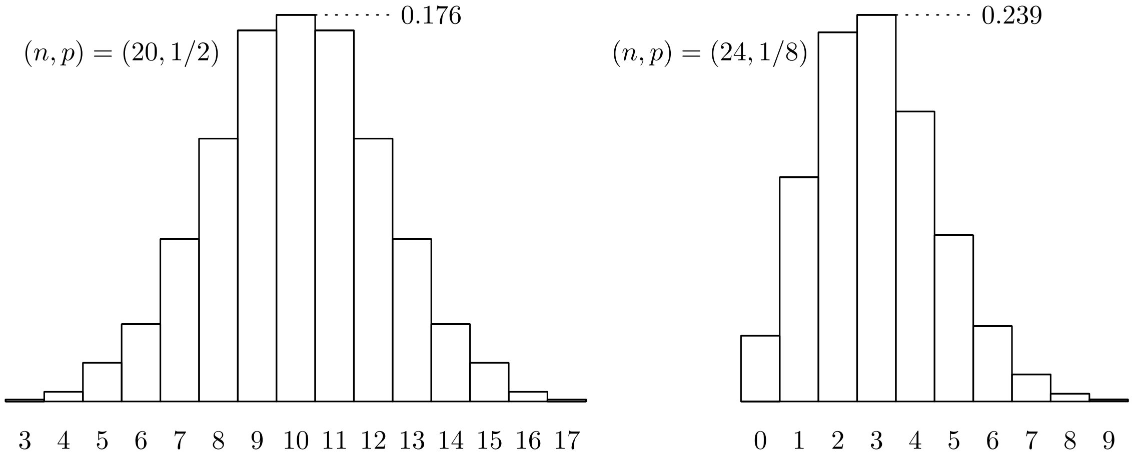
Binomial distrbution, CLT, Martingales, and Brownian Motion[1]

Nobuo Yoshida[2]



[1]Notes for a presentation at Shinshu University on February 24, 2021.

[2][e-mail] `noby@math.nagoya-u.ac.jp`, [URL] `http://www.math.nagoya-u.ac.jp/~noby`

Contents

0.1 Elementaly distributions . . . 3
0.2 The Law of Large Numbers . . . 9
0.3 Characteristic functions . . . 11
0.4 Weak Convergence . . . 16
0.5 Martingales . . . 18
0.6 Brownian Motion . . . 24
0.7 Continuity of the Brownian Motion . . . 25
0.8 Germ triviality . . . 26

The topics above are selected from [Yos].

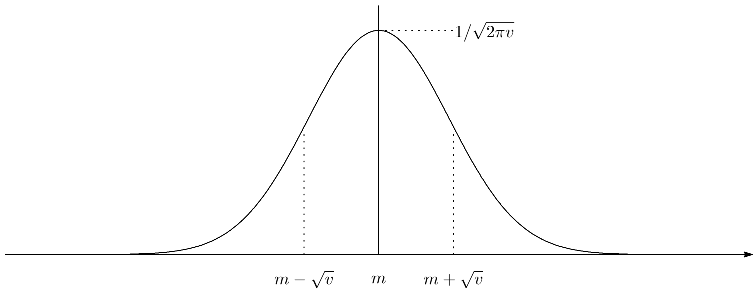
0.1 Elementaly distributions

Example 0.1.1 (**Normal distribution**) Let $m \in \mathbb{R}$ and $v > 0$.
► A r.v. $X : \Omega \to \mathbb{R}$ is called a (m, v)-**normal** r.v. if

$$P(X \in B) = \frac{1}{\sqrt{2\pi v}} \int_B \exp\left(-\frac{(x-m)^2}{2v}\right) dx \quad \text{for } B \in \mathcal{B}(\mathbb{R}). \tag{0.1}$$

The law of an (m, v)-normal r.v. is denoted by $N(m, v)$. It is not difficult to see that

$$EX = m, \quad \text{var}\, X = v.$$



In particular, $N(0, 1)$ is called the **standard normal** distribution. $N(m, v)$ and $N(0, 1)$ is related as follows.

$$Y \approx N(0,1) \iff X \stackrel{\text{def}}{=} m + \sqrt{v}Y \approx N(m, v). \tag{0.2}$$

Remark: By setting $m = 0$ and $B = \mathbb{R}$ in (0.1),

$$\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2v}\right) dx = \sqrt{2\pi v}.$$

Then, formally plugging $v = \mathbf{i}/2$ in the above identity, we obtain **Fresnel integral**:

$$\int_{-\infty}^{\infty} \exp\left(\mathbf{i}x^2\right) dx = \sqrt{\pi \mathbf{i}} = \sqrt{\frac{\pi}{2}}(1 + \mathbf{i}),$$

i.e.,

$$\int_{-\infty}^{\infty} \cos(x^2) dx = \int_{-\infty}^{\infty} \sin(x^2) dx = \sqrt{\frac{\pi}{2}}.$$

Exercise Justify the above manupulation: $v = \mathbf{i}/2$.

Example 0.1.2 (**Poisson distribution**) Let $c > 0$.

► A r.v. $N : \Omega \to \mathbb{N}$ is called a c-**Poisson r.v.** if

$$P(N \in B) = \pi_c(B) \stackrel{\text{def.}}{=} \sum_{n \in B} \frac{e^{-c} c^n}{n!}, \quad B \subset \mathbb{N}. \tag{0.3}$$

A probability measure π_c defined above is called c-**Poisson distribution**. It is not hard to see that

$$EN = \text{var}\, N = c. \tag{0.4}$$

Let N_1 and N_2 be independent r.v.'s. $c_1, c_2 > 0$ and $c = c_1 + c_2$. We prove that

$$N_j \approx \pi_{c_j} \ (j = 1, 2) \implies N_1 + N_2 \approx \pi_c. \tag{0.5}$$

We stat by noting that

1) $$\frac{c^r}{r!} = \sum_{\substack{k,\ell \geq 0 \\ k+\ell = r}} \frac{c_1^k}{k!} \frac{c_2^\ell}{\ell!},$$

which can be seen as follows. For $t \in \mathbb{R}$,

$$\sum_{r \geq 0} t^r \frac{c^r}{r!} = e^{tc} = e^{tc_1} e^{tc_2} = \sum_{k \geq 0} t^k \frac{c_1^k}{k!} \sum_{\ell \geq 0} t^\ell \frac{c_2^\ell}{\ell!} = \sum_{n \geq 0} t^n \sum_{\substack{k,\ell \geq 0 \\ k+\ell = n}} \frac{c_1^k}{k!} \frac{c_2^\ell}{\ell!}.$$

