Characteristics and Convergence

Characteristic Functions

Defn.: Let X be a random variable on a probability space $(\Omega, \mathcal{B}, P)$. The characteristic function of X is the function $\phi_X : \mathbb{R} \to \mathbb{C}$ defined by $\phi_X(t) = E(e^{itX})$.

Elementary Property of characteristic function.

(a) $\phi_X(0) = 1$, (b) $|\phi_X(t)| \le 1$, (c) $\phi_X(-t) = \overline{\phi_X(t)}$, the complex conjugate of $\phi_X(t)$.

Characteristic functions of standard distribution

Distribution	Characteristic Function $\phi(t)$, $t \in \mathbb{R}$
Bernoulli(p)	$1 - p + pe^{it}$
Bionomial (n, p)	$(1 - p + pe^{it})^n$
Uniform$(\{1, 2, \cdots n\})$	$\dfrac{e^{it}(1 - e^{it})}{n(1 - e^{int})}$
Poisson(λ)	$e^{\lambda(e^{it} - 1)}$
Uniform(a, b)	$\dfrac{e^{ibt} - e^{iat}}{i(b-a)t}$
Normal (m, σ^2)	$e^{imt - \frac{t^2\sigma^2}{2}}$
Geometric(p)	$\dfrac{pe^{it}}{1 - (1-p)e^{it}}$
Exponential(λ)	$\dfrac{\lambda}{\lambda - it}$

Th. (Inversion Theorem): Let X be a random variable with characteristic function $\varphi_X(\cdot)$. Assume that $\int_{\mathbb{R}} |\varphi_X(t)|\,dt < \infty$. Then X has a density function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t)\,dt.$$

Proof:

Lemma: Let X be a real valued random variable with characteristic function $\varphi_X(\cdot)$. Let $Z \stackrel{d}{=} N(0,1)$ be independent of X. For each $\sigma > 0$, the random variable $X_\sigma = X + \sigma Z$ has a density f_σ given by

$$f_\sigma(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t)\, e^{-\frac{\sigma^2 t^2}{2}}\,dt,$$

for all $x \in \mathbb{R}$.

Proof of Lemma: For $\sigma > 0$. Using the independence of X and Z, we have

$$P(X_\sigma \le \alpha) = \int_{\mathbb{R}} F_Z\left(\frac{\alpha - x}{\sigma}\right) d\mu_X(x)$$

$[\because X_\sigma = X + \sigma Z]$

$$= \int_{\mathbb{R}} \int_{-\infty}^{\frac{\alpha - x}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}}\,da\,d\mu_X(x)$$

$$= \int_{\mathbb{R}} \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}\,dy\,d\mu_X(x)$$

$\left[\text{Put } a = \frac{y-x}{\sigma}\right]$

$$= \int_{-\infty}^{\alpha} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}\,d\mu_X(x)\,dy.$$

$$= \int_{-\infty}^{\alpha} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}\,d\mu_X(x)\,da$$

Let $Y \overset{d}{=} N(0, \frac{1}{\sigma^2})$ be a random variable independent of X and Z. So $\phi_Y(t) = e^{-\frac{t^2}{2\sigma^2}}$, $t \in \mathbb{R}$. Using the independence of X and Y and the definition of a characteristic function, we have

(3)

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dP_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \phi_Y(x-a) \, dP_X(x)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} E\left(e^{i(X-a)Y}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} E\left(e^{iXY} \cdot e^{-iaY}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} E\left(\phi_X(Y) e^{-iaY}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \frac{d\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iay} \phi_X(y) e^{-\frac{a^2 y^2}{2}} dy$$

$$\left[\because Y \equiv N(0, \tfrac{1}{\sigma^2})\right]$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iay} \phi_X(y) e^{-\frac{a^2 y^2}{2}} dy \qquad (2)$$

From (1) & (2) we get

$$P(X_\sigma \le \alpha) = \int_{-\infty}^{\alpha} \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iay} \phi_X(y) e^{-\frac{a^2 y^2}{2}} dy \, da$$

$$= ~~dy \int_{-\infty}^{\alpha} \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iay} \phi_X(y) e^{-\frac{a^2y^2}{2}} da\, dy~~$$

$$= \int_{-\infty}^{\alpha} f_\sigma(a) \, da$$

$\therefore f_\sigma(x)$ is the density function

Proof of thm: Let f ~~: R → R~~ be defined by

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\,\phi_X(t)\,dt, \quad x \in \mathbb{R}.$$

