Función característica

Sean X e Y v.a. en $(\Omega, \mathcal{A}, \mathcal{P})$, $Z = X + i\,Y$ es una variable aleatoria compleja, es decir, $Z : \Omega \to C$

$$Z(w) = X(w) + i\,Y(w) \qquad \forall w \in \Omega$$

La esperanza de Z se define:

$$E(Z) = E(X) + i\,E(Y)$$

si $E(X) < \infty$, $E(Y) < \infty$.

Consideremos la función $e^{i\,x} = cos(x) + isen(x)$, la variable aleatoria compleja

$$e^{i\,X} = cos(X) + i\,sen(X)$$

tiene esperanza finita cualquiera sea X pues $sen(x)$ y $cos(x)$ están acotadas.

Definición: Sea X una v.a. con distribución F, la función característica de X es la función

$$\varphi_X : R \to C$$

definida como

$$\varphi_X(t) = E(e^{i\,t\,X}) = E(cos(t\,X)) + i\,E(sen(t\,X))$$

Propiedades: 1) $|\varphi_X(t)| \leq 1 \qquad \forall t \in R$

2) $\varphi_X(0) = 1$

3) $\overline{\varphi_X(t)} = \varphi_X(-t)$

4) φ_X es uniformemente continua en R

5) Si X e Y son v.a. independientes,

$$\varphi_{X+Y}(t) = \varphi_X(t)\,\varphi_Y(t) \qquad \forall t \in R$$

6) Si $X_1, \ldots, X_n$ son v.a. independientes,

$$\varphi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n \varphi_{X_i}(t) \qquad \forall t \in R$$

7) X tiene distribución simétrica alrededor de 0 si y sólo si $\varphi_X(t)$ es real $\forall t$

8) Si $Y = a\,X + b$,

$$\varphi_Y(t) = e^{i\,t\,b}\,\varphi_X(a\,t)$$

9) Si $E(|X|^n) < \infty$, entonces φ_X tiene n derivadas continuas ($\varphi_X \in \mathcal{C}^n$) y

$$\varphi_X^{(k)}(t) = i^k E(X^k\, e^{i\,t\,X}) \qquad \forall \quad k = 1, \ldots, n$$

En particular, $\varphi_X^{(k)}(0) = i^k\, E(X^k)$, o sea que la función característica actúa como función generadora de momentos.

Ejemplos: **1)** $X \sim \mathcal{P}(\lambda)$

$$\varphi_X(t) = E(e^{i\,t\,X}) = \sum_{n=0}^{\infty} \frac{e^{itn}\, e^{-\lambda} \lambda^n}{n!} =$$

$$= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda\, e^{i\,t})^n}{n!} = e^{-\lambda}\, e^{\lambda\, e^{it}} = e^{\lambda\,(e^{it}-1)}$$

Entonces, $i\,E(X) = \varphi_X'(0)$

pero, $\varphi_X'(t) = e^{\lambda(e^{it}-1)}\, \lambda\, e^{it}\, i$

Entonces, $\varphi_X'(0) = \lambda\, i$ y por lo tanto,

$$i\,E(X) = \lambda\, i \Rightarrow E(X) = \lambda$$

Por otra parte, $i^2\, E(X^2) = \varphi''_X(0)$

pero,

$$\varphi''_X(t) = e^{\lambda\,(e^{it}-1)}\lambda e^{it}\, i\,\lambda\, i\, e^{it} + e^{\lambda\,(e^{it}-1)}\lambda\, i^2\, e^{it} =$$

$$= e^{\lambda\,(e^{it}-1)}\, e^{it}\, i^2\,\lambda\,(\lambda\, e^{it}+1)$$

Entonces,$\varphi''_X(0) = i^2\,\lambda\,(\lambda+1)$

$$i^2\, E(X^2) = i^2\,\lambda\,(\lambda+1) \Rightarrow E(X^2) = \lambda^2+\lambda$$

y $V(X) = \lambda^2+\lambda-\lambda^2 = \lambda.$

Del mismo modo, podemos obtener los restantes momentos de X.