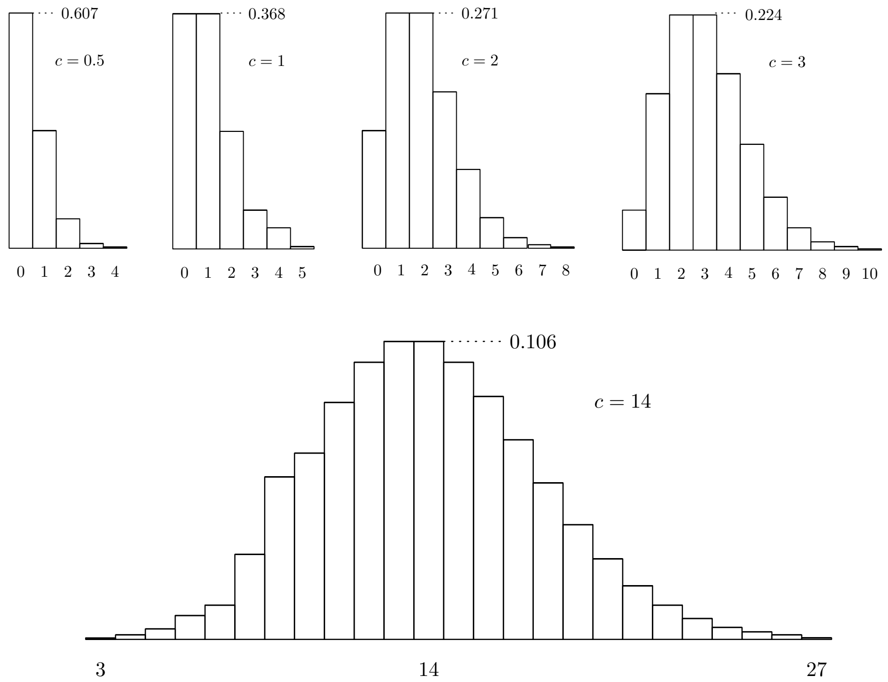
By comparing the coefficient of t^r, we get 1).

We now conclude (0.5) as follows:

$$\begin{aligned} P(N_1 + N_2 = r) &= \sum_{\substack{k,\ell \geq 0 \\ k+\ell = r}} P(N_1 = k, N_2 = \ell) = \sum_{\substack{k,\ell \geq 0 \\ k+\ell = r}} P(N_1 = k) P(N_2 = \ell) \\ &= \sum_{\substack{k,\ell \geq 0 \\ k+\ell = r}} \frac{e^{-c_1} c_1^k}{k!} \frac{e^{-c_2} c_2^\ell}{\ell!} = e^{-c} \sum_{\substack{k,\ell \geq 0 \\ k+\ell = r}} \frac{c_1^k}{k!} \frac{c_2^\ell}{\ell!} \stackrel{1)}{=} e^{-c} \frac{c^r}{r!}. \end{aligned}$$

\(^□^)/

Here are histograms of $\pi_c(n) \overset{\text{def}}{=} \frac{e^{-c}c^n}{n!}$ $(n \in \mathbb{N})$.



When c is large, the histogram looks like that of $N(c, c)$. This is a manifestation of the **central limit theorem**:

$$\frac{N_c - c}{\sqrt{c}} \xrightarrow{\text{w}} N(0, 1), \quad \text{as } c \to \infty.$$

Example 0.1.3 (**Binomial distribution**) Let $p \in [0,1]$ and $n = 1, 2, ...$ A probability measure $\mu_{n,p}$ on $\{0, 1, .., n\}$ defined as follows is called the (n, p)**-binomial distribution**, and will henceforth be denoted by $\mathrm{Bin}(n, p)$:

$$\mu_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, ..., n. \tag{0.6}$$

Note in particular that

$$\mu_{1,p}(k) = \begin{cases} p & \text{if } k = 1, \\ 1-p & \text{if } k = 0. \end{cases} \tag{0.7}$$

Let $\{X_j\}_{j=1}^n$ be i.i.d. with $X_j \approx \mathrm{Bin}(1, p)$ Then,

$$S_n \stackrel{\text{def}}{=} X_1 + \ldots + X_n \approx \mathrm{Bin}(n, p). \tag{0.8}$$

To prove this, note first that for $j = 1, ..., n$,

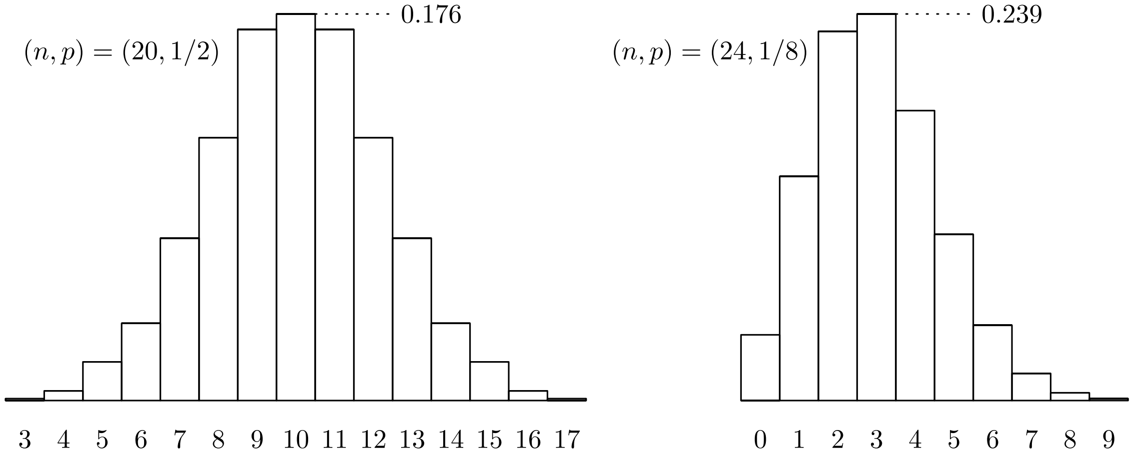
1) $$P(X_j = k) = \mu_{1,p}(k) = \begin{cases} p & \text{if } k = 1, \\ 1-p & \text{if } k = 0. \end{cases}$$

Therefore, we have for any $k = 0, 1, ..., n$ that

$$\begin{aligned} P(S_n = k) &= \sum_{\substack{k_1,...,k_n=0,1 \\ k_1+...+k_n=k}} P(X_1 = k_1, ..., X_n = k_n) \\ &= \sum_{\substack{k_1,...,k_n=0,1 \\ k_1+...+k_n=k}} \underbrace{P(X_1 = k_1) \cdots P(X_n = k_n)}_{\stackrel{1)}{=} p^k(1-p)^{n-k}} = \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

Question Let Z be a r.v. defined on a probability space $(\Omega, \mathcal{F}, P)$ such that $Z \approx \mathrm{Bin}(n, p)$. Is it always true that there exist iid $X_j \approx \mathrm{Bin}(1, p)$ $(j = 1, ..., n)$ defined on $(\Omega, \mathcal{F}, P)$ such that $Z = X_1 + ... + X_n$?

