• <u>Definition</u> (Probability Measure, Distribution, Distribution Function)

1. Let the probability measure space be $(\Omega, \mathcal{F}, P)$, where P is the probability measure.
2. A random variable is a measurable real function X s.t.
   $X : (\Omega, \mathcal{F}) \mapsto (R^1, \mathcal{R}^1)$, where $\mathcal{R}^1$ is the Borel sets of R^1.
3. The distribution of the r.v. X is defined by probability measure $\mu = PX^{-1}$ on $(R^1, \mathcal{R}^1)$ defined by
   $\mu(A) = P[X \in A] = PX^{-1}(A) = P[\omega : X(\omega) \in A]$, $A \in \mathcal{R}^1$.
4. The distribution function of X is defined by
   $F(x) = \mu(-\infty, x] = P[X \leq x]$ for real x

• <u>Definition</u> (Expected Values)

The expected value or expectation of $g(X)$ is

$$E[g(X)] = \int g(X)\,dP = \int_\Omega g(X(\omega))\,P(d\omega) \quad \text{w.r.t the probability measure } P.$$
$$= \int_R g(x)\,\mu(dx) \quad \text{w.r.t. the distribution } \mu$$
$$= \int_R g(x)\,dF(x) \quad \text{if } F \text{ is the distribution function of } X$$
$$= \int_R g(x) f(x)\,dx \quad \text{if } X \text{ has density } f$$

- <u>Proposition</u> (Monotone Continuity of P)

The measure P is continuous for monotone sequences

(i) If $A_n \uparrow A$, where $A_n \in \mathcal{B}$, then $P(A_n) \uparrow P(A)$

(ii) If $A_n \downarrow A$, where $A_n \in \mathcal{B}$, then $P(A_n) \downarrow P(A)$.

Pf: (i) Assume

$A_1 \subset A_2 \subset A_3 \cdots$

and define

$B_1 = A_1$, $B_2 = A_2 \setminus A_1, \cdots, B_n = A_n \setminus A_{n-1}$;

Then $\{B_n\}$ is a disjoint sequence of events and

$\bigcup_{i=1}^{n} B_i = A_n$, $\bigcup_{i=1}^{\infty} B_i = \bigcup_i A_i = A$ by $A_1 \subset A_2 \subset \cdots$

By σ-additivity,

$$P(A) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \uparrow \sum_{i=1}^{n} P(B_i)$$

$$= \lim_{n\to\infty} \uparrow P\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} \uparrow P(A_n).$$

To prove (ii), note if $A_n \downarrow A$, then $A_n^c \uparrow A^c$ and by part (i)

$$P(A_n^c) = 1 - P(A_n) \uparrow P(A^c) = 1 - P(A)$$

so that $P(A_n) \downarrow P(A)$.

#

• <u>Proposition</u> (Fatou's Lemma and Continuity of P)

Suppose $A_n \in \mathcal{B}$, $n \geq 1$.

(i) <u>Fatou Lemma:</u>

$$P(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} P(A_n)$$
$$\leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_{n\to\infty} A_n)$$

(ii) Continuity of P:

If $A_n \to A$, then $P(A_n) \to P(A)$

Pf: (ii). Suppose (i) is true.

If $A_n \to A$, then $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A$

$$P(A) = P(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} P(A_n)$$
$$\leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_{n\to\infty} A_n) = P(A)$$

Hence

$\lim_{n\to\infty} P(A_n) = P(A)$, i.e. $P(A_n) \to P(A)$.

(i) $P(\liminf_{n\to\infty} A_n) = P\left(\lim_{n\to\infty} \uparrow \left(\bigcap_{k\geq n} A_k\right)\right)$

$= \lim_{n\to\infty} \uparrow P\left(\bigcap_{k\geq n} A_k\right)$ by the monotone continuity of P.

$\leq \lim_{n\to\infty} \inf_{k\geq n} P(A_k)$ by $A_k \supset \bigcap_{k\geq n} A_k$ for all $k \geq n$.

$= \liminf_{n\to\infty} P(A_n)$

Likewise,

$P(\limsup_{n\to\infty} A_n) = P\left(\lim_{n\to\infty} \downarrow \left(\bigcup_{k\geq n} A_k\right)\right)$

$= \lim_{n\to\infty} \downarrow P\left(\bigcup_{k\geq n} A_k\right)$ by the monotone continuity of P.

$\geq \lim_{n\to\infty} \sup_{k\geq n} P(A_k)$ by $A_k \subset \bigcup_{k\geq n} A_k$ for all $k \geq n$.

$= \limsup_{n\to\infty} P(A_n)$.

#.

• <u>Proposition</u> (Distribution Function F).

A function $F: \mathbb{R} \mapsto [0,1]$ satisfying (i), (ii), (iii) is called a (probability) distribution function (df). Define

$F(x) = \mu((-\infty, x]) = P[X \le x]$

(i) F is right continuous,

(ii) F is monotone non-decreasing.

(iii) F has limits at $\pm\infty$,

$$F(\infty) \equiv \lim_{x \uparrow \infty} F(x) = 1$$

$$F(-\infty) \equiv \lim_{x \downarrow -\infty} F(x) = 0.$$

Pf: (i), we may show F is right continuous as follows:

Let $x_n \downarrow x$, we need to prove $F(x_n) \downarrow F(x)$.

$\because (-\infty, x_n] \downarrow (-\infty, x]$, by the monotone continuity of P,

$F(x_n) = P((-\infty, x_n]) \downarrow P((-\infty, x]) = F(x)$.

(ii) If $x < y$, then $(-\infty, x] \subset (-\infty, y]$,

by the monotonicity of P,

$F(x) = P((-\infty, x]) \le P((-\infty, y]) = F(y)$

(iii) $F(\infty) = \lim_{x_n \uparrow \infty} F(x_n)$ for any sequence $x_n \uparrow \infty$

$= \lim_{x_n \uparrow \infty} \uparrow P((-\infty, x_n])$

$= P(\lim_{x_n \uparrow \infty} \uparrow (-\infty, x_n])$ by monotone continuity of P

$= P(\bigcup_n (-\infty, x_n]) = P(-\infty, \infty)$

$= P(\mathbb{R}) = P(\Omega) = 1$

$F(-\infty) = \lim_{x_n \downarrow -\infty} F(x_n) = \lim_{x_n \downarrow -\infty} \downarrow P((-\infty, x_n])$

$= P(\lim_{x_n \downarrow -\infty} \downarrow (-\infty, x_n])$ by monotone continuity of P.