The hypothesis on $\phi_X(\cdot)$ implies that f is a bounded complex function. Fix $\sigma > 0$. Consider X_σ as in above lemma and denote its density function by $f_\sigma(\cdot)$. Now

$$|f(x) - f_\sigma(x)| = \frac{1}{2\pi}\left|\int_{\mathbb{R}} e^{-itx}\,\phi_X(t)\left(1 - e^{-\frac{\sigma^2 t^2}{2}}\right)dt\right|$$

$$\leq \frac{1}{2\pi}\int_{\mathbb{R}} |\phi_X(t)|\left(1 - e^{-\frac{\sigma^2 t^2}{2}}\right)dt$$

Now for all $t \in \mathbb{R}$, $|\phi_X(t)|\left(1 - e^{-\frac{\sigma^2 t^2}{2}}\right) \to 0$ as $\sigma \to 0$ and is less than or equal to $|\phi_X(t)|$. The dominated convergence theorem shows ~~that~~

$$\sup_{x \in \mathbb{R}} |f_\sigma(x) - f(x)| \to 0$$

as $\sigma \to 0$.

So, f is real-valued.

Let $a \leq b \in \mathbb{R}$. Define a sequence of functions $g_n : \mathbb{R} \to \mathbb{R}$ by

$$g_n(x) = \begin{cases} n(x-a) & \text{if } x \in [a, a+\frac{1}{n}] \\ 1 & \text{if } x \in [a+\frac{1}{n}, b] \\ -n(x-b-\frac{1}{n}) & \text{if } x \in [b, b+\frac{1}{n}] \\ 0 & \text{otherwise.} \end{cases}$$

Now, $g_n \to g$ with $g = 1_{(a,b]}$ and $X_\sigma \to X$ as $\sigma \to 0$. Using above lemma and applying dominated convergence we obtain

$$P(a < X \le b) = E(g(X))$$

$$= \lim_{n\to\infty} E(g_n(X))$$

$$= \lim_{n\to\infty} \lim_{\sigma\to 0} E(g_n(X_\sigma))$$

$$= \lim_{n\to\infty} \lim_{\sigma\to 0} \int_{\mathbb{R}} g_n(x) f_\sigma(x)\, dx$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x) f(x)\, dx$$

$$= \int_{\mathbb{R}} g(x) f(x)\, dx$$

As ~~the~~ above holds for arbitrary $a \le b, \in \mathbb{R}$ we can conclude that f is the density of X.

Th. (Uniqueness Theorem) Two random variables X and Y have the same distribution if and only if $\phi_X(t) = \phi_Y(t)$ for all $t \in \mathbb{R}$

Proof:

Lemma: Let μ_1 and μ_2 be two probability measures on $(\Omega, \mathcal{B})$ and

$$C = \left\{ f : \mathbb{R} \to \mathbb{R} : f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, a, \sigma \in \mathbb{R} \right\}.$$

Suppose $\int f\, d\mu_1 = \int f\, d\mu_2$, for all $f \in C$,

then $\mu_1 = \mu_2$.

Proof of the

From the definition of characteristic function, it is trivial that if X and Y have the same distribution, then their characteristic functions are the same.

For the converse we will show that $E(f(X)) = E(f(Y))$, $\forall f \in C$ where C is as in above lemma, this will imply that X and Y have the same distribution. Let μ_X denote the distribution of X and $f \in C$. Then f is the density of a $N(a, \sigma^2)$ random variable.

So, $E(f(X)) = \int f(x)\, d\mu_X(x)$

$$= \int \left(\frac{1}{2\pi} \int e^{-iat - \frac{t^2}{2\sigma^2}} e^{-ixt}\, dt \right) d\mu_X(x)$$

(by inversion formula)

$$= \frac{1}{2\pi} \int e^{-iat - \frac{t^2}{2\sigma^2}} \int e^{-ixt}\, d\mu_X(x)\, dt$$

$$= \frac{1}{2\pi} \int e^{-iat - \frac{t^2}{2\sigma^2}} \phi_X(-t)\, dt$$

Similarly, $E(f(Y)) = \frac{1}{2\pi} \int e^{-iat - \frac{t^2}{2\sigma^2}} \phi_Y(-t)\, dt$

Since the characteristic function are equal i.e.

$\phi_X(t) = \phi_Y(t)$ $\forall t \in \mathbb{R}$

$\therefore$ $E(f(X)) = E(f(Y))$

or, $\int f(x)\, d\mu_X(x) = \int f(y)\, d\mu_Y(y)$

By above lemma, $\mu_X = \mu_Y$.