2) $X \sim N((0,1)$

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{i\,t\,x}\frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\,t\,x}\, e^{-x^2/2}\, dx$$

Pero,

$$e^{itx-x^2/2} = e^{itx-x^2/2+t^2/2}\, e^{-t^2/2} = e^{(t+ix)^2/2}\, e^{-t^2/2}$$

Entonces,

$$\varphi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(t+ix)^2/2}\, e^{-t^2/2}\, dx = e^{-t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(t+ix)^2/2}\, dx$$

Probaremos que

$$h(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(t+ix)^2/2}\, dx = 1$$

$$h'(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (t + i\, x) e^{(t+ix)^2/2}\, dx =$$

$$= lim_{k\to\infty} \int_{-k}^{k} \frac{1}{\sqrt{2\pi}} (t + i\, x) e^{(t+ix)^2/2}\, dx =$$

$$= lim_{k\to\infty} \frac{1}{\sqrt{2\pi}} (-i) e^{(t+ix)^2/2} |_{-k}^{k} =$$

$$= lim_{k\to\infty} \frac{(-i)}{\sqrt{2\pi}} \left[e^{(t+ik)^2/2} - e^{(t-ik)^2/2} \right] =$$

$$= lim_{k\to\infty} \frac{(-i)}{\sqrt{2\pi}} e^{(t^2-k^2)/2} \left[e^{tik} - e^{-tik} \right]$$

Pero,

$$|e^{tik} - e^{-tik}| \leq 2 \qquad \text{y} \qquad lim_{k\to\infty} e^{(t^2-k^2)/2} = 0$$

Entonces, $h'(t) = 0 \Rightarrow h(t)$ es constante. Pero

$$h(0) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx = 1$$

Entonces, $h(t) = 1 \quad \forall t$ y

$$\varphi_X(t) = e^{-t^2/2}$$

Si $Y \sim N(\mu, \sigma^2)$, entonces, $Y = \sigma X + \mu$ con $X \sim N(0,1)$, entonces

$$\varphi_Y(t) = \varphi_{\sigma X+\mu}(t) = e^{it\mu}\, \varphi_X(\sigma t) = e^{it\mu}\, e^{-\sigma^2 t^2/2}$$

Fórmula de Inversión: Sea X una v.a. con distribución F y función característica φ, y sean a y b dos puntos de continuidad de F, $a < b$, entonces

$$F(b) - F(a) = \frac{1}{2\pi} \lim_{k\to\infty} \int_{-k}^{k} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

Corolario:

$$F_X = F_Y \Longleftrightarrow \varphi_X = \varphi_Y \quad (\varphi_X(t) = \varphi_Y(t) \quad \forall t \in R)$$

Ejemplo: Sea $\varphi(t) = cos(at) \quad (a > 0)$
Veremos que φ es una función característica, hallando la v.a. X correspondiente.

Como $\varphi(t)$ es real, X es simétrica. Además,

$$\varphi(t) = cos(at) = E(cos(tX))$$

pues la parte imaginaria es nula. Como $cos(at) = cos(-at)$, X es la v.a. discreta tal que

$$P(X = a) = P(X = -a) = 1/2.$$

Convergencia en distribución

Definición: Sean X y $X_1, X_2, \ldots$ v.a. con funciones de distribución F y $F_1, F_2, \ldots$ respectivamente. Diremos que X_n converge en distribución a X cuando $n \to \infty$ ($X_n \xrightarrow{\mathcal{D}} X$) si

$$F_n(x) \to F(x) \qquad \forall x \quad \text{punto de continuidad de } F$$

Nota: Se dice también que F_{X_n} converge débilmente a $F_X (F_{X_n} \xrightarrow{w} F_X)$.

Porqué se pide convergencia sólo en los puntos de continuidad?

Ejemplo: Sea $(\Omega, \mathcal{A}, \mathcal{P})$ un espacio de probabilidad y $\forall n \in N$ sea

$$X_n = \frac{1}{n} \qquad \text{o sea} \quad X_n(w) = \frac{1}{n} \qquad \forall w \in \Omega$$

y sea $X = 0$. Intuitivamente, X_n debería converger a X según cualquier criterio de convergencia. De hecho $X_n \xrightarrow{pp} X$ y por lo tanto $X_n \xrightarrow{p} X$. Pero

$$F_{X_n}(x) = \begin{cases} 1 & \text{si} \quad x \geq 1/n \\ \\ 0 & \text{si} \quad x < 1/n \end{cases} \qquad F_X(x) = \begin{cases} 1 & \text{si} \quad x \geq 0 \\ \\ 0 & \text{si} \quad x < 0 \end{cases}$$

