The Radon Nikodyn Theorem

(in essense tells you why $P(A) = \int_A f(x)\,dx$ or $P(A) = \sum_{x \in A} p(x)$ and why the distinction b/w discrete & continuous is artificial)

Example : $X \geq 0$ meas funct^n

μ is a finite measure

Then $\Phi(A) = \int_A X \cdot d\mu$ is a measure on $(\Omega, \mathcal{K})$ where $A \in \mathcal{K}$

since $\Phi(\emptyset) = 0$

$$\Phi(A \cup B) = \int_{A\cup B} X \cdot d\mu = \int_A X\,d\mu + \int_B X \cdot d\mu = \Phi(A) + \Phi(B)$$

A & B are disjoint

we can define $\Phi(A) = \int_A f(x)\,d\mu(x)$

DEF : Let Φ and μ be 2 meas on $\mathcal{A}$.

Φ is called an <u>absolutely continuous meas</u> w.r.t μ (denoted by $\Phi << \mu$)

if $\mu(A) = 0 \Rightarrow \Phi(A) = 0 \quad \forall A \in \mathcal{A}$

[μ is called a <u>dominating measure</u>]

{ <u>Imp</u> Say we have a nice meas & Φ is similar meas to μ then all our calculations simplify }

DEF : Φ and μ are 2 meas on $\mathcal{A}$

Φ is called <u>singular to μ</u> (denoted by $\Phi \perp \mu$)

if $\exists$ an $N \in \mathcal{A}$ st (i) $\mu(N) = 0$ and (ii) $\Phi(N^c) = 0$

$\boxed{\text{THEOREM}}$: <u>Lebesgue Decomposition Theorem</u>

Let μ and Φ be 2 σ-finite measures on $(\Omega, \mathcal{A})$. Then $\exists$ a unique decomposition of Φ as

(1) $\underline{\Phi = \Phi_{ac} + \Phi_s}$

$$\text{where} \quad \Phi_{ac} << \mu$$
$$\Phi_s \perp \mu$$

(2) Also $\exists$ a funct^n $\underline{Z_0 \geq 0}$, <u>finite</u> & <u>unique a.e wrt μ</u>

st $\underline{\Phi_{ac}(A)} = \underline{\int_A Z_0 \, d\mu}$ $\forall$ $A \in \mathcal{A}$

[The problem working with the unknown meas Φ reduces to a part with a familiar meas μ through Φ_{ac} and some remainder part]

<u>Proof</u>: <u>Case 1</u> $\Phi << \mu$ (assume given)

Define $\Phi_{ac} \equiv \Phi$ and $\Phi_s \equiv 0$

then $\Phi = \Phi_{ac} + \Phi_s$

If we define $N \equiv \Omega^c = \phi$ then $\mu(N) = 0$ & $\Phi_s(\Omega) = 0$

But we do not know if Φ_{ac} can be written as an integral

- Idea is to find a Z_0 and define $\Phi_{ac}(A) = \int_A Z_0 \, d\mu$

  then $\mu(A)=0 \Rightarrow \Phi_{ac}(A)=0$ also {by absolute continuity of the integral}

  Then define $\Phi_s(A) = \Phi(A) \dot{-} \Phi_{ac}(A)$

  and we can show $\Phi_s(A)=0 \quad \forall A \in \mathcal{A}$ {lecture 14}

- Define $\mathcal{Z} = \{ Z : Z \geq 0, Z \in \mathcal{L}_1(\mu)$ and $\int_A Z \, d\mu \leq \Phi(A), \forall A \in \mathcal{A} \}$

{Note $0 \in$ the set since $\Phi(A) \geq 0$ always because $\Phi(A)$ is also a meas}

(i) $0 \in \mathcal{Z} \Rightarrow \mathcal{Z} \neq \phi$

(ii) $\mathcal{Z}$ is a lattice (if $Z_1 \in \mathcal{Z}$ & $Z_2 \in \mathcal{Z}$ then $\max(Z_1, Z_2) \in \mathcal{Z}$ also)

Let $A_1 \equiv \{ \omega \in A \mid Z_1(\omega) > Z_2(\omega) \}$

$A_2 \equiv A \setminus A_2$

Now $$\int_A (Z_1 \vee Z_2) \, d\mu = \int_{A_1} Z_1 \, d\mu + \int_{A_2} Z_2 \, d\mu$$

$$\leq \Phi(A_1) + \Phi(A_2)$$ {since $Z_1, Z_2 \in \mathcal{Z}$ & by how $\mathcal{Z}$ is defined}

$$= \Phi(A)$$ $\{\because A_1 \cup A_2 = A\}$

$$\therefore \int_A (Z_1 \vee Z_2) \, d\mu \leq \Phi(A) \Rightarrow Z_1 \vee Z_2 \in \mathcal{Z}$$

- Let $V_n = Z_1 \vee Z_2 \vee \cdots \vee Z_n \in \mathcal{Z}$

  V_n is $\uparrow$ seq

- Define $c \equiv \sup_{z \in \mathcal{Z}} \int_{\Omega} z \cdot d\mu \leq \sup_{z \in \mathcal{Z}} \Phi(\Omega)$ $\{$ by def of $\mathcal{Z}$

$< \infty$

- We now construct a seq of nos which converge to c (always possible)
Choose $\{V_n\}_{n \geq 1} \in \mathcal{Z}$ st $\int_{\Omega} V_n \cdot d\mu \xrightarrow[n \to \infty]{} C$

- Let $z_0 \equiv \lim_{n \to \infty} V_n$ (since $V_n \uparrow$ it has a limit; finite/∞)

Hence z_0 is unique a.e and $z_0 \geq 0$ $\{$ by construction $\}$

- We need to show $z_0 \in \mathcal{Z}$ [eg $\{1, \frac{1}{2}, \frac{1}{3}, \cdots \frac{1}{n} \cdots\}$ has lim $= 0$ which does not $\in$ to the set]

Consider $\int_A z_0 \cdot d\mu = \int_A \lim_n V_n \, d\mu$

$= \lim_{n \to \infty} \int_A V_n \cdot d\mu$ (by MCT)

$\leq \lim_{n \to \infty} \Phi(A)$ [by construction]

$= \Phi(A)$

ie $\int_A z_0 \cdot d\mu \leq \Phi(A)$

$\therefore$ $z_0 \in \mathcal{Z}$

- $\int_{\Omega} z_0 \, d\mu = \lim_{n \to \infty} \int_{\Omega} V_n \, d\mu = C < \infty$ $\{$by the way we defined $C\}$

$\Rightarrow Z_0$ is finite

- Now define $\Phi_{ac}(A) = \int_A Z_0 \cdot d\mu$

$\Phi_{ac} << \mu$

- For this Z_0 we can further show that (we will not prove)

$$\Phi_S(A) = \Phi(A) - \Phi_{ac}(A) = 0 \quad \forall A \in \mathcal{A}$$

QED case 1

Case 2 $\quad \Phi$ is any general measure

- Define $\quad \nu = \Phi + \mu$

We always have $\Phi << \nu$ and $\mu << \nu$

given $\nu(A) = 0 \Rightarrow \Phi(A) + \mu(A) = 0$

$\Rightarrow \Phi(A) = 0$ and $\mu(A) = 0$ $\left\{ \because \begin{array}{l} \Phi(A) \geq 0 \\ \mu(A) \geq 0 \end{array} \text{ always} \right\}$

$\Rightarrow \mu << \nu$ and $\Phi << \nu$

$\therefore$ By case 1, $\exists$ $X_0 \geq 0$, $Y_0 \geq 0$, finite and unique a.e ν st

$\Phi(A) = \int_A X_0 \cdot d\nu$ and $\mu(A) = \int_A Y_0 \cdot d\nu$

Define $D = \{\omega \mid Y_0(\omega) = 0\}$ and so $D^c = \{\omega \mid Y_0(\omega) > 0\}$

Define $\Phi_{ac}(A) = \Phi(AD^c)$ and $\Phi_S(A) = \Phi(AD)$

$$\Phi_{ac}(A) + \Phi_s(A) = \Phi(AD) + \Phi(AD^c) = \Phi(A)$$

<u>Check</u> $\Phi_{ac} << \mu$ and $\Phi_s \perp \mu$

Summary of Results

The Radon - Nikodyn Theorem (for finite measures)

Let Φ and μ be 2 finite measures on $(\Omega, \mathcal{A})$

$\Phi << \mu \iff \exists$ an a.e μ unique, μ integrable functⁿ $Z_0 \geq 0$ and finite valued st

$$\Phi(A) = \int_A Z_0 \, d\mu \qquad \forall \quad A \in \mathcal{A}$$

Terminology : The function Z_0 is called the Radon - Nicodyn derivative of Φ wrt μ and is denoted by

$$Z_0 = \frac{d\Phi}{d\mu} \quad \text{OR} \quad \frac{\partial \Phi}{\partial \mu}$$

So $$\Phi(A) = \int_A \frac{d\Phi}{d\mu} \, d\mu$$ $\left\{ \begin{array}{l} \text{similar to} \\ f(x) = \int f'(x) \end{array} \right.$

Note The R-N derivative of Φ wrt μ (of measures) is a function

LINK B/W RADON-NICODYN DERIVATIVES & ORDINARY DERIVATIVE

DEF : A functⁿ $F : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ is called absolutely continuous if $\forall\ \varepsilon > 0, \exists\ \delta_\varepsilon$ st $\sum_{k=1}^{n} |d_k - c_k| < \delta_\varepsilon \Rightarrow \sum_{k=1}^{n} |F(d_k) - F(c_k)| < \varepsilon$

$(\forall \ k \in I)$

Note Abs continuity $\Rightarrow$ continuity (converse need not hold)

Eg: say $|d_k - c_k| \le \frac{1}{k^2}$ $\overset{\text{say}}{\Longrightarrow}$ $|F(d_k) - F(c_k)| \le 1/k$

$$\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty \qquad\qquad \sum_{k=1}^{\infty} 1/k = \infty$$

Example : Any Lipschitz funct^n on I is absolutely continuous

Note: f is a Lipschitz funct^n if $\exists$ an M st $|f(x) - f(y)| \le M|x-y|$

- f is differentiable with bdd derivatives it is Lipschitz

$$f(x) - f(y) = f'(x^*)\,(x-y) \qquad \text{for } x^* \in [x,y]$$

$$\Rightarrow |f(x) - f(y)| = |f'(x^*)|\,|x-y|$$

$$\le M|x-y|$$

So Differentiable + Bdd derivatives $\Rightarrow$ Lipschitz $\Rightarrow$ Abs Cont $\Rightarrow$ continuous

THEOREM The Fundamental Theorem of Calculus

Let F be an absolutely continuous funct^n on $[a,b]$ {need not be diffrentiab

Let λ be the Lebesgue measure on $[a,b]$ then

1) The derivative of F (denoted by F') exists a.e λ on $[a,b]$ and is integrable wrt λ

$$\boxed{F(x) - F(a) = \int_a^x F'(y)\,dy}$$

{ Note $\int d\lambda(y) \equiv \int dy$ when λ is the Lebesgue meas }

2) Also if we can write

$$\underline{F(x) - F(a)} = \underline{\int_a^x f(y)\, d\lambda(y)}$$ for $\underline{\text{some } f}$ that is

$\underline{\text{integrable wrt } \lambda}$ on $[a,b]$ then $\underline{F' = f}$ a.e λ on $[a,b]$

↳ In addition F is abs continuous and of Bounded Variation.

In essence :

F is abs cont $\Longleftrightarrow$ F is the integral of its derivative

$\boxed{\text{Proposition}}$ (Link b/w R-N and ordinary derivatives) — No proof

Let F be an abs continuous function on $[a,b]$; F is ↑

and λ is the Lebesgue meas on $[a,b]$

Let μ_F be the associated L-S meas $[\mu_F([a,b]) = F(b) - F(a)$ by the correspondence theorem $]$

then

1) $\underline{\mu_F << \lambda}$ for any absolutey cont F ↑ & μ_F

2) The R-N derivative of μ_F wrt λ (denoted by $\underline{f = \frac{d\mu_F}{d\lambda}}$)

satisfies $\underline{\mu_F((a,x)) = F(x) - F(a)} = \underline{\int_a^x f\, d\lambda}$ and

$$\underline{F' = f \quad \text{a.e } \lambda}$$

[The R-N derivative of μ_F wrt λ is the derivative F' of F]

Statement : $\mu_F << \lambda \iff F$ is absolutely continuous

Assume Φ or μ_F is abs continuous wrt λ it means they want to use the Lebesgue measure to obt

$$P(A) = \int_A f(x)\, d\lambda(x)$$

$$\mu_F((a,x)) = F(x) - F(a) = \int_a^x f(x)\, d\lambda(x)$$

Again we need to select a F which is a modelling assumption. ...

Proposition

Let X be a r.v with df F_X (which is abs continuous)

Assume $F_X << \lambda$, λ is the lebesgue measure.

Let f_X be the density of F_X

Let $Y = g(X)$, for a given g st g^{-1} is $\uparrow$ and absolutely continuous

Then

$$F_Y << \lambda$$

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy}(g^{-1}(y))$$

[Note : Here g^{-1} is $\uparrow$ and so it has a derivative a.e. λ]

Proof : $F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$

$$= P(X \leq g^{-1}(y))$$

$$= F_X(g^{-1}(y))$$

$$\therefore \quad F_Y'(y) = F_X'(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

To now show f_Y is the RN derivative of F_Y

By the fundamental theorem of calculus.

$$F_Y(b) - F_Y(a) = \int_a^b F_Y'(y) \, d\lambda(y)$$

{ F_Y is abs cont since $F_Y = F_X \circ g^{-1}$

ie $F_Y((a,b]) = \int_a^b F_Y'(y) \, dy$ (by correspondence theorem.

By the R-N theorem: $F_Y << \lambda$ & the RN derivative of F_Y wrt λ,

$$\frac{dF_Y}{d\lambda} = F_X'(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

Also the previous result tells us $\dfrac{dF_Y}{d\lambda} = f_Y = F_Y'$

(Proposition of link b/w RN & the ordinary derv.)

$\boxed{\text{RESULT:}}$ Let $X : (\Omega, \mathcal{K}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, \mu)$

f is any meas funct^n $: \mathbb{R} \to \mathbb{R}$

Theorem of the <u>unconscious statistician</u> (Change of variable theorem

$$\int_{\Omega} \underline{f(X(\omega))\, dP(\omega)} = \int_{-\infty}^{\infty} \underline{f(t)\, d\mu(t)} \qquad \text{—— } (*)$$

$$\left[\text{ie } E_P(f(X)) = E_\mu(f)\right]$$

$\boxed{\text{THEOREM:}}$ <u>(Change of Variable)</u>

<u>λ</u> is a <u>lebesgue meas</u> on $\mathbb{R}$ & μ defined as above.

Assume <u>$\mu << \lambda$</u>

Then
$$\underline{\int_{-\infty}^{\infty} f(t)\, d\mu(t) = \int_{-\infty}^{\infty} f(t) \left(\frac{d\mu(t)}{d\lambda}\right) d\lambda(t)} \qquad \text{—— } (**)$$

All you are doing is changing the measures.

$(*)$ & $(**)$ $\Rightarrow$ $\displaystyle\int_{\Omega} f(X(\omega))\, dP(\omega) = \int_{-\infty}^{\infty} f(t) g(t)\, d\lambda(t)$ $\left\{ \text{where } g(t) = \dfrac{d\mu(t)}{d\lambda} \right.$

ie $\displaystyle E(f(X)) = \int_{-\infty}^{\infty} f(t)\, g(t)\, dt$

[Here the density is on $(\mathbb{R}, \mathcal{B}, \mu)$ ie $F_X \equiv P \circ X^{-1}$ and the RN derivative $\frac{d\mu}{d\lambda}$ where μ is the meas associated with F_X]

NOTE: There are dominating measures for discrete r.v also namely a meas called the counting measure

Ex Let λ_0 be the lebesgue meas on $[0,1]$

Let P_1 be the Pois(γ) distn on $\mathbb{N}$, for some $\gamma > 0$.

$$P \equiv \frac{\lambda_0 + P_1}{2}$$

λ is the lebesgue meas on $\mathbb{R}$

Find the lebesgue decomp of P wrt λ

ie obt P_{ac} & P_s st $P = P_{ac} + P_s$ &

& $P_{ac} << \lambda$

& $P_s \perp \lambda$

Sol

Let $P_{ac} \equiv \frac{\lambda_0}{2}$

$P_s = \frac{P_1}{2}$

$$\lambda_0(A) = \lambda(A \cap [0,1]) \quad \forall A \in \mathcal{B}$$

$$P_{ac}(A) = \int_{A \cap [0,1]} \frac{1}{2} \, d\lambda(t) \quad \xRightarrow{\text{RN theorem}} \quad P_{ac} << \lambda$$

$$P_s(A) = \frac{1}{2} P_1(A) = \frac{1}{2} \sum_{k \in A} \frac{e^{-\gamma} \gamma^k}{k!} \qquad \forall A \in \mathcal{B}$$

Now for $A = \mathbb{N}$, $\lambda(\mathbb{N}) = 0$ and

$$P_s(\mathbb{N}^c) = 0$$ $\left\{ \text{since } P_s \equiv \frac{P_1}{2} \text{ and } P_1 \text{ is a poisson r.v which is only defined on } \mathbb{N} \right.$

So $P_s \perp \lambda$

$\boxed{\text{PRODUCT MEASURES}}$ (Important for Midterms)

In essence : To prove you can change the order of integration

DEF : $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ are 2 measurable spaces

$\mathcal{A} \times \mathcal{A}' \equiv \sigma[\mathcal{F}]$ is the <u>product field</u>

$$\mathcal{F} \equiv \Big\{ \sum_{i=1}^{m} A_i \times A_i' \ ; \ m \geq 1 \ , \ A_i \in \mathcal{A} \ , \ A_i' \in \mathcal{A}' \Big\}$$

disjoint union

$(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}')$ is called the <u>product space</u>

$\boxed{\text{THEOREM :}}$ <u>Existance of Product Measures</u>

$(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \nu)$ are 2 measure spaces

Define Φ on the field $\mathcal{F}$ st

$$\Phi \Big(\sum_{i=1}^{m} A_i \times A_i' \Big) = \sum_{i=1}^{m} \mu(A_i)\, \nu(A_i')$$

Then 1) Φ is well defined (say $\sum_{i=1}^{n} B_i \times B_i' = B = \sum_{i=1}^{m} A_i \times A_i'$ both give same $\Phi(B)$)

Φ is a σ-finite meas on $\mathcal{F}$

2) Φ extends uniquely to a σ-finite meas on $(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}')$

The extension, still denoted by Φ is called the product measure.

<u>Example</u> : $(\mathbb{R}, \mathcal{B}, \lambda)$

$$(\mathbb{R}^n, \mathcal{B}_n, \lambda_n) = \prod_{i=1}^{n} (\mathbb{R}, \mathcal{B}, \lambda)$$

$\mathcal{B}_n = \sigma[\mathcal{U}_n]$ where $\mathcal{U}_n$ = all disjoint unions of open sets in $\mathbb{R}^n$.

λ_n is called (and still denoted by λ) the product meas on $\mathbb{R}^n$

$\lambda([0,2]^n) = \lambda([0,2]) \times \lambda([0,2]) \times \cdots \lambda([0,2])$ {n times

$= 2^n$

[volume of an n-dimensional cube]

We need product measures to define n-dimensional integrals

eg $\iint f(x,y)\, d\lambda(x,y) = \iint f(x,y)\, dx\, dy$.

DEF : X on $\Omega \times \Omega'$

a) For each fixed $\omega \in \Omega$, the function $X_\omega(\omega') = X(\omega, \omega')$ is called the $\underline{\omega\text{-section of } X}$

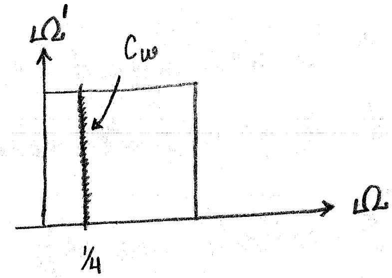
(ie it is now a function of only one variable)

b) Let $C \subseteq \Omega \times \Omega'$

$C_\omega = \{\omega' \in \Omega' \mid (\omega, \omega') \in C\}$ is called the ω-section of C

eg $C \subseteq \Omega \times \Omega' \equiv$ unit square

$\omega = 1/4$



The ω section of C is the line C_ω

THEOREM:

$(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \nu)$ are 2 measure spaces

1) $C_{\omega'} \in \mathcal{A}$, $C_\omega \in \mathcal{A}'$ (ie they are measurable)

2) Let $\Phi \equiv \mu \times \nu$ be the product measure

$$\iint_C d\Phi = \Phi(C) = \int_{\Omega'} \mu(C_{\omega'})\, d\nu(\omega') = \int_{\Omega} \nu(C_\omega)\, d\mu(\omega) \quad \forall\, C \in \mathcal{A} \times \mathcal{A}'$$

(ie it is like tak (i) " integrate over all possible slices)

Proof : all $C \in \mathcal{A} \times \mathcal{A}'$ for which (2)

least the empty set belongs]

To show Proof for ① By the way $\mathcal{M}$ & $\mathcal{C}_\omega$ are defined ?

$\mathcal{A} \times \mathcal{A}' \equiv \sigma[\mathcal{F}] \quad \left\{ \mathcal{F} = \left\{ \sum_{i=1}^{m} A_i \times A_i' ,\ A_i \in \mathcal{A} ,\ A_i' \in \mathcal{A}' ,\ m \geq 1 \right\} \right.$ —— (*)

By proposition 1.1.6 ⊛ holds if $\mathcal{M}$ is a monotone class.