Modes of Convergence

Def^n: A sequence $\{X_n : n = 1, 2, \cdots\}$ of random variables on $(\Omega, \mathcal{B}, P)$ is said to

(i) converge almost everywhere to X, $(X_n \xrightarrow{a.e} X)$, if there exists a P-null set N such that $\{X_n(\omega)\}$ converges to $X(\omega)$ whenever $\omega \notin N$;

(ii) converge in probability to X, $(X_n \xrightarrow{P} X)$, if for every $\epsilon > 0$, $\lim_{n\to\infty} P(\{\omega : |(X_n - X)(\omega)| > \epsilon\}) = 0$;

(iii) converges in the rth mean to X, $(X_n \xrightarrow{r} X)$, if $E(|X_n - X|^r) \longrightarrow 0$; and

(iv) converge in distribution to X, $(X_n \xrightarrow{d} X)$, if $F_{X_n}(x) \longrightarrow F_X(x)$ for all continuity points x of F_X. This mode of convergence is also referred to as weak convergence.

Note: ~~Above definition~~

(***) **Theorem 1:** Let X and $\{X_n : n \geq 1\}$ be random variables on $(\Omega, \mathcal{B}, P)$.

(a) $X_n \xrightarrow{a.e} X$ implies that $X_n \xrightarrow{P} X$.

(b) $X_n \xrightarrow{r} X$ for some $r \geq 1$, implies that $X_n \xrightarrow{P} X$.

(c) $X_n \xrightarrow{P} X$ implies that $X_n \xrightarrow{d} X$.

Proof:

(a) we are given that

$$1 = P(\lim_{n\to\infty} X_n = X) = P\left(\bigcap_{\epsilon>0} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^{\epsilon}\right), \quad (1)$$

where $A_n^{\epsilon} = \{|X_n - X| < \epsilon\}$. Let $B_m^{\epsilon} = \bigcap_{n=m}^{\infty} A_n^{\epsilon}$.

Let $\epsilon_0 > 0$ be given. Since $\{B_m^{\epsilon_0}\}$ is an increasing sequence of sets, by continuity from ~~above~~ below $P(B_m^{\epsilon_0}) \uparrow P(\bigcup_{m=1}^{\infty} B_m^{\epsilon_0})$.

i.e. $$P\left(\bigcup_{m=1}^{\infty} B_m^{\epsilon_0}\right) = \lim_{m\to\infty} P(B_m^{\epsilon_0})$$

or, $$1 = \lim_{m\to\infty} P(B_m^{\epsilon_0}) \quad (\text{by } 1)$$

Hence for all $\delta > 0$, $\exists N$ such that

$$P(B_m^{\epsilon_0}) > 1-\delta \text{ for all } m \geq N.$$

As, $B_m^{\epsilon_0} \subseteq A_m^{\epsilon_0}$, we have shown that for all $\delta > 0$, $\exists N$,

$$P(\{|X_n - X| < \epsilon_0\}) > 1-\delta, \quad \forall m \geq N$$

As ϵ_0 was arbitrary, $X_n \xrightarrow{P} X$.

(b) $$E(|X_n - X|^p) \geq E(|X_n - X|^p)\, 1_{\{|X_n - X| \geq \epsilon\}}$$

$$\geq \epsilon^p P(|X_n - X| \geq \epsilon)$$

(by Tchebychev's inequality)

$\therefore$ if $E(|X_n - X|^p) \to 0$ as $n \to \infty$

then $P(|X_n - X| \geq \epsilon) \to 0$ as $n \to \infty$.

$\therefore X_n \xrightarrow{P} X$

(c) Let $\epsilon > 0$. By definition, $F_{X_n}(t) = P(X_n \le t)$.

Hence $F_{X_n}(t) = P(X_n \le t, |X_n - X| > \epsilon) + P(X_n \le t, |X_n - X| \le \epsilon)$

$\le P(|X_n - X| > \epsilon) + P(X_n \le t, |X_n - X| \le \epsilon)$

$\le P(|X_n - X| > \epsilon) + P(X \le t + \epsilon)$

$= P(|X_n - X| > \epsilon) + F_X(t+\epsilon)$ —— (1)

Similarly,

$F_X(t-\epsilon) \le F_{X_n}(t) + P(|X_n - X| \ge \epsilon)$ —— (2)

From (1), $\limsup_{n\to\infty} F_{X_n}(t) \le F_X(t+\epsilon)$ $(\because X_n \xrightarrow{P} X)$

From (2), $\liminf_{n\to\infty} F_{X_n}(t) \ge F_X(t-\epsilon)$ $(\because X_n \xrightarrow{P} X)$

$\therefore F_X(t-\epsilon) \le \liminf_{n\to\infty} F_{X_n}(t) \le \limsup_{n\to\infty} F_X(t) \le F_X(t+\epsilon)$

Take $\epsilon \to 0$ and use the continuity points t of F_X we have

$\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$.

