

Funciones características y sus aplicaciones

Abraham Rojas Vega
Introducción a los Procesos Estocásticos
FC-UNI

12-11-2016

Definición y propiedades

Dada una variable aleatoria X con valores en \mathbb{R}^d , su **función característica** $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$ se define por:

$$\begin{aligned}\varphi_X(t) &= \mathbf{E}[e^{i\langle t, X \rangle}] \\ &= \int_{\mathbb{R}^d} e^{i\langle t, X \rangle} \mathbf{P}_X(dx) = \int_{\mathbb{R}^d} e^{i\langle t, X \rangle} dF_X \\ &= \int_{\mathbb{R}^d} e^{i\langle t, X \rangle} f_X(x) dx\end{aligned}$$

donde $\mathbf{P}_X = \mathbf{P} \circ X^{-1}$ (medida de probabilidad inducida por X)

La última igualdad es válida cuando X admite función de densidad f_X .

También se puede definir la función característica de cualquier función de distribución, usando la tercera igualdad

Propiedades

Teorema:

Sean X, Y variables aleatorias que toman valores en \mathbb{R}^d . Se cumple:

- $|\varphi_X| \leq 1$ y $\varphi_X(0) = 1$.
- $\varphi_X(-t) = \overline{\varphi_X(t)}$ (φ es hermitiana)
- $\varphi_{aX+b} = \varphi_X(at)e^{i\langle b, t \rangle}$
- Si X e Y son independientes entonces $\varphi_{X+Y} = \varphi_X \varphi_Y$
- $0 \leq 1 - \Re[\varphi_X(2t)] \leq 4(1 - \Re[\varphi_X(t)])$ para todo $t \in \mathbb{R}^d$

Demostración (para no perder la costumbre).

El primer y segundo ítem son consecuencia directa de la definición.

En el tercer ítem, observamos que $e^{i\langle t, X \rangle}$ y $e^{i\langle t, Y \rangle}$ son independientes, se tiene: $\varphi_{X+Y}(t) = \mathbf{E}[e^{i\langle t, X \rangle} \cdot e^{i\langle t, Y \rangle}] = \mathbf{E}[e^{i\langle t, X \rangle}] \cdot \mathbf{E}[e^{i\langle t, Y \rangle}] = \varphi_X \varphi_Y$

Para el cuarto ítem, basta ver que:

$$1 - \cos(\langle 2t, X \rangle) = 2(1 - \cos(\langle t, X \rangle)) \leq 4(1 - \cos(\langle t, X \rangle))$$



Notación con multi-índices

Consideremos $\alpha = (\alpha_1, \dots, \alpha_d), \beta = (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$ (denominados **multi-índices**), pongamos $|\beta| = \beta_1 + \dots + \beta_d$.

Dado $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ y $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$, definimos:

$$x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}$$

$$D^\beta \phi = \frac{\partial^{|\beta|} \phi}{\partial x_1^{\beta_1} \dots \partial x_d^{\beta_d}}$$

Espacio de Schwartz ($\mathcal{S}(\mathbb{R}^d)$)

Es el conjunto de funciones $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ infinitamente derivables tales que

$$| \sup_{x \in \mathbb{R}^d} x^\alpha D^\beta \phi(x) | < \infty$$

para cualquier $\alpha, \beta \in (\mathbb{N}_0)^d$.

Puede probarse que la transformada de Fourier se torna en un homeomorfismo lineal del espacio de Schwartz sobre sí mismo. Hay que tener en cuenta que la transformada de Fourier no siempre está definida en $L^1(\mathbb{R}^d)$

Transformadas de Fourier

Sea $\phi \in L^1(\mathbb{R}^d)$. Se define la:

- **transformada de Fourier** mediante:

$$\mathcal{F}[\phi](y) = \int_{\mathbb{R}^d} \phi(x) e^{-i\langle x, y \rangle} dx$$

- **transformada inversa de Fourier** mediante:

$$\mathcal{F}^{-1}[\phi](y) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \phi(x) e^{i\langle x, y \rangle} dx$$

Observamos que, salvo constantes, la función característica coincide con la transformada inversa de Fourier de f_X (función de densidad).