Let $C_n \in \mathcal{M}$ and $C_n \uparrow C$ then we need to show $C \equiv \lim_{n\to\infty} C_n \in \mathcal{M}$

ie to show $\Phi(C) = \int_{\Omega'} \mu(C_{\omega'})\, d\nu(\omega')$

We know $C_n \in \mathcal{M} \Rightarrow \Phi(C_n) = \int_{\Omega'} \mu(C_{n,\omega'})\, d\nu(\omega')$

(i) Now $C_n \uparrow C \Rightarrow I_{C_n}(\omega, \omega') \uparrow I_C(\omega, \omega')$

$\rightarrow I_{C_{n,\omega'}}(\omega) \uparrow I_{C,\omega'}(\omega)$ —— (**)

C_n's are measurable sets $\Rightarrow C$ is also measurable set

(ii) ∴ I_{C_n} and I_C are measurable functions (simple funct^n)

(iii) Also every section of the measurable sets are measurable and so the functions I_{C_n} and I_C have measurable sections.

Let $h_n(\omega') \equiv \mu(C_{n,\omega'}) = \int_{\Omega} I_{C_{n,\omega'}}(\omega)\, d\mu(\omega)$ —— (***)

∴ $h_n(\omega') \uparrow \int_{\Omega} \lim I_{C_{n,\omega'}}(\omega)\, d\mu(\omega) = \int_{\Omega} I_{C,\omega'}(\omega)\, d\mu(\omega) \equiv h(\omega')$

ie $\quad h_n(\omega') \uparrow h(\omega') \quad$ ie $\quad \mu(C_{n,\omega'}) \uparrow \mu(C_{\omega'})$ —(****)

$$\left.\begin{array}{l} 0 \leq \Phi(C_n) \leq \mu(\Omega)\,\nu(\Omega') < \infty \\ C_n \uparrow \;\Rightarrow\; \Phi(C_n) \uparrow \end{array}\right\} \Rightarrow \Phi(C_n) \text{ converges. and so the limit exists}$$

$$\left\{ \begin{array}{l} \Phi(C_n) = \int_{\Omega'} \mu(C_{n,\omega'})\, d\nu(\omega') \leq \int_{\Omega'} \mu(\Omega)\, d\nu(\omega') \\ \qquad = \mu(\Omega) \int_{\Omega'} d\nu(\omega') = \mu(\Omega)\,\nu(\Omega') \end{array} \right\}$$

Now

$$\Phi(C) \stackrel{def}{\equiv} \lim_{n\to\infty} \Phi(C_n) = \lim_{n\to\infty} \int_{\Omega'} \underbrace{\mu(C_{n,\omega'})}_{h_n(\omega')}\, d\nu(\omega') \quad \left\{ \begin{array}{l} \Phi(C) = \Phi(\lim C_n) = \lim \Phi(C_n) \\ \text{since } C_n \uparrow C \end{array} \right.$$

$$= \int_{\Omega'} \underbrace{\mu(C_{\omega'})}_{h(\omega)}\, d\nu(\omega') \quad \left\{ \text{by MCT using } (****) \right.$$

So $\Phi(C) = \int_{\Omega'} \mu(C_{\omega'})\, d\nu(\omega')$

$\Rightarrow \Phi(C) \in \mathcal{M}$ [since we can write it in the form of (2)]

COROLLARY $\boxed{\text{FUBINI'S THEOREM}}$

$(\Omega, \mathcal{A}, \mu)$ and $(\Omega', \mathcal{A}', \nu)$ are 2 σ-finite spaces

$\Phi \equiv \mu \times \nu$ is the product meas on $(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}')$

If $X : \Omega \times \Omega' \to \mathbb{R}$ is Φ integrable

Then

$$\int_{\Omega \times \Omega'} X(\omega, \omega')\, d\Phi(\omega, \omega') = \int \left[\int X(\omega, \omega')\, d\mu(\omega) \right] d\nu(\omega')$$

$\boxed{\text{corollary : Tonelli}}$

Let X be an $\mathcal{A} \times \mathcal{A}'$ meas funct$^{\underline{n}}$

If $$\int \left(\int |X| \, d\mu \right) d\nu < \infty$$

OR $$\int \left(\int |X| \, d\nu \right) d\mu < \infty \qquad \text{OR} \qquad X \geq 0$$

then
$$\int X \, d\Phi = \int \left[\int X(\omega, \omega') \, d\mu(\omega) \right] d\nu(\omega')$$
$$= \int \left(\int X(\omega, \omega') \, d\nu(\omega') \right) d\mu(\omega)$$

END

CHAPTER 7 : DISTRIBUTIONS & QUANTILE FUNCTIONS

RECALL: $\boxed{\text{Distribution Function}}$

$$X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}) \quad \text{is a r.v}$$

The distⁿ functⁿ of X,

$$F_X(x) = P(X \leq x)$$

st
1) $F_X \equiv F \uparrow$
2) Right continuous
3) $F(-\infty) = 0$ and $F(\infty) = 1$

$\Big\}$ — (*)

$\boxed{\text{PROPOSITION}}$: Let F be a function satisfying (*)
Then $\exists$ a probability space $(\Omega, \mathcal{A}, P)$ and a r.v
$X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ st the distⁿ functⁿ of X is F.
[denoted by $X \sim F$ or $X \cong F$]

Proof: Given a funct F st F is $\uparrow$, rt continuous & $F(-\infty) = 0$, $F(\infty) = 1$
Take $(\Omega, \mathcal{A}) \equiv (\mathbb{R}, \mathcal{B})$
By the correspondence theorem : $\exists$ a P on $\mathcal{B}$ st

$$P((a, b]) = F(b) - F(a)$$

So $P((-\infty, x]) = F(x) \quad \{\because F(-\infty) = 0$ — (1)

We want to construct a funct $X : (\mathbb{R}, \mathcal{B}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, ?)$

st $F_X(x) = F(x) \quad \forall x$ — (3)

where F is the given functⁿ and F_X is the distⁿ functⁿ of X ($F_X(x) = P(X \leq x) = P \circ X^{-1}((-\infty, x])$ — (2)

Finding X that satisfies ③ means finding X st

$$P((-\infty, x]) = P \circ X^{-1}((-\infty, x]) \qquad \{\text{by ① \& ②}$$

So we take $X(\omega) = \omega$ the identity functⁿ.

THEOREM: (The decomposition of a df F) – Proof not req

Any df F can be written as:

$$F = F_d + F_c$$

$$= F_d + F_s + F_{ac}$$

where (i) F_d is the discrete part – a step function

(ii) F_c is a continuous function

this is a break up of F_c {

(iii) $F_s \perp \lambda$ (the lebesgue meas)

(iv) $F_{ac}(x) = \int_{-\infty}^{x} f_{ac}(y)\,dy$ for some $f_{ac} \geq 0$, finite a.e λ.

[ie the meas associated with F_{ac} is $\ll \lambda$]

* Note: Any discrete meas which are $\perp$ wrt the lebesgue meas λ

* There are also continuous df which are $\perp$ wrt λ.

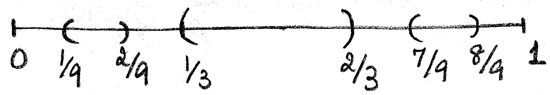
Consider the Normal df. It is abs continuous w.r.t λ (lebesgue)

ie $\Phi(A) = \int_A f(x)\,dx$

But it is not true that all continuous functions are a.c w.r.t λ (lebesgue)

** Example : Continuous df $\perp$ λ (Lebesgue meas)

Construct $F \perp \lambda$ ie find an N st $\lambda(N^c) = 0$ and $\mu_F(N) = 0$

Cantor set 

Take out $(1/3, 2/3)$ $(1/9, 2/9)$ $(7/9, 8/9)$ all middle thirds

$$C = \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}, \ a_n \in \{0,1\} \ \forall n \right\}$$

= all left over point

$\lambda(C) = 0$ { shown before

$F : [0,1] \longrightarrow [0,1]$

$$F(x) = \begin{cases} \sum_{n=1}^{\infty} \frac{a_n}{2^n} & \text{for } x \in C \\ \text{constant} & x \in C^c \end{cases}$$

Book Pg 109

This funct^n is $\uparrow$

$F(C) = [0,1]$ { any pt in $[0,1]$ can be written as a sum of an infinite series of $\frac{1}{2^n}$ ie diadic representation

image of C

$\Rightarrow$ F is continuous

$\because$ If a funct^n is $\uparrow$ and takes all the values in the codomain then the funct^n must be continuous – Darboux prop in real analysis.

Now take $N = C^c \Rightarrow \lambda(N^c) = \lambda(C) = 0$

$$\mu_F(C^c) = \sum_{I \equiv (a,b)} \mu_F(I) = \sum_{(a,b) \in C^c} \mu_F((a,b))$$

$$= \sum_{(a,b) \in C^c} F(b) - F(a)$$

$\left\{ \because F(a,b) = F(a,b] \right.$ since F is rt continuous

$$= \sum_{(a,b) \in C^c} \text{const} - \text{const}$$

$\{ \because$ by construction of F

$$= 0.$$

[Proof of this counter example is not req for the exam.]

DEF : For any df F we define the quantile function as

$$K(t) = \underbrace{F^{-1}(t)}_{} \equiv \inf \{ x : F(x) \geq t \} \quad \text{for any } t \in (0,1)$$

notation (does not necessarily mean inverse function

$\boxed{\text{Theorem 1}}$ The inverse transformation

Let $\xi \sim \text{Unif}(0,1)$.

Define $X \equiv K(\xi) \equiv F^{-1}(\xi)$ then

(1) $\{ \omega \in \Omega \mid X(\omega) \leq x \} = \{ \omega \in \Omega \mid \xi(\omega) \leq F(x) \}$

(2) $I_{[X \leq \cdot]} \equiv I_{[\xi \leq F(\cdot)]}$

(3) $X \equiv K(\xi) \equiv F^{-1}(\xi) \sim F$

Proof (i) To show "$\supseteq$"

Let $\omega_0 \in \Omega$ st $F(x) \geq \xi(\omega_0)$

We will show $X(\omega_0) \leq x$

Now $F(x) \geq \xi(\omega_0) \implies F^{-1}(\xi(\omega_0)) \leq x$ { Here $t = \xi(\omega_0) \in (0,1)$ Use def of $F^{-1}(t)$ and meaning of the infimum

$$\Updownarrow$$

$X(\omega_0) \leq x$ { by def of X

$\implies \omega_0 \in \{\omega \in \Omega \mid X(\omega) \leq x\}$

"$\supseteq$"

Let $\omega_0 \in \Omega$ be st $X(\omega_0) \leq x$

We will show $F(x) \geq \xi(\omega_0)$

Now $X(\omega_0) \equiv F^{-1}(\xi(\omega_0)) \leq x$

$\implies \inf\{y \mid F(y) \geq \xi(\omega_0)\} \leq x$ ——— (*)

$y_0 = \inf A$ when (i) $y_0 \leq y \quad \forall\, y \in A$

(ii) $\forall\, \varepsilon > 0 \quad \exists\, y_\varepsilon \in A$ st $y_\varepsilon - \varepsilon < y_0$

Let $A \equiv \{y \mid F(y) \geq \xi(\omega_0)\}$

$\inf A \equiv y_0$

$\implies y_0 \leq x$ { from (*)

We also know $\forall\, \varepsilon > 0$, $\exists$ a $y_\varepsilon \in A$ st $y_0 + \varepsilon \geq y_\varepsilon$ { from def of inf

Also $x \geq y_0$ [ie $x - y_0 \geq 0$]

$\implies x + \varepsilon \geq y_\varepsilon$

Now F is $\uparrow \implies F(x+\varepsilon) \geq F(y_\varepsilon)$

$\geq \xi_0$

$\Rightarrow \quad F(x+\varepsilon) \geqslant \xi$

$\Rightarrow \lim_{\varepsilon \to 0} F(x+\varepsilon) \geqslant \xi \quad \Rightarrow \quad F(x) \geqslant \xi$ $\quad \{$ F is rt continuou

(3) To show $X = F^{-1}(\xi) \sim F$

ie to show $P(X \leq t) = F(t)$

Now
$$\begin{aligned} P(X \leq t) &= P(\{\omega \in \Omega \mid X(\omega) \leq t\}) \\ &= P(\{\omega \in \Omega \mid \xi(\omega) \leq F(t)\}) \quad \text{by ①} \\ &= P(\xi \leq F(t)) \\ &= G(F(t)) \\ &= F(t) \end{aligned}$$

$\{$ where G is the dist^n funct^n of a Uniform r.v so $G(t) = t$

Counting Measure

$$\mu(A) = \# \text{ in } A \quad ; \text{ if } A \text{ is finite}$$
$$= \infty \quad ; \text{ ow}$$

eg $\Omega = [0,1]$

$\mu(\{0, 1/2\}) = 2$

$\mu([0, 1/2]) = \infty$

$\mu(\phi) = 0$

You can check that the counting meas is indeed a measure.

$\mu(\,) \longrightarrow [0, \infty)$

$\mu(\phi) = 0$

$\mu(\Sigma' A_i) = \Sigma \mu(A_i)$

You can show all discrete random variables are dominated by counting measures

$X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, \mu)$

$\mu \equiv P \circ X^{-1}$ is the meas induced by X on $\mathbb{R}$

Densities are the RN derivatives of the induced measures

ie $\dfrac{d(P \circ X^{-1})}{d\lambda}$

To find $P(X \in A) = P \circ X^{-1}(A)$ we need to assume a density on X say X ~ Normal and $P \circ X^{-1} << \lambda$ $\{ \lambda \equiv$ lebesgue meas$\}$

Then $\dfrac{d(P \circ X^{-1})}{d\lambda} = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

Now if $X \sim Bin(n,p)$

Let $A \equiv \{1\}$

$P(X \in A) = P \circ X^{-1}(A) = \binom{n}{1} p q^{n-1} \neq 0$

But $\lambda(A) = 0$

So we cannot define RN derivaties w.r.t the Lebesgue measure since $P \circ X^{-1} \not<< \lambda$

Now $\mu \equiv$ counting meas on $\mathbb{Z}$ = set of all integers

$$\mu(A) = \begin{cases} \# \text{ of integers in } A & ; \text{ if } A \text{ is finite} \\ \infty & ; \text{ ow} \end{cases}$$

eg $\mu(A) = 0$ when $A = [1/3, 1/2]$

and $P \circ X^{-1}(A) = 0$ also for $X \sim Bin(n,p)$

The induced measures of the Binomial and Poisson r.v. are a.c w.r.t the counting meas. μ on $\mathbb{Z}$

If $\mu(A) = 0 \Rightarrow A$ has no integers

$\Rightarrow P(X \in A) = 0$ for $X \sim$ Pois OR Bin

ie $P \circ X^{-1}(A) = 0$

ie $\mu(A)=0 \implies P \circ X^{-1}(A) = 0$

$\therefore (P \circ X^{-1}) << \mu$

So the RN derivative w.r.t μ (the counting meas on $\mathbb{Z}$) exists

$$\frac{d(P \circ X^{-1})}{d\mu} = \binom{n}{x} p^x q^{n-x} \qquad \text{for } X \sim \text{Bin}(n,p).$$

$$\therefore P \circ X^{-1}(A) = \int_A \binom{n}{x} p^x q^{n-x} \cdot d\mu \qquad \mu \equiv \text{counting meas on } \mathbb{Z}$$

eg $$\int_{A=[0,3]} f(x)\,d\mu = \int_{\{0\}} f(x)\,d\mu + \int_{(0,1)} f(x)\,d\mu + \int_{\{1\}} f(x)\,d\mu + \int_{(1,2)} + \int_{\{2\}} + \int_{(2,3)} + \int_{\{3\}}$$

$$= \int_{\{0\}} f(0)\,d\mu + 0 + \int_{\{1\}} f(1)\,d\mu + 0 + \int_{\{2\}} f(2)\,d\mu + 0 + \int_{\{3\}} f(3)\,d\mu$$

$$= f(0) + f(1) + f(2) + f(3)$$

Representation of Random Variables

$X \sim F$

① If $\xi \sim Unif(0,1) \Rightarrow F^{-1}(\xi) \sim F$

So $F^{-1}(\xi)$ is a representation of X

② We can represent X as integrals w.r.t "random measures"

- Let δ_a be the dirac measure

  ie $\delta_a(A) = \begin{cases} 1 & \text{if } a \in A \\ 0 & \text{ow} \end{cases}$

- Any function can be written as $f(a) = \int f(x) d\delta_a(x)$

So $X(\omega_0) = \cancel{\int x(x)\,\delta_{x(\omega_0)}(x)}$ $\quad \int x \, d\underbrace{\delta_{X(\omega_0)}}(x)$

is called a random measure

- $\delta_a \longleftrightarrow F(x) = I_{[a,\infty)}(x)$ by correspondence theorem.

So $\delta_{X(\omega)} \longleftrightarrow F(u) = I_{[X(\omega),\infty]}(u) = I_{[x \geq X(\omega)]}(u)$

So $X(\omega) = \int_{-\infty}^{\infty} x \, dI_{[x \geq X]}$

③ Now $X(\omega) = F^{-1}(\xi(\omega))$

$$= \int_0^1 F^{-1}(t)\, d\delta_{\xi(\omega)}(t)$$

$$\delta_{\xi(\omega)} \longleftrightarrow F = I_{[\xi(\omega),\infty)} = I_{[t \geq \xi(\omega)]}$$

$$\therefore \quad X(\omega) = \int_0^{1.} F^{-1}(t)\, dI_{[t \geq \xi]}(t)$$

To summarize

$$X = \int_{-\infty}^{\infty} x\, dI_{[x \geq X]}(x) = \int_0^1 F^{-1}(t)\, dI_{[t \geq \xi]}(t)$$

We can now obt diff ways to obt $E(X)$

1) $X \sim F$ then $\mu = \int x\, dF(x)$

2) $X \equiv F^{-1}(\xi)$ where $\xi \sim Unif(0,1)$

$$\mu = \int_0^1 F^{-1}(t)\, dt$$

{ just using usual transf concepts

3) $X = \int_0^1 F^{-1}(t)\, dI_{[t \geq \xi]}$ and $\mu = \int_0^1 F^{-1}(t)\, dt$

$$\therefore X - \mu = \int_0^1 F^{-1}(t)\, d\left(I_{[t \geq \xi]} - t\right)$$

$$= -\int_0^1 \left(I_{[t \geq \xi]} - t\right) dF^{-1}(t) \qquad \left\{ \begin{array}{l} \text{integration by parts} \\ u = F^{-1} \ \& \ v = \left(I_{[t \geq \xi]} - t\right) \end{array} \right.$$

Since

$$u_+(b)v(b) - u(a)v_-(a) = u_+(1)v(1) - u(0)v_-(0)$$

$$= F^{-1}(1)\left(I_{[\xi \leq 1]} - 1\right) - F^{-1}(0)\left(I_{[\xi \leq 0]} - 0\right)$$

$$= F^{-1}(1)\,(1-1) - F^{-1}(0)\,(0-0)$$

$$= 0 - 0$$

$$= 0$$

$$\therefore \int_a^b u\,dv + \int_a^b v\,du = 0$$

WEAK LAW OF LARGE NUMBERS (WLLN)

THEOREM Let $X_1, X_2 \ldots X_n$ be iid with mean μ.

Assume $Var(X_1) = \sigma^2 < \infty$ then

$\bar{X}_n \xrightarrow{P} \mu$ where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ { weak law of large # }

Proof $$P(|\bar{X}_n - \mu| > \varepsilon) \le \frac{Var(\bar{X}_n)}{\varepsilon^2}$$ (by Chebyshev

$$= \frac{\sigma^2}{n\varepsilon^2}$$

$$\longrightarrow 0 \quad \text{as } n \to \infty$$

Theorem: (Khinchin)

Let $X_1, X_2 \ldots X_n$ be iid with $X \sim F_X$ (mean $= \mu$). Assume $E|X_1| < \infty$ Then $\bar{X}_n \xrightarrow{P} \mu$

[Note if the variables itself were bdd then $Var(X) \le E(X^2) \le c^2$.]

Proof • Let $Y_{nk} = X_k \, I_{[-n \le X_k \le n]}$ for $1 \le k \le n$ and n. arbitrary fixed

ie Y_{nk}'s are X_k's truncated at n

• Let $\mu_n = E(Y_{nk})$

We will show $\bar{Y}_n \xrightarrow{P} \mu$ where $\bar{Y}_n = \frac{\sum_{k=1}^{n} Y_{nk}}{n}$

What we want is $\bar{X}_n \xrightarrow{P} \mu$

$\boxed{\text{Def}}$ <u>(Khinchin Equivalence)</u> Two sequences of r.v $\{X_n\}$ and $\{Y_n\}$ are called Khinchin equivalent if $\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$

<u>Step 3</u>
We will show $\{X_n\}$ and $\{Y_n\}$ are K-equivalent.

$\Downarrow$

<u>Step 4</u>
$\bar{X}_n$ and $\bar{Y}_n$ converge in P if one of them converges and they do so to the same limit.