Note: The following example show that the converses of the above statements are not true.

Example: Let $\Omega = [0,1]$, $\mathcal{B} = \mathcal{B}_{[0,1]}$, $P(dx) = dx$

(a) Let $X_n = 1_{[\frac{j}{2^k}, \frac{j+1}{2^k}]}$, if $n = 2^k + j$, for some $j = 0, 1, 2, \ldots, 2^k - 1$ and $k = 1, 2, \ldots$. If we let $A_n = \{X_n > 0\}$, then clearly $P(A_n) \to 0$.

Consequently $X_n \xrightarrow{P} 0$ but $X_n(\omega) \not\to 0$ for all $\omega \in \Omega$.

(b) Let $X_n = n 1_{(0,\frac{1}{n})}$ and $X \equiv 0$. For any $\epsilon > 0$, then

$$P(|X_n| > \epsilon) \leq \frac{1}{n} \quad \forall n.$$

Hence $X_n \xrightarrow{P} 0$. Clearly, $X_n \stackrel{r}{\not\to} X \;\forall r \geq 1$.

Theorem 2: (Slutsky's theorem) Let $\{X_n, X, Y_n\}$ be random variables on a probability space $(\Omega, \mathcal{B}, P)$. Let $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$, where $c \in \mathbb{R}$. Then,

(1) $X_n + Y_n \xrightarrow{d} X + c$

(2) $X_n Y_n \xrightarrow{d} cX$

(3) $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$, if $c \neq 0$.

Proof: (1) Let $\epsilon > 0$ be given. Write $F_n = F_{X_n + Y_n}$. Choose t such that t, $t-c+\epsilon$, $t-c-\epsilon$ are all continuity points of F_X.

Now, $F_n(t) = P(X_n + Y_n \leq t)$

$$\leq P(X_n + Y_n \leq t, |Y_n - c| < \epsilon) + P(|Y_n - c| \geq \epsilon)$$

$$\leq P(X_n \leq t - c + \epsilon) + P(|Y_n - c| \geq \epsilon)$$

$$\limsup_{n\to\infty} F_n(t) \leq \limsup_{n\to\infty} P(X_n \leq t - c + \epsilon) + \limsup_{n\to\infty} P(|Y_n - c| \geq \epsilon)$$

or, $\limsup_{n\to\infty} F_n(t) \leq F(t - c + \epsilon)$

$(\because Y_n \xrightarrow{P} c)$

Now $P(X_n+Y_n>t) \le P(X_n+Y_n>t, |Y_n-c|<\epsilon)$
$+ P(|Y_n-c| \ge \epsilon)$

$\le P(X_n+c+\epsilon>t) + P(|Y_n-c| \ge \epsilon)$

or, $1-P(X_n+Y_n \le t) \le 1-P(X_n+c+\epsilon \le t)$
$+ P(|Y_n-c| \ge \epsilon)$

or, $P(X_n \le t-c-\epsilon) \le P(X_n+Y_n \le t)$
$+ P(|Y_n-c| \ge \epsilon)$

Hence $F(t-c-\epsilon) \le \liminf_{n\to\infty} F_n(t)$

$[\because Y_n \xrightarrow{P} c]$ —— (2)

From (1) and (2) we get,

$F(t-c-\epsilon) \le \liminf_{n\to\infty} F_n(t) \le \limsup_{n\to\infty} F_n(t)$

$\le F(t-c+\epsilon)$

Since $t-c$ is ~~the~~ a continuity point of F and $\epsilon>0$ is arbitrary

$\therefore \lim_{n\to\infty} F_n(t) = \lim_{n\to\infty} F_{X_n+Y_n}(t) = F(t-c)$

(b) Let $\frac{t}{c}$, $\frac{t}{c-\epsilon}$, $\frac{t}{c+\epsilon}$ be continuity points of F. ~~and $\epsilon>0$~~ where $\epsilon>0$.

