Revista IECOS, Journal of Economics, Applied Statistics and Social Science
Vol. 24, N° 2, diciembre de 2023, pp. 79-100
DOI: https://doi.org/10.21754/iecos.v24i2.2005
ISSN: 2788-7480 (En Línea)
ISSN: 2961-2845 (Impresa)

Convergencia de procesos aleatorios unidimensionales

Convergence of one-dimensional random processes

Roberto Vila
Universidade de Brasília, Brasília, Brasil

Recibido: 09/12/2023
Aceptado: 20/12/2023
Publicado: 31/12/2023

Departamento de Estadística, Universidade de Brasília, Brasília, Brasil
Correspondencia: rovig161@gmail.com

https://orcid.org/0000-0003-1073-0114

Revista de la Facultad de Ingeniería Económica, Ingeniería Estadística y Ciencias Sociales de la Universidad Nacional de Ingeniería

RESUMEN

En este trabajo desenvolvemos extensivamente algunos de los resultados obtenidos en la referencia (Cioletti et al., 2017). Usamos la distancia de Wasserstein para obtener algunos teoremas del tipo límite central para procesos aleatorios unidimensionales que tienen dependencia asociada positiva.

Palabras claves: *Distancia de Wasserstein; Proceso aleatorio, Asociado positivo.*

ABSTRACT

In this paper we extensively develop some of the results obtained in reference (Cioletti et al., 2017). We use the Wasserstein distance to obtain some central limit type theorems for one-dimensional random processes having positive associated dependence.

Keywords: *Wasserstein distance; Random process, Positive associate.*

1. INTRODUCCIÓN

La distancia de Wasserstein $W_r(\mu, v)$, también conocida como Monge-Kantorovich-Rubinstein (Kantorovič & Rubinšteĭn, 1958; Jordan et al., 1998), distancia de Mallows (1972) o distancia de transporte óptima en optimización (Ambrosio, 2003), es responsable de medir la discrepancia entre dos medidas de probabilidades μ y v. Esta métrica ha sido aplicada con éxito en una amplia variedad de campos, por ejemplo, Gray (2009), Rachev y Rüschendorf (1998), Sommerfeld y Munk (2018), Otto (2001) y Villani (2003, 2009). Sobre los números reales, esta distancia cuantifica la discrepancia entre dos funciones de distribución (acumulativas) F y G. Si F y G son las funciones de distribución de dos variables aleatorias X e Y, respectivamente. El Teorema de representación de Dorea y Ferreira (2012) nos permite escribir, $W_r^r(F, G) = \mathbb{E}_H\,|X^* - Y^*|^r$, si $r \geq 1$, donde $H(x, y) = F(x) \wedge G(y)$ es el límite superior de Frèchet. Además, puede ser demostrado que la representación $W_r^r(F, G) = \mathbb{E}_K\,|X' - Y'|^r$, donde $K(x, y) = 0\ \vee [F(x) + G(y) - 1]$ es el límite inferior de Frèchet, es válida en el caso que $0 < r < 1$. Usando esas representaciones y el conocido Teorema de Bickel y Freedman (1981),

$$(1) \qquad W_r(F_n, F) \underset{n}{\rightarrow} 0 \iff F_n \overset{d}{\rightarrow} F; \quad y; \quad \int |x|^r\, dF_n(x) \underset{n}{\rightarrow} \int |x|^r\, dF(x),$$

el cual proporciona una estrecha relación con la convergencia en distribución $\left(\overset{d}{\rightarrow}\right)$, en este trabajo haremos uso de la distancia de Wasserstein (Vaserstein, 1969) para analizar el comportamiento asintótico de procesos aleatorios unidimensionales que tienen dependencia asociada positiva. Ejemplos de procesos aleatorios que exhiben este tipo de comportamiento, por medio de W_r, usualmente son encontrados en la Mecánica Estadística, por ejemplo, en modelos ferromagnéticos tipo Ising con espines discretos y continuos, para mayores detalles, vea referencia (Cioletti et al., 2017).

2. LA DISTANCIA DE WASSERTEIN

En esta sección presentamos algunos conceptos de asociación positiva y de distancia de Wasserstein. En seguida, enunciamos algunos resultados preliminares que utilizaremos a lo largo de la exposición de este trabajo. Así mismo, cerramos esta sección con algunas definiciones adicionales.

2.1. ASOCIACIÓN POSITIVA

Denote por $\mathbb{Z}$ al conjunto de los números enteros. Consideraremos procesos aleatorios del siguiente tipo $\{X_j : j \in \mathbb{Z}\}$, los cuales son definidos sobre algún espacio de probabilidad $(\Omega, \mathfrak{F}, \mathbb{P})$, y están asociados positivamente de acuerdo a la siguiente definición.

Definición 1. *Un proceso aleatorio* $\{X_j : j \in \mathbb{Z}\}$, *es asociado positiva si, dadas dos funciones coordenadas no decrecientes* $f, g : \mathbb{R}^n \to \mathbb{R}$ *y* $j_1, \ldots, j_n \in \mathbb{Z}$, *tenemos*

$$cov\left(f\left(X_{j1}, \ldots, X_{jn}\right), g\left(X_{j1}, \ldots, X_{jn}\right)\right) \geq 0,$$

(Gabriel, 2017)

siempre que la covarianza exista.

Decimos que una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es no decreciente si $f(x_1, \ldots, x_n) \leq f(y_1, \ldots, y_n)$, siempre que $x_j \leq y_j$, para todo $j = 1, \ldots, n$. Algunos ejemplos de procesos asociados positivamente son los siguientes:

Ejemplo 2. *Cualquier conjunto de variables aleatorios independientes está asociado positivamente* (Esary et al., 1967).

