Capítulo 3

El Teorema Central del Límite

3.1. El Teorema Central del Límite

3.1.1. TCL para sucesiones iid

Lema 3.1 *Para $i = 1, \dots, n$ supongamos que $a_i \in \mathbb{C}$, $b_i \in \mathbb{C}$ con $|a_i| \leq 1$ y $|b_i| \leq 1$. Entonces*

$$\Big| \prod_{i=1}^n a_i - \prod_{i=1}^n b_i \Big| \leq \sum_{i=1}^n |a_i - b_i|.$$

Demostración. Para $n = 2$ tenemos

$$a_1 a_2 - b_1 b_2 = a_1(a_2 - b_2) + (a_1 - b_1) b_2.$$

Tomando módulos y usando el hecho de que los módulos de a_i y b_i están acotados por 1 se obtiene el resultado. Para n cualquiera se usa inducción. ■

Teorema 3.1 (TCL para variables iid) *Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a.i.i.d. con $\mathrm{E}(X_n) = \mu$ y $\mathrm{Var}(X_n) = \sigma^2$. Sea $N \sim \mathcal{N}(0, 1)$. Si $S_n = X_1 + \cdots + X_n$ entonces*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N.$$

Demostración. Suponemos, sin pérdida de generalidad, que $\mathrm{E}(X_n) = 0$ y $\mathrm{Var}(X_n) = 1$. Sean

$$\varphi_n(t) = \mathrm{E}\, e^{itS_n/\sqrt{n}}, \qquad \varphi(t) = \mathrm{E}\, e^{itX_1},$$

entonces

$$\varphi_n(t) = (\mathrm{E}\, e^{itX_1/\sqrt{n}})^n = \varphi^n(t/\sqrt{n}).$$

Como existen los dos primeros momentos podemos hacer un desarrollo de φ:

$$\varphi\big(\frac{t}{\sqrt{n}}\big) = 1 + \frac{it\,\mathrm{E}(X_1)}{\sqrt{n}} + \frac{i^2 t^2\,\mathrm{E}(X_1^2)}{2n} + o\big(\frac{t^2}{n}\big) = 1 + 0 - \frac{t^2}{2n} + o\big(\frac{t^2}{n}\big)$$

donde

$$\big|o\big(\frac{t^2}{n}\big)\big| \leq \mathrm{E}\Big(\frac{|tX_1|^3}{n^{3/2}3!} \wedge \frac{2|tX_1|^2}{2n}\Big).$$

Veamos que

$$n o\big(\frac{t^2}{n}\big) \leq \mathrm{E}\Big(\frac{|tX_1|^3}{n^{1/2}3!} \wedge |tX_1|^2\Big) \to 0 \quad (n \to \infty).$$

Para esto observemos que

$$\frac{|tX_1|^3}{\sqrt{n}3!} \wedge |tX_1|^2 \leq |tX_1|^2 \in L^1$$

y también

$$\frac{|tX_1|^3}{\sqrt{n}3!} \wedge |tX_1|^2 \leq \frac{|tX_1|^3}{\sqrt{n}3!} \to 0, \quad (n \to \infty).$$

Por el TCD

$$\mathrm{E}\Big(\frac{|tX_1|^3}{\sqrt{n}\,3!} \wedge |tX_1|^2\Big) \to 0.$$

Ahora, usando el lema 3.1

$$\Big|\varphi^n\big(\frac{t}{\sqrt{n}}\big) - \big(1 - \frac{t^2}{2n}\big)^n\Big| \leq n\big|\varphi\big(\frac{t}{\sqrt{n}}\big) - \big(1 - \frac{t^2}{2n}\big)\big| = no\big(\frac{t^2}{n}\big) \to 0,$$

y ahora, como

$$\Big(1 - \frac{t^2}{2n}\Big)^n \to e^{t^2/2},$$

que es la f.c. de la distribución $\mathcal{N}(0,1)$, obtenemos el resultado. ■

3.1.2. El TCL de Lindeberg-Feller

Ahora generalizamos el TCL al caso de sumandos que no tienen la misma distribución. Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a.i. pero no necesariamente idénticamente distribuidas, supongamos que X_k tiene f.d. F_k, f.c. φ_k, y que $\mathrm{E}(X_k) = 0,\ \mathrm{Var}(X_k) = \sigma_k$. Definimos

$$s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2 = \mathrm{Var}(\sum_{i=1}^n X_i).$$

Decimos que X_k satisface la condición de Lindeberg si para todo t, cuando $n \to \infty$, tenemos

$$\frac{1}{s_n^2}\sum_{k=1}^n \mathrm{E}\left(X_k^2 \mathbf{1}_{\{|X_k/s_n|>t\}}\right) = \frac{1}{s_n^2}\sum_{k=1}^n \int_{|x|>ts_n} x^2\, dF_k(x) \to 0. \tag{3.1}$$

Observación 3.1

1. La condición de Lindeberg implica

$$\max_{k\leq n} \frac{\sigma_k^2}{s_n^2} \to 0, \quad (n \to \infty). \tag{3.2}$$

Para ver esto observemos que

$$\begin{aligned}
\frac{\sigma_k^2}{s_n^2} &= \frac{1}{s_n^2}\,\mathrm{E}(X_k^2) \\
&= \frac{1}{s_n^2}\,\mathrm{E}(X_k^2\mathbf{1}_{\{|X_k/s_n|\leq t\}}) + \frac{1}{s_n^2}\,\mathrm{E}(X_k^2\mathbf{1}_{\{|X_k/s_n|>t\}}) \\
&\leq t^2 + \frac{1}{s_n^2}\,\mathrm{E}(X_k^2\mathbf{1}_{\{|X_k/s_n|>t\}}) \to t^2.
\end{aligned}$$

De modo que para todo $t > 0$

$$\lim_{n\to\infty}\big(\max_{k\leq n}\frac{\sigma_k^2}{s_n^2}\big) \leq t^2.$$

Haciendo ahora $t \downarrow 0$ se obtiene el resultado.

2. La condición (3.2) implica
$$\max_{k\leq n} P(|X_k|/s_n > \varepsilon) \to 0 \tag{3.3}$$
por la desigualdad de Chebysheff. La condición (3.1) se conoce como infinitesimalidad asintótica uniforme.

Teorema 3.2 (Lindeberg-Feller) *La condición de Lindeberg (3.1) implica*
$$\frac{S_n}{s_n} \xrightarrow{d} \mathcal{N}(0,1).$$

Observación 3.2 Aunque no daremos la demostración, el recíproco también es cierto en el siguiente sentido: Si
$$\frac{1}{s_n^2}\max_{k\leq n}\sigma_k^2 \to 0 \qquad \text{y} \qquad \frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1),$$
entonces la condición de Lindeberg vale.

Demostración. (1) Comenzamos con un resultado preliminar. Sea $\{Y_n,\ n\geq 1\}$ una sucesión iid con distribución común F y f.c. φ. Sea N independiente de la sucesión con distribución de Poisson de parámetro c. Definimos $\chi_n = \sum_{i=1}^n Y_i$. Obtenemos la f.c. de χ_N de la siguiente manera: Para $t\in\mathbb{R}$,
$$\begin{aligned}
\mathrm{E}(e^{it\chi_N}) &= \sum_{k=0}^\infty \mathrm{E}(e^{it\chi_N}\mathbf{1}_{\{N=k\}}) = \sum_{k=0}^\infty \mathrm{E}(e^{it\chi_k}\mathbf{1}_{\{N=k\}})\\
&= \sum_{k=0}^\infty \mathrm{E}(e^{it\chi_k})P(N=k) = \sum_{k=0}^\infty \varphi^k(t)\frac{e^{-c}c^k}{k!}\\
&= e^{-c}e^{c\varphi(t)} = e^{c(\varphi(t)-1)}.
\end{aligned}$$
Concluimos, además, que $e^{c(\varphi(t)-1)}$ también es una f.c.

(2) Para ver que $S_n/s_n \xrightarrow{d} \mathcal{N}(0,1)$ por el teorema de continuidad tenemos que demostrar que la f.c. de S_n/s_n satisface
$$\varphi_{S_n/s_n}(t) = \prod_{k=1}^n \varphi_k(t/s_n) \to e^{-t^2/2} \tag{3.4}$$
que es la f.c. de $\mathcal{N}(0,1)$. Si demostramos que
$$\Big|\exp\Big\{\sum_{k=1}^n(\varphi_k(t/s_n)-1)\Big\} - \prod_{k=1}^n\varphi_k(t/s_n)\Big| \to 0 \tag{3.5}$$
entonces bastaría mostrar que
$$\sum_{k=1}^n(\varphi_k(t/s_n)-1) + t^2/2 \to 0. \tag{3.6}$$
Veamos primero la demostración de (3.5). Recordemos que $\exp\{\varphi_k(t/s_n)-1\}$ es una f.c. y por lo tanto su módulo está acotado por 1. Usando el lema 3.1,
$$\begin{aligned}
\Big|\exp\Big\{\sum_{k=1}^n(\varphi_k(t/s_n)-1)\Big\} - \prod_{k=1}^n\varphi_k(t/s_n)\Big| &= \Big|\prod_{k=1}^n e^{\varphi_k(t/s_n)-1} - \prod_{k=1}^n\varphi_k\Big|\\
&\leq \sum_{k=1}^n\big|e^{\varphi_k(t/s_n)-1} - \varphi_k(t/s_n)\big|\\
&= \sum_{k=1}^n\big|e^{\varphi_k(t/s_n)-1} - 1 - (\varphi_k(t/s_n)-1)\big|.
\end{aligned}$$

Observamos ahora que para $z \in \mathbb{C}$,

$$\begin{aligned} |e^z - 1 - z| &= \Big| \sum_{k=2}^{\infty} \frac{z^k}{k!} \Big| \leq \sum_{k=2}^{\infty} |z|^2 = \frac{|z|^2}{1-|z|} \\ &\leq \begin{cases} 2|z|^2, & \text{si } |z| \leq 1/2, \\ \delta|z|, & \text{si } |z| \leq \frac{\delta}{2} < \frac{1}{2}. \end{cases} \end{aligned} \tag{3.7}$$

Para t fijo usamos (2.5) con $n = 1$ para obtener las desigualdades

$$|\varphi_k(t/s_n) - 1| \leq \frac{t^2}{2s_n^2}\sigma_k^2 \tag{3.8}$$

$$\leq \frac{t^2}{2} \max_{k \leq n} \frac{\sigma_k^2}{s_n^2}. \tag{3.9}$$

Recordando (3.2), concluimos de (3.9) que dado $\delta > 0$, si n es suficientemente grande, para $k = 1, \dots, n$

$$|\varphi_k(t/s_n) - 1| \leq \frac{\delta}{2}. \tag{3.10}$$

Sea $z_k = \varphi_k(t/s_n) - 1$ y

$$\Big| \exp\Big\{ \sum_{k=1}^{n} (\varphi_k(t/s_n) - 1) \Big\} - \prod_{k=1}^{n} \varphi_k(t/s_n) \Big| \leq \sum_{k=1}^{n} |e^{z_k} - 1 - z_k| \leq \sum_{k=1}^{n} \delta |z_k|$$

para n grande, porque $\max_{k \leq n} |z_k| \leq \delta/2$ para n grande, y usando (3.8) obtenemos la cota

$$\leq \delta \frac{t^2}{2s_n^2} \sum_{k=1}^{n} \sigma_k^2 = \delta \frac{t^2}{2}$$

y como δ es arbitrario, obtenemos (3.5). Por lo tanto basta probar (3.6).