Here are histograms of $k \mapsto \mu_{n,p}(k)$ for $(n,p) = (20, 1/2)$ and $(n,p) = (24, 1/8)$.



The histogram on the left looks like that of the normal distribution, which can be explained by the **de Moivre-Laplace theorem**: Suppose that $n, k \to \infty$ and $\frac{k-np}{n^{2/3}} \to 0$. Then,

$$\mu_{n,p}(k) \sim \frac{1}{\sqrt{2\pi vn}} \exp\left(-\frac{(k-np)^2}{2vn}\right), \quad \text{where} \quad v = p(1-p). \tag{0.9}$$

On the other hand, the histogram on the right looks like that of Poisson distribution, which can be explained by law of small numbers: Suppose that $n \to \infty$, $p \to 0$, $np \to c > 0$. Then,

$$\binom{n}{k} p^k (1-p)^{n-k} \longrightarrow \frac{e^{-c}c^k}{k!}, \quad k \in \mathbb{N}. \tag{0.10}$$

\(^□^)/

Example 0.1.4 (**Gamma distributions**) Let $a, c > 0$.

▶ We define (c, a)**-gamma distribution** $\gamma_{c,a} \in \mathcal{P}((0,\infty))$ by

$$\gamma_{c,a}(B) = \frac{c^a}{\Gamma(a)} \int_B x^{a-1} e^{-cx} dx, \quad \text{for } B \in \mathcal{B}((0,\infty)). \tag{0.11}$$

Here, we have introdued the Gamma function as usual:

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \quad a \in \mathbb{C}, \ \mathrm{Re}(a) > 0. \tag{0.12}$$

$\gamma_{c,a}$ is also denoted by $\gamma(c, a)$. It is not difficult to see that

$$EX = a/c, \quad \mathrm{var}\, X = a/c^2. \tag{0.13}$$

0.2 The Law of Large Numbers

Theorem 0.2.1 (The Law of Large Numbers) *Let* $S_n = X_1 + ... + X_n$*, where* $\{X_n\}_{n\geq 1}$ *are i.i.d. with* $E|X_n| < \infty$*. Then,*

$$\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} EX_1, \quad P\text{-}a.s. \tag{0.14}$$

Example 0.2.2 (**Uniqueness of the Laplace transform**) Let $\mu_1, \mu_2 \in \mathcal{P}([0,\infty))$. Then $\mu_1 = \mu_2$ if

$$\int_{[0,\infty)} e^{-\lambda x} d\mu_1(x) = \int_{[0,\infty)} e^{-\lambda x} d\mu_2(x) \quad \text{for all } \lambda \geq 0. \tag{0.15}$$

Proof: Let $f \in C_{\mathrm{b}}([0,\infty) \to [0,\infty))$ be arbitrary. We first prove the following approximation:

1) $\lim\limits_{n \nearrow \infty} f_n(x) = f(x)$ for all $x \geq 0$,

where

$$f_n(x) = e^{-nx} \sum_{k=0}^{\infty} \frac{(nx)^k}{k!} f\left(\frac{k}{n}\right), \quad n \in \mathbb{N}.$$

To prove 1), we may assume $x > 0$, since $f_n(0) = f(0)$. For $x > 0$, we let

$$S_n = X_1 + \ldots + X_n$$

where X_n are iid, $\approx \pi_x$ (cf. (0.3)). Then,

2) $S_n \stackrel{(0.5)}{\approx} \pi_{nx}$.

Moreover, by the law of large numbers (Theorem 0.2.1),

$$S_n/n \stackrel{n\to\infty}{\longrightarrow} EX_1 \stackrel{(0.4)}{=} x, \text{ a.s.}$$

and hence by the bounded convergence theorem,

$$f_n(x) \stackrel{2)}{=} E[f(S_n/n)] \stackrel{n\to\infty}{\longrightarrow} f(x).$$

We now use 1) to prove that $\mu_1 = \mu_2$. It is enough to prove that

3) $\int_{[0,\infty)} f d\mu_1 = \int_{[0,\infty)} f d\mu_2$.

Indeed, by differentiating (0.15) k times at in λ and then setting $\lambda = n \in \mathbb{N}$, we have that

$$\int_{[0,\infty)} x^k e^{-nx} d\mu_1(x) = \int_{[0,\infty)} x^k e^{-nx} d\mu_2(x) \quad \text{for all } k, n \in \mathbb{N}.$$

By multiplying both hands-sides of the above identity by $\frac{n^k}{k!} f\left(\frac{k}{n}\right)$, and adding over $k \in \mathbb{N}$, we arrive at:

4) $\int_{[0,\infty)} f_n d\mu_1 = \int_{[0,\infty)} f_n d\mu_2$.

Since $\sup_{x\geq 0} |f_n(x)| \leq \sup_{x \geq 0} |f(x)|$, we obtain 3) from 2) and 4) via the bounded convergence theorem. \(^□^)/

0.3 Characteristic functions

For $\nu \in \mathcal{P}(\mathbb{R}^d)$, we define its **Fourier transform** by

$$\widehat{\nu}(\theta) \stackrel{\text{def}}{=} \int \exp(\mathbf{i}(\theta \cdot x)) d\nu(x), \quad \theta \in \mathbb{R}^d.$$

Proposition 0.3.1 (Characteristic function) *For $\nu \in \mathcal{P}(\mathbb{R}^d)$ and a r.v. $X : \Omega \to \mathbb{R}^d$, the following are equivalent:*

a) $E \exp(\mathbf{i}(\theta \cdot X)) = \widehat{\nu}(\theta)$ *for all* $\theta \in \mathbb{R}^d$*;*

b) $X \approx \nu$.

