- **Definition** (Densities, Absolutely Continuous)
  - Suppose δ is a "nonnegative" measurable function, define a measure ν by
    $$\nu(A) = \int_A \delta \, d\mu, \quad A \in \mathcal{F},$$
    then ν has "density" δ w.r.t. μ
  - The measure ν is "absolutely continuous" w.r.t. μ; $\nu \ll \mu$:
    $\mu(A) = 0$ implies $\nu(A) = 0$ for each A in $\mathcal{F}$.
  - $d\nu = \delta d\mu$; or $\frac{d\nu}{d\mu} = \delta$,
    If ν has density δ w.r.t μ, then for nonnegative or integrable f,
    $\int f d\nu = \int f \delta \, d\mu$, and $\int_A f d\nu = \int_A f \delta d\mu$.

- **Theorem** (Radon-Nikodym Thm)
  If μ and ν are σ-finite measures s.t. "$\nu \ll \mu$", then there exists a "nonnegative" f, a density, s.t.
  $$\nu(A) = \int_A f d\mu \quad \text{for all } A \in \mathcal{F}.$$
  For two such densities f and g, $\mu[f \neq g] = 0$.
  Remark:
  - If $\nu(A) = \int_A f d\mu$, f is nonnegative, then certainly $\nu \ll \mu$; The Radon-Nikodym thm goes in the opposite direction.
  - The uniqueness of the density f is up to sets of μ-measure 0.

• Definition (Conditional Expectation)

Suppose that X is an "integrable" r.v. on $(\Omega, \mathcal{F}, P)$,

$\mathcal{G}$ is σ-field in $\mathcal{F}$.

There exist a r.v. $E[X|\mathcal{G}]$,

called the "conditional expectation of X given $\mathcal{G}$",

having these two properties:

(i) $E[X|\mathcal{G}]$ is measurable $\mathcal{G}$ and integrable

(ii) $E[X|\mathcal{G}]$ satisfies the functional equation

$$\int_G E[X|\mathcal{G}]\,dP = \int_G X\,dP, \quad G \in \mathcal{G}.$$

Pf: • Consider first the case of nonnegative X.

Define a measure ν on $\mathcal{G}$ by $\nu(G) = \int_G X\,dP$.

$\because$ X is integrable, $\therefore$ ν is finite, and

$\because$ $P(G)=0$ implies $\nu(G)=0$, hence $\nu \ll P$.

By Radon-Nikodym thm, there is a nonnegative function f,

measurable $\mathcal{G}$, s.t. $\nu(G) = \int_G f\,dP$.

$\because$ ν is finite, $\therefore$ f is integrable; and

$$\int_G f\,dP = \int_G X\,dP, \quad G \in \mathcal{G}$$

$\therefore$ This f has properties (i) and (ii). Denote f by $E[X|\mathcal{G}]$.

• If X is not nonnegative,

$E[X^+|\mathcal{G}] - E[X^-|\mathcal{G}]$ clearly has the required properties.

#

Remark:

• There will in general be many such r.v.'s $E[X|\mathcal{G}]$;

any one of them is called "a version of the conditional expectation",

any two versions are equal with probability 1 (P restricted to $\mathcal{G}$).

• If $[X_t, t\in T]$ is a collection of r.v.'s,

$E[X|X_t, t\in T]$ is by definition $E[X|\mathcal{G}]$ with $\sigma[X_t, t\in T]$ in the role of $\mathcal{G}$.

• <u>Proposition</u> (Trivial Conditional Expectation)

(i) $E[X \| \{\emptyset, \Omega\}] = E[X]$ with probability 1;

(ii) $E[X \| \mathcal{G}] = X$ with probability 1, if X is measurable $\mathcal{G}$.

Remark:

- As $\mathcal{G}$ increases, condition (i) becomes weaker, (ii) becomes stronger
- For (i), the observer learns nothing from the experiment $\mathcal{G}$;

  for (ii), to know the outcome of the experiment $\mathcal{G}$ is to know r.v. X.

Pf: (i) If $\mathcal{G}$ is $\{\emptyset, \Omega\}$, the smallest σ-field,

every function measurable $\mathcal{G}$ is constant for all $\omega \in \Omega$,

$\because$ $E[X]$ is constant, $E[X]$ is measurable $\mathcal{G}$ and integrable; and

$\int_G E[X] \, dP = \int_G X \, dP$, $G \in \mathcal{G}$

$\therefore$ $E[X \| \{\emptyset, \Omega\}] = E[X]$ with probability 1

(ii) Suppose that X is measurable $\mathcal{G}$,

(which will hold if $\mathcal{G}$ coincides with the whole σ-field $\mathcal{F}$)

$\because$ X is measurable $\mathcal{G}$ and integrable by hypothesis; and

$\int_G X \, dP = \int_G X \, dP$, $G \in \mathcal{G}$

$\therefore$ $E[X \| \mathcal{G}] = X$ with probability 1, if X is measurable $\mathcal{G}$.

#

- <u>Theorem</u> (Uniqueness of Conditional Expectation, π-System)

Let $\mathcal{P}$ be a π-system generating the σ-field $\mathcal{G}$; $\mathcal{G} = \sigma(\mathcal{P})$, and suppose that Ω is a finite or countable union of sets in $\mathcal{G}$.

A integrable function f is an version of $E[X|\mathcal{G}]$ if

(i) f is measurable $\mathcal{G}$, and

(ii) $\int_G f\,dP = \int_G X\,dP$ for all G in $\mathcal{P}$.

Then $f = X$ on $\sigma(\mathcal{P}) = \mathcal{G}$; That is,

$f = E[X|\mathcal{G}] = X$ a.s., P restricted to $\mathcal{G}$.

- <u>Theorem</u> (Properties of Conditional Expectation)

All the equalities and inequalies below hold with probability 1:

Suppose that X, Y, X_n are integrable.

(i) If $X=a$ with probability 1, then $E[X|\mathcal{G}]=a$.

(ii) For constants a and b, $E[aX+bY|\mathcal{G}]=aE[X|\mathcal{G}]+bE[Y|\mathcal{G}]$

(iii) If $X \le Y$ with probability 1, then $E[X|\mathcal{G}] \le E[Y|\mathcal{G}]$

(iv) $|E[X|\mathcal{G}]| \le E[|X||\mathcal{G}]$.

(v) If $\lim_n X_n = X$ with probability 1, $|X_n| \le Y$, and Y is integrable, then $\lim_n E[X_n|\mathcal{G}] = E[X|\mathcal{G}]$ with probability 1.

Pf: (i) If $X=a$ with probability 1.

$\because$ a is measurable $\mathcal{G}$ and integrable; and

$\int_G a\,dP = \int_G X\,dP$, $G \in \mathcal{G}$,

$\therefore$ $E[X|\mathcal{G}]=a$ with probability 1.

(ii) Since $E[X|\mathcal{G}]$ and $E[Y|\mathcal{G}]$ are measurable $\mathcal{G}$ and integrable, then $aE[X|\mathcal{G}]+bE[Y|\mathcal{G}]$ is measurable $\mathcal{G}$ and integrable; and

$$
\begin{aligned}
&\int_G (aE[X|\mathcal{G}]+bE[Y|\mathcal{G}])\,dP \\
&= a\int_G E[X|\mathcal{G}]\,dP + b\int_G E[Y|\mathcal{G}]\,dP \\
&= a\int_G X\,dP + b\int_G Y\,dP \qquad \text{by def of } E[\cdot|\mathcal{G}] \\
&= \int_G (aX+bY)\,dP \\
&= \int_G E[aX+bY|\mathcal{G}]\,dP
\end{aligned}
$$

for all G in $\mathcal{G}$.

Hence $E[aX+bY|\mathcal{G}]=aE[X|\mathcal{G}]+bE[Y|\mathcal{G}]$ with probability 1.

(next pg. cont.)

(iii) If $X \le Y$ with probability 1, then

$$\int_G (E[X|\mathcal{G}] - E[Y|\mathcal{G}])\,dP$$
$$= \int_G (X-Y)\,dP \le 0 \text{ for all } G \text{ in } \mathcal{G}.$$

Since $E[X|\mathcal{G}] - E[Y|\mathcal{G}]$ is measurable $\mathcal{G}$,

$E[X|\mathcal{G}] - E[Y|\mathcal{G}] \le 0$ with probability 1,

This holds because if we consider G' where $E[X|\mathcal{G}] - E[Y|\mathcal{G}] > 0$,

then $\int_{G'} (E[X|\mathcal{G}] - E[Y|\mathcal{G}])\,dP = 0$ implies that $P(G') = 0$.

(iv) Since $-|X| \le X \le |X|$, then

$-E[|X||\mathcal{G}] \le E[X|\mathcal{G}] \le E[|X||\mathcal{G}]$ with probability 1,

so $|E[X|\mathcal{G}]| \le E[|X||\mathcal{G}]$ with probability 1.

(v) $\because \lim_n X_n = X$ with probability 1,

$\therefore P[\lim_n \sup_{k\ge n} |X_k - X| = 0] = 1$, that is,

Consider $Z_n = \sup_{k\ge n} |X_k - X|$, now $Z_n \downarrow 0$ with probability 1.