Then $F_{X_nY_n}(t) = P(X_nY_n \le t)$

$\le P(X_nY_n \le t, |Y_n-c| \le \epsilon)$

$+ P(|Y_n-c|>\epsilon)$

$\therefore \limsup_{n\to\infty} P(X_nY_n \le t) \le \limsup_{n\to\infty} P(X_n \le \frac{t}{c-\epsilon})$

$(\because Y_n \xrightarrow{P} c)$

or, $\limsup_{n\to\infty} P(X_n Y_n \le t) \le F\left(\frac{t}{c-\epsilon}\right)$ ——— (1)

Similarly, $P(X_n Y_n \le t) + P(|Y_n - c| \ge \epsilon)$

$\ge P\left(X_n \le \frac{t}{c+\epsilon}\right)$

or, $\liminf_{n\to\infty} P(X_n Y_n \le t) \cancel{+0}$

$\ge \liminf_{n\to\infty} P\left(X_n \le \frac{t}{c+\epsilon}\right)$

$\left(\because Y_n \xrightarrow{P} c \Rightarrow P(|Y_n - c| \ge \epsilon) \to 0 \text{ as } n\to\infty\right)$

$\therefore \liminf_{n\to\infty} P(X_n Y_n \le t) \ge F\left(\frac{t}{c+\epsilon}\right)$ ——— (2)

From (1) & (2) we get

$F\left(\frac{t}{c+\epsilon}\right) \le \liminf_{n\to\infty} P(X_n Y_n \le t) \le \limsup_{n\to\infty} P(X_n Y_n \le t)$

$\le F\left(\frac{t}{c-\epsilon}\right)$

Since $\frac{t}{c}$ is the continuity point of F and $\epsilon > 0$ is arbitrary,

$\therefore \lim_{n\to\infty} F_{X_n Y_n}(t) = F\left(\frac{t}{c}\right)$.

(c) Let ct, $(c+\epsilon)t$, $(c-\epsilon)t$ be all continuity points of F where $\epsilon > 0$ be given.

Then, $P\left(\frac{X_n}{Y_n} \le t\right) \le P\left(\frac{X_n}{Y_n} \le t, |Y_n - c| < \epsilon\right)$

$+ P(|Y_n - c| \ge \epsilon)$

$\le P(X_n \le (c+\epsilon)t) + P(|Y_n - c| \ge \epsilon)$

$\therefore \limsup_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right) \le \limsup_{n\to\infty} P(X_n \le (c+\epsilon)t)$

$\left[\because Y_n \xrightarrow{P} c\right]$

so, $\limsup_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right) \le F((c+\epsilon)t)$ ——— (1)

Similarly, $P(X_n \le (c-\epsilon)t) \le P\left(\frac{X_n}{Y_n} \le t\right) + P(|Y_n - c| > \epsilon)$

$\therefore \liminf_{n\to\infty} P(X_n \le (c-\epsilon)t) \le \liminf_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right)$

$\left[\because Y_n \xrightarrow{P} c\right]$

or, $F((c-\epsilon)t) \le \liminf_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right)$ ——— (2)

From (1) & (2),

$F((c-\epsilon)t) \le \liminf_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right) \le \limsup_{n\to\infty} P\left(\frac{X_n}{Y_n} \le t\right)$

$\le F((c+\epsilon)t)$

since ct is a continuity point of F and $\epsilon > 0$ is arbitrary

so, $\lim_{n\to\infty} F_{\frac{X_n}{Y_n}}(t) = F(ct)$.

———— o ————

Theorem (3): (Egoroff's theorem): Let $\{X_n : n \in \mathbb{N}\}$ and X be random variables on a probability space $(\Omega, \mathcal{B}, P)$. If the sequence of random variables $X_n \xrightarrow{a.e} X$, then for every $\epsilon > 0$ there exists $E \in \mathcal{B}$ such that $P(E) < \epsilon$ and $X_n \longrightarrow X$ uniformly on E' (or $\Omega - E$).

Proof: since all the convergence are 'translation invariant', i.e. $X_n \to X \Leftrightarrow X_n - X \to 0$, We may assume without loss of generality that $X = 0$. Suppose $X_n \to 0$ a.e. Let, $F(m,n) = \{w : |X_k(w)| \geq \frac{1}{m}$ for some $k \geq n\}$ and for any m, $\{F(m,n) : n = 1, 2, \ldots\}$ is a decreasing sequence of sets and by continuity from above

$$\lim_{n\to\infty} P(F(m,n)) = P\left(\bigcap_{n=1}^{\infty} F(m,n)\right) = 0.$$

~~Therefore~~

Therefore we may find an integer N_m such that $P(F(m,n)) < \frac{\epsilon}{2^m}$ for all $n \geq N_m$

or, $P(F(m, N_m)) < \frac{\epsilon}{2^m}$.

Set $E = \bigcup_{m=1}^{\infty} F(m, N_m)$ and $P(E) < \epsilon . \sum_{m=1}^{\infty} \frac{1}{2^m} = \epsilon$.

and that if $w \notin E$, then

$|X_k(w)| < \frac{1}{m}$. $\forall k \geq N_m$

Thus the sequence $\{X_n\}$ converges uniformly on E'.

