Note 6

Martingale CLT and backwards martingales.

see also A. Bovier's script for SS17, Chapter 2 [pdf].

1 Martingale CLT

Theorem 1. *Let* $(M_n)_{n\geqslant 0}$ *be a martingale with* $M_0 = 0$ *and let*

$$\sigma_n^2 = \sum_{i=1}^n \mathbb{E}[(\Delta M_i)^2] = \mathbb{E}\{[M]_n\} = \mathbb{E}[\langle M\rangle_n].$$

Assume moreover that, as $n\to\infty$,

$$\sigma_n^{-2} \max_{0\leqslant k\leqslant n} \mathbb{E}[(\Delta M_k)^2] \downarrow 0; \tag{1}$$

$$\sigma_n^{-2}\sum_{i=1}^n \mathbb{E}[(\Delta M_i)^2 \mathbb{1}_{|\Delta M_i|>\varepsilon\sigma_n} | \mathcal{F}_{i-1}] \downarrow 0, \qquad \text{for all } \varepsilon>0; \tag{2}$$

$$\sigma_n^{-2}\langle M\rangle_n \to 1, \qquad \text{in probability}. \tag{3}$$

Then

$$\sigma_n^{-1} M_n \xrightarrow{d} \mathcal{N}(0,1).$$

($\xrightarrow{d}$ denotes convergence in distribution)

Proof. For convenience let $\tilde{M}_k = \tilde{M}_k^n = \sigma_n^{-1} M_k$ for $n\geqslant 0$ and $k = 0,\ldots,n$. For ease of notation we will leave n implicit in the estimates below and just write $(\tilde{M}_k)_{k=0,\ldots,n}$. Is it however important to remark that the quantities denoted with tilda depends implicitly on n.

To prove convergence in distribution to a normal r.v. it is enough to prove that for all $u\in\mathbb{R}$

$$\mathbb{E}[e^{iu\tilde{M}_n}] \to e^{-u^2/2}, \qquad \text{as } n\to\infty.$$

Let us first assume that

$$\langle\tilde{M}\rangle_n \leqslant C \tag{4}$$

for some finite constant C. We will remove subsequently this limitation. Now let us decompose

$$\begin{aligned} |\mathbb{E}[e^{iu\tilde{M}_n} - e^{-u^2/2}]| &= |\mathbb{E}[e^{iu\tilde{M}_n}(1-e^{-u^2(1-\langle\tilde{M}\rangle_n)/2})] + \mathbb{E}[e^{-u^2/2}(e^{iu\tilde{M}_n+u^2\langle\tilde{M}\rangle_n/2}-1)]| \\ &\leqslant \mathbb{E}[|1-e^{-u^2(1-\langle\tilde{M}\rangle_n)/2}|] + |\mathbb{E}[e^{iu\tilde{M}_n+u^2\langle\tilde{M}\rangle_n/2}-1]| \\ &\leqslant \mathbb{E}[|1-e^{-u^2(1-\langle\tilde{M}\rangle_n)/2}|] + \sum_{k=1}^n \underbrace{|\mathbb{E}[e^{iu\tilde{M}_k+u^2\langle\tilde{M}\rangle_k/2} - e^{iu\tilde{M}_{k-1}+u^2\langle\tilde{M}\rangle_{k-1}/2}]|}_{=:D_k}. \end{aligned}$$

By assumption (3) we have $\langle\tilde{M}\rangle_n \to 1$ in probability and therefore $\mathbb{E}\big[\big|1-e^{-u^2(1-\langle\tilde{M}\rangle_n)/2}\big|\big]\to 0$, using that $x\mapsto|1-e^{-u^2x/2}|$ is bounded and continuous in 0. So it will be enough to prove that $\sum_k D_k\to 0$. Now, we have