- Then by (ii), (iii), (iv),

$$|E[X_n|\mathcal{G}] - E[X|\mathcal{G}]|$$
$$= |E[X_n - X|\mathcal{G}]| \le E[|X_n - X||\mathcal{G}]$$
$$\le E[\sup_{k\ge n} |X_k - X||\mathcal{G}] = E[Z_n|\mathcal{G}] \quad \because |X_n - X| \le \sup_{k\ge n} |X_k - X|$$

- It suffices to show that $E[Z_n|\mathcal{G}] \downarrow 0$ with probability 1

$\because$ the sequence $Z_n = \sup_{k\ge n} |X_k - X|$ is nonincreasing,

$\therefore$ by (iii) the sequence $E[Z_n|\mathcal{G}]$ is nonincreasing, a.s.,

and hence has a limit Z.

- The problem is to prove $Z = 0$ with probability 1, or,

Z being nonnegative, that $E[Z] = 0$.

$\because 0 \le Z_n = \sup_{k\ge n} |X_k - X| \le \sup_{k\ge n} (|X_k| + |X|) = 2Y \quad \because |X| = \lim_n |X_n| \le Y$ a.s.

$\therefore$ by the defining relation and the DCT,

$$E[Z] = \int E[Z|\mathcal{G}]\,dP \le \int E[Z_n|\mathcal{G}]\,dP = E[Z_n] \to 0.$$

#

- <u>Theorem</u> (Product Rule of Conditional Expectation)

  If X is measurable $\mathcal{G}$, and if Y and XY are integrable, then

  $$E[XY|\mathcal{G}] = X\,E[Y|\mathcal{G}] \text{ with probability } 1$$

  Remark:
  - For an observer with information in $\mathcal{G}$, X is effectively a constant if it is measurable $\mathcal{G}$.
  - X has not been assumed integrable.

  pf: • First suppose $X = I_{G_0}$ and $G_0 \in \mathcal{G}$.

  $\because$ $I_{G_0}E[Y|\mathcal{G}]$ is measurable $\mathcal{G}$; and.

  $$\int_G I_{G_0}E[Y|\mathcal{G}]\,dP$$
  $$= \int_{G\cap G_0} E[Y|\mathcal{G}]\,dP = \int_{G\cap G_0} Y\,dP$$
  $$= \int_G I_{G_0}Y\,dP = \int_G E[I_{G_0}Y|\mathcal{G}]\,dP.$$

  $\therefore$ the defining relation holds if X is the indicator of the element of $\mathcal{G}$.

  • It follows that the defining relation holds if X is a simple function measurable $\mathcal{G}$.

  • For the general X that is measurable $\mathcal{G}$, there exists simple functions X_n, measurable $\mathcal{G}$, such that $|X_n| \le |X|$ and $\lim_n X_n = X$.

  Since $|X_nY| \le |XY|$ and $|XY|$ is measurable, the dominated convergence thm implies that

  $$\lim_n E[X_nY|\mathcal{G}] = E[XY|\mathcal{G}] \text{ with probability } 1.$$

  • $\because$ $E[X_nY|\mathcal{G}] = X_n E[Y|\mathcal{G}]$ since X_n is simple functions measurable $\mathcal{G}$,

  $\therefore$ $\lim_n X_n E[Y|\mathcal{G}] = X E[Y|\mathcal{G}]$

  • Note that $|X_n E[Y|\mathcal{G}]| = |E[X_nY|\mathcal{G}]| \le E[|X_nY|\,|\mathcal{G}] \le E[|XY|\,|\mathcal{G}]$, so that the limit $XE[Y|\mathcal{G}]$ is integrable, thus $E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$ with probability 1 in general.

  #

- <u>Theorem</u> (Smoothing of Conditional Expectation)

If X is integrable and the σ-fields $\mathcal{G}_1$ and $\mathcal{G}_2$ satisfy $\mathcal{G}_1 \subset \mathcal{G}_2$, then

$$E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1] = E[E[X|\mathcal{G}_1]|\mathcal{G}_2].$$

with probability 1.

Remark:

- Taking expectation can be thought of as an averaging or smoothing operation.
- If $\mathcal{G}_1 \subset \mathcal{G}_2$, taking iterated expectation in either order gives $E[X|\mathcal{G}_1]$.
- If $\mathcal{G}_2 = \mathcal{F}$, then $E[X|\mathcal{G}_2] = X$, the result is trivial;

  If $\mathcal{G}_1 = \{\emptyset, \Omega\}$ and $\mathcal{G}_2 = \mathcal{G}$, then $E[E[X|\mathcal{G}]] = E[X]$.

Pf: (1) We first show $E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1]$ with probability 1.

$\because$ $E[E[X|\mathcal{G}_2]|\mathcal{G}_1]$ is measurable $\mathcal{G}_1$, and

since $E[X|\mathcal{G}_2]$ is integrable, so is $E[E[X|\mathcal{G}_2]|\mathcal{G}_1]$; and

for $G \in \mathcal{G}_1$, then $G \in \mathcal{G}_2$,

$$\int_G E[E[X|\mathcal{G}_2]|\mathcal{G}_1]\,dP$$

$$= \int_G E[X|\mathcal{G}_2]\,dP = \int_G X\,dP = \int_G E[X|\mathcal{G}_1]\,dP.$$

$\therefore$ the result follows.

(2) We then show $E[E[X|\mathcal{G}_1]|\mathcal{G}_2] = E[X|\mathcal{G}_1]$.

Since $\mathcal{G}_1 \subset \mathcal{G}_2$, $E[X|\mathcal{G}_1]$, being measurable $\mathcal{G}_1$, is also measurable $\mathcal{G}_2$,

so $E[E[X|\mathcal{G}_1]|\mathcal{G}_2] = E[X|\mathcal{G}_1]$.

#

- <u>Theorem</u> (Jensen's Inequality for Conditional Expectation)

If φ is convex function on the line and "X and $\varphi(X)$ are both integrable", then

$$\varphi(E[X|\mathcal{G}]) \le E[\varphi(X)|\mathcal{G}] \text{ with probability } 1.$$

Pf: ◦ For each x_0 take a support line through $(x_0, \varphi(x_0))$:

$$\varphi(x_0) + A(x_0)(x - x_0) \le \varphi(x).$$

The slope $A(x_0)$ can be taken as the right-hand derivative of φ, so that it is nondecreasing in x_0. Now

◦ $$\varphi(E[X|\mathcal{G}]) + A(E[X|\mathcal{G}])(X - E[X|\mathcal{G}]) \le \varphi(X)$$

• Suppose that $E[X|\mathcal{G}]$ is bounded.

Since φ is convex on R^1, then φ and A are bounded on bounded sets.

• then all three terms here are integrable, so taking conditional expectations

$$E[\varphi(E[X|\mathcal{G}])|\mathcal{G}] + E[A(E[X|\mathcal{G}])(X - E[X|\mathcal{G}])|\mathcal{G}]$$

$$= \varphi(E[X|\mathcal{G}]) + A(E[X|\mathcal{G}])(E[X|\mathcal{G}] - E[X|\mathcal{G}])$$

$$= \varphi(E[X|\mathcal{G}]) \le E[\varphi(X)|\mathcal{G}]. \quad (1)$$

• To prove the result in general,

let $G_n = [|E[X|\mathcal{G}]| \le n]$.

Then $E[I_{G_n} X|\mathcal{G}] = I_{G_n} E[X|\mathcal{G}]$ is bounded, and so (1) holds for $I_{G_n} X$:

• $$\begin{aligned} \varphi(I_{G_n} E[X|\mathcal{G}]) &\le E[\varphi(I_{G_n} X)|\mathcal{G}] \\ &= E[\varphi(I_{G_n} X + I_{G_n^c} \cdot 0)|\mathcal{G}] \\ &\le E[I_{G_n}\varphi(X) + I_{G_n^c}\varphi(0)|\mathcal{G}] \quad \because \varphi \text{ is convex} \\ &= I_{G_n} E[\varphi(X)|\mathcal{G}] + I_{G_n^c}\varphi(0) \\ &\to E[\varphi(X)|\mathcal{G}] \text{ as } n \to \infty \end{aligned}$$

• Since $\varphi(I_{G_n} E[X|\mathcal{G}]) \to \varphi(E[X|\mathcal{G}])$ as $n \to \infty$ by continuity of φ, the result follows. #

- <u>Definition</u> (Martingales)

Let $X_1, X_2, \cdots$ be a sequence of r.v.'s on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_1, \mathcal{F}_2, \cdots$ be a sequence of σ-fields in $\mathcal{F}$.

The sequence $\{(X_n, \mathcal{F}_n) : n = 1, 2, \cdots\}$ is a "martingale" if

(i) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$;

(ii) X_n is measurable $\mathcal{F}_n$;

(iii) $E[|X_n|] < \infty$.

(iv) with probability 1, $E[X_{n+1} | \mathcal{F}_n] = X_n$.