Theorem (4) (Skorokhod's Theorem) Let $X, X_1, X_2, \ldots$ be a sequence of random variables. The following are equivalent:

(1) $X_n \xrightarrow{d} X$

(2) there exists a probability space $(\Omega, \mathcal{B}, P)$ and random variables

$Y, Y_1, Y_2, \ldots$ such that $Y \stackrel{d}{=} X$, $Y_n \stackrel{d}{=} X_n$ $\forall n \in \mathbb{N}$ and $Y_n \xrightarrow{a.e.} Y$.

Proof: (2) $\Rightarrow$ (1) is obvious by Th. (1).

Conversely we shall assume (1). For $x \in \mathbb{R}$, let us write $F_n = F_{X_n}$ and $F = F_X$. Consider $\Omega = [0,1]$, $\mathcal{B} = \mathcal{B}_{[0,1]}$, and $P =$ Lebesgue measure on $[0,1]$. On this probability space define

$$Y_n(\omega) = \inf\{x \in \mathbb{R} : \omega \leq F_n(x)\}, \quad \forall n \in \mathbb{N}.$$

and

$$Y(\omega) = \inf\{x \in \mathbb{R} : \omega \leq F(x)\}.$$

Note that $\{\omega \in \Omega : Y(\omega) \leq y\} = \{\omega \in \Omega : \omega \leq F(y)\}$ and $\{\omega \in \Omega : Y_n(\omega) \leq y\} = \{\omega \in \Omega : \omega \leq F_n(y)\}$.

Thus $F_Y = F$ and $F_{Y_n} = F_n$ i.e. $Y \stackrel{d}{=} X$, $Y_n \stackrel{d}{=} X_n$ $\forall n \in \mathbb{N}$

Let $\omega \in \Omega$, let $\epsilon > 0$ be such that $a = Y(\omega) - \epsilon$ is a continuity point of F. So,

$$Y(\omega) > a \Rightarrow F(a) < \omega$$

$$\Rightarrow \exists\, m, \text{ such that } F_n(a) < \omega \ \forall n \geq m$$

$$\left[\because F_n(a) \to F(a)\right]$$

$$\Rightarrow \exists\, m, \text{ such that } Y_n(\omega) > a \ \forall n \geq m$$

$$\therefore \liminf_{n\to\infty} Y_n(\omega) \geq a = Y(\omega) - \epsilon.$$

The discontinuity points of F being countable, we have $\liminf_{n\to\infty} Y_n(\omega) \geq Y(\omega)$, $\forall \omega \in \Omega$. —— (1)

Let $\omega_0 \in \Omega$ be such that $\omega < \omega_0$. Let $\delta > 0$ be such that $b = Y(\omega_0) + \delta$ is a continuity point of F. So

$$Y(\omega_0) < b \Rightarrow F(b) \geq \omega_0$$

$\Rightarrow \exists\, m_2$ such that $F_n(b) \geq w_0 - \delta \quad \forall n \geq m_2$

$[\because F_n(b) \to F(b)]$

$\Rightarrow \exists\, m_2$ such that $Y_n(w_0 - \delta) \leq b \quad \forall n \geq m_2$

$\therefore \limsup_{n\to\infty} Y_n(w_0 - \delta) \leq b = Y(w_0) + \delta$

As the discontinuity points of F are countable, we first choose $\delta < w_0 - w$ to obtain

$\limsup_{n\to\infty} Y_n(w) \leq Y(w_0) + \delta$ ($\because Y_n$ is an increasing function)

We next choose δ_n such that $\delta_n \to 0$ to obtain $\limsup_{n\to\infty} Y_n(w) \leq Y(w_0)$ ——— (2)

If w is a continuity point of Y, then allowing w_0 to decrease to w yields

$\limsup_{n\to\infty} Y_n(w) \leq Y(w)$ ——— (3)

As the discontinuity points of Y are countable, we conclude from (1) & (3) that

$\lim_{n\to\infty} Y_n(w) = Y(w)$ a.e.

Central Limit Theorem

Theorem 1 (continuity theorem): Let the random variables $X, \{X_n\}_{n\geq 1}$ have characteristic function $\phi_X, \{\phi_{X_n}\}_{n\geq 1}$. The following are equivalent:

1. $X_n \xrightarrow{d} X$

2. $E(g(X_n)) \to E(g(X))$ for all bounded Lipschitz continuous functions,

3. $\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t)$, for all $t \in \mathbb{R}$.

Proof:

- (1) ⇒ (2) : Apply dominated convergence theorem.