► *The expectation on the left-hand side of a) above is called the* **characteristic function** *(***ch.f.** *for short) of* X*.*

Example 0.3.2 (ch.f. of a Poisson r.v.) Let $\pi_c(n) = \frac{e^{-c}c^n}{n!}$, $n \in \mathbb{N}$, $c > 0$, cf. (0.3) and N be a r.v. $\approx \pi_c$, . We then see for any $z \in \mathbb{C}$ that

$$E[z^N] = e^{-c} \sum_{n \geq 0} z^n \frac{c^n}{n!} = \exp((z-1)c).$$

This shows (by setting $z = \exp(\mathbf{i}\theta)$) in particular that

$$\widehat{\pi}_c(\theta) = E \exp(\mathbf{i}\theta N) = \exp((e^{\mathbf{i}\theta} - 1)c). \tag{0.16}$$

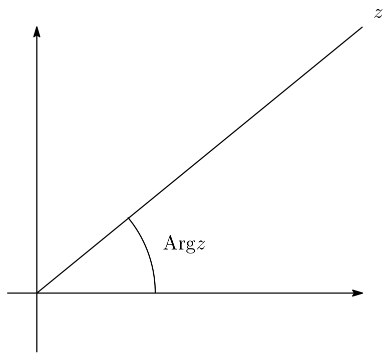
Example 0.3.3 ($\star$) **(ch.f. of a Gamma r.v.)** For $z \in \mathbb{C}\backslash\{0\}$, we define $\text{Arg}\, z \in (-\pi, \pi]$ (argument of z) by

$$z = |z| \exp(\mathbf{i}\text{Arg}\, z),$$

and $\text{Log} z \in \mathbb{C}$ by

$$\text{Log}\, z = \log|z| + \mathbf{i}\text{Arg}\, z.$$

By definition, Arg z is the angle, signed counter-clockwise, from the positive real axis to the vecor representing z.



Finally we set:

$$z^s = \exp(s\text{Log}\, z), \quad \text{for } z \in \mathbb{C}\backslash\{0\} \text{ and } s \in \mathbb{C}.$$

Let X be a real r.v. such that $X \approx \gamma_{c,a}$. We will show that

1) $$E\exp(-zX) = \left(1 + \frac{z}{c}\right)^{-a} \quad \text{for any } z \in \mathbb{C} \text{ with } \text{Re}\, z > -c.$$

Then, it follows from 1) that for $\theta \in \mathbb{R}$,

$$\begin{aligned} \widehat{\gamma_{c,a}}(\theta) &= \left(1 - \frac{\mathbf{i}\theta}{c}\right)^{-a} = \left|1 - \frac{\mathbf{i}\theta}{c}\right|^{-a} \exp\left(-a\mathbf{i}\text{Arg}\left(1 - \frac{\mathbf{i}\theta}{c}\right)\right) \\ &= \left(1 + \frac{\theta^2}{c^2}\right)^{-a/2} \exp\left(\mathbf{i}a\text{Arctan}\,\frac{\theta}{c}\right). \end{aligned} \tag{0.17}$$

To prove 1), note first that both hand-sides are holomorphic in z for $\text{Re}\, z > -c$. Therefore, it is enough to prove it for all $z = t \in (-c, \infty)$. Then,

$$\begin{aligned} E\exp(-tX) &\overset{(0.11)}{=} \frac{c^a}{\Gamma(a)} \int_0^\infty x^{a-1} e^{-(t+c)x} dx \\ &\overset{x=y/(t+c)}{=} \frac{c^a}{\Gamma(a)} \left(\frac{1}{t+c}\right)^a \underbrace{\int_0^\infty y^{a-1} e^{-y} dy}_{=\Gamma(a)} = \left(1 + \frac{t}{c}\right)^{-a}. \end{aligned}$$

This proves 1).

Example 0.3.4 ($\star$) (**Stieltjes' counterexample to the moment problem**) We consider the following question. Suppose that a function $f \in C([0,\infty))$ satisfies

$$\int_0^\infty x^n |f(x)| dx < \infty, \quad \text{and} \quad \int_0^\infty x^n f(x) dx = 0 \quad \text{for all } n \in \mathbb{N}.$$

Then $f \equiv 0$?

Stieltjes gave a counterexample $f(x) \stackrel{\text{def}}{=} \exp(-x^{1/4}) \sin x^{1/4}$ to this question (1894). We can use (0.17) to verify that the above function is indeed a counterexample. In fact, we see from (0.17) that $\widehat{\gamma_{1,4n+4}}(1) \in \mathbb{R}$ for all $n \in \mathbb{N}$. Thus, taking the imaginary part, we have

$$0 = \int_0^\infty x^{4n+3} e^{-x} \sin x dx = \frac{1}{4} \int_0^\infty x^n \exp(-x^{1/4}) \sin x^{1/4} dx.$$

Example 0.3.5 ($\star$) (**Euler's complementary formula for the Gamma function**) We will use (0.17) to prove the following identity due to Euler:

$$\frac{1}{\Gamma(1+a)\Gamma(1-a)} = \frac{\sin(\pi a)}{\pi a}, \quad a \in (0,1). \tag{0.18}$$

Let $f_a(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x} \mathbf{1}_{x>0}$ (the density of $\gamma(1,a)$). We have by the Plancherel formula that:

1) $$\int_0^\infty f_{1+a}(x) f_{1-a}(x)dx = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f_{1+a}}(\theta)\widehat{f_{1-a}}(-\theta)d\theta.$$

Since

$$f_{1+a}(x) f_{1-a}(x) = \frac{1}{\Gamma(1+a)\Gamma(1-a)} e^{-2x}\mathbf{1}_{x>0},$$

we see that

2) $$\int_0^\infty f_{1+a}(x) f_{1-a}(x)dx = \frac{1}{2\Gamma(1+a)\Gamma(1-a)}$$

On the other hand,

$$\begin{aligned} \widehat{f_{1+a}}(\theta)\widehat{f_{1-a}}(-\theta) &\overset{(0.17)}{=} \frac{1}{1+\theta^2}\exp(\mathbf{i}(1+a)\text{Arctan}\,\theta - \mathbf{i}(1-a)\text{Arctan}\,\theta) \\ &= (\text{Arctan}\,\theta)' \exp(2\mathbf{i}a\text{Arctan}\,\theta). \end{aligned}$$

Thus,

3) $$\left\{ \begin{aligned} \int_{\mathbb{R}} \widehat{f_{1+a}}(\theta)\widehat{f_{1-a}}(-\theta)d\theta &\overset{t=\text{Arctan}\,\theta}{=} \int_{-\pi/2}^{\pi/2} \exp(2\mathbf{i}at)dt \\ &= \frac{\exp(\mathbf{i}a\pi) - \exp(-\mathbf{i}\pi a)}{2\mathbf{i}a} = \frac{\sin(\pi a)}{a} \end{aligned} \right.$$

By 1)–3), we obtain (0.18).

0.4 Weak Convergence

Proposition 0.4.1 (Weak convergence of r.v.'s) *For $n = 0, 1, ...,$ let X_n be $\mathbb{R}^d$-valued r.v.'s and that $X_n \approx \mu_n \in \mathcal{P}(\mathbb{R}^d)$. Then, the following are equivalent:*

a) $E \exp(\mathbf{i}\theta \cdot X_n) \longrightarrow E \exp(\mathbf{i}\theta \cdot X_0)$ *for all* $\theta \in \mathbb{R}^d$.

b) $\mu_n \stackrel{\mathrm{w}}{\longrightarrow} \mu_0$.