$= P(\bigcap_n (-\infty, x_n]) = P(\emptyset) = 0$.

#

• <u>Proposition</u> (Borel-Cantelli Lemma)

Let $\{A_n\}$ be any events. If

$$\sum_n P(A_n) < \infty,$$

then

$$P([A_n \text{ i.o.}]) = P(\limsup_{n\to\infty} A_n) = 0.$$

Note: the basic tool for proving almost sure convergence!

Pf: $P([A_n \text{ i.o.}]) = P(\lim_{n\to\infty} \bigcup_{j\geq n} A_j)$

$$= \lim_{n\to\infty} P(\bigcup_{j\geq n} A_j) \quad \text{by monotone continuity of } P.$$

$$\leq \limsup_{n\to\infty} \sum_{j\geq n} P(A_j) \quad \text{by subadditivity of } P.$$

$$= 0$$

since $\sum_n P(A_n) < \infty$ implies $\sum_{j\geq n} P(A_j) \to 0$ as $n\to\infty$.

#

• <u>Example</u>

Suppose $\{X_n, n\geq 1\}$ are Bernoulli r.v.'s with

$$P[X_n=1] = p_n = 1 - P[X_n=0]$$

We assert that $P[\lim_{n\to\infty} X_n = 0] = 1$ if $\sum_n p_n < \infty$.

If $\sum_n p_n = \sum_n P[X_n=1] < \infty$,

then by the Borel-Cantelli Lemma,

$$P([X_n=1] \text{ i.o.}) = 0$$

$$\Rightarrow P((\limsup_{n\to\infty} [X_n=1])^c) = P(\liminf_{n\to\infty} [X_n=0]) = 1$$

That is, with probability 1, the r.v.'s $\{X_n\}$ must converge to 0.

Hence $X_n \to 0$ with probablity 1, a.s..

- <u>Proposition</u> (Borel Zero-One Law)

If $\{A_n\}$ is a sequence of "independent" events, then

$$P([A_n \text{ i.o.}]) = \begin{cases} 0 & \text{iff } \sum_n P(A_n) < \infty, \\ 1 & \text{iff } \sum_n P(A_n) = \infty. \end{cases}$$

Note: The Borel-Cantelli Lemma does not require "independence", but the Borel Zero-One Law does.

Pf: From the Borel-Cantelli Lemma, if

$$\sum_n P(A_n) < \infty,$$

then $P([A_n \text{ i.o.}]) = 0$.

Conversely, suppose $\sum_n P(A_n) = \infty$, then

$$\begin{aligned}
P([A_n \text{ i.o.}]) &= P(\limsup_{n\to\infty} A_n) \\
&= 1 - P(\liminf_{n\to\infty} A_n^c) = 1 - P(\lim_{n\to\infty} \bigcap_{k\geq n} A_k^c) \\
&= 1 - \lim_{n\to\infty} P(\bigcap_{k\geq n} A_k^c) = 1 - \lim_{n\to\infty} P(\lim_{m\to\infty}\downarrow \bigcap_{k\geq n}^{m} A_k^c) \quad \text{by monotone continuity} \\
&= 1 - \lim_{n\to\infty}\lim_{m\to\infty} P(\bigcap_{k\geq n}^{m} A_k^c) \quad \text{by monotone continuity of } P \\
&= 1 - \lim_{n\to\infty}\lim_{m\to\infty} \prod_{k\geq n}^{m} (1 - P(A_k)) \quad \text{by independence of } \{A_k\}.
\end{aligned}$$

To prove $\lim_{n\to\infty}\lim_{m\to\infty} \prod_{k\geq n}^{m}(1-P(A_k)) = 0$.

use the inequality $1-x < e^{-x}$, $0<x<1$. (*)

Now we have

$$\begin{aligned}
\lim_{m\to\infty} \prod_{k\geq n}^{m}(1-P(A_k)) &\leq \lim_{m\to\infty} \prod_{k\geq n}^{m} e^{-P(A_k)} = \lim_{m\to\infty} e^{-\sum_{k\geq n}^{m} P(A_k)} \\
&= e^{-\sum_{k\geq n}^{\infty} P(A_k)} = e^{-\infty} = 0 \quad \text{since } \sum_n P(A_n) = \infty.
\end{aligned}$$

This is true for all n, so

$$\lim_{n\to\infty}\lim_{m\to\infty} \prod_{k\geq n}^{m}(1-P(A_k)) = 0.$$

To verify (*), for $0<x<1$,

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k} \geq x \Rightarrow \frac{1}{1-x} \geq e^{x}, \; e^{-x} \leq 1-x.$$

\#

- <u>Example</u>

Suppose $\{X_n, n\geq 1\}$ are independent in addition to being Bernoulli, with

$$P[X_k=1] = p_k = 1 - P[X_k=0].$$

Then we assert that

$$P[X_n \to 0] = 1 \text{ iff } \sum_n p_n < \infty.$$

$\because$ $P[X_n \to 0] = P\left(\liminf_{n\to\infty} [X_n = 0]\right) = 1$

$\therefore$ we need to show

$$P\left(\left(\liminf_{n\to\infty} [X_n=0]\right)^c\right) = P\left(\limsup_{n\to\infty} [X_n=1]\right) = P([X_n=1] \text{ i.o.}) = 0$$

iff

$$\sum_n P[X_n=1] = \sum_n p_n < \infty.$$

by the Borel Zero-One Law.

#

- **Theorem** (A Second Pf of the Second Borel-Cantelli Lemma)

For a sequence $A_1, A_2, \dots$ of events, let $N_n = I_{A_1} + \cdots + I_{A_n}$, $p_k = P(A_k)$.

Note that $[A_n \text{ i.o.}] = [\omega : \sup_n N_n(\omega) = \infty]$.