(3) Ahora demostramos (3.6). Tenemos

$$\sum_{k=1}^{n} (\varphi_k(t/s_n) - 1) + t^2/2 = \sum_{k=1}^{n} \mathrm{E}\left(e^{itX_k/s_n} - 1 - i\frac{t}{s_n}X_k - \frac{1}{2}\big(\frac{it}{s_n}\big)^2 X_k^2 \right).$$

Denotemos por A el argumento de la esperanza en la expresión anterior. Tenemos la descomposición

$$\sum_{k=1}^{n} \left[\mathrm{E}(A\mathbf{1}_{\{|X_k|/s_n \leq \varepsilon\}}) + \mathrm{E}(A\mathbf{1}_{\{|X_k|/s_n > \varepsilon\}}) \right] = I + II.$$

Ahora veremos que I y II son pequeños. Para I tenemos, usando (2.6) con $n = 2$,

$$\begin{aligned} |I| &\leq \sum_{k=1}^{n} \mathrm{E}(|A| \mathbf{1}_{\{|X_k/s_n| \leq \varepsilon\}}) \\ &\leq \sum_{k=1}^{n} \mathrm{E}\left(\frac{1}{3!} \Big| \frac{t}{s_n} X_k \Big|^3 \mathbf{1}_{\{|X_k/s_n| \leq \varepsilon\}} \right) \\ &= \frac{|t|^3}{6} \sum_{k=1}^{n} \mathrm{E}\left(\Big| \frac{X_k}{s_n} \Big|^3 \mathbf{1}_{\{|X_k/s_n| \leq \varepsilon\}} \right) \\ &\leq \frac{|t|^3}{6} \varepsilon \sum_{k=1}^{n} \mathrm{E}\left(\Big| \frac{X_k}{s_n} \Big|^2 \mathbf{1}_{\{|X_k/s_n| \leq \varepsilon\}} \right) \\ &\leq \frac{|t|^3}{6} \varepsilon \sum_{k=1}^{n} \frac{\sigma_k^2}{s_n^2} = \varepsilon \frac{|t|^3}{6} \end{aligned}$$

donde hemos usado el hecho de que

$$\sum_{k=1}^{n} \frac{\sigma_k^2}{s_n^2} = 1.$$

Ahora veamos que II es pequeño. Usando (2.6) tenemos

$$\begin{aligned}
|II| &\leq \sum_{k=1}^{n} \mathrm{E}(|A| \mathbf{1}_{\{|X_k/s_n|>\varepsilon\}}) \\
&\leq 2 \sum_{k=1}^{n} \frac{1}{2} \mathrm{E}\left(\Big| \frac{tX_k}{s_n} \Big|^2 \mathbf{1}_{\{|X_k/s_n|>\varepsilon\}} \right) \\
&\leq \frac{t^2}{s_n^2} \sum_{k=1}^{n} \mathrm{E}\left(X_k^2 \mathbf{1}_{\{|X_k/s_n|>\varepsilon\}} \right) \to 0
\end{aligned}$$

por la condición de Lindeberg. Esto completa la demostración de (3.6). ■

A continuación presentamos una condición suficiente para la condición de Lindeberg que es más sencilla de verificar.

Corolario 3.1 (Condición de Lyapunov) *Sea* $\{X_k,\ k \geq 1\}$ *una sucesión de v.a.i. con* $\mathrm{E}(X_k) = 0,\ \mathrm{Var}(X_k) = \sigma_k^2 < \infty,\ s_n^2 = \sum_{k=1}^{n} \sigma_k^2$. *Si para algún* $\delta > 0$

$$\frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} \mathrm{E}\,|X_k|^{2+\delta} \to 0$$

entonces se satisface la condición de Lindeberg y por lo tanto el TCL vale.

Observación 3.3 Un caso útil es cuando $\delta = 1$:

$$\frac{1}{s_n^3} \sum_{k=1}^{n} \mathrm{E}\,|X_k|^3 \to 0.$$

Demostración. Tenemos

$$\begin{aligned}
\frac{1}{s_n^2} \sum_{k=1}^{n} \mathrm{E}\left(X_k^2 \mathbf{1}_{\{|X_k/s_n|>t\}} \right) &= \sum_{k=1}^{n} \mathrm{E}\left(\Big| \frac{X_k}{s_n} \Big|^2 \mathbf{1}_{\{|X_k/s_n|>t\}} \right) \\
&\leq \sum_{k=1}^{n} \mathrm{E}\left(\Big| \frac{X_k}{s_n} \Big|^2 \Big| \frac{X_k}{ts_n} \Big|^\delta \mathbf{1}_{\{|X_k/ts_n|>1\}} \right) \\
&\leq \frac{1}{t^\delta s_n^{2+\delta}} \sum_{k=1}^{n} \mathrm{E}\,|X_k|^{2+\delta} \to 0.
\end{aligned}$$

■

Ejemplo 3.1
Consideramos ahora el comportamiento asintótico de los records. Sea $\{X_n,\ n \geq 1\}$ una sucesión de variables iid con distribución común continua F y sea $\mathbf{1}_k = \mathbf{1}_{\{X_k \text{ es un record}\}}$, $\mu_n = \sum_1^n \mathbf{1}_k$, de modo que μ_n es el número de records en $X_1, \dots, X_n$. La continuidad de F garantiza que en la sucesión con probabilidad cero ocurren dos valores iguales.

El siguiente resultado, que se debe a Rényi, muestra que las variables $\mathbf{1}_k$ son independientes de Bernoulli con probabilidad de éxito $P(\mathbf{1}_k = 1) = 1/k$. Primero introducimos la definición del rango relativo de una variable en una muestra: Si tenemos una muestra $X_1, \dots, X_n$ el rango relativo de X_n, denotado R_n, se define como

$$R_n = \sum_{j=1}^{n} \mathbf{1}_{\{X_j \geq X_n\}},$$

de modo que R_n vale 1 si X_n es un record, vale 2 si X_n es el segundo mayor valor, y así sucesivamente.

Teorema 3.3 (Rényi) *Sea $(X_n, n \geq 1)$ una sucesión i.i.d. con f.d. común continua F.*
(a) Las variables $(R_n, n \geq 1)$ son independientes y

$$P(R_n = k) = \frac{1}{n}, \quad \textit{para } k = 1, \dots, n.$$

(b) Las variables $(\mathbf{1}_n, n \geq 1)$ son independientes de Bernoulli con

$$P(\mathbf{1}_k = 1) = \frac{1}{k}$$

Demostración. (b) se obtiene a partir de (a). Para la primera parte observamos que hay $n!$ ordenamientos de las variables $X_1, \dots, X_n$. Por simetría todos los ordenamientos tienen igual probabilidad $1/n!$.

La observación clave en la demostración es que cada realización de las variables $R_1, \dots, R_n$ determina un ordenamiento de las variables $X_1, \dots, X_n$: El rango R_n nos dice que lugar ocupa X_n en relación a las primeras $n-1$ variables ordenadas, y de esta manera podemos reconstruir los ordenamientos sucesivos, y por lo tanto el ordenamiento final, a partir de los rangos relativos. Veamos un ejemplo. Supongamos que $R_1 = 1, R_2 = 2, R_3 = 2, R_4 = 3$. El primer rango siempre vale 1; $R_2 = 2$ nos dice que la segunda variable ocupa el segundo lugar en el ordenamiento de las dos primeras, es decir, que $X_2 < X_1$; $R_3 = 2$ nos dice que X_3 ocupa el segundo lugar en el ordenamiento de las tres primeras, y por lo tanto debe ser menor que la mayor (X_1) y mayor que la menor (X_2), es decir, $X_2 < X_3 < X_1$. Finalmente, $R_4 = 3$ dice que hay dos variables mayores que X_4 (X_3 y X_1) y una menor (X_2), así que el orden final es

$$X_2 < X_4 < X_3 < X_1.$$

Ahora, como las variables X_i tienen todas la misma distribución, todos los ordenamientos tienen igual probabilidad, que es $1/n!$, y como cada ordenamiento corresponde a una realización de las variables $R_1, \dots, R_n$ tenemos que

$$P(R_1 = r_1, \dots, R_n = r_n) = \frac{1}{n!}$$

para $r_i \in \{1, 2, \dots, i\},\ i = 1, 2, \dots, n$.