Remark:

- Interpretations:

  (i): the $\mathcal{F}_n$ form a filtration

  (ii): the X_n are adapted to the filtration

  (iii): ensure that $E[X_{n+1} | \mathcal{F}_n]$ exists.

  (iv): X_n is a version of $E[X_{n+1} | \mathcal{F}_n]$.

- A martingale represents a "fair game":

  If X_n represents the fortune of a gambler after the nth play, and $\mathcal{F}_n$ represents his information about the game at the nth time, then his expected fortune after the next play is the same as his present fortune.

- The definition makes sense if n ranges over $1, \dots, N$:

  (ii), (iii) are required for $1 \le n \le N$;

  (i), (iv) are required for $1 \le n < N$.

- The definition makes sense if n ranges over an arbitrary ordered set.

- <u>Proposition</u> (Some Filtration implies $\mathcal{G}_n = \sigma(X_1, \cdots, X_n)$ always works)

If the sequence $X_1, X_2, \cdots$ is a martingale relative to some sequence $\mathcal{F}_1, \mathcal{F}_2, \cdots$; then the sequence of σ-fields $\mathcal{G}_n = \sigma(X_1, \cdots, X_n)$ always work.

Hence $E[X_{n+1} | \mathcal{F}_n]$ reduces to $E[X_{n+1} | X_1, \cdots, X_n]$.

Pf: Obviously, $\mathcal{G}_n \subset \mathcal{G}_{n+1}$ and X_n is measurable $\mathcal{G}_n$;

$E[|X_n|] < \infty$ by hypothesis;

Since $\sigma(X_1, \cdots, X_n) \subset \mathcal{F}_n$ iff X_n is measurable $\mathcal{F}_n$ for each n, the $\sigma(X_1, \cdots, X_n)$ are the smallest σ-fields.

$$\begin{aligned} E[X_{n+1} | \mathcal{G}_n] &= E[E[X_{n+1} | \mathcal{F}_n] | \mathcal{G}_n] \\ &= E[X_n | \mathcal{G}_n] = X_n. \end{aligned}$$

Hence $\{(X_n, \mathcal{G}_n), n=1,2,\cdots\}$ is a martingale.

#

- <u>Proposition</u> (Martingales have constant expections)

If $\{(X_n, \mathcal{F}_n) : n=1,2,\cdots\}$ is a martingale, then

(i) For $A \in \mathcal{F}_n$,

$$\int_A X_{n+1} \, dP = \int_A X_n \, dP$$

(ii) $E[X_{n+k} | \mathcal{F}_n] = X_n$ with probability 1 for $k \geq 1$.

(iii) $E[X_1] = E[X_2] = \cdots$, that is, the X_n have constant expectations.

Pf: (i) For $A \in \mathcal{F}_n$, X_n is a version of $E[X_{n+1} | \mathcal{F}_n]$,

$$\int_A X_n \, dP = \int_A E[X_{n+1} | \mathcal{F}_n] \, dP = \int_A X_{n+1} \, dP \qquad (1)$$

$\because$ $E[X_{n+1} | \mathcal{F}_n]$ is measurable $\mathcal{F}_n$, and hence measurable $\mathcal{F}_{n+1}$.

(ii) Since the $\mathcal{F}_n$ are nested, $A \in \mathcal{F}_n$ implies that

$$\int_A X_n \, dP = \int_A X_{n+1} \, dP = \cdots = \int_A X_{n+k} \, dP$$

Therefore, X_n, being measurable $\mathcal{F}_n$, is a version of $E[X_{n+k} | \mathcal{F}_n]$.

(iii) For $A = \Omega$, (1) gives

$E[X_n] = E[X_{n+1}] = \cdots$.

#

- **Definition** (Martingale Differences)

Suppose $\{(X_n, \mathcal{F}_n) : n=1,2,\cdots\}$ is a martingale.

Let the "martingale difference" be

$$\Delta_n = X_n - X_{n-1} \quad (\Delta_1 = X_1).$$

So $E[X_{n+1} | \mathcal{F}_n] = X_n$ is the same thing as

$E[X_{n+1} | \mathcal{F}_n] - E[X_n | \mathcal{F}_n] = 0$, $\because X_n$ is measurable $\mathcal{F}_n$

$E[X_{n+1} - X_n | \mathcal{F}_n] = 0$, $\because$ by linearity of conditional expectation.

$\Rightarrow E[\Delta_{n+1} | \mathcal{F}_n] = 0$.

Remark:

Since $X_n = \Delta_1 + \cdots + \Delta_k$ and $\Delta_k = X_k - X_{k-1}$,

the sets $X_1, \cdots, X_n$ and $\Delta_1, \cdots, \Delta_k$ generate the same σ-field:

$$\sigma(X_1, \cdots, X_n) = \sigma(\Delta_1, \cdots, \Delta_k).$$

- **Example** (Sums of (Indep) r.v.'s with Mean 0 is a Martingale)

① Let $\Delta_1, \Delta_2, \cdots$ be independent, integrable r.v.'s s.t. $E[\Delta_n] = 0$ for $n \geq 2$.

If $\mathcal{F}_n$ is the σ-field $\sigma(\Delta_1, \cdots, \Delta_n)$, then

$E[\Delta_{n+1} | \mathcal{F}_n] = E[\Delta_{n+1}] = 0$ $\because$ by independence of Δ_n

So let $X_n = \Delta_1 + \cdots + \Delta_n$, $\mathcal{F}_n = \sigma(\Delta_1, \cdots, \Delta_n) = \sigma(X_1, \cdots, X_n)$,

we know that $\{(X_n, \mathcal{F}_n) : n=1,2,\cdots\}$ is a martingale.

② If Δ is another r.v. independent of the Δ_n, and

if $\mathcal{F}_n$ is replaced by $\sigma(\Delta, \Delta_1, \cdots, \Delta_n)$, then

$X_n = \Delta_1 + \cdots + \Delta_n$ are still martingale relative to $\mathcal{F}_n$.

That is, it is natural and convenient to allow

σ-field $\mathcal{F}_n$ larger than the minimum one.

③ Method of centering by conditional means

Let $\{\xi_n\}$ be arbitrary sequence of integrable r.v.'s. Define $\mathcal{F}_n = \sigma(\xi_1, \cdots, \xi_n)$.

Then $\{\xi_n - E[\xi_n | \mathcal{F}_{n-1}]\}$ is a fair sequence since $E[\xi_n - E[\xi_n | \mathcal{F}_{n-1}] | \mathcal{F}_{n-1}] = 0$.

So the $X_n = \sum_{k=1}^{n} (\xi_k - E[\xi_k | \mathcal{F}_{k-1}])$ is a martingale. #

- **Example** (Conditional Expectation is a Martingale)

Suppose that Z is an "integrable" r.v. on $(\Omega, \mathcal{F}, P)$, and that $\mathcal{F}_n$ are "nondecreasing σ-field's" in $\mathcal{F}$, i.e. $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$.

If $X_n = E[Z \mid \mathcal{F}_n]$, then the first three conditions are satisfied, and

$$E[X_{n+1} \mid \mathcal{F}_n] = E\big[E[Z \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n\big] = E[Z \mid \mathcal{F}_n] = X_n.$$

Thus $X_n = E[Z \mid \mathcal{F}_n]$ is a martingale relative to $\mathcal{F}_n$. #

- **Example** (Products of Indep. r.v.'s with Mean 1 is a Martingale)

① Let $Y_1, Y_2, \ldots$ be independent, integrable r.v.'s s.t. $E[Y_n] = 1$.

If $X_n = \prod_{i=1}^{n} Y_i$ and $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n) = \sigma(X_1, \ldots, X_n)$, then the first three conditions are satisfied, and

$$E[X_{n+1} \mid \mathcal{F}_n] = E[X_n Y_{n+1} \mid \mathcal{F}_n] = X_n E[Y_{n+1} \mid \mathcal{F}_n] = X_n E[Y_{n+1}] = X_n$$

Thus $X_n = \prod_{i=1}^{n} Y_i$ is a martingale relative to $\mathcal{F}_n$.

② Generating functions, Laplace transforms, chf's etc.

Let the Z_n be iid r.v.'s with range $\{0, 1, 2, \ldots\}$ and we use the generating function as our typical transform.

Define $\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n)$, $n \geq 1$, and let the generating function of the Z be $\phi(s) = E s^{Z_1}$, $0 \leq s \leq 1$.

Set $S_n = \sum_{i=1}^{n} Z_i$, and $M_n = \frac{s^{S_n}}{\phi^n(s)} = \prod_{i=1}^{n} \frac{s^{Z_i}}{\phi(s)}$, where $E\left[\frac{s^{Z_i}}{\phi(s)}\right] = 1$,

then $\{(M_n, \mathcal{F}_n) : n = 1, 2, \ldots\}$ is a martingale.

③ Likelihood ratios

Suppose the Y_n are iid r.v.'s and suppose the true density of Y_1 is f_0.

Let f_1 be some other probability density.