If the A_n are independent and $\sum p_n = \infty$, then

$$P[A_n \text{ i.o.}] = P[\omega : \sup_n N_n(\omega) = \infty] = 1$$

This is the second Borel-Cantelli lemma.

Pf: Put $p_k = P(A_k)$ and $m_n = p_1 + \cdots + p_n$.

From $E[I_{A_k}] = p_k$, $Var[I_{A_k}] = p_k(1-p_k) \le p_k$ follow

$E[N_n] = \sum_1^n p_k = m_n$, $Var[N_n] = \sum_1^n p_k(1-p_k) \le m_n$. by the A_n indep.

If $m_n > x$, then

$$\begin{aligned} P[N_n \le x] &= P[N_n - m_n \le x - m_n] \quad \text{by } x - m_n < 0 \\ &= P[N_n - m_n \le -|m_n - x|] \quad \text{by } m_n - x > 0 \\ &\le P[|N_n - m_n| \ge m_n - x] \\ &\le \frac{Var[N_n]}{(m_n - x)^2} \le \frac{m_n}{(m_n - x)^2}. \quad \text{by Chebychev's inequality} \end{aligned}$$

If $\sum p_n = \infty$, so that $m_n \to \infty$, thus

$\lim_n P[N_n \le x] = 0$ for each x.

Since $P[\sup_k N_k \le x] \le P[N_n \le x]$,

$P[\sup_k N_k \le x] = 0$, and hence

$$P[\sup_k N_k < \infty] = P\left(\bigcup_x [\sup_k N_k \le x]\right)$$

$$\le \sum_x P[\sup_n N_n \le x] = 0.$$

Hence $P[A_n \text{ i.o.}] = P[\sup_n N_n = \infty] = 1$.

#

- **Theorem** (A Refinement of the Second Borel-Cantelli Lemma)

If $\sum P(A_n)=\infty$ and

$$\liminf_n \frac{\sum_{j,k\le n} P(A_j\cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2}\le 1, \quad (*)$$

then $P[A_n \text{ i.o.}]=1$

Remark:

As the proof will show, the ratio in $(*)\ge 1$;

if $(*)$ holds, then $\liminf_n \frac{\sum_{j,k\le n} P(A_j\cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2}=1$.

Pf: • Let $\theta_n=\frac{\sum_{j,k\le n} P(A_j\cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2}=\frac{\sum_{j,k\le n} P(A_j\cap A_k)}{m_n^2}$

where $m_n=E[N_n]=\sum_{k\le n} E[I_{A_k}]=\sum_{k\le n} P(A_k)$.

• $Var[N_n]=E[N_n^2]-m_n^2$

$=\sum_{j,k\le n} E[I_{A_j}\cdot I_{A_k}]-m_n^2$

$=\sum_{j,k\le n} P(A_j\cap A_k)-m_n^2$

$=(\theta_n-1)\,m_n^2\ge 0$

So that $\theta_n-1\ge 0$, i.e. $\theta_n\ge 1$.

• If $m_n>x$, then

$P[N_n\le x]=P[N_n-m_n\le -(m_n-x)]$ by $m_n>x$.

$\le P[|N_n-m_n|\ge m_n-x]$

$\le \frac{Var[N_n]}{(m_n-x)^2}=\frac{(\theta_n-1)\,m_n^2}{(m_n-x)^2}$ by Chebychev's inequality.

If $\sum p_n=\infty$, so that $m_n\to\infty$, thus $m_n^2/(m_n-x)^2\to 1$.

Since $\theta_n-1\ge 0$ and by the hypothesis $\liminf_n \theta_n\le 1$,

thus $\liminf_n(\theta_n-1)=0$, and hence

$\liminf_n P[N_n\le x]=0$ for each x.

Since $P[\sup_k N_k\le x]\le P[N_n\le x]$, then $P[\sup_k N_k\le x]=0$, hence

$P[\sup_k N_k<\infty]=P\left(\bigcup_x[\sup_k N_k\le x]\right)$

$\le \sum_x P[\sup_k N_k\le x]=0$

Hence $P[A_n \text{ i.o.}]=P[\sup_n N_n=\infty]=1$. #

- <u>Corollary</u> (Cor of the Refinement of the Second Borel-Cantelli Lemma)

If, as in the second Borel-Cantelli lemma, the A_n are independent (or even if they are merely pairwise independent), then $\Sigma P(A_n) = \infty$ implies that

$$\liminf_n \frac{\sum_{j,k\le n} P(A_j \cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2} \le 1$$

then $P[A_n \text{ i.o.}] = 1$ by the refinement of the second Borel-Cantelli lemma.

Pf: Put $p_k = P(A_k)$. Suppose $\Sigma p_n = \infty$.

Suppose the A_n are independent (or pairwise independent), by the refinement of the second Borel-Cantelli lemma, it suffices to check that

$$\liminf_n \frac{\sum_{j,k\le n} P(A_j \cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2} \le 1.$$

$$\frac{\sum_{j,k\le n} P(A_j \cap A_k)}{\left(\sum_{k\le n} P(A_k)\right)^2} = \frac{\sum_{k\le n} p_k + \sum_{\substack{j,k\le n \\ j\ne k}} p_j p_k}{m_n^2} \quad \text{by } \begin{cases} \text{pairwise independent} \\ \text{independent} \end{cases}$$

$$= \frac{m_n^2 + \sum_{k\le n} p_k - \sum_{k\le n} p_k^2}{m_n^2} \quad \text{by } m_n^2 = (p_1 + \cdots + p_n)^2 = \sum_{k\le n} p_k^2 + \sum_{\substack{j,k\le n \\ j\ne k}} p_j p_k$$

$$= 1 + \frac{\sum_{k\le n}(p_k - p_k^2)}{m_n^2}$$

$$\le 1 + \frac{m_n}{m_n^2} = 1 + \frac{1}{m_n} \quad \text{by } \sum_{k\le n}(p_k - p_k^2) \le \sum_{k\le n} p_k = m_n$$

$$\to 1 \quad \text{by } m_n = \sum_{k=1}^{n} p_k \to \sum_n p_n = \infty$$

Hence this proves the second Borel-Cantelli lemma again.