<u>Step 1</u> Show $\bar{Y}_n - \mu_n \xrightarrow{P} 0$ (then we will show $Y_n \xrightarrow{P} \mu$)

$$P\left(|\bar{Y}_n - \mu_n| > \sqrt{\varepsilon}\right) \leq \frac{Var(\bar{Y}_n)}{\varepsilon} \quad \text{by Chebyshev's ineq.}$$

$$= \frac{Var(Y_{n_1})}{n\,\varepsilon}$$

$$\leq \frac{E(Y_{n_1}^2)}{n\,\varepsilon}$$

{ <u>Note</u> here $E(Y_{n_1}^2) \leq n^2$ so why don't we truncate at some constant? We cannot because we want Y_n & X_n to be K-equivalent

$$P\left(|\bar{Y}_n - \mu_n| > \sqrt{\varepsilon}\right) \leq \frac{1}{n\varepsilon} E\left[Y_{n_1}^2 I_{[|X_1| \leq \varepsilon\sqrt{n}]} + Y_{n_1}^2 I_{[|X_1| > \varepsilon\sqrt{n}]}\right]$$

$$P\left(|\bar{y}_n-\mu_n|>\varepsilon\right)\leq \frac{1}{n\varepsilon}\left\{ E\left(X_1^2 I_{[-n\leq X_1\leq n]} I_{[-\varepsilon\sqrt{n}\leq X_1\leq \varepsilon\sqrt{n}]}\right) + E\left(X_1^2 I_{[-n\leq X_1\leq n]} I_{[|X_1|>\varepsilon\sqrt{n}]}\right)\right\}$$

$$\leq \frac{1}{n\varepsilon}(\varepsilon\sqrt{n})^2 + \frac{1}{n\varepsilon} E\left(X_1^2 I_{[\varepsilon\sqrt{n}<|X_1|\leq n]}\right)$$

$$= \varepsilon + \frac{1}{n\varepsilon} E\left(X_1^2 I_{[\varepsilon\sqrt{n}<|X_1|\leq n]}\right)$$

$$\leq \varepsilon + \frac{1}{n\varepsilon} E\left(|X_1|\,|X_1|\, I_{[\varepsilon\sqrt{n}<|X_1|\leq n]}\right)$$

$$\leq \varepsilon + \frac{n}{n\varepsilon} E\left(|X_1|\, I_{[\varepsilon\sqrt{n}<|X_1|\leq n]}\right)$$

$$= \varepsilon + \frac{1}{\varepsilon}\int_{[\varepsilon\sqrt{n}<|X_1|\leq n]} |x|\, dF_X(x)$$

$$\leq \varepsilon + \frac{1}{\varepsilon}\int_{[|X_1|>\varepsilon\sqrt{n}]} |x|\, dF_X(x)$$

Now $P\left(|X_1|>\varepsilon\sqrt{n}\right) \leq \dfrac{E|X_1|}{\varepsilon\sqrt{n}}$ (Markov's Ineq

$\longrightarrow 0$ as $n\to\infty$ (since $E|X_1|<\infty$

$$\therefore \int_{[|X_1| > \varepsilon\sqrt{n}]} |x| \, dF_X(x) \longrightarrow 0$$ by absolute continuity of the integral

$$\therefore P\left(|\bar{Y}_n - \mu_n| > \sqrt{\varepsilon} \right) \leq \varepsilon + \varepsilon = 2\varepsilon$$

$$\therefore \bar{Y}_n - \mu_n \xrightarrow{P} 0$$

<u>Step 2</u> Show $\mu_n \longrightarrow \mu$

$$\mu_n = E\left(X_1 \, I_{[-n \leq X_1 \leq n]}\right)$$

$$\mu = E(X_1)$$

$$= E\left(X_1 \, I_{[-n \leq X_1 \leq n]} \right) + E\left(X_1 \, I_{[|X_1| > n]} \right)$$

$$= \mu_n + E\left(X_1 \, I_{[|X_1| > n]} \right)$$

$$E\left(X_1 \, I_{[|X_1| > n]} \right) = \int_{[|X_1| > n]} |x| \, dF_X(x) \longrightarrow 0$$

since $P(|X_1| > n) \leq \dfrac{E|X_1|}{n^2} \longrightarrow 0$

$$\therefore \mu = \mu_n + \varepsilon$$

ie $\mu_n \longrightarrow \mu$

Prop 2.1 – Reading needed for <u>step 4</u>.

Step 3 We will prove later.

Lemma 1 : a) If $X \geq 0$, discrete with values $0, 1, 2, 3 \ldots$ then

$$\boxed{E(X) = \sum_{n=1}^{\infty} P(X \geq n)}$$

b) For any r.v X we have

$$\boxed{\sum_{n=1}^{\infty} P(|X| \geq n) \leq E|X| \leq \sum_{n=0}^{\infty} P(|X| \geq n)}$$

[ie even when you have a continuous r.v., we can approx it from below and above]

Proof (b): Always: $[|X|] \leq |X| \leq [|X|] + 1$ $\quad$ { $[\cdot]$ = integer part

$$\therefore \quad E[|X|] \leq E|X| \leq E([|X|] + 1)$$

Now $\quad E([|X|]) = \sum_{n=1}^{\infty} P([|X|] \geq n)$ $\quad$ from ⓐ since $|X| \geq 0$ & $[\cdot]$ is discrete

$$= \sum_{n=1}^{\infty} P(|X| \geq n)$$

{ by def & showing both events are same.

Now $\quad E([|X|] + 1) = E([|X|]) + 1$

$$= \sum_{n=1}^{\infty} P(|X| \geq n) + P(|X| \geq 0)$$

$$= \sum_{n=0}^{\infty} P(|X| \geq n)$$

$\overline{\lim} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m\geq n}^{\infty} A_m = \{\omega : \omega \in A_n \text{ infinitely often}\}$

BOREL-CANTELLI LEMMAS

LEMMA 1: For any seq of events A_n : $\underline{\sum_{n=1}^{\infty} P(A_n) < \infty \rightarrow P(A_n \text{ i.o}) = 0}$

Proof: Now $P(A_n \text{ io}) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{m\geq n}^{\infty} A_m\right)$

$$= \lim_{n\to\infty} P\left(\bigcup_{m\geq n}^{\infty} A_m\right)$$

Here $\mu(\Omega) = P(\Omega) < \infty$

$\left\{ \begin{array}{l} \because B_n = \bigcap_{m\geq n}^{\infty} A_m \text{ is a } \downarrow \text{seq} \\ \therefore \mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) \end{array} \right.$

$$\leq \lim_{n\to\infty} \sum_{m\geq n}^{\infty} P(A_m)$$

$\{\because P(\cup A_m) \leq \sum P(A_m)$

Now $\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow \lim_{n\to\infty} \sum_{m\geq n}^{\infty} P(A_m) = 0$ $\left\{ \begin{array}{l} \because \text{if a series conv}^g \\ \text{its tail goes to zero} \end{array} \right.$

$$\therefore P(A_n \text{ i.o}) \leq 0 \Rightarrow P(\{\omega : \omega \in A_n \text{ io}\}) = 0$$

<u>NOTE</u> Used to show $X_n \xrightarrow{a.s} X$

LEMMA 2 : $\underline{\sum_{n=1}^{\infty} P(A_n) = \infty \implies P(A_n \text{ i.o.}) = 1}$ for <u>A_n's indep</u>

Proof To show $P(A_n \text{ i.o}) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{m\geq n}^{\infty} A_m\right) = 1$

Now $\{A_n \text{ i.o}\}^c = \left\{ \bigcup_{n=1}^{\infty} \bigcap_{m\geq n}^{\infty} A_m^c \right\}$

$\therefore P(\{A_n \text{ i.o}\}^c) = P\left(\bigcup^{\infty} \bigcap^{\infty} A_m^c\right)$

$$= \lim_{n\to\infty} P\left(\bigcap_{m\geq n}^{\infty} A_m^c\right)$$

$$= \lim_{n\to\infty} P\left(\lim_{N\to\infty} \bigcap_{m\geq n}^{N} A_m^c\right)$$

$$= \lim_{n\to\infty} \lim_{N\to\infty} P\left(\bigcap_{m\geq n}^{N} A_m^c\right) \qquad \left\{ \because \left(\bigcap_{m\geq n}^{N} A_m^c\right) \text{is } \downarrow \text{ seq.} \right.$$

$$= \lim_{n\to\infty} \lim_{N\to\infty} \prod_{m\geq n}^{\infty} \left(1 - P(A_m)\right) \qquad \left\{ \begin{array}{l} A_n\text{'s are} \\ \text{indep events} \end{array} \right.$$

$$\leq \lim_{n\to\infty} \lim_{N\to\infty} \left[e^{-\sum_{m\geq n}^{N} P(A_m)} \right] \qquad \left\{ \because\ 1-x \leq e^{-x} \quad \forall x \right.$$

$$= \lim_{n\to\infty} e^{-\sum_{m\geq n}^{\infty} P(A_m)}$$

Now $\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow \lim_{n\to\infty} \sum_{m\geq n} P(A_m) = \infty$ $\left\{ \begin{array}{l} \text{Tail diverges} \\ \text{since the } \sum \text{seq} = \infty. \end{array} \right.$

$$\Rightarrow \lim_{n\to\infty} e^{-\sum_{m\geq n}^{\infty} P(A_m)} = e^{-\infty} = 0$$

$$\therefore P\left(\{A_n \text{ i.o}\}^c\right) = 0$$

$$\Rightarrow P\left(\{A_n \text{ i.o}\}\right) = 1$$

LEMMA 3 : $X_1, X_2, \ldots X_n \ldots$ are iid as X

$$\check{X}_n \equiv X_n I_{[|X_n| < n]}$$

Then $E|X| < \infty \iff \sum_{n=1}^{\infty} P(X_n \neq \check{X}_n) < \infty$ [ie K-equivalent]

$$\iff P(|X_n| \geq n \ \ i.o) = 0$$

Proof :

Now since $\sum_{n=1}^{\infty} P(|X| \geq n) \leq E|X| \leq \sum_{n=1}^{\infty} P(|X| \geq n) + P(|X| \geq 0)$

$$\therefore E|X| < \infty \iff \sum_{n=1}^{\infty} P(|X| \geq n) < \infty$$

$$\iff \sum_{n=1}^{\infty} P(|X_n| \geq n) < \infty \quad \{\because X_n\text{'s are iid } X$$

Now for $A_n \equiv \{|X_n| \geq n\}$ indep the Borel-cantelli lemmas give us

ⓐ $\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(A_n \ i.o) = 0$

ⓑ $\left[\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow P(A_n \ i.o) = 1\right] \xrightarrow{\text{negation}} \left[P(A_n \ i.o) < 1 \Rightarrow \sum_{n=1}^{\infty} P(A_n) < \infty\right]$

$$\Rightarrow \left[P(A_n \ i.o) = 0 \Rightarrow \sum_{n=1}^{\infty} P(A_n) < \infty\right]$$

by Kolmogorov's 0,1 law $\{A_n \ i.o\}$ is an event which has prob zero or one always for A_n's indep

So we have $P(A_n \text{ i.o}) = 0 \iff \sum_{n=1}^{\infty} P(A_n) < \infty$

$$\therefore \quad E|X| < \infty \iff \sum_{n=1}^{\infty} P(|X_n| \geq n) < \infty$$

$$\iff P(\{|X_n| \geq n\} \text{ i.o}) < \infty$$

Now $\{|X_n| \geq n\} = \{X_n \neq \check{X}_n\}$ {by way we define $\check{X}_n$

$$\therefore P(|X_n| \geq n) = P(X_n \neq \check{X}_n)$$

$$\therefore \sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(X_n \neq \check{X}_n)$$

So $E|X| < \infty \iff \sum_{n=1}^{\infty} P(X_n \neq \check{X}_n) < \infty$

$$\iff P(\{|X_n| \geq n \text{ io}\}) < \infty$$

QED.

THEOREM: KOLMOGOROV's SLLN

$X_1, X_2 \dots X_n$ are iid X Let $\mu = E(X)$

Then (i) $E|X| < \infty \Rightarrow \bar{X}_n \xrightarrow{a.s} \mu$

(ii) $E|X| = \infty \Rightarrow \overline{\lim}\ \bar{X}_n \overset{a.s}{=} \infty$

Proof: Let $Y_n \equiv X_n I_{[|X| < n]}$

$$E|X| < \infty \iff \sum_{n=1}^{\infty} P(|X| \geq n) < \infty$$

$\left\{ \text{Since} \quad \sum_{n=1}^{\infty} P(|X| \geq n) < E|X| < \sum_{n=0}^{\infty} P(|X| \geq n) \quad \text{(Proposition)} \right.$

$$\iff \sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$$

$\Rightarrow$ The seq $\{X_n\}$ and $\{Y_n\}$ are k-equivalent

$\Rightarrow \bar{X}_n \xrightarrow{a.s} \mu \iff \bar{Y}_n \xrightarrow{a.s} \mu$ $\{$Prop 2.1 pg 206

Now let $\mu_i \equiv E(Y_i)$ for any i

let $\bar{\mu}_n = \dfrac{\mu_1 + \mu_2 + \dots + \mu_n}{n}$

We showed before (in WLLN) that $\mu_n \longrightarrow \mu$

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty}\left(\frac{a_{n+1}-a_n}{b_{n+1}-b_n}\right)$$

$b_n \uparrow \infty$

Take $a_n \equiv \mu_1 + \cdots + \mu_n$

$b_n \equiv n$

$$\lim_{n\to\infty} \bar{\mu}_n = \lim_{n\to\infty} \frac{(\mu_1 + \cdots + \mu_n + \mu_{n+1}) - (\mu_1 + \cdots \mu_n)}{(n+1) - (n)}$$

$$= \lim_{n\to\infty} \mu_{n+1}$$

$$= \mu.$$

- $\therefore \quad \bar{\mu}_n \longrightarrow \mu$

- Now $\bar{Y}_n = \dfrac{\sum_{i=1}^{n} (Y_i - \mu_i)}{n} + \bar{\mu}_n$

- To show $\bar{Y}_n \longrightarrow \mu$ we only need to show $\dfrac{\sum_{i=1}^{n}(Y_i - \mu_i)}{n} \xrightarrow{a.s.} 0$

since $\bar{\mu}_n \longrightarrow \mu$ has been proved.

- **Kroneckers lemma says**: If $\{Z_i\}$ is a seq of r.v

$$\sum_{i=1}^{n} Z_i \xrightarrow{a.s} Z \implies \frac{1}{b_n}\sum_{i=1}^{n} b_i Z_i \xrightarrow{a.s} 0 \quad \text{for any } b_n \uparrow \infty$$

Now $\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu_i) = \frac{1}{n}\sum_{i=1}^{n}\frac{(y_i-\mu_i)\,i}{i}$

$$= \sum_{i=1}^{n}\frac{i Z_i}{n} \qquad \left\{ \begin{array}{l} \text{let } Z_i = (y_i-\mu_i)/i \\ b_i = i \end{array} \right.$$

$\therefore$ It is enough to show $\sum_{i=1}^{n} Z_i \xrightarrow{a.s} Z$ (some)

• Prop pg 30 : If $\mu(\Omega)<\infty$ then $X_n \xrightarrow{a.e} X \Leftrightarrow$

① $\mu\left(\bigcup_{m\geq n}^{\infty} |X_m - X_n| \geq \varepsilon\right) \longrightarrow 0 \quad \forall\, \varepsilon>0$

② $\mu\left(\max_{n\leq m\leq N} |X_m-X_n| \geq \varepsilon\right) \leq \varepsilon \quad \forall\, N\geq n\geq n_\varepsilon \quad \forall\, \varepsilon>0$

$$P\left(\max_{n\leq m\leq N} |Z_m - Z_n|\right) = P\left(\max_{n<m\leq N}\left|\sum_{i=n+1}^{m}\frac{y_i-\mu_i}{i}\right| \geq \varepsilon\right)$$

$$\leq \frac{1}{\varepsilon^2}\sum_{i=n+1}^{N} \operatorname{Var}\left(\frac{y_i-\mu_i}{i}\right)$$

$$= \frac{1}{\varepsilon^2}\sum_{i=n+1}^{N}\frac{\sigma_i^{\,2}}{i^2}$$

$$\leq \frac{1}{\varepsilon^2}\sum_{i=n+1}^{\infty}\left(\frac{\sigma_i^{\,2}}{i^2}\right)$$

$$\longrightarrow 0 \quad \text{if} \quad \sum_{n=1}^{\infty}\frac{\sigma_n^{\,2}}{n^2} < \infty$$

Consider $\sum\limits_{n=1}^{\infty} \frac{Var(y_n)}{n^2} \leq \frac{\sum\limits_{n=1}^{\infty} E(y_n^2)}{n^2}$

$$= \frac{1}{n^2} \sum_{n=1}^{\infty} \int_{[|x| \leq n]} x^2 dF$$

$$= \frac{1}{n^2} \sum_{n=1}^{\infty} \sum_{k=1}^{n} \int_{[k-1 \leq |x| \leq k]} x^2 dF$$

$$= \sum_{k=1}^{n} \sum_{n=k}^{\infty} \frac{1}{n^2} \int_{[k-1 \leq |X| \leq k]} x^2 dF$$

$\left\{ \begin{array}{l} \text{change order} \\ \text{of the sum.} \end{array} \right.$

$$= \sum_{k=1}^{n} \left(\int_{[k-1 \leq |X| \leq k]} x^2 dF \sum_{n=k}^{\infty} \frac{1}{n^2} \right)$$

Now

$$\sum_{n=k}^{\infty} \frac{1}{n^2} = \sum_{n=k}^{\infty} \int_{n}^{n+1} \frac{dx}{n^2} \qquad \text{—— (1)}$$

Now $\frac{1}{n^2} \leq \frac{2}{x^2}$ for $n \leq x \leq n+1$ for $n \geq 3$ —— (2)

(1) & (2) $\Rightarrow$ $\sum\limits_{n=k}^{\infty} \frac{1}{n^2} \leq \sum\limits_{n=k}^{\infty} \int_{n}^{n+1} \frac{2 dx}{x^2}$

$$= \int_{k}^{\infty} \frac{2 dx}{x^2} = \frac{2}{k}$$

$$\therefore \quad \sum_{n=k}^{\infty} \frac{1}{n^2} \leq \frac{2}{k} \qquad \text{——} (*)$$

So

$$\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^2} \leq \sum_{k=1}^{\infty} \int_{[k-1 \leq |X| \leq k]} x^2 dF \; \frac{2}{k}$$

$$= \sum_{k=1}^{\infty} \frac{2}{k} \int_{[k-1 \leq |X| \leq k]} |x| \; |x| \; dF$$

$$\leq \sum_{k=1}^{\infty} \frac{2}{k} \; k \int_{[k-1 \leq |X| \leq k]} |x| \; dF$$

$$= \sum_{k=1}^{\infty} 2 \int_{[k-1 \leq |X| \leq k]} |x| \; dF$$

$$= 2 \int_0^{\infty} |x| \; dF$$

$< \infty$ since $E|X| < \infty$

— QED

PROPOSITION (Kolmogorov)

$X_1, X_2, \dots X_k \dots$ are indep r.v with $X_k \cong (0, \sigma_k^2)$

Let $S_n = X_1 + \dots + X_n$ then

$$\boxed{P\left(\max_{1 \le k \le n} |S_k| \ge \lambda\right) \le \frac{\sum_{k=1}^{n} \sigma_k^2}{\lambda^2} \equiv \frac{Var(S_n)}{\lambda^2}}$$

Proof:

$$A_k \equiv \left[\left(\max_{1 \le j < k} |S_j|\right) < \lambda \le |S_k|\right]$$

k is the 1st (time) index for which $|S_k| \ge \lambda$

[in Stochastic theory k = first passage time]

Let $A = \bigcup_{k=1}^{n} A_k$ $\quad\{ A_k$'s are disjoint

$= \max_{1 \le k \le n} |S_k| \ge \lambda$

To show $P(A) \le \dfrac{Var(S_n)}{\lambda^2}$

ie $Var(S_n) \ge \lambda^2 P(A)$

Now $Var(S_n) = \int S_n^2 \, dP$ $\quad\{\because E(S_n) = 0$

$\ge \int_A S_n^2 \, dP$

$$= \sum_{k=1}^{n} \int_{A_k} (S_n^2)\, dP \qquad \left\{ \because A = \sum_{k=1}^{n} A_k \right.$$

$$= \sum_{k=1}^{n} \int_{A_k} (S_n - S_k + S_k)^2\, dP$$

$$= \sum_{k=1}^{n} \left[\int (S_n - S_k)^2 I_{A_k}\, dP + 2 \underbrace{\int (S_n - S_k) S_k I_{A_k}\, dP} + \int S_k^2 I_{A_k}\, dP \right]$$

$$\geq \sum_{k=1}^{n} \left[0 + 2 \int \underbrace{S_k I_{A_k}} \underbrace{(X_{k+1} + X_{k+2} + \cdots + X_n)} dP + \int_{A_k} S_k^2\, dP \right.$$

indep

$$= \sum_{k=1}^{n} 2\, E(S_k I_{A_k}) \underbrace{E(S_n - S_k)}_{\substack{\shortparallel \\ 0}} + \int_{A_k} S_k^2\, dP$$

$$= \sum_{k=1}^{n} \int_{A_k} S_k^2\, dP$$

$$\geq \sum_{k=1}^{n} \int_{A_k} \lambda^2\, dP \qquad \left\{ \because \text{ on } A_k \;;\; S_k^2 \geq \lambda^2 \right.$$

$$= \sum_{k=1}^{n} \lambda^2 P(A_k)$$

$$= \lambda^2 \sum_{k=1}^{n} P(A_k)$$

$$= \lambda^2 P(A) \qquad \left\{ \because A = \sum_{k=1}^{n} A_k \right.$$

KOLMOGOROV SLLN : $\bar{X}_n \xrightarrow{a.s} \mu$

needs indep and identical X_i's with $E|X|<\infty$ & $E(X_i)=\mu$

if the identical assumption is not there the SLLN fails

<u>Example 1</u> [The identically distributed cond^n req in Kolmogorov SLLN]

$X_1, \dots X_n \dots$ are indep r.v st

$$P(X_n = -1) = 1 - 1/n^2$$

$$P(X_n = n^2-1) = 1/n^2$$

Now
$$E(X_i) = -1 \times \left(1 - \frac{1}{i^2}\right) + (i^2-1)\frac{1}{i^2}$$

$$= -1 + \frac{1}{i^2} + 1 - \frac{1}{i^2}$$

$$= 0$$

$$E(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = 0$$

We shall now show that $\bar{X}_n \stackrel{a.s}{\not\to} 0$

Define $Y_n = -1 \quad \forall\ n \geq 1$

If $\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$ then $\{X_n\}$ and $\{Y_n\}$ are k-equivalent

