Conditional probability & conditional measures

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://squash.1gainesville.com/

1 April, 2019 (at *23:52*)

ABSTRACT: Abs. cty, Radon-Nikodym thm. Elementary martingale thy. In progress: As of 1Apr2019

Bonjour. As additional notation [♡1] use $\equiv$ to mean '*identically equals*'; on the probability space, we mean this a.e. Use r.var or r.v. for '*random variable*'. Use r.walk for '*random walk*'.

Sets & Fields. Use $\in$ for "is an element of". E.g, letting $\mathbb{P}$ be the set of primes, then, $5 \in \mathbb{P}$ yet $6 \notin \mathbb{P}$. Changing the emphasis, $\mathbb{P} \ni 5$ ("$\mathbb{P}$ *owns* 5") yet $\mathbb{P} \not\ni 6$.

For subsets A and B of the same space, Ω, the ***inclusion relation*** $A \subset B$ means:

$$\forall \omega \in A, \text{ necessarily } B \ni \omega.$$

And this can be written $B \supset A$. Use $A \subsetneqq B$ for *proper* inclusion, i.e, $A \subset B$ yet $A \neq B$.

The *difference set* $B \smallsetminus A$ is $\{\omega \in B \mid \omega \notin A\}$. Employ A^c for the ***complement*** $\Omega \smallsetminus A$. Use $A \triangle B$ for ***symmetric difference*** $[A \smallsetminus B] \cup [B \smallsetminus A]$. Furthermore

$A \boxdot B$,	Sets A & B have *at least* <u>one</u> point in common; they intersect.
$A \sqcap B$,	The sets have *no* common point; disjoint.

The symbol "$A \boxdot B$" both asserts intersection <u>and</u> represents the set $A \cap B$. For a collection $\mathcal{C} = \{E_j\}_j$ of sets in Ω, let the ***disjoint union*** $\bigsqcup_j E_j$ or $\bigsqcup(\mathcal{C})$ represent the union $\bigcup_j E_j$ *and also* assert that the sets are pairwise disjoint.

If there is a *measure* on the space then

$A \overset{\text{a.e}}{\sqcap} B$,	means their intersection is a nullset; it is empty a.e. (i.e *almost everywhere*)

In contrast, $A \overset{\text{a.e}}{\boxdot} B$ means that the sets intersect in positive mass.

A measurable space $(X, \mathcal{X})$, is a set X together with a ***field*** (a σ-algebra) $\mathcal{X}$ of subsets. Suppose we have a collection $\vec{\mathcal{G}} := \{\mathcal{G}_j\}_{j \in \mathcal{J}}$ of subfields. Given a subcollection $\mathcal{B} \subset \mathcal{J}$, define two new fields

$$\textbf{1:} \qquad \begin{aligned} \bigwedge\nolimits_{j \in \mathcal{B}} \mathcal{G}_j &:= \bigcap\nolimits_{j \in \mathcal{B}} \mathcal{G}_j \quad \text{and} \\ \bigvee\nolimits_{j \in \mathcal{B}} \mathcal{G}_j &:= \operatorname{Fld}\Bigl(\bigcup\nolimits_{j \in \mathcal{B}} \mathcal{G}_j\Bigr). \end{aligned}$$

(Field $\bigvee_{\mathcal{B}} \mathcal{G}_j$ is called the ***join*** of the $\mathcal{G}_j$ fields.) A natural partial-order $\leqslant$ is induced on $\mathcal{J}$ by

$$j \leqslant k \;\Leftarrow\; \mathcal{G}_j \subset \mathcal{G}_k.$$

Our $\mathcal{J}$ can be extended to be a *complete lattice* by, for each subset $\mathcal{B} \subset \mathcal{J}$, adjoining the two fields $\bigwedge_{\mathcal{B}} \mathcal{G}_j$ and $\bigvee_{\mathcal{B}} \mathcal{G}_j$.

Absolute continuity. Our measurable space is $(X, \mathcal{X})$, on which we have two measures μ and ν. Say that ν is ***absolutely continuous*** w.r.t μ (written $\nu \ll \mu$) if $\forall E \in \mathcal{X}$:

$$E \text{ a } \mu\text{-nullset} \quad \Longrightarrow \quad E \text{ a } \nu\text{-nullset}.$$

Stronger, say that "ν is ***uniformly*** abs-cts w.r.t μ" if: $\forall \varepsilon, \exists \delta$ such that $\forall E$:

$$\mu(E) \leqslant \delta \quad \Longrightarrow \quad \nu(E) \leqslant \varepsilon.$$

Write this as $\nu \overset{\text{strg}}{\ll} \mu$.

[♡1] **Phrases:** WLOG: '*Without loss of generality*'. TFAE: '*The following are equivalent*'. ITOf: '*In Terms Of*'. OTForm: '*of the form*'. FTSOC: '*For the sake of contradiction*'. Use iff: '*if and only if*'.

IST: '*It Suffices to*' as in ISTShow, ISTExhibit.

Use w.r.t: '*with respect to*' and s.t: '*such that*'.

Latin: e.g: *exempli gratia*, '*for example*'. i.e: *id est*, '*that is*'. N.B: *Nota bene*, '*Note well*'. QED: *quod erat demonstrandum*, meaning "end of proof".

Number Sets: An expression such as $k \in \mathbb{N}$ (read as "k is an element of $\mathbb{N}$" or "k in $\mathbb{N}$") means that k is a natural number; a ***natnum***.

$\mathbb{N}$ = natural numbers = $\{0, 1, 2, \dots\}$.

$\mathbb{Z}$ = integers = $\{\dots, -2, -1, 0, 1, \dots\}$. For the set $\{1, 2, 3, \dots\}$ of positive integers, the ***posints***, use $\mathbb{Z}_+$. Use $\mathbb{Z}_-$ for the negative integers, the ***negints***.

$\mathbb{Q}$ = rational numbers = $\{\frac{p}{q} \mid p \in \mathbb{Z} \text{ and } q \in \mathbb{Z}_+\}$. Use $\mathbb{Q}_+$ for the positive ***ratnums*** and $\mathbb{Q}_-$ for the negative ratnums.

$\mathbb{R}$ = reals. The ***posreals*** $\mathbb{R}_+$ and the ***negreals*** $\mathbb{R}_-$.

$\mathbb{C}$ = complex numbers, also called the ***complexes***.

For $\omega \in \mathbb{C}$, let "$\omega > 5$" mean "ω *is* ***real*** <u>*and*</u> $\omega > 5$". [Use the same convention for $\geqslant, <, \leqslant$, and also if 5 is replaced by any real number.]

Mathematical objects: Seq: '*sequence*'. poly(s): '*polynomial(s)*'. irred: '*irreducible*'. Coeff: '*coefficient*' and var(s): '*variable(s)*' and parm(s): '*parameter(s)*'. Expr.: '*expression*'. Fnc: '*function*' (so ratfnc: means rational function, a ratio of polynomials). cty: '*continuity*'. cts: '*continuous*'. diff'able: '*differentiable*'. CoV: '*Change-of-Variable*'. CoI: '*Constant of Integration*'. LoI: '*Limit(s) of Integration*'. RoC: '*Radius of Convergence*'.

Soln: '*Solution*'. Thm: '*Theorem*'. Prop'n: '*Proposition*'. CEX: '*Counterexample*'. eqn: '*equation*'. RhS: '*RightHand Side*' of an eqn or inequality. LhS: '*lefthand side*'. Sqrt or Sqroot: '*square-root*', e.g, "the sqroot of 16 is 4". Ptn: '*partition*', <u>*but*</u> pt: '*point*', as in "a fixed-pt of a map".

FTC: '*Fund. Thm of Calculus*'. IVT: '*intermediate-Value Thm*'. MVT: '*Mean-Value Thm*'.

The ***logarithm*** fnc, defined for $x>0$, is $\log(x) := \int_1^x \frac{\mathrm{d}v}{v}$. Its inverse-fnc is $\exp()$. For $x>0$, then, $\exp(\log(x)) = x = \mathrm{e}^{\log(x)}$. For real t, naturally, $\log(\exp(t)) = t = \log(\mathrm{e}^t)$. PolyExp: '*Polynomial-times-exponential*'. E.g, $F(t) := [3 + t^2] \cdot \mathrm{e}^{4t}$ is a polyExp.