Algunas propiedades

Teorema:

Sean $\phi \in \mathcal{S}(\mathbb{R}^d)$, $a, \alpha \in \mathbb{R}^d$, $\alpha \neq 0$; se cumple:

$$\mathcal{F}[\phi \circ (I - a)] = e^{-i\langle a, y \rangle} \mathcal{F}[\phi] \quad (1)$$

$$\mathcal{F}[\phi \circ (\lambda I)] = \frac{1}{|\lambda|^d} \mathcal{F}[\phi] \left(\frac{y}{\lambda} \right) \quad (2)$$

Donde I es la función identidad

$$\int_{\mathbb{R}^d} \mathcal{F}[\phi](y) \psi(y) dy = \int_{\mathbb{R}^d} \mathcal{F}[\phi](y) \psi(y) dy \quad (3)$$

Fórmula de inversión de Fourier

Teorema:

Si $\phi \in \mathcal{S}(\mathbb{R}^d)$ entonces, para todo $x \in \mathbb{R}^d$:

$$\phi(x) = \mathcal{F}^{-1}[\mathcal{F}[\phi]]$$

De hecho, $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ es un isomorfismo lineal.

Teorema: (Identidad de Parseval)

Dados $g, h \in \mathcal{S}(\mathbb{R}^d)$, se cumple

$$\langle g, h \rangle = \frac{1}{(2\pi)^d} \langle \mathcal{F}[g], \mathcal{F}[h] \rangle$$

donde se ha considerado el producto interno de $L^2(\mathbb{R}^d)$

Veremos que si la f.c es integrable entonces el pdf existe y es la transformada inversa de la f.c.

Fórmula de inversión para variables aleatorias (reales)

Teorema:

Sea φ la f.c de la v.a. X con función de distribución F . Si $a < b$:

$$\begin{aligned} & \lim_{T \uparrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-i\theta a} - e^{i\theta b}}{i\theta} \varphi(\theta) d\theta \\ &= \frac{1}{2} \mathbf{P}_X(\{a\}) + \mathbf{P}_X(< a, b >) + \frac{1}{2} \mathbf{P}_X(\{b\}) \\ &= \frac{1}{2} [F(b) + F(b-)] - \frac{1}{2} [F(a) + F(a-)] \end{aligned}$$

Es más, si se aumenta la condición $\int_{-\infty}^{\infty} |\varphi(\theta)| < \infty$ entonces X tiene función de densidad continua, dada por:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\theta x} \varphi(\theta) d\theta$$

Teorema de Continuidad (de Lévy)

Definición:

Una sucesión (F_n) de funciones de distribución reales **converge débilmente** a la función de distribución real F cuando $\lim F_n(x) = F(x)$ en todo punto de continuidad x de F .

Teorema:

Sea (F_n) una sucesión de funciones de distribución reales y φ_n la función características de F_n .

Supongamos que (φ_n) converge puntualmente a cierta función φ , siendo esta continua en 0.

Entonces existe una función de distribución F tal que $F_n \longrightarrow F$ débilmente y $\varphi_F = \varphi$.

Demostración Teor. de Continuidad

Primero algunos preliminares:

Teorema: (Lema de Helly-Bray)

Sea (F_n) una sucesión de p.d.f. Entonces existe una función F continua por la derecha y no decreciente tal que $0 \leq F \leq 1$ y una subsucesión (n_i) tal que

$$\lim_{n \rightarrow \infty} F_{n_i} = F(x) \text{ cuando } F \text{ es continua en } x$$

Definición:

*Una sucesión de p.d.f's (F_n) es **magra** si, dado $\varepsilon > 0$, existe $K > 0$ tal que*

$$F(K) - F(-K-) > 1 - \varepsilon$$

donde F es la función que aparece en el lema anterior; la cual será una p.d.f. (podemos construir una v.a.)

AFIRMACIÓN. La sucesión (F_n) del enunciado de nuestro teorema es magra. Vamos a asumir eso.

Entonces por el lema de Helly-Bray, existe una subcesión (F_{n_k}) y una p.d.f. tal que $F_{n_k} \longrightarrow F$ débilmente.

Sea $\theta \in \mathbb{R}$ tenemos: $\varphi_{n_k}(\theta) \longrightarrow \varphi_F(\theta)$. Por tanto $\varphi = \varphi_F$

Ahora supongamos que (F_n) no converge débilmente a F . Entonces algún punto x de continuidad de F podemos encontrar la subsucesión (\hat{F}_n) y $\eta > 0$ tales que $|(\hat{F}_n(x) - F(x))| \geq \eta$ para todo n .

Como (\hat{F}_n) es magra entonces posee una subsucesión (\hat{F}_{n_j}) tal que existe una p.d.f \hat{F} cumpliéndose: $(\hat{F}_{n_j}) \longrightarrow \hat{F}$ débilmente.