Observamos ahora que

$$\begin{aligned} P(R_n = r_n) &= \sum_{r_1, \dots, r_{n-1}} P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = r_n) \\ &= \sum_{r_1, \dots, r_{n-1}} \frac{1}{n!} \end{aligned}$$

y cada r_i toma i valores posibles, por lo que la suma tiene $1 \cdot 2 \cdots (n-1) = (n-1)!$ sumandos y

$$P(R_n = r_n) = \frac{(n-1)!}{n!} = \frac{1}{n}, \quad n = 1, 2, \dots$$

En consecuencia

$$P(R_1 = r_1, \dots, R_n = r_n) = \frac{1}{n!} = P(R_1 = r_1) P(R_2 = r_2) \cdots P(R_n = r_n)$$

y esto demuestra la independencia. ∎

Veamos a continuación una ley fuerte para las variables μ_n. Recordamos primero dos resultados que necesitamos para la demostración.

Teorema 3.4 (Criterio de Convergencia de Kolmogorov) *Supongamos que* $(X_n,\ n \geq 1)$ *es una sucesión de v.a.i. Si*

$$\sum_n \operatorname{Var}(X_n) < \infty,$$

entonces

$$\sum_{n\geq 1}(X_n - \operatorname{E}(X_n)) \quad \textit{converge c.p.1.}$$

Demostración. Sin pérdida de generalidad podemos suponer que $\operatorname{E}(X_j) = 0$. Entonces la suma de las varianzas es

$$\sum_{n\geq 1} \operatorname{E} X_n^2 < \infty.$$

Esto implica que (S_n) es de Cauchy en L^2 porque para $m < n$

$$||S_n - S_m||_2^2 = \operatorname{Var}(S_n - S_m) = \sum_{j=m+1}^{n} \operatorname{E} X_j^2 \to 0,$$

cuando $m, n \to \infty$. En consecuencia (S_n) también es de Cauchy en probabilidad porque

$$P(|S_n - S_m| \geq \varepsilon) \leq \varepsilon^{-2} \operatorname{Var}(S_n - S_m) = \varepsilon^{-2} \sum_{j=m+1}^{n} \operatorname{Var}(X_j) \to 0$$

cuando $n, m \to \infty$. Por el Teorema de Lévy (S_n) es convergente c.p.1. ■

Lema 3.2 (Kronecker) *Sean* x_k *y* a_k *dos sucesiones de números reales tales que* $0 < a_n \uparrow \infty$. *Si*

$$\sum_{k=1}^{\infty} \frac{x_k}{a_k} \quad \textit{converge,}$$

entonces

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=1}^{n} x_k = 0.$$

Demostración. Sea $r_n = \sum_{k=n+1}^{\infty} x_k/a_k$ de modo que $r_n \to 0$ cuando $n \to \infty$. Dado $\varepsilon > 0$ existe $N_0 = N_0(\varepsilon)$ tal que si $n \geq N_0$ tenemos $|r_n| \leq \varepsilon$. Además

$$\frac{x_n}{a_n} = r_{n-1} - r_n$$

de modo que

$$\begin{aligned}\sum_{k=1}^{n} x_k &= \sum_{k=1}^{n} (r_{k-1} - r_k) a_k \\ &= \sum_{j=1}^{n-1} (a_{j+1} - a_j) r_j + a_1 r_0 - a_n r_n.\end{aligned}$$

Entonces, para $n \geq N_0$,

$$\begin{aligned}
\frac{1}{a_n}\Big|\sum_{k=1}^{n} x_k\Big| &\leq \frac{1}{a_n}\Big(\sum_{j=1}^{N_0-1}(a_{j+1}-a_j)|r_j| + \sum_{j=N_0}^{n-1}(a_{j+1}-a_j)|r_j|\Big) + \Big|\frac{a_1 r_0}{a_n}\Big| + \Big|\frac{a_n r_n}{a_n}\Big| \\
&= \frac{Const}{a_n} + \frac{\varepsilon}{a_n}\sum_{j=N_0}^{n-1}(a_{j+1}-a_j)|r_j| + |r_n| \\
&= o(1) + \frac{\varepsilon(a_n - a_{N_0})}{a_n} + \varepsilon \\
&\leq 2\varepsilon + o(1).
\end{aligned}$$

y esto demuestra el resultado. ■

Proposición 3.1 *El número de records en una sucesión i.i.d. crece logaritmicamente y se tiene con probabilidad 1*

$$\lim_{n\to\infty} \frac{\mu_n}{\log n} = 1.$$

Para la demostración usaremos el siguiente resultado de Análisis:

$$\log n - \sum_{j=1}^{n} \frac{1}{j} \to \gamma$$

cuando $n \to \infty$ donde $\gamma > 0$ es la constante de Euler.

Demostración. Tenemos que las $\mathbf{1}_j$ son independientes,

$$\mathrm{E}(\mathbf{1}_j) = P(\mathbf{1}_j = 1) = \frac{1}{j},$$

$$\mathrm{Var}(\mathbf{1}_j) = \mathrm{E}(\mathbf{1}_j)^2 - (\mathrm{E}\,\mathbf{1}_j)^2 = \frac{1}{j} - \frac{1}{j^2} = \frac{j-1}{j}.$$

Usando esto obtenemos

$$\sum_{j=2}^{\infty} \mathrm{Var}\Big(\frac{\mathbf{1}_j}{\log j}\Big) = \sum_{j=2}^{\infty} \frac{1}{(\log j)^2}\,\mathrm{Var}(\mathbf{1}_j) = \sum_{j=2}^{\infty} \frac{j-1}{j^2(\log j)^2} < \infty$$

Por lo tanto el criterio de convergencia de Kolmogorov implica que

$$\sum_{j=2}^{\infty}\Big(\frac{\mathbf{1}_j - \mathrm{E}(\mathbf{1}_j)}{\log j}\Big) = \sum_{j=2}^{\infty} \frac{\mathbf{1}_j - 1/j}{\log j}$$

converge y por el lema de Kronecker

$$\frac{1}{\log n}\sum_{j=1}^{n}\Big(\mathbf{1}_j - \frac{1}{j}\Big) \to 0 \quad c.p.1,$$

pero

$$\frac{1}{\log n}\sum_{j=1}^{n}\Big(\mathbf{1}_j - \frac{1}{j}\Big) = \frac{1}{\log n}\Big(\sum_{j=1}^{n}\mathbf{1}_j - \sum_{j=1}^{n}\frac{1}{j}\Big) = \frac{1}{\log n}\Big(\mu_n - \sum_{j=1}^{n}\frac{1}{j}\Big).$$

Por lo tanto

$$\frac{\mu_n}{\log n} - 1 = \frac{\mu_n - \sum_{j=1}^{n} j^{-1}}{\log n} + \frac{\sum_{j=1}^{n} j^{-1} - \log n}{\log n} \to 0.$$

■

Probaremos ahora que

$$\frac{\mu_n - \log n}{\sqrt{\log n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Para ver esto recordemos que $\mathrm{E}(\mathbf{1}_k) = 1/k$ y $\mathrm{Var}(\mathbf{1}_k) = \frac{1}{k} - \frac{1}{k^2}$. Por lo tanto

$$s_n^2 = \mathrm{Var}(\mu_n) = \sum_{k=1}^{n} \big(\frac{1}{k} - \frac{1}{k^2}\big) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k^2} \sim \log n,$$

y en consecuencia $s_n^3 \sim (\log n)^{3/2}$. Por otro lado

$$\mathrm{E}\,|\mathbf{1}_k - \mathrm{E}(\mathbf{1}_k)|^3 = \mathrm{E}\,\big|\mathbf{1}_k - \frac{1}{k}\big|^3 = \big|1 - \frac{1}{k}\big|^3 \frac{1}{k} + \big(\frac{1}{k}\big)^3 \big(1 - \frac{1}{k}\big) \leq \frac{1}{k} + \frac{1}{k^3},$$

y por lo tanto

$$\frac{1}{s_n^3} \sum_{k=1}^{n} \mathrm{E}\,|\mathbf{1}_k - \mathrm{E}(\mathbf{1}_k)|^3 \leq \frac{1}{(\log n)^{3/2}} \sum_{k=1}^{n} \big(\frac{1}{k} + \frac{1}{k^3}\big) \sim \frac{\log n}{(\log n)^{3/2}} \to 0.$$

De modo que la condición de Lyapunov es válida y

$$\frac{\mu_n - \mathrm{E}(\mu_n)}{\sqrt{\mathrm{Var}(\mu_n)}} \xrightarrow{d} \mathcal{N}(0,1).$$

Observamos que $\sqrt{\mathrm{Var}(\mu_n)} \sim s_n \sim \sqrt{\log n}$ y

$$\frac{\mathrm{E}(\mu_n) - \log n}{\sqrt{\log n}} = \frac{1}{\sqrt{\log n}} \Big(\sum_{k=1}^{n} \frac{1}{k} - \log n\Big) \sim \frac{\gamma}{\sqrt{\log n}} \to 0,$$

donde γ es la constante de Euler. Por el teorema de convergencia a familias

$$\frac{\mu_n - \log n}{\sqrt{\log n}} \xrightarrow{d} \mathcal{N}(0,1).$$

3.1.3. Extensiones

En esta sección presentamos sin demostración algunas extrensiones del Teorema Central de Límite que demostramos anteriormente.

Cantidad Aleatoria de Sumandos

Teorema 3.5 *Sea $X_1, X_2, \dots$ v.a.i.i.d. con media 0 y varianza finita σ^2 y sea $S_n = X_1 + \cdots + X_n$ para $n \geq 1$. Sea $\{N(t), t \geq 0\}$ una familia de variables aleatorias positivas, con valores enteros y tales que para algún $0 < \theta < \infty$ se tiene que*

$$\frac{N(t)}{t} \xrightarrow{P} \theta \quad \textit{cuando } t \to \infty.$$

Entonces

$$\frac{S_{N(t)}}{\sigma\sqrt{N(t)}} \xrightarrow{d} \mathcal{N}(0,1), \quad \textit{cuando } t \to \infty.$$

$$\frac{S_{N(t)}}{\sigma\sqrt{\theta t}} \xrightarrow{d} \mathcal{N}(0,1), \quad \textit{cuando } t \to \infty.$$

Demostración. Ver Gut, *Probability: A Graduate Course*, Cap. 7.

Velocidad de Convergencia

El TCL nos dice que si n es grande, $\bar{X}_n - \mu \approx (\sigma/\sqrt{n})N$ donde $N \sim \mathcal{N}(0,1)$. Un apregunta natural es ¿Cuán buena es esta aproximación? es decir, ¿con qué velocidad se da la convergencia en el TCL?