$$\begin{aligned} D_k &= \big|\mathbb{E}\big[e^{iu\tilde{M}_{k-1}+u^2\langle\tilde{M}\rangle_{k-1}/2}\mathbb{E}\big[\big(e^{iu\Delta\tilde{M}_k+u^2\theta_k^2/2}-1\big)|\mathscr{F}_{k-1}\big]\big]\big|\leqslant \\ &\leqslant e^{u^2C/2}\mathbb{E}\big[\big|\mathbb{E}\big[\big(e^{iu\Delta\tilde{M}_k+u^2\theta_k^2/2}-1\big)|\mathscr{F}_{k-1}\big]\big|\big] \end{aligned}$$

where we used (4) to bound $\big|e^{iu\tilde{M}_{k-1}+u^2\langle\tilde{M}\rangle_{k-1}/2}\big|\leqslant e^{u^2C/2}$ and the notation

$$\theta_k^2 := \langle\tilde{M}\rangle_k-\langle\tilde{M}\rangle_{k-1}=\mathbb{E}\big[(\Delta\tilde{M}_k)^2|\mathscr{F}_{k-1}\big].$$

Using the elementary Taylor remaind estimates

$$e^{i\zeta}=1+i\zeta-\frac{1}{2}\zeta^2+R_1(\zeta),\qquad |R_1(\zeta)|\leqslant\min(\zeta^2,|\zeta|^3),$$

$$e^{\zeta^2/2}=1+\frac{1}{2}\zeta^2+R_2(\zeta),\qquad |R_2(\zeta)|\leqslant\zeta^4e^{\zeta^2/2},$$

we get

$$\mathbb{E}\big[\big(e^{iu\Delta\tilde{M}_k+u^2\theta_k^2/2}-1\big)|\mathscr{F}_{k-1}\big]$$

$$=\mathbb{E}\Big[\Big(1+iu\Delta\tilde{M}_k-\frac{1}{2}(u\Delta\tilde{M}_k)^2+R_1(u\Delta\tilde{M}_k)\Big)\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)-1\Big|\mathscr{F}_{k-1}\Big]$$

$$=\mathbb{E}\Big[\Big(1+iu\Delta\tilde{M}_k-\frac{1}{2}(u\Delta\tilde{M}_k)^2+R_1(u\Delta\tilde{M}_k)\Big)\Big|\mathscr{F}_{k-1}\Big]\underbrace{\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)}_{\hat{\in}\mathscr{F}_{k-1}}-1$$

(by the definition of θ_k),

$$=\mathbb{E}\Big[\Big(1-\frac{1}{2}(u\Delta\tilde{M}_k)^2+R_1(u\Delta\tilde{M}_k)\Big)\Big|\mathscr{F}_{k-1}\Big]\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)-1$$

(since $\mathbb{E}[\Delta\tilde{M}_k|\mathscr{F}_{k-1}]=0$, due to the martingale property of M.)

$$=\Big(1-\frac{1}{2}u^2\theta_k^2+\mathbb{E}[R_1(u\Delta\tilde{M}_k)|\mathscr{F}_{k-1}]\Big)\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)-1$$

(by the definition of θ_k),

$$=\Big(1-\frac{1}{2}u^2\theta_k^2+\mathbb{E}[R_1(u\Delta\tilde{M}_k)|\mathscr{F}_{k-1}]\Big)\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)-1$$

$$=\Big(1+\frac{1}{2}(u\theta_k)^2+R_2(u\theta_k)\Big)\mathbb{E}[R_1(u\Delta\tilde{M}_k)|\mathscr{F}_{k-1}]+\Big(1-\frac{1}{2}u^2\theta_k^2\Big)R_2(u\theta_k)-\frac{1}{4}(u\theta_k)^4$$

where we note that the term $\propto u^2\theta_k^2$ has coefficient zero. Now for ε small enough

$$\big|R_1(u\Delta\tilde{M}_k)\big|\leqslant\min\big((u\Delta\tilde{M}_k)^2,|u\Delta\tilde{M}_k|^3\big)\leqslant u^2\mathbb{1}_{|\Delta\tilde{M}_k|>\varepsilon}(\Delta\tilde{M}_k)^2+\varepsilon|u|^3\mathbb{1}_{|\Delta\tilde{M}_k|\leqslant\varepsilon}|\Delta\tilde{M}_k|^2$$