Example. Let X be a denumerable set $\{p_1, p_2, \ldots\}$. Define

$$\mu(\{p_n\}) := 1/2^n \quad \text{and} \quad \nu(\{p_n\}) := 7\,.$$

Then $\nu \ll \mu$, but not uniformly. (Pt-atoms are not necessary; replace $\{p_n\}$ by the interval $(\frac{1}{2^{n+1}}, \frac{1}{2^n}]$.)

Looking ahead to the Radon-Nikodym derivative, note that $\frac{d\nu}{d\mu}(p_n) = 7 \cdot 2^n$. □

2: Prop'n. *$\nu \overset{strg}{\ll} \mu$ implies $\nu \ll \mu$. If $\nu(X) < \infty$, then the converse holds.* ◊

Proof. If $\nu \overset{\text{strg}}{\ll} \mu$ fails then there is an epsilon, say 7, and a sequence of sets so that

$$\mu(E_n) \to 0, \text{ as } n \to \infty, \text{ but each } \nu(E_n) \geqslant 7.$$

WLOG $\sum_n \mu(E_n)$ is finite. So by Borel-Cantelli, $\mu(G) = 0$ where

$$G := \bigcap_{k=1}^{\infty} U_k \quad \text{with} \quad U_k := \bigcup_{n=k}^{\infty} E_n.$$

Evidently each $\nu(U_k) \geqslant \nu(E_k) \geqslant 7$, and $U_1 \supset U_2 \supset \ldots$. Since $\nu(U_1) \leqslant \nu(X) < \infty$, we obtain the following equality

$$\nu(G) \;=\; \lim_{k\to\infty} \nu(U_k) \;\geqslant\; 7\,.$$

Hence $\nu \not\ll \mu$. ♦

3: Lemma. *Suppose $h\colon X \to \mathbb{R}$ is $\mathcal{X}$-measurable and*

$$\forall G \in \mathcal{X}: \quad \int_G h \, d\mu \;=\; 0\,.$$

Then h is constant-zero μ-a.e. ◊

Proof. By restricting h to the set $\{h \geqslant 0\}$, WLOG $h \geqslant 0$. Let Λ_n be the set of x with $h(x) \geqslant 1/n$. Integrating shows that

$$0 \;=\; \int_{\Lambda_n} h \;\geqslant\; \frac{1}{n} \cdot \mu(\Lambda_n)\,.$$

Hence Λ_n is a nullset. Hence $\bigcup_1^\infty \Lambda_n$ is null. ♦

Measures λ_0, λ_1 on $(X, \mathcal{X})$ are ***mutually singular***, written $\lambda_0 \perp \lambda_1$, if there is a (measurable) partition $X = A_0 \sqcup A_1$ so that $\lambda_0(A_1)$ and $\lambda_1(A_0)$ are each zero.

4: Lebesgue-Radon-Nikodym Thm. *On $(X, \mathcal{X})$ suppose we have a signed-measure ν and positive measure μ, each σ-finite. Then exists a unique pair of σ-finite signed-measures λ and ρ so that:*

$$\nu \;=\; \lambda + \rho\,, \text{ with } \lambda \perp \mu \text{ and } \rho \ll \mu\,.$$

*Furthermore, there is an μ-a.e-unique μ-integrable ($\mathcal{X}$-measurable) fnc $h\colon X \to \mathbb{R}$ so that $\boxed{\rho = \int h \, d\mu}$. The notation for this h is $\frac{d\rho}{d\mu}$; the **"Radon-Nikodym derivative** of ρ w.r.t μ".* ◊

Note that each measurable fnc f has unique decomposition into its ***positive part*** f^+ and ***negative part*** f^- (each as measurable as f), where

$$\text{5:} \quad \begin{aligned} f^+ \geqslant 0\,,\ f^- > 0 \quad \text{and} \quad f^+ - f^- \;&=\; f \\ \text{Further,} \quad f^+ + f^- \;&=\; |f|\,. \end{aligned}$$

6: Prop'n. *Let $\mathcal{Y} := \mathrm{Fld}(f)$, where $f \geqslant 0$. Then there exists a non-decreasing sequence*

$$\dagger\colon \qquad f_n \nearrow f \quad (\textit{convergence ptwise})$$

*of $\mathcal{Y}$-meas. **step functions** f_n. We can arrange that each f_n is bounded, and has only finitely-many steps. Or, allowing ∞-ly many steps, we can improve (†) to <u>uniform</u> convergence.* ◊

Proof. For a posint k, set $h_k(x) := \frac{1}{k} \cdot \lfloor k \cdot f(x) \rfloor$ and let $f_n := h_{2^n}$. (Finitely-many steps: cut off at $\pm n$.) ♦

Conditional Expectation. We work now on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, with subfields $\mathcal{G}, \mathcal{H} \subset \mathcal{F}$. An integrable random variable Y has a ***conditional expectation***, written $\mathsf{E}(\mathsf{Y} \mid \mathcal{G})$ or $\mathsf{E}_{\mathcal{G}}(\mathsf{Y})$, which is a r.var itself. It is characterized by:

^{CE}1: $\mathsf{E}(\mathsf{Y} \mid \mathcal{G})$ is integrable and $\mathcal{G}$-measurable.

^{CE}2: For each set $G \in \mathcal{G}$: $\int_G \mathsf{E}(\mathsf{Y} \mid \mathcal{G}) = \int_G \mathsf{Y}$.

If $\mathsf{Y}_0, \mathsf{Y}_1$ are each cond-expectations of Y w.r.t $\mathcal{G}$, then $\mathsf{Y}_0 \stackrel{\text{a.e}}{=} \mathsf{Y}_1$. [The $\mathcal{G}$-measurable difference $\mathsf{Y}_0 - \mathsf{Y}_1$ has zero-integral against each $G \in \mathcal{G}$. Now apply (3).]

Below, $\mathsf{E}(h) = \mathsf{E}(\mathsf{E}_{\mathcal{G}}(h)) = 0$, so (3) gives:

7: $[h \stackrel{\text{a.e}}{\geqslant} 0 \quad \& \quad \mathsf{E}_{\mathcal{G}}(h) \stackrel{\text{a.e}}{=} 0] \implies h \stackrel{\text{a.e}}{=} 0$.
If $\mathsf{E}_{\mathcal{G}}(\mathsf{Z}^2) \stackrel{\text{a.e}}{=} 0$ *then* $\mathsf{Z} \stackrel{\text{a.e}}{=} 0$.

8: Fact. *The conditional expectation operator has these properties:*

^{CE}**3:** *Linear:* $\mathsf{E}_{\mathcal{G}}(3\mathsf{Y} + 5\mathsf{Z}) = 3\mathsf{E}_{\mathcal{G}}(\mathsf{Y}) + 5\mathsf{E}_{\mathcal{G}}(\mathsf{Z})$.

^{CE}**4:** *Absorbing: For fields* $\mathcal{H} \supset \mathcal{G}$,

$$\mathsf{E}_{\mathcal{H}}(\mathsf{E}_{\mathcal{G}}(\mathsf{Y})) = \mathsf{E}_{\mathcal{G}}(\mathsf{Y}) = \mathsf{E}_{\mathcal{G}}(\mathsf{E}_{\mathcal{H}}(\mathsf{Y})).$$

^{CE}**5:** *Positive: If* $\mathsf{Y} \stackrel{\text{a.e}}{\geqslant} 0$ *then* $\mathsf{E}_{\mathcal{G}}(\mathsf{Y}) \stackrel{\text{a.e}}{\geqslant} 0$. *I.e, Order-preserving:* $\mathsf{Y}_1 \stackrel{\text{a.e}}{\geqslant} \mathsf{Y}_1 \implies \mathsf{E}_{\mathcal{G}}(\mathsf{Y}_1) \stackrel{\text{a.e}}{\geqslant} \mathsf{E}_{\mathcal{G}}(\mathsf{Y}_1)$. *Further,* $\mathsf{E}_{\mathcal{G}}(|\mathsf{Z}|) \geqslant |\mathsf{E}_{\mathcal{G}}(\mathsf{Z})|$.