- (2) ⇒ (3) : is obvious by definition of the characteristic function.

- (3) ⇒ (2) : Define $Y_n^{(k)} = X_n + \frac{1}{k} Z$ and $Y^{(k)} = X + \frac{1}{k} Z$

where Z is an $N(0,1)$ random variable which is independent of all X_n and X. Let $\phi_{n,k}(\cdot)$ be the characteristic function of $Y_n^{(k)}$ and $\phi_k(\cdot)$ be the characteristic function of $Y^{(k)}$. Let $f_{Y_n^{(k)}}$ and $f_{Y^{(k)}}$ be the density function of $Y_n^{(k)}$ and $Y^{(k)}$ respectively. Independence of Z and X_n implies that

$$\phi_{n,k}(t) = e^{-\frac{t^2}{2k^2}} \phi_{X_n}(t) \qquad \left(\because \phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) \right.$$

if X & Y are independent

(13)

and similarly we have,

$$\phi_K(t) = e^{-\frac{t^2}{2K^2}} \phi_X(t).$$

Now, $\left| f_{Y_n^{(K)}}(a) - f_{Y^{(K)}}(a) \right|$

$$= \left| \int_{\mathbb{R}} (\phi_{n,K}(t) - \phi_K(t)) e^{-iat}\, dt \right|$$

[by Inversion formula]

$$\leq \int_{-\infty}^{\infty} |\phi_{n,K}(t) - \phi_K(t)|\, dt$$

$$= \int_{-\infty}^{\infty} e^{-\frac{t^2}{2K^2}} |\phi_{X_n}(t) - \phi_X(t)|\, dt$$

$$\longrightarrow 0$$

[$\because \lim_{n\to\infty} \phi_{X_n} = \phi_X$] and dominated convergence th.

So, we conclude ~~that~~ $Y_n^{(K)} \xrightarrow{d} Y^{(K)}$.

Let g be a bounded Lipschitz continuous function on $\mathbb{R}$.

$$|E(g(X_n)) - E(g(X))| \leq |E(g(X_n)) - E(g(Y_n^{(K)}))|$$
$$+ |E(g(Y_n^{(K)})) - E(g(Y^{(K)}))| +$$
$$|E(g(Y^{(K)})) - E(g(X))|$$

$$\leq c\left(E|X_n - Y_n^{(K)}| + E|X - Y^{(K)}|\right)$$
$$+ |E(g(Y_n^{(K)})) - E(g(Y^{(K)}))|$$

$$= \frac{2c\,E|Z|}{K} + |E(g(Y_n^{(K)})) - E(g(Y^{(K)}))|$$

since $Y_n^{(k)} \xrightarrow{d} Y^{(k)}$ and the fact (1) $\Rightarrow$ (2)

we have $E(g(X_n)) \longrightarrow E(g(X))$. Hence (3) $\Rightarrow$ (2).

- (2) $\Rightarrow$ (1) : Let $x \in \mathbb{R}$ and $k \in \mathbb{N}$. Define

$$f_k(y) = \begin{cases} 1 & y \leq x \\ k(x-y)+1 & x \leq y \leq x+\frac{1}{k} \\ 0 & y \geq x+\frac{1}{k} \end{cases}$$

and

$$g_k(y) = \begin{cases} 1 & y \leq x-\frac{1}{k} \\ k(x-y) & x-\frac{1}{k} \leq y \leq x \\ 0 & y \geq x \end{cases}$$

Then $g_k \leq 1_{(-\infty, x]} \leq f_k$.

Now

$$\limsup_{n\to\infty} F_{X_n}(x) = \limsup_{n\to\infty} E\, 1_{(-\infty,x]}(X_n)$$
$$\leq \limsup_{n\to\infty} E f_k(X_n)$$
$$= E f_k(X), \quad \text{——— (1)}$$

($\because$ f_k is a bounded Lipschitz function)

Similarly,

$$\liminf_{n\to\infty} F_{X_n}(x) = \liminf_{n\to\infty} E\, 1_{(-\infty,x]}(X_n)$$
$$\geq \liminf_{n\to\infty} E g_k(X_n)$$
$$= E g_k(X). \quad \text{——— (2)}$$

Observe that $f_k \to 1_{(-\infty,x]}$ and $g_k \to 1_{(-\infty,x]}$ and they are uniformly bounded by 1.

Letting $k \to \infty$ in (i) and (ii), by dominated convergence theorem, we have
$$\limsup_{n\to\infty} F_{X_n}(x) \le P(X \le x) \text{ and}$$
$\liminf_{n\to\infty} F_{X_n}(x) \ge P(X<x)$. Hence at all continuity points x of F we have shown that $F_{X_n}(x) \longrightarrow F_X(x)$.