Ejemplo 3. *Variables aleatorias con distribución Gaussiana multivariada y con covarianza positiva están asociadas positivamente* (Pitt, 1982).

Ejemplo 4. *Sean* $\{X_i' : i \geq 1\}$ *independiente e idénticamente distribuidos y sea Y independiente de* $\{X_i' : i \geq 1\}$, *entonces,* $\{X_i = X_i' + Y : i \geq 1\}$ *es asociado positivo* (Barlow & Proschan, 1975).

Lema 5. *Sea* $\{X_j : i \in \mathbb{Z}\}$ *un proceso aleatorios asociado positiva; Para* $m_j \geq 1$, *si* $f_j : \mathbb{R}^{m_j} \to \mathbb{R}$ *son funciones coordenadas no decrecientes, entonces* $\{ f_j\left(X_{i1}, \ldots, X_{i_{m_j}}\right) : i_1, \ldots, i_{m_j} \in \mathbb{Z} \}$, *también, es asociado positivo* (Oliveira, 2012).

Ahora, con el Lema 5 a nuestra disposición, es sencillo generar nuevas familias de variables aleatorias asociadas positivamente a partir de un conjunto de variables aleatorias con esta propiedad, al aplicar transformaciones monótonas.

Ejemplo 6. Si X_n , $n \in \mathbb{N}$, son variables aleatorios asociados positivas, entonces, la secuencia de sumas parciales $S_n = X_1 + \cdots + X_n$, $n \in \mathbb{N}$ esta asociada positivamente. Esto es una consecuencia inmediata del Lema 5.

Ejemplo 7. Dadas las variables aleatorias $X_1, \dots, X_n$, definan las estadísticas ordenadas $X_{k:n} =$ "el k-ésimo más pequeño entre $X_1, \dots, X_n$". Estas estadísticas de orden son transformaciones no decrecientes de $X_1, \dots, X_n$, consecuentemente, estas estadísticas de orden están asociadas positivamente, lo mismo se aplica a $X_1, \dots, X_{n:n}$.

Ejemplo 8. Dada una secuencia de variables aleatorias Y_n, con $n \in \mathbb{N}$ fijo, defina $X_n = máx\ \{Y_n,\ Y_{n+1}, \dots, Y_{n+m}\}$, con $m \in \mathbb{N}$ fijo. Si los Y_n están asociadas, también lo están los X_n.

2.2. DISTANCIA DE WASSERSTEIN

En esta parte definimos el concepto de distancia de Wasserstein (Mallows, 1972; Newman, 1980) y establecemos una equivalencia con la definición de distancia Mallows que aparecen en las referencias (Bickel & Freedman, 1981; Dorea & Ferreira, 2012; Mallows, 1972) (ver Lema 10).

Sea $(\Omega, \mathfrak{F})$ un espacio medible correspondiente a un experimento aleatorio dado. Denotando por $\mathfrak{B}(\Omega)$ a la σ-álgebra de los borelianos de Ω, definimos la colecci´on de todas las medidas de probabilidad sobre $\mathfrak{B}(\Omega)$ por $\mathcal{M}_1(\Omega)$.

Supongamos que estamos encargados del "transporte de mercancías" entre productores e consumidores, cuyas distribuciones espaciales son modeladas por las medidas de probabilidad μ y ν. Si los productores y consumidores estan localizados a una distancia mayor, más difícil será nuestro trabajo. Luego, nos gustaría resumir el "grado de dificultad" con apenas una cantidad. Para ello, es natural considerar el "costo óptimo de transporte" entre las medidas μ y ν como

$$C(\mu, \nu) = \underset{\pi \in \Pi(\mu,\nu)}{ínf} \int_{\mathbb{R}^2} c(x, y)\, d\pi(x, y),$$

(Gabriel, 2017)

donde $c(x, y)$ denota el costo de transporte de una unidad de masa de x para y y el conjunto

(2) $$\Pi(\mu,\nu) = \{\pi \in \mathcal{M}_1(\mathbb{R}^2) : Proy_1\, \pi = \mu \quad y \quad Proy_2\, \pi = \nu\}$$

(Gabriel, 2017)

está constituido por todos los acoplamientos π de $\mu, \nu \in \mathcal{M}_1\,(\mathbb{R})$, conocidos como planos de transporte.
Aquí, $(Proy_1\, \pi)(A\, x\, \mathbb{R})$ y $(Proy_2 \pi)(B) = \pi\,(\mathbb{R}\, x\, B)$, para cada borelianos A y B en $\mathbb{R}$, son las proyecciones sobre la medida π. En terminos simples, $\Pi(\mu,\nu)$ es el conjunto de todas las medidas de probabilidad $\pi \in \mathcal{M}_1\,(\mathbb{R}^2)$ con marginales μ y ν, respectivamente. En general *C* no es una distancia. En el caso que *c* sea una distancia, entonces *C* es una distancia (métrica), también.

Definición 9 (Distancia de Wasserstein). *Sea (*$\mathbb{R}, d$*) un espacio m´etrico, con métrica dada por d :* $\mathbb{R}^2$ *→ [0, ∞). La distancia de Wasserstein de orden r > 0 entre dos medidas de probabilidad* $\mu, \nu \in \mathcal{M}_1\,(\mathbb{R})$ *se define mediante la siguiente fórmula* (Mallows, 1972; Villani, 2009):

$$W_r(\mu,\nu) = \left\{ \begin{matrix} ínf \\ \pi \in \Pi(\mu,\nu) \end{matrix} \int_{\mathbb{R}^2} d(x,y)^r \, d\pi(x,y), \right\}^{\frac{1}{r}},$$

donde $\Pi(\mu,\nu)$ *es el conjunto definido en* (2).