^{CE}**6:** *For each* $p \in [0, \infty]$: $\mathsf{E}_{\mathcal{G}}()$ *is an* $\mathbb{L}^p$*-contraction. Indeed, its p-norm is 1.* ◊

9: Lemma. *Fix a subfield* $\mathcal{Y}$ *of our probability space. Suppose that* Y, Z *are integrable random vars, whose product* $\mathsf{Y} \cdot \mathsf{Z}$ *is integrable. If* Y *is* $\mathcal{Y}$*-measurable, then*

$$\mathsf{E}(\mathsf{Y} \cdot \mathsf{Z} \mid \mathcal{Y}) \stackrel{\text{a.e}}{=} \mathsf{Y} \cdot \mathsf{E}(\mathsf{Z} \mid \mathcal{Y}).$$ ◊

Proof. WLOG $\mathsf{Y}, \mathsf{Z} \geqslant 0$. WLOG Y is a step-fnc, measurable w.r.t $\mathcal{Y}$. Et.c. ♦

Example. Find a seq $\vec{\mathsf{Y}} \xrightarrow{\text{a.e}} 0$, with $0 \leqslant \mathsf{Y}_n$ and $\mathsf{E}(\mathsf{Y}_n) \leqslant 17$, together with a subfield $\mathcal{H}$ so that

No subseq. of $\vec{\mathsf{X}}$ a.e-converges,

where $\mathsf{X}_n := \mathsf{E}(\mathsf{Y}_n \mid \mathcal{H})$.

Soln. Let H and V each be copies of $(0, 1]$, and let $\Omega := H \times V$ equipped with area measure. Let $\mathcal{H}$ be the Borel field of H; now stretch it across Ω.

Let $(B_n)_{n=1}^{\infty}$ be an iid-seq of subsets of H –say, B_n is the set of points in $(0, 1]$ whose n^{th} bit is '1'. Let $I_n := (0, 1/n] \subset V$. Define

$$\mathsf{Y}_n := [\mathbf{1}_{B_n} \times \mathbf{1}_{I_n}] \cdot n.$$

So $\mathsf{E}(\mathsf{Y}_n)$ is $\frac{1}{2}$. And

$$\mathsf{E}(\mathsf{Y}_n \mid \mathcal{H}) = \mathbf{1}_{B_n} \times \mathbf{1}_V.$$

Since $[n \mapsto \mathbf{1}_{B_n}]$ is iid, no subsequence converges in the a.e.-sense. □

Probabilistic interpretations

What is the expected time, E, to first get "heads", when flipping a p-coin? Letting $q := 1-p$ be the probability of Tails, we have the recurrence

$$E = 1 + [p \cdot 0 + q \cdot E].$$

Its non-negative solns are $E = \frac{1}{p}, +\infty$. But E equals $\sum_{N=1}^{\infty} q^{N-1} pN$, which is finite. So

10: *Independently flipping a p-coin, the expected number of flips till "heads" is* $1/p$ *flips.*

SMartingales

We now let $\mathcal{J}$ denote an ordered set $(\mathcal{J}, \leqslant)$. A ***filtration*** $\vec{\mathcal{G}}$ (over $\mathcal{J}$) is an indexed collection of fields s.t

$$j \leqslant k \implies \mathcal{G}_j \subset \mathcal{G}_k, \quad \text{for all } j,k \in \mathcal{J}.$$

A $\mathcal{J}$***-martingale*** $(\vec{\mathsf{Y}}, \vec{\mathcal{G}})$ has *integrable* r.vars $\vec{\mathsf{Y}}$ (indexed by $\mathcal{J}$) so that $j \leqslant k$ implies

11: $$\mathsf{Y}_j = \mathsf{E}(\mathsf{Y}_k \mid \mathcal{G}_j).$$

Our indexing set $\mathcal{J}$ will usually be $[0 .. \infty)$ or $[0 .. \infty]$. Whenever $\mathcal{J} = [0 .. \infty)$ we will automatically define a field

$$\mathcal{G}_\infty := \text{Fld}(\bigcup_{j \in \mathcal{J}} \mathcal{G}_j).$$

(We do not need the generality of (1).) However, there may not exist a reasonable *random variable* Y_∞; the main goal of this section is studying when $\lim_{j \to \infty} \mathsf{Y}_j$ exists (in various senses) and when the limit r.var gives us a martingale in that $\mathsf{E}(\mathsf{Y}_\infty \mid \mathcal{G}_j) = \mathsf{Y}_j$.

We sometimes use $\vec{\mathsf{Y}}$ to abbreviate $(\vec{\mathsf{Y}}, \vec{\mathcal{G}})$, where the $\vec{\mathcal{G}}$ fields are known. If they aren't, then we let

$$\mathcal{G}_k := \bigvee_{j : j \leqslant k} \text{Fld}(\mathsf{Y}_j);$$

this is the smallest field making all the preceding random variables measurable.

Replacing (11) by $\mathsf{Y}_j \leqslant \mathsf{E}(\mathsf{Y}_k \mid \mathcal{G}_j)$ gives a ***submartingale***, and by $\mathsf{Y}_j \geqslant \mathsf{E}(\mathsf{Y}_k \mid \mathcal{G}_j)$, a ***supermartingale***. I'll abbreviate the three notions by MG, subMG and superMG. We'll use Chung's term ***smartingale*** (or ***sMG***) for a process $\vec{\mathsf{Y}}$ which is any one of these three types.

Stopping-times

Henceforth our indexset $\mathcal{J}$ is $\mathbb{N} = [0..\infty)$ or $\dot{\mathbb{N}} := [0..\infty]$. We have a filtration $\vec{\mathcal{G}}$, and automatically a $\mathcal{G}_\infty$ field.

A ***stopping time*** τ (relative to $\vec{\mathcal{G}}$) is a *past-measurable* fnc $\tau{:}\Omega\to\dot{\mathbb{N}}$. That is, for each $N \in \dot{\mathbb{N}}$,

$$12: \qquad \{\tau \leqslant N\} \in \mathcal{G}_N .$$

Use *ST* and *STs* to abbrev. '*stopping time(s)*'. Condition (12) is equivalent to

$$12': \qquad \{\tau = j\} \in \mathcal{G}_j ,$$

due to the nesting of the fields, since $\{\tau \leqslant N\}$ equals $\bigcup_{j\leqslant N} \{\tau = j\}$.

13: Fact. *Take $\mathcal{G}, \mathcal{H}$ fields, and $A \in \mathcal{G}$. Then*

$$\mathcal{H}^{\langle A\rangle} := \{D \in \mathcal{H} \mid D \cap A \in \mathcal{G}\}$$

is a subfield of $\mathcal{H}$. ♢

Defn. A filtration $\vec{\mathcal{G}}$ and a *ST* $\alpha()$ give rise to a *new* field

$$\mathcal{G}_\alpha := \{D \in \mathcal{G}_\infty \mid \text{For each } N \in \dot{\mathbb{N}}\text{: } D \cap \{\alpha \leqslant N\} \in \mathcal{G}_N \} .$$

It <u>*is*</u> a field since, from (13), this $\mathcal{G}_\alpha$ equals $\bigcap_{N\in\dot{\mathbb{N}}} \mathcal{G}_\infty^{\langle A_N\rangle}$, where A_N is $\{\alpha \leqslant N\}$. Easily

$$\mathcal{G}_\alpha = \{D \in \mathcal{G}_\infty \mid \text{For each } N \in \dot{\mathbb{N}}\text{: } D \cap \{\alpha = N\} \in \mathcal{G}_N \} .$$

Exer. *E0.* Suppose $\alpha()$ is a constant *ST*, say, $\alpha \equiv 5$. Then $\mathcal{G}_\alpha$ indeed is $\mathcal{G}_5$. □

14: Fact. *For each $K \in \dot{\mathbb{N}}$: $\{\alpha \leqslant K\} \in \mathcal{G}_\alpha$.*

(I.e, stopping-time $\alpha()$ is $\mathcal{G}_\alpha$-measurable.) ♢

Proof. For $N \geqslant K$ note $\{\alpha \leqslant K\} \cap \{\alpha \leqslant N\} = \{\alpha \leqslant K\} \in \mathcal{G}_K \subset \mathcal{G}_N$.