#.

• <u>Thm</u> (Bernstein's Version of The Weierstrass Approximation Thm).

Let $f: [0,1] \to \mathbb{R}$ be continuous, and
define the "Bernstein polynomial of degree n" by

$$B_n(x) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}, \quad 0 \le x \le 1.$$

Then

$$B_n(x) \to f(x) \text{ uniformly for } x \in [0,1].$$

Pf: • Let $\{\delta_k, k=1,\cdots,n\}$ be iid Bernoulli r.v.'s with
$P[\delta_i = 1] = x = 1 - P[\delta_i = 0]$, hence have finite mean x.

• Define $S_n = \sum_{i=1}^{n} \delta_i$ s.t. $S_n \sim \text{binomial}(n, x)$.

• The weak law of large numbers implies that

$$\frac{S_n}{n} \xrightarrow{P} x.$$

• $\because$ f is continuous on $[0,1]$, a compact set

$\therefore$ $f\left(\frac{S_n}{n}\right) \xrightarrow{P} f(x)$ and f is bounded, uniformly continuous
the dominated convergence thm implies that

$$E f\left(\tfrac{S_n}{n}\right) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} = B_n(x) \to f(x) \text{ as } n \to \infty.$$

• Define the modulus of continuity

$$\omega(\delta) = \sup_{\substack{|x-y| \le \delta \\ 0 \le x,y \le 1}} |f(x) - f(y)|,$$

the uniformly continuity of f implies $\lim_{\delta \downarrow 0} \omega(\delta) = 0$.

$$\begin{aligned}
\sup_{0 \le x \le 1} |B_n(x) - f(x)| &= \sup_x \left|E\left(f\left(\tfrac{S_n}{n}\right)\right) - f(x)\right| \le \sup_x E\left|f\left(\tfrac{S_n}{n}\right) - f(x)\right| \\
&\le \sup_x \left\{ E\left(\left|f\left(\tfrac{S_n}{n}\right) - f(x)\right| 1_{\left[\left|\frac{S_n}{n} - x\right| \le \epsilon\right]}\right) + E\left(\left|f\left(\tfrac{S_n}{n}\right) - f(x)\right| 1_{\left[\left|\frac{S_n}{n} - x\right| > \epsilon\right]}\right) \right\} \\
&\le \omega(\epsilon) P\left[\left|\tfrac{S_n}{n} - x\right| \le \epsilon\right] + 2\|f\| \sup_x P\left[\left|\tfrac{S_n}{n} - x\right| > \epsilon\right]. \\
&\quad \because \left|f\left(\tfrac{S_n}{n}\right) - f(x)\right| \le \left|f\left(\tfrac{S_n}{n}\right)\right| + |f(x)| \le 2\|f\|, \\
&\le \omega(\epsilon) + \frac{2\|f\|}{\epsilon^2} \sup_x \operatorname{Var}\left(\tfrac{S_n}{n}\right) = \omega(\epsilon) + \frac{2\|f\|}{\epsilon^2} \sup_x \frac{x(1-x)}{n} \\
&\le \omega(\epsilon) + \frac{2\|f\|}{\epsilon^2} \frac{1}{4n}.
\end{aligned}$$

$$\limsup_{n \to \infty} \sup_x |B_n(x) - f(x)| \le \omega(\epsilon) \to 0 \text{ as } \epsilon \to 0.$$

#

- <u>Definition</u> (Tail σ-field)

  Let $\{X_n\}$ be a sequence of r.v.'s and define

  $$\mathcal{F}_n' = \sigma(X_{n+1}, X_{n+2}, \cdots), \; n=1,2,\cdots$$

  The "tail σ-field" $\mathcal{T}$ is defined as

  $$\mathcal{T} = \bigcap_n \mathcal{F}_n' = \lim_{n\to\infty} \downarrow \sigma(X_{n+1}, X_{n+2}, \cdots)$$

  These are events which depend on the tail of the $\{X_n\}$ sequence.

  Λ: tail event if $\Lambda \in \mathcal{T}$

  X: tail r.v. if $X : (\Omega, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ s.t. $X^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{T}$.

- <u>Theorem</u> (Kolmogorov Zero-One Law).

  If $\{X_n\}$ are "independent" r.v.'s with tail σ-field $\mathcal{T}$,

  then $\Lambda \in \mathcal{T}$ implies $P(\Lambda) = 0$ or 1.

  Pf: • Suppose $\Lambda \in \mathcal{T}$. We show that Λ is independent of itself s.t

  $$P(\Lambda) = P(\Lambda \cap \Lambda) = P(\Lambda)P(\Lambda), \text{ therefore } P(\Lambda) = 0 \text{ or } 1.$$

  • We define $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$ s.t. $\mathcal{F}_n \uparrow$ and $\mathcal{F}_\infty = \sigma(X_1, X_2, \cdots)$.

  Note that

  $$\Lambda \in \mathcal{T} \subset \mathcal{F}_n' = \sigma(X_{n+1}, X_{n+1}, \cdots) \subset \sigma(X_1, X_2, \cdots) = \mathcal{F}_\infty.$$

  • For all n, we have $\Lambda \in \mathcal{F}_n'$,

  Since $\mathcal{F}_n \perp\!\!\!\perp \mathcal{F}_n'$, we have $\Lambda \perp\!\!\!\perp \mathcal{F}_n, \forall n$,

  and therefore $\Lambda \perp\!\!\!\perp \bigcup_n \mathcal{F}_n$.

  Let $\mathcal{C}_1 = \{\Lambda\}$, $\mathcal{C}_2 = \bigcup_n \mathcal{F}_n$. Then $\mathcal{C}_i$ is a π-system, $i=1,2$, $\mathcal{C}_1 \perp\!\!\!\perp \mathcal{C}_2$ and

  $$\sigma(\mathcal{C}_1) = \{\phi, \Lambda, \Lambda^c, \Omega\} \perp\!\!\!\perp \sigma(\mathcal{C}_2) = \mathcal{F}_\infty.$$

  Now $\Lambda \in \sigma(\mathcal{C}_1)$ and $\Lambda \in \mathcal{F}_\infty$,

  thus Λ is independent of itself.