El siguiente resultado fue probado, de manera independiente, por Berry y Essen.

Teorema 3.6 (Berry-Essen) *Sean $X_1, X_2, \dots$ v.a.i.i.d. con $S_n = X_1 + \dots + X_n$. Sea $\mu = \mathrm{E}(X_1)$, $\sigma^2 = \mathrm{Var}(X_1)$ y supongamos que $\gamma^3 = \mathrm{E}\,|X|^3 < \infty$. Sea F_n la función de distribución de $(S_n - n\mu)/\sigma\sqrt{n}$, entonces*

$$\sup_x \big|F_n(x) - \Phi(x)\big| \leq C\frac{\gamma^3}{\sigma^3\sqrt{n}},$$

donde c es una constante numérica.

En el caso de variables independientes pero no idénticamente distribuidas tenemos el siguiente resultado.

Teorema 3.7 *Sean $X_1, X_2, \dots$ v.a.i. con media cero y $S_n = X_1 + \dots + X_n$. Sea $\sigma_k^2 = \mathrm{Var}(X_k)$ y supongamos que $\gamma_k^3 = \mathrm{E}\,|X_k|^3 < \infty$ para todo k. Sea $s_n^2 = \sum_1^n \sigma_k^2$, $\beta_n = \sum_1^n \gamma_k^3$ y F_n la función de distribución de S_n/s_n, entonces*

$$\sup_x \big|F_n(x) - \Phi(x)\big| \leq C\frac{\beta_n^3}{s_n^3},$$

donde C es una constante numérica.

Demostración. Ver Gut, *Probability: A Graduate Course*, Cap. 7.

Teorema Central de Límite Local

El TCL dice que la f.d. de sumas de variables aleatorias S_n debidamente normalizadas convergen a la f.d. $\mathcal{N}(0,1)$. La pregunta que nos hacemos ahora es si el resultado vale a nivel de las densidades. En general, no es cierto que si $F_n(x) \to F(x)$ entonces $F_n'(t) \to F'(t)$ cuando $n \to \infty$ pero bajo condiciones bastante generales tenemos el siguiente resultado, que se conoce como el TCL local.

Teorema 3.8 (TCL local) *Sean $X_1, X_2, \dots$ v.a.i.i.d. con media cero y varianza 1, y supongamos que su función característica φ satisface*

$$\int_{\mathbb{R}} |\varphi(t)|^r dt < \infty$$

para algún entero $r \geq 1$. Entonces la densidad g_n de $U_n = S_n/\sqrt{n}$ existe para $n \geq r$ y

$$g_n(x) \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad n \to \infty$$

uniformemente en $x \in \mathbb{R}$.

Demostración. Ver Grimmet & Stirzaker, Cap. 5.

Dependencia

El TCL puede extenderse a ciertas clases de sucesiones dependientes y hay numerosos resultados en este sentido. Vamos a mencionar sólo uno de ellos para sucesiones que satisfacen una condición 'mixing' o de mezcla.

Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias y consideremos dos conjuntos, $A \in \sigma(X_1, \dots X_k)$ y $B \in \sigma(X_{k+n}, X_{k+n+1}, \dots)$ para $k \geq 1, n \geq 1$. Si existe una sucesión α_n tal que para cualesquiera A, B que satisfagan la condición anterior se tiene que

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_n$$

y $\alpha_n \to 0$, decimos que la sucesión $X_n, n \geq 1$ es α-mixing. La idea es que si n es grande, los conjuntos A y B dependen de variables que están muy separadas en la sucesión, y en consecuencia deben ser (casi) independientes.

Teorema 3.9 (TCL para sucesiones α-mixing) *Sea $X_n, n \geq 1$ una sucesión alpha-mixing. Si existe $\delta > 0$ tal que $\mathrm{E}\,|X_1|^{2+\delta} < \infty$ y*

$$\sum_{n=0}^{\infty} \alpha_n^{\delta/(2+\delta)} < \infty$$

entonces vale el TCL.

Para la demostración y extensiones a otros tipos de condiciones mixing ver P. Doukhan, Mixing. Properties and Examples. Lecture Notes in Statistics 85, Springer.

Una condición sencilla que implica α-mixing es la m dependencia. Una sucesión $X_n, n \geq 1$ es m-dependiente si $(X_1, \dots, X_k)$ y $X_{k+n}, \dots)$ son independientes siempre que $n > m$. En este caso la sucesión es α-mixing con $\alpha_n = 0$ para $n > m$.

Por ejemplo, si Y_1, Y_2, son i.i.d. y definimos $X_n = f(Y_n, \dots, Y_{n+m})$ para una función $f : \mathbb{R}^{m+1} \to \mathbb{R}$ entonces $(X_n, n \geq 1)$ es estacionaria y m-dependiente.

Sistemas Triangulares

Supongamos que para cada n, las variables

$$X_{n1}, \dots, X_{nr_n}$$

son independientes. Una colección de este tipo se conoce como un arreglo triangular de variables aleatorias. El TCL de Lindeberg considera el caso especial en el cual $r_n = n$ y $X_{nk} = X_k$ para todo n. Supongamos que todas las variables están centradas y sea

$$\sigma_{nk}^2 = \mathrm{E}(X_{nk}^2), \qquad s_n^2 = \sum_{k=1}^{r_n} \sigma_{nk}^2. \tag{3.11}$$

Supongamos que $s_n > 0$. La condición de Lindeberg en este caso es

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \frac{1}{s_n^2} \int_{|X_{nk}| \geq \varepsilon s_n} X_{nk}^2 \, dP = 0 \tag{3.12}$$

para $\varepsilon > 0$.

Teorema 3.10 *Supongamos que para cada n las variables $X_{n1}, \dots, X_{nr_n}$ son independientes y satisfacen (3.11). Si (3.12) vale para todo $\varepsilon > 0$ entonces $S_n/s_n \xrightarrow{d} \mathcal{N}(0,1)$.*

Demostración. Ejercicio.

Supongamos ahora que $\mathrm{E}\,|X_{nk}|^{2+\delta} < \infty$ para algún $\delta > 0$ y que la condición de Lyapunov se satisface:

$$\lim_{n} \sum_{k=1}^{r_n} \frac{1}{s_n^{2+\delta}} \mathrm{E}\left(|X_{nk}|^{2+\delta}\right) = 0, \tag{3.13}$$

entonces la condición de Lindeberg vale y el teorema es válido.

Teorema 3.11 *Supongamos que para cada n las variables $X_{n1}, \dots, X_{nr_n}$ son independientes y satisfacen (3.11). Si (3.13) vale para algún $\delta > 0$ entonces $S_n/s_n \xrightarrow{d} \mathcal{N}(0,1)$.*

3.1.4. Grandes Desviaciones

La LFGN nos dice que la suma S_n de n v.a.i.i.d. es aproximadamente $n\mu$, donde μ es la media común. El TCL dice que las desviaciones de S_n de $n\mu$ son típicamente de orden $\sqrt{n}$, es decir, son de orden menor que el valor medio. Sin embargo, la distancia de S_n a $n\mu$ puede ser de orden mayor que $\sqrt{n}$, por ejemplo n^α, donde $\alpha > 1/2$, y en particular puede ser de orden n. La teoría de Grandes Desviaciones estudia el comportamiento asintótico de

$$P(|S_n - n\mu| > n^\alpha) \quad \text{cuando } n \to \infty$$

para $\alpha > 0.5$.

Sea $X_1, X_2, \dots$ una sucesión de v.a.i.i.d. con $S_n = X_1 + \dots + X_n$, $n \geq 1$ Supongamos que $\mathrm{E}(X) = \mu$ y sumas parciales $S_n = X_1 + \dots + X_n$ y queremos estimar $P(S_n > na)$ para $a > \mu$. La función generadora de momentos $M(t) = \mathrm{E}(e^{tX})$ y su logaritmo, la función generadora de cumulantes $\Lambda(t) = \log M(t)$ juegan un papel en este problema. La función Λ tiene las siguientes propiedades: $\Lambda(0) = 0$, $\Lambda'(0) = \mu$ si $M'(0)$ existe y $\Lambda(t)$ es convexa donde esté definida. Esto se puede ver porque la segunda derivada $\lambda''(t)$ es no-negativa.

Definimos

$$\Lambda^*(a) = \sup_{t\in\mathbb{R}}(at - \Lambda(t)), \quad a \in \mathbb{R}.$$

Teorema 3.12 (Grandes Desviaciones) *Sean $X_1, X_2, \dots$ v.a.i.i.d. con media μ y supongamos que su f.g.m. $M(t) = \mathrm{E}(e^{tX})$ es finita en alguna vecindad del origen. Sea $a > \mu$ tal que $P(X > a) > 0$. Entonces $\Lambda^*(a) > 0$ y*

$$\frac{1}{n}\log P(S_n > na) \to \Lambda^*(a), \quad \textit{cuando } n \to \infty.$$

Demostración. Ver Grimmet & Stirzaker, Cap. 5.

Bajo las condiciones del teorema, $P(S_n > na)$ decae exponencialmente como $e^{-n\Lambda(a)}$.

3.2. Distribuciones Estables

Si $X_1, X_2, \dots$ son i.i.d. con media μ y varianza σ^2 finitas, sabemos por los resultados anteriores que

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

Por lo tanto la sucesión $a_n^{-1}(S_n - b_n)$ sólo puede converger a una distribución normal. Si no requerimos que las varianzas sean finitas, es posible obtener límites que no son normales. Por ejemplo, supongamos que X_i tiene distribución de Cauchy con densidad

$$f(x) = \frac{\theta}{\pi(x^2+\theta^2)}, \quad x \in \mathbb{R},\ \theta > 0.$$

La función característica correspondiente es $\varphi(t) = e^{\theta|t|}$, por lo tanto S_n tiene f.c. $\varphi^n(t) = e^{-n\theta|t|}$ y $n^{-1}S_n$ tiene f.c. $\varphi^n(t/n) = e^{-\theta|t|}$. En consecuencia $n^{-1}S_n \xrightarrow{d} X$ donde X tiene distribución de Cauchy con el mismo parámetro θ. Como $\mathrm{E}\,|X| = \infty$ esto no contradice los resultados previos.