When $N < K$ then $\{\alpha \leqslant K\} \cap \{\alpha \leqslant N\} = \{\alpha \leqslant N\} \in \mathcal{G}_N$. ♦

15: Lemma. *When $\alpha \leqslant \beta$ are STs then $\mathcal{G}_\alpha \subset \mathcal{G}_\beta$.* ♢

Proof. For each $N \in \dot{\mathbb{N}}$ we have that

$$16: \qquad \{\alpha \leqslant N\} \supset \{\beta \leqslant N\} .$$

Fix a set $D \in \mathcal{G}_\alpha$. Given N and letting

$$I := D \cap \{\beta \leqslant N\} ,$$

our goal is $I \in \mathcal{G}_N$. Happily,

$$\begin{aligned} I &= I \cap \{\alpha \leqslant N\} , \quad \text{by (16)},\\ &= [D \cap \{\alpha \leqslant N\}] \cap \{\beta \leqslant N\} . \end{aligned}$$

This lies in $\mathcal{G}_N \vee \mathcal{G}_N \overset{\text{note}}{=\!=} \mathcal{G}_N$. ♦

Examples of Martingales. Below we describe several MGs in terms of gambling. The probability space can be thought of as $\Omega := (0,1]$ or as a cantor set.

17: *The pre-divorced gambler.* The gambler has \$1 in his pocket, enters a casino and –at each stage– bets *all* his money on a fair game. He stops the first time that he is broke –which is the first time that he loses! His fortune r.v. at time n is

$$\mathsf{X}_n := 2^n \cdot \mathbf{1}_{(0,1/2^n]} .$$

Evidently we have almost-sure convergence $\mathsf{X}_n \xrightarrow{\text{a.e}} 0$ (but not $\mathbb{L}^1$ convergence). He comes home to his wife flat-broke. Moreover, he skulks home –on average– after two bets! (This, from (10).) □

18: *Win or Double-up.* This gambler starts with no money, $\mathsf{Y}_0 \equiv 0$; he his going to borrow to bet. He bets a buck: if wins, quits, else doubles his bet to \$2. If he wins, he quits, else he doubles-up again. Etc.

Evidently $\vec{\mathsf{Y}}$ is a disguised version of $\vec{\mathsf{X}}$; indeed

$$\mathsf{Y}_n = 1 - \mathsf{X}_n .$$

So $\vec{\mathsf{Y}} \xrightarrow{\text{a.e}} 1$, and $\vec{\mathsf{Y}}$ has the same convergence properties as $\vec{\mathsf{X}}$.

While this looks good for the gambler, we will later show that, in expectation, he must have infinitely deep pockets to implement this scheme. □

19: *Insanity that never quits.* Fix posints $H_n \nearrow \infty$ so that each $H_N \geqslant 3 \cdot \sum_{j=1}^{N-1} H_j$.

Write the prob-space as $\Omega := \{\pm 1\}^{\mathbb{Z}_+}$; a cantor set. This gambler borrows money from The Mob, and he never quits. At stage j he bets H_j dollars. So his (cumulative) fortune is

$$\mathsf{Z}_{N+1}(\omega) = \sum_{j=1}^{N} \omega_j \cdot H_j .$$

For an ω with $\omega_N = +1$ infinitely-often, evidently

$$\limsup_N \mathsf{Z}_N(\omega) = +\infty .$$

The liminf is $-\infty$ when $\omega_N = -1$ infinitely; evidently each of these events happens almost-surely (off the end-points of the cantor set). So *this* MG diverges almost-surely, in a spectacular way. (And –when The Mafia comes to collect its loan– things will spectacular as well.) □

Exer. *E1.* Create a mean-zero MG $\vec{\mathsf{Z}}$ such that $\mathsf{X} := \lim_n \mathsf{Z}_n$ exists-a.e. Arrange that $0 \leqslant \mathsf{X} < \infty$ and $\mathsf{E}(\mathsf{X}) = \infty$. □

Convention. When a filtration $\vec{\mathcal{G}}$ is known, agree to allow $\boxed{\mathsf{E}_j() \text{ to abbreviate } \mathsf{E}_{\mathcal{G}_j}()}$.

Doob decomposition of subMG. Some results about smartingales can be reduced to MGs.

20: Theorem. *Consider a subMG* $(\vec{\mathsf{S}}, \vec{\mathcal{G}})$. *Then there exists a MG* $\vec{\mathsf{Y}}$, *adapted to* $\vec{\mathcal{G}}$, *and an integrable positive process* $\vec{P}$ *so that*

d1: $\mathsf{S}_n = \mathsf{Y}_n + P_n$ (for $n = 0, 1, 2, \dots$).

d2: $0 = P_0 \leqslant P_1 \leqslant P_2 \leqslant P_3 \leqslant \dots$

d3: Each P_j *is measurable w.r.t* $\mathcal{G}_{j-1}$.

The $\vec{\mathsf{Y}}, \vec{P}$ *pair is unique.*

If $\vec{\mathsf{S}}$ *is* $\mathbb{L}^1$*-bounded, then so are* $\vec{\mathsf{Y}}$ *and* $\vec{P}$. *Indeed* $\mathsf{E}(|\vec{P}|) \leqslant 2B$, *where* $B := \mathsf{E}(|\vec{\mathsf{S}}|)$. ◊

Proof. We establish Uniqueness: For $j \geqslant 1$ certainly $\mathsf{E}_j(\mathsf{Y}_j - \mathsf{Y}_{j-1}) \equiv 0$, since $\vec{\mathsf{Y}}$ is a MG. Thus $\mathsf{E}_{j-1}(\mathsf{S}_j - \mathsf{S}_{j-1})$ equals $\mathsf{E}_{j-1}(P_j - P_{j-1})$. Courtesy (d3),

$$P_j - P_{j-1} = \mathsf{E}_{j-1}(\mathsf{S}_j) - \mathsf{S}_{j-1} .$$

Since $P_0 \equiv 0$, summing the telescoping series gives

21: $$P_N = \sum_{j \in [1..N]} [\mathsf{E}_{j-1}(\mathsf{S}_j) - \mathsf{S}_{j-1}] .$$

Thus $\vec{P}$ is uniquely determined, hence so is $\vec{\mathsf{Y}}$.

Existence. Define P_N by (21). Then $P_0 \equiv 0$ and $P_N \geqslant P_{N-1}$ since $\mathsf{E}_{N-1}(\mathsf{S}_N) - \mathsf{S}_{N-1} \geqslant 0$. And RhS(21) is $\mathcal{G}_{N-1}$-measurable, hence (d3).

As a finite sum, P_N is integrable; so Y_N too is integrable, when defined by (d1). To verify MG-ness we compute

$$\begin{aligned} \mathsf{Y}_N - \mathsf{Y}_{N-1} &= \mathsf{S}_N - \mathsf{S}_{N-1} - [P_N - P_{N-1}] \\ &= \text{same} - [\mathsf{E}_{N-1}(\mathsf{S}_N) - \mathsf{S}_{N-1}] \\ &= \mathsf{S}_N - \mathsf{E}_{N-1}(\mathsf{S}_N) . \end{aligned}$$

Conditioning this on $\mathcal{G}_{N-1}$ indeed gives 0.

$\mathbb{L}^1$***-boundedness.*** Observe that

$$\begin{aligned} \mathsf{E}_0(P_N) &= \sum_{j \in [1..N]} \mathsf{E}_0(\mathsf{E}_{j-1}(\mathsf{S}_j - \mathsf{S}_{j-1})) \\ &= \sum_j [\mathsf{E}_0(\mathsf{S}_j) - \mathsf{E}_0(\mathsf{S}_{j-1})] , \end{aligned}$$

which equals $\mathsf{E}_0(\mathsf{S}_N) - \mathsf{S}_0$. And $\int |P_N| = \int P_N$ i.e $\int \mathsf{E}_0(P_N)$, i.e $[\int \mathsf{S}_N] - \int \mathsf{S}_0$. ♦

Sampling. Henceforth, fix a MG $(\vec{\mathsf{Y}}, \vec{\mathcal{G}})$ over indexset $\mathcal{J}$. A $\mathcal{ST}$ τ is "***$\mathcal{J}$-stopping-time***" if the event $\{\tau() \notin \mathcal{J}\}$ is null. More strongly, a $\mathcal{ST}$ τ is ***$\mathcal{J}$-bounded*** if there exists $N_0 \in \mathcal{J}$ with $\tau() \leqslant N_0$. (So either $\mathcal{J} \ni \infty$ or else τ is bounded by some integer.) A $\mathcal{J}$-$\mathcal{ST}$ τ yields a random variable Y_τ defined, at each $\omega \in \Omega$, to be $[\mathsf{Y}_{\tau(\omega)}](\omega)$.