  #.

• <u>Lemma</u> (Tail r.v. X is constant a.s.)

Let $\mathcal{G}$ be a σ-field s.t. for $\Lambda \in \mathcal{G}$, $P(\Lambda)=0$ or 1, and X be a r.v. measurable w.r.t $\mathcal{G}$. Then there exists c s.t. $P[X=c]=1$.

Pf: Let $F(x)=P[X\le x]$, which is non-decreasing.

Since $[X\le x]\in\sigma(X)\subset\mathcal{G}$,

$F(x)=0$ or 1, $\forall x\in\mathbb{R}$.

Let $c=\sup_{x\in\mathbb{R}}\{x: F(x)=0\}$.

The df F must have a jump of size 1 at c and thus $P[X=c]=1$.

#

• <u>Corollary</u> (Cor. of the Kolmogorov Zero-One Law)

Let $\{X_n\}$ be "indep." r.v.'s and $S_n=\sum_{i=1}^{n}X_i$. Then

(i) $P[\sum_n X_n \text{ converges}]=0$ or 1

(ii) $\limsup_{n\to\infty} X_n$, $\liminf_{n\to\infty} X_n$ are constants w.p. 1

(iii) $P[\omega: \frac{S_n(\omega)}{n}\to 0]=0$ or 1

Pf: (i) $\because \sum_{n=1}^{\infty}X_n(\omega)$ converges iff $\sum_{n=m+1}^{\infty}X_n$ converges for any m.

$\therefore [\sum_{n=1}^{\infty}X_n(\omega) \text{ converges}]=[\sum_{n=m+1}^{\infty}X_n \text{ converges}]\in\mathcal{F}_m'$, $\forall m$.

Then $[\sum_{n=1}^{\infty}X_n(\omega) \text{ converges}]\in\mathcal{T}$

(ii) $\because$ limsup of the sequence $\{X_n\}$ = limsup of the sequence $\{X_m, X_{m+1},\cdots\}$, $\forall m$

so we have $\limsup_{n\to\infty} X_n\in\mathcal{T}$.

(iii) For any m,

$$\lim_{n\to\infty}\frac{S_n(\omega)}{n}=\lim_{n\to\infty}\frac{\sum_{i=1}^{n}X_i(\omega)}{n}=\lim_{n\to\infty}\frac{\sum_{i=m+1}^{n}X_i(\omega)}{n} \text{ for any } m.$$

So for any m, $\lim_{n\to\infty}\frac{S_n(\omega)}{n}\in\mathcal{F}_m'$, and hence $\lim_{n\to\infty}\frac{S_n(\omega)}{n}\in\mathcal{T}$.

#

- **Definition** (Simple Functions)

A simple real function is one with finite range; that is,

$$f = \sum_{i=1}^{n} x_i I_{A_i}$$

where the A_i form a finite decomposition of Ω; that is.

$$\bigcup_{i=1}^{n} A_i = \Omega \text{ and } A_i \cap A_j = \emptyset \text{ for } i \neq j$$

f is measurable $\mathcal{F}$ if each $A_i \in \mathcal{F}$.

- **Theorem** (Approximation Sequence of Real Measurable Function)

If f is real and measurable $\mathcal{F}$, there exists a sequence $\{f_n\}$ of simple functions, each measurable $\mathcal{F}$, s.t.

$$0 \le f_n(\omega) \uparrow f(\omega) \quad \text{if } f(\omega) \ge 0$$

$$0 \ge f_n(\omega) \downarrow f(\omega) \quad \text{if } f(\omega) \le 0$$

Pf: Define

$$f_n = -n 1_{[f \le -n]} + \sum_{k=1}^{n2^n} \frac{-(k-1)}{2^n} 1_{\left[\frac{-k}{2^n} \le f \le \frac{-(k-1)}{2^n}\right]} + \sum_{k=1}^{n2^n} \frac{(k-1)}{2^n} 1_{\left[\frac{(k-1)}{2^n} \le f \le \frac{k}{2^n}\right]} + n 1_{[f \ge n]}$$

This sequence has the desired property. #

- Definition (The Integral)
  - Let f be real, nonnegative measurable function on a measurable space $(\Omega, \mathcal{F}, \mu)$.
    The integral of f is defined by
    $$\int f \, d\mu = \sup \sum_i \left[\inf_{\omega \in A_i} f(\omega)\right] \mu(A_i),$$
    the supremum extends over all finite decompositions $\{A_i\}$ of Ω into $\mathcal{F}$-sets.
  - For general f (not necessarily nonnegative),
    its "positive part" is
    $$f^+(\omega) = f(\omega) \vee 0, \quad \text{so } f^+ \geq 0.$$
    its "negative part" is
    $$f^-(\omega) = (-f(\omega)) \vee 0, \quad \text{so } f^- \geq 0.$$
    Then $f = f^+ - f^-$.
    The integral of f is defined by
    $$\int f \, d\mu = \int f^+ d\mu - \int f^- d\mu. \quad (1)$$
    Note:
    - f is integrable if
      $\int f^+ d\mu < \infty$ and $\int f^- d\mu < \infty$; thus has (1) as its definite integral.
    - f is not integrable if
      1. $\int f^+ d\mu = \infty$ and $\int f^- d\mu < \infty$; but has definite integral $+\infty$;
      2. $\int f^+ d\mu < \infty$ and $\int f^- d\mu = \infty$; but has definite integral $-\infty$;
      3. $\int f^+ d\mu = \infty$ and $\int f^- d\mu = \infty$; no definite integral.
  - f is integrable iff $|f|$ is integrable
    Pf: f is integrable
    $\Leftrightarrow \int f^+ d\mu < \infty$ and $\int f^- d\mu < \infty$
    $\Leftrightarrow \int f^+ d\mu + \int f^- d\mu = \int (f^+ + f^-) d\mu = \int |f| d\mu < \infty$
    $\Leftrightarrow |f|$ is integrable.
    #

- <u>Definition</u> (Integration over Sets)

The integral of f over a set A in $\mathcal{F}$ is defined by

$$\int_A f\,d\mu = \int I_A f\,d\mu$$

Note:

- The definition applies if f is defined only on A (Set $f=0$ outside A).
- $\int_A f\,d\mu = 0$ if $\mu(A)=0$.
- Properties

(i) If f and g are nonnegative and

$$\int_A f\,d\mu = \int_A g\,d\mu \text{ for all } A \text{ in } \mathcal{F},$$

and if μ is σ-finite, then $f=g$ a.e.