Algunos casos particulares de la distancia de Wasserstein son conocidos, por ejemplo (Gabriel, 2017):

- $\mathcal{W}_1(\mu,\nu) = \sup_{||\psi||_{Lip} \leq 1} \{\int_{\mathbb{R}} \psi d(\mu - \nu)\}$, esa expresión es llamada "fórmula de dualidad para la distancia de Kantorovich-Rubinstein", para mayores detalles, ver Villani (2003). Aquí el supremo es tomado sobre todas las funciones Lipschitzianas (limitadas) ψ que están dentro de la bola unitaria, según la norma $||\psi||_{Lip} = máx\,\{k_1\,(\psi), k_2\,(\psi)\}$, donde $k_1(\psi) = \frac{sup_{x \neq y} |\psi(x) - \psi(y)|}{|x - y|}$ y $k_2(\psi) = 2\, sup_x |\psi(x)|$.

- $\mathcal{W}_r(\delta_x\,, \delta_y) = d(x,y), r < 0$, donde δ_x y δ_y son medidas de delta de Dirac concentradas en los puntos fijos x e y, respectivamente.

- Si d es una métrica discreta, es decir $d(x,y) = \mathbb{1}_{\{x \neq y\}}$, entonces $\mathcal{W}_r(\mu,\nu) = \left(\frac{1}{2}\right) ||\mu - \nu||_{VT}$ ver [21]), donde $||\mu - \nu||_{VT} =$

$2sup_{A\in\mathfrak{F}} |\mu(A) - \nu(A)|$ denota la distancia de variación total entre μ y ν.

El siguiente resultado nos ofrece una caracterización de la distancia de Wasserstein en el caso que $\mathbb{R}$ está equipado con la métrica euclidiana. La prueba de este resultado puede ser encontrado en el Apéndice A. En algunas referencias, vea por ejemplo (Bickel & Freedman, 1981; Dorea & Ferreira, 2012; Mallows, 1972), esta medida es conocido como distancia Malllows.

Lema 10. *En la Definición 9, considere $d(x,y) = |x - y|, \mu = F$ y $\nu = G$ dos funciones de distribución (acumulativa). Entonces, la distancia de Wasserstein de orden $r > 0$ entre F y G es dada por*

$$W_r(F,G) = \left\{ \underset{\pi \in \Pi(\mu,\nu)}{ínf} \mathbb{E}|X - Y|^r \right\}^{\frac{1}{r}},$$

donde el ínfimo es tomado sobre todos los pares de variables aleatorios (X,Y) cuyas distribuciones marginales son F y G, respectivamente.

Tenga en cuenta que, estrictamente hablando, W_r , como definido anteriormente, no es una distancia sobre el espacio de las funciones de distribución, ya que esta definición admite la posibilidad $W_r(F,G) = \infty$. Pero esto no crea ningún inconveniente, para que esta definición tenga sentido las distribuciones F y G deben tener un momento (absoluto) de orden *r* finito. Formalmente, definimos el espacio de distribuciones que tienen esta propiedad por

$$\mathcal{P}_r = \left\{ F : \mathbb{R} \to [0\,,1] : \begin{array}{c} F \; es \; una \; función \; de \; distribución \; tal \; que \\ \int_{\mathbb{R}} |x|^r \, dF(x) < \infty \end{array} \right\}. \quad (3)$$

Este espacio fue introducido por Bickel y Freedman (1981) para mostrar que, para $r \geq 1$ la función $W_r : \mathcal{P}_r \, x \, \mathcal{P} \to [0,\infty)$ en el Lema 10 es una métrica.

Observación 11. *De aquí en adelante, en este trabajo, usaremos la definición de la distanción de Wassertein providenciada por el Lema 10. Esto es, consideraremos la Definición 9 consideramos la métrica euclidiana $d(x,y) = |x - y|$.*

2.3. RESULTADOS Y DEFINICIÓN PRELIMINARES

A continuación recopilamos algunos resultados, propiedades y definiciones necesarias para las pruebas de este trabajo.

Lema 12 (Newman & Wright, 1981). *Sea* $\{X_j : j \in \mathbb{Z}\}$ *un proceso aleatorio asociado positivo; Si todos los* $X_j's$ *poseen un segundo momento finitom entonces las funciones características* $\phi_j(r_j) = \mathbb{E}(exp\{ir_jX_j\})$ *y*

$$\phi(r_1, \dots r_n) = \mathbb{E}(\exp\left\{i \sum_{j=1}^{n} r_j\, X_j\right\}\}$$

satisfacen

$$\left|\phi\left((r_1, \dots r_n) - \prod_{j=1}^{n} \phi_j(r_j)\right)\right| \leq \frac{1}{2} \sum_{1 \leq j \neq k \leq n} |r_j r_k| cov(X_j, X_k).$$

El lema 12 nos informa que, para procesos aleatorios asociados positivos cuyas combinaciones lineales de las covarianzas poseen un determinado decaimiento a medida que n crece, el proceso puede ser considerado asintóticamente independiente.