For simplicity suppose $f_0(y) > 0$ for all y, then

$$X_n = \frac{\prod_{i=1}^{n} f_1(Y_i)}{\prod_{i=1}^{n} f_0(Y_i)} \text{ is a martingale since}$$

$$E\left[\frac{f_1(Y_n)}{f_0(Y_n)}\right] = \int \frac{f_1(y_n)}{\cancel{f_0(y_n)}} \cancel{f_0(y_n)}\, \mu(dy_n) = \int f_1(y_n)\, \mu(dy_n) = 1.$$

#

- <u>Example</u> (Densities or Radon-Nikodym Derivatives are a Martingale)

Let ν be a finite measure on $\mathcal{F}$, and
let $\mathcal{F}_1, \mathcal{F}_2, \dots$ be a nondecreasing sequence of σ-fields in $\mathcal{F}$.
Suppose P dominates ν when both are restricted to $\mathcal{F}_n$ — that is,
suppose that $A \in \mathcal{F}_n$ and $P(A)=0$ together implies $\nu(A)=0$.
There is then a density or Radon-Nikodym derivative X_n of ν w.r.t. P
when both restricted to $\mathcal{F}_n$.
X_n is measurable $\mathcal{F}_n$ and is integrable w.r.t. P since ν is finite,
and it satisfies

$$\int_A X_n \, dP = \nu(A), \quad A \in \mathcal{F}_n.$$

If $A \in \mathcal{F}_n$, then $A \in \mathcal{F}_{n+1}$ as well, so that

$$\int_A X_{n+1} \, dP = \nu(A), \text{ then}$$

$$\int_A X_{n+1} \, dP = \int_A X_n \, dP, \quad A \in \mathcal{F}_n.$$

Thus the densities or Radon-Nikodym derivatives X_n are a martingale.

#

• <u>Example</u> (Branching Process $\{Z_n\}$ v.s. Martingale $\{Z_n/m^n\}$)

Suppose $Z_0, Z_1, \cdots$ is a branching process — that is,

define $Z_0(\omega)=1$, $Z_n(\omega) = N_{n,1}(\omega) + \cdots + N_{n,Z_{n-1}(\omega)}(\omega)$;

$Z_n(\omega) = 0$ if $Z_{n-1}(\omega) = 0$;

where N_{nk}, $n, k = 1, 2, \cdots$ is an "independent" array of "identically distributed" r.v.'s assuming the values of $1, 2, \cdots$.

with $E[N_{nk}] = m$. Set $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$, then

$$E[Z_n \mid \mathcal{F}_{n-1}] = E[Z_n \mid Z_{n-1}] = Z_{n-1} m.$$

$\Rightarrow$ So that $X_n = Z_n/m^n$, $n = 0, 1, 2, \dots$, is a martingale since

$$E[X_n \mid \mathcal{F}_{n-1}] = E[Z_n/m^n \mid Z_{n-1}] = Z_{n-1} m / m^n = Z_{n-1}/m^{n-1} = X_{n-1}.$$

#

- Definition (Submartingales and Supermartingales)

(1) Random variables X_n are a "submartingale" relative to σ-fields $\mathcal{F}_n$ if (i), (ii), (iii) of the definition of martingale hold, and (iv) with probability 1, $E[X_{n+1} | \mathcal{F}_n] \geq X_n$.

Remark:

- If the X_n are a submartingale relative some $\mathcal{F}_n$, then the special sequence $\mathcal{G}_n = \sigma(X_1, \cdots, X_n)$ always work; hence $E[X_{n+1} | \mathcal{F}_n]$ implies $E[X_{n+1} | X_1, \cdots, X_n]$.
- If $\{(X_n, \mathcal{F}_n) : n = 1, 2, \cdots\}$ is a submartingale, then
  ① For $A \in \mathcal{F}_n$,
  $$\int_A X_{n+1} \, dP \geq \int_A X_n \, dP$$
  ② $E[X_{n+k} | \mathcal{F}_n] \geq X_n$ with probability 1 for $k \geq 1$
  ③ $E[X_1] \leq E[X_2] \leq \cdots$

(2) Random variables X_n are a "supermartingale" relative to σ-fields $\mathcal{F}_n$ if (i), (ii), (iii) of the definition of martingale hold, and the inequality in (iv) becomes reversed; and hence the inequalities in ①, ②, ③ become reversed as well.

(3) Martingale can be viewed as a submartingale or a supermartingale.

• <u>Example</u> (Sums of Indep r.v.'s with mean ≥ 0 is a Submartingale)

Suppose that the Δ_n are independent and integrable, but assume $E[\Delta_n] \geq 0$ for $n \geq 2$.

If $\mathcal{F}_n$ is the σ-field $\sigma(\Delta_1, \cdots, \Delta_n)$, then

$$E[\Delta_{n+1} \| \mathcal{F}_n] = E[\Delta_{n+1}] \geq 0 \quad \because \text{ by independence of } \Delta_n$$

So the partial sum $X_n = \Delta_1 + \cdots + \Delta_n$ form a submartingale.

#

• <u>Example</u> (Martingale X_n implies Submartingale $|X_n|$)

Suppose that the X_n are a martingale relative to the $\mathcal{F}_n$.

Then $|X_n|$ is measurable $\mathcal{F}_n$ and integrable. Since

$$E[|X_{n+1}| \| \mathcal{F}_n] \geq |E[X_{n+1} | \mathcal{F}_n]| = |X_n|.$$

Thus the $|X_n|$ are submartingale relative to $\mathcal{F}_n$.

#

- <u>Example</u> (Martingale Transforms)
- Suppose the X_n is a martingale
  Start it at $n=0$, X_0 = the gambler's initial fortune.
  The difference $\Delta_n = X_n - X_{n-1}$ = the amount the gambler wins on the nth play.
- Suppose instead that Δ_n = the amount he wins if he puts up unit stakes.
  If instead of unit stakes he wagers the amount W_n on the nth play,
  $W_n \Delta_n$ = his gain on that play, where $W_n \geq 0$.
- "W_n is measurable $\mathcal{F}_{n-1}$" to exclude prevision:
  Before the nth play the information available to the gambler is that in $\mathcal{F}_{n-1}$.
- Suppose W_n is "bounded", then $W_n \Delta_n$ is integrable, and measurable $\mathcal{F}_n$ if Δ_n is.
  Hence $E[W_n \Delta_n | \mathcal{F}_{n-1}] = W_n E[\Delta_n | \mathcal{F}_{n-1}] = 0$, thus
  $X_0 + W_1 \Delta_1 + \cdots + W_n \Delta_n$ is a martingale relative to $\mathcal{F}_n$.
  That is, transforming $X_n = X_0 + \Delta_1 + \cdots + \Delta_n$ into $X_0 + W_1 \Delta_1 + \cdots + W_n \Delta_n$ preserves the martingale property.
- Similarly, if X_n is a supermartingale or submartingale, the same argument shows that $X_0 + W_1 \Delta_1 + \cdots + W_n \Delta_n$ is again a supermartingale or submartingale.

#

• Definition (Stopping Times)

• Let τ be a r.v. taking on "nonnegative integers" as values.

τ is a "stopping time" w.r.t. $\{\mathcal{F}_n\}$ if

$[\tau = k] \in \mathcal{F}_k$ for each finite k; or

$[\tau \le k] \in \mathcal{F}_k$ for each finite k.

• Define "pre-τ σ-field"

$$\mathcal{F}_\tau = [A \in \mathcal{F} : A \cap [\tau \le k] \in \mathcal{F}_k,\ 1 \le k < \infty]$$
$$= [A \in \mathcal{F} : A \cap [\tau = k] \in \mathcal{F}_k,\ 1 \le k < \infty]$$

1. This is a σ-field.
2. Since $[\tau = j] \in \mathcal{F}_\tau$ for finite j, τ is measurable $\mathcal{F}_\tau$.
3. If $\tau(\omega) < \infty$ for all ω and $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$,

   the information in $\mathcal{F}_\tau$ consists of the values $\tau(\omega), X_1(\omega), \dots, X_{\tau(\omega)}(\omega)$.

   pf: $I_A(\omega) = I_A(\omega')$ for all A in $\mathcal{F}_\tau$ iff $X_i(\omega) = X_i(\omega')$ for $i \le \tau(\omega) = \tau(\omega')$. #

• **Theorem** (Stopped Martingales / Supermartingales)

Let τ be a stopping time w.r.t $\{\mathcal{F}_n\}$ if $[\tau=k]\in\mathcal{F}_k$.

The gambler's fortune at time n for this stopping rule is

$$X_n^* = \begin{cases} X_n & \text{if } n\le\tau \\ X_\tau & \text{if } n\ge\tau \end{cases} \Rightarrow X_n^* = X_n 1_{[\tau>n]} + X_\tau 1_{[\tau\le n]}$$

Here X_τ (which has value $X_{\tau(\omega)}(\omega)$ at ω) is his "ultimate fortune".

Suppose the X_n is a martingale relative to the $\mathcal{F}_n$, then

the X_n^* is a martingale relative to the $\mathcal{F}_n$;

PS. The same kind of argument works for "supermartingales".