Theorem 2 (Central limit theorem): Let $X, X_1, X_2, \dots$ be independent and identically distributed (i.i.d) random variables. Assume that $\mu = E(X) < \infty$ and $\sigma^2 = E(X-\mu)^2 < \infty$. Define $S_n = \sum_{i=1}^{n} X_i$. Then
$$\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \xrightarrow{d} Z,$$
where Z is a standard normal random variable.

Proof: Let $Y_n = \dfrac{S_n - n\mu}{\sqrt{n}\,\sigma}$ and $Z \stackrel{d}{=} N(0,1)$.

Then by the above th. 1 we have to show that $\lim_{n\to\infty} \phi_{Y_n}(t) = \phi_Z(t)$, $t \in \mathbb{R}$.

Now,
$$(\phi_{Y_n}(t)) = E(\exp(itY_n)) = E\left(\exp \frac{it}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\mu\right)\right)$$
$$= \left[E\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(X-\mu)\right)\right)\right]^n$$
$(\because X, X_1, X_2, \dots$ are i.i.d$)$

and $\phi_Z(t) = e^{-\frac{t^2}{2}}$.

since $\left| e^{it} - \sum_{k=0}^{n} \frac{(it)^k}{k!} \right| \leq \min\left(\frac{|t|^{n+1}}{(n+1)!}, \frac{2|t|^n}{n!} \right)$

$$\therefore \left| \exp\left(i\frac{t}{\sigma\sqrt{n}}(X-\mu)\right) - 1 - \frac{it}{\sigma\sqrt{n}}(X-\mu) + \frac{t^2(X-\mu)^2}{2\sigma^2 n} \right|$$

$$\leq \min\left(\frac{|t(X-\mu)|^3}{6\sigma^3 n^{3/2}}, \frac{t^2(X-\mu)^2}{\sigma^2 n} \right)$$

so, $\left| E\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(X-\mu)\right) - 1 - \frac{it}{\sigma\sqrt{n}}(X-\mu) + \frac{t^2(X-\mu)^2}{2\sigma^2 n}\right) \right|$

$$\leq E\min\left(\frac{|t(X-\mu)|^3}{6\sigma^3 n^{3/2}}, \frac{t^2(X-\mu)^2}{\sigma^2 n} \right)$$

$$= \frac{t^2}{2\sigma^2 n} g\left(\frac{t}{\sqrt{n}}\right) \text{ ——— } (1)$$

~~with $g(a) = E(\min(\ldots$~~

with $g(a) = E\left(\min\left(\frac{|a(X-\mu)^3|}{3\sigma}, 2(X-\mu)^2\right)\right)$.

Let $\epsilon > 0$ be given. Then the finite variance hypothesis of the theorem and dominated convergence theorem imply the existence of an n_0 large enough so that

$$n \geq n_0 \Rightarrow \left| g\left(\frac{t}{\sqrt{n}}\right) \right| < \sigma^2 \epsilon \text{ ——— } (2)$$

$\therefore$ From (1) & (2)

$$\left| E\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(X-\mu)\right)\right) - 1 - \frac{it}{\sigma\sqrt{n}} E(X-\mu) + \frac{t^2}{2\sigma^2 n} E(X-\mu)^2 \right|$$

$$\leq \frac{t^2}{2n}\epsilon$$

or, $\left| E\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(X-\mu)\right)\right) - 1 - 0 + \frac{t^2}{2n} \right| \leq \frac{t^2}{2n}\epsilon$

$[\because \mu = E(X) \ \& \ \sigma^2 = E(X-\mu)^2]$

$$\therefore \; 1-\frac{t^2}{2n}(1+\epsilon) \leq E\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(X-\mu)\right)\right) \leq 1-\frac{t^2}{2n}(1-\epsilon)$$

or, $$\left(1-\frac{t^2}{2n}(1+\epsilon)\right)^n \leq \phi_{Y_n}(t) \leq \left(1-\frac{t^2}{2n}(1-\epsilon)\right)^n$$

or, $$e^{-\frac{t^2}{2}(1+\epsilon)} \leq \liminf_{n\to\infty} \phi_{Y_n}(t) \leq \limsup_{n\to\infty} \phi_{Y_n}(t) \leq e^{-\frac{t^2}{2}(1-\epsilon)}.$$

Hence $$\lim_{n\to\infty} \phi_{Y_n}(t) = e^{-\frac{t^2}{2}} = \phi_Z(t).$$