► *The sequence* $(X_n)_{n \geq 0}$ *is said to* **converge weakly** *(or* **converge in law** *) to* X_0 *if one (therefore all) of the above conditions is satisfied. We will henceforth denote this convergence by*

$$X_n \stackrel{\mathrm{w}}{\longrightarrow} X_0 \quad \text{or} \quad X_n \stackrel{\mathrm{w}}{\longrightarrow} \mu_0$$

Example 0.4.2 Let $(N_c)_{c>0}$ be r.v.'s such that $\pi_c(k) \stackrel{\rm def}{=} P(N_c = k) = e^{-c}c^k/k!$ for all $k \in \mathbb{N}$ and $c > 0$. We will prove the following two facts, of which the first is probabilistic, the second purely analytic:

a) $\dfrac{N_c - c}{\sqrt{c}} \stackrel{\rm w}{\longrightarrow} N(0,1)$, as $c \to \infty$ (**Central limit theorem**).

b) $n! \stackrel{n\to\infty}{\sim} \sqrt{2\pi n}\,(n/e)^n$ (**Stirling's formula**).

Proof: a) Note that

$$\exp(\mathbf{i}\theta) = 1 + \mathbf{i}\theta - \frac{\theta^2}{2} + O(|\theta|^3) \text{ as } \theta \to 0,$$

and hence that

1) $$\exp\left(\mathbf{i}\frac{\theta}{\sqrt{c}}\right) = 1 + \frac{\mathbf{i}\theta}{\sqrt{c}} - \frac{\theta^2}{2c} + O\left(\frac{|\theta|^3}{c^{3/2}}\right) \text{ as } c \to \infty \text{ for any } \theta \in \mathbb{R}.$$

Since $\widehat{\pi_c}(\theta) \stackrel{(0.16)}{=} \exp(c(\exp(\mathbf{i}\theta) - 1))$, we have

2) $$\begin{cases} E\exp\left(\mathbf{i}\theta\dfrac{N_c - c}{\sqrt{c}}\right) &= \widehat{\pi_c}\left(\frac{\theta}{\sqrt{c}}\right)\exp\left(-\mathbf{i}\sqrt{c}\theta\right) \\ &= \exp\left(c\left(\exp\left(\mathbf{i}\dfrac{\theta}{\sqrt{c}}\right) - 1 - \mathbf{i}\dfrac{\theta}{\sqrt{c}}\right)\right) \\ &\stackrel{1)}{=} \exp\left(c\left(-\dfrac{\theta^2}{2c} + O\left(\dfrac{\theta^3}{c^{3/2}}\right)\right)\right) \stackrel{c\to\infty}{\longrightarrow} \exp\left(-\dfrac{\theta^2}{2}\right). \end{cases}$$

Recall that $\exp\left(-\frac{\theta^2}{2}\right)$ is the Fourier transform of $N(0,1)$. We see the desired weak convergence from 2) and Proposition 0.4.1.

b) We have that

$$\widehat{\pi_c}(\theta) = \sum_{k\geq 0}\exp(\mathbf{i}k\theta)\pi_c(k), \quad \theta \in \mathbb{R}$$

Multiplying $\exp(-\mathbf{i}n\theta)/(2\pi)$ to the both hands sides of the above identity and integrating them over $\theta \in [-\pi, \pi]$, we obtain

3) $$\pi_c(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\widehat{\pi_c}(\theta)\exp(-\mathbf{i}n\theta)d\theta.$$

Moreover, since $1 - \cos\theta \geq \frac{2\theta^2}{\pi^2}$, $|\theta| \leq \pi$, we have

4) $$\left|\widehat{\pi_c}\left(\tfrac{\theta}{\sqrt{c}}\right)\right| = \exp\left(-c\left(1 - \cos\tfrac{\theta}{\sqrt{c}}\right)\right) \leq \exp\left(-\tfrac{2\theta^2}{\pi^2}\right), \quad |\theta| \leq \pi\sqrt{c}.$$

Finally, note that

5) $$\begin{cases} \dfrac{\sqrt{n}}{n!}\left(\dfrac{n}{e}\right)^n &= \sqrt{n}\pi_n(n) \stackrel{3)}{=} \dfrac{\sqrt{n}}{2\pi}\displaystyle\int_{-\pi}^{\pi}\widehat{\pi_n}(\theta)\exp(-\mathbf{i}n\theta)d\theta \\ &= \dfrac{1}{2\pi}\displaystyle\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\widehat{\pi_n}\left(\dfrac{\theta}{\sqrt{n}}\right)\exp(-\mathbf{i}\sqrt{n}\theta)d\theta \end{cases}$$

By 2), 4) and the dominated convergence theorem, we conclude that

$$\text{the RHS 5)} \stackrel{n\to\infty}{\longrightarrow} \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-\tfrac{\theta^2}{2}\right)d\theta = \frac{1}{\sqrt{2\pi}}.$$

This proves b). \(^□^)/

0.5 Martingales

We suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, and that $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$.

Proposition 0.5.1 (**Conditional expectation**) *Let* $X \in L^1(P)$.

a) *There exists a unique* $Y \in L^1(\Omega, \mathcal{G}, P|_{\mathcal{G}})$ *such that*

$$E[X : A] = E[Y : A] \text{ for all } A \in \mathcal{G}. \tag{0.19}$$

The r.v. Y *is called the* **conditional expectation** *of* X *given* $\mathcal{G}$*, and is denoted by* $E[X|\mathcal{G}]$.

b) *For* $X, X_n \in L^1(P)$ $(n \in \mathbb{N})$,

$$E[\alpha X_1 + \beta X_2|\mathcal{G}] = \alpha E[X_1|\mathcal{G}] + \beta E[X_2|\mathcal{G}], \quad \text{a.s. for } \alpha, \beta \in \mathbb{R}, \tag{0.20}$$

$$X_1 \leq X_2, \quad \text{a.s.} \implies E[X_1|\mathcal{G}] \leq E[X_2|\mathcal{G}], \quad \text{a.s.}, \tag{0.21}$$

$$|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}], \quad \text{a.s.}, \tag{0.22}$$

$$X \text{ is } \mathcal{G}\text{-measurable} \iff E[X|\mathcal{G}] = X, \quad \text{a.s.} \tag{0.23}$$

$$E[X : A] = EXP(A), \forall A \in \mathcal{G} \iff E[X|\mathcal{G}] = EX, \quad \text{a.s.} \tag{0.24}$$

$$X \text{ is independent of } \mathcal{G} \implies E[X|\mathcal{G}] = EX, \quad \text{a.s.} \tag{0.25}$$

$$X_n \stackrel{n\to\infty}{\longrightarrow} X \text{ in } L^1(P) \iff E[|X_n - X||\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} 0 \text{ in } L^1(P). \tag{0.26}$$

We assume that

- $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathbb{T} \subset \mathbb{R}$;
- $(\mathcal{F}_t)_{t\in\mathbb{T}}$ is a filtration;
- $X = (X_t)_{t\in\mathbb{T}}$ is a sequence of real r.v.'s defined on $(\Omega, \mathcal{F}, P)$.