Asuma que X, Y y (X, Y) tienen distribuciones F, G Y H, respectivamente, donde $\forall(x, y) \in \mathbb{R}^2$,

$$H(x, y) = F(x) \wedge G(y) = mín\{F(x), G(y)\}.$$

El siguiente resultado (Teorema 13) facilita la evaluación de $W_r(F, G)$, pues en este caso la distancia de Wassersteins $W_r(F, G)$, con $r \geq 1$, es alcanzado por el *r*-ésimo momento de $|X - Y|$ con respecto a la distribución H.

Teorema 13 (Dorea & Ferreira, 2012). *Para* $r \geq 1$ *la distancia de Wasserstein del Lema 10 puede ser escrita como*

$$\begin{aligned} W_r^r(F, G) &= \mathbb{E}|F^{-1}(U) - G^{-1}(U)|^r \\ &= \int_0^1 |F^{-1}(u) - G^{-1}(u)|^r du \\ &= \mathbb{E}_H|X - Y|^r = \int_{\mathbb{R}^2} |x - y|^r dH(x, y), \end{aligned}$$

donde U es uniformemente distribuida sobre el intervalo (0,1) y

$$F^{-1}(u) = ínf\{x \in \mathbb{R} : F(x) \geq u\}, \quad 0 < u < 1,$$

(Gabriel, 2017)

denota la inversa generalizada.

La representación de la distancia de Wasserstein del Teorema 13 no es valida para $0 < r < 1$, incluso cuando el momento de orden *r* es finito, como se muestra en el siguiente ejemplo:

Ejemplo 14. *Sean* $X \in \{1,2\}$ *e* $Y \in \{3,4\}$ *dos variables aleatorias discretas con respectivas funciones de probabilidades dadas por*

$$\mathbb{P}(X = 1) = \frac{2}{3} = 1 - \mathbb{P}(X = 2), \mathbb{P}(Y = 3) = \frac{3}{4} = 1 - \mathbb{P}(Y = 4).$$

Denota por F *y* G *a las funciones de distribución de* X *e* Y *, respectivamente.*

Supongamos que $(X,Y) \stackrel{d}{=} 0 \vee [F + G - 1]$ *y que* $(X^*, Y^*) \stackrel{d}{=} F \wedge G$*, es decir, las funciones de distribución de* (X,Y) *y* (X^*, Y^*)*, respectivamente, son definidas como,* $\forall (x, y) \in \mathbb{R}^2$,

$$F_{X,Y}(x,y) = máx\,\{0, F(x) + G(y) - 1\},$$
$$F_{X^*,Y^*}(x,y) = mín\{F(x), G(y)\}.$$

Vemos las distribuciones de probabilidades de (X,Y) *y* (X^*,Y^*) *son explícitamente dadas en la Tabla 1:*

Tabla 1

Distribuciones de probabilidades de (X,Y) *y* (X^*,Y^*).

$X \backslash Y$	3	4
1	5/12	1/4
2	1/3	0

$X^* \backslash Y^*$	3	4
1	2/3	0
2	1/12	1/4

Observe que

$$\mathbb{E}|X^* - Y^*|^r - \mathbb{E}|X - Y|^r = \frac{2^{r+1} - 3^r - 1}{4}.$$

Una vez que la función $r \to \frac{2^{r+1}-3^r-1}{4}$ *es positiva,* $\forall\ 0 < r < 1$*, tenemos que* $\mathbb{E}|X^* - Y^*|^r - \mathbb{E}|X - Y|^r > 0$.

Por lo tanto, para cualquier $0 < r < 1$ *no es posible obtener una representación de la distancia de Wasserstein como la del Teorema 13.*

Observación 15. *Una generalización del Ejemplo 14 puede ser encontrada en la referencia* (Dorea & Ferreira, 2012).

Ahora, asuma que, X, Y y (X, Y) tienen distribuciones $F, G\ y\ K$, respectivamente, donde, $\forall(x, y) \in \mathbb{R}^2$,

$$K(x, y) = 0 \vee [F(x) + G(y) - 1] = máx\{0, F(x) + G(y) - 1\}.$$

El próximo resultado (Teorema 16) facilita la evaluación de $W_r(F, G)$, pues en este caso $W_r(F, G)$, con $0 < r < 1$, tienen una expresión cerrada en función de *r*-ésimo momento de $|X - Y|$ con respecto K.

Teorema 16. *Para* $0 < r < 1$ *la distancia de Wassertein del Lema 10 puede ser escrita como*

$$\begin{aligned} W_r^r(F, G) &= \mathbb{E}|F^{-1}(U) - G^{-1}(1 - U)|^r \\ &= \int_0^1 |F^{-1}(u) - G^{-1}(1 - u)|^r du \\ &= \mathbb{E}_K|X - Y|^r = \int_{\mathbb{R}^2} |x - y|^r dK(x, y), \end{aligned}$$

donde U es uniformemente distribuida sobre el intervalo (0,1).

Ejemplo 17. *Sean* X, Y, (X, Y) *y* (X^*, Y^*) *las variables aleatorias consideradas en el Ejemplo 14. Una vez que, para cualquier* $r \geq 1$,

$$\mathbb{E}|X^* - Y^*|^r - \mathbb{E}|X - Y|^r = \frac{2^{r+1} - 3^r - 1}{4} < 0,$$

concluimos que, para $r \geq 1$, *no es posible obtener una representación de la distancia de Wasserstein como la del Teorema 16.*

Definición 18 (Convergencia). *Sean* $\{F_n\}_{n \in \mathbb{N}}$ *y* F *funciones de distribución. Diremos que* F_n *converge em distância de Wasserstein a* F*, si*

$$W_r(F_n, F) \underset{n}{\rightarrow} 0$$

El siguiente lema básicamente nos brinda una conexión directa entre la convergencia en distancia de Wasserstein $\mathcal{W}_r$ y la convergencia en distribución. Para ello es necesario recordar la definición del conjunto $\mathcal{P}_r$ en (3).