and

$$|R_2(u\theta_k)| \leqslant (u\theta_k)^4 e^{(u\theta_k)^2/2} \leqslant (u\theta_k)^4 e^{u^2 C/2}$$

where we used that by eq. (4) we have

$$\sup_k \theta_k^2 \leqslant \sum_{k=1}^{\infty} \theta_k^2 \leqslant C.$$

Therefore

$$D_k \leqslant \mathbb{E}\left[\left|\left(1+\frac{1}{2}(u\theta_k)^2 + R_2(u\theta_k)\right)\mathbb{E}[R_1(u\Delta\tilde{M}_k)|\mathscr{F}_{k-1}]\right|\right]$$

$$+\mathbb{E}\left[\left|\left(1-\frac{1}{2}u^2\theta_k^2\right)R_2(u\theta_k)\right|\right] + \mathbb{E}\left[\left|\frac{1}{4}(u\theta_k)^4\right|\right] \lesssim_{u,C} \mathbb{E}[|\mathbb{E}[R_1(u\Delta\tilde{M}_k)|\mathscr{F}_{k-1}]|] + \mathbb{E}[|\theta_k|^4]$$

$$\lesssim_{u,C} \mathbb{E}[\mathbb{1}_{|\Delta\tilde{M}_k|>\varepsilon}(\Delta\tilde{M}_k)^2] + \varepsilon\mathbb{E}[|\Delta\tilde{M}_k|^2] + \mathbb{E}[|\theta_k|^4].$$

But now observe that

$$\sum_k \mathbb{E}[\theta_k^4] = \sum_k \mathbb{E}[(\mathbb{E}[(\Delta\tilde{M}_k)^2|\mathscr{F}_{k-1}])^2]$$

$$\leqslant \varepsilon^2 \sum_k \mathbb{E}[\mathbb{1}_{|\Delta\tilde{M}_k|\leqslant\varepsilon}(\Delta\tilde{M}_k)^2] + C\mathbb{E}\left[\sum_k \mathbb{E}[\mathbb{1}_{|\Delta\tilde{M}_k|>\varepsilon}(\Delta\tilde{M}_k)^2|\mathscr{F}_{k-1}]\right].$$

This and the bound (4) then imply that

$$\sum_k D_k \lesssim_{u,C} \sum_k \mathbb{E}[\mathbb{1}_{|\Delta\tilde{M}_k|>\varepsilon}(\Delta\tilde{M}_k)^2] + \varepsilon^2 \mathbb{E}\left[\sum_k |\theta_k|^2\right] \lesssim_{u,C} \sum_k \mathbb{E}[\mathbb{1}_{|\Delta\tilde{M}_k|>\varepsilon}(\Delta\tilde{M}_k)^2] + \varepsilon^2$$

From which we conclude that $\sum_k D_k \to 0$ as $n \to \infty$ and subsequently $\varepsilon \to 0$. In order to complete the proof we have now to remove the assumption (4). Take $C > 1$, let $A_m = \{\langle\tilde{M}\rangle_m \leqslant C\}$ and observe that $A_m \in \mathscr{F}_{m-1}$, $A_m \subseteq A_{m+1}$ and that $\mathbb{P}(A_n) \to 1$ as $n \to \infty$ since $\langle\tilde{M}\rangle_n \to 1$ in probability. Now define a process $(\hat{M}_m)_m$ such that $\hat{M}_0 = 0$ and $\Delta\hat{M}_m = \mathbb{1}_{A_m}\Delta\tilde{M}_m$. This is a martingale transform of $(\tilde{M}_m)_m$ since the process $(\mathbb{1}_{A_m})_m$ is previsible, therefore $(\hat{M}_m)_m$ is a martingale and $\langle\hat{M}\rangle_n = \sum_{k=1}^n \mathbb{E}[(\Delta\hat{M})_k|\mathscr{F}_{k-1}] = \sum_{k=1}^n \mathbb{1}_{A_k}\mathbb{E}[(\Delta\tilde{M})_k^2|\mathscr{F}_{k-1}] \leqslant \mathbb{1}_{A_n}\langle\tilde{M}\rangle_n \leqslant C$. We can apply the results above to $\hat{M}$ and conclude that, as $n \to \infty$, $\mathbb{E}[e^{iu\hat{M}_n}] \to e^{-u^2/2}$. But on the set A_n we have $\tilde{M}_n = \hat{M}_n$ and therefore