(ii) If f and g are integrable and

$$\int_A f\,d\mu = \int_A g\,d\mu \text{ for all } A \text{ in } \mathcal{F}, \text{ then } f=g \text{ a.e.}$$

(iii) If f and g are integrable and

$\int_A f\,d\mu = \int_A g\,d\mu$ for all $A \in \mathcal{P}$ where

$\mathcal{P}$ is a π-system generating $\mathcal{F}$ and

Ω is a finite or countable union of $\mathcal{P}$-sets, then $f=g$ a.e.

Remark:

The results about integrals are proved

1. for indicator functions
2. then for simple functions.
3. then for nonnegative functions
4. then for integrable functions.

- <u>Theorem</u> (Monotone Convergence Thm, MCT).

If $0 \le X_n \uparrow X$, then

$E(X_n) \uparrow E(X)$, or equivalently, $\lim_{n\to\infty} E(X_n) = E(\lim_{n\to\infty} X_n)$.

Note: MCT allows the interchange of limits and expectation

Pf: Suppose we have $X_n, X \in \bar{\mathcal{E}}_+ := \{X \ge 0,\ X:(\Omega, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))\}$.

We may find simple functions $Y_m^{(n)} \in \mathcal{E}_+ := \{X \ge 0,\ X \text{ is simple r.v.}\}$,

to act as approximation to $X_{(n)}$ s.t.

$$Y_m^{(n)} \uparrow X_n,\ m \to \infty.$$

We need to find a sequence of simple functions $\{Z_m\}$ approximating X,

$$Z_m \uparrow X,\ m \to \infty,$$

which can be expressed in terms of the approximation $\{X_n\}$.

So define

$$Z_m = \bigvee_{n \le m} Y_m^{(n)} \quad \because Y_m^{(n)} \ge 0 \Rightarrow Z_m \in \mathcal{E}_+$$

Note that Z_m is non-decreasing since

$$Z_m \le \bigvee_{n \le m} Y_{m+1}^{(n)} \quad \text{since } Y_m \le Y_{m+1}$$

$$\le \bigvee_{n \le m+1} Y_{m+1}^{(n)} = Z_{m+1}$$

For $n \le m$,

① $Y_m^{(n)} \le Y_m^{(m)} \le \bigvee_{j \le m} Y_m^{(j)} = Z_m$;

② $Z_m \le \bigvee_{j \le m} X_j = X_m$ since $Y_m^{(j)} \le X_j$ which is monotone in j

$\Rightarrow Y_m^{(n)} \le Z_m \le X_m$

We conclude that for all n

$$X_n = \lim_{m\to\infty} Y_m^{(n)} \le \lim_{m\to\infty} Z_m \le \lim_{m\to\infty} X_m$$

So

$$X = \lim_{n\to\infty} X_n \le \lim_{m\to\infty} Z_m \le \lim_{m\to\infty} X_m = X.$$

(next pg. cont)

Hence $X = \lim_{n\to\infty} X_n = \lim_{m\to\infty} Z_m$ and

it follows that $\{Z_m\}$ is a simple function approximation to X

Because expectation is monotone on $\mathcal{E}_+$ simple r.v.'s,

$$E(X_n) = \lim_{m\to\infty} \uparrow E(Y_m^{(n)}) \quad \text{by expectation definition.}$$

$$\le \lim_{m\to\infty} \uparrow E(Z_m) \le \lim_{m\to\infty} \uparrow E(X_m) \quad \text{from above}$$

$\because$ $Z_m \in \mathcal{E}_+$ and $\{Z_m\}$ is a simple function approximation to X,

$$E(X) = E(\lim_{m\to\infty} \uparrow Z_m) \equiv \lim_{m\to\infty} \uparrow E(Z_m) \quad \text{by expectation definition.}$$

$\therefore$ $E(X_n) \le E(X) = \lim_{m\to\infty} \uparrow E(Z_m) \le \lim_{m\to\infty} \uparrow E(X_m)$,

So $\lim_{n\to\infty} E(X_n) \le E(X) \le \lim_{m\to\infty} E(X_m)$,

hence the result follows.

#

<u>Corollary</u> (Series Version of MCT)

If $X_k \ge 0$, $k \ge 1$, then

$$E\left(\sum_{k=1}^{\infty} X_k\right) = \sum_{k=1}^{\infty} E(X_k).$$

So that the expectation and infinite sum can be exchanged

Pf: $E\left(\sum_{k=1}^{\infty} X_k\right) = E\left(\lim_{n\to\infty} \sum_{k=1}^{n} X_k\right)$

$$= \lim_{n\to\infty} E\left(\sum_{k=1}^{n} X_k\right) \quad \because \sum_{k=1}^{n} X_k \uparrow \sum_{k=1}^{\infty} X_k \text{ and then by MCT}$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} E(X_k)$$

$$= \sum_{k=1}^{\infty} E(X_k).$$

#

- <u>Theorem</u> (Fatou's Lemma)

  If $X_n \ge 0$, then

  $$E(\liminf_{n\to\infty} X_n) \le \liminf_{n\to\infty} E(X_n).$$

  More generally, if there exists $Z \in L_1$ s.t. $X_n \ge Z$, then the inequality holds.