We will show $\sum_{n=1}^{\infty} P(X_n \neq -1) < \infty$

$$\sum_{n=1}^{\infty} P(X_n \neq -1) = \sum_{n=1}^{\infty} P(X_n = (n^2-1))$$

$$= \sum_{n=1}^{\infty} 1/n^2$$

$$< \infty$$

Now if $\{X_n\}$ and $\{Y_n\}$ are k-equivalent then if one of $\bar{X}_n$ or $\bar{Y}_n \xrightarrow{as}$ then they both do and to the same limit

(Proposition 1 - pg 6 lec 19)

Now as $n \to \infty$ $\quad \bar{Y}_n \xrightarrow{a.s} -1$

$$\Rightarrow \quad \bar{X}_n \xrightarrow{a.s} -1$$

$$\therefore \quad \bar{X}_n \overset{a.s}{\not\longrightarrow} 0$$

Aside

$$\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty \quad \Rightarrow \quad \lim_{n\to\infty} P(X_n \neq Y_n) = 0$$

$\Downarrow$ Borel Cantelli

$$P(X_n \neq Y_n \text{ io}) = 0$$

Example 2 $X_1, X_2 \dots X_n$ are iid $Exp(1)$

Show $\overline{\lim_n} \dfrac{X_n}{\log n} = 1$ a.s

Sol^n We will show "$\geq$" and "$\leq$" $\Rightarrow$ "$=$"

"$\geq$". $\forall\ \varepsilon > 0$ $P(X_n > (1+\varepsilon)\log n) = e^{-(1+\varepsilon)\log n}$ $\{X_i \sim exp(1)$

$$= n^{-(1+\varepsilon)}$$

$$\sum_{n=1}^{\infty} P(X_n > (1+\varepsilon)\log n) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}}$$

$$< \infty$$

$$\left\{ \sum_{n=1}^{\infty} \frac{1}{n^p} \begin{array}{ll} < \infty & \text{if } p > 1 \\ = \infty & \text{if } p \leq 1 \end{array} \right.$$

ie $\displaystyle\sum_{n=1}^{\infty} P\left(\frac{X_n}{\log n} > (1+\varepsilon)\right) < \infty$

$$\Rightarrow P\left(\left\{\omega \,\middle|\, \frac{X_n}{\log n} > (1+\varepsilon) \ \text{io}\right\}\right) = 0$$

$\{\because$ Borell Cantelli lemma

$$\Rightarrow P\left(\left\{\omega \,\middle|\, \overline{\lim} \frac{X_n}{\log n} \geq 1+\varepsilon\right\}\right) = 0$$

$$\left\{\omega \,\middle|\, \frac{X_n(\omega)}{\log n} > 1+\varepsilon\right\} = \left\{\omega \,\middle|\, \frac{X_n(\omega)}{\log n} \geq 1+\varepsilon\right\}$$

$$\boxed{\text{If } X_n > a \Rightarrow \overline{\lim}\, X_n \geq a}$$

$$\Rightarrow \quad \overline{\lim_{n\to\infty}} \frac{X_n}{\log n} < 1+\varepsilon \quad a.s.$$

$$\Rightarrow \quad \overline{\lim} \frac{X_n}{\log n} \leq 1 \quad a.s \quad \text{taking } \varepsilon \longrightarrow 0$$

[Taking $1+\varepsilon$ was to ensure convergence]

" $\geq$ " Now $P(X_n > (1-\varepsilon)\log n) = e^{-(1-\varepsilon)\log n}$

$$= \frac{1}{n^{1-\varepsilon}}$$

$$\therefore \sum_{n=1}^{\infty} P(X_n > (1-\varepsilon)\log n) = \sum_{n=1}^{\infty} \frac{1}{n^{1-\varepsilon}}$$

$$= \infty \qquad \{\because 1-\varepsilon < 1$$

$$\Rightarrow \quad P(X_n > (1-\varepsilon)\log n \ \ io) = 1 \qquad \{\because \text{Borel cantelli}$$

$$\Rightarrow \quad P\left(\omega \,\middle|\, \frac{X_n}{\log n} \geq 1-\varepsilon \ \ io\right) = 1$$

$$\Rightarrow \quad P\left(\overline{\lim} \frac{X_n}{\log n} \geq 1-\varepsilon\right) = 1$$

$$\Rightarrow \quad \overline{\lim} \left(\frac{X_n}{\log n}\right) \geq 1-\varepsilon \quad a.s$$

$$\Rightarrow \quad \overline{\lim} \left(X_n/\log n\right) \geq 1 \quad a.s \qquad \text{take } \varepsilon \longrightarrow 0$$

GENERAL SLLN & WLLN

Let $X_1, X_2 \ldots X_n \ldots$ be indep r.v [not necessarily identical]

a) If $\sum_{k=1}^{n} \frac{\sigma_k^2}{b_n^2} \longrightarrow 0$ where $\sigma_k^2 = Var(X_k)$

then $$\frac{1}{b_n} \sum_{k=1}^{n} (X_k - \mu_k) \xrightarrow{P} 0$$

b). If $\sum_{k=1}^{\infty} \frac{\sigma_k^2}{b_k^2} \longrightarrow 0$ where $\sigma_k^2 = Var(X_k)$ & $\{b_n\} \uparrow \infty$

then $$\sum_{k=1}^{n} \frac{(X_k - \mu_k)}{b_n} \xrightarrow{a.s} 0$$

[No proof for midterm]

Now back to example 1 Indep X_n's st $P(X_n = -1) = 1 - \frac{1}{n^2}$

$P(X_n = n^2 - 1) = \frac{1}{n^2}$

We can now find a b_n st GSLLN holds.

Now $E(X_n) = 0$

$$E(X_k^2) = 1\left(1-\frac{1}{n^2}\right) + (n^2-1)^2\frac{1}{n^2}$$

$$= \left(1-\frac{1}{n^2}\right) + \frac{(n^4-2n^2+1)}{n^2}$$

$$= 1-\frac{1}{n^2} + n^2 - 2 + \frac{1}{n^2}$$

$$= n^2-2+1$$

$$= n^2-1$$

We must find a b_k st $\sum_{k=1}^{\infty} \frac{\sigma_k^2}{b_k^2} < \infty$ & $\{b_n\} \uparrow \infty$

Now $\sigma_k^2 = k^2-1$

we want $\sum_{k=1}^{\infty}\left(\frac{k^2-1}{b_k^2}\right) < \infty$ & $\{b_k\} \uparrow \infty$

If $b_k = (k)^{1+\varepsilon}$ for any $\varepsilon > 1/2$

$\Rightarrow b_k^2 = k^{2+\varepsilon_1}$ for any $\varepsilon_1 > 1$

$\therefore$ by GSLLN we get $\frac{1}{n^{1+\varepsilon}} \sum_{i=1}^{n} X_i \xrightarrow{a.s} 0$ for any $\varepsilon > \frac{1}{2}$

$\boxed{\text{HOW LLN HELPS WITH MC ESTIMATION}}$

Let $h : [0,1] \longrightarrow [0,1]$ be a continuous

Let $(\xi_1, Z_1) \ldots (\xi_n, Z_n)$ be iid pairs of $\mathcal{U}(0,1)$ r.v

ie $\xi_i \sim Unif(0,1)$
$Z_i \sim Unif(0,1)$

Let $X_k \equiv I_{[h(\xi_k) \geq Z_k]}$

Then $\bar{X}_n \xrightarrow{a.s} \int_0^1 h(t)dt$

$\boxed{\text{Proof}}$ X_k's are functions of iid r.v $\Rightarrow$ X_k's are iid

$$E|X_k| = E\left(I_{[h(\xi_k) \geq Z_k]}\right) < \infty$$

Now $E(X_1) = E\, I_{[h(\xi_k) \geq Z_k]}$

$$= \int_0^1 \int_0^1 I_{[h(t) \geq s]}\, dt\, ds \qquad \left\{ \begin{array}{l}\text{Theorem of} \\ \text{unconscious stat}\end{array}\right.$$

$$= \int_0^1 \int_0^{h(t)} ds\, dt = \int_0^1 h(t)dt$$

$\therefore$ By LLN $\bar{X}_n \xrightarrow{a.s} \int_0^1 h(t)dt$

Ques: How fast does the convergence take place.? How many uniforms do we need to generate?

Theorem: Glivenko-Cantelli Theorem:

Let $X_1, X_2 \ldots X_n$ be iid with df F

$$F_n(x) = F_n(x,\omega) \equiv \frac{1}{n}\sum_{k=1}^{n} I_{[-\infty, x]}(X_k(\omega)) = \frac{1}{n}\sum_{k=1}^{n} I\cdot[X_k \le x]$$

F_n is the empirical dist^n function (also a random function) Then

$$\|F_n - F_\infty\| = \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \xrightarrow[n\to\infty]{} 0$$

(A uniform law of large numbers)

Given $X \sim F$

We want to obtain $F(x) = P(X \le x)$

We have a sample $X_1, X_2 \ldots X_n$ iid X

An approximation $\hat{F}(x) = \sum_{i=1}^{n} \frac{I(X_i \le x)}{n}$

Note: for a fixed x (arbitrary)

$$E[F_n(x)] = \frac{1}{n}\sum_{k=1}^{n} E(I[X_k \le x]) = E[X_1 \le x] = P(X_1 \le x) = F(x)$$

By LLN for any fixed $x \in \mathbb{R}$: $F_n(x) \xrightarrow{a.s} F(x)$

This holds for any $x \in \mathbb{R}$ so the uniform

Here we are looking at sup over $\mathbb{R}$ an infinite family

We simplify it to $[0,1]$

Proof:

Claim: $F_n(x) = \frac{1}{n} \sum_{k=1}^{n} I_{[\xi_k \le F(x)]}$ where $\xi_i \sim Unif(0,1)$

$\left\{ \because I_{[X \le \cdot]} = I_{[\xi \le F(\cdot)]} \right.$

Now let G_n be the empirical $dist^n$ $funct^n$ of $\xi_1, \dots \xi_n$

ie $G_n(t) = \frac{1}{n} \left(\sum_{i=1}^{n} I_{[\xi_i \le t]} \right)$

Let $h(\cdot)$ be the identity $funct^n$

then $G_n(F(x)) - h(F(x)) = F_n(x) - F(x)$

$$\therefore \sup_{x \in \mathbb{R}} \left| F_n(x) - F(x) \right| = \sup_{t \in [0,1]} \left| G_n(t) - h(t) \right| \quad \left\{ t \equiv F(x) \in [0,1] \right.$$

$$= \sup_{t \in [0,1]} \left| G_n(t) - t \right|$$

We will show $\sup_{t \in [0,1]} \left| G_n(t) - t \right| \xrightarrow[n \to \infty]{a.e} 0$

We will now discretize the uncountable set $[0,1]$ to a finite # of points so that sup reduces to a max.

Let $\varepsilon > 0$ we can always find M st $\frac{1}{M} < \varepsilon$

Let $[0,1] \equiv$ $0, \frac{1}{M}, \frac{2}{M}, \dots, \frac{M-1}{M}, \frac{M}{M} \equiv 1$

We can show $\sup_{t \in [0,1]} \left| G_n(t) - t \right| \leq \max_{0 \leq K \leq M} \left| G_n\left(\frac{K}{M}\right) - \frac{K}{M} \right| + \frac{1}{M}$

$\therefore$ we need to show $\max_{0 \leq K \leq M} \left| G_n\left(\frac{K}{M}\right) - \frac{K}{M} \right| \xrightarrow{a.s} 0$

We will show each $\left| G_n\left(\frac{K}{M}\right) - \frac{K}{M} \right| \xrightarrow{a.s} 0 \quad \forall K$

Now $G_n(t) = \frac{1}{n} \sum_{i=1}^{n} I_{[0 \leq \xi_i \leq t]}$

$$= \frac{1}{n} \sum_{i=1}^{n} I_{[0,t]}(\xi_i)$$

$$\therefore G_n\left(\frac{K}{M}\right) = \frac{1}{n} \sum_{i=1}^{n} I_{[0, K/M]}(\xi_i) \equiv \frac{1}{n} \sum_{i=1}^{n} (Y_i)$$

Now $Y_i = 1$ when $\xi_i \in [0, K/M]$
$= 0$ ow

$Y_i = 1$ with prob. K/M
0 with prob. $(1 - K/M)$ $\quad \xi \sim$ Unif.

$\therefore E(Y_i) = K/M$

$G_n\left(\frac{K}{M}\right) = \bar{Y}_n$

By SLLN $\bar{Y}_n \xrightarrow{a.s} \frac{K}{M} \implies \bar{Y}_n - \frac{K}{M} \xrightarrow{a.s} 0$

$\implies \left| \bar{Y}_n - K/M \right| \xrightarrow{a.s} 0$

$\therefore \quad \left| G_n\left(\frac{K}{M}\right) - \frac{K}{M} \right| \xrightarrow{a.s} 0 \qquad \forall K$

$\Rightarrow \quad \max_{0 \le K \le M} \left| G_n\left(\frac{K}{M}\right) - \frac{K}{M} \right| \xrightarrow{a.s} 0$

Theorem: Law of Iterated Logarithms

Let $X_1, X_2, \ldots X_n \ldots$ be iid r.v. Assume $E(X_1) = 0$ and $Var(X_1) = \sigma^2 < \infty$

$$\overline{\lim} \frac{S_n}{\sqrt{2n \log\log n}} = \sigma \quad a.s$$

$$\underline{\lim} \frac{S_n}{\sqrt{2n \log\log n}} = -\sigma \quad a.s$$

where $S_n = X_1 + X_2 + \cdots + X_n$

Suppose $X_1 \ldots X_n \ldots$ are iid $(\mu, 1)$ with $Var(X_1) = 1 < \infty$

1) $\dfrac{\sum_{k=1}^{n} (X_k - \mu)}{n} \xrightarrow{a.s} 0$ by SLLN

2) $\dfrac{\sum_{k=1}^{n} X_k - \mu}{n^{1/r}} \xrightarrow{a.s} 0$ for $1 \le r < 2$

3) $$\overline{\lim} \frac{\sum_{k=1}^{n}(X_k-\mu)}{\sqrt{n}} = \left(\sqrt{2\log\log n} =\right) \infty \quad a.s$$

$$\underline{\lim} \frac{\sum_{k=1}^{n}(X_k-\mu)}{\sqrt{n}} = -\infty \quad a.s$$ [Talks about actual values.

4) $$\frac{\sum_{k=1}^{n}(X_k-\mu)}{\sqrt{n}} \xrightarrow{d} N(0,1)$$ {Talks about freq of values

$\Rightarrow$ $$\frac{\sum_{k=1}^{n}(X_k-\mu)}{\sqrt{2n\log\log n}} \xleftrightarrow{a.s} [-\sigma, \sigma]$$

values are a.s in $[\sigma, \sigma]$

* This is why we divide S_n by 'n' if we want an estimate of μ

CONVERGENCE IN DIST^N - CLT

Say W_n is a r.v

Let $Z \sim N(0,1)$ with df Φ

W_n is said to be asymp Normal when

$$F_{W_n}(z) \xrightarrow[n\to\infty]{} \Phi(z)$$

for all $z \in$ set of continuity pts of Φ
(here it is $\forall z$ since 'Z' is continuous)

We will prove $\| F_{W_n} - \Phi \|_\infty = \sup_z \left| F_{W_n}(z) - \Phi(z) \right| \longrightarrow 0$

$\boxed{\text{Theorem (Berry - Esseen CLT)}}$

Suppose $X_{nk} \cong (\mu_{nk}, \sigma_{nk}^2)$ for $k = 1, 2 \ldots n$ are indep

Let $Y_k = \dfrac{X_{nk} - \mu_{nk}}{\left(\sum^n \sigma_{nk}^2\right)^{1/2}}$

Let $W_n = \sum_{k=1}^{n} Y_k \cong (0,1)$

Let $\gamma_n = \dfrac{\sum_{k=1}^{n} E\,|X_{nk} - \mu_{nk}|^3}{\left(\sum_{k=1}^{n} \sigma_{nk}\right)^{3/2}}$

If $\gamma_n \xrightarrow[n\to\infty]{} 0$ then $\| F_{W_n} - \Phi \|_\infty \xrightarrow[n\to\infty]{} 0$

Note: You can show $\| F_{W_n} - \Phi \|_\infty \leq 9\gamma_n$

and the rate of convergence of F_{W_n} to Φ is $9\gamma_n$

In applications we can work with convergence of γ_n

Sometimes however the 3rd abs moments may not exist

so we need easier cond$^{\underline{n}}$ $\longrightarrow$ Lindberg Feller CLT

Theorem: Lindberg-Feller CLT

Suppose $X_{nk} \cong (\mu_{nk}, \sigma_{nk}^2)$ $k = 1, 2 \ldots n$ are indep

Let $Y_k = \dfrac{X_{nk} - \mu_{nk}}{\left(\sum_{k=1}^{n} \sigma_{nk}^2\right)^{1/2}}$

$Z_n = \sum_{k=1}^{n} Y_k \cong (0,1)$

$S_n^2 = \sum_{k=1}^{n} \sigma_{nk}^2$

Then the following are equivalent

1) $\underline{\| F_{Z_n} - \Phi \|_\infty \longrightarrow 0}$ & $\underbrace{\underline{\max_{1 \le k \le n} P\left(\frac{|X_{nk} - \mu_{nk}|}{s_n} \ge \varepsilon \right)} \longrightarrow 0}$

u.a.n [uniformly asymp negligible]

2) $\underbrace{\underline{LF_n^\varepsilon}} \equiv \underline{\left\{ \sum_{k=1}^{n} \int_{\underbrace{[|x-\mu_{nk}|/s_n \ge \varepsilon]}_{\text{tail behaviours}}} \left(\frac{x-\mu_{nk}}{s_n} \right) dF_{n_k}(x) \right\}} \longrightarrow 0 \quad \forall \varepsilon > 0$

Lindberg Feller cond$^{\underline{n}}$

Proof : We first show ② $\Rightarrow$ ①

Step 1 We show ② $\Rightarrow$ $\max_{1 \le k \le n} P\left(\frac{|X_{n_k} - \mu_{n_k}|}{s_n} \ge \varepsilon \right) \longrightarrow 0.$

Now $P\left(|X_{n_k} - \mu_{n_k}| \ge \varepsilon s_n \right) \le \frac{Var(X_{n_k})}{\varepsilon^2 s_n^2}$

$\therefore$ it is enough to show $\max_{1 \le k \le n} \left(\frac{\sigma_{nk}^2}{\varepsilon^2 s_n^2} \right) \longrightarrow 0$

Let $Y_k \equiv \frac{X_{n_k} - \mu_{n_k}}{s_n}$. & F_k be the df of Y_k.

then we have by ② $\sum_{k=1}^{n} \int_{[|y| \ge \varepsilon]} y^2 dF_k(y) \longrightarrow 0$

$$\Rightarrow \quad \max_{1\le k\le n} \int_{[|Y|\ge\varepsilon]} y^2 dF_k(y) \le \sum_{k=1}^{n} \int_{[|Y|\ge\varepsilon]} y^2 dF_k(y) \longrightarrow 0$$

Now $\quad Var(Y_k) = \dfrac{\sigma_{nk}^2}{S_n^2} \le E(Y_k^2)$

$$\therefore \quad \max_{1\le k\le n} Var(Y_k) = \max_{1\le k\le n}\left(\frac{\sigma_{nk}^2}{S_n^2}\right) \le \max_{1\le k\le n} E(Y_k^2)$$

$\therefore \quad \max\limits_{1\le k\le n}\left(\dfrac{\sigma_{nk}^2}{S_n^2}\right) \longrightarrow 0$ if we show $\max\limits_{1\le k\le n} E(Y_k^2) \longrightarrow 0$

Now $\quad \max\limits_{1\le k\le n} E(Y_k^2) = \max\limits_{1\le k\le n}\left\{ \displaystyle\int_{[|Y|\le\varepsilon]} y^2 dF_k(y) + \int_{[|Y|>\varepsilon]} y^2 dF_k(y)\right\}$

$$\le \max_{1\le k\le n}\left\{\varepsilon^2 + \int_{[|Y|>\varepsilon]} y^2 dF_k(y)\right\}$$

$$\le \varepsilon^2 + \varepsilon \qquad \{\text{using } **$$

$$< \varepsilon$$

<u>Step 2</u>: To show ② $\Rightarrow \| F_{Z_n} - \Phi \| \xrightarrow[n\to\infty]{} 0$

We have $\quad Z_n \equiv \sum\limits_{k=1}^{n} Y_k \cong (0,1)$

Also $E(y_k)=0$ & $\sum_{k=1}^{n} E(y_k^2) = 1$

<u>Idea</u> : Find a seq Z_n' st 1) $Z_n' \xrightarrow{d} Z \; [\equiv N(0,1)]$ (a)

2) $Z_n' - Z_n \xrightarrow{P} 0$ (b)

$\Rightarrow Z_n \xrightarrow{d} Z$ by Slutsky's lemma.