$$|\mathbb{E}[e^{iu\tilde{M}_n}] - e^{-u^2/2}| \leqslant \underbrace{|\mathbb{E}[e^{iu\hat{M}_n}] - e^{-u^2/2}|}_{\to 0} + \underbrace{|\mathbb{E}[\mathbb{1}_{A_n}(e^{iu\tilde{M}_n} - e^{iu\hat{M}_n})]|}_{=0} + \underbrace{|\mathbb{E}[\mathbb{1}_{A_n^c}(e^{iu\tilde{M}_n} - e^{iu\hat{M}_n})]|}_{\leqslant \mathbb{P}(A_n^c) \to 0} \to 0$$

and we have proven the required convergence for $(\tilde{M}_n)_n$. □

2 Backwards martingales

Let $\mathbb{Z}_- = \{-n : n \in \mathbb{N}\}$ and $(\mathscr{G}_n)_{n\leqslant 0}$ an increasing sequence of σ-algebras indexed by $\mathbb{Z}_-$. Note that $\mathscr{G}_n \subseteq \mathscr{G}_m$ for $n \leqslant m \leqslant 0$. A backward martingale $(X_n)_{n\leqslant 0}$ is a process adapted to $(\mathscr{G}_n)_{n\leqslant 0}$ with $X_0 \in L^1$ and for all $n \leqslant -1$ such that

$$\mathbb{E}[X_{n+1}|\mathscr{G}_n] = X_n.$$

By the tower property we have $X_n = \mathbb{E}[X_0|\mathscr{G}_n]$ and therefore $X_n \in L^1$ automatically and moreover $(X_n)_{n\leqslant 0}$ is a uniformly integrable family of random variables.

Theorem 2. *Let X be a backwards martingale with $X_0 \in L^p$ for some $p \geqslant 1$. Then $X_n \to X_{-\infty} = \mathbb{E}[X_0|\mathscr{G}_{-\infty}]$ as $n \to -\infty$, almost surely and in L^p. Here $\mathscr{G}_{-\infty} = \cap_{n\leqslant 0}\mathscr{G}_n$.*

Proof. The proof parallel those of the martingale case. In particular it uses a variation of Doob's upcrossing inequality for backwards martingales. Note that for all $N \geqslant 0$ we have that $(X_{-N+k})_{k=0,\ldots,N}$ is a (forward) martingale wrt the filtration $(\mathscr{G}_{-N+k})_{k=0,\ldots,N}$ so the usual upcrossing inequality applies to it and is not difficult from that to adapt the proofs of the martingale case to this setting. (We leave the details to the reader) □

3 Some applications

Recall that the *tail* σ-algebra $\mathscr{T}$ of a stochastic process $(X_n)_{n\geqslant 1}$ is defined as $\mathscr{T} = \cap_{n\geqslant 1}\mathscr{T}_n$ where $\mathscr{T}_n = \sigma(X_k\colon k \geqslant n)$, and it contains all the events which do not depend on any finite subset of $(X_n)_{n\geqslant 1}$.

Example 3. The event $\{\omega \in \Omega\colon X_n(\omega) \geqslant \lambda$ infinitely often (in n)$\} = \cap_{n\geqslant 1}\cup_{k\geqslant n}\{X_k \geqslant \lambda\}$ is in $\mathscr{T}_n$ for every n and therefore in the tail algebra $\mathscr{T}$. Similarly, random variables like $\limsup_n X_n$ or $\liminf_n X_n$ are also $\mathscr{T}$-measurable.