22: Lemma. *If τ is a $\mathcal{J}$-$\mathcal{ST}$ then $\mathsf{Y}_\tau \in \mathcal{G}_\tau$.* ◇

Proof. Take a Borel set $S \subset \mathbb{R}$. Fixing an $N \in \mathcal{J}$, we want to show that

$$\{\mathsf{Y}_\tau() \in S\} \cap \{\tau = N\}$$

is in $\mathcal{G}_N$. But this intersection equals

$$\{\mathsf{Y}_N \in S\} \cap \{\tau = N\} \stackrel{\text{note}}{\in} \mathcal{G}_N \vee \mathcal{G}_N\,.$$ ♦

23: *Integrability.* A Y_τ could have $\mathsf{E}(\mathsf{Y}_\tau) \neq \mathsf{E}(\mathsf{Y}_0)$: Let $\vec{\mathsf{Y}}$ be the std random-walk on $\mathbb{Z}$, and let τ stop at 7. So $\mathsf{E}(\mathsf{Y}_\tau) = 7 \neq 0 = \mathsf{E}(\mathsf{Y}_0)$.

Worse is a r.walk $\vec{\mathsf{Z}}$ and $\mathcal{ST}$ with $\mathsf{E}(\mathsf{Z}_\beta) = +\infty$: Set $\mathsf{Z}_0 := 0$. Let Z_1 jump to $\pm n$, each with prob$=\frac{1}{2}/2^n$, for $n = 1, 2, \ldots$. Depending on the value of $n := |\mathsf{Z}_1|$, our $\mathcal{ST}$ β stops at the first visit to position 3^n. So $\mathsf{E}(\mathsf{Z}_\beta)$ is $\sum_{n=1}^{\infty} [3/2]^n$. Even worse, we could modify β so arrange that Z_β simply *fails* to have an expectation.

What goes wrong in these examples is that the $\mathcal{ST}$ β is not $\mathcal{J}$-bounded. Fortunately:

24: *Imagine that β is a $\mathcal{J}$-bounded $\mathcal{ST}$ for $\mathcal{J}$-martingale $\vec{\mathsf{Y}}$. Then Y_β is integrable.*

This is implicit in the next proof, of Doob's thm, near the end. □

Generalizing the below: The next thm, as stated, applies to a MG. However, the proof goes through to show: *If $\vec{\mathsf{Y}}$ is a smartingale, then $(\mathsf{Y}_\alpha, \mathcal{G}_\alpha), (\mathsf{Y}_\beta, \mathcal{G}_\beta)$ is a two-term smartingale of the same type.* □

25: Doob's Optional Sampling Theorem. *Suppose that $\alpha \leqslant \beta$ are $\mathcal{J}$-bounded $\mathcal{ST}$s. Then*

25′: $$\mathsf{E}(\mathsf{Y}_\beta \mid \mathcal{G}_\alpha) = \mathsf{Y}_\alpha\,.$$ ◇

Proof. Said differently, we need to establish that $(\mathsf{Y}_\alpha, \mathsf{Y}_\beta)$ is a two-term martingale. We'll do this in two steps; by reducing to $(\mathsf{Y}_0, \mathsf{Y}_\beta)$, then to $(\mathsf{Y}_0, \mathsf{Y}_{17})$.

Fix an $K \in \mathcal{J}$; ISTShow (25′) when restricted to the set $\Omega' := \{\alpha = K\}$, since Ω' is in $\mathcal{G}_\alpha$, courtesy (14). So WLOG $\alpha \equiv K$. Since Y_K is integrable, we can subtract it to define new sequences, for $n \geqslant K$, by

$$\widetilde{\mathsf{Y}}_{n-K} := \mathsf{Y}_n - \mathsf{Y}_K \quad \text{and}$$
$$\widetilde{\mathcal{G}}_{n-K} := \mathcal{G}_n\,.$$

Renaming $(\widetilde{\mathsf{Y}}_k, \widetilde{\mathcal{G}}_k)$ to $(\mathsf{Y}_k, \mathcal{G}_k)$ gives:

$$\textit{WLOG} \;\; \alpha \equiv 0 \;\; \textit{and} \;\; \mathsf{Y}_0 \equiv 0.$$

Our goal[♡2] is $\mathsf{E}(\mathsf{Y}_\beta \mid \mathcal{G}_0) \stackrel{\text{a.e}}{=} \mathsf{Y}_0$. (For the sequel, we don't need that $\mathsf{Y}_0 \equiv 0$, but the reader may find this extra knowledge helpful in understanding the argument.) Restating, for each set $\Gamma \in \mathcal{G}_0$ we desire

$$\int_\Gamma \mathsf{E}(\mathsf{Y}_\beta \mid \mathcal{G}_0) \stackrel{?}{=} \int_\Gamma \mathsf{Y}_0\,.$$

Conditioning on Γ, then, we need but show that $\int_\Omega \mathsf{E}(\mathsf{Y}_\beta \mid \mathcal{G}_0) = \int_\Omega \mathsf{Y}_0$. Consequently

‡: $$\int_\Omega \mathsf{Y}_\beta \stackrel{?}{=} \int_\Omega \mathsf{Y}_0$$

is our goal.[♡3]

It is now time to use that β is $\mathcal{J}$-bounded. WLOG $\beta() \leqslant 17$. In consequence

$$\int_\Omega \mathsf{Y}_\beta = \sum_{j \leqslant 17} \int_{\{\beta=j\}} \mathsf{Y}_\beta = \sum_{j \leqslant 17} \int_{\{\beta=j\}} \mathsf{Y}_j = \sum_{j \leqslant 17} \int_{\{\beta=j\}} \mathsf{Y}_{17}\,;$$

this latter, since $\{\beta = j\}$ is in $\mathcal{G}_j$. The upshot is that

$$\int_\Omega \mathsf{Y}_\beta = \int_\Omega \mathsf{Y}_{17} = \int_\Omega \mathsf{Y}_0\,,$$

since –by hypothesis– the pair $(\mathsf{Y}_0, \mathsf{Y}_{17})$ is a two-term martingale. ♦

[♡2] IOWords, we have reduced the problem to showing that $(\mathsf{Y}_0, \mathsf{Y}_\beta)$ is a two-term martingale.

[♡3] This used that $\int \mathsf{E}(\mathsf{Y}_\beta \mid \mathcal{G}_0) = \int \mathsf{Y}_\beta$, which goes all the way back to knowing that, originally, $\mathcal{G}_\beta \supset \mathcal{G}_\alpha$.