We define $Y_k' \equiv Y_k \, I\,[|Y_k| \leq \varepsilon_n]$ for some ε_n {we will define late

$Z_n' \equiv \sum_{k=1}^{n} Y_k'$

<u>Step (b)</u>

$$P(|Z_n'-Z_n| \geq \varepsilon_n) = P\left(\left|\sum_{k=1}^{n} Y_k' - \sum_{k=1}^{n} Y_k\right| \geq \varepsilon_n\right)$$

$$\leq P\left(\sum_{k=1}^{n} y_k' \neq \sum_{k=1}^{n} y_k\right)$$

{ $\Downarrow$ can be shown by contrad

$$\leq P\left(\exists\; 1 \leq k \leq n \text{ st } Y_k' \neq Y_k\right)$$

$$= P\left(\bigcup_{k=1}^{n} [y_k' \neq y_k]\right)$$

$$\leq \sum_{k=1}^{n} P(y_k' \neq y_k)$$

$$= \sum_{k=1}^{n} P(|Y_k| > \varepsilon_n)$$ { by def of Y_k'

Now $E\left(|Y_k|^2 I_{[|Y_k| \geq \varepsilon_n]}\right) \geqslant \varepsilon_n^2 P(|Y_k| \geq \varepsilon_n)$

$$\therefore P(|Y_k| \geq \varepsilon_n) \leq \frac{E\left(|Y_k|^2 I_{[|Y_k| \geq \varepsilon_n]}\right)}{\varepsilon_n^2}$$

$$\therefore P(|Z_n' - Z_n| > \varepsilon_n) \leq \frac{\sum_{k=1}^{n} \left(E|Y_k|^2 I_{[|Y_k| \geq \varepsilon_n]}\right)}{\varepsilon_n^2}$$

$$= \frac{1}{\varepsilon_n^2} LF_n^{\varepsilon_n}$$

Now chose an ε_n st $\dfrac{LF_n^{\varepsilon_n}}{\varepsilon_n^2} \longrightarrow 0$

$$\therefore Z_n' - Z_n \xrightarrow{P} 0$$

To show $Z_n' \xrightarrow{d} Z$ ie to show $\| F_{Z_n'} - \Phi \|_\infty \longrightarrow 0$

we will show the condition of the Berry-Esseen theorem for this truncated variable sums hold.

Let $\mu_k^* = E\left(Y_k I_{[|Y_k| \leq \varepsilon_n]}\right) = E(Y_k')$

$\sigma_k^{2*} = Var\left(Y_k I_{[|Y_k| \leq \varepsilon_n]}\right) = V(Y_k')$

Let $U_k = \dfrac{Y_k' - \mu_k^*}{\left[\sum_{k=1}^{n} (\sigma_k^{2*})\right]^{1/2}}$

Now $\sum_{k=1}^{n} U_k \simeq (0,1)$

$\therefore \quad U_k \left(\sum_{k=1}^{n} \sigma_k^{2*}\right)^{1/2} + \mu_k^* = Y_k'$

$\Rightarrow \left(\sum_{k=1}^{n} \sigma_k^{2*}\right)^{1/2} \sum_{k=1}^{n} U_k + \sum_{k=1}^{n} \mu_k^* = \sum_{k=1}^{n} Y_k' \equiv Z_n'$

$$\left.\begin{array}{ll} \text{If } ① & \sum_{k=1}^{n} \mu_k^* \longrightarrow 0 \\ \& \ ② & \left(\sum_{k=1}^{n} \sigma_k^{2*}\right)^{1/2} \longrightarrow 1 \\ \& \ ③ & \sum_{k=1}^{n} U_k \xrightarrow{d} N(0,1) \end{array}\right\} \xRightarrow{\text{by Slutsky's}} \quad \| Z_n' - \Phi \|_\infty \longrightarrow 0$$

By the Berry-Esseen theorem if $\gamma_n \longrightarrow 0$ the $\|\sum_{k=1}^{n} U_k - \Phi\|_\infty \longrightarrow 0$

where $\gamma_n = \sum_{k=1}^{n} \dfrac{E\,|\,Y_k' - \mu_k^*\,|^3}{\left(\sum_{k=1}^{n} \sigma_k^{*2}\right)^{3/2}}$

Now $E|Y_k' - \mu_k^*|^3 = E\left|\left(Y_k I_{[|Y_k| \le \varepsilon_n]} - \mu_k^*\right)\right|^3$

$$\le 4\left\{E\left|Y_k I_{[|Y_k|\le\varepsilon_n]}\right|^3 + |\mu_k^*|^3\right\}$$ {by the C_2 ineq

$$\sum_{k=1}^n E|Y_k' - \mu_k^*|^3 \le 4 \underbrace{\sum_{k=1}^n E\left|Y_k I_{[|Y_k|\le\varepsilon_n]}\right|^3}_{\text{II}} + \underbrace{\sum_{k=1}^n |\mu_k^*|^3}_{\text{I}}$$

To show $\gamma_n \longrightarrow 0$ we need to show both I & II $\longrightarrow 0$ (since the denomenator in γ_n goes to 1)

$$\text{I} = \sum_{k=1}^n |\mu_k^*|^3 = \sum_{k=1}^n E\left|Y_k I_{[|Y_k|\le\varepsilon_n]}\right|^3$$

Recall $Y_k = \dfrac{X_{nk} - \mu_{nk}}{s_n}$ so $E(Y_k)=0$ & $\sum_{k=1}^n V(Y_k) = \sum_{k=1}^n E(Y_k^2) = 1$

$$\sum_{k=1}^n |\mu_k^*|^3 = \sum_{k=1}^n E\left|Y_k I_{[|Y_k|\le\varepsilon_n]}\right|\left(E\left|Y_k I_{[|Y_k|\le\varepsilon_n]}\right|\right)^2$$

$$\le \varepsilon_n \sum_{k=1}^n \left[E\left(Y_k I_{[|Y_k|\le\varepsilon_n]}\right)\right]^2$$

$$\le \varepsilon_n \sum_{k=1}^n E(Y_k^2)\, E\left(I_{[|Y_k|\le\varepsilon_n]}\right)$$ {by cauchy schwartz $|\langle u,v\rangle|^2 \le \|u\|^2 \|v\|^2$

$$\le \varepsilon_n \sum_{k=1}^n E(Y_k^2)$$ { $\because E\left[I_{[|Y_k|\le\varepsilon_n]}\right] \le 1$

$$\therefore \sum_{k=1}^{n} |\mu_k^*|^3 \leq \mathcal{E}_n \longrightarrow 0 \qquad \left\{ \because \sum_{k=1}^{n} E(Y_k^2) = 1 \right.$$

— QED

APPLICATION OF L-F CLT

What happens when the lindberg-feller condⁿ fails?

ie $\underline{LF_n^{\varepsilon} \not\rightarrow 0}$

$\exists$ seq of undep and zero mean r.v $X_1, X_2 \cdots X_n \cdots$ st.

for $S_n = X_1 + \cdots + X_n$

& $s_n^2 = Var(S_n)$ with the LF condⁿ failing

BUT

$$\underbrace{\frac{S_n}{s_n^2} \xrightarrow{d} N(0, a^2)} \quad \text{with} \quad \underline{a^2 < 1}$$

and $$\underbrace{\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \longrightarrow 0} \quad \text{with} \quad \sigma_k^2 \stackrel{d}{=} (Var[X_k])$$

$\boxed{\text{Note}}$ The diff b/w this & the L-F theorem is that $\xrightarrow{d}$ is not to $N(0,1)$ anymore.

Let $Y_1, Y_2 \dots Y_n \dots$ iid $N(0,1)$

Let $U_1, \dots U_n \dots$ indep $\simeq (0,1)$ r.v

$$U_n \in \begin{pmatrix} -nc & 0 & nc \\ \frac{1}{2n^2} & \left(1-\frac{1}{n^2}\right) & \frac{1}{2n^2} \end{pmatrix} \begin{matrix} \to \text{values} \\ \to \text{prob} \end{matrix}$$ for some fixed $c > 0$

Let $X_k = Y_k + U_k$

$\begin{cases} X_k\text{'s} \equiv Y_k\text{'s most of the time} \\ \text{but sometimes it is off by a lot} \end{cases}$

Given $V_n \xrightarrow{d} V \not\Rightarrow Var(V_n) \longrightarrow Var(V)$

Let $V_n = \frac{S_n}{\sigma_n} \Rightarrow Var\left(\frac{S_n}{\sigma_n}\right) = Var(V_n) = 1$

Let $Var(V) = a^2 < 1$

$\therefore Var(V_n) \not\longrightarrow Var(V)$

Sometimes however you can use the Var of V_n as an estimator of a^2

What is always true is $Var(V_n) \geq Var(V)$

Let W_n, V be k-dimensional vectors

Let $a \in \mathbb{R}^k$

Assume $c_n [W_n - a] \xrightarrow{d} \mathbb{V}$

Let $\Phi : \mathbb{R}^k \longrightarrow \mathbb{R}^m$

then

$$\underbrace{c_n [\Phi(W_n) - \Phi(a)]}_{\mathbb{R}^m} \xrightarrow{d} \underbrace{\Phi'(a)}_{m \times k} \underbrace{\mathbb{V}}_{k \times 1}$$

where $\Phi(x_1 \ldots x_k) = (\Phi_1(x_1 \ldots x_k) \ldots \Phi_m(x_1 \ldots x_k))$

$$\Phi'(x_1 \ldots x_k) = \begin{bmatrix} \frac{\partial \Phi_1(\underset{\sim}{x})}{\partial x_1} & \frac{\partial \Phi_1(\underset{\sim}{x})}{\partial x_2} & \cdots & \frac{\partial \Phi_1(\underset{\sim}{x})}{\partial x_k} \\ \vdots & & & \\ \frac{\partial \Phi_m(\underset{\sim}{x})}{\partial x_1} & \frac{\partial \Phi_m(\underset{\sim}{x})}{\partial x_2} & \cdots & \frac{\partial \Phi_m(\underset{\sim}{x})}{\partial x_k} \end{bmatrix}_{m \times k}$$

The bootstrap

Let $X_1, X_2 \cdots X_n$ be iid F with mean μ

Let F_n be the empirical dist^n based on the sample

$$\bar{X}_n = \int x \, dF_n(x) = \frac{1}{n} \int x \, d \sum_{i=1}^{n} \delta_{X_i} = \frac{1}{n} \sum_{i=1}^{n} \int x \, d\delta_{X_i}$$

$$= \sum_{i=1}^{n} \frac{X_i}{n}$$

$$= E_{F_n}(X)$$

$$S_n^2 = Var_{F_n}(X) = E_{F_n}[X - E_{F_n}(X)]^2$$

<u>Def</u>: A bootstrap sample is an iid sample from F_n (sampling with replacement from $(X_1, X_2 \cdots X_n)$)

Let $(X_{n1}^*, X_{n2}^*, \ldots X_{nn}^*)$ be a bootstrap sample.

$$F_n^*(x) = \frac{1}{n} \sum_{k=1}^{n} I[X_k^* \le x]$$

$$\bar{X}_n^* = \frac{1}{n} \sum_{k=1}^{n} X_{nk}^*$$

$$S_n^* = \frac{1}{(n-1)} \sum_{k=1}^{n} (X_{n_k}^* - \bar{X}_n^*)^2$$

$$\bar{Z}_n^* = \frac{\sqrt{n}\,(\bar{X}_n^* - \bar{X}_n)}{S_n}$$

measures closeness of a bootstrap mean to original sample mean

$$T_n^* = \frac{\sqrt{n}\,(\bar{X}_n^* - \bar{X}_n)}{S_n^*}$$

$$D \equiv \max_{1 \le k \le n} \frac{X_k - \bar{X}_n}{\sqrt{n}\, S_n}$$

$\boxed{\text{Theorem}}$:

1) $\bar{Z}_n^* \xrightarrow{d} N(0,1)$ iff $D_n \xrightarrow{P} 0$

2) If $D_n \xrightarrow{P} 0 \implies T_n^* \xrightarrow{d} N(0,1)$

Lec 26

If $C_n U_n \xrightarrow{d} U$ (some). for $C_n \to \infty$ then.

$$U_n \xrightarrow{P} 0$$

$\therefore$ if we let $U_n = \left(\frac{\bar{X}_n^* - \bar{X}_n}{S_n}\right)$

$$C_n = \sqrt{n}$$

we have $C_n U_n \xrightarrow{d} U \equiv N(0,1)$

$$\implies U_n \xrightarrow{P} 0$$

Note If $U_n \not\xrightarrow{P} 0$ then there is no point in trying to find the asy

Proof : We will use the Berry Esseen theorem

$$\frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{m_{ii}}} = \sum_{k=1}^{n} \frac{a_{ki}\,\varepsilon_k}{\sigma} \qquad \text{(can be shown)}$$

$$= \sum_{k=1}^{n} \frac{a_k\,\varepsilon_k}{\sigma} \equiv \sum_{k=1}^{n}{}' X_{nk} \qquad \text{supressing index 'i'}$$

where a_k's are fixed #'s based on the design

$$E(X_{nk}) = E\left(\frac{a_k\,\varepsilon_k}{\sigma}\right) = 0 = \mu_{nk}$$

$$V(X_{nk}) = Var(a_k\,\varepsilon_k/\sigma) = a_k^2 = \sigma_{nk}^2$$

Fact $\sum_{k=1}^{n} a_k^2 = 1$

Let $$Y_k = \frac{X_{nk} - \mu_{nk}}{\left(\sum_{k=1}^{n}\sigma_{nk}^2\right)^{1/2}} = \frac{X_{nk} - 0}{\left(\sum_{k=1}^{n} a_k^2\right)^{1/2}} = X_{nk}$$

$$\therefore \sum_{k=1}^{n} X_{nk} = \sum_{k=1}^{n} Y_k \equiv W_n.$$

Now $$\gamma_n = \frac{\sum_{k=1}^{n} E\,|X_{nk} - \mu_{nk}|^3}{\left(\sum_{k=1}^{n}\sigma_{nk}^2\right)^{3/2}} = \frac{\sum E|X_{nk} - 0|^3}{(\sum a_k^2)^{3/2}} = \sum_{k=1}^{3} E\,|X_{nk}|^3$$

$$= \sum_{k=1}^{3} E\left|\frac{a_k\,\varepsilon_k}{\sigma}\right|^3 \le \underset{\substack{\downarrow \\ \text{Notes}}}{} \; \mathcal{C}_3 \max_{1\le k\le n} (h_{kk})^{1/2}$$

$\therefore$ by Berry-Eseeen

$$\| G_{n_i} - \Phi \|_\infty = \| F_{W_n} - \Phi \|_\infty \leq 9 \gamma_n$$

Lindberg Feller CLT says $X_{nk} \simeq F_{nk}(\mu_{nk}, \sigma_{nk}^2)$

$$s_n^2 \equiv \sum_{k=1}^{n} \sigma_{nk}^2$$

$$Z_n \equiv \frac{\sum_{k=1}^{n} (X_{nk} - \mu_{nk})}{s_n}$$

then the foll are equivalent

(i) $\| F_{Z_n} - \Phi \|_\infty \longrightarrow 0$

$$\max_{1 \le k \le n} P\left(\frac{|X_{nk} - \mu_{nk}|}{s_n} \geq \varepsilon \right) \longrightarrow 0$$

(ii) $$LF_n^\varepsilon = \sum_{k=1}^{n} \int_{\left\{\left|\frac{X - \mu_{nk}}{s_n}\right| \geq \varepsilon\right\}} \left(\frac{(x - \mu_{nk})}{s_n} \right)^2 dF_{nk}(x) \longrightarrow 0 \quad \forall \varepsilon > 0$$

If (ii) holds $\Rightarrow$ (i) holds

$\Rightarrow M_n \xrightarrow{P} 0$ ie $\max_{1 \le k \le n} \left(\frac{|X_{nk} - \mu_{nk}|}{s_{dn}} \right) \xrightarrow{P} 0$

and $\max_{1 \le k \le n} \frac{\sigma_k^2}{s_{dn}^2} \longrightarrow 0$ where $s_{dn}^2 = \sum_{k=1}^{n} \sigma_{nk}^2$

$$U_n = \begin{pmatrix} -nc & 0 & nc \\ \frac{1}{2n^2} & \left(1-\frac{1}{n^2}\right) & \frac{1}{2n^2} \end{pmatrix}$$

$$\sum_{n=1}^{\infty} P(|U_n| \neq 0) = \sum_{n=1}^{\infty} P(|U_n| \geq \varepsilon) = \sum_{n=1}^{\infty} \left(\frac{1}{2n^2} + \frac{1}{2n^2}\right)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^2}$$

$$< \infty$$

Now U_n's are indep so Borel Cantelli 1 gives us

$$P(|U_n| \neq 0 \text{ i.o.}) = 0$$

$\therefore$ given any ω $\exists$ $N(\omega)$ st $U_n = 0$ $\forall$ $n \geq N(\omega)$ (a.s) —— $(*)$

Now consider $\sqrt{n} \dfrac{(U_1 + \cdots + U_n)}{n} = \dfrac{U_1 + \cdots + U_n}{\sqrt{n}}$ —— $(**)$

$\longrightarrow 0$ as $n \longrightarrow \infty$

$$\left\{ \frac{U_1 + \cdots + U_n}{\sqrt{n}} = \underbrace{\left(\frac{U_1 + U_2 \cdots + U_N}{\sqrt{n}}\right)}_{\substack{\downarrow \\ 0 \\ \text{since } n > N}} + \underbrace{\left(\frac{U_{N+1} \cdots U_n}{\sqrt{n}}\right)}_{\substack{\downarrow \\ 0 \text{ by } (*)}} \right\}$$

Let Y_n be a iid seq $(0,1)$ indep of U_n.

Let $X_n = U_n + Y_n$

we will show

1. LF cond^n fails
2. CLT holds for X_n
3. $\max\limits_{1 \le k \le n} \frac{\sigma_k^2}{Sd_n^2} \longrightarrow 0$

(2) $Var(X_n) = Var(U_n) + Var(Y_n)$

$$= \frac{2 n^2 c^2}{2 n^2} + 1 \qquad = c^2 + 1$$

$$S_n = \sum_{k=1}^{n} X_k$$

$$\sigma_n^2 = Var(S_n) = n(c^2+1)$$

$$\frac{S_n}{\sigma_n} = \frac{n\bar{U}_n + n\bar{Y}_n}{\sqrt{n}\sqrt{c^2+1}} = \frac{n\bar{U}_n}{\sqrt{n}\sqrt{c^2+1}} + \frac{n\bar{Y}_n}{\sqrt{n}\sqrt{c^2+1}}$$

Now $\sqrt{n}\,\bar{Y}_n \xrightarrow{d} N(0,1)$ by CLT

$$\therefore \frac{\sqrt{n}\,\bar{Y}_n}{\sqrt{c^2+1}} \xrightarrow{d} N\left(0, \frac{1}{\sqrt{c^2+1}}\right)$$

Also $\sqrt{n}\,\bar{U}_n \longrightarrow 0$ we showed in $(**)$ $\Rightarrow \left(\sqrt{n}\,\bar{U}_n / \sqrt{c^2+1}\right) \longrightarrow 0$

$$\therefore \quad \frac{S_n}{\sigma_n} \xrightarrow{d} N\left(0, \left(\frac{1}{\sqrt{c^2+1}}\right)^2\right) \equiv N(0, a^2) \qquad a < 1.$$

{by slutsky}

① $$LF_n^{\varepsilon} = \sum_{k=1}^{n} \int_{\left\{\left|\frac{x-\mu_{nk}}{s_n}\right| \geq \varepsilon\right\}} \left(\frac{x-\mu_{nk}}{s_n}\right)^2 dF_{nk}(x)$$

$$= \sum_{k=1}^{n} \int_{\left\{\frac{|x|}{\sqrt{n}/a} \geq \varepsilon\right\}} \left(\frac{x-0}{\sqrt{n}/a}\right)^2 dF_{nk}(x)$$

$$\left\{\begin{array}{l} E(X_{nk}) = \mu_{nk} = 0 \\ s_n = \sum Var(X_{nk}) \\ \quad = \sqrt{n(c^2+1)} \\ a^2 = 1/(c^2+1) \end{array}\right.$$

$$= \frac{a^2}{n} \sum_{k=1}^{n} \int_{\left\{|x| \geq \frac{\varepsilon\sqrt{n}}{a}\right\}} x^2 dF_{nk}(x).$$

$$= \frac{a^2}{n} \cdot (nc^2) + o(1)$$ ← Show - Bonus problem

$$= a^2 c^2$$

$$= \frac{c^2}{c^2+1}$$

$$\not\to 0$$

③ $$\max_{1\le k\le n} \frac{\sigma_k^2}{\sigma_n^2} = \max_{1\le k\le n}\left(\frac{Var(X_k)}{Var(S_n)}\right) = \frac{1+c^2}{n(1+c^2)} = \frac{1}{n} \xrightarrow[n\to\infty]{} 0$$ ③

$$\max_{1\le k\le n} \frac{\sigma_k^2}{\sigma_n} = \max_{1\le k\le n}\left(\frac{Var(X_k)}{\sqrt{Var(S_n)}}\right) = \frac{(1+c^2)}{\sqrt{n}\sqrt{(1+c^2)}} \xrightarrow{n\to\infty} 0$$

This example also shows

$$\text{if } X_n \xrightarrow{d} X \quad \not\Rightarrow \quad Var(X_n) \longrightarrow Var(X)$$

$$\text{Here } Var\left(\frac{\sum X_n}{\sigma_n^2}\right) = \frac{\sigma_n^2}{\sigma_n^2} = 1$$

$$Var(X) = \frac{1}{1+c^2}$$

CONTINUOUS MAPPING THEOREM

$$\left.\begin{array}{l} X_n \xrightarrow{d} X \\ X_n \xrightarrow{a.s} X \\ X_n \xrightarrow{P} X \end{array}\right\} \text{all imply} \quad g(X_n) \longrightarrow g(X) \quad \left[\text{in } d/a.s/P \text{ resp}\right]$$

if 'g' is a continuous function

Prop

If $$\left.\begin{array}{l} C_n U_n \xrightarrow{d} U \\ C_n \longrightarrow \infty \end{array}\right\} \Rightarrow U_n \xrightarrow{P} 0$$

Proof

$$\left.\begin{array}{l} \dfrac{1}{C_n} \longrightarrow 0 \\ C_n U_n \xrightarrow{d} U \end{array}\right\} \overset{\text{slutskys}}{\Longrightarrow} C_n U_n \frac{1}{C_n} \xrightarrow{d} 0.U$$

ie $U_n \xrightarrow{d} 0$

$$U_n \xrightarrow{d} 0 \iff U_n \xrightarrow{P} 0 \quad \left\{\begin{array}{l} X_n \xrightarrow{P} c \\ \Updownarrow \\ X_n \xrightarrow{d} c \end{array}\right. \quad \text{for any const 'c'}$$

- A seq V_n is $o_p(a_n)$ if $\dfrac{V_n}{a_n} \xrightarrow{P} 0$

- A seq V_n is $O_p(a_n)$ if given any $\varepsilon > 0$ $\exists$ $M(\varepsilon)$ & $N(\varepsilon)$ st

$$P\left(\left|\frac{V_n}{a_n}\right| > M(\varepsilon) \right) < \varepsilon \quad \forall \quad n \geq N(\varepsilon)$$

$\boxed{\text{Proof}}$ Given $C_n U_n \xrightarrow{d} U$

To show $X_n = C_n U_n = O_p(1)$

To show given any $\varepsilon > 0$ $\exists$ M and N st

$$P(|X_n| > M) < \varepsilon \qquad \forall\, n \geq N(\varepsilon)$$

Proof: $P(|U| > M) < \varepsilon$ for any r.v.