Definición 3.1 La distribución de una v.a. X es estable en sentido amplio si dadas X, X_1, $X_2, \dots$ v.a.i.i.d. con $S_n = X_1 + \dots + X_n$ para $n \geq 1$ existen constantes $a_n > 0$ y $b_n \in \mathbb{R}$ para $n \geq 1$ tales que

$$S_n \stackrel{d}{=} a_nX + b_n, \qquad \text{para todo } n.$$

La distribución es estrictamente estable si $b_n = 0$ para todo n.

Observamos que la distribución normal con media 0 es estrictamente estable con $a_n = \sqrt{n}$ y que la distribución de Cauchy estándar es estrictamente estable con $a_n = n$.

En general, sean $X, X_1, X_2, \ldots$ v.a.i. con la misma distribución estable y supongamos para simplificar que la distribución es simétrica. Entonces podemos interpretar S_{2n} de dos maneras:

$$S_{2n} = \sum_{k=1}^{n} X_k + \sum_{k=n+1}^{2n} X_k \quad \text{y} \quad S_{2n} = \sum_{k=1}^{2n} X_k$$

y recordando que $S_n \stackrel{d}{=} a_n X$ obtenemos que

$$a_2 a_n X \stackrel{d}{=} a_2 S_n \stackrel{d}{=} S_{2n} \stackrel{d}{=} a_{2n} X,$$

lo cual implica que $a_{2n} = a_2 a_n$. De manera similar si dividimos la suma S_{mn} en bloques del mismo tamaño, obtenemos que la sucesión $\{a_n,\ n \geq 1\}$ es multiplicativa:

$$a_{mn} = a_m a_n.$$

La ecuación anterior se conoce como la ecuación de Hamel y es posible demostrar que su solución

$$n^{1/\alpha} \quad \text{para } \alpha > 0. \tag{3.14}$$

Por otro lado como $S_n \stackrel{d}{=} n^{1/\alpha} X$, la función característica de una distribución estable simétrica debe satisfacer

$$\varphi(t) = \big(\varphi(t/n^{1/\alpha})\big)^n \quad \text{para todo } n \in \mathbb{N}. \tag{3.15}$$

Es posible demostrar que es este caso la función característica es

$$\varphi(t) = e^{-c|t|^\alpha}, \tag{3.16}$$

donde $c > 0$ es una constante. Es fácil verificar que esta f.c. satisface (3.15).

Para $\alpha > 2$ (3.16) implica que $\varphi(t) = 1 + o(t^2)$ para t cerca del origen, lo cual implica que la varianza vale 0, y por unicidad φ no puede ser una f.c. Por lo tanto los únicos valores posibles son $0 < \alpha \leq 2$.

Resumimos las propiedades principales de las distribuciones estables simétricas en la siguiente proposición

Proposición 3.2 *Si X tiene distribución estable simétrica con índice α y f.c. φ entonces*

- $0 < \alpha \leq 2$.
- $\varphi(t) = e^{c|t|^\alpha},\ c > 0$.
- $\mathrm{E}\,|X|^r < \infty$ *para* $r < \alpha < 2$.
- $\mathrm{E}\,|X|^r = \infty$ *para* $r \geq \alpha,\ 0 < \alpha < 2$.

En el caso general es posible demostrar que la f.c. es de la forma

$$\varphi(t) = \exp\big\{i\mu t - c|t|^\alpha\big(1 + i\beta\frac{t}{|t|}\tan\frac{\pi\alpha}{2}\big)\big\} \tag{3.17}$$

para $0 < \alpha \leq 2,\ \alpha \neq 1,\ |\beta| \leq 1,\ \mu \in \mathbb{R},\ c > 0$, y

$$\varphi(t) = \exp\big\{i\mu t - c|t|^\alpha\big(1 + i\beta\frac{t}{|t|}\frac{2}{\pi}\log|t|\big)\big\} \tag{3.18}$$

para $\alpha = 1,\ |\beta| \leq 1,\ \mu \in \mathbb{R},\ c > 0$.

El parámetro α es el *índice*, y β es el parámetro de asimetría. Observamos que la distribución es simétrica respecto a μ cuando $\beta = 0$.

Además, como la f.c. es integrable, concluimos que las distribuciones estables simétricas son absolutamente continuas, y que las densidades pueden obtenerse usando la fórmula de inversión. Sin embargo, sólo para tres casos $\alpha = 1/2, 1$ y 2 se tienen expresiones cerradas. Para el resto de los valores las densidades se expresan como series infinitas.

Veamos que una variable con f.c. de la forma (3.17) o (3.18) tiene que ser estable. Si $X_1, \dots, X_n$ son i.i.d. con f.c. de alguna de la formas descritas. Sea g el exponente de la función característica y $\lambda = 1/\alpha$ en el caso (3.17) y $\lambda = 1$ en el caso (3.18). entonces, en el primer caso

$$g(n^\lambda t) = itn^\lambda \mu - cn|t|^\alpha \big(1 + i\beta \frac{t}{|t|} \tan \frac{\pi\alpha}{2}\big) = ng(t) - it\mu(n - n^\lambda),$$

y en el segundo,

$$\begin{aligned} g(n^\lambda t) = g(nt) &= itn\mu - cn|t|\big(1 + i\beta \frac{t}{|t|} \frac{2}{\pi}(\log n + \log|t|)\big) \\ &= ng(t) - itc\beta\frac{2}{\pi} n \log n. \end{aligned}$$

Por lo tanto, en ambos casos

$$\varphi^n(t) = \exp\{ng(t)\} = \exp\{g(n^\lambda t)\}\exp\{ib_n t\} = \varphi(n^\lambda t)\exp\{ib_n t\}$$

donde $b_n = \mu(n - n^\lambda)$ en el primer caso y $b_n = c\beta(2/\pi)n \log n$ en el segundo caso. Por lo tanto $S_n \stackrel{d}{=} a_n X + b_n$ con $a_n = n^\lambda = n^{1/\alpha}$ y X es estable.

Teorema 3.13 *La variable x es estable si y sólo si hay una sucesión $(X_n, n \geq 1)$ de v.a.i.i.d. con $S_n = X_1 + \cdots + X_n$ tal que*

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} X \tag{3.19}$$

where $a_n > 0$, $b_n \in \mathbb{R}$.

Demostración. Si X es estable, sean $X_1, \dots, X_n$ i.i.d. con la misma distribución que X. Entonces $S_n = X_1 + \cdots + X_n \stackrel{d}{=} a_n X + b_n$ por la definición y por lo tanto (3.19) vale.

Supongamos ahora que $X_1, X_2, \dots$ son i.i.d. con $V_n = a_n^{-1}(S_n - b_n) \xrightarrow{d} X$. Si X es una variable degenerada, entonces es estable, asi que podemos suponer que X no es degenerada. Sea $r \in \mathbb{N}$ fijo, definimos

$$\begin{aligned} S_n^{(1)} &= X_1 + \cdots + X_n, \\ S_n^{(2)} &= X_{n+1} + \cdots + X_{2n}, \\ &\vdots \\ S_n^{(r)} &= X_{(r-1)n+1} + \cdots + X_{rn}. \end{aligned}$$

Llamemos

$$Z_n^{(j)} = \frac{S_n^{(j)} - b_n}{a_n}$$

y $W_n^{(r)} = Z_n^{(1)} + \cdots Z_n^{(r)}$. Observamos que las variables $Z_n^{(j)}$ son independientes y todas tienen la misma distribución: $Z_n^{(j)} \stackrel{d}{=} Z_n^{(1)}$ para todo j. Por lo tanto $Z_n^{(j)} \xrightarrow{d} X$ para todo j. En consecuencia

$$W_n^{(r)} \xrightarrow{d} Z_1 + \cdots Z_r,$$

donde las variables $Z_1, \dots, Z_r$ son i.i.d. con $Z_i \stackrel{d}{=} X$.

Pero también podemos escribir

$$\begin{aligned} W_n^{(r)} &= \frac{X_1 + \cdots + X_{rn} - rb_n}{a_n} = \frac{a_{rn}}{a_n}\Big(\frac{X_1 + \cdots X_{rn} - b_{rn}}{a_{rn}}\Big) + \frac{b_{rn} - rb_n}{a_n} \\ &= \alpha_n^{(r)} V_{rn} + \beta_n^{(r)} \end{aligned}$$

donde $\alpha_n^{(r)} = a_{rn}/a_n > 0$, $\beta_n^{(r)} = (b_{rn} - rb_n)/a_n$ y $V_{rn} = (S_{rn} - b_{rn})/a_{rn}$. Resumiendo,

$$V_{rn} = \frac{W_n^{(r)} - \beta_n^{(r)}}{\alpha_n^{(r)}}, \quad V_{rn} \xrightarrow{d} X, \quad W_n^{(r)} \xrightarrow{d} Z_1 + \cdots + Z_r.$$

Por el teorema de convergencia a familias $\alpha_n^{(r)} \to \alpha_r > 0$, $\beta_n^{(r)} \to \beta_r$ y

$$X \stackrel{d}{=} \frac{Z_1 + \cdots + Z_r - \beta_r}{\alpha_r}$$

de modo que X es estable. ■

Observación 3.4 Si r y s son enteros positivos

$$\alpha_n^{(rs)} = \frac{a_{rsn}}{a_n} = \frac{a_{rn}}{a_n}\frac{a_{rsn}}{a_{rn}} = \alpha_n^{(r)}\alpha_{rn}^{(s)},$$

y haciendo $n \to \infty$ obtenemos $\alpha_{rs} = \alpha_r \alpha_s$.

3.2.1. Dominios de Atracción.

Definición 3.2 Sea $X, X_1, X_2, \dots$ variables i.i.d. con sumas parciales $\{S_n,\ n \geq 1\}$. Decimos que X, o equivalentemente su f.d. F_X, pertenece al dominio de atracción de la distribución (no degenerada) G si existen sucesiones de normalización $a_n > 0,\ b_n \in \mathbb{R}$ para $n \geq 1$ tales que

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} G \quad (n \to \infty).$$

Usaremos la notación $F_X \in \mathcal{D}(G)$ o alternativamente $X \in \mathcal{D}(G)$.