Luego $\hat{\varphi}_{n_j} \longrightarrow \hat{\varphi}$ puntualmente. Por tanto $\hat{\varphi} = \varphi_{\hat{F}} = \varphi = \varphi_F$. Por los teoremas de inversión concluimos que $\hat{F} = F$, en particular:

$\hat{F}_{n_j}(x) \longrightarrow \hat{F}(x) = F(x)$.

□

Teorema del límite central

Teorema:

Sea (X_n) una secuencia de v.a. reales i.i.d. donde $\mathbf{E}[X_1] = 0$ y $\sigma^2 := \mathbf{V}[X_1] \in (0, \infty)$.

Si definimos $S_n := X_1 + \dots + X_n$ y

$$G_n := \frac{S_n}{\sigma\sqrt{n}}$$

Entonces, para todo $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbf{P}[G_n \leq x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$$

Demostración de Teor. del límite central

Usaremos los siguientes resultados técnicos:

Proposición:

Si φ es la f.c. de una v.a. con $\mathbf{E} = 0$ y $\mathbf{V} < \infty$ entonces:

$$\varphi(t) = 1 - \frac{1}{2}t^2\mathbf{V} + o(t^2) \text{ cuando } t \longrightarrow 0 \quad (4)$$

Si $z \in \mathbb{C}$ y $|z| \leq 0,5$ entonces

$$|\log(1+z) - z| \leq |z|^2 \quad (5)$$

Sea $t \in \mathbb{R}$. Entonces por 4:

$$\varphi_{G_n}(t) = \varphi_{S_n}\left(\frac{t}{\sqrt{n}}\right) = \varphi\left(\frac{t}{\sqrt{n}}\right) = \left[1 - \frac{1}{2}\frac{t^2}{n}\mathbf{V} + o\left(\frac{t^2}{n}\right)\right]^n \text{ cuando } n \longrightarrow \infty$$

Por 5 tenemos:

$$\log \varphi_{G_n}(t) = n \log \left[1 - \frac{1}{2} \frac{t^2}{n} \mathbf{V} + o\left(\frac{t^2}{n}\right) \right] = n \left[-\frac{1}{2} \frac{t^2}{n} \mathbf{V} + o\left(\frac{t^2}{n}\right) \right] \longrightarrow \frac{1}{2} t^2 \text{ cuando } n \longrightarrow \infty$$

Por tanto $\varphi_{G_n}(t) \longrightarrow \exp\left(-\frac{1}{2}t^2\right)$. Observamos $t \mapsto \exp\left(-\frac{1}{2}t^2\right)$ es la función característica de la distribución normal con media 0 y varianza \mathbf{V} .
Aplicando el teorema de continuidad a G_n ... ¡ya está!



Aplicación al cálculo de momentos

Teorema:

Sea φ la f.c de la v.a.r. X . Si $\mathbf{E}[|X^n|] < \infty$ entonces φ es n veces continuamente diferenciable, con derivadas:

$$\varphi^{(k)}(t) = \mathbf{E}[(iX)^k e^{itX}]$$

Otro interesante resultado:

Teorema:

Sea X una v.a.r. y φ su f.c. Si φ es $2n$ diferenciable en 0, se tiene:

$$\mathbf{E}[X^{2n}] = (-1)^n \varphi^{(2n)}(0) < \infty$$

Una aplicación de estos resultados es mostrar que la distribución de Cauchy no tiene esperanza, pues su función característica no es diferenciable en $t = 0$.

Fórmula de Stirling

La fórmula está dada por:

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1$$

Esta nos permite hacer la siguiente aproximación:

$$\ln n! = n \ln n - n + O(\ln n)$$

Es útil para cuando se necesita hacer aproximaciones con factoriales, como en Mecánica Estadística.

Una versión acotada de la fórmula es:

$$\sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} \leq n! \leq e n^{n+\frac{1}{2}} e^{-n}$$

Esta expresión es útil en aproximaciones que acarreen factoriales, pues las cotas son derivables.

Método de los momentos

Teorema:

Sea X una v.a.r. donde

$$\alpha := \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbf{E}[|X|^n]^{1/n} < \infty$$

Entonces la f.c. φ de X es analítica y la distribución de X está determinada por sus momentos $\mathbf{E}[X^n]$, siendo esta representación "única". En particular, esto se cumple cuando $\mathbf{E}[e^{t|X|}] < \infty$ para algún $t > 0$.