Pf: (i) $\mathcal{F}_n\subset\mathcal{F}_{n+1}$ by hypothesis.

(ii) Since $[\tau>n]=[\tau\le n]^c\in\mathcal{F}_n$,

$$[X_n^*\in H] = \bigcup_{k=0}^{n}[\tau=k, X_k\in H]\cup[\tau>n, X_n\in H]\in\mathcal{F}_n,$$

hence X_n^* is measurable $\mathcal{F}_n$.

(iii)
$$\begin{aligned} E[|X_n^*|] &= \int |X_n^*|\,dP \\ &= \int_{[\tau<n]}|X_\tau|\,dP + \int_{[\tau\ge n]}|X_n|\,dP \\ &= \sum_{k=0}^{n-1}\int_{[\tau=k]}|X_k|\,dP + \int_{[\tau\ge n]}|X_n|\,dP \\ &\le \sum_{k=0}^{n}\int|X_k|\,dP = \sum_{k=0}^{n}E[|X_k|]<\infty \text{ by hypothesis.} \end{aligned}$$

(iv)
$$\int_A X_n^*\,dP = \int_{A\cap[\tau\le n]}X_\tau\,dP + \int_{A\cap[\tau>n]}X_n\,dP$$

$$\begin{aligned} \int_A X_{n+1}^*\,dP &= \int_{A\cap[\tau\le n+1]}X_\tau\,dP + \int_{A\cap[\tau>n+1]}X_{n+1}\,dP \\ &= \int_{A\cap[\tau\le n]}X_\tau\,dP + \int_{A\cap[\tau=n+1]}X_{n+1}\,dP + \int_{A\cap[\tau>n+1]}X_{n+1}\,dP \\ &= \int_{A\cap[\tau\le n]}X_\tau\,dP + \int_{A\cap[\tau>n]}X_{n+1}\,dP \end{aligned}$$

Since for $A\in\mathcal{F}_n$, $\int_A X_n\,dP = \int_A X_{n+1}\,dP$, hence

$\int_A X_n^*\,dP = \int_A X_{n+1}^*\,dP$ for $A\in\mathcal{F}_n$.

The result follows.

#

- <u>Proposition</u> (Stopped Martingales, Moment Convergence)

Suppose the X_n is a martingale, and let τ be a stopping time.

Define the "stopped martingale" is

$$X_n^* = \begin{cases} X_n & \text{if } n \le \tau \\ X_\tau & \text{if } n \ge \tau \end{cases}$$

- Since $X_n^* = X_\tau$ for $n \ge \tau$, $X_n^* \to X_\tau$.
- It is not always possible to integrate to the limit:

  Let $X_n = a + \Delta_1 + \cdots + \Delta_n$, $(X_0 = a)$,

  where the Δ_n are iid with $\Delta_n = \pm 1$ w.p. $\frac{1}{2}$. i.e. X_n is a random walk.

  If τ is the smallest n for which $\Delta_1 + \cdots + \Delta_n = 1$, then

  $X_\tau = a+1$, so $E[X_\tau] = a+1$;

  $E[X_n^*] = E[X_0] = a$ since the X_n^* is a martingale with constant mean.

  Note. $E[X_n^*] \to a \neq E[X_\tau] = a+1$.
- If the X_n are "uniformly bounded" or "uniformly integrable",

  then the X_n^* are uniformly bounded or uniformly integrable, too,

  hence by $X_n^* \to X_\tau$, we have $E[X_n^*] \to E[X_\tau]$, moments convergent.

  Since the X_n^* is a martingale, we have

  $E[X_n^*] = E[X_0]$ for all n, and hence $E[X_0] = E[X_\tau]$.

#

- <u>Theorem</u> (The Martingale Stopping Thm)

If either:

(i) X_n^* are "uniformly bounded" or "uniformly integrable", or;

(ii) τ is bounded, or;

(iii) $E[\tau] < \infty$, and there is an $M < \infty$ s.t.

$$E[|X_{n+1} - X_n| \mid X_1, \ldots, X_n] < M, \text{ i.e. } E[|\Delta_{n+1}| \mid \mathcal{F}_n] < M$$

then $E[X_n^*] \to E[X_\tau]$ as $n \to \infty$. i.e. the moments converge.

Hence $E[X_0] = E[X_\tau]$ by $\{X_n^*\}$ is a martingale, so $E[X_n^*] = E[X_0]$

• <u>Theorem</u> (Convex Functions of Martingales/Submartingales)

(i) If the X_n is a "martingale" relative to the $\mathcal{F}_n$,
if φ is convex, and if "the $\varphi(X_n)$ is integrable",
then the $\varphi(X_n)$ is a submartingale relative to the $\mathcal{F}_n$.

(ii) If the X_n is a submartingale relative to the $\mathcal{F}_n$,
if φ is "nondecreasing" and convex, and if "the $\varphi(X_n)$ is integrable",
then the $\varphi(X_n)$ is a submartingale relative to the $\mathcal{F}_n$.

Pf: (i) In the martingale case,

$X_n = E[X_{n+1} \mid \mathcal{F}_n]$, then

$$\varphi(X_n) = \varphi(E[X_{n+1} \mid \mathcal{F}_n])$$
$$\leq E[\varphi(X_{n+1}) \mid \mathcal{F}_n] \quad \because \varphi \text{ is convex.}$$

Hence the $\varphi(X_n)$ is a submartingale.

(ii) In the submartingale case,

$$X_n \leq E[X_{n+1} \mid \mathcal{F}_n]$$
$$\varphi(X_n) \leq \varphi(E[X_{n+1} \mid \mathcal{F}_n]) \quad \because \varphi \text{ is nondecreasing.}$$
$$\leq E[\varphi(X_{n+1}) \mid \mathcal{F}_n] \quad \because \varphi \text{ is convex.}$$

Hence the $\varphi(X_n)$ is a submartingale.

#.

- **Theorem** (Optimale Sampling Thm)

If the X_n is a "submartingale" w.r.t. the $\mathcal{F}_n$, and τ_1, τ_2 are stopping times satisfying "$1 \le \tau_1 \le \tau_2 \le n$", then X_{τ_1}, X_{τ_2} is a "submartingale" w.r.t. "$\mathcal{F}_{\tau_1}, \mathcal{F}_{\tau_2}$". i.e. $E[X_{\tau_1}] \le E[X_{\tau_2}]$

P.S. X_{τ_1}, X_{τ_2} is a martingale if the X_n is. i.e. $E[X_{\tau_1}] = E[X_{\tau_2}]$.

Pf: • Define τ_i is a stopping time if $[\tau_i \le k] \in \mathcal{F}_k$

• Define pre-τ_i σ-fields $\mathcal{F}_{\tau_i}$ are, $i=1,2$,

$$\mathcal{F}_{\tau_i} = \{A \in \mathcal{F} : A \cap [\tau_i \le k] \in \mathcal{F}_k,\ 1 \le k < \infty\}$$

(i) $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$:

If $A \in \mathcal{F}_{\tau_1}$, then $A \cap [\tau_1 \le k] \in \mathcal{F}_k$, and

$A \cap [\tau_2 \le k] = A \cap [\tau_1 \le k] \cap [\tau_2 \le k] \in \mathcal{F}_k$, $\because \tau_1 \le \tau_2$

hence $A \in \mathcal{F}_{\tau_2}$.

* (ii) X_{τ_i} is measurable $\mathcal{F}_{\tau_i}$:

Since $\tau_i(\omega) \le n < \infty$, $\mathcal{F}_n \equiv \sigma(X_1, \dots, X_n)$, then

$I_A(\omega) = I_A(\omega')$ for all A in $\mathcal{F}_{\tau_i}$ iff $X_i(\omega) = X_i(\omega')$ for $i \le \tau(\omega) = \tau(\omega')$:

The information in $\mathcal{F}_{\tau_i}$ consists of the values $\tau_i(\omega), X_1(\omega), \dots, X_{\tau_i(\omega)}(\omega)$.

(iii) $E[|X_{\tau_i}|] < \infty$:

Since $X_n = X_{\tau_i}$ if $n \ge \tau_i$, $|X_{\tau_i}| \le \sum_{k=1}^{n} |X_k| \in L_1$ $\because X_n \in L_1$.

(iv) $E[X_{\tau_2} \mid \mathcal{F}_{\tau_1}] \ge X_{\tau_1}$:

It suffices to show that

$\int_A X_{\tau_2}\, dP \ge \int_A X_{\tau_1}\, dP$ for $A \in \mathcal{F}_{\tau_1}$. i.e. $\int_A (X_{\tau_2} - X_{\tau_1})\, dP \ge 0$, $A \in \mathcal{F}_{\tau_1}$.

But $A \in \mathcal{F}_{\tau_1}$ implies that

$A \cap [\tau_1 < k \le \tau_2] = A \cap [\tau_1 \le k-1] \cap [\tau_2 \le k-1]^c \in \mathcal{F}_{k-1}$.