Por lo tanto, si $x > 0$, $F_{X_n}(x) \longrightarrow 1 = F_X(x)$

si $x < 0$, $F_{X_n}(x) \longrightarrow 0 = F_X(x)$

pero si $x = 0$, $F_{X_n}(x) \longrightarrow 0 \neq 1 = F_X(x)$

Teorema: 1) si $X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X$

2) si $X_n \xrightarrow{\mathcal{D}} X \quad \text{y} \quad P(X = c) = 1 \Rightarrow X_n \xrightarrow{p} X$

$$\boxed{X_n \xrightarrow{pp} X \Rightarrow X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X}$$

El objetivo de los Teoremas que se enuncian a continuación es probar que

$$X_n \xrightarrow{\mathcal{D}} X \Longleftrightarrow \varphi_{X_n}(t) \to \varphi_X(t) \quad \forall t \in R$$

Teorema de Helly–Bray: Sean F y $F_1, F_2, \ldots$ funciones de distribución. Si F_n converge débilmente a F, entonces

$$\lim_{n\to\infty} \int g(x)dF_n(x) = \int g(x)dF(x)$$

para toda función g continua y acotada ($g : R \to R$).

Observaciones: 1) Si $X_n \xrightarrow{\mathcal{D}} X$, por Helly–Bray, $E(g(X_n))$ converge a $E(g(X))$, para toda g continua y acotada. En particular, como $sen(x)$ y $cos(x)$ son continuas y acotadas, resulta que

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t) \qquad \forall t \in R$$

2) La recíproca del Teorema es válida, o sea si $\lim_{n\to\infty} \int g(x)dF_n(x) = \int g(x)dF(x)$ para toda g continua y acotada, entonces F_n converge débilmente a F y, de hecho, a veces se toma como definición de convergencia débil.

Teorema de continuidad de Paul–Levy: Sean $F_1, F_2, \ldots$ funciones de distribución y $\varphi_1, \varphi_2, \ldots$ sus correspondientes funciones características. Si φ_n converge puntualmente a una función φ y si φ es continua en 0, entonces
a) existe una función de distribución F tal que $F_n \xrightarrow{w} F$
b) φ es la función característica de F

Notas: 1) Estos dos teoremas implican que

$$X_n \xrightarrow{\mathcal{D}} X \Longleftrightarrow \varphi_{X_n} \to \varphi_X$$

2) El Teorema de Paul–Levy es más fuerte que la "suficiencia" pues dice que el límite de una sucesión de funciones características es una función característica si es continua en 0.

Corolario: a) Sean $X_1, X_2, \ldots$ v.a. Si $\varphi_{X_n}(t) \to \varphi_X(t)$ $\forall t \in R$ y si φ es continua en 0, entonces φ es función característica de alguna v.a. X, o sea $\varphi = \varphi_X$ y $X_n \xrightarrow{\mathcal{D}} X$.
b) Si $\varphi_{X_n}(t) \to e^{-t^2/2}$ $\forall t$, entonces $X_n \xrightarrow{\mathcal{D}} Z$ con $Z \sim N(0,1)$, es decir $F_n \xrightarrow{w} \phi$.

Teorema Central del Límite

Teorema Central del Límite para v.a. i.i.d: Sean $X_1, X_2, \ldots$ v.a. i.i.d. con $E(X_i) = \mu$, $V(X_i) = \sigma^2 < \infty$, y sea $S_n = X_1 + X_2 + \ldots + X_n$, entonces

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n\,\mu}{\sigma\,\sqrt{n}} \xrightarrow{\mathcal{D}} Z$$

con $Z \sim N(0,1)$.

Demostración: Supondremos $\mu = 0$. Por el Teorema de Paul–Levy, basta probar que

$$\varphi_{S_n/\sigma\sqrt{n}}(t) \to e^{-t^2/2} \qquad \forall t \in R$$

$$\varphi_{S_n/\sigma\sqrt{n}}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right) = \prod_{i=1}^{n} \varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left[\varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

Como $E(X_i^2) < \infty$, entonces φ_{X_i} tiene dos derivadas continuas y, aplicando la fórmula de Taylor

$$\varphi_{X_i}(t) = \varphi_{X_i}(0) + \varphi_{X_i}^{(1)}(0)\,t + \varphi_{X_i}^{(2)}(\theta)\,t^2/2$$

donde $|\theta| \le |t|$.

$$\varphi_{X_i}(t) = \varphi_{X_i}(0) + \varphi_{X_i}^{(1)}(0)\, t + \varphi_{X_i}^{(2)}(0)\, \frac{t^2}{2} + \frac{t^2}{2}\, [\varphi_{X_i}^{(2)}(\theta) - \varphi_{X_i}^{(2)}(0)]$$

donde $\varphi_{X_i}^{(2)}(\theta) - \varphi_{X_i}^{(2)}(0) \rightarrow_{t\rightarrow 0} 0$.