26: Corollary. *$(\vec{\mathsf{Y}}, \vec{\mathcal{G}})$ a $\mathcal{J}$-subMG, and $N \in \mathcal{J}$. For each posreal b:*

$$\mathsf{P}(S) \quad \leqslant \quad \tfrac{1}{b} \cdot \mathsf{E}(|\mathsf{Y}_N|)\,,$$

where S is event $\{[\sup_{j \leqslant N} \mathsf{Y}_j] \geqslant b\}$. ◊

Proof. WLOG $b = 7$. Let τ be the $[1..N]$-infimum of those j with $\mathsf{Y}_j \geqslant 7$. Thus

$$7 \cdot \mathsf{P}(S) \;=\; \int_S 7 \;\leqslant\; \int_S \mathsf{Y}_\tau \;\leqslant\; \int_S \mathsf{Y}_N\,;$$

this latter, since $S \in \mathcal{G}_\tau$ and $(\mathsf{Y}_\tau\,, \mathsf{Y}_N)$ is a subMG. ♦

27: Application. *Suppose MG $\vec{\mathsf{Y}}$ has pointwise bound*

$$*: \qquad \forall n: \quad |\mathsf{Y}_{n+1} - \mathsf{Y}_n| \;\leqslant\; 7\,.$$

Suppose β is an integrable ST. Then Y_β is integrable and $\mathsf{E}(\mathsf{Y}_\beta) = \mathsf{E}(\mathsf{Y}_0)$. ◊

Proof. The tool we use is: *If $\vec{\mathsf{Z}}$ integrable and it $\mathbb{L}^1$-converges to a r.v. Z_∞, then $\mathsf{E}(\mathsf{Z}_n) \to \mathsf{E}(\mathsf{Z}_\infty)$.*

Automatically $\mathsf{Z}_N := \mathsf{Y}_{\beta \wedge N}$ is integrable. For each $k > N$, by $(*)$, the difference $|\mathsf{Y}_k - \mathsf{Y}_N| \leqslant 7 \cdot [k - N] \leqslant 7 \cdot k$. Estimating the $\mathbb{L}^1$-norm,

$$\|\mathsf{Y}_k - \mathsf{Z}_N\| \;\leqslant\; \sum_{k:k>N} \int_{\{\beta()=k\}} |\mathsf{Y}_k - \mathsf{Y}_N| \;\leqslant\; \sum_{k:k>N} \int_{\{\beta=k\}} 7k \;=\; 7 \cdot \int_{\{\beta>N\}} \beta\,.$$

This last goes to zero, since $\mathsf{E}(\beta) < \infty$.

So ISTShow that $\mathsf{E}(\mathsf{Z}_N) \stackrel{?}{=} \mathsf{E}(\mathsf{Y}_0)$. Here is the only place that we use the MG property: Doob's Optional Sampling, (25), tells us that the pair $(\mathsf{Y}_0, \mathsf{Y}_{\beta \wedge N})$ is a two-term MG, since $0 \leqslant \beta \wedge N$ are bounded stopping-times. ♦

Exer. A2. Consider an independent random-walk on the integers, where each step-probability depends on both position and time.

A ***3-spread*** $D()$ is a mean-zero random variable with support on $J := [-3..{+3}]$. That is,

$$\sum_{j \in J} \mathsf{P}(D{=}j) \;=\; 1 \quad \text{and}$$

$$\mathsf{E}(D) \stackrel{\text{note}}{=\!=} \sum_{j \in J} j \cdot \mathsf{P}(D{=}j) \;=\; 0\,.$$

For each *time* $n \in \mathbb{Z}_+$ and *position* $p \in \mathbb{Z}$, we have a 3-spread $D_{n,p}$, and all these random variables are mutually independent. Define random-walk $\vec{\mathsf{S}}$ by $\mathsf{S}_0 \equiv 0$ (we start at the origin) and

$$\mathsf{S}_{n+1} \;:=\; \mathsf{S}_n + D_{n+1\,,\,\mathsf{S}_n}\,.$$

Let $\tau()$ be the stopping time where the r.walk first hits position "5". Prove that $\mathsf{E}(\tau)$ is infinite. □

Soln. For each natnum N let

$$\mathcal{G}_N \;:=\; \bigvee_{\substack{j \in [1..N) \\ p \in \mathbb{Z}}} \mathrm{Fld}(D_{j,p})\,.$$

So Trivial $= \mathcal{G}_0 \subset \mathcal{G}_1 \subset \ldots$. The independence implies $D_{N+1,\,p} \perp \mathcal{G}_N$. Restated

$$†: \qquad \mathsf{E}_N(\,D_{N+1,\,p}) \;\stackrel{\text{a.e}}{=}\; 0\,.$$

Measurability: Note $\mathsf{S}_0 \in \mathcal{G}_0$. To show each $\mathsf{S}_N \in \mathcal{G}_N$, we will confirm

$$[\mathsf{S}_7 \in \mathcal{G}_7] \;\Longrightarrow\; [\mathsf{S}_8 \in \mathcal{G}_8]\,,$$

the induction step. For each integer p let

$$B_p \;:=\; \{\omega \mid \mathsf{S}_7(\omega) = p\}$$

Each $D_{8,p} \in \mathcal{G}_8$, so $\mathsf{S}_7 + D_{8,p} \in \mathcal{G}_8$. And $B_p \in \mathcal{G}_7 \subset \mathcal{G}_8$, so the product $[\mathsf{S}_7 + D_{8,p}] \cdot \mathbf{1}_{B_p}$ is $\mathcal{G}_8$-measurable. As a result,

$$\sum_{p \in \mathbb{Z}} [\mathsf{S}_7 + D_{8,p}] \cdot \mathbf{1}_{B_p} \quad \stackrel{\text{note}}{=\!=} \quad \mathsf{S}_8$$

is $\mathcal{G}_8$-measurable.

Integrability: $\mathrm{Range}(\mathsf{S}_N) \subset [-3^N..3^N]$, whence S_N is bounded, hence integrable.

Martingale-ness. ISTDemonstrate that

$$‡: \qquad \mathsf{E}_7(\mathsf{S}_8) \;\stackrel{\text{a.e}}{=}\; \mathsf{S}_7\,.$$

Fix p. Because $B_p \in \mathcal{G}_7$, ISTEstablish (‡) *on set* B_p. There, $\mathsf{S}_8 = \mathsf{S}_7 + D_{8,p}$; so $\mathsf{E}_7(\mathsf{S}_8) = \mathsf{S}_7 + \mathsf{E}_7(D_{8,p})$. Now (†) completes the argument. ♦

Inequalities

Below, J always denotes a subinterval of J. A fnc $f{:}J{\to}\mathbb{R}$ is ***convex*** (for emphasis, some say "convex-up") if the set $\{(x, y) \mid x \in J \;\&\; y \geqslant f(x)\}$ is a convex subset of the plane.

Henceforth, let $\mathcal{A}$ be the set of linear (well, *affine*) fncs $L{:}\mathbb{R}{\to}\mathbb{R}$. Use $\mathcal{B} = \mathcal{B}_f \subset \mathcal{A}$ for the subset of fncs L lying below, i.e, $L() \leqslant f()$. Let $\mathcal{Q} \subset \mathcal{A}$ be the set of linear fncs with *rational slope* and that pass through some rational point.

28: Lemma. *$f: J \to \mathbb{R}$ convex, on an open interval J. Then, pointwise,*

$$f() \;=\; \sup_{L \in \mathcal{B}_f} L()\,.$$

Indeed, $f = \sup_{L \in \mathcal{C}} L$ holds for a certain countable subcollection $\mathcal{C} \subset \mathcal{B}$. If $J = \mathbb{R}$ and f is linear, then $\mathcal{C} := \{f\}$. Otherwise, let $\mathcal{C} := \mathcal{Q} \cap \mathcal{B}_f$. ♢

Proof. Exercise. ♦

29: Jensen's Inequality (Thm). *Take $f: J \to \mathbb{R}$ convex, on an open interval J. Suppose Y is an integrable r.v. with range in J. Then*

$$f(\mathsf{E}(\mathsf{Y} \mid \mathcal{G})) \;\overset{\text{a.e}}{\leqslant}\; \mathsf{E}(f(\mathsf{Y}) \mid \mathcal{G})\,,$$

for each field $\mathcal{G}$ on the probability space. ♢

Proof. Take a set $\mathcal{C}$ of linears with $f = \sup_{L \in \mathcal{C}} L$. Let $\mathsf{E}(\cdot)$ denote $\mathsf{E}(\cdot \mid \mathcal{G})$. Fixing a version of $\mathsf{E}(\mathsf{Y})$, we can let $L(\mathsf{E}(\mathsf{Y}))$ be the *definition* of $\mathsf{E}(L(\mathsf{Y}))$. Taking sups gives this pointwise equality,

$$\dagger: \qquad f(\mathsf{E}(\mathsf{Y})) \;=\; \sup_{L \in \mathcal{C}} \mathsf{E}(L(\mathsf{Y}))\,.$$

For each L we have, since $\mathsf{E}()$ is a positive operator,

$$\mathsf{E}(L(\mathsf{Y})) \;\overset{\text{a.e}}{\leqslant}\; \mathsf{E}(f(\mathsf{Y}))\,.$$

While we can a choose a version of $\mathsf{E}(f(\mathsf{Y}))$ making the "a.e" nullset actually empty, it is unclear how to do make this choice work for *every* $L \in \mathcal{C}$. We'd like to be able to say

$$\ddagger: \qquad \sup_{L \in \mathcal{C}} \mathsf{E}(L(\mathsf{Y})) \;\overset{\text{a.e}}{\leqslant}\; \mathsf{E}(f(\mathsf{Y}))\,.$$

However, if $\mathcal{C}$ is uncountable then we seem to in danger of an uncountable union of nullsets.