If $\Delta_k = X_k - X_{k-1}$, then

$$\int_A (X_{\tau_2} - X_{\tau_1})\, dP = \int_A \sum_{k=1}^{n} I_{[\tau_1 < k \le \tau_2]} \Delta_k\, dP \qquad \text{by } \tau_i \le n,\ i=1,2.$$

$$= \sum_{k=1}^{n} \int_{A \cap [\tau_1 < k \le \tau_2]} \Delta_k\, dP \ge 0$$

by the submartingale property. #

• Theorem (Extention of Kolmogorov's Inequality)

If $X_1, \dots, X_n$ is a "submartingale", then for $\alpha > 0$,

$$P\left[\max_{i\le n} X_i \ge \alpha\right] \le \frac{1}{\alpha} E[|X_n|].$$

Remark:

this extends Kolmogorov's inequality:

If $S_1, S_2, \dots$ are partial sums of "indep" r.v.'s with "mean 0",

the S_n is a martingale;

If the variances are finite, then the S_n^2 is a submartingale,

the inequality for this submartingale S_n^2 is exactly

kolmogorov's inequality:

$$P\left[\max_{i\le n} S_i^2 \ge \alpha^2\right] = P\left[\max_{i\le n} |S_i| \ge \alpha\right] \le \frac{1}{\alpha^2} E[S_n^2] = \frac{1}{\alpha^2} Var[S_n].$$

Pf: • Let $\tau_2 = n$; let τ_1 be the smallest k s.t. $X_k \ge \alpha$, if there is one, and n otherwise.

• If $M_k = \max_{i\le k} X_i$, then

$[M_n \ge \alpha] \cap [\tau_1 \le k] = [M_k \ge \alpha] \in \mathcal{F}_k$, and hence $[M_n \ge \alpha] \in \mathcal{F}_{\tau_1}$.

• $\alpha P[M_n \ge \alpha]$

$\le \int_{[M_n \ge \alpha]} X_{\tau_1} \, dP \quad \because X_{\tau_1} \ge \alpha$

$\le \int_{[M_n \ge \alpha]} X_n \, dP \quad \because$ optional sampling thm, $\tau_1 \le \tau_2 = n$

$\le \int_{[M_n \ge \alpha]} X_n^+ \, dP \quad \because X_n = X_n^+ - X_n^- \ge X_n^+$

$\le E[X_n^+] \le E[|X_n|] \quad \because |X_n| = X_n^+ + X_n^- \ge X_n^+.$

#.

• Remark:

Improvement: If $X_1, \dots, X_n$ is a martingale,

then $|X_1|, \dots, |X_n|$ is a submartingale,

the inequality for this submartingale $|X_n|$ is

$$P\left[\max_{i\le n} |X_i| \ge \alpha\right] \le \frac{1}{\alpha} E[|X_n|].$$

- **Definition** (Upcrossing)

Let $[\alpha, \beta]$ be an interval ($\alpha < \beta$) and let $X_1, \dots, X_n$ be r.v.'s.
Inductively define r.v.'s $\tau_1, \tau_2, \dots$:

$$\tau_1 = \min\{j : 1 \le j \le n \text{ and } X_j \le \alpha, \text{ and is } n \text{ if there is no such } j\},$$

$$\tau_{2k} = \min\{j : \tau_{2k-1} < j \le n \text{ and } X_j \ge \beta, \text{ and is } n \text{ if there is no such } j\},$$

$$\tau_{2k-1} = \min\{j : \tau_{2k-2} < j \le n \text{ and } X_j \le \alpha, \text{ and is } n \text{ if there is no such } j\}.$$

For number U of upcrossings of $[\alpha, \beta]$ by $X_1, \dots, X_n$ is the largest i s.t. $X_{\tau_{2i-1}} \le \alpha < \beta \le X_{\tau_{2i}}$.

- Theorem (Upcrossing Inequality)

For a submartingale $X_1, \dots, X_n$, the number U of upcrossing of $[\alpha, \beta]$ satisfies

$$E[U] \le \frac{E[|X_n|] + |\alpha|}{\beta - \alpha}$$

Pf: • Let $Y_k = \max\{0, X_k - \alpha\}$ and $\theta = \beta - \alpha$, i.e. $Y_k = (X_k - \alpha) I_{[X_k > \alpha]}$.

Since $\max(\cdot)$ is nondecreasing and convex, and

$\max\{0, X_k - \alpha\} \le \max\{0, |X_k - \alpha|\} = |X_k - \alpha| \le |X_k| \in L^1$,

$Y_1, \dots, Y_n$ is a submartingale.

• The τ_k are unchanged if the definitions

$X_j \le \alpha$ is replaced by $Y_j = 0$ and $X_j \ge \beta$ by $Y_j \ge \theta$,

and so U is also the number of upcrossings of $[0, \theta]$ by $Y_1, \dots, Y_n$:

If k is even and τ_{k-1} is a stopping time, then for $j < n$,

$[\tau_k = j] = \bigcup_{i=1}^{j-1} [\tau_{k-1} = i,\ Y_{i+1} < \theta, \dots, Y_{j-1} < \theta,\ Y_j \ge \theta] \in \mathcal{F}_j$, and

$[\tau_k = n] = [\tau_k \le n-1]^c \in \mathcal{F}_n$,

and so τ_k is also a stopping time.

With a similar argument for odd k,

this shows that τ_k are all stopping times.

• Since the τ_k are strictly increasing until they reach n, $\tau_n = n$.

$Y_n = Y_{\tau_n} \ge Y_{\tau_n} - Y_{\tau_1}$

$= \sum_{k=2}^{n} (Y_{\tau_k} - Y_{\tau_{k-1}}) = \Sigma_e + \Sigma_o$

where Σ_e and Σ_o are the sums over the even k and the odd k.

• By the optimal sampling thm,

the Y_{τ_n} is a submartingale w.r.t. the $\mathcal{F}_{\tau_n}$,

$E[\Sigma_o] = \sum_{\text{odd } k} E[Y_{\tau_k} - Y_{\tau_{k-1}}] \ge 0$, hence $E[Y_n] \ge E[\Sigma_e]$.

• Note $Y_{\tau_{2i}} - Y_{\tau_{2i-1}}$ appears in the sum Σ_e, and

since $Y_{\tau_{2i-1}} = 0 < \theta \le Y_{\tau_{2i}}$, so $Y_{\tau_{2i}} - Y_{\tau_{2i-1}} \ge \theta$,

Since there are U of $(Y_{\tau_{2i}} - Y_{\tau_{2i-1}})$'s, $\Sigma_e \ge \theta U$, hence $E[Y_n] \ge E[\Sigma_e] \ge \theta E[U]$.

$(\beta - \alpha) E[U] \le \int_{[X_n > \alpha]} (X_n - \alpha)\, dP \le E[|X_n|] + |\alpha|$

since $X_n - \alpha \le |X_n - \alpha| \le |X_n| + |\alpha|$.

#

- <u>Theorem</u> (The Martingale Convergence Thm)

Let $X_1, X_2, \dots$ be a "submartingale". If "$K = \sup_n E[|X_n|] < \infty$",
then $X_n \to X$ with probability 1, where X is a r.v. with $E[|X|] \le K$.

Pf: Suppose the X_n is a submartingale,

- fix α and β for the moment, and
let U_n be the number of upcrossings of $[\alpha, \beta]$ by $X_1, \dots, X_n$.
By the upcrossing thm,

$$E[U_n] \le \frac{E[|X_n|] + |\alpha|}{\beta - \alpha} \le \frac{K + |\alpha|}{\beta - \alpha}.$$

Since U_n is nondecreasing and $E[U_n]$ is bounded,
it follows by the monotone convergence thm that

$$\lim_n E[U_n] = E[\lim_n U_n] = E[\sup_n U_n] < \infty,$$

So $\sup_n U_n$ is integrable and hence is finite with probability 1.

- Let $X_* = \liminf_n X_n$ and $X^* = \limsup_n X_n$; they may be infinite.
If $X_* < \alpha < \beta < X^*$, then U_n must go to ∞.
Since U_n is bounded with probability 1,

$$P[X_* < \alpha < \beta < X^*] = 0.$$

Now $[X_* < X^*] = \bigcup [X_* < \alpha < \beta < X^*]$
where the union extends all pairs of rationals α and β.
Thus $P[X_* < X^*] = 0$ and hence $X_* = X^*$ with probability 1,
and X_n converges to their common value X, which may be $\pm\infty$.

- By Fatou's lemma,

$$E[|X|] = E[\liminf_n |X_n|] \le \liminf_n E[|X_n|] \le K.$$

Then X is integrable and hence is finite with probability 1.