  Pf: • If $X_n \ge 0$, then

  $$E(\liminf_{n\to\infty} X_n) = E\left(\lim_{n\to\infty} \uparrow \left(\bigwedge_{k\ge n} X_k\right)\right)$$

  $$= \lim_{n\to\infty} \uparrow E\left(\bigwedge_{k\ge n} X_k\right) \quad \text{by MCT} : \bigwedge_{k\ge n} X_k \ge 0$$

  $\because \bigwedge_{k\ge n} X_k \le X_k, \forall k \ge n, \therefore E\left(\bigwedge_{k\ge n} X_k\right) \le E(X_k), \forall k \ge n$

  $$\le \lim_{n\to\infty} \inf_{k\ge n} E(X_k) = \liminf_{n\to\infty} E(X_n).$$

  • For the case: $X_n \ge Z$, we have $X_n - Z \ge 0$, and

  $$E(\liminf_{n\to\infty} (X_n - Z)) \le \liminf_{n\to\infty} E(X_n - Z)$$

  so $E(\liminf_{n\to\infty} X_n) - \cancel{E(Z)} \le \liminf_{n\to\infty} E(X_n) - \cancel{E(Z)}$,

  then the result follows.

  #

- <u>Corollary</u> (More Fatou Lemma)

  If $X_n \le Z$ s.t. $Z \in L_1$. Then

  $$\limsup_{n\to\infty} E(X_n) \le E(\limsup_{n\to\infty} X_n)$$

  Pf: If $X_n \le Z$, then $-X_n \ge -Z \in L_1$, and

  the Fatou Lemma implies

  $$E(\liminf_{n\to\infty} (-X_n)) \le \liminf_{n\to\infty} E(-X_n)$$

  $$E(-\liminf_{n\to\infty} (-X_n)) \ge -\liminf_{n\to\infty} E(-X_n)$$

  The result follows by $-\liminf - = \limsup$

  #

• <u>Proposition</u> (More Fatou's Lemma)

(a) If $X_n \xrightarrow{a.s.} X$, then for every $r > 0$,

$$E(|X|^r) \le \liminf_{n\to\infty} E(|X_n|^r)$$

(b) If $X_n \xrightarrow{P} X$, or $X_n \Rightarrow X$, then the above also holds.

pf: (a) Suppose $X_n \xrightarrow{a.s.} X$,

$$E(|X|^r) = E(\liminf_n |X_n|^r)$$

$$\le \liminf_n E(|X_n|^r) \quad \text{by Fatou's lemma: } |X_n|^r \ge 0.$$

(b) Suppose $X_n \xrightarrow{P} X$, hence $X_n \xrightarrow{d} X$.

By Skorohod's thm, we know that

there exist $Y_n, n \ge 1, Y$ s.t. $Y_n \overset{d}{=} X_n, n \ge 1$, $Y \overset{d}{=} X$, $Y_n \xrightarrow{a.s.} Y$.

From (a) we get that for every $r > 0$,

$$E(|Y|^r) \le \liminf_{n\to\infty} E(|Y_n|^r)$$

$\because$ $Y_n \overset{d}{=} X_n, n \ge 1$, $Y \overset{d}{=} X$,

$\therefore$ $E(|X|^r) \le \liminf_{n\to\infty} E(|X_n|^r)$.

#

- <u>Theorem</u> (DCT and Lebesgue Dominated Convergence)

(a) Dominated Convergence Thm (DCT)

If $X_n \xrightarrow{a.s.} X$ and $\exists$ a dominating r.v. Z s.t. $|X_n| \le Z \in L_1$, then

$E(X_n) \to E(X)$

(b) Lebesgue Dominated Convergence Thm

If $X_n \xrightarrow{P} X$ and $\exists$ a dominating r.v. Z s.t. $|X_n| \le Z \in L_1$, then

$E(X_n) \to E(X)$

(c) If $X_n \Rightarrow X$ and $\exists$ a dominating r.v. Z s.t. $|X_n| \le Z \in L_1$, then

$E(X_n) \to E(X)$

pf: (a) If $|X_n| \le Z$, then $-Z \le X_n \le Z$, and $Z \in L_1$, $-Z \in L_1$.

$\because X_n \xrightarrow{a.s.} X \quad \therefore \lim_{n\to\infty} X_n = \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n$

The Fatou's Lemma implies that

$$E(X) = E(\liminf_{n\to\infty} X_n) \le \liminf_{n\to\infty} E(X_n)$$

$$\le \limsup_{n\to\infty} E(X_n) \le E(\limsup_{n\to\infty} X_n) = E(X).$$

Thus all inequalities are equality.

(b) It suffices to show that

every convergent subsequence $\{E(X_{n_k})\}$ converges to $E(X)$.

Suppose $E(X_{n_k})$ converges. Then since $X_n \xrightarrow{P} X$,

$\{X_{n_k}\}$ contains an a.s. convergent subsequence $\{X_{n_{k(i)}}\}$ s.t. $X_{n_{k(i)}} \xrightarrow{a.s.} X$, $i \to \infty$.

Then DCT implies $E(X_{n_{k(i)}}) \to E(X)$. So $E(X_n) \to E(X)$.

(c) Since $\exists$ r.v. Z s.t. $|X_n| \le Z \in L_1 \Rightarrow \{X_n\}$ is ui

Since $X_n \Rightarrow X$, by Skorohod's theorem,

$\exists\, Y_n, n \ge 1, Y$ with $Y_n \stackrel{d}{=} X_n$, $Y \stackrel{d}{=} X$ s.t. $Y_n \xrightarrow{a.s.} Y$, thus $Y_n \xrightarrow{P} Y$.

$\{X_n\}$ is ui $\Rightarrow \{Y_n\}$ is ui by $Y_n \stackrel{d}{=} X_n$

$\because Y_n \xrightarrow{P} Y$ and $\{Y_n\}$ is ui $\Leftrightarrow Y_n \xrightarrow{L_1} Y$ (Result)

$\therefore E(X_n) = E(Y_n) \to E(Y) = E(X)$. #

(c) If X, Y are random variables with distribution functions $F(x), G(x)$ which have no common discontinuities, then

$$E(F(Y)) + E(G(X)) = 1.$$

Interpret the sum of expectations on the left as a probability.