Courtesy (28), we can use a *countable* $\mathcal{C}$. Now $(\dagger,\ddagger)$ together give the lemma. ♦

30: Corollary. *$f: J \to \mathbb{R}$ convex-up on an open interval J, and $\vec{\mathsf{Y}}$ is a process with range in J. Then*

$\vec{\mathsf{Z}}$ is a subMG, where $\mathsf{Z}_n := f(\mathsf{Y}_n)$,

if either: $\vec{\mathsf{Y}}$ is a MG –or– $\vec{\mathsf{Y}}$ is a subMG and f is non-decreasing. ♢

Proof. Fix n and let $\mathsf{E}(\cdot)$ mean $\mathsf{E}(\cdot \mid \mathcal{G}_n)$. So

$$\begin{aligned} \mathsf{E}(\mathsf{Z}_{n+1}) &\overset{\text{def}}{=\!=} \mathsf{E}(f(\mathsf{Y}_{n+1})) \\ &\geqslant f(\mathsf{E}(\mathsf{Y}_{n+1}))\,, \quad \text{by Jensen's}\,, \\ &\overset{*}{\geqslant} f(\mathsf{Y}_n) \overset{\text{def}}{=\!=} \mathsf{Z}_n\,. \end{aligned}$$

When $\vec{\mathsf{Y}}$ a MG then $(*)$ is equality. But for a subMG $\mathsf{E}(\mathsf{Y}_{n+1}) \geqslant \mathsf{Y}_n$, and here is where we use that f is non-decreasing. ♦

31: Cauchy-Schwarz Inequality. *Suppose Y, Z are square-integrable r.vars. Then $\mathsf{Y} \cdot \mathsf{Z}$ is integrable and*

$$\dagger: \qquad \mathsf{E}_{\mathcal{G}}(\mathsf{YZ})^2 \;\overset{\text{a.e}}{\leqslant}\; \mathsf{E}_{\mathcal{G}}(\mathsf{Y}^2) \cdot \mathsf{E}_{\mathcal{G}}(\mathsf{Z}^2)\,.$$ ♢

Proof. (Integrability of YZ follows from truncation.)

If $G := \{\mathsf{E}_{\mathcal{G}}(\mathsf{Z}^2) = 0\}$ has positive-mass, then condition on it. (Permissible, since $G \in \mathcal{G}$.) By (7), WLOG $\mathsf{Z} \equiv 0$. Hence the product $\mathsf{YZ} \equiv 0$. So $\mathsf{E}_{\mathcal{G}}(\mathsf{YZ}) \equiv 0$. Thus $(\dagger)$.

Let $\mathsf{E}() := \mathsf{E}_{\mathcal{G}}()$. WLOG the event $\{\mathsf{S} > 0\}$ is all of Ω, where

$$\mathsf{S} \;:=\; \mathsf{E}(\mathsf{Z}^2) \;\text{ and }\; \mathsf{M} \;:=\; \mathsf{E}(\mathsf{YZ})\,.$$

(The symbols come from square and mixed product.) Let $h := \mathsf{SY} - \mathsf{MZ}$. And h^2 is non-negative, so $0 \overset{\text{a.e}}{\leqslant} \mathsf{E}(h^2)$. Courtesy (9) and $\mathsf{S}, \mathsf{M} \in \mathcal{G}$, our $\mathsf{E}(h^2)$ equals

$$\begin{aligned} &\mathsf{S}^2 \cdot \mathsf{E}(\mathsf{Y}^2) + \mathsf{M}^2 \cdot \mathsf{E}(\mathsf{Z}^2) - 2\mathsf{SM} \cdot \mathsf{E}(\mathsf{YZ}) \\ \overset{\text{note}}{=\!=}\; &[\mathsf{S} \cdot \mathsf{E}(\mathsf{Y}^2) - \mathsf{M}^2]\mathsf{S}\,. \end{aligned}$$

Dividing by the positive S gives $(\dagger)$, in the form $0 \overset{\text{a.e}}{\leqslant} \mathsf{S} \cdot \mathsf{E}(\mathsf{Y}^2) - \mathsf{M}^2$. ♦

Convergence

Below, convergence of a sequence of reals means convergence in $[-\infty, \infty]$.

A process $\vec{\mathsf{Y}}$ is $\mathbb{L}^1$***-bounded*** if

$$B := \sup_n \int |\mathsf{Y}_n| \qquad \text{is finite.}$$

Write this bound B as $\mathsf{E}(|\vec{\mathsf{Y}}|)$.

32: Prop'n. *Suppose that an $\mathbb{L}^1$-bounded process $\vec{\mathsf{Y}}$ a.e-converges (in $[-\infty, +\infty]$) and call the limit X. Then X has the same bound, $\mathsf{E}(|\mathsf{X}|) \leqslant \mathsf{E}(|\vec{\mathsf{Y}}|)$.* ♢

Proof. We get a.e-convergence $|\mathsf{Y}_n| \xrightarrow{\text{a.e}} |\mathsf{X}|$ of the absolute values, so Fatou tells us that $\mathsf{E}(|\mathsf{X}|) \leqslant \liminf_n \mathsf{E}(|\mathsf{Y}_n|)$. ♦

Doob's notion of upcrossings. When does a seq. of reals, $\vec{y} = (y_j)_{j=0}^{\infty}$, converge? Certainly "Yes" if, for each pair of rationals $a < b$, there are only *finitely many* index-pairs $\alpha < \beta$ with $y_\alpha \leqslant a < b \leqslant y_\beta$. To count such ***upcrossings***, define times

33: $$\alpha_0 \leqslant \beta_0 \leqslant \alpha_1 \leqslant \beta_1 \leqslant \alpha_2 \leqslant \beta_2 \leqslant \ldots$$

by artificially letting $\beta_{-1} := 0$. For $j = 0, 1, \ldots$ let

$\alpha_j := n$ be the smallest $n \in [\beta_{j-1} .. \infty]$ with $y_n \leqslant a$;
$\beta_j := n$ be the smallest $n \in [\alpha_j .. \infty]$ with $y_n \geqslant b$.

(Here "smallest" means *infimum*; it is ∞ if no such n exists.) We say that $\vec{y}$ "***upcrosses*** the $[a, b]$-band as time goes from α_j to β_j".

Given a *process* $\vec{\mathsf{Y}}$ and sample-pt $\omega \in \Omega$, let $U^{a,b}(\omega) \in [0 .. \infty]$ count the number of upcrossings of sequence $\vec{\mathsf{Y}}(\omega)$. (The count $U^{a,b}(\omega)$ is the number of j having $\beta_j(\omega)$ finite.) For $\vec{\mathsf{Y}}$ to a.e-converge, then, we need but show that each $U^{a,b}$ is *a.e-finite.*

So showing each $\boxed{\mathsf{E}(U^{a,b}) < \infty}$, will suffice.

34: MCT (Martingale Convergence Thm). *For an $\mathbb{L}^1$-bnded smartingale $\vec{\mathsf{Y}}$, the almost-everywhere $\lim_n \mathsf{Y}_n =: \mathsf{X}$ exists. Indeed, $\mathsf{E}(|\mathsf{X}|) \leqslant \mathsf{E}(|\vec{\mathsf{Y}}|)$.* ♢

Reductions. It is enough to show, for a.e ω, that $\lim \vec{\mathsf{Y}}(\omega) =: \mathsf{X}$ exists in $[-\infty, +\infty]$. For then (32) tells us that X is finite-a.e and $\mathsf{E}(|\mathsf{X}|)$ has the same bound.