Now given $C_n U_n \xrightarrow{d} U$ ie $F_n(x) \longrightarrow F(x)$ $\forall\, x \in$ continuity pts of F

ie $\left| P(|C_n U_n| \leq x) - P(|U| \leq x) \right| \leq \varepsilon \qquad \forall\, n \geq N(\varepsilon)$

ie $\left| P(|C_n U_n| > x) - P(|U| > x) \right| \leq \varepsilon$

$\therefore$ $P(|C_n U_n| > x) - P(|U| > x) \leq \left| \qquad \right| \leq \varepsilon.$

so $P(|C_n U_n| > M) - P(|U| > M) \leq \varepsilon$ $\left\{ \begin{array}{l} \text{we can find an } M \\ \text{in the set of continuity} \\ \text{pts of } F \end{array} \right.$ – Portmanteau th.

$$P(|C_n U_n| > M) \leq P(|U| > M) + \varepsilon$$
$$< \varepsilon + \varepsilon$$

The Portmanteaus theorem ensures that you can find a M st it is in the set of continuity pts of F & such that $P(|u| > M) \leq \varepsilon$

1) Sheffé's theorem
2) moments
3) Portmanteaus theorem

$\Big\}$ all give conditions for convergence in dist$^{\underline{n}}$

Delta Method

$$c_n(W_n - a) \xrightarrow{d} V$$

$$\Longrightarrow \quad c_n\left(g(W_n) - g(a)\right) \xrightarrow{d} g'(a)\,V$$

for any 'g' which is differentiable at 'a'

Proof: By def of differentiability

$$g(x) - g(a) = g'(a)(x-a) + R(x-a)$$

where $\lim_{x \to a} \frac{R(x-a)}{|x-a|} = 0$ ie $R(x-a) = o(|x-a|)$

$$\therefore \quad g(W_n) - g(a) = g'(a)(W_n - a) + R(W_n - a)$$

provided $R(W_n - a) = o_p(|W_n - a|)$

$$c_n\left(g(W_n) - g(a)\right) = g'(a)\,c_n(W_n - a) + c_n R(W_n - a)$$

Now as $n \to \infty$ $c_n(W_n - a) \xrightarrow{d} V$.

If we can show $c_n R(W_n - a) \xrightarrow{P} 0$

then by slutsky's $c_n(g(W_n) - g(a)) \xrightarrow{d} g'(a)\,V$

Consider $c_n R(W_n - a) = c_n \dfrac{R(W_n - a)}{|W_n - a|}\,|W_n - a|$

Now $\quad R(W_n - a) = o_p(|W_n - a|)$ (can be shown)

$$\therefore \quad \frac{R(W_n - a)}{|W_n - a|} = o_p(1)$$

Since $\quad C_n(W_n - a) \xrightarrow{d} V \Rightarrow C_n|W_n - a| \xrightarrow{d} |V|$ $\Big\{$ continuous mapping theorem

We also showed if $\quad X_n \xrightarrow{d} X \Rightarrow X_n = O_p(1)$

$$\therefore \quad \frac{R(W_n - a)}{|W_n - a|} \; C_n|W_n - a| = o_p(1)\, O_p(1)$$

$\therefore$ now we show $o_p(1)\, O_p(1) = o_p(1)$ we are done

let $X_n \xrightarrow{P} 0$ & $Y_n = O_p(1)$

we show $\quad X_n \cdot Y_n \xrightarrow{P} 0$

ie to show $P(|X_n Y_n - 0| > \varepsilon) \longrightarrow 0 \quad$ for any $\varepsilon > 0$

Now since $\quad Y_n = O_p(1) \Rightarrow P(|Y_n| \geq M) \leq \varepsilon \quad \forall\, n \geq N_1$

for any $\varepsilon > 0$
(by definition)

Also $X_n \xrightarrow{P} 0 \implies P(|X_n| > \varepsilon) \longrightarrow 0 \quad \forall\ n \geq N_2$ for $\varepsilon > 0$ & some N_2

Now
$$P(|X_n Y_n| > \varepsilon) = \left\{ \begin{array}{l} \underbrace{P(|X_n Y_n| > \varepsilon \cap |Y_n| \geq M)}_{ⓐ} \\ + \underbrace{P(|X_n Y_n| > \varepsilon \cap |Y_n| < M)}_{ⓑ} \end{array} \right\}$$

ⓑ : $M|X_n| > |X_n Y_n| > \varepsilon$

$$(|Y_n| < M) \cap (|X_n Y_n| > \varepsilon) \implies |X_n| > \varepsilon/M$$

$$\therefore P\{|Y_n| < M \cap |X_n Y_n| > \varepsilon\} \leq P(|X_n| > \varepsilon/M) \quad \forall\ n \geq N_2$$

$\downarrow$ as $n \to \infty$

$0 \quad [\because X_n \xrightarrow{P} 0]$

ⓐ : $P(|X_n Y_n| > \varepsilon_1 \cap |Y_n| \geq M) \leq P(|Y_n| \geq M)$ $\{ P(AB) \leq P(B)$ always

$< \varepsilon \quad \forall\ n \geq N_1$

$$\therefore P(|X_n Y_n| > \varepsilon) \leq \varepsilon \quad \forall\ n \geq \max(N_1, N_2)$$

Thus $o_p(1) O_p(1) = o_p(1)$

To show $R(U_n) = o_p(U)$ when $R(h) = o(|h|)$

Let $t(h) = \frac{R(h)}{|h|}$ when $h \neq 0$

$= 0$ when $h = 0$

since $R(h) = o(|h|) \Rightarrow \frac{R(h)}{|h|} \longrightarrow 0$ as $h \to 0$

$\therefore$ $t(h)$ is a continuous funct^{n}

Now $U_n \xrightarrow{P} 0 \Rightarrow t(U_n) \xrightarrow{P} t(0)$ by continuous mapping theorem

ie $\frac{R(U_n)}{|U_n|} \xrightarrow{P} 0$

CHARACTERISTIC FUNCTION

Def Let X be a r.v with df F. The characteristic function is $\Phi(t) \equiv \Phi_X(t) \equiv E(e^{itX})$

$$= \int_{-\infty}^{\infty} e^{itX} dF(x) \qquad \{\text{fourier transf of } F$$

$$= \int_{-\infty}^{\infty} \cos(tx)\, dF(x) + i \int_{-\infty}^{\infty} \sin(tx)\, dF(x)$$

Note : The functⁿ cos and sin are bdd and so the integrals always exist. ie the functions always exist

Ex 1 : ① $X \sim f(x) = \frac{1}{\pi} \frac{1}{(1+x^2)}$

$$\Phi_X(t) = e^{-t}$$

② $X \sim N(0,1)$

$$\Phi_X(t) = e^{-t^2/2}$$

Remark : There is a 1-1 correspondance b/w characteristic functions and the cdf F.

$\boxed{\text{THEOREM 1}}$: Every df F on the real line has a unique chf Φ

$\boxed{\text{THEOREM 2}}$: For any Φ there is a unique F (construct inversion)

$\boxed{\text{INVERSION FORMULA FOR DENSITIES}}$ (wrt lebesgue meas)

If X has a chf $\Phi_X(\cdot)$ st $\int_{-\infty}^{\infty} |\Phi_X(t)|\, dt < \infty$

then X has a uniformly continuous density given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, \Phi_X(t)\, dt \qquad \left\{ \begin{array}{l} \text{inverse fourier} \\ \text{transform} \end{array} \right.$$

THEOREM (Cramèr - Levy)

1) If $\{\Phi_n\}$ is any seq of funct^n st $\underline{\Phi_n \longrightarrow \Phi}$

where Φ is continuous at '0' then $\underline{F_n \longrightarrow F}$

Moreover $\underline{\Phi}$ is the <u>chf</u> <u>associated with F</u>

2) If $\underline{F_n \longrightarrow F}$ then $\underline{\Phi_n \longrightarrow \Phi}$ <u>uniformly</u> over any <u>finite interval</u> $|t| \leq T < \infty$

<u>NOTE</u> : This is used to show convergence in dist^n of $\{X_n\}$ to some X with df F

Example : Say we wanted to know $\int \cos(tX_n)\, dF_n \longrightarrow g(t) \Rightarrow$ what is this

<u>Proof</u> (2) : $X_n \xrightarrow{d} X$ (given)

By Skorokhod's construction $Y_n \xrightarrow{a.s} Y$

$$\therefore \quad e^{it(Y_n - Y)} \xrightarrow[n \to \infty]{a.s} 1$$

{ you can always construct some such Y_n's on some other prob space
but these Y_n's have the same dist^n as X_n's

$$|\Phi_n(t) - \Phi(t)| \leq \int |e^{ity_n} - e^{ity}|\, dP$$

$$= \int |e^{ity}|\, |e^{it(y_n - y)} - 1|\, dP$$

$$= \int |\cos(ty) + i\sin(ty)|\, |e^{it(y_n-y)} - 1|\, dP$$

$$= \int \sqrt{\cos^2(ty) + \sin^2(ty)}\, |e^{it(y_n-y)} - 1|\, dP$$

$$= \int |e^{it(y_n-y)} - 1|\, dP$$

$$\leq \int \sup_{|t|\leq T} |e^{it(y_n-y)} - 1|\, dP$$

$$\lim_{n\to\infty} |\Phi_n(t) - \Phi(t)| \leq \int \lim_{n\to\infty} \sup_{|t|\leq T} |e^{it(y_n-y)} - 1|\, dP$$

$$\longrightarrow 0$$

$$\left\{ \begin{array}{l} \text{by DCT} \\ \sup_{|t|\leq T} |e^{it(y_n-y)} - 1| \leq 1 + 1 \end{array} \right.$$

-qed

CHARACTERISTIC FUNCTION OF VECTOR'S

$\underset{\sim}{X} = (X_1 \dots X_k) : \Omega \longrightarrow \mathbb{R}^k$

Let $P_{\underset{\sim}{X}}$ be the induced measure (in 1-1 corr with F)

$$\Phi_{\underset{\sim}{X}}(t) = E\left[e^{it^T \underset{\sim}{X}}\right]$$

$$= E\left[e^{i(t_1 X_1 + t_2 X_2 + \dots + t_k X_k)}\right]$$

Theorem: $\Phi_{\underset{\sim}{X}}$ determines uniquely $P_{\underset{\sim}{X}}$

net $\quad y = \lambda^T X = \sum_{j=1}^{k} \lambda_j X_j \quad ; \quad y \in \mathbb{R}^1$

$$\Phi_y(t) = E\left(e^{ity}\right)$$

$$= E\left(e^{it\lambda^T X}\right)$$

$$= E\left[e^{i(t\lambda_1 X_1 + t\lambda_2 X_2 + \dots + t\lambda_k X_k)}\right]$$

1) Now knowing $\Phi_{\underset{\sim}{X}}(\underset{\sim}{t})$ for all $\underset{\sim}{t} \in \mathbb{R}^k$ means knowing $P_{\underset{\sim}{X}}$

2) knowing $\Phi_{\lambda^T \underset{\sim}{X}}(t)$ for all $t \in \mathbb{R}^1$ and all $\underset{\sim}{\lambda}$ means knowing $P_{\lambda^T \underset{\sim}{X}}$

3) But knowing $\Phi_{\underset{\sim}{X}}(\underset{\sim}{t})$ for all $\underset{\sim}{t} \in \mathbb{R}^k$ is knowing $\Phi_{\lambda^T \underset{\sim}{X}}(t)$ for all $t \in \mathbb{R}^1$ and $\underset{\sim}{\lambda} \in \mathbb{R}^k$ with $\|\lambda\| = 1$.

This means studying $\underset{\sim}{X}$ is the same as studying $\underset{\sim}{\lambda}^T \underset{\sim}{X}$ (in dist^n)

$\boxed{\text{Theorem}}$ $\underset{\sim}{X_n} = (X_{1n} \cdots X_{kn}) \in \mathbb{R}^k$ (the components need not be indep.

If $\Phi_{\lambda^T X_n}(t) \longrightarrow \Phi_{\lambda^T \underset{\sim}{X}}(t)$ $\forall\ t \in \mathbb{R}^1$ & $\lambda \in \mathbb{R}^k$ with $\|\lambda\| = 1$.

then $\underset{\sim}{X_n} \xrightarrow{d} \underset{\sim}{X}$

(This is called the <u>Cramèr-Wald device</u>)

Remark: By Cramèr Levy theorem: $\Phi_{\lambda^T \underset{\sim}{X_n}}(t) \longrightarrow \Phi_{\lambda^T \underset{\sim}{X}}(t)$

$$\Rightarrow \lambda^T \underset{\sim}{X_n} \xrightarrow{d} \lambda^T \underset{\sim}{X}$$

$\therefore$ the Cramer-Wald device holds true in the following case

$$\lambda^T \underset{\sim}{X_n} \xrightarrow{d} \lambda^T \underset{\sim}{X} \Rightarrow \underset{\sim}{X_n} \xrightarrow{d} \underset{\sim}{X} \qquad \forall\ \lambda \in \mathbb{R}^k \ \& \ \|\lambda\| = 1$$

$\therefore$ you do not have to deal with chf of $\lambda^T \underset{\sim}{X_n}$ you can also work directly with $\lambda^T \underset{\sim}{X_n}$ and show $\xrightarrow{d}$ for that linear comb^n

1. Complex Number

$$z = a + ib$$

$$\bar{z} = a - ib \quad \text{complex conjugate}$$

$$|z| = \sqrt{a^2 + b^2} \quad \Rightarrow \quad z \cdot \bar{z} = |z|^2$$

2. Complex Function

$$f(t) = u(t) + i v(t)$$

$$\int_a^b f(t)\,dt = \int_a^b u(t)\,dt + i \int_a^b v(t)\,dt$$

3. $\cos(-\theta) = \cos(\theta)$ $\quad \sin(-\theta) = -\sin(\theta)$

4 If $X \sim$ df F

$$\Phi_X(t) = E(e^{itX}) \quad \text{is the characteristic funct}^n$$

$$= \int_{-\infty}^{\infty} e^{itX}\,dP$$

$$= \int_{-\infty}^{\infty} e^{itx}\,dF(x)$$

$$= \int_{-\infty}^{\infty} e^{itx} f(x)\,dx$$

$\Phi_X(t) < \infty$ always since

$$|\Phi_X(t)| = \left| \int_{-\infty}^{\infty} e^{itX}\,dF(x) \right| \leq \int_{-\infty}^{\infty} |e^{itX}|\,dF(x) = \int_{-\infty}^{\infty} |\cos tx + i \sin tx|\,dF(x)$$

$$= \int_{-\infty}^{\infty} \sqrt{\cos^2(tx) + \sin^2(tx)} \; dF(x)$$

$$= \int_{-\infty}^{\infty} 1 \; dF(x)$$

$$= 1$$

So Though $|\Phi_X(t)| < \infty$ (always exists) $\forall \; t \in \mathbb{R}$

$\int_{-\infty}^{\infty} \Phi_X(t) \, dt$ is not always $< \infty$

In fact if $\int_{-\infty}^{\infty} |\Phi_X(t)| \, dt < \infty$ then X has a density

---

PROPERTIES

1. $\Phi_X(0) = 1$

$$\Phi_X(0) = E(e^{i0X}) = E(e^0) = E(1) = 1.$$

2. $|\Phi_X(t)| \leq 1$

$$|\Phi_X(t)| = \left| \int_{-\infty}^{\infty} e^{itx} dF(x) \right| \leq \int_{-\infty}^{\infty} |e^{itx}| dF(x)$$

$$= \int_{-\infty}^{\infty} |\cos(tx) + i\sin(tx)| \, dF(x)$$

$$= \int_{-\infty}^{\infty} \sqrt{\cos^2(tx) + \sin^2(tx)} \, dF(x) = 1$$

3. $\Phi_{aX+b}(t) = e^{itb}\,\Phi_X(at)$

$$\Phi_{aX+b}(t) = E\left(e^{i(aX+b)t}\right)$$

$$= E\left(e^{itaX+itb}\right)$$

$$= E\left(e^{itaX}\,e^{itb}\right)$$

$$= e^{itb}\,E\left(e^{i(at)X}\right)$$

$$= e^{itb}\,\Phi_X(at)$$

4. $\Phi_X(-t) = \overline{\Phi}_X(t)$

$$\Phi_X(-t) = E\left(e^{-itX}\right) = \int_{-\infty}^{\infty}\cos(-tx)\,dF(x) + i\int_{-\infty}^{\infty}\sin(-tx)\,dF(x)$$

$$= \int_{-\infty}^{\infty}\cos(-tx)\,dF(x) - i\int_{-\infty}^{\infty}\sin(tx)\,dF(x)$$

$$\begin{cases}\cos(-\theta)=\cos(\theta)\\ \sin(-\theta)=-\sin(\theta)\end{cases}$$

$$\Phi_X(t) = E\left[\cos(tX) + i\sin(tX)\right]$$

$$= \int_{-\infty}^{\infty}\cos(tx)\,dF(x) + i\int_{-\infty}^{\infty}\sin(tx)\,dF(x)$$

$$\overline{\Phi}_X(t) = \int_{-\infty}^{\infty}\cos(tx)\,dF(x) - i\int_{-\infty}^{\infty}\sin(tx)\,dF(x)$$

5. If $X_1 \ldots X_n \ldots$ are indep

then $\Phi_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} \Phi_{X_i}(t)$

Proof : $\Phi_{\Sigma X_i}(t) = E\left(e^{it \Sigma X_i}\right)$

$$= E\left(\prod_{i=1}^{n} e^{itX_i}\right)$$

$$= \prod_{i=1}^{n} E\left(e^{itX_i}\right) \qquad \{\text{by independence}$$

$$= \prod_{i=1}^{n} \Phi_{X_i}(t)$$

7. If X and X' are iid with chf Φ then

$Y \equiv X - X'$ has chf $|\Phi|^2$

(Note $Y \equiv X - X'$ is a symmetric r.v)

Proof : $\Phi_Y(t) = \Phi_{X-X'}(t) = \Phi_X(t) \cdot \Phi_{-X'}(t)$

$$= \Phi_X(t)\, \Phi_{X'}(-t)$$

$$= \Phi_X(t)\, \Phi_X(-t) \qquad \{\text{by iid}$$

$$= \underbrace{\Phi_X(t)}_{\text{complex \#}}\, \underbrace{\overline{\Phi}_X(t)}_{\text{complex conj}} = |\Phi_X(t)|^2$$

6. Φ is a real valued # iff $X \cong -X$ (ie X & $-X$ are iid)

Proof :

If $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2$

$$z_1 = z_2 \iff a_1 = a_2 \text{ and } b_1 = b_2.$$

If $z_1 = \bar{z}_1 \iff b = -b$

$$\iff b = 0$$

$$\iff z_1 \text{ is real.}$$

—— (*)

$$\Phi_X(t) = \int_{-\infty}^{\infty} \cos(tx)\, dF(x) + i \int_{-\infty}^{\infty} \sin(tx)\, dF(x)$$

"$\Leftarrow$" Now X and $-X$ are iid.

$$\Rightarrow \Phi_X(t) = \Phi_{-X}(t) = \Phi_X(-t) = \bar{\Phi}_X(t)$$

$$\Rightarrow \Phi_X(t) = \bar{\Phi}_X(t)$$

$$\Rightarrow \Phi_X(t) \text{ is real valued}$$

"$\rightarrow$" Φ_X is real valued $\Rightarrow \Phi_X(t) = \bar{\Phi}_X(t)$ {by (*)

$\Rightarrow \Phi_X(t) = \Phi_X(-t)$ {Prop 4

$\Rightarrow \Phi_X(t) = \Phi_{-X}(t)$

$\Rightarrow X \cong -X$ (r.v are uniquely det by chf

COMPLEMENTARY PAIRS OF R.V

(1) $U \sim f_U(\cdot) \rightarrow$ continuous

$W \sim f_W(\cdot) \rightarrow$ bdd & continuous with chf $\Phi_W(\cdot)$

and

(2) $$f_U(t) = c\,\Phi_W(-t)$$ for some constant c and $\forall t$

such pairs exist

Note : (2) $\Rightarrow$ Φ_W must be real valued (since density funct^n cannot be complex)

$\rightarrow$ W is symmetric

Also roles of U and W cannot be exchanged.