Theorem 4. *(Kolmogorov's 0-1 law) Let $(X_n)_{n\geqslant 1}$ be a family of independent random variables. Then $\mathscr{T}$ is trivial, i.e. if $A \in \mathscr{T}$ then $\mathbb{P}(A) \in \{0, 1\}$.*

Proof. Take $A \in \mathscr{T}$. Let $\mathscr{G}_n = \sigma((X_k)_{k\leqslant n})$ and $Z_n = \mathbb{E}[\mathbb{1}_A|\mathscr{G}_n]$ for all $n \geqslant 0$. Then $(Z_n)_{n\leqslant 0}$ a UI martingale for the filtration $(\mathscr{G}_n)_{n\geqslant 0}$ and $\mathscr{G}_n$ is independent of $\mathscr{T}_{n+1} \supseteq \mathscr{T}$. Therefore $Z_n = \mathbb{P}(A)$ and at the same time $Z_n \to Z_\infty = \mathbb{E}[\mathbb{1}_A|\mathscr{G}_\infty]$ a.s. but $\mathscr{G}_\infty \supseteq \mathscr{T}_n \supseteq \mathscr{T}$ for all $n \geqslant 0$ and as a consequence

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A|\mathscr{G}_\infty] = \mathbb{1}_A, \qquad \text{a.s.}$$

and we can conclude $\mathbb{P}(A) \in \{0, 1\}$. (In particular observe that if $\omega \in A$ then $\mathbb{1}_A(\omega) = 1 = P(A)$ while if $\mathbb{P}(A) = 0$ then for all $\omega \in A^c$ we have $0 = \mathbb{1}_A(\omega) = \mathbb{P}(A)$). □

Theorem 5. *(Strong law of large numbers) Let $(X_n)_{n\geqslant 1}$ an i.i.d. sequence with $X_1 \in L^1$. Let $S_n = X_1 + \cdots + X_n$ for $n \geqslant 1$, then $S_n / n \to \mathbb{E}[X_1]$ almost surely and in L^1.*

Proof. Let $\mathscr{G}_n = \sigma(S_n, X_{n+1}, X_{n+2}, \ldots)$ for $n \geqslant 1$, then $\mathscr{G}_m \subseteq \mathscr{G}_n$ for $n \leqslant m$. The process $(Z_n)_{n\leqslant 1}$ defined by $Z_n = S_{-n} / (-n)$ is a backwards martingale wrt. the filtration $(\mathscr{F}_n = \mathscr{G}_{-n})_{n\leqslant 0}$, indeed for all $n \geqslant 0$,

$$\mathbb{E}[Z_{-n+1}|\mathscr{F}_{-n}] = \frac{1}{n-1}\mathbb{E}[S_{n-1}|S_n, X_{n+1}, X_{n+2}, \ldots] = \frac{1}{n-1}\mathbb{E}[S_n - X_n|S_n, X_{n+1}, X_{n+2}, \ldots]$$

$$= \frac{1}{n-1}\left(S_n - \frac{S_n}{n}\right) = \frac{S_n}{n} = Z_{-n}$$

since $\mathbb{E}[X_n|S_n, X_{n+1}, X_{n+2}, \ldots] = \mathbb{E}[X_n|S_n] = \mathbb{E}[X_k|S_n]$ for all $k = 1, \ldots, n$ by the symmetry under exchange of the law of $(X_1, \ldots, X_n)$, from which we deduce also that

$$\mathbb{E}[X_n|S_n, X_{n+1}, X_{n+2}, \ldots] = \frac{\mathbb{E}[X_1|S_n] + \cdots + \mathbb{E}[X_n|S_n]}{n} = \frac{\mathbb{E}[S_n|S_n]}{n} = \frac{S_n}{n}.$$

Moreover we have $Z_1 = X_1 \in L^1$, therefore the backwards martingale is uniformly integrable. By the backwards martingale convergence theorem we deduce that $Z_{-n} \to Z_{-\infty}$ in L^1 and a.s. where $Z_{-\infty} = \mathbb{E}[X_1|\mathscr{F}_{-\infty}] = \mathbb{E}[X_1|\mathscr{G}_\infty]$. Note that $\mathscr{G}_\infty \subseteq \mathscr{T} = \cap_{n\geqslant 0}\sigma(X_k : k \geqslant n)$, therefore by Kolmogorov's 0-1 law the r.v. $Z_{-\infty}$ is almost surely constant and

$$Z_{-\infty} = \mathbb{E}[Z_{-\infty}] = \mathbb{E}[Z_{-1}] = \mathbb{E}[X_1]. \qquad \square$$