La primera observación es que si $\operatorname{Var}(X) < \infty$, el TCL nos dice que X pertenece al dominio de atracción de la distribución normal, basta tomar $b_n = n\,\mathrm{E}(X)$ y $a_n = \sqrt{x \operatorname{Var}(X)}$. En particular, la distribución normal pertenece a su dominio de atracción y en general lo mismo pasa para las leyes estables.

Teorema 3.14 *Una v.a. con f.d. F pertenece al dominio de atracción de una distribución estable si y sólo si existe una función de variación lenta en infinito L tal que*

$$U(x) = \mathrm{E}\left[X^2 \mathbf{1}_{\{|X| \leq x\}}\right] \sim x^{2-\alpha} L(x), \quad (x \to \infty), \tag{3.20}$$

y, además, para $\alpha \in (0,2)$,

$$\frac{P(X > x)}{P(|X| > x)} \to p \qquad \frac{P(X < -x)}{P(|X| > x)} \to 1 - p \quad (x \to \infty). \tag{3.21}$$

Es posible demostrar que (3.20) es equivalente a

$$\frac{x^2 P(|X| > x)}{U(x)} \to \frac{2-\alpha}{\alpha} \quad (x \to \infty), \text{ para } 0 < \alpha \leq 2, \tag{3.22}$$

$$P(|X| > x) \sim \frac{2-\alpha}{\alpha} \frac{L(x)}{x^\alpha} \quad (x \to \infty), \text{ para } 0 < \alpha < 2, \tag{3.23}$$

Esto nos da la siguiente formulación alternativa del teorema anterior

Teorema 3.15 *Una v.a. X con f.d. F pertenece al dominio de atracción de*
(a) la distribución normal sii U es de variación lenta.
(a) una distribución estable con índice $\alpha \in (0,2)$ sii (3.23) y (3.21) valen.

Como consecuencia de la demostración de los teoremas anteriores se obtiene que $a_n \sim n^{1/\alpha}(L(a-n))^{1/\alpha}$. Poniendo $L^*(n) \sim (L(a_n))^{1/\alpha}$ se tiene que

$$a_n \sim n^{1/\alpha} L^+(n)$$

para una función de variación lenta L^+.

Para $\alpha = 2$ la condición (3.20) dice que U debe ser de variación lenta. Un caso en el cual esto es cierto es cuando la varianza es finita: $(\infty) < \infty$, en cuyo caso $a_n \sim c\sqrt{n}$. El otro caso es cuando la varianza es infinita: $U(\infty) = \infty$. Esto ocurre, por ejemplo, si $U(x) \sim \log x$ o $U(x) \sim \log\log x$ cuando $x \to \infty$. En estos casos la distribución límite es normal pero la normalización será de la forma $a_n \sim \sqrt{n}L(n)$.

Ejemplo 3.2
Sea $X, X_1, X_2, \dots$ v.a.i.i.d. con densidad

$$f(x) = \frac{1}{2x^2}, \quad |x| > 1.$$

La distribución es simétrica, la media es infinita y las colas decrecen como las de la distribución de Cauchy, así que esperaríamos que la distribución pertenezca al dominio de atracción de la distribución estable simétrica con índice $\alpha = 1$.

Integrando obtenemos

$$P(X > x) = \frac{1}{2x}, \quad P(X < -x) = \frac{1}{2|x|}, \quad P(|X| > x = \frac{1}{x}, \quad U(x) = x - 1,$$

de modo que (3.20)-(3.23) se satisfacen con $p = 1/2$ y $L(x) = 1$, de modo que la conjetura era correcta.

Ejemplo 3.3
Sea $X, X_1, X_2, \dots$ v.a.i.i.d. con densidad

$$f(x) = \frac{1}{|x|^3}, \quad |x| > 1.$$

En este caso la varianza es infinita: $\int_1^\infty \frac{x^2}{x^3}\,dx = +\infty$.

Comenzamos por investigar (3.20):

$$U(x) = \int_{|y| \leq x} y^2 f(y)\,dy = 2\int_0^\infty \frac{1}{y}\,dy = 2\log x,$$

de modo que U es de variación lenta en infinito, es decir, X pertenece al dominio de atracción de la distribución normal. Por lo tanto, para constantes normalizadoras apropiadas $a : n$,

$$\frac{S_n}{a_n} \xrightarrow{d} \mathcal{N}(0,1) \quad (n \to \infty).$$

Como la varianza es infinita, la constante de normalización es $a_n \sim \sqrt{n 2 \log(a_n)}$. Con una primera aproximación, $a_n \approx \sqrt{n}$ y obtenemos

$$a_n \sim \sqrt{n 2 \log \sqrt{n}} = \sqrt{n \log n},$$

y tenemos

$$\frac{S_n}{\sqrt{n \log n}} \xrightarrow{d} \mathcal{N}(0,1) \quad (n \to \infty).$$

3.3. Distribuciones Infinitamente Divisibles

En el ejemplo 2.2 del capítulo anterior vimos que la distribución de Poisson se puede obtener como límite de sucesiones de variables binomiales para las cuales la probabilidad de éxito $p = p(n)$ tiende a cero de modo que $np(n) \to \lambda > 0$.

Este resultado se puede considerar como una ley límite para sumas de variables independientes: Si ponemos $S_n = X_{n1}, X_{n2}, \dots, X_{nn}$ donde las variables $X_{ni}, i = 1, \dots, n$ son independientes con distribución de Bernoulli con probabilidad de éxito $p(n)$. La diferencia con los resultados anteriores es que ahora no estamos considerando una sucesión de variables aleatorias sino que para cada n tenemos una colección distinta de variables $X_{ni}, i = 1, \dots, n$, es decir, tenemos un arreglo triangular como los considerados en la sección 3.1.3. Veamos cómo es el comportamiento general de este tipo de estructuras.

Consideramos sistemas triangulares de la forma $X_{ni}, 1 \leq i \leq k_n, n \geq 1, k_n \to \infty$ cuando $n \to \infty$. usualmente tendremos $k_n = n$ pero los resultados son válidos en el caso más general. Suponemos que las variables de una fila son independientes, ponemos $T_n = X_{n1} + \cdots + X_{nn}$ y queremos investigar la convergencia en ley de estas variables.

Vamos a considerar el caso en el cual los elementos de cada fila son idénticamente distribuidas: Para cada n, $X_{n1}, X_{n2}, \dots, X_{nn}$ son i.i.d. y queremos caracterizar las leyes límite.

Definición 3.3 Una variable aleatoria X (o su función de distribución F o su función característica φ) es infinitamente divisible si y sólo si para cada n, X tiene la misma distribución que la suma de n variables independientes e idénticamente distribuidas. En términos de la función característica, $\varphi = (\varphi_n)^n$ donde φ_n es una función característica.

Teorema 3.16 *Una variable X es infinitamente divisible (i.d.) si y sólo si existe un arreglo triangular con $X_{n1}, \dots, X_{nn}$ i.i.d. para cada n, tal que*

$$T_n = \sum_{k=1}^{n} X_{nk} \xrightarrow{d} X.$$

Demostración. Supongamos que tenemos un arreglo triangular con $X_{nk}, 1 \leq k \leq n$ i.i.d. para cada n y $T_n \xrightarrow{d} X$. Fijamos un entero positivo r y ponemos

$$T_{rn} = Z_n^{(1)} + \cdots + Z_n^{(r)},$$

donde

$$\begin{aligned}
Z_n^{(1)} &= X_{rn,1} + \cdots + X_{rn,n}, \\
Z_n^{(2)} &= X_{rn,n+1} + \cdots + X_{rn,2n}, \\
&\vdots \\
Z_n^{(r)} &= X_{rn,(r-1)n+1} + \cdots + X_{rn,rn}.
\end{aligned}$$

Como $T_{rn} \xrightarrow{d} X$ cuando $n \to \infty$, tenemos que $\{T_{rn}, n = 1, 2, \dots\}$ es relativamente compacta, y por el teorema de Prohorov (teorema 1.21) la sucesión es tensa. Pero, usando la independencia tenemos que

$$\begin{aligned}(P(Z_n^{(r)} > z))^r &= P(Z_n^{(1)} > z, \dots, Z_n^{(n)} > z) \\ &\leq P(T_{rn} > rz)\end{aligned}$$

y de manera similar

$$(P(Z_n^{(r)} < -z))^r \leq P(T_{rn} < -rz)$$

Esto muestra que la sucesión $(Z_n^{(1)})_{n\geq 1}$ es tensa y en consecuencia es relativamente compacta. Por lo tanto tenemos una subsucesión $(Z_{n_k}^{(1)})$ que converge en distribución a una variable aleatoria Y. Pero como las $Z_n^{(i)}, i = 1, \dots, r$ son i.i.d., para cualquier i se tiene también que $(Z_{n_k}^{(i)})$ que converge en distribución a la variable aleatoria Y. Por lo tanto $T_{rn} \xrightarrow{d} Y_1 + \cdots + Y_r$, donde $Y_1, \dots, Y_r$ son i.i.d. con $Y_i \stackrel{d}{=} Y$. Pero $T_{rn} \xrightarrow{d} X$ y en consecuencia $X \stackrel{d}{=} Y_1 + \cdots + Y_r$. ■

Es posible demostrar que el resultado es cierto si en lugar de pedir que las variables de cada fila sea i.i.d. se pide independencia y una condición del tipo

$$\max_{1\leq i\leq n} P(|X_{ni}| \geq \varepsilon) \to 0 \quad \text{cuando } n \to \infty \text{ para todo } \varepsilon > 0.$$

Ejemplos 3.4

1. Las leyes estables son infinitamente divisibles. Basta observar que si $X_1 + \cdots + X_n \stackrel{d}{=} a_n X + b_n$ entonces
$$X \stackrel{d}{=} \sum_{n=1}^{n} \Big(\frac{X_i}{a_n} - \frac{b_n}{na_n}\Big)$$

2. Una variable aleatoria con distribución de Poisson es infinitamente divisible. Supongamos que Y tiene distribución de Poisson de parámetro $\lambda > 0$. La función característica de Y es
$$\varphi(t) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{\lambda(e^{it}-1)}$$
y en consecuencia si $Y_i \sim Pois(\lambda_i)$ entonces $Y_1 + \cdots + Y_n \sim Pois(\lambda_1 + \cdots + \lambda_n)$. En particular si $Y \sim Pois(\lambda)$ entonces $Y \stackrel{d}{=} Y_1 + \cdots + Y_n$ donde las Y_i son independientes y tienen distribución $Pois(\lambda/n)$.