Definition 0.5.2 $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is called a **martingale** if the following hold true.

- (**adapted**) X_t is $\mathcal{F}_t$-measurable for all $t \in \mathbb{T}$;
- (**integrable**) $X_t \in L^1(P)$ for all $t \in \mathbb{T}$;
- (**martingale property**)

$$E[X_t|\mathcal{F}_s] = X_s \text{ a.s. if } s, t \in \mathbb{T} \text{ and } s < t. \tag{0.27}$$

If the equality in (0.27) is replaced by $\geq$ (resp. $\leq$), X is called a **submartingale** (resp. **supermartingale**).

Example 0.5.3 Let $\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \in \mathbb{T})$, Q be a signed measure on $(\Omega, \mathcal{F}_\infty)$, and $P_t = P|_{\mathcal{F}_t}$, $Q_t = Q|_{\mathcal{F}_t}$. Suppose that $Q_t \ll P_t$ for all $t \in \mathbb{T}$. Then, $X_t \stackrel{\text{def}}{=} \frac{dQ_t}{dP_t}$, $t \in \mathbb{T}$ is a martingale.
Proof: X_t is $\mathcal{F}_t$-measurable and $X_t \in L^1(P)$. Let $s, t \in \mathbb{T}$, $s < t$ and $A \in \mathcal{F}_s$. Then, since $A \in \mathcal{F}_t$,

$$E[X_t : A] = Q_t(A) = Q(A) = Q_s(A) = E[X_s : A].$$

Thus, $E[X_t|\mathcal{F}_s] = X_s$, a.s. \(^□^)/

Now, a naive qustion arises.

Question 1 Is an arbitrary martingale X_t expressed as $X_t = dQ_t/dP_t$ by a sined measure Q as in Example 0.5.3?

But the answer is clearly negative. Indeed, if $X_t = dQ_t/dP_t$ for a sined measure Q, then

$$\sup_{t \geq 0} E|X_t| = \sup_{t \geq 0} |Q_t| \leq |Q|, \tag{0.28}$$

where $|Q_t|$ and $|Q|$ above are total variations. Therefore, the martingale X_t should be at least L^1-bounded. We now arrive at a less obvious question:

Question 2 Is an arbitrary L^1-bounded martingale X_t expressed as $X_t = dQ_t/dP_t$ by a sined measure Q as in Example 0.5.3?

I am grateful to Francis Comets for bringing the following lemma to my interest.

Lemma 0.5.4 *Suppose that the set* $\mathbb{T} \subset \mathbb{R}$ *is unbounded from above and that* $X = (X_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ *is a submartingale such that* $\sup_{t \in \mathbb{T}} E[X_t^+] < \infty$.

a) *There exists a martingale* $Y = (Y_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ *such that* $X_t^+ \leq Y_t$ *for all* $t \in \mathbb{T}$.

b) (**Krickeberg decomposition**) *There exists a nonnegative supermartingale* $Z = (Z_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ *such that* $X_t = Y_t - Z_t$ *for all* $t \in \mathbb{T}$. *In particular,* Z *is a martingale if* X *is a martingale.*

Proof: a) We start by observing that

1) $t, u, v \in \mathbb{T},\ t \leq u < v \implies E[X_u^+ | \mathcal{F}_t] \leq E[X_v^+ | \mathcal{F}_t]$, a.s.

Indeed, $(X_t^+, \mathcal{F}_t)_{t \in \mathbb{T}}$ is a submartingale. Thus,

$$X_u^+ \leq E[X_v^+ | \mathcal{F}_u], \text{ a.s.}$$

We obtain 1) by taking the conditional expextations of the both hands sides of the above identity.
By 1), the limit $Y_t \stackrel{\text{def}}{=} \lim_{u \to \infty} E[X_u^+ | \mathcal{F}_t] \in [0, \infty]$ exists and $X_t^+ \leq Y_t$ for all $t \in \mathbb{T}$. We verify that

2) $Y = (Y_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ is a martingale.

First, $Y_t \in L^1(P)$ for all $t \in \mathbb{T}$, since by 1) and the monotone convergence theorem,

$$EY_t = \lim_{u \to \infty} E[E[X_u^+ | \mathcal{F}_t]] = \lim_{u \to \infty} E[X_u^+] < \infty.$$

Next, if $s, t \in \mathbb{T}$ and $s < t$, then, by the monotone convergence theorem for conditional expectations,

$$E[Y_t | \mathcal{F}_s] = \lim_{u \to \infty} E[E[X_u^+ | \mathcal{F}_t] | \mathcal{F}_s] = \lim_{u \to \infty} E[X_u^+ | \mathcal{F}_s] = Y_s, \text{ a.s.}$$

b) $Z_t \stackrel{\text{def}}{=} Y_t - X_t$, $t \in \mathbb{T}$ is a nonnegative supermartingale. In particular, Z is a martingale if X is a martingale. \(^□^)/

Let $X = (X_t)_{t \in \mathbb{T}}$ be a process. We write $\mathcal{F}_t^X = \sigma(X_s \ ; \ s \in \mathbb{T} \cap [0,t])$ $t \in \mathbb{T}$, and $\mathcal{F}_\infty^X = \sigma(\mathcal{F}_t^X \ ; \ t \in \mathbb{T})$. For a signed measure Q on $(\Omega, \mathcal{F}_\infty^X)$, let $|Q|$ be its variation, $Q^\pm = (|Q| \pm Q)/2$ (Jordan decomposition) and $Q_t = Q|_{\mathcal{F}_t^X}$.