Theorem 6. *(Kakutani's product martingale theorem) Let $(X_n)_{n\geqslant 1}$ a sequence of independent, positive and mean 1 random variables. Let $M_0 = 1$ and $M_n = X_1 \cdots X_n$. Then $(M_n)_{n\geqslant 0}$ is a positive martingale and $M_n \to M_\infty$ a.s. as $n \to \infty$. Let $a_n = \mathbb{E}[X_n^{1/2}]$, then $a_n \in (0,1]$ and*

a) if $\prod_n a_n > 0$ then $M_n \to M_\infty$ in L^1 and $\mathbb{E}[M_\infty] = 1$;

b) if $\prod_n a_n = 0$ then $M_n \to 0$ a.s.

Proof. As $(M_n)_{n\geqslant 0}$ is a positive super-martingale it converges a.s. to a limit which we denote $M_\infty \in L^1$. Cauchy–Schwarz inequality gives $0 < a_n \leqslant 1$. Let $N_n = M_n^{1/2} / (\prod_{k\leqslant n} a_k)$. Then $(N_n)_{n\geqslant 1}$ is a positive martingale, which again converges a.s. to a limit N_∞. In case a) we have

$$\sup_n \mathbb{E}[N_n^2] = \sup_n \frac{\mathbb{E}[M_n]}{(\prod_{k\leqslant n} a_k)^2} = \sup_n \frac{1}{(\prod_{k\leqslant n} a_k)^2} < +\infty.$$

By Doob's inequality $\sup_n N_n \in L^2$ and therefore $\sup_n M_n = \sup_n [N_n^2 (\prod_{k\leqslant n} a_k)^2] \leqslant (\sup_n N_n)^2 \in L^1$. We conclude that $(M_n)_{n\geqslant 1}$ is UI since $M_n \leqslant \sup_n M_n \in L^1$ and therefore it converges a.s. and in L^1 and $\mathbb{E}[M_\infty] = \mathbb{E}[M_1] = 1$.

In case b) we have $M_\infty = \lim_n M_n = N_n^2 (\prod_{k\leqslant n} a_k)^2 = N_\infty^2 \lim_n (\prod_{k\leqslant n} a_k)^2 = 0$ a.s.. $\square$

Kakutani's theorem is related to the likelyhood ratio test. Assume you have a sequence of i.i.d. observations $(X_n)_{n\geqslant 1}$ and you want to test the null hypothesis H_0 that these observations are samples of a given law μ against the hypothesis H_1 that they are drawn from a different law ν which one assumes absolutely continuous wrt. μ with density f, that is $\mathrm{d}\nu = f \mathrm{d}\mu$. Consider the quantity

$$T_n = \sum_{i=1}^n \log f(X_i)$$

and note that the process $(M_n)_{n\geqslant 1}$ defined by $M_n = e^{T_n} = \prod_{i=1}^n f(X_i)$ is a martingale wrt. the probability $\mathbb{P}$ under which each X_i has law μ. Moreover $\mathbb{E}[M_n] = 1$. In this case $a_n = a_1 = \mathbb{E}[(\log f(X_1))^{1/2}]$ and by Jensen's inequality we have

$$a_1^2 < \mathbb{E}[\log f(X_1)] = 1,$$

as soon as the function $\log f(X_1)$ is not a.s. constant, that is when $\mu \neq \nu$. Then Kakutani's theorem allows to conclude that, almost surely $M_n \to 0$ which implies that $T_n \to -\infty$.