3. Si Y es infinitamente divisible, también lo es $aY + b$. Por lo tanto si $Y \sim Pois(\lambda)$, $aY + b, a \neq 0$ es infinitamente divisible y su función característica es $\exp\{ibt + \lambda(e^{iat} - 1)\}$.

4. La distribución gamma es infinitamente divisible. La densidad de esta distribución es
$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}} \quad \text{para } x \geq 0.$$
La función característica de X es
$$\varphi(t) = (1 - i\beta t)^{-\alpha} = [(1 - i\beta t)^{-\alpha/n}]^n,$$
de modo que X es la suma de n variables independientes con distribución gamma de parámetros α/n y β.

Veamos ahora algunas propiedades generales de las distribuciones infinitamente divisibles.

Lema 3.3 *Si φ_1 y φ_2 son f.c.i.d. también lo es $\varphi_1\varphi_2$. Si φ es i.d., su conjugada compleja $\bar{\varphi}$ y $|\varphi|^2$ también son i.d.*

Demostración.
Si $\varphi_i = (\varphi_{ni})^n$ para $i = 1, 2$, entonces $\varphi_1\varphi_2 = (\varphi_{n1}\varphi_{n2})^n$. Pero $\varphi_{n1}\varphi_{n2}$ es la función característica de la suma de dos v.a.i. con f.c. φ_{n1} y φ_{n2}, de modo que $\varphi_{n1}\varphi_{n2}$ es i.d.

Por otro lado si X tiene f.c. φ, $-X$ tiene f.c. $\bar{\varphi}$. Por lo tanto si $\varphi = (\varphi_n)^n$ entonces $\bar{\varphi} = (\bar{\varphi}_n)^n$ también es i.d. Finalmente, $|\varphi|^2 = \varphi\bar{\varphi}$, también es i.d. ■

Teorema 3.17 *Si φ_n para $n \geq 1$ es una sucesión de f.c.i.d. y $\varphi_n \to \varphi$ cuando $n \to \infty$ entonces φ es i.d.*

Demostración. Sea Z_n una v.a. con f.c. φ_n para $n \geq 1$. Para todo entero positivo r fijo tenemos que $Z_n \stackrel{d}{=} Z_n^{(1)} + \cdots + Z_n^{(r)}$ donde las $Z_n^{(i)}$, $i = 1, \ldots, r$ son i.i.d. Si Z es una v.a. con f.c. φ entonces $Z_n \stackrel{d}{\to} Z$ por el teorema de continuidad. En consecuencia (Z_n) es relativamente compacta y por lo tanto tensa. Al igual que en la demostración del teorema 3.16 vemos que $(Z_n^{(1)})$ es tensa. Por lo tanto existe una subsucesión $(Z_{n_i}^{(1)})$ que converge en distribución a una variable aleatoria Y_1. Al igual que en el teorema 3.16 concluimos que $Z_n \stackrel{d}{\to} Y_1 + \cdots + Y_r$, donde las Y_i son i.i.d. con $Y_i \stackrel{d}{=} Y_1$. Pero también tenemos que $Z_n \stackrel{d}{\to} Z$. Por lo tanto $Z \stackrel{d}{=} Y_1 + \cdots + Y_r$ y φ es i.d. ■

Supongamos que φ es una f.c.i.d. ¿Es posible representar a φ de dos maneras distintas como la n-ésima potencia de funciones características diferentes? Supongamos que φ_n y ψ_n son dos funciones características tales que $\varphi_n^n = \psi_n^n$. Como ambas son funciones características tenemos que $\varphi_n(0) = 1 = \psi_n(0)$. Supongamos que φ_n y ψ_n nunca se anulan en $\mathbb{R}$ y veamos que bajo esta condición $\varphi_n = \psi_n$.

Como $(\varphi_n/\psi_n)^n = 1$, φ_n/ψ_n es una función continua de $\mathbb{R}$ a $\{\exp\{i2\pi k/n\}, k = 0, 1, \ldots, n-1\}$. Como $\mathbb{R}$ es un conjunto conexo y su imagen bajo una función continua también lo es, la imagen de $\mathbb{R}$ bajo φ_n/ψ_n debe ser constante, y como φ_n y ψ_n tienen un valor común, esa constante debe ser 1. Por lo tanto la representación es única, siempre y cuando podamos demostrar que una f.c.i.d. nunca se anula.

Teorema 3.18 *Una f.c.i.d., nunca se anula.*

Demostración.
Sean φ y φ_n f.c. tales que $\varphi = \varphi_n^n$. Por el lema (3.3) $f = |\varphi|^2$ y $f_n = |\varphi_n|^2$ son también f.c. y son reales positivas. Para cada $t \in \mathbb{R}$, $f(t)$ es real y positiva y por lo tanto su n-ésima raíz real positiva es única y la denotamos por $(f(t))^{1/n}$. Como tenemos que $f(t) = (f_n(t))^n$ y $f_n(t)$ es positivo, tenemos que, para todo t,

$$f_n(t) = (f(t))^{1/n}.$$

Pero sabemos que $0 \leq f(t) \leq 1$ y por lo tanto $\lim_{n\to\infty}(f(t))^{1/n}$ vale 0 ó 1, según $f(t) = 0$ ó $f(t) \neq 0$. Por lo tanto $\lim_{n\to\infty} f_n(t)$ existe para todo t y la función límite, que llamaremos $g(t)$, puede tomar a lo sumo dos valores, 0 y 1. Por otro lado, como f es continua en $t = 0$ con $f(0) = 1$, existe un $t_0 > 0$ tal que $f(t)$ no se anula para $|t| \leq t_0$, y en consecuencia $g(t) = 1$ para $|t| \leq t_0$. Por lo tanto la sucesión de f.c. f_n converge a la función g, que hemos visto que es continua en 0. Por el teorema de continuidad g debe ser una f.c. y por lo tanto es una función continua. En consecuencia g vale 1 siempre y para todo t

$$|\varphi(t)|^2 = f(t) \neq 0$$

■

El siguiente resultado, que se conoce como la representación de Lévy-Khinchin, caracteriza la clase de las distribuciones infinitamente divisibles.

Teorema 3.19 *La variable aleatoria X tiene distribución infinitamente divisible si y sólo si*

$$\varphi_X(t) = \exp\{iat + \int_{\infty}^{\infty} \Big(e^{itx} - 1 - \frac{itx}{1+x^2}\Big)\frac{1+x^2}{x^2} dH(x)\} \tag{3.24}$$

donde $a \in \mathbb{R}$ y H es una medida finita. El integrando se define en $x = 0$ por continuidad y su valor es $-t^2/2$.

Para presentar las ideas básicas de la prueba vamos a considerar un caso más simple, suponiendo que la variables tienen varianza finita. Haremos la demostración en dos partes.

Teorema 3.20 *Supongamos que la variable aleatoria X tiene segundo momento finito. Si X tiene distribución infinitamente divisible entonces su función característica es*

$$\varphi_X(t) = \exp\{i\mu t + \int_{\infty}^{\infty} \Big(e^{itx} - 1 - itx\Big)\frac{1}{x^2} dG(x)\} \tag{3.25}$$

donde $\mu \in \mathbb{R}$ y G es una medida finita. El integrando se define en $x = 0$ por continuidad y su valor es $-t^2/2$.

Antes de ver la demostración veamos tres casos particulares

Ejemplo 3.5
Si G tiene solo una masa σ^2 en el origen, entonces (3.25) vale

$$\exp\{i\mu t - \sigma^2 t^2/2\}$$

que es la f.c. de una distribución normal con media μ y varianza σ^2.

Ejemplo 3.6
Si G tiene masa puntual λ en 1 y $\mu = \lambda$, entonces (3.25) vale

$$\exp\{\lambda(e^{it} - 1)\}$$

que es la f.c. de una distribución de Poisson de parámetro λ.

Supongamos que $\mu = 0$ y G tiene masa λx^2 en $x \neq 0$. Entonces (3.25) vale

$$\exp\{\lambda(e^{itx} - 1 - itx)\}$$

que es la función característica de $x(Z_\lambda - \lambda)$ con $Z_\lambda \sim Pois(\lambda)$. Por lo tanto (3.25) es la f.c. de una distribución F que es infinitamente divisible: tomamos como F_n la f.d. de $x(Z_{\lambda/n} - \lambda/n)$.

Ejemplo 3.7
Si $\varphi_j(t)$ la función correspondiente a la fórmula (3.25) con $\mu = 0$ y medida G_j y si $G = \sum_{j=1}^k G_j$ entonces (3.25) corresponde a $\varphi_1(t) \cdots \varphi_k(t)$. Por los dos ejemplos anteriores vemos que si G consiste de una cantidad de masas puntuales finitas como las de los ejemplos anteriores, (3.25) es una f.c. Es fácil verificar en el ejemplo anterior que la distribución correspondiente a $\varphi(t)$ tiene media 0 y varianza $G(\mathbb{R})$, y como medias y varianzas (en el caso independiente) se suman, lo mismo es cierto en este ejemplo.

Ante de hacer la demostración del teorema recordemos algunas cosas sobre logaritmos de números complejos. Para un número complejo $z = re^{i\theta}$, donde $r = |z|$ y $\theta = \arg(z)$ el logaritmo se define como $\log(z) = \mathrm{Log}(r) + i\theta$, donde Log denota el logaritmo de un número real. El logaritmo es una función multivaluada para todos los $z \in \mathbb{C}$ salvo $z = 0$.