#

- <u>Example</u> (Martingale Convergence Thm: $K = \sup_n E[|X_n|] < \infty$ is essential)
  - If $X_n = \Delta_1 + \cdots + \Delta_n$ where the Δ_n are iid with $P[\Delta_n = \pm 1] = \frac{1}{2}$.
  - Since $\sum_n P[|\Delta_n| > \frac{1}{2}] = \sum_n 1 = \infty$,
    it follows by the three-series thm that
    $X_n = \sum_{k=1}^{n} \Delta_k$ does not converge a.s.
  - Since the Δ_n are iid with $E[\Delta_n] = 0$,
    $X_n = \sum_{k=1}^{n} \Delta_k$ is a martingale;
    $|X_n|$ is a submartingale with $E[|X_n|]$ nondecreasing. $\because |\cdot|$ is convex.
    So that $K = \sup_n E[|X_n|] = \lim_n E[|X_n|]$
  - Since the Δ_n are iid with $E[\Delta_n] = 0$ and $Var[\Delta_n] = 1$,
    it follows by the Lindeberg-Lévy thm that
    $\frac{X_n}{\sqrt{n}} \Rightarrow N$.
  - For show $\lim_n E[\frac{|X_n|}{\sqrt{n}}] = E[|N|]$.
    we need to show $\frac{X_n}{\sqrt{n}}$ is uniformly integrable first.

$$\int_{[\frac{|X_n|}{\sqrt{n}} > \alpha]} \frac{|X_n|}{\sqrt{n}} dP$$

$$= \frac{1}{\sqrt{n}} \int |X_n| I_{[\frac{|X_n|}{\sqrt{n}} > \alpha]} dP$$

$$\le \frac{1}{\sqrt{n}} \left(\int X_n^2 dP \cdot P[|X_n| > \alpha\sqrt{n}] \right)^{\frac{1}{2}} \quad \text{by Schwartz's inequality}$$

$$\le \frac{1}{\sqrt{n}} \left(n \cdot \frac{E[X_n^2]}{\alpha^2 n} \right)^{\frac{1}{2}} = \frac{1}{\sqrt{n}} \sqrt{\frac{n}{\alpha^2}} = \frac{1}{\alpha} \quad \text{by Markov's inequality}$$

$$\lim_{\alpha\to\infty} \sup_n \int_{[\frac{|X_n|}{\sqrt{n}} > \alpha]} \frac{|X_n|}{\sqrt{n}} dP \le \lim_{\alpha\to\infty} \sup_n \frac{1}{\alpha} = \lim_{\alpha\to\infty} \frac{1}{\alpha} = 0.$$

Hence $\lim_n E[\frac{|X_n|}{\sqrt{n}}] = E[|N|]$

$$= \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} x e^{-\frac{x^2}{2}} dx = \sqrt{\frac{2}{\pi}}.$$

  - Thus $E[|X_n|] \sim \sqrt{\frac{2}{\pi}} \sqrt{n}$ and hence $K = \lim_n E[|X_n|] = \infty$.

#

- <u>Example</u> (Convergence of Nonnegative Martingale)

- If the X_n form a "nonnegative martingale", then

  $E[|X_n|] = E[X_n] = E[X_1]$,

  So $k = \sup_n E[|X_n|] = E[X_1] < \infty$ always holds,

  thus the martingale convergence thm necessarily holds!

  $X_n \to X$ with probability 1,

  where X is a r.v. satisfying $E[|X|] < \infty$.

- But $E[X_n] \to E[X]$ may fail:

  Suppose $Z_0, Z_1, \ldots$ is a branching process,

  define $Z_0 = 1$, $Z_n(\omega) = N_{n,1} + \cdots + N_{n,Z_{n-1}}$;

  $Z_n = 0$ if $Z_{n-1} = 0$;

  where N_{nk}, $n, k = 1, 2, \ldots$ is an independent array

  of identically distributed r.v.'s assuming the values of $1, 2, \ldots$

  with $E[N_{nk}] = m$. Set $\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n)$, then

  the $X_n = Z_n / m^n$ is a martingale

- Since the X_n are nonnegative, with the arguments above,

  $X_n = Z_n / m^n \to X$, a.s.

  where X is nonnegative and integrable.

- If $m < 1$, then, since Z_n is an integer, $Z_n = 0$ for large n,

  and hence the population dies out.

  In this case, $X = 0$ with probability 1

  Since $E[X_n] = E[X_0] = 1$, this shows that

  $E[X_n] \to 1 \neq 0 = E[X]$.

#

• Lemma (The $E[Z|\mathcal{F}_n]$ are ui)

If Z is integrable and $\mathcal{F}_n$ are "arbitrary σ-fields" in $\mathcal{F}$, where $\mathcal{F}_n$ need not be nondecreasing.

Then the r.v.'s $E[Z|\mathcal{F}_n]$ are uniformly integrable.

Pf: • Since $|E[Z|\mathcal{F}_n]| \le E[|Z||\mathcal{F}_n]$,

Z may be assumed nonnegative.

(i) Let $A_{\alpha n} = [E[Z|\mathcal{F}_n] \ge \alpha]$. Since $A_{\alpha n} \in \mathcal{F}_n$,

$$\int_{A_{\alpha n}} E[Z|\mathcal{F}_n]\,dP = \int_{A_{\alpha n}} Z\,dP$$

Define $\nu(A) = \int_A Z\,dP$, then

$P(A)=0$ implies that $\nu(A)=0$.

Thus ν is absolutely continuous w.r.t. P; $\nu \ll P$,

so there is a δ s.t. $P(A)<\delta$ implies that $\nu(A) = \int_A Z\,dP < \epsilon$.

But $P[E[Z|\mathcal{F}_n] \ge \alpha]$

$\le \alpha^{-1} E[E[Z|\mathcal{F}_n]]$ by Markov's inequality

$= \alpha^{-1} E[Z] < \delta$ for large α

Hence the $E[Z|\mathcal{F}_n]$ are uniformly continuous;

(ii) $\sup_n E[E[Z|\mathcal{F}_n]] = \sup_n E[Z] = E[Z] < \infty$,

Hence the $E[Z|\mathcal{F}_n]$ are uniformly bounded;

$\Rightarrow$ it follows that the $E[Z|\mathcal{F}_n]$ is ui. #

- <u>Definition</u> ($\mathcal{F}_n \uparrow \mathcal{F}_\infty$)

  Define $\mathcal{F}_n \uparrow \mathcal{F}_\infty$ as

  (i) the $\mathcal{F}_n$ are σ-fields satisfying $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$; and

  (ii) the union of $\mathcal{F}_n$ generates the σ-field $\mathcal{F}_\infty$:

  $$\sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right) = \mathcal{F}_\infty.$$

- <u>Theorem</u> (Convergence Thm for $E[Z|\mathcal{F}_n]$ when $\mathcal{F}_n \uparrow \mathcal{F}_\infty$)

  If $\mathcal{F}_n \uparrow \mathcal{F}_\infty$ and Z is integrable, then

  $E[Z|\mathcal{F}_n] \to E[Z|\mathcal{F}_\infty]$ with probability 1

  Pf: • Note that the r.v.'s $X_n = E[Z|\mathcal{F}_n]$ form a ① martingale w.r.t. the $\mathcal{F}_n$; and that the X_n are ② uniformly integrable (ui).

  • Since $E[|X_n|] = E[|E[Z|\mathcal{F}_n]|] \le E[E[|Z||\mathcal{F}_n]] = E[|Z|] < \infty$,

  then ③ $K = \sup_n E[|X_n|] = E[|Z|] < \infty$,

  it follows by the martingale convergence thm that

  $X_n \to X$ with probability 1

  where X is an integrable r.v..

  • The problem is identify X with $E[Z|\mathcal{F}_\infty]$

  Because the X_n are ui, it is possible that

  $$\int_A X\,dP = \int_A \lim_n X_n\,dP = \lim_n \int_A X_n\,dP.$$

  If $A \in \mathcal{F}_k$ and $n \ge k$, then $A \in \mathcal{F}_n$, and

  $$\int_A X_n\,dP = \int_A E[Z|\mathcal{F}_n]\,dP = \int_A Z\,dP.$$

  $$\int_A X\,dP = \lim_n \int_A Z\,dP = \int_A Z\,dP \text{ for all } A \text{ in the } \pi\text{-system } \bigcup_{n=1}^{\infty} \mathcal{F}_n$$

  • Since X_n is measurable $\mathcal{F}_n \subset \mathcal{F}_\infty$, X is measurable $\mathcal{F}_\infty$

  it follows that

  $$\int_A X\,dP = \int_A Z\,dP = \int_A E[Z|\mathcal{F}_\infty]\,dP \text{ for all } A \text{ in } \mathcal{F}_\infty,$$

  That is, $X = E[Z|\mathcal{F}_\infty]$ with probability 1,

  then the result follows. #

- <u>Definition</u> (Reversed Martingale)

A left-infinite sequence $\ldots, X_{-2}, X_{-1}$ is a martingale relative to σ-fields $\ldots, \mathcal{F}_{-2}, \mathcal{F}_{-1}$ if

(i) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for $n < -1$;

(ii) X_n is measurable $\mathcal{F}_n$ for $n \leq -1$;

(iii) $E[|X_n|] < \infty$ for $n \leq -1$;

(iv) with probability 1,

$$E[X_{n+1} | \mathcal{F}_n] = X_n \text{ for } n < -1.$$

Remark:

This martingale has a last r.v. X_{-1}.

