets*
- *The closed sets*
- *The intervals $]-\infty, a]$ where $a \in \mathbb{D}$ (the decimals).*

Proposition 3.4.3. *On $[0, 1[$ the following are the same tribes:*

- *The tribe generated from the intervals*
- *The tribe generated from $[0, a]$ for $a \in \mathcal{D} \cap [0, 1[$.*
- *The tribe $\{B \cap [0, 1[, B \in \mathcal{B}(\mathbb{R})\} = i^\star\mathcal{B}$. Where $i : [0, 1[\to \mathbb{R}$ is the inclusion.*

3.5 Product tribes

Definition 3.5.1. Let $f_i : \Omega \to (X_i, \mathcal{T}_i)$ for $i \in \mathcal{I}$ a family of applications. The tribe generated by the f_i is the tribe of Ω generated by $\cup_{i\in\mathcal{I}} f_i^\star(\mathcal{T}_i)$. It is the smallest tribe for which all the functions f_i are measurable.

Definition 3.5.2 (Product tribe)**.** Let $(X_i, \mathcal{T}_i)$ be measurable spaces for $i \in \mathcal{I}$. The product tribe on $\Omega = \prod_{i\in\mathcal{I}} X_i$ is the tribe generated by the projections: $\pi_i : \Omega \to X_i$.

Remark. *If $\mathcal{I}$ is finite or countable the product tribe is the tribe generated by the products $\prod_i Z_i$.*

Proposition 3.5.1.

- *Take $f_i : \Omega \to (X_i, \mathcal{T}_i)$ and $F : \Omega \to (\prod X_i, \prod \mathcal{T}_i)$ the application which coordinates are given by the f_i then the tribe generated by the f_i is $F^\star(\prod \mathcal{T}_i)$.*
- *The application F is measurable if and only if its coordinates are.*

Proposition 3.5.2 (Tribes of $\mathbb{R}^d$.)**.** *The following tribes are identical:*

- *The tribe generated by the open sets*
- *The tribe generated by the closed sets*
- *The product of the Borelian tribes*
- *The tribe generated by the product of intervals*
- *The tribe generated by the products $]-\infty, a_1] \times \cdots \times]-\infty, a_d]$ where $a_i \in \mathbb{D}$.*