Teorema 19 (Bickel & Freedman, 1981). *Si* $r \geq 1$, $F_n \in \mathcal{P}_r, n = 1, 2, \ldots$ *y* $F \in \mathcal{P}_r$, *entonces* $W_r(F_n, F) \underset{n}{\rightarrow} 0$ *si y solamente si,*

- $F_n \xrightarrow{d} F$
- $\int_{-\infty}^{\infty} |x|^r dF_n(x) \underset{n}{\to} \int_{-\infty}^{\infty} |x|^r dF(x)$.

En otras palabras, el Teorema de Bickel y Freedman arma que convergencia en distancia de Wasserstein es un concepto más fuerte que convergencia en distribución. Para variaciones y extensiones de este resultado, consulte las referencias (Dorea & Ferreira, 2012; Shorack & Wellner, 2009).

Definición 20. *Un proceso aleatorio* $\{X_j : j \in \mathbb{Z}\}$ *es estacionario (fuerte) si para todo* $m \geq 1$ *y* $l \in \mathbb{Z}$,

$$(X_{i1}, \dots X_{im}) \overset{d}{=} (X_{i_1+l}, \dots, X_{i_m+l}), \qquad \forall\, i_1, \dots, i_m \in \mathbb{Z},$$

Donde "$\overset{d}{=}$" *denota igualdad en distribución.*

La demostración del siguiente resultado puede ser encontrado en detalle en el Apéndice B de este trabajo.

Proposición 21. *Sea* $\{X_j : j \in \mathbb{Z}\}$ *un proceso aleatorio estacionario. Para* $m_j \geq 1$, *si* $f_i : \mathbb{R}^n \to \mathbb{R}$ *son funciones medibles, entonces* $\left\{f_j\left(X_{i_1}, \dots, X_{i_{m_j}}\right) : i_1, \dots, i_{m_j} \in \mathbb{Z}\right\}$ *también es estacionario.*

Teorema 22 (Newman, 1980). *Sea* $\{X_j : j \in \mathbb{Z}\}$ *un proceso aleatorio estacionario y asociado positivo. Supongamos que la varianza es finita y estrictamente positiva,* $0 < var X_0 < +\infty$, *y que*

$$\sigma^2 \equiv var(X_0) + 2\sum_{i=1}^{\infty} cov(X_i, X_0) < +\infty$$

Entonces,

$$\frac{S_{[k,k+n)} - n\mathbb{E}(X_0)}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1), \qquad \forall k \in \mathbb{Z}. \qquad (4)$$

Proposición 23. Convergencia de una serie monótona (Yeh, 2006). *Si para todos los números naturales* j *y* k, $a_{j,k}$ *es un número real no negativo y* $a_{j,k} \leq a_{j+1,k}$, *entonces*

$$\lim_{j\to\infty} \sum_k a_{j,k} = \sum_k \lim_{j\to\infty} a_{j,k}.$$

3. TEOREMAS DEL LÍMITE CENTRAL

Sea $\{X_j : j \in \mathbb{Z}\}$ un proceso aleatorio estacionario en el sentido de la Definición 20. Para procesos estocásticos es natural, cuando lidiamos con teoremas límites, considerar bloques de n variables aleatorias consecutivas,

$$S_n = \sum_{j=1}^{n} X_j \quad y \quad S_{[k,k+n)} = \sum_{j=k}^{k+n-1} X_j.$$

Claramente, bajo el supuesto de estacionariedad tenemos $S_{[k,k+n)} \overset{d}{=} S_n$, para todo $k \in \mathbb{Z}$, esto es, $S_{[k,k+n)}$ y S_n tienen la misma función de distribución. Para verificar esto basta considerar la función (medible) $f_j(x_1; \dots; x_n) = x_1 + x_2 + \cdots + x_n$ en la Proposición 21.

Cabe mencionar que la asociatividad positiva y la estacionariedad aseguran que, σ^2 en (6) pueda ser escrita como

$$\sigma^2 = \mathcal{X} = \sup_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} cov\left(X_k, X_j\right)$$

y que este a bien definida. En Mecánica Estadística $\mathcal{X}$ es conocida como la susceptiblidad correspondiente al proceso aleatorio $\{X_j : j \in \mathbb{Z}\}$.
Defina la variable aleatoria

$$V_{[k,k+n)} \equiv \frac{S_{[k,k+n)} - n\mathbb{E}(X_0)}{\sqrt{n}\sigma} \qquad (5)$$

con su respectiva función de distribución (acumulativa), dada por

$$F_{[k,k+n)}(x) \equiv \mathbb{P}\big(V_{[k,k+n)} \le x\big), \quad x \in \mathbb{R},$$

lo cual por simplicidad denotamos como $V_{[k,k+n)} \overset{d}{=} F_{[k,k+n)}$.

El primer resultado (Teorema 24), extraído de la referencia (Cioletti et al., 2017), se desprende del Teorema del límite central (TLC) de Newman (Teorema 22).