Consider also the more general situation where $(X_n)_{n\geqslant 1}$ are independent but not identically distributed and let $\mathbb{P}$ the law under which each of them has distribution μ while let $\mathbb{Q}$ the law under which each of them has law $\mathrm{d}\nu_n = f_n \mathrm{d}\mu$ for some sequence of density functions $(f_n)_n$. Then Kakutani theorem gives sufficient and necessary conditions under which $\mathbb{Q} \ll \mathbb{P}$. Indeed note that

$$\mathrm{d}\mathbb{Q}|_{\mathscr{F}_n}=Z_n\mathrm{d}\mathbb{P}|_{\mathscr{F}_n}$$

with $Z_n = \prod_{k=1}^n f_k(X_k)$. If $\prod_n \mathbb{E}[(\log f_n(X_n))^{1/2}] > 0$ the limit $Z_n \to Z_\infty$ exist a.s. and in L^1 and we have $Z_n = \mathbb{E}_{\mathbb{P}}[Z_\infty|\mathscr{F}_n]$ and therefore, for all $A \in \mathscr{F}_n$ we also have

$$\mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_A Z_n] = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_A Z_\infty]$$

from which we can conclude (how?) that $\mathrm{d}\mathbb{Q} = Z_\infty \mathrm{d}\mathbb{P}$. On the other hand if $\mathbb{Q} \ll \mathbb{P}$ then $\mathrm{d}\mathbb{Q} = Z_\infty \mathrm{d}\mathbb{P}$ and $Z_n = \mathbb{E}[Z_\infty|\mathscr{F}_n]$ which then implies that $Z_n \to Z_\infty$ a.s. and $\prod_n \mathbb{E}[(\log f_n(X_n))^{1/2}] > 0$. So

Corollary 7. *(In the conditions above)*

$$\mathbb{Q} \ll \mathbb{P} \Leftrightarrow \prod_n \mathbb{E}[(\log f_n(X_n))^{1/2}] > 0.$$

4 Densities, Radon-Nikodým derivatives

If f is a measurable non-negative function on the measure space $(\Omega, \mathscr{F}, \mu)$ then we can consider the set function $\mu_f(A) = \int_A f \mathrm{d}\mu = \int \mathbb{1}_A f \mathrm{d}\mu$. This function is actually a measure on $(\Omega, \mathscr{F})$ which has the property to be absolutely continuous wrt. μ. We write $\mathrm{d}\mu_f = f \mathrm{d}\mu$.

We say that a measure ν is absolutely continuous wrt. a measure μ (and write $\nu \ll \mu$) iff $\nu(A) = 0$ whenever $\mu(A) = 0$ for $A \in \mathscr{F}$.

Theorem 8. (Radon-Nikodým) *If μ, ν are two σ-finite measures on the measure space $(\Omega, \mathscr{F})$ then the following statements are equivalent:*

a) $\nu \ll \mu$;

b) $\mathrm{d}\nu = f \mathrm{d}\mu$ *for some non-negative measurable function f unique μ-almost everywhere.*

The (equivalence class) function f which realizes the representation above is called the Radon-Nikodým derivative of ν wrt. μ and denoted by $f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

Proof. It is enough to prove that a) $\Rightarrow$ b). Moreover by an easy partitioning of Ω we can reduce to prove the statement for finite measures ν, μ which we can take to be probability measures. And furthermore we can restrict us further to consider only the case when $\nu \leqslant \mu$. This is done by considering the measure $\tilde{\mu} = \mu + \nu$ for which $\nu \leqslant \tilde{\mu}$ and $\nu \ll \tilde{\mu}$. Provided the statement is true in this case we have $d\nu = g \mathrm{d}\tilde{\mu}$ and $\mathrm{d}\mu = f \mathrm{d}\tilde{\mu}$ and therefore since $\mu(f = 0) = 0$ implies $\tilde{\mu}(f = 0) = 0$ we also have $\mathrm{d}\mu = (f + \mathbb{1}_{f=0}) \mathrm{d}\tilde{\mu}$ and

$$d\nu = g \mathrm{d}\tilde{\mu} = g \mathrm{d}\tilde{\mu} = g(f + \mathbb{1}_{f=0})^{-1} \mathrm{d}\mu.$$