Lemma 0.5.5 *Let $Y = (Y_t, \mathcal{F}_t^X)_{t \in \mathbb{T}}$ be a nonnegative, mean-one martingale. Then, there exists a unique probability measure P^Y on $(\Omega, \mathcal{F}_\infty^X)$ such that*

$$P^Y(A) = E[Y_t : A] \text{ for all } t \in \mathbb{T} \text{ and } A \in \mathcal{F}_t^X.$$

Proof: For each $t \in \mathbb{T}$, let $\widetilde{P}_t(A) = E[Y_t : A]$ for $A \in \mathcal{F}_t^X$. Then, the family of measures $(\mathcal{F}_t^X, \widetilde{P}_t)$, $t \in \mathbb{T}$ are consistent in the sense that $\widetilde{P}_t|_{\mathcal{F}_s^X} = \widetilde{P}_s$ if $s, t \in \mathbb{T}$, $s < t$. Thus, by Kolmogorov's extension theorem, there exists a unique probability measure P^Y on $(\Omega, \mathcal{F}_\infty^X)$ such that $P^Y|_{\mathcal{F}_t^X} = \widetilde{P}_t$ for all $t \in \mathbb{T}$. \(^□^)/

Proposition 0.5.6 *Suppose that the set* $\mathbb{T}$ *is ubbounded from above, and that* $X = (X_t, \mathcal{F}_t^X)_{t\in\mathbb{T}}$ *is a martingale. Then, the following conditions are equivalent.*

a) X *is a difference of two nonnegative* $(\mathcal{F}_t^X)$*-martingales.*

b1) *There exists a signed measure* Q *on* $(\Omega, \mathcal{F}_\infty^X)$ *such that for all* $t \in \mathbb{T}$, $|Q|_t \ll P_t$ *and* $dQ_t/dP_t = X_t$.

b2) *There exists a signed measure* Q *on* $(\Omega, \mathcal{F}_\infty^X)$ *such that for all* $t \in \mathbb{T}$, $Q_t \ll P$ *and* $dQ_t/dP_t = X_t$.

c) $\sup_{t\in\mathbb{T}} E|X_t| < \infty$.

Proof: of Proposition 0.5.6: a) $\Rightarrow$ b1): Suppose that X is a difference of two nonnegative $(\mathcal{F}_t^X)$-martingales Y_t and Z_t. Then, by Lemma 0.5.5, there exist finite measures Q^Y, Q^Z on $(\Omega, \mathcal{F}_\infty^X)$ such that for all $t \in \mathbb{T}$, $Q_t^Y \ll P_t$, $Q_t^Z \ll P_t$, $Y_t = dQ_t^Y/dP_t$, $Z_t = dQ_t^Z/dP_t$. Set $Q = Q^Y - Q^Z$. Then, $|Q| \le Q^Y + Q^Z$ and hence $|Q|_t \le (Q^Y + Q^Z)_t \ll P_t$. Moreover,

$$dQ_t/dP_t = d(Q_t^Y - dQ_t^Z)/dP_t = dQ_t^Y/dP_t - dQ_t^Z/dP_t = Y_t - Z_t = X_t.$$

b1) $\Rightarrow$ b2): This follows from the inequality $|Q_t| \le |Q|_t$.
b2) $\Rightarrow$ c): $E|X_t| = |Q_t|(\Omega) \le |Q|(\Omega) < \infty$.
c) $\Rightarrow$ a): This follows from Lemma 0.5.4. \(^□^)/

0.6 Brownian Motion

The Brownian motion came into the history in 1827, when R. Brown, a British botanist, observed that pollen grains suspended in water perform a contunual swarming motion. In 1905, A. Einstein derived (0.30) below from the moleculer physics point of view. A mathematically rigorous construction with a proof of the continuity (cf. B3) below) was given by N. Wiener (1923).

We fix a probability space $(\Omega, \mathcal{F}, P)$ in this subsection. In the sequel, we will repeatedly refer to a finite time series of the form

$$0 = t_0 < t_1 < ... < t_n, \quad n \geq 1. \tag{0.29}$$

Definition 0.6.1 (**Brownian motion**) Let $B = (B_t : \Omega \to \mathbb{R}^d)_{t \geq 0}$ be a family r.v.'s. We consider the following conditions.

B1) For any time series (0.29), the following r.v.'s are independent.

$$B(0), B(t_1) - B(0), \ldots, B(t_n) - B(t_{n-1}).$$

B2) For any $0 \leq s < t$,

$$B_t - B_s \approx N(0, (t-s)I_d), \tag{0.30}$$

where I_d is the identity matrix of degree d,

B3) There is an $\Omega_B \in \mathcal{F}$ such that $P(\Omega_B) = 1$ and $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega_B$.

B4) $B_0 = x$, for a nonrandom vector $x \in \mathbb{R}^d$,

► B is called a d-dimensional **Brownian motion** (BM^d for short) if the conditions B1)–B3) are satisfied.

► B is called a d-dimensional **Brownian motion** started at x (BM^d_x for short), if the conditions B1)–B4) are satisfied.

► B is called a d-dimensional **pre-Brownian motion** (pre-BM^d for short), if the conditions B1), B2) are satisfied. A d-dimensional **pre-Brownian motion** is said to be started at x, if it saitesfies B4) and is abbreviated by pre-BM^d_x.

0.7 Continuity of the Brownian Motion

Referring to Definition 0.6.1, given the distribution of B_0, the distribution of $B = (B_t)_{t\geq 0}$ is determined by properties B1) and B2). Then,

Question 1 Do all pre-Brownian motions have continuous path?

Example 0.7.1 Let B be BM^1_0, and U be a r.v. uniformly distributed on $(0, 1)$, which is independent of B. Now, define $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ by

$$\widetilde{B}_t = \begin{cases} B_t, & \text{if } t \neq U, \\ 0, & \text{if } t = U. \end{cases}$$

Since $P(\mathrm{t} = \mathrm{U}) = 0$ for any fixed $t \geq 0$, B and $\widetilde{B}$ have the same law, and hence the latter is a pre-BM^1_0. However, $\widetilde{B}$ is discontinuous a.s.

In fact, Example 0.7.1 does more job than to construct a discontinuous pre Brownian motion. The following remark is due to Kouji Yano:

Proposition 0.7.2 *The following "event" is not $\sigma(B)$-measurable:*

$$C \stackrel{\text{def}}{=} \{B_t \textit{ is continuous in } t \geq 0\}.$$

Proof (sketch): The map $B \mapsto \widetilde{B}$ preserves the law of the Brownian motion. Thus, if C is $\sigma(B)$-measurable, then, it should be the case that $P(B \in C) = P(\widetilde{B} \in C)$, a contradiction $(1 = 0)$! \(^□^)/

0.8 Germ triviality

Let B be a BM^d. We define the **right-continuous enlargement** $(\mathcal{F}_t)_{t\geq 0}$ of the canonical filtration $(\mathcal{F}_t^0)_{t\geq 0}$ as follows;

$$\mathcal{F}_t^0 = \sigma(B_s\ ;\ s \leq t), \quad \text{and} \quad \mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}^0. \tag{0.31}$$

In particular, $\mathcal{F}_0$ is called the **germ σ-algebra**. The technical advantage of introducing $\mathcal{F}_t$ ("an infinitesimal peeking in the future") is to enlarge $\mathcal{F}_t^0$ to get the right-continuity:

$$\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t, \quad \forall t \geq 0. \tag{0.32}$$

Indeed,

$$\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \bigcap_{\varepsilon>0}\bigcap_{\delta>0} \mathcal{F}_{t+\varepsilon+\delta}^0 = \bigcap_{\varepsilon,\delta>0} \mathcal{F}_{t+\varepsilon+\delta}^0 = \mathcal{F}_t.$$

Note that $\mathcal{F}_t$ is strictly larger than $\mathcal{F}_t^0$. For example, the r.v. $X = \overline{\lim_{n\to\infty}}\, B^1(t+\frac{1}{n})$ is $\mathcal{F}_t$-measurable, but not $\mathcal{F}_t^0$-measurable.