Teorema 24 (Cioletti et al., 2017). *Sea $\{X_j : j \in \mathbb{Z}\}$ un proceso estacionario y asociado positivo. Supongamos que la varianza es finita y estrictamente positiva $0 < varX_0 < +\infty$, y que*

$$\sigma^2 = var(X_0) + 2\sum_{i=1}^{\infty} cov(X_i, X_0) < +\infty. \qquad (6)$$

Para $0 < r \leq 2$, tenemos

$$W_r\left(F_{[k,k+n)}, \Phi\right) \underset{n}{\rightarrow} 0,$$

donde Φ denota la función de distribución de la distribución normal estándar $N(0,1)$, la cual, para simplificar, usualmente escribiremos como $\Phi \overset{d}{=} N(0,1)$.

Demostración. Sabemos que, por estacionariedad, $S_{[k,k+n)} \overset{d}{=} S_n$. Consecuentemente, ambas variables tienen los mismos momentos de orden finito, y por tanto, la misma varianza. Luego

$$var\left(S_{[k,k+n)}\right) = cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right).$$

Usando la bilinealidad y la simetría de la covarianza, la expresión anterior puede ser escrita como

$$\begin{aligned} &= \sum_{i=1}^{n}\sum_{j=1}^{n} cov\left(X_i, X_j\right) = \sum_{i=1}^{n} var(X_i) + 2\sum_{1\leq i\leq j\leq n} cov\left(X_i, X_j\right) \\ &= n\, var(X_0) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} cov\left(X_0, X_{j-i}\right) \\ &= n\, var(X_0) + 2\sum_{i=1}^{n-1}\sum_{j=1}^{n-i} cov\left(X_0, X_j\right), \qquad (7) \end{aligned}$$

donde en la cuarta igualdad, nuevamente, usamos la estacionalidad del proceso. Note que los elementos de la suma $\sum_{i=1}^{n-1}\sum_{j=1}^{n-j} cov\left(X_0, X_j\right)$ pueden ser reordenados de la siguiente forma:

$$\begin{aligned} \sum_{i=1}^{n-1}\sum_{j=1}^{n-i} cov\left(X_i, X_j\right) &= \sum_{i=1}^{n-1} (n-i)\, cov(X_0, X_i) \\ &= \sum_{i=1}^{n} (n-i)\, cov(X_0, X_i). \end{aligned}$$

Usando este reordenamiento, (7) es

$$= nvar(X_0) + 2\sum_{i=1}^{n}(n-i)\ cov(X_0, X_i).$$

Luego, tenemos

$$\frac{var\left(S_{[k,k+n)}\right)}{n} = var(X_0) + 2\sum_{i=1}^{\infty} a_{i,n}, \qquad (8)$$

donde $a_{i,n}$ es definida por

$$a_{i,n} = \mathbb{1}_{\{n\geq i\}}\left(\frac{n-i}{n}\right) cov(X_0, X_i).$$

Vea que, por asociatividad, $a_{i,n} \geq 0, a_{i,n} \leq a_{i+1,n}$ y que $a_{i,n} \underset{n}{\rightarrow} cov(X_0, X_i)$. Luego, usando la Proposición 23 sobre convergencia de series monótonas, sigue que

$$\lim_{n\to\infty}\sum_{i=1}^{\infty} a_{i,n} = \sum_{i=1}^{\infty}\lim_{n\to\infty} a_{i,n} = \sum_{i=1}^{\infty} cov(X_0, X_i).$$

Sustituyendo el límite anterior en (8), de (6) se deduce que

$$\frac{var\left(S_{[k,k+n)}\right)}{n} \underset{n}{\rightarrow} \sigma^2.$$

De este modo,

$$(9) \qquad \mathbb{E}\left(V_{[k,k+n)}^2\right) = \mathbb{E}\left(\frac{S_{[k,k+n)} - n\mathbb{E}(X_0)}{\sqrt{n}\sigma}\right)^2 \underset{n}{\rightarrow} 1 = \mathbb{E}(Z^2),$$

donde $Z \overset{d}{=} \Phi$. Como $\mathbb{E}\left(V_{[k,k+n)}^2\right)$ es convergente, $\mathbb{E}\left(V_{[k,k+n)}^2\right)$ es acotada, y claramente $V_{[k,k+n)} \in \mathcal{P}_2$. Dado que se cumple la convergencia en distribución (4), Teorema de Newman (1980), concluimos del Teorema 19 de Bickel y Freedman que $W_2\left(V_{[k,k+n)}, \Phi\right) \underset{n}{\rightarrow} 0$.

A continuación, para extender la convergencia para $0 < r < 2$ hacemos uso del Teorema de representación 13 de Dorea y Ferreira. Existe una variable aleatoria $Z^* \overset{d}{=} \Phi$ tal que la distribución conjunta de $\left(V_{[k,k+n)}, Z^*\right)$ está dada por $H(x, y) = F_{[k,k+n)}(x) \wedge \Phi(y)$ y

$$\mathbb{E}\left(V_{[k,k+n)} - Z^*\right)^2 = W_2^2\left(F_{[k,k+n)}, \Phi\right) \underset{n}{\rightarrow} 0. \qquad (10)$$

Usando el Lema 10 y la Desigualdad de Lyapunov tenemos, para $0 < r \leq 2$,

$$W_r^r\left(F_{[k,k+n)}, \Phi\right) = \inf_{\left(V_{[k,k+n)}, Z^*\right)} \mathbb{E}\left|V_{[k,k+n)} - Z^*\right|^r$$

$$\leq \mathbb{E}\left|V_{[k,k+n)} - Z^*\right|^r \leq \mathbb{E}\left(V_{[k,k+n)} - Z^*\right)^2 \underset{n}{\rightarrow} 0,$$

donde hemos usado la convergencia en (10).