CONVOLUTION THEOREM

Let X and Y be indep r.v

$$P(X+Y \le r) = \iint_{\{x+y\le r\}} dF_X(x)\, dF_Y(y)$$

uses indep, product meas etc, induced meas.

any r.v has distn funct n but not necessarily a density

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} dF_Y(y)\; dF_X(x)$$

{ Fubini's

$$= \int_{-\infty}^{\infty} \underline{F_Y(r-x)\, dF_X(x)}$$

$$= \underline{F_Y * F_X (r)}$$

{ convolution

Now suppose Y has a density

$$F_{X+Y}(z) = \int_{-\infty}^{\infty} F_Y(z-x)\, dF_X(x)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_Y(y)\, dy\, dF_X(x)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_Y(v-x)\; dv\; dF_X(x)$$

using $v = y + x$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_Y(y-x)\; dy\; dF_X(x)$$

Use $v = y$

$$= \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f_Y(y-x)\, dF_X(x) \right) dy \qquad \text{(Fubini)}$$

$\therefore$ $F_{X+Y}(z)$ has density

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_Y(z-x)\, dF_X(x)$$

Thus given 2 indep random variables X, Y st Y has a density but X does not ; the r.v $X+Y$ still has a density

Note: Actually to show $f_{X+Y}(z)$ is the density of $X+Y$ we must show

$$P\left((X+Y)^{-1}(A)\right) = \int_A f_{X+Y}\, d\lambda \qquad \text{for } A \in \mathcal{B}$$

we have only proved this for $A = (-\infty, z)$

$\Longrightarrow$ can show that it holds for all $A \in \mathcal{B}$

Example

Let X be a r.v whose density does not exist

Let U is a r.v which has a density f_U

$\therefore \frac{1}{n}U$ also has a density

$$\frac{1}{n} \xrightarrow{P} 0 \Rightarrow \frac{1}{n}U \xrightarrow{d} 0 \text{ and } \frac{1}{n}U \xrightarrow{P} 0. \quad \text{—(*)}$$

$$X \xrightarrow{d} X \quad \text{——(**)}$$

$$\therefore \quad X + \frac{1}{n}U \xrightarrow{d} X \qquad \text{(by slutsky (*) \& (**))}$$

Now X has no density, $\frac{1}{n}U$ has a density

$\Rightarrow$ by convolution theorem $\left(X + \frac{1}{n}U\right)$ has a density also

However X was selected to be a r.v whose density does not exist

$\therefore$ If $X_n \xrightarrow{d} X$ and X_n's have densities need not imply X has a density

THEOREM)

To show chf uniquely defines a distⁿ on a line

$X_1 \sim F_{X_1} \qquad X_2 \sim F_{X_2} \qquad$ with chf $\Phi_{X_1}(\cdot) = \Phi_{X_2}(\cdot)$

$\Longleftrightarrow \quad X_1, X_2$ both $\sim F$

Proof. "$\Leftarrow$" If $\left.\begin{matrix} X_1 \sim F \\ X_2 \sim F \end{matrix}\right\} \rightarrow$ by def of chf $\Phi_{X_1}(\cdot) = \Phi_{X_2}(\cdot)$

"$\Longrightarrow$"

Inversion Theorem (establishes the uniqueness of the chf)

If X is a r.v with df F (we do not assume that a density exists) and chf Φ_X then

$$F_X(x_1) - F_X(x_2) = \lim_{a \to 0} \int_{x_1}^{x_2} f_a(x)\, dx \qquad \forall\ x_1, x_2 \text{ continuity pts of } F$$

where
$$f_a(t) = \int_{-\infty}^{\infty} e^{-itv}\, \Phi_X(v)\, c\, f_w(av)\, dv$$

If X does not have density and Y is a r.v with a density then $X+Y$ has a density

So W here is a r.v which has a complementary pair and is being added to X so that $(X+W)$ has a density

To show $\Phi_{X_1}(\cdot) = \Phi_{X_2}(\cdot) \Rightarrow X_1 \simeq X_2$

same as showing $X_1 \not\simeq X_2 \Rightarrow \Phi_{X_1}(\cdot) \neq \Phi_{X_2}(\cdot)$

Now

$$F_{X_1}(r_1) - F_{X_1}(r_2) = \lim_{a \to 0} \int_{r_1}^{r_2} \int_{-\infty}^{\infty} e^{itv} \Phi_{X_1}(v)\, c\, f_W(av)\, dv$$

$$F_{X_2}(r_1) - F_{X_2}(r_2) = \lim_{a \to 0} \int_{r_1}^{r_2} \int_{-\infty}^{\infty} e^{itv} \Phi_{X_2}(v)\, c\, f_W(av)\, dv$$

If $X_1 \not\simeq X_2$ then $\exists$ r_1^* and r_2^* (continuity pts) st

$$F_{X_1}(r_1^*) - F_{X_1}(r_2^*) \neq F_{X_2}(r_1^*) - F_{X_1}(r_1^*)$$

Now suppose if possible $\Phi_{X_1}(\cdot) = \Phi_{X_2}(\cdot)$

$$\Rightarrow F_{X_1}(r_1) - F_{X_1}(r_2) = F_{X_2}(r_1) - F_{X_2}(r_2) \quad \forall\ r_1, r_2 \text{ continuity pts}$$

$$\Rightarrow F_{X_1}(r_1^*) - F_{X_1}(r_2^*) = F_{X_2}(r_1^*) - F_{X_2}(r_2^*)$$

we have a contradiction $\Rightarrow$ our supposition is wrong

ie $\Phi_{X_1}(\cdot) \neq \Phi_{X_2}(\cdot)$

q.e.d.

THEOREM: Inversion Formula

(Given a chf how to recover the density)

General:

$$F_X(r_2) - F_X(r_1) = \lim_{a \to 0} \int_{r_1}^{r_2} f_a(t)\,dt$$

where $$f_a(t) = \int_{-\infty}^{\infty} e^{-itv}\, \Phi_X(v)\, c\, f_W(av)\, dv$$

Note the W need not be unique it can be any complementary pair; however the limit is unique.

Statement: If X has a chf $\Phi_X(\cdot)$ st $\int_{-\infty}^{\infty} |\Phi_X(t)|\,dt < \infty$

then X has a uniformly continuous density given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, \Phi_X(t)\, dt$$

Proof: Steps

1. $F_X(r_2) - F_X(r_1) = \lim_{a \to 0} \int_{r_1}^{r_2} f_a(t)\,dt$ $\underline{\text{show}}$ $\int_{r_1}^{r_2} f_X(x)\,dx$

2. Show uniformly continuous

1.)

$$\lim_{a\to 0} f_a(t) = \lim_{a\to 0} \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v) c\, f_W(av)\, dv$$

$$= \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v)\, c \lim_{a\to 0} f_W(av)\, dv \qquad \text{[by DCT}$$

$$\because \lim_{a\to 0} e^{-itv} \Phi_X(v) c\, f_W(av) = e^{-itv} \Phi_X(v) c\, f_W(0)$$

$$\leq \underbrace{|\Phi_X(v)||c f_W(0)|}_{\in \mathcal{L}_1}$$

since $|\Phi_X(v)| \in \mathcal{L}_1$
and W is complementary r.v
so it has bdd continuous density

$$\therefore \lim_{a\to 0} f_a(t) = \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v) \lim_{a\to 0} c\, f_W(av)\, dv$$

$$= \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v)\, c\, f_W(0)\, dv$$

$$= c\, f_W(0) \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v)\, dv$$

Let $W \sim N(0,1)$ and $c = \frac{1}{\sqrt{2\pi}}$ (W is any complementary r.v

$\therefore \quad \lim_{a \to 0} f_a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itv} \Phi_X(v)\, dv = f_X(t)$

Also we know

$$F_X(r_1) - F_X(r_2) = \lim_{a \to 0} \int_{r_1}^{r_2} f_a(t)\, dt$$

Claim $\lim_{a \to 0} f_a(t) = f_X(t)$ uniformly (bonus!!)

Now since $f_a(t) \longrightarrow f_X(t)$ uniformly

$$\Rightarrow \int_{r_1}^{r_1} f_a(t)\, dt \longrightarrow \int_{r_1}^{r_2} f_X(t)\, dt$$

$$\therefore \quad F_X(r_1) - F_X(r_2) = \lim_{a \to 0} \int_{r_1}^{r_2} f_a(t)\, dt$$

$$= \int_{r_1}^{r_2} f_X(t)\, dt$$

2. To show uniformly continuous

Any funct^n $f(x)$ is continuous at x_0 st

$\forall \varepsilon > 0$, $\exists$ $\delta_{(\varepsilon, x_0)}$ st $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$

A function $f(x)$ is uniformly continuous if $\forall \varepsilon > 0$ $\exists$ $\delta(\varepsilon)$ st
$|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$

Consider

$$\left|f_X(x) - f_X(y)\right| = \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(e^{-itx} - e^{-ity}\right) \Phi_X(t)\, dt \right|$$

$$= \left| \frac{1}{2\pi} e^{-ity} \int_{-\infty}^{\infty} \left(e^{-it(x-y)} - 1\right) \Phi_X(t)\, dt \right|$$

$$\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| e^{-it(x-y)} - 1 \right| \left| \Phi_X(t) \right| dt$$

$$\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| e^{-it(x-y)} - 1 \right| \left| \Phi_X(t) \right| dt$$

Now $e^{-itx} \longrightarrow 1$ as $x \longrightarrow 0$

$\therefore$ given $\varepsilon > 0$, $\exists\, \delta$ st $\left| e^{-itx} - 1 \right| < \varepsilon$ whenever $|x| < \delta$

$\therefore$ given $\varepsilon > 0$, $\exists\, \delta$ st $\left| e^{-it(x-y)} - 1 \right| < \varepsilon$ whenever $|x-y| < \delta$

So $$\left|f_X(x) - f_Y(y)\right| < \frac{\varepsilon}{2\pi} \int_{-\infty}^{\infty} \left|\Phi_X(t)\right| dt$$

$$< \frac{\varepsilon}{2\pi} \cdot M \qquad \left\{ \text{since } \int_{-\infty}^{\infty} |\Phi_X(t)|\, dt < \infty \right.$$

CLASSICAL CENTRAL LIMIT THEOREM

Theorem: $X_{n1}, \dots X_{nn} \dots$ are iid F with mean $\mu < \infty$ and var $\sigma^2 < \infty$

$$T_n \equiv X_{n1} + \dots + X_{nn}$$

$$\bar{X}_n = \frac{T_n}{n}$$

Then

$$\sqrt{n}\,(\bar{X}_n - \mu) = \frac{1}{\sqrt{n}}\,(T_n - \mu n)$$

$$= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_{nk} - \mu) \xrightarrow[n\to\infty]{d} N(0, \sigma^2)$$

Proof

$$\Phi_{\sqrt{n}(\bar{X}_n - \mu)}(t) = \Phi_{\left(\frac{\sum_{k=1}^{n}(X_{nk}-\mu)}{\sqrt{n}}\right)}(t)$$

$$= \prod_{k=1}^{n} \left\{ \Phi_{\frac{(X_{nk}-\mu)}{\sqrt{n}}}(t) \right\} \quad \text{by indep}$$

$$= \prod_{k=1}^{n} \left\{ \Phi_{X_{nk}-\mu}(t/\sqrt{n}) \right\}$$

$X \simeq (0, \sigma^2)$ then $\Phi_X(t) - \left(1 - \frac{\sigma^2 t^2}{2}\right) = o(t^2)$ as $t \to 0$

Pg 354 → Ineq 4.3

$$= \left[1 - \frac{\sigma^2}{2}\left(\frac{t}{\sqrt{n}}\right)^2 + \underbrace{o\left(\frac{t}{\sqrt{n}}\right)}\right]^n$$

$$= \left[1 - \frac{\sigma^2}{2}\left(\frac{t}{\sqrt{n}}\right)^2 + \overbrace{\left(\frac{t}{\sqrt{n}}\right)^2 r\left(\frac{t}{\sqrt{n}}\right)}\right]^n \quad \left\{ \text{where } r(u) \to 0 \text{ as } u \to 0 \right.$$

$$\equiv (1 + b_n)^n \quad \left\{ b_n = -\frac{\sigma^2 t^2}{n} + \frac{t^2}{n} r\left(\frac{t}{\sqrt{n}}\right) \right.$$

Now to show $(1+b_n)^n \longrightarrow \Phi_X(t)$ where $X \sim N(0, \sigma^2)$

$$\lim_{n\to\infty} (1+b_n)^n = \lim_{n\to\infty} \left[(1+b_n)^{1/b_n}\right]^{n b_n} \quad \left\{ \text{where } b_n \xrightarrow[n\to\infty]{} 0 \right.$$

$$= e^{\lim_{n\to\infty} (b_n n)} \quad \left\{ \text{where } (1+b_n)^{1/b_n} \longrightarrow e \right.$$

Now $$\lim_{n\to\infty} n b_n = \lim_{n\to\infty} n\left[-\frac{\sigma^2 t^2}{2n} + \frac{t^2}{n} r\left(t/\sqrt{n}\right)\right]$$

$$= \lim_{n\to\infty} \left\{-\frac{\sigma^2 t^2}{2} + t^2 r\left(t/\sqrt{n}\right)\right\}$$

$$= -\frac{\sigma^2 t^2}{2}$$

$$\therefore \quad (1+b_n)^n \xrightarrow[n\to\infty]{} e^{-\frac{\sigma^2 t^2}{2}}$$

$$\Rightarrow \quad \sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2) \qquad \left\{ \text{by Cramer-Lévy theorem.} \right.$$

THEOREM: $\boxed{\text{POISSON LIMIT THEOREM}}$

$X_{n1}, X_{n2}, \ldots, X_{nn}$ are indep Bernoulli (λ_{nk})

st $\quad \lambda_n = \sum_{k=1}^{n} \lambda_{nk} \longrightarrow \lambda$ ——— (1)

$\quad \sum_{k=1}^{n} \lambda_{nk}^2 \longrightarrow 0$ ——— (2)

then $\quad T_n \equiv (X_{n1} + X_{n2} + \cdots + X_{nn}) \xrightarrow{d} \text{Pois}(\lambda)$

Remark:

1) If $\lambda_{n1} = \lambda_{n2} = \cdots = \lambda_{nk} = \frac{U_n}{n}$ and $U_n \longrightarrow \lambda$

then (1) & (2) hold.

eg: If λ_{nk}'s $= \frac{k}{n^3}$

2) If $X_1, \cdots X_n \ldots$ are iid Ber (p) ie iid r.v

then $T_n = \text{Bin}(np, np(1-p))$

If $np \longrightarrow \lambda \quad \Rightarrow \quad T_n \longrightarrow \text{Pois}(\lambda)$

Proof: $\Phi_{X_{nk}}(t) = 1 + \lambda_{nk}(e^{it} - 1)$ $\quad \{ X_{nk} \sim Ber(\lambda_{nk})$

$$\Phi_{T_n} = \prod_{k=1}^{n} \Phi_{X_{nk}}(t)$$

$$= \prod_{k=1}^{n} \left[1 + \lambda_{nk}(e^{it} - 1) \right]$$

Pg 353 Lemma 4.3:

If (i) $\sum_{k=1}^{n} \theta_{nk} \longrightarrow \theta$

and (ii) $\max_{1 \le k \le n} |\theta_{nk}| \longrightarrow 0$

and (iii)

$$\Rightarrow \prod_{i=1}^{n} (1 + \theta_{nk}) \longrightarrow e^{\theta}$$

Consider $$\sum_{k=1}^{n} \lambda_{nk}(e^{it} - 1) = (e^{it} - 1) \sum_{k=1}^{n} \lambda_{nk}$$

$$\longrightarrow (e^{it} - 1)\lambda$$

cond^n (ii) & (iii) hold for us.

$$\therefore \quad \Phi_{T_n}(t) \longrightarrow e^{\lambda(e^{it} - 1)}$$

$\Rightarrow \quad T_n \longrightarrow Pois(\lambda)$ $\Big\{$ Cramer - Lévy and uniqueness.

Note : Pg 352 defines $\log(z)$ where z is complex So here using logs would not simplify the proof.

THEOREM : MULTIVARIATE CLT

$\underset{\sim}{X}_1 \cdots \underset{\sim}{X}_n$ are iid $(\underset{\sim}{\mu}, \Sigma)$; $X_j \in \mathbb{R}^m$

then $Z_n \equiv \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\underset{\sim}{X}_j - \underset{\sim}{\mu}) \xrightarrow{d} N_m(\underset{\sim}{0}, \Sigma)$

Proof : Let $\underset{\sim}{\lambda} \in \mathbb{R}^m$ (use cramer-wald device)

$Y_j \equiv \lambda^T(\underset{\sim}{X}_j - \underset{\sim}{\mu}) \cong (0, \lambda^T \Sigma \lambda)$

Y_j's are iid vectors

$$\sqrt{n}\,\overline{Y}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \lambda^T (x_j - \mu)$$

$$= \frac{\lambda^T}{\sqrt{n}} \sum_{j=1}^{n} (x_j - \mu)$$

$$= \lambda^T Z_n$$

Now if $\lambda^T Z_n \xrightarrow{d} Z \quad \forall \lambda$ then $Z_n \xrightarrow{d} Z$

Now $\quad Y_j \simeq (0, \underset{1\times d}{\lambda^T} \underset{d\times d}{\Sigma} \underset{d\times 1}{\lambda})$ are iid

so by univariate CLT $\quad \sqrt{n}\ \overline{Y}_n \xrightarrow{d} N(0, \lambda^T \Sigma \lambda)$

ie $\quad \lambda^T Z_n \xrightarrow{d} N(0, \lambda^T \Sigma \lambda)$

$$\Longrightarrow \quad \Phi_{\lambda^T Z_n}(t) \longrightarrow \exp\left\{-\frac{t^2(\lambda^T \Sigma \lambda)}{2}\right\}$$

Now if $Z \sim N(0, \Sigma)$

$$\Phi_{\lambda^T Z}(t) = \exp\left\{-t^2\left(\frac{\lambda^T \Sigma \lambda}{2}\right)\right\}$$

$$\therefore \quad \Phi_{\lambda^T Z_n}(t) \longrightarrow \Phi_{\lambda^T Z}(t) \qquad \forall \ \underset{\sim}{\lambda} \in \mathbb{R}^m$$

$$\Rightarrow \quad Z_n \xrightarrow{d} Z \qquad \{\because \text{ by Cramer-Wald}$$

CONDITIONAL EXPECTATION AND CONDITIONAL PROB

Def: $(\Omega, \mathcal{A}, P)$ is a prob space

$\mathcal{D}$ is a σ-field on $\mathcal{A}$

Let Y be a r.v on $(\Omega, \mathcal{A}, P)$ with $E|Y| < \infty$

The conditional expectation

$E(Y|\mathcal{D})$ is a $\mathcal{D}$-meas funct^n st

① $$\int_D E(Y|\mathcal{D})(\omega)\, dP(\omega) = \int_D Y(\omega)\, dP(\omega) \quad \forall\ D \in \mathcal{D}.$$

In particular

$\mathcal{D} \equiv \mathcal{F}(X)$ where X is a r.v on $(\Omega, \mathcal{A}, P)$

ie $\mathcal{D} \equiv \mathcal{F}(X) = X^{-1}(\mathcal{B})$

So $E(Y|\mathcal{F}(X)) \equiv E(Y/X)$ is the notation which is used.

PROPERTY:
We show $\underline{E(Y|\mathcal{D})(\cdot)}$ $\underline{\text{exists}}$ and $\underline{\text{is unique a.s}}$

Proof: Case I: $Y \geq 0$

Let us define a meas on $\mathcal{D}$ by

$$\nu(D) \equiv \int_D Y\, dP_{|\mathcal{D}} \quad \forall\ D \in \mathcal{D}.$$

P is a meas on $\mathcal{A} \Longrightarrow P_{/\mathcal{D}}$ is also a measure.

$$\Rightarrow \quad \nu << P_{/\mathcal{D}}$$

Thus by Radon-Nicodyn theorem:

$\exists$ uniquely a.s $P_{/\mathcal{D}}$ a $\mathcal{D}$-meas funct$^{\underline{n}}$ h st

$$\nu(D) = \int_D h \, dP_{/\mathcal{D}} = \int_D h \, dP$$

We showed $\exists$ a funct$^{\underline{n}}$ h which is $\mathcal{D}$-meas and unique a.s $P_{/\mathcal{D}}$ st

$$\int_D h \, dP = \int_D Y \, dP \qquad \forall \; D \in \mathcal{D}$$

NOTE: h is a funct$^{\underline{n}}$ on $\mathcal{D}$ whereas Y is a function on $\mathcal{A}$.

$$h \equiv \frac{d\nu}{dP_{/\mathcal{D}}} \equiv E(Y/\mathcal{D})$$ is called cond$^{\underline{n}}$ expectation.

PROP 2:

Now $\underline{E(y/\mathcal{F}(x)) \equiv E(y/x) \text{ is } \mathcal{F}(x) \text{ measurable}}$.