The following fact is well-known.

Proposition 0.8.1 (**Germ triviaility/Blumenthal zero-one law**) *For* BM$_x^d$ *for some* $x \in \mathbb{R}^d$,

$$A \in \mathcal{F}_0 \implies P(A) \in \{0, 1\}.$$

Question 1 How much larger is $\mathcal{F}_t$ than $\mathcal{F}_t^0$?

Question 2 Can germ triviality be explained from a general property for $\mathcal{F}_t$ $(t \geq 0)$?

Question 3 Does germ triviality remain true for pre-Brownian motions?

Proposition 0.8.2 (**Markov property**) *Let* $s \geq 0$ *and* $G \in \mathcal{T}_s \stackrel{\text{def}}{=} \sigma(B_t\ ;\ t \geq s)$. *Then,*

$$P(G|\mathcal{F}_s) = P(G|B_s), \quad a.s. \tag{0.33}$$

Proposition 0.8.2 can be used to show that the right-continuous enlargement of $\mathcal{F}_t$ is larger than $\mathcal{F}_t^0$ by null sets:

Proposition 0.8.3 *Let B be a* BM^d*, $t \geq 0$. Then,*

a)

$$\mathcal{F}_t = \mathcal{F}_t^0 \vee \sigma(\mathcal{N}_t), \tag{0.34}$$

where $\mathcal{N}_t$ denotes the totality of $\mathcal{F}_t$-measurable null sets.

b) *In particular, if B is a* BM^d_x *for some $x \in \mathbb{R}^d$, then, $\mathcal{F}_0 = \sigma(\mathcal{N}_0)$ and hence $P(A) \in \{0,1\}$ for $A \in \mathcal{F}_0$* (**germ triviality**)*.*

Proof: a) It is clear that $\mathcal{F}_t \supset \mathcal{F}_t^0 \vee \sigma(\mathcal{N}_t)$. We will show the opposite inclusion. Let

$$G \in \mathcal{G}_t \stackrel{\text{def}}{=} \bigcap_{\varepsilon>0} \sigma(B_{t+s}\ ;\ 0 \leq s \leq \varepsilon).$$

Since $\mathcal{G}_t \subset \mathcal{F}_t \cap \mathcal{T}_t$, we see from (0.33) that

$$\mathbf{1}_G = P(G|\mathcal{F}_t) \stackrel{(0.33)}{=} P(G|B_t), \ \text{ a.s.}$$

Thus, $\mathbf{1}_G$ is a.s. equals to an $\sigma(B_t)$-measurable function. This implies that

$$\mathcal{G}_t \subset \sigma(B_t) \vee \sigma(\mathcal{N}_t).$$

Hence

$$\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{G}_t \subset \mathcal{F}_t^0 \vee \sigma(\mathcal{N}_t).$$

b) Suppose in particular that B is a BM^d_x for some $x \in \mathbb{R}^d$. Then $\mathcal{F}_0^0 = \{\emptyset, \Omega\}$, and hence $\mathcal{F}_0 = \sigma(\mathcal{N}_0)$, which consists only of events A with $P(A) \in \{0,1\}$. \(^□^)/

Remark:
The germ triviality is not true in gereral for pre-Brownian motions. In fact, let B be BM^1_0, and U be a r.v. uniformly distributed on $(0,1)$, which is independent of B. Now, define $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ by

$$\widetilde{B}_t = \begin{cases} B_t & \text{if } t \neq U/n \text{ for any } n \in \mathbb{N}, \\ U & \text{if } t = U/n \text{ for some } n \in \mathbb{N}. \end{cases}$$

Since $P(t = U/n \text{ for some } n \in \mathbb{N}) = 0$ for any fixed $t \geq 0$, B and $\widetilde{B}$ have the same law, and hence the latter is a pre-BM^1_0. However, the germ σ-algebra of $\widetilde{B}$ contains $\sigma(U)$.

References

[Bil95] Billingsley, P: “Probability and Measure”, 3rd Ed. Wiley Series in Probability and Statistics, 1995.

[Chu74] Chung, K.-L. :“A Course in Probability Theory”, Harcourt Brace & World

[Dud89] Dudley, R.: “Real Analysis and Probability”, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Calfornia.

[Dur84] Durrett, R. :“Brownian Motion and Martingales in Analysis”, Wadsworth & Brooks/Cole Advanced Books & Software, Belmont, Calfornia.

[Dur95] Durrett, R. :“Probability–Theory and Examples”, 2nd Ed., Duxbury Press, 1995.

[Fol76] Folland, G. B.: Introduction to partial differential equations, Princeton University Press, 1976.

[IkWa89] Ikeda, N. and Watanabe, S. : Stochastic Differential Equations and Diffusion Processes (2nd ed.), North-Holland, Amsterdam / Kodansha,Tokyo (1989).

[KS91] Karatzas, I. and Shreve, S. E.: Brownian Motion and Stochastic Calculus, Second Edition. Springer Verlag (1991).

[Kre89] Kreyszig, E. : Introductory Functional Analysis with Applications (Wiley Classics Library)

[Law91] Lawler, G. F.: Intersections of Random Walks: Birkhäuser.

[Leb72] Lebedev, N. N.: Special Functions & Their Applicatinns, Dover, 1972.

[LeG16] Le Gall, J.-F.: Brownian Motion, Martingales, and Stochastic Calculus, Springer Verlag (2016).

[MP10] Möters, P., Peres, Y. “Brownian Motion” Cambridge University Press (2010).

[RS80] Reed, M. and Simon, B. “Method of Modern Mathematical Physics II” Academic Press 1980.

[Rud87] Rudin, W.: “Real and Complex Analysis—3rd ed.” McGraw-Hill Book Company, 1987.

[Spi76] Spitzer, F.: “Principles of Random Walks”, Springer Verlag, New York, Heiderberg, Berlin (1976).

[Yos] Yoshida, N.: "A Course in Probability", http://www.math.nagoya-u.ac.jp/~noby/pdf/prob.pdf