Para derivar la convergencia de orden superior para W_r, se requerirán condiciones de momento adicionales en X_j. Para $k \in \mathbb{Z}$, sea $u_k(.)$ el coeficiente de Cox-Grimmet, definido por

$$u_k(n) = \sum_{j \in \mathbb{Z}: |k-j| \geq n} cov(X_k, X_j), \;\; n \geq 0.$$

Tenga en cuenta que, según el Lema 12, el proceso aleatorio $\{X_j - \mathbb{E}(X_j): j \in \mathbb{Z}\}$ también es estacionario y está asociado positivamente. Esto nos permite plantear una desigualdad de momentos de Birkel (1988) adaptada a nuestras necesidades.

Lema 25 (Cioletti et al., 2017). *Sean $2 < r < r^*$ y $\{X_j : j \in \mathbb{Z}\}$ un proceso estacionario y asociado positivo. Supongamos que $\mathbb{E}|X_0|^{r^*} < +\infty$ y que para algunas constantes $C_1 > 0$ y*

$$\theta \geq \frac{r^*(r-2)}{2(r^*-r)}$$

tenemos $u(n) \leq C_1 n^{-\theta}$. Entonces existe una constante $C_2 = C_2(r, r^*) > 0$ tal que

$$\sup_{k \in \mathbb{Z}} \mathbb{E}\left|S_{[k,k+n)} - n\mathbb{E}(X_1)\right|^r \leq C_2 n^{\frac{r}{2}}. \qquad (11)$$

Demostración. La prueba se desprende inmediatamente del Corolario 2.21 en Oliveira (2012).
Note que, según el Teorema 24, tenemos satisfechas las condiciones del Lema 25 para $r = 2$. De hecho, por (6) tenemos $u(n) \leq C_1$ y (11) sigue de (9).

El siguiente teorema, extraído de la referencia (Cioletti et al., 2017), nos brinda una extensión de la convergencia en distancia de Wasserstein para ordenar r mayores que 2.

Teorema 26 (Cioletti et al., 2017). *Sean $2 < r < r^*$ y $\{X_j : i \in \mathbb{Z}\}$ un proceso estacionario y asociado positivo. Supongamos que $\mathbb{E}|X_0|^{r^*} < +\infty$ y que para algunas constantes $C_1 > 0$ y*

$$\theta \geq \frac{r^*(r-2)}{2(r^*-r)}$$

se satisface $u(n) \leq C_1 n^{-\theta}$. Entonces, si σ^2, dada por (6), es tal que $0 < \sigma^2 < +\infty$, tenemos

$$W_r(F_{[k,k+n)}, \Phi) \underset{n}{\rightarrow} 0 \quad y \quad \mathbb{E}\left|V_{[k,k+n)}\right|^r \underset{n}{\rightarrow} \mathbb{E}|Z|^r,$$

donde $V_{[k,k+n)} \overset{d}{=} F_{[k,k+n)}$ y $V_{[k,k+n)}$ se define mediante (5)*; y $Z \overset{d}{=} \Phi \overset{d}{=} N(0,1)$.*

Demostración. Tenga en cuenta que para $\mathrm{r} < 2$, el Teorema de Newman (Teorema 22) implica $\mathrm{V}_{[\mathrm{k},\mathrm{k}+\mathrm{n})} \overset{\mathrm{d}}{\rightarrow} \mathrm{Z}$. Luego, para completar la prueba del teorema, necesitamos demostrar que

$$V_{[k,k+n)} \in \mathcal{P}_r \text{ y } \mathbb{E}\left|V_{[k,k+n)}\right|^r \underset{n}{\rightarrow} \mathbb{E}|Z|^r, \mathrm{Z} \overset{d}{=} \Phi \qquad (12)$$

En este caso la convergencia $W_r\left(F_{[k,k+n)}, \Phi\right) \underset{n}{\rightarrow} 0$ sigue inmediatamente por aplicar el Teorema de Bickel y Freedman (Teorema 19).
Si demostramos que la secuencia $\left\{|V_{[k,k+n)}|^r\right\}_{n\geq 1}$ es uniformemente integrable, entonces, por usar resultados estándar en la literatura, tendríamos la validez de la convergencia $\mathbb{E}\left|V_{[k,k+n)}\right|^r \underset{n}{\rightarrow} \mathbb{E}|Z|^r$ en (12).

Para demostrar que $\left\{|V_{[k,k+n)}|^r\right\}_{n\geq 1}$ es uniformemente integrable, basta probar la siguiente integración uniforme: para algún $r < r' < r^*$,

$$K \equiv \sup_{n\geq 1} \sup_{k\in\mathbb{Z}} \mathbb{E}|V_{[k,k+n)}|^{r'+\epsilon} < \infty,$$

lo cual implicaría que $V_{[k,k+n)} \in \mathcal{P}_{r'} \subset \mathcal{P}_r$. De hecho, usando la Desigualdad de Lyapunov tenemos

$$\int_{\{|V_{[k,k+n)}|^r \geq \ell\}} \left|V_{[k,k+n)}\right|^r(\omega) d\mathbb{P}(\omega) = \mathbb{E}\left[\mathbb{1}_{\left\{\left|V_{[k,k+n)}\right|^r \geq \ell\right\}} \left|V_{[k,k+n)}\right|^r\right]$$
$$\leq \frac{1}{\ell^{\frac{1}{r}}} \mathbb{E}\left[\mathbb{1}_{\left\{|V_{[k,k+n)}| \geq \ell^{\frac{1}{r}}\right\}} \left|V_{[k,k+n)}\right|^{r'}\right]$$

$$\leq \frac{K}{\ell^{\frac{1}{r}}}.$$

Bajo la condición $K < \infty$, al tomar $\sup_{n\geq 1}$ y luego $\lim_{\ell\to\infty}$ en la desigualdad anterior, la integrabilidade uniforme de $\left\{\left|V_{[k,k+n)}\right|^{r}\right\}_{n\geq 1}$ sigue inmediatamente.