Si ϑ denota el valor principal de $\arg(z)$, es decir, es el valor que satisface $-\pi < \vartheta \leq \pi$, cualquier otro valor del argumento se puede escribir como $\theta = \vartheta + 2\pi n$ para $n \in \mathbb{Z}$. En consecuencia

$$\log(z) = \mathrm{Log}(r) + i(\vartheta + 2\pi n) \quad n \in \mathbb{Z}$$

El *valor principal* de $\log(z)$ es el valor que se obtiene de la fórmula anterior con $n = 0$, y es el valor que usaremos en las demostraciones que vienen cuando tomemos el logaritmo de una función característica.

El logaritmo tiene el siguiente desarrollo en series de potencias para $|z| < 1$:

$$\log(1+z) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{z^n}{n}$$

y en particular, $\log(1-z) = -z + o(1)$ cuando $|z| \to 0$.

Hemos visto que si $\varphi(t)$ es una f.c.i.d., nunca se anula y es continua, y por lo tanto tiene un logaritmo continuo. Si pedimos que $\log\varphi(0) = \log 1 = 0$, el logaritmo queda determinado de manera única. Además, si $\varphi = \varphi_n^n$, donde φ_n es una f.c., entonces

$$\varphi_n^n = \big(\exp\big(\frac{1}{n}\log\varphi\big)\big)^n$$

y por lo tanto

$$\varphi_n(t) = \exp\big(\frac{1}{n}\log\varphi(t)\big) \tag{3.26}$$

Demostración del Teorema 3.20. Si X es i.d. para cada n existen v.a.i. $(X_{n,k}, k = 1, \dots, n)$ con f.d. F_n y f.c. φ_n tales que $\varphi_X(t) = (\varphi_n(t))^n$. Además $\varphi_X \neq 0$ de modo que los logaritmos están bien definidos. Tenemos

$$\log\varphi_X(t) = n\log\varphi_n(t) = n\log(1-(1-\varphi_n(t))).$$

Pero para cualquier T, a partir de la ecuación (3.26) vemos que $\varphi_n(t) \to 1$ cuando $n \to \infty$, uniformemente en el intervalo $|t| < T$. Por lo tanto tenemos

$$\begin{aligned}
\log\varphi_X(t) &= n(\varphi_n(t) - 1) + o(1) = n\int_{-\infty}^{\infty}(e^{itx} - 1)dF_n(x) + o(1)\\
&= it\,\mathrm{E}(X) + \int_{-\infty}^{\infty}(e^{itx} - 1 - itx)\frac{1}{x^2}nx^2 dF_n(x) + o(1)\\
&= it\,\mathrm{E}(X) + \int_{-\infty}^{\infty}(e^{itx} - 1 - itx)\frac{1}{x^2}dG_n(x) + o(1)
\end{aligned}$$

cuando $n \to \infty$ donde

$$G_n(x) = n\int_{-\infty}^{x} t^2\, dF_n(t)$$

Esta G_n es una f.d. generalizada, es decir, es no-decreciente y continua por la derecha con $G_n(-\infty) = 0$ y $G_n(\infty) = n\,\mathrm{E}(X_{n,1}^2) = \mathrm{E}(X^2)$.

Por el teorema de selección de Helly (T. 1.19) existe una subsucesión $(G_{n_k}, k \geq 1)$ que converge vagamente a G, de modo que para cualquier intervalo acotado $[a, b]$, con $a < 0 < b \in \mathcal{C}(G)$, usando el teorema 1.20 se tiene que

$$\int_a^b (e^{itx} - 1 - itx)\frac{1}{x^2}dG_{n_k}(x) \to \int_a^b (e^{itx} - 1 - itx)\frac{1}{x^2}dG(x), \quad k \to \infty.$$

Más aún, teniendo en cuenta que

$$\Big|\frac{e^{itx} - 1 - itx}{x^2}\Big| \leq \frac{2|t|}{|x|} \to 0 \quad (|x| \to \infty)$$

vemos que el integrado es una función que se anula en ∞ y por la caracterización de la convergencia vaga tenemos que

$$\log \varphi_X(t) = it\,\mathrm{E}(X) + \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)\frac{1}{x^2} dG(x).$$

Esto demuestra el teorema e identifica a μ como $\mathrm{E}(X)$. ■

Teorema 3.21 *Si*

$$\varphi(t) = \exp\Big\{ \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)\frac{1}{x^2}\, dG(x)\Big\} \tag{3.27}$$

donde G es una medida finita, entonces φ es una f.c.i.d.

Demostración.

Recordemos que si G tiene masa puntual $\lambda > 0$ en 1 y $\mu = 1$ entonces X tiene distribución de Poisson de parámetro λ. Si en cambio G tiene masa λx^2 en x y $\mu = 0$, entonces

$$\varphi(t) = \exp\{\lambda(e^{itx} - 1 - itx)\}$$

que es la f.c. de una transformación lineal de una distribución de Poisson: Si $Y \sim Pois(\lambda)$ entonces $\varphi(t)$ es la función característica de $\lambda(Y - x)$.

Definimos la medida $G_{n,k}$ como la medida discreta con masa $\alpha_{n,k} = G(k2^{-n}, (k+1)2^{-n}]$ colocada en $k2^{-n}$ para $k = 0, \pm1, \pm2, \dots, \pm2^{2n}, n \geq 1$. Entonces

$$\begin{aligned}\varphi_{n,k}(t) &= \exp\Big\{ \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)\frac{1}{x^2}\, dG_{n,k}(x)\Big\}\\ &= \exp\Big\{(e^{itk2^{-n}} - 1 - itk2^{-n})\frac{\alpha_{n,k}}{(k2^{-n})^2}\, dG_{n,k}(x)\Big\}\end{aligned}$$

que es la f.c. de

$$\frac{\alpha_{n,k}}{(k2^{-n})^2}(Y - k2^{-n})$$

donde $Y \sim Pois(\alpha_{n,k}/k2^{-n})$.

Sea $Y_{n,k}, k = 0, \pm1, \pm2, \dots, \pm2^{2n}, n \geq 1$ v.a.i. con f.c. $\varphi_{n,k}$ y sea $Y_n = \sum_k Y_{n,k}$, con f.c. φ_n con representación (3.27) con medida G_n que satisface $G_n = \sum_k G_{n,k}$. Las variables $Y_{n,k}$ son i.d. por ser Poisson generalizadas y por lo tanto, también lo es Y_n. Es fácil ver que $G_n \xrightarrow{v} G$ cuando $n \to \infty$ y además $G_n(\mathbb{R}) \leq G(\mathbb{R}) < \infty$. De nuevo observamos que el integrando $(e^{itx} - 1 - itx)/x^2 \in \mathcal{C}_0$ y por lo tanto $\varphi_n(t) \to \varphi(t)$ cuando $n \to \infty$. Si demostramos que $\varphi(t)$ es continua en 0, por el teorema de continuidad φ sería una f.c. y como es el límite de f.c.i.d., también sería i.d.

Para ver que φ es continua en 0 queremos ver que $\varphi(h) \to 1$ cuando $h \to 0$ y esto equivale a ver que

$$\int_{-\infty}^{\infty} (e^{ihx} - 1 - ihx)\frac{1}{x^2} dG(x) \to 0$$

cuando $h \to 0$.

$$\begin{aligned}\Big|\int_{-\infty}^{\infty} (e^{ihx} - 1 - ihx)\frac{1}{x^2} dG(x)\Big| &\leq \int_{-\infty}^{\infty} \Big|(e^{ihx} - 1 - ihx)\frac{1}{x^2}\Big| dG(x)\\ &\leq \int_{-\infty}^{\infty} \frac{h^2x^2}{x^2} dG(x) = h^2 G(\mathbb{R})\end{aligned}$$

y esto tiende a 0 cuando $h \to 0$. Por lo tanto φ es una f.c.i.d. ■

Proposición 3.3 *Si φ es una f.c. de la forma (3.27), la distribución correspondiente tiene media 0 y varianza $G(\mathbb{R})$.*

Demostración. Consideremos las variables $Y_n = \sum_k Y_{n,k}$ definidas en la demostración del teorema anterior. Por lo cometarios del ejemplo 3.7 $\mathrm{Var}(Y_{n,k}) = G_{n,k}(\mathbb{R})$ y $\mathrm{Var}(Y_n) = G_n(\mathbb{R}) = \sum_k G_{n,k}(\mathbb{R})$, que satisfacen $G_n(\mathbb{R}) \le G(\mathbb{R}) < \infty$. Por el teorema 1.12

$$\mathrm{E}(Y^2) \le \liminf_{k\to\infty} \mathrm{E}(Y_n^2) \le \sup_n G_n(\mathbb{R}) \le G(\mathbb{R}) < \infty,$$

de modo que la distribución límite tiene segundo momento finito. Usando la fórmula (2.9) para la función característica (3.27) obtenemos el resultado. ■

Corolario 3.2 *Si Y es una v.a. con f.c. (3.27) entonces $Y + \mu$ tiene f.c. (3.25). Recíprocamente, si Z tiene f.c. (3.25) entonces existe una v.a. i.d. Y tal que $Z = Y + \mu$ y por lo tanto Z es i.d.*

Sabemos que los límites de f.c.i.d. también son i.d. El siguiente teorema nos da condiciones para la convergencia.

Teorema 3.22 *Sean $Y_n, n \ge$ v.a. con varianzas finitas que corresponden a distribuciones infinitamente divisibles con f.c.*

$$\varphi_n(t) = \exp\Big\{ i\mu_n t + \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)\frac{1}{x^2} dG_n(x) \Big\}.$$

Entonces

$$Y_n \xrightarrow{d} Y \quad y \quad \mathrm{Var}(Y_n) \to \mathrm{Var}(Y) \quad (n \to \infty),$$

donde

$$\varphi_Y(t) = \exp\Big\{ i\mu t + \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)\frac{1}{x^2} dG(x) \Big\}.$$

si y sólo si

$$G_n \xrightarrow{v} G, \quad G_n(\mathbb{R}) \to G(\mathbb{R}), \quad \mu_n \to \mu$$

cuando $n \to \infty$. Las dos primeras condiciones equivaldrían a convergencia en distribución si la masa total fuese 1.