- <u>Theorem</u> (The Reversed Martingale Convergence Thm)

For a reversed martingale, that is, a left-infinite sequence $\dots, X_{-2}, X_{-1}$ is a "martingale",

(i) $\lim_{n\to\infty} X_{-n} = X$ exists and is integrable, and

(ii) $E[X] = E[X_{-n}]$ for all n.

Remark:

- The martingale $\dots, X_{-2}, X_{-1}$ has a last r.v. X_{-1}, so it is unnecessary to assume $k = \sup_n E[|X_n|] < \infty$.
- For reversed submartingale,

  $\lim_{n\to\infty} X_{-n} = X \in [-\infty, \infty)$ exists;

  but we cannot conclude that X is integrable.

Pf: • Suppose $\dots, X_{-2}, X_{-1}$ is a martingale.

Let $X_* = \liminf_n X_{-n}$ and $X^* = \limsup_n X_{-n}$; they may be infinite.

$$[X_* < X^*] = \bigcup [X_* < \alpha < \beta < X^*]$$

where the union extends all pairs of rationals α and β.

• Let U_n be the number of upcrossings of $[\alpha, \beta]$ by $X_{-n}, \dots, X_{-1}$.

By the upcrossing thm,

$$E[U_n] \le \frac{E[|X_{-1}|] + |\alpha|}{\beta - \alpha} < \infty.$$

Since U_n is nondecreasing and $E[U_n]$ is bounded, by the monotone convergence thm,

$$\lim_n E[U_n] = E[\lim_n U_n] = E[\sup_n U_n] < \infty$$

So $\sup_n U_n$ is integrable and hence is finite with probability 1.

If $X_* < \alpha < \beta < X^*$, then U_n must go to ∞.

Since U_n is bounded with probability 1,

$$P[X_* < X^*] \le \Sigma\, P[X_* < \alpha < \beta < X^*] = 0$$

and hence $X_* = X^*$ with probability 1,

Therefore $\lim_{n\to\infty} X_{-n} = X$ with probability 1.

- By the property for martingales:

  $$E[X_{n+k} | \mathcal{F}_n] = X_n,$$

  it follows that $X_{-n} = E[X_{-1} | \mathcal{F}_{-n}]$ for $n = 1, 2, \dots$.

  Since for integrable Z, $\mathcal{F}_n$ arbitrary σ-fields,

  the $E[Z | \mathcal{F}_n]$ are uniformly integrable,

  it follows that the X_{-n} are uniformly integrable.

  Therefore X is integrable:

  $$E[|X|] = E[\liminf_n |X_{-n}|] \le \liminf_n E[|X_{-n}|] < \infty,$$

- and we have that

  $$E[X] = E[\lim_n X_{-n}] = \lim_n E[X_{-n}]$$

  $$= E[X_{-n}] \text{ for all } n$$

  since $\dots, X_{-2}, X_{-1}$ is a martingale with a constant mean. #

- <u>Remark</u>:

  For reversed submartingale,

  the proof in Thm that X is integrable would not work:

  $$X_{-n} \le E[X_{-1} | \mathcal{F}_{-n}] \text{ for } n = 1, 2, \dots,$$

  so the X_{-n} are not ui.

- **Definition** ($\mathcal{F}_n \downarrow \mathcal{F}_0$)

  If the $\mathcal{F}_n$ are σ-fields.

  Define $\mathcal{F}_n \downarrow \mathcal{F}_0$ by

  (i) $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots$; and

  (ii) $\cap_{n=1}^{\infty} \mathcal{F}_n = \mathcal{F}_0$ is also a σ-field.

- **Theorem** (Convergence Thm for $E[Z|\mathcal{F}_n]$ when $\mathcal{F}_n \downarrow \mathcal{F}_0$)

  If $\mathcal{F}_n \downarrow \mathcal{F}_0$ and Z is integrable, then

  $E[Z|\mathcal{F}_n] \to E[Z|\mathcal{F}_0]$ with probability 1.

Pf: • If $X_{-n} = E[Z|\mathcal{F}_n]$, then

$$E[X_{-n+1}|\mathcal{F}_n] = E[E[Z|\mathcal{F}_{n-1}]|\mathcal{F}_n]$$
$$= E[Z|\mathcal{F}_n] \quad \because \mathcal{F}_{n-1} \supset \mathcal{F}_n$$
$$= X_{-n}. \quad \because \text{by definition of } X_{-n}.$$

Hence $\ldots, X_{-2}, X_{-1}$ form a martingale relative to $\ldots, \mathcal{F}_2, \mathcal{F}_1$.

• By the reversed martingale convergence thm,

$X_{-n} = E[Z|\mathcal{F}_n] \to X$ where X is integrable,

• and we have that the $X_{-n} = E[Z|\mathcal{F}_n]$ are ui.

As the limit of the $E[Z|\mathcal{F}_n]$ for $n \geq k$,

X is measurable $\mathcal{F}_k$; k being arbitrary,

X is measurable $\mathcal{F}_0$.

• By ui, $A \in \mathcal{F}_0$ implies that

$$\int_A X dP = \lim_n \int_A E[Z|\mathcal{F}_n] dP$$
$$= \lim_n \int_A E[E[Z|\mathcal{F}_n]|\mathcal{F}_0] dP \quad \because \text{by def of } E[\cdot|\mathcal{F}_0].$$
$$= \lim_n \int_A E[Z|\mathcal{F}_0] dP \quad \because \mathcal{F}_0 \subset \mathcal{F}_n$$
$$= \int_A E[Z|\mathcal{F}_0] dP,$$

Thus X is a version of $E[Z|\mathcal{F}_0]$. #

• <u>Theorem</u> (Dominated Convergence Thm for $E[Y_n | \mathcal{F}_n]$)

Suppose $Y_n \to_{a.s.} Y$ and $|Y_n| \le Z$ where Z is integrable.

If $\mathcal{F}_n \uparrow \mathcal{F}_\infty$, then

$$E[Y_n | \mathcal{F}_n] \to E[Y | \mathcal{F}_\infty] \text{ a.s.}$$

Pf: $\because |E[Y_n | \mathcal{F}_n] - E[Y | \mathcal{F}_\infty]|$

$$\le |E[Y_n | \mathcal{F}_n] - E[Y | \mathcal{F}_n]| + |E[Y | \mathcal{F}_n] - E[Y | \mathcal{F}_\infty]|$$

$\therefore$ it suffices to show that

(i) $|E[Y_n | \mathcal{F}_n] - E[Y | \mathcal{F}_n]| \to 0$ a.s.

(ii) $|E[Y | \mathcal{F}_n] - E[Y | \mathcal{F}_\infty]| \to 0$ a.s.

(i) By result before we have that

$$E[Y_n | \mathcal{F}_n] \to E[Y | \mathcal{F}_n] \text{ a.s.},$$

thus (i) follows.

(ii) By result before we have that

$$E[Y | \mathcal{F}_n] \to E[Y | \mathcal{F}_\infty] \text{ a.s.},$$

thus (ii) follows.

Hence the result follows.

#.

- <u>Example</u> (SLLN by the Reversed Martingale Convergence Thm)

Let $X_1, X_2, \cdots$ be iid with mean μ.
Then the partial sum process $S_n = X_1 + \cdots + X_n$ satisfies
$\bar{X}_n = S_n/n \to \mu$ a.s. and L_1 as $n \to \infty$.

Pf: • Let $\mathcal{A}_{-n} = \sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n, X_{n+1}, X_{n+2}, \cdots)$

Define $Y_{-n} = E[X_1 | \mathcal{A}_{-n}]$, then
the Y_{-n} is a reversed martingale on $\ldots, -2, -1$ and

• by the reversed martingale convergence thm,

$$E[X_1 | \mathcal{A}_{-n}] \to E[X_1 | \mathcal{A}_0] \text{ a.s. and } L_1, \quad \mathcal{A}_0 = \bigcap_n \mathcal{A}_{-n}, \text{ as } n \to \infty.$$

• Since

$$\begin{aligned} E[X_1 | \mathcal{A}_{-n}] &= E[X_1 | S_n, X_{n+1}, \ldots] \\ &= E[X_1 | S_n] \quad \text{by the } X_n \text{ are iid.} \\ &= \sum_{k=1}^{n} E[X_k | S_n]/n \quad \text{by symmetry} \\ &= E[S_n | S_n]/n \\ &= S_n/n. \end{aligned}$$

• Thus

$$S_n/n = E[X_1 | \mathcal{A}_{-n}] \to E[X_1 | \mathcal{A}_0] \text{ a.s. and } L_1 \text{ as } n \to \infty.$$

• But $\lim_n (S_n/n)$ is a tail r.v. and thus
it is a.s. constant by the Kolmogorov 0-1 law.
Hence $E[X_1 | \mathcal{A}_0]$ is a.s. a constant.

• But $E[E[X_1|\mathcal{A}_0]] = E(X_1) = \mu$,
so that the constant must be μ; that is
$E[X_1 | \mathcal{A}_0] = \mu$ a.s.

Hence the result follows.

#