Pero,

$\varphi_{X_i}(0) = 1 \qquad \varphi_{X_i}^{(1)}(0) = i\mu = 0 \qquad \varphi_{X_i}^{(2)}(0) = i^2 E(X_i^2) = -\sigma^2$

Entonces

$$\varphi_{X_i}(t) = 1 - \frac{\sigma^2 t^2}{2} + \frac{t^2}{2}\epsilon(t)$$

donde $lim_{t\rightarrow 0}\epsilon(t) = 0$.

Para t fijo,

$$\left[\varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2}{2n} + \frac{t^2}{2\sigma^2 n}\epsilon(\frac{1}{\sigma\sqrt{n}})\right]^n =$$

$$= \left[1 - \frac{t^2}{2n}\left(1 - \frac{1}{\sigma^2}\,\epsilon(\frac{1}{\sigma\sqrt{n}})\right)\right]^n \rightarrow_{n\rightarrow\infty} e^{-t^2/2}$$

pues $1 - \frac{1}{\sigma^2}\epsilon\left(\frac{1}{\sigma\sqrt{n}}\right) \rightarrow_{n\rightarrow\infty} 1$.

Corolario (TCL de De Moivre y Laplace): Sea S_n el número de éxitos en n ensayos Bernoulli independientes con probabilidad p de éxito $(0 < p < 1)$, entonces

$$\frac{S_n - n\,p}{\sqrt{n\,p\,(1-p)}} \xrightarrow{\mathcal{D}} Z$$

con $Z \sim N(0,1)$.

Ejemplos: 1) Al sumar números, una calculadora aproxima cada número al entero más próximo. Los errores de aproximación se suponen independientes y con distribución $U(-0.5, 0.5)$.
a) Si se suman 1500 números, cuál es la probabilidad de que la magnitud del error total exceda 15?

$$P(|S_{1500}| > 15) = 1 - P(|S_{1500}| \leq 15) =$$
$$= 1 - P(-15 \leq S_{1500} \leq 15) =$$
$$= 1 - P\left(\frac{-15}{\sqrt{1500/12}} \leq \frac{S_{1500}}{\sqrt{1500/12}} \leq \frac{15}{\sqrt{1500/12}}\right)$$

pues $E(X_i) = 0$ y $V(X_i) = 1/12$. Entonces,

$$P(|S_{1500}| > 15) \overset{(a)}{=} 1 - \phi(15/11.18) + \phi(-15/11.18) =$$

$$1 - \phi(1.34) + \phi(-1.34) = 0.18$$

b) Cuántos números pueden sumarse a fin de que la magnitud del error total sea menor que 10, con probabilidad 0.90?

$$n \quad \text{tal que} \quad P(|S_n| \leq 10) \geq 0.90$$

$$P(-10 \leq S_n \leq 10) \geq 0.90$$

$$P\left(\frac{-10}{\sqrt{n/12}} \leq \frac{S_n}{\sqrt{n/12}} \leq \frac{10}{\sqrt{n/12}}\right) \geq 0.90$$

$$\phi\left(\frac{10}{\sqrt{n/12}}\right) - \phi\left(\frac{-10}{\sqrt{n/12}}\right) \geq 0.90$$

$$2\,\phi\left(\frac{10}{\sqrt{n/12}}\right) - 1 \geq 0.90$$

$$\frac{10}{\sqrt{n/12}} \geq 1.64$$

$$\sqrt{n} \leq 21.12 \qquad n \leq 446$$

2) Sea $X \sim Bi(60, 1/3)$. Calcular en forma aproximada la probabilidad de que X sea mayor o igual que 25.

$$P(X \geq 25) = P(X \geq 24.5) =$$

$$P\left(\frac{X-20}{\sqrt{13.33}} \geq \frac{24.5-20}{\sqrt{13.33}}\right) \stackrel{(a)}{=} 1-\phi(1.23) = 1-0.89 = 0.11$$

Proposición: Sean X y $X_1, X_2, \ldots$ v.a. y $g : R \rightarrow R$ una función continua, entonces

a) $X_n \stackrel{pp}{\rightarrow} X \quad \Rightarrow \quad g(X_n) \stackrel{pp}{\rightarrow} g(X)$

b) $X_n \stackrel{p}{\rightarrow} X \quad \Rightarrow \quad g(X_n) \stackrel{p}{\rightarrow} g(X)$

c) $X_n \stackrel{\mathcal{D}}{\rightarrow} X \quad \Rightarrow \quad g(X_n) \stackrel{\mathcal{D}}{\rightarrow} g(X)$