(d) Even if F and G have common jumps, if $X \perp\!\!\!\perp Y$ then

$$E(F(Y)) + E(G(X)) = 1 + P[X = Y].$$

6. Suppose $X \in L_1$ and A and A_n are events.

✓ (a) Show

$$\int_{[|X|>n]} X dP \to 0.$$

✓ (b) Show that if $P(A_n) \to 0$, then

$$\int_{A_n} X dP \to 0.$$

Hint: Decompose

$$\int_{A_n} |X| dP = \int_{A_n[|X|\le M]} |X| dP + \int_{A_n[|X|>M]} |X| dP$$

for large M.

✓ (c) Show $E(|X|) = \int_{\Omega} |X| dP = 0$ iff $P[|X|>0] = 0$, i.e. $|X| = 0$ a.s.

$$\int_A |X| dP = 0 \text{ iff } P(A \cap [|X| > 0]) = 0.$$

(d) If $X \in L_2$, show $\mathrm{Var}(X) = 0$ implies $P[X = E(X)] = 1$ so that X is equal to a constant with probability 1.

(e) Suppose that $(\Omega, \mathcal{B}, P)$ is a probability space and $A_i \in \mathcal{B}$, $i = 1, 2$. Define the distance $d : \mathcal{B} \times \mathcal{B} \mapsto \mathbb{R}$ by

$$d(A_1, A_2) = P(A_1 \Delta A_2).$$

Check the following continuity result: If $A_n, A \in \mathcal{B}$ and

$$d(A_n, A) \to 0$$

then

$$\int_{A_n} X dP \to \int_A X dP$$

so that the map

$$A \mapsto \int_A X dP$$

is continuous.

Ex 5-6

Pf: (a) $\int_{[|X|>n]} X\,dP = E(X 1_{[|X|>n]})$

$|E(X 1_{[|X|>n]})| \le E(|X| 1_{[|X|>n]})$

$\because |X| 1_{[|X|>n]} \downarrow |X| \cdot 0 = 0$ as $n \to \infty$

and $|X| 1_{[|X|>n]} \le |X| \in L_1$,

$\therefore$ by DCT

$|E(X 1_{[|X|>n]})| \le E(|X| 1_{[|X|>n]}) \to 0$ as $n \to \infty$

Hence

$E(X 1_{[|X|>n]}) \to 0$ as $n \to \infty$

(b) $\int_{A_n} |X|\,dP$

$= E(|X| 1_{A_n \cap [|X| \le M]}) + E(|X| 1_{A_n \cap [|X|>M]})$

$\le M P(A_n) + E(|X| 1_{[|X|>M]})$

$\to 0 + E(|X| 1_{[|X|>M]})$ as $n \to \infty$

$\to 0$ as $M \to \infty$ from (a)

(c) ① show $P(A \cap [|X|>0]) = 0 \Rightarrow E(|X| 1_A) = 0$

$E(|X| 1_A) = E(|X| 1_{A \cap [|X|>0]}) + E(\overset{0}{\cancel{|X|}} 1_{A \cap [|X|=0]})$

$= E(|X| 1_{A \cap [|X|>0]})$

$\because 1_{A \cap [|X| \ge \frac{1}{n}]} \uparrow 1_{A \cap [|X|>0]}$

and $|1_{A \cap [|X| \ge \frac{1}{n}]}| \le 1$,

$\therefore P(A \cap [|X| \ge \frac{1}{n}]) = E(1_{A \cap [|X| \ge \frac{1}{n}]}) \to E(1_{A \cap [|X|>0]}) = P(A \cap [|X|>0]) = 0$ by DCT

Hence

$\int_{A \cap [|X| \ge \frac{1}{n}]} |X|\,dP = E(|X| 1_{A \cap [|X| \ge \frac{1}{n}]}) \to E(|X| 1_{A \cap [|X|>0]}) = 0$ from (b).

② show $E(|X| 1_A) = 0 \Rightarrow P(A \cap [|X|>0]) = 0$

$\because \frac{1}{n} 1_{A \cap [|X| \ge \frac{1}{n}]} \le |X| 1_{A \cap [|X| \ge \frac{1}{n}]} \le |X| 1_A$

$\therefore \frac{1}{n} P(A \cap [|X| \ge \frac{1}{n}]) \le E(|X| 1_A) = 0 \Rightarrow P(A \cap [|X| \ge \frac{1}{n}]) = 0$.

$\because 1_{A \cap [|X| \ge \frac{1}{n}]} \uparrow 1_{A \cap [|X|>0]}$ & $|1_{A \cap [|X| \ge \frac{1}{n}]}| \le 1$

$\therefore$ by DCT, $0 = P(A \cap [|X| \ge \frac{1}{n}]) \to P(A \cap [|X|>0]) = 0$.

#

- <u>Theorem</u> (Fubini's Thm).

If $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ are probability spaces,
let $\pi(A \times B) = \mu(A) \cdot \nu(B)$ be the product measure
for measurable rectangles $A \times B$, $A \in \mathcal{X}$, $B \in \mathcal{Y}$. Then

(1) for nonnegative f, the functions

(*) $\int_Y f(x,y)\,\nu(dy)$, $\int_X f(x,y)\,\mu(dx)$ are measurable $\mathcal{X}, \mathcal{Y}$, respectively;

(**) $\int_{X \times Y} f(x,y)\,\pi(d(x,y)) = \int_X \left[\int_Y f(x,y)\,\nu(dy)\right] \mu(dx)$

$$= \int_Y \left[\int_X f(x,y)\,\mu(dx)\right] \nu(dy)$$

(2) If f (not necessarily nonnegative) is integrable w.r.t. π, the product measure, then (*) and (**) again hold.

Remark:

- Fubini's thm implies that double integrals can be calculated as the iterated integrals.
- Application of Fubini's thm:
  usually follows a two step procedure,

  ① check $\int_X \left[\int_Y |f(x,y)|\,\nu(dy)\right] \mu(dx)$ or $\int_Y \left[\int_X |f(x,y)|\,\mu(dx)\right] \nu(dy) < \infty$

  ② If the result of ① is finite,
  then the double integral (w.r.t. π) of $|f(x,y)|$ must be finite,
  so that f is integrable w.r.t. π;
  then the value of the double integral of f is found by computing one of the iterated integrals of f;
  If the result of ① is infinite,
  f is not integrable π.