Let us therefore consider the case $\nu \leqslant \mu$ and we can certainly assume that $\mu \neq 0$. Let $\mathscr{R}$ be the family of partitions $\mathscr{P}$ of Ω made of measurable sets. This family $\mathscr{R}$ is a directed set once endowed with the order relation $\mathscr{P} \leqslant \mathscr{Q}$ iff $\sigma(\mathscr{P}) \subseteq \sigma(\mathscr{Q})$. For any $\mathscr{P} \in \mathscr{R}$ we let

$$f_{\mathscr{P}} := \sum_{A \in \mathscr{P}} \frac{\nu(A)}{\mu(A)} \mathbb{1}_A$$

where we understand that $\frac{\nu(A)}{\mu(A)} = 0$ if $\mu(A) = 0$. We have that $f_{\mathscr{P}}\colon \Omega \to [0,1]$ and that $\nu(A) = \int_A f_{\mathscr{P}} \mathrm{d}\mu$ for all $A \in \sigma(\mathscr{P})$ and moreover if $\mathscr{P} \leqslant \mathscr{Q}$ we have

$$\int_A f_{\mathscr{P}} \mathrm{d}\mu = \int_A f_{\mathscr{Q}} \mathrm{d}\mu, \qquad \forall A \in \sigma(\mathscr{P}). \tag{5}$$

This property implies that $(f_{\mathscr{P}})_{\mathscr{P} \in \mathscr{R}}$ is a *martingale* indexed by the directed set $\mathscr{R}$. Recall that $\mathscr{R}$ is directed if is a partially ordered set such that for $x, y \in \mathscr{R}$ there exists always $z \in \mathscr{R}$ such that $x \prec z$ and $y \prec z$. In particular partitions of a given set form a naturally directed set wrt. the refinements. It is not difficult to generalise our results on martingales from $\mathbb{N}$ to any directed set $\mathscr{R}$. We will study martingales later on, but for the moment we are going to need the following fundamental result: a bounded martingale converges μ-almost surely and in $L^1(\mu)$. If we call g such a limit and pass to the limit in $\mathscr{Q}$ in the relation (5), then we have

$$\int_A f_{\mathscr{P}} \mathrm{d}\mu = \int_A g \mathrm{d}\mu, \qquad \forall A \in \sigma(\mathscr{P}).$$

For any $A \in \mathscr{F}$ we can now consider a partition $\mathscr{P}$ which contains A, then we have

$$\nu(A) = \int_A f_{\mathscr{P}} \mathrm{d}\mu = \int_A g \mathrm{d}\mu$$

which therefore implies the claim. Let us show the uniqueness of g. Assume there exists another function $\hat{g}$ such that $\mathrm{d}\nu = \hat{g} \mathrm{d}\mu$, then for any $C > 0$:

$$\int \mathbb{1}_{\{C > \hat{g} > g > -C\}} (\hat{g} - g) \mathrm{d}\mu = \nu(\{C > \hat{g} > g > -C\}) - \nu(\{C > \hat{g} > g > -C\}) = 0$$

which implies that $\mathbb{1}_{\{C > \hat{g} > g > -C\}} (\hat{g} - g) = 0$ μ a.-e. and that $\mu(\{C > \hat{g} > g > -C\}) = 0$. Since this is true for all C and for exchanging $\hat{g}$ with g we have $\mu(\hat{g} \neq g) = 0$. □

Lemma 9. *Let μ, ν be σ-finite measures on $(\Omega, \mathscr{F})$ such that $\nu \ll \mu$. If f is a $\mathscr{F}$-measurable, ν-integrable function then, for all $A \in \mathscr{F}$:*

$$\int_A f \mathrm{d}\nu = \int_A f \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \mathrm{d}\mu.$$

Proof. We may assume μ is finite and $f \geqslant 0$. By monotone convergence is also possible to assume that f is bounded. Then let $\mathscr{H}$ be the family of all non-negative bounded $\mathscr{F}$-measurable functions for which the claim is true. Then $\mathscr{H}$ contains 1 by definition of RN derivative, is a vector space and it is stable by increasing limits thanks to the monotone convergence theorem. Therefore by the monotone class theorem $\mathscr{H}$ contains all the $\mathscr{F}$ measurable bounded functions since it is clear that it contains all the indicators of measurable sets. □