DEMOSTRACIÓN. Si $|h| < 1/3\alpha$, por la fórmula de Stirling :

$$\begin{aligned} \limsup \mathbf{E}[|X|^n] \cdot |h|^n \cdot /n! &= \limsup \sqrt{2\pi n} (\mathbf{E}[|X|^n]^{1/n} \cdot |h| \cdot e/n)^n \leq \\ \limsup \sqrt{2\pi n} (e/3)^n &= 0 \end{aligned}$$

Es decir, la f.c. puede expandirse como serie de Taylor en cualquier $t \in \mathbb{R}$ con radio de convergencia $1/3\alpha$. En particular, φ es analítica y determinada por los coeficientes de la serie, tomando $t = 0$, estos coeficientes se tornan en los momentos.

Más aplicaciones

- Se usa en la definición de la **segunda función característica**:

$$H(t) := \log \mathbf{E}[e^{itX}] \approx \sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!} \approx \mu it - \sigma^2 \frac{t^2}{2} + \dots$$

Los **cumulantes** de la v.a. X viene los coeficientes κ_n de la expansión de Taylor (cuando existen).

La ventaja de trabajar de cumulantes radica en que la cumulantes de dos v.a. es el cumulante de la suma de estas, cuando son independientes.

En Física Estadística, muchas cantidades que son proporcionales al tamaño del sistema están relacionadas con cumulantes de v.a.'s.

- Puede usarse en el proceso de ajustar una distribución de probabilidad a un conjunto de datos. Existe el concepto de **función característica empírica**, la cual se calcula a partir de los datos. Para más información:
<http://www.utstat.toronto.edu/andrey/pubs/EmpiricalCF.pdf>
- En la “ida” de la demostración del teorema de las tres series (de Kolmogorov):

Teorema:

Sean (X_n) v.a.i.i.d. Entonces $\sum X_n$ converge c.s. si y solo si para todo $k > 0$ se cumplen las tres siguientes condiciones:

- 1 $\sum \mathbf{P}[|X_n| > K] < \infty$
- 2 $\sum \mathbf{E}[X_n^K]$ converge
- 3 $\sum \mathbf{V}[X_n^k] < \infty$

donde

$$X_n^K(w) = \begin{cases} X_n(w) & \text{si } |X_n(w)| \leq K \\ 0 & \text{si } |X_n(w)| > K \end{cases}$$

- Estimar probabilidades de cola usando aproximaciones de punto silla.

Teorema:

Sea X una variable aleatoria y φ su función característica. Entonces, para $\varepsilon > 0$ se tiene:

$$\mathbf{P}[|X| > \frac{2}{\varepsilon}] \leq \frac{1}{\varepsilon} \int_{-\varepsilon}^{\varepsilon} (1 - \varphi(t)) dt$$

- Probar resultados como el siguiente:

Teorema:

Si X e Y son independientes y $X + Y$ tienen distribución normal entonces tanto X como Y tienen distribución normal.

Algunos ejemplos

Distribution	Characteristic function $\varphi(t)$
Degenerate δ_a	e^{ita}
Bernoulli $\text{Bern}(p)$	$1 - p + pe^{it}$
Binomial $B(n, p)$	$(1 - p + pe^{it})^n$
Negative binomial $\text{NB}(r, p)$	$\left(\frac{1 - p}{1 - pe^{it}}\right)^r$
Poisson $\text{Pois}(\lambda)$	$e^{\lambda(e^{it} - 1)}$
Uniform $U(a, b)$	$\frac{e^{itb} - e^{ita}}{it(b - a)}$
Laplace $L(\mu, b)$	$\frac{e^{it\mu}}{1 + b^2 t^2}$
Normal $N(\mu, \sigma^2)$	$e^{it\mu - \frac{1}{2}\sigma^2 t^2}$
Chi-squared χ_k^2	$(1 - 2it)^{-k/2}$
Cauchy $C(\mu, \theta)$	$e^{it\mu - \theta t }$
Gamma $\Gamma(k, \theta)$	$(1 - it\theta)^{-k}$
Exponential $\text{Exp}(\lambda)$	$(1 - it\lambda^{-1})^{-1}$
Geometric $\text{Gf}(p)$ (number of failures)	$\frac{p}{1 - e^{it}(1 - p)}$
Geometric $\text{Gt}(p)$ (number of trials)	$\frac{p}{e^{-it} - (1 - p)}$
Multivariate normal $N(\mu, \Sigma)$	$e^{it^T \mu - \frac{1}{2} t^T \Sigma t}$
Multivariate Cauchy $\text{MultiCauchy}(\mu, \Sigma)$ [10]	$e^{it^T \mu - \sqrt{t^T \Sigma} t}$

Bibliografía

- Achim Klenke. Probability Theory. Second Edition. Universitext. Springer.
- David Williams. Probability with Martingales. Cambridge Mathematical Textbooks. Cambridge University Press.