WLOG $\vec{\mathsf{Y}}$ is a subMG (replace Y_j by $-\mathsf{Y}_j$ to convert a superMG a subMG).

The upcrossing count of $\vec{\mathsf{Y}}$ over band-$[a, b]$ is that same as that of process

$$\mathsf{Z}_n := [\mathsf{Y}_n - a] \cdot \tfrac{1}{b-a}$$

upcrossing band-$[0, 1]$. Furthermore, $\vec{\mathsf{Z}}$ is still a subMG (since $b-a$ is positive) and $\vec{\mathsf{Z}}$ is still $\mathbb{L}^1$-bounded.

So ISTShow, for an *arbitrary* $\mathbb{L}^1$-bnded subMG, that $\mathsf{E}(U^{0,1})$ is finite. Verifying that

†: $\mathsf{E}(U) \overset{?}{\leqslant} \mathsf{E}(|\vec{\mathsf{Y}}|)$, *where U is the upcrossing count of the $[0, 1]$-band,*

would certainly suffice. □

Proof of MCT. Courtesy Jensen's Inequality, in form (30), we may assume that

35: $$\forall n : \quad \mathsf{Y}_n \geqslant 0,$$

simply by removing the negative part: Replace Y_n with $\mathrm{Max}(0, \mathsf{Y}_n)$. And $\vec{\mathsf{Y}}$ stays $\mathbb{L}^1$-bounded.

Fix N and let U_N count the number of upcrossings of $[0, 1]$ by $(\mathsf{Y}_1, \ldots, \mathsf{Y}_N)$; so cut-off the STs of (33) at N by redefining

$$\alpha_j := \mathrm{Min}(\alpha_j, N) \quad \text{and} \quad \beta_j := \mathrm{Min}(\beta_j, N).$$

(For each $j > \frac{N}{2}$, now, our $\alpha_j = \beta_j = N$.) Our noble goal (†) can be transmogrified into

††: $$\mathsf{E}(U_N) \overset{?}{\leqslant} \mathsf{E}(\mathsf{Y}_N).$$

Astronomy. Decompose Y_N as a telescoping sum,

$$\mathsf{Y}_N = \mathbf{P} + \mathbf{I} + \mathsf{Y}_{\alpha_0},$$

where the Positive and Integral–non-negative parts (names to be justified) are

$$\mathbf{P} := \sum_{j \in [0..N]} [\mathsf{Y}_{\beta_j} - \mathsf{Y}_{\alpha_j}];$$
$$\mathbf{I} := \sum_{k \in [0..N)} [\mathsf{Y}_{\alpha_{k+1}} - \mathsf{Y}_{\beta_k}].$$

For arbitrary stopping times $\sigma() \leqslant \tau() \leqslant N$ on our subMG, remark that

$$\int [\mathsf{Y}_\tau - \mathsf{Y}_\sigma] = \int [\mathsf{E}(\mathsf{Y}_\tau \mid \mathcal{G}_\sigma) - \mathsf{Y}_\sigma] \geqslant \int 0.$$

It follows that $\mathsf{E}(\mathbf{I})$ is non-negative.[♡4] Also non-negative is $\mathsf{E}(\mathsf{Y}_{\alpha_0})$, by (35). Thus $\mathsf{E}(\mathbf{P}) \leqslant \mathsf{E}(\mathsf{Y}_N)$.

[♡4] The same is true for $\mathsf{E}(\mathbf{P})$, but we don't want to discard $\mathbf{P}$.

So (††) will follow if we can establish this pointwise inequality,

$$‡: \qquad \forall\omega\in\Omega: \qquad U_N(\omega) \overset{?}{\leqslant} \mathbf{P}(\omega)\,.$$

To this heroic end, fix a sample point ω and now interpret $\mathbf{P}, U_N, \mathsf{Y}_n$ as *numbers*, rather than as random variables.

Let K be the smallest index j for which (α_j, β_j) is not an upcrossing. Thus

$$\text{For } j\in[0\,..\,K)\text{:}\quad \mathsf{Y}_{\beta_j} - \mathsf{Y}_{\alpha_j} \;\geqslant\; 1-0 \;=\; 1\,.$$
$$\text{For } j\in(K\,..\,N]\text{:}\quad \mathsf{Y}_{\beta_j} - \mathsf{Y}_{\alpha_j} \;=\; \mathsf{Y}_N - \mathsf{Y}_N \;=\; 0\,.$$

Summing, we see that $\mathbf{P} \geqslant U_N + [\mathsf{Y}_{\beta_K} - \mathsf{Y}_{\alpha_K}]$.

Our heart's desire now is to corroborate $\mathsf{Y}_{\beta_K} \geqslant \mathsf{Y}_{\alpha_K}$. We can dispense with the case where α_K is already N since, there, $\alpha_K = \beta_K$.

Thus $\alpha_K < N$ and so $\mathsf{Y}_{\alpha_K} = 0$. Since K did not give an upcrossing, it must be that $\mathsf{Y}_{\beta_K} < 1$ (and $\beta_K = N$). But how could we ever establish that

$$\mathsf{Y}_N \overset{?}{\geqslant} 0\,,$$

if we didn't have (35) at our disposal? We know that $\mathsf{Y}_N < 1$. But without our Jensen's Inequality step, this Y_N could be arbitrarily negative. Although (35) was used elsewhere in the proof, it is here where it is *crucially* used. ♦

Downcrossings. How can $\mathbf{I}$ have non-negative integral? After all, it is a sum of differences such as $\mathsf{Y}_{\alpha_7} - \mathsf{Y}_{\beta_6}$; and isn't that always a downcrossing?

Well for some ω, yes, $\mathsf{Y}_{\alpha_7} - \mathsf{Y}_{\beta_6}$ is a downcrossing and hence is $\leqslant 1$. Other ω have $\beta_6 = N$, so $\mathsf{Y}_{\alpha_7} - \mathsf{Y}_{\beta_6} = \mathsf{Y}_N - \mathsf{Y}_N$ is zero. But some ω *start* a downcrossing, $\mathsf{Y}_{\beta_6} \geqslant 1$, but never finish it. So $\alpha_7 = N$ and Y_{α_7} can be any posreal. *Here* is the case where the difference $\mathsf{Y}_{\alpha_7} - \mathsf{Y}_{\beta_6}$ can be arbitrarily positive —and this allows the integral $\mathsf{E}(\mathbf{I})$ to be positive. □

36: Theorem. *Fix a MG $(\vec{\mathsf{Y}}, \vec{\mathcal{G}})_{\mathbb{N}}$ and a r.var Z.*

Suppose $\mathsf{Y}_n \overset{\text{in } \mathbb{L}^1}{\longrightarrow} \mathsf{Z}$. Then $\mathsf{Z}\in\mathbb{L}^1$ and $\mathsf{E}(\mathsf{Z}\mid\mathcal{G}_N) = \mathsf{Y}_N$, for each N.

Conversely, recall $\mathcal{G}_\infty := \bigvee_{j<\infty}\mathcal{G}_j$ and suppose $\forall n : \mathsf{E}(\mathsf{Z}\mid\mathcal{G}_n) = \mathsf{Y}_n$. Then $\mathsf{Y}_n \overset{\text{in } \mathbb{L}^1}{\longrightarrow} \mathsf{Z}'$, where $\mathsf{Z}' := \mathsf{E}(\mathsf{Z}\mid\mathcal{G}_\infty)$. ◊

Proof. Take a $B\in\mathcal{G}_N$. For each $k>N$,

$$0 \;\leqslant\; \left|\int_B \mathsf{Y}_N - \int_B \mathsf{Z}\right| \;=\; \left|\int_B \mathsf{Y}_k - \int_B \mathsf{Z}\right| \;\leqslant\; \int_B |\mathsf{Y}_k - \mathsf{Z}|$$

Now sending $k\nearrow\infty$ corroborates $\int_B \mathsf{Y}_N = \int_B \mathsf{Z}$. ♦

Filename: `Problems/Analysis/Measures/co.martct-.latex`
As of: *Sunday 05Mar2006.* Typeset: *1Apr2019 at 23:52.*