If Z is a r.v that is $\mathcal{F}(x)$ meas $\exists$ g on $(\mathbb{R}, \mathcal{B})$ st
$Z = g(X)$

So $\underline{E(y/\mathcal{F}(x))(\omega) \equiv E(y/x)(\omega) = g(X(\omega))}$

Def 1': $g(x) = E(Y \mid X = x)$ is a $\mathcal{B}$-meas funct^{n} on $\mathbb{R}$ st

$$\int_B E(Y \mid X = x)\, dP_X(x) = \int_{X^{-1}(B)} Y(\omega)\, dP(\omega) \qquad \forall\ B \in \mathcal{B}$$

COND^{N} PROB

Def 2: $P(A \mid \mathcal{D}) = E(I_A \mid \mathcal{D})$ for any $A \in \mathcal{A}$

NOTE: (1) $P(A \mid \mathcal{D})$ is a funct^{n} which is $\mathcal{D}$-meas unique a.s $P_{|\mathcal{D}}$

st

$$\int_D P(A \mid \mathcal{D})\, dP_{|\mathcal{D}} = \int_D I_A\, dP = P(A \cap D)$$

So

$$P(A \cap D) = \int_D P(A \mid \mathcal{D})\, dP_{|\mathcal{D}} \qquad \forall\ D \in \mathcal{D}$$

(2) let $\mathcal{D} = \mathcal{F}(X)$

$$P(A \mid \mathcal{F}(X)) \equiv P(A \mid X) \quad \text{(notation)}$$

NOTE: here the prob is a funct^{n} of $\mathcal{D}$

However when $\mathcal{D}$ is given fixed then you have a prob still denoted by $P(A|\mathcal{D})$

(3) $P(A|X)(\omega) = g(X(\omega))$ where $g(x) \equiv P(A|X=x)$
notation

satisfies

$$P(A \cap X^{-1}(B)) = \int_B P(A|X=x)\,dP_X(x) \qquad \text{for } B \in \mathcal{B}$$

PROPERTIES

(1) $0 \leq P(A|\mathcal{D}) \leq 1$ a.s P

(2) $P\left(\sum_{i=1}^{\infty} A_i \,|\, \mathcal{D}\right) = \sum_{i=1}^{\infty} P(A_i|\mathcal{D})$ a.s P

(3) $P(\phi|\mathcal{D}) = 0$ a.s P

(4) $A_1 \subset A_2 \implies P(A_1|\mathcal{D}) \leq P(A_2|\mathcal{D})$ a.s P

So the funct^n $P(A|\mathcal{D})$ behaves like a probability

INDEPENDENCE

$(\Omega, \mathcal{K}, P)$ is a prob space

1)

Let $\mathcal{A}_1 \dots \mathcal{A}_n$ be sub σ-fields of $\mathcal{K}$.

We say $\mathcal{A}_1 \dots \mathcal{A}_n$ are indep if

$$P(A_1 \cap A_2 \dots \cap A_n) = \prod_{i=1}^{n} P(A_i) \qquad \forall \quad A_i \in \mathcal{A}_i$$

2)

Let $X_1 \dots X_n$ be r.v.

For each X_i we have the sub σ-field generated by X_i

ie $\quad \mathcal{F}(X_i) = X_i^{-1}(\mathcal{B})$

$X_1 \dots X_n$ are indep if $\mathcal{F}(X_1) \dots \mathcal{F}(X_n)$ are indep σ-fields

** Reading Pg 152.

$\boxed{\text{Lecture 33}}$

We have $(\Omega, \mathcal{A}, P)$ a prob space.

Y is a random variable on $(\Omega, \mathcal{A}, P)$ st $E|Y| < \infty$

$\mathcal{D}$ is a sub σ-field of $\mathcal{A}$

Conditional expectation $E(Y/\mathcal{D})$ is a function ie a r.v which is $\mathcal{D}$ measurable.

(i) given any Y and $\mathcal{D}$ then $E(Y/\mathcal{D})$ exists and is unique a.s (proof using R-N theorem)

(ii) $$\int_D E(Y/\mathcal{D})\, dP(\omega) = \int_D Y(\omega)\, dP \qquad \forall\ D \in \mathcal{D}$$

$\boxed{\text{NOTE}}$ To prove uniqueness in (i) you use the fact that
$$\int_D h(\omega)\, dP(\omega) = \int_D h'(\omega)\, dP(\omega) \qquad \forall\ D \in \mathcal{D}$$
where both $h(\omega)$ and $h'(\omega)$ are $\mathcal{D}$-measurable

$$\Rightarrow \quad h(\omega) = h'(\omega)$$

[we call $E(Y/\mathcal{D})(\omega) \equiv h(\omega)$

<u>Property 1</u> : $\quad E(aY + bX \mid \mathcal{D}) = a E(Y \mid \mathcal{D}) + b E(X \mid \mathcal{D})$ a.s

Let $P_{|\mathcal{D}}$ denote the measure on the measurable space $(\Omega, \mathcal{D})$

st $\quad P_{|\mathcal{D}}(D) = P(D) \quad \forall\ D \in \mathcal{D}$

We need to show

$$\int_D E(aY + bX \mid \mathcal{D})\, dP_{|\mathcal{D}} = \int_D \left[a E(Y \mid \mathcal{D}) + b E(X \mid \mathcal{D})\right] dP_{|\mathcal{D}}$$

$\forall\ D \in \mathcal{D}$

Consider

$$LHS = \int_D (aY + bX)\, dP$$ property of conditional expectat^n

$\forall\ D \in \mathcal{D}$

$$= \int_D aY\, dP + \int_D bX\, dP$$

$$= \int_D a E(Y \mid \mathcal{D})\, dP_{|\mathcal{D}} + \int_D b E(X \mid \mathcal{D})\, dP_{|\mathcal{D}}$$ {from def^n of cond^n expec

– qed

Property 2: $E(Y) = E[E(Y|\mathcal{D})]$

$$E(Y) = \int_{\Omega} Y \, dP$$

$$E[E(Y|\mathcal{D})] = \int_{\Omega} E(Y|\mathcal{D}) \, dP$$

def of Expectat^n

by def^n of cond^n expect

$$= \int_{\Omega} Y \, dP$$

$\left\{ \begin{array}{l} \because \Omega \in \mathcal{D} \text{ becauce} \\ \mathcal{D} \text{ is a } \sigma\text{-field} \end{array} \right.$

$$= E(Y)$$

NOTE $\quad E(Y) = E\left(Y \mid \{\phi, \Omega\}\right)$ ← check this result

Property 3: If $Y \geq X$ a.s then

$$E(Y|\mathcal{D}) \geq E(X|\mathcal{D})$$

ie $\quad E(Y-X|\mathcal{D}) \geq 0.$

To show this we can show $\int_A E(Y-X|\mathcal{D})\, dP \geq 0 \quad \forall A \in \mathcal{A}$

Here we use instead R-N theorem

Consider

$$\int_D E(Y-X \mid \mathcal{D})\, dP_{/\mathcal{D}} = \int_D (Y-X)\, dP$$

$$\equiv \nu(D) \qquad \nu \text{ is a meas.}$$

where $\nu << P$ and R-N derivative $\frac{d\nu}{dP}$ (by R-N theorem

$$\Rightarrow \frac{d\nu}{dP} \geq 0 \text{ a.s}$$

Property 4: MCT for conditional expectation

$X_n \geq 0$ and $X \geq 0$ and $X_n \uparrow X$ a.s then.

$$E(X_n \mid \mathcal{D}) \uparrow E(X \mid \mathcal{D})$$

Proof: $X_{n+1} \geq X_n$ since X_n's $\uparrow$

$\therefore E(X_{n+1} \mid \mathcal{D}) \geq E(X_n \mid \mathcal{D})$ so $E(X_n \mid \mathcal{D})$ is $\uparrow$ seq

Also $X \geq X_n \quad \forall n$

$$\Rightarrow E(X \mid \mathcal{D}) \geq \cdots \geq E(X_{n+1} \mid \mathcal{D}) \geq E(X_n \mid \mathcal{D})$$

the seq of r.v $E(X_n \mid \mathcal{D})$ is bdd above by $E(X \mid \mathcal{D})$ (a.s.)

$\Rightarrow$ limit exists

$$\int_D \lim_{n\to\infty} E(X_n \mid \mathcal{D})(\omega)\, dP_{|\mathcal{D}}(\omega) = \lim_{n\to\infty} \int_D E(X_n \mid \mathcal{D})\, dP_{|\mathcal{D}}$$ { by MCT applied to $E(X_n \mid \mathcal{D})$)

$$= \lim_{n\to\infty} \int_D X_n \, dP$$ { by def of condⁿ expectatⁿ

$$= \int_D \lim_{n\to\infty} X_n \, dP$$ { MCT applied to $\{X_n\}$

$$= \int_D X \, dP$$

$$= \int_D E(X \mid \mathcal{D})\, dP \qquad \forall\, D \in \mathcal{D}$$

$$\Rightarrow \quad E(X_n \mid \mathcal{D}) \longrightarrow E(X \mid \mathcal{D})$$

Property : If Y is $\mathcal{D}$ meas, $XY \in \mathcal{L}_1(P)$ then

$$E(XY \mid \mathcal{D}) = Y E(X \mid \mathcal{D}) \quad \text{a.s } P_{\mid \mathcal{D}}$$

[Here Y is like a "known" when $\mathcal{D}$ is known.]

Proof : We will prove for the indicator junctⁿ

Let $Y = I_{D^*}$ for some $D^* \in \mathcal{D}$

We will show $\int E(XY \mid \mathcal{D})\, dP = \int Y E(X \mid \mathcal{D})\, dP$

$$\Rightarrow \quad E(XY \mid \mathcal{D}) = Y E(X \mid \mathcal{D})$$

Consider for $D \in \mathcal{D}$ arbitrary

$$\int_D E(XY \mid \mathcal{D})\, dP_{\mid \mathcal{D}} = \int_D E(X I_{D^*} \mid \mathcal{D})\, dP$$

$$= \int_D X I_{D^*}\, dP \qquad \{\text{by def}$$

$$= \int_{DD^*} X\, dP$$

$$= \int_{DD^*} E(X \mid \mathcal{D})\, dP \qquad \{\text{by def} \because DD^* \in \mathcal{D}$$

$$= \int_D I_{D^*} E(X \mid \mathcal{D})\, dP = \int_D Y E(X \mid \mathcal{D})\, dP$$

qed.

Property : Y, X_1, X_2 are r.v

If $\mathcal{F}(Y, X_1)$ is indep of $\mathcal{F}(X_2)$ then $E(Y|X_1, X_2) = E(Y|X_1)$

Proof: To show $\int_D E(Y|\mathcal{F}(X_1,X_2))\,dP_{12} = \int_D E(Y|\mathcal{F}(X_1))\,dP_{12}$

Now $\mathcal{F}(X_1, X_2) = \sigma[D]$ st $D \equiv D_1 \cap D_2$ $\quad \forall D \in \mathcal{F}(X_1,X_2)$.

$\equiv X_1^{-1}(B_1) \cap X_2^{-1}(B_2)$; $B_1, B_2 \in \mathcal{B}$.

Let $\nu(D) = \int_D Y\,dP$ with Y- $\mathcal{D}$ measurable is a meas

$\therefore$ our problem reduces to showing the equality of 2 measures

$$\nu_1(D) \equiv \int_D E(Y|\mathcal{F}(X_1,X_2))\,dP = \int_D E(Y|\mathcal{F}(X_1))\,dP \equiv \nu_2(D) \quad \forall D \in \mathcal{F}(X_1,X$$

By the caratheodory extension theorem it is enough to show

$\nu_1(D) = \nu_2(D)$ $\quad \forall D \in D_1 \cap D_2$ $\left\{ \text{on the generators of } \mathcal{F}(X_1, X_2) \right)$

Consider

$$\nu_1(D) = \int_D E(Y|X_1,X_2)\,dP = \int_D Y\,dP = \int_{D_1 \cap D_2} Y\,dP$$

$$= \int I_{D_1 \cap D_2}\, Y\,dP$$

$$= \int I_{D_2} I_{D_1} Y\,dP$$

$$= \int I_{X_2^{-1}(B_2)} \; I_{X_1^{-1}(B_1)} \; Y \, dP$$

$$= \int I_{X_2^{-1}(B_2)} \, dP \quad \int I_{X_1^{-1}(B_1)} \, y \, dP$$

$\left\{ \begin{array}{l} \because (X_1, y) \text{ is indep of } X_2 \\ \text{so is any funct}^{\underline{n}} \\ E(uv) = E(u)E(v) \end{array} \right.$

$$= \int I_{D_2} \, dP \; \int_{D_1} y \, dP$$

$$= \int I_{D_2} \, dP \quad \int_{D_1} E(y \mid X_1) \, dP$$

$\{$ by def$^{\underline{n}}$ of cond$^{\underline{n}}$ expectation

$$= \int I_{D_2} \, I_{D_1} \; E(y \mid X_1) \, dP$$

$\{$ by indep

$$= \int_{D_1 \cap D_2} E(y \mid X_1) \, dP$$

$$= \int_{D} E(y \mid X_1) \, dP$$

$$= \nu_2(D)$$

$$\mathcal{H} \equiv \mathcal{L}_2(P) = \{X \text{ on } (\Omega, \mathcal{A}, P) \mid E(X^2) < \infty\}$$

$\mathcal{H}$ is called a Hilbert space with inner product

$$\langle X, Y \rangle = E(XY)$$

Let $\mathcal{D} \subset \mathcal{A}$

$$\mathcal{H}_{\mathcal{D}} = \{X \in \mathcal{L}_2(P) : \mathcal{F}(X) \subset \mathcal{D}\}$$

ie the space of r.v X which are $\mathcal{D}$-meas.

We define an operator

$$P_{\mathcal{D}} : \mathcal{H} \longrightarrow \mathcal{H}_1 \subseteq \mathcal{H} \quad \text{st} \quad P_{\mathcal{D}}(y) \equiv E(y/\mathcal{D})$$

We can show $P_{\mathcal{D}}$ is a projection operator

To show

$$\langle y - E(y|X), X \rangle = 0$$

ie $E[[y - E(y|X)]X] = 0$

Consider $E(Xy - XE(y|X))$

$$= E(Xy) - E[XE(y|X)]$$

$$= E(Xy) - E[E(Xy|X)]$$

$$= E(Xy) - E(Xy)$$

$$= 0.$$

Recitation – Chp 34.

$\mathcal{H} = \{ X \text{ on } (\Omega, \mathcal{A}, P) \mid E(X^2) < \infty \}$

$\mathcal{H}_{\mathcal{F}(X_1)} = \{ X_m \text{ on } (\Omega, \mathcal{A}, P) \mid X_m\text{'s are } \mathcal{F}(X_1)\text{-meas}, E(X_i^2) < \infty \}$

$\mathcal{H}$ is a vector space

$\mathcal{H}_{\mathcal{F}(X_1)}$ is a subspace of $\mathcal{H}$

$P_{\mathcal{D}}(Y) \equiv E(Y \mid \mathcal{D})$ is a proj$^{\underline{n}}$ operator from $\mathcal{H}$ to $\mathcal{H}_{\mathcal{D}}$ (linear funct$^{\underline{n}}$)

$\langle X, Y \rangle \equiv E\big((X - E(X))(Y - E(Y))\big)$

<u>Properties</u>

(i) $P_{\mathcal{D}}(\cdot)$ is a ppo onto $\mathcal{H}_{\mathcal{D}}$

$P_{\mathcal{D}}^2 = P_{\mathcal{D}}$ $\quad \begin{cases} P_{\mathcal{D}}^2 = P_{\mathcal{D}}[E(Y \mid \mathcal{D})] = E(E(Y \mid \mathcal{D}) \mid \mathcal{D}) = E(Y \mid \mathcal{D}) = P_{\mathcal{D}} \\ \because \text{ if } X \text{ is } \mathcal{D} \text{ meas} \Rightarrow E(X \mid \mathcal{D}) = E(X) \end{cases}$

(ii) $(I - P_{\mathcal{D}})$ is a proj on $\mathcal{H}_{\mathcal{D}}^{\perp}$

$\langle (I - P_{\mathcal{D}})Y, P_{\mathcal{D}} Y \rangle = 0$

(iii) $\langle P_{\mathcal{D}} Y, Z \rangle = \langle Y, P_{\mathcal{D}} Z \rangle$

Def: We say $\mathcal{H}_1 \perp \mathcal{H}_2$ if $\langle g_1(X_1), g_2(X_2) \rangle \equiv E[g_1(X_1)\, g_2(X_2)] = 0$

$\forall\ g_1 \in \mathcal{H}_1 \quad g_2 \in \mathcal{H}_2$

operty : X_1 is indep of X_2 $\longleftrightarrow$ $\mathcal{H}_1 \perp \mathcal{H}_2$

Proof : $\langle g_2(X_2) - E(g_2(X_2) | g_1(X_1)) ; g_1(X_1) \rangle$

$$= \langle (I - P_{\mathcal{D}}) g_2(X_2) , g_1(X_1) \rangle$$

$$= 0$$

$g_1(X_1) \in \mathcal{H}_1$

$(I - P_{\mathcal{D}}) g_2(X_2) \in \mathcal{H}_1^{\perp}$

$\Rightarrow \langle g_2(X_2), g_1(X_1) \rangle = \langle E(g_2(X_2) | g_1(X_1)), g_1(X_1) \rangle$

Given X_1 indep of X_2

$$\langle g_2(X_2), g_1(X_1) \rangle = \langle E(g_2(X_2) | g_1(X_1)), g_1(X_1) \rangle$$

$$= \langle E(g_2(X_2)), g_1(X_1) \rangle \qquad \{ X_1 \text{ indep of } X_2$$

$$= E\{[E(g_2(X_2)) - E(g_2(X_2))][g_1(X_1) - E(g_1(X_1))]\}$$

$$= E\{0 \cdot [g_1(X_1) - E(g_1(X_1))]\}$$

$$= 0$$

$\Rightarrow \mathcal{H}_1 \perp \mathcal{H}_2$

Given $\mathcal{H}_1 \perp \mathcal{H}_2$ ie $\langle g_1(X_1), g_2(X_2) \rangle = 0$

$$\Rightarrow E\{[g_1(X_1) - E(g_1(X_1))][g_2(X_2) - E(g_2(X_2))]\} = 0$$

$$\Rightarrow E[g_1(X_1) g_2(X_2)] - E(g_1(X_1)) E(g_2(X_2)) = 0$$

$$\Rightarrow E[g_1(X_1) g_2(X_2)] = E(g_1(X_1)) E(g_2(X_2)) \qquad \forall g_1, g_2$$

Property : $\langle (I-P_{\mathcal{D}})Y , Z \rangle = 0$ for any $Y \in \mathcal{H}$, $Z \in \mathcal{H}_{\mathcal{D}}$

Proof : $\langle (I-P_{\mathcal{D}})Y, Z \rangle = \langle Y - E(Y|\mathcal{D}) , Z \rangle$

$$= \langle Y, Z \rangle - \langle E(Y|\mathcal{D}), Z \rangle$$

$$= E(YZ) - E(Z\, E(Y|\mathcal{D}))$$

$$= E(YZ) - E(E(ZY|\mathcal{D})) \qquad \left\{ \because Z \text{ is } \mathcal{D} \text{ meas.} \right.$$

$$= E(YZ) - E(YZ)$$

$$= 0$$

Note : $\langle (I-P_{\mathcal{D}})Y , P_{\mathcal{D}} Y \rangle = 0$

$$\langle Y - E(Y|X_1) , X_1 \rangle = 0$$

Property : $\mathcal{H}_1 \equiv \mathcal{H}_{\mathcal{F}(X_1)} \equiv \{ g_1(X_1) \text{ which are } \mathcal{F}(X_1) \text{ meas st } E(X_1^2) < \infty \}$

$\mathcal{H}_2 \equiv \mathcal{H}_{\mathcal{F}(X_2)} \equiv \{ g_2(X_2) \text{ which are } \mathcal{F}(X_2) \text{ meas st } E(X_2^2) < \infty \}$

$P_1(Y) \equiv E(Y|X_1)$ proj from $\mathcal{H}$ to $\mathcal{H}_1$

$P_2(Y) \equiv E(Y|X_2)$ proj from $\mathcal{H}$ to $\mathcal{H}_2$

$\mathcal{H}_+ \equiv \mathcal{H}_1 + \mathcal{H}_2$

$$\equiv \{ g_1(X_1) + g_2(X_2) : g_i \text{ is measurable } \& \ E[g_i^2(X_i)] < \infty \}$$

$E[Y|g(X_1)\, g(X_2)] = E[Y|g_1(X_1)] + E[Y|g_2(X_2)]$ if X_1 & X_2 are indep.

Sol^n Let P_+ denote the proj of $\mathcal{H}$ to $\mathcal{H}_+$

$\Rightarrow \langle Y - P_+ Y , P_+ Y \rangle = 0 \quad \forall \ Y \in \mathcal{H}$

We need to show $E(Y \mid g(X_1), g(X_2)) = E(Y \mid g(X_1)) + E(Y \mid g(X_2))$

ie to show $P_+(Y) = P_1(Y) + P_2(Y)$

ie to show $P_1 + P_2$ is a proj on $\mathcal{H}_+$

$Y \in \mathcal{H}$
$Z \in \mathcal{H}_+$

Consider $\langle (I - P_1 - P_2) Y, Z \rangle = \langle Y - E(Y \mid X_1) - E(Y \mid X_2), Z \rangle$

$$= \langle Y - E(Y \mid X_1) - E(Y \mid X_2), g_1(X_1) + g_2(X_2) \rangle$$

$$= \langle Y - E(Y \mid X_1), g_1(X_1) \rangle + \langle Y - E(Y \mid X_2), g_2(X_2) \rangle$$

$$- \langle E(Y \mid X_1), g(X_1) \rangle - \langle E(Y \mid X_2), g(X_2) \rangle$$

$$= \langle (I - P_1) Y, g_1(X_1) \rangle + \langle (I - P_2) Y, g_2(X_2) \rangle$$

$$- \langle E(Y \mid X_2), g_1(X_1) \rangle - \langle E(Y \mid X\ ), g_2(X_2) \rangle$$

$$= 0.$$

$$\langle E(Y \mid X_1), g_2(X_2) \rangle = E\left[g_2(X_2)\, E(Y \mid X_1) \right]$$

$$= E(g_2(X_2))\, E(E(Y \mid X_1))$$ {by indep of X_1 & X_2

$$= 0$$ {by assumption