Ejemplos:

1) $X_n \stackrel{\mathcal{D}}{\rightarrow} Z \sim N(0,1) \quad \Rightarrow \quad X_n^2 \stackrel{\mathcal{D}}{\rightarrow} W \sim \chi_1^2$

2) $X_n \stackrel{\mathcal{D}}{\rightarrow} Z \sim N(0,1) \quad \Rightarrow \quad c\,X_n \stackrel{\mathcal{D}}{\rightarrow} Y \sim N(0,c^2)$

3) $X_n \stackrel{p}{\rightarrow} c \quad (c > 0) \quad \Rightarrow \quad ln(X_n) \stackrel{p}{\rightarrow} ln(c)$

Nota: Si g no está definida o no es continua en toda la recta ($g(x) = ln(x)$ o $g(x) = 1/x$), se puede probar que la proposición anterior es cierta si, para algún abierto A de R, g es continua en A y $P(X \in A) = 1$.

Teorema de Slutsky: Sean $X, X_1, X_2, \ldots$ e $Y_1, Y_2, \ldots$ v.a. tales que $X_n \xrightarrow{\mathcal{D}} X$, $Y_n \xrightarrow{p} c$ (c constante), entonces:

a) $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$

b) $X_n - Y_n \xrightarrow{\mathcal{D}} X - c$

c) $X_n Y_n \xrightarrow{\mathcal{D}} c\,X$

d) Si $c \neq 0$ y $P(Y_n \neq 0) = 1$, entonces

$$\frac{X_n}{Y_n} \xrightarrow{\mathcal{D}} \frac{X}{c}$$

Proposición: Sean $Y_1, Y_2, \ldots$ v.a. tales que

$$\sqrt{n}(Y_n - \mu) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$$

Si $g(y)$ es una función derivable en μ, entonces

$$\sqrt{n}(g(Y_n) - g(\mu)) \xrightarrow{\mathcal{D}} N(0, \sigma^2\,(g'(\mu))^2)$$

Nota: Vale si $g'(\mu) = 0$, si se piensa que $N(0,0)$ equivale a la masa puntual en 0.

Ejemplo: Si $X_1, X_2, \ldots$ son v.a. i.i.d. con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 < \infty$ entonces, por el TCL:

$$\frac{S_n - n\,\mu}{\sqrt{n}\,\sigma} \xrightarrow{\mathcal{D}} N(0,1)$$

$$\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right)\xrightarrow{\mathcal{D}} N(0,1)$$

$$\sqrt{n}\ (\bar{X}_n-\mu)\xrightarrow{\mathcal{D}} N(0,\sigma^2)$$

Sea $g(x)=x^2$, entonces $g'(\mu)=2\,\mu$ y

$$\sqrt{n}\ \left(\bar{X}_n^2-\mu^2\right)\xrightarrow{\mathcal{D}} N(0,4\,\sigma^2\mu^2)$$

Teorema Central del Límite de Lindeberg: Sean $X_1, X_2, \ldots$ v.a. independientes tales que $E(X_n)=\mu_n$ y $V(X_n)=\sigma_n^2<\infty$ con por lo menos un $\sigma_n>0$. Sea $S_n=X_1+\ldots+X_n$, entonces para que

$$\frac{S_n-E(S_n)}{\sqrt{V(S_n)}}\xrightarrow{\mathcal{D}} Z\sim N(0,1)$$

es suficiente que se satisfaga la siguiente condición:

$$\forall\epsilon>0 \qquad lim_{n\to\infty}\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon\, s_n}(x-\mu_k)^2 dF_k(x)=0$$

donde $F_k=F_{X_k}$ y $s_n=\sqrt{V(S_n)}=\sqrt{\sigma_1^2+...+\sigma_n^2}$.

Nota: La condición de Lindeberg significa que los términos

$$\frac{X_k-\mu_k}{s_n} \qquad \text{de la suma} \qquad \frac{S_n-E(S_n)}{s_n}$$

son uniformemente pequeños para n grande.
Por ejemplo, la condición de Lindeberg implica que

$$max_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2} \rightarrow 0$$

cuando $n \rightarrow \infty$, o sea que las varianzas de cada sumando X_k son uniformemente pequeñas cuando n es grande y, por lo tanto, ningún sumando tiene demasiado peso en la suma $\frac{S_n - E(S_n)}{s_n}$.