Sea $\psi(r) = \frac{r^*(r-2)}{[2(r^*-r)]}$. Se deduce que existen $r^* > r' > r$ tales que $\theta > \psi(r')$. Simplemente tome $r' = \frac{2r^*(1+\theta)}{2\theta + r^*}$. Además, $\psi(r') \geq \psi(r)$, pues ψ es creciente. Del Lema 25 tenemos, para $C_2 = C_2(r', r^*) > 0$,

$$\sup_{k\in\mathbb{Z}} \mathbb{E}\left|S_{[k,k+n)} - n\mathbb{E}(X_1)\right|^{r'} = \sup_{k\in\mathbb{Z}} \mathbb{E}\left|V_{[k,k+n)}\sqrt{n}\sigma\right|^{r'} \leq C_2 n^{\frac{r'}{2}}.$$

Resulta que,

$$\sup_{k\in\mathbb{Z}} \mathbb{E}\left|V_{[k,k+n)}\right|^{r'} \leq C_2 \frac{n^{\frac{r'}{2}}}{\left(\sqrt{n}\sigma\right)^{r'}} = \frac{C_2}{\sigma^{r'}} < +\infty, \quad \forall n \geq 1.$$

Esto concluye la prueba del teorema.

Como aplicación de los Teoremas 24, 26 y el Lema 28 (vea Apéndice C), tenemos el siguiente resultado de convergencia:

Corolario 27. *Si el proceso aleatorio $\{X_j : j \in \mathbb{Z}\}$ satisface las hipótesis del Teorema 24 (o Teorema 26), entonces*

$$d_K\left(F_{[k,k+n)}, \Phi\right) \underset{n}{\to} 0,$$

donde $d_K(F,G) = sup_{x\in\mathbb{R}}|F(x) - G(y)|$ es la distancia de Kolmogorov enter F y G.

AGRADECIMIENTO

RV agradece la invitación de los organizadores del evento "I Congreso Internacional de Investigación de Ingeniería Económica, Estadística y Ciencias Sociales".

APÉNDICE A. DEMOSTRACIÓN DEL LEMA 10

La prueba de este resultado requiere el uso del Teorema de cambio de variables.

Nuestro objetivo principal es demostrar que

$$W_r^r(F,G) \leq \inf_{(X,Y)} \mathbb{E}|X-Y|^r \qquad (13)$$

y que

$$W_r^r(F,G) \geq \inf_{(X,Y)} \mathbb{E}|X-Y|^r \qquad (14)$$

Primero, demostraremos la desigualdad en (13). Para ello, considere un vector aleatorio $Z=(X,Y)$, definido sobre algún espacio de probabilidad $(\Omega,\mathfrak{F},\mathbb{P})$, cuyas distribuciones marginales de X e Y son dadas por F y G, respectivamente. Si $\widetilde{H}$ denota la distribución conjunta de **Z**, entonces las distribuciones marginales de $\widetilde{H}$ son dadas por:

$$\lim_{y\to\infty} \widetilde{H}(x,y) = \mathbb{P}(X\in(-\infty\,,x], Y\in\mathbb{R}) = F(x), \qquad (15)$$
$$\lim_{y\to\infty} \widetilde{H}(x,y) = \mathbb{P}(X\in\mathbb{R}, Y\in(-\infty\,,y]) = G(x).$$

Considere la función Borel medible $g:\mathbb{R}^2\to\mathbb{R}$ definida por $g(x,y)=|x-y|^r$. Usando la definición de esperanza y aplicando el Teorema de cambio de variables, obtenemos

$$\mathbb{E}|X-Y|^r = \int_\Omega g\big(X(\omega),Y(\omega)\big)d\mathbb{P}(\omega)$$
$$= \int_{\mathbb{R}^2} g(x,y)d(\mathbb{P}\ o\ Z^{-1})(x,y).$$

Afirmamos que $\pi=\mathbb{P}\ o\ Z^{-1}\in\Pi(F,G)$. De hecho, esto es inmediato, pues

$$\pi\big((-\infty,x]\times\mathbb{R}\big) = \mathbb{P}\ o\ Z^{-1}\big((-\infty,x]\times\mathbb{R}\big)$$
$$= \mathbb{P}\ (X\in(-\infty,x], Y\in\mathbb{R}) = F(x),$$

donde en la última igualdad usamos (15). De forma similar, tenemos

$$\pi(R\ \times\ (-\infty,y])\ =\ G(y),$$

lo cual prueba la afirmación.

Los argumentos anteriores muestran que, para cada vector aleatorio (X,Y) cuyas distribuciones marginales están dadas por F y G, existe una medida de probabilidad $\pi\ \in\ \Pi(F,G)$, tal que

$$\mathbb{E}|X-Y|^r = \int_{\mathbb{R}^2} |x-y|^r d\pi(x,y) \geq W_r^r(F,G).$$

Como consecuencia de esta igualdad y de la definición del ínfimo, obtenemos
$$W_r^r(F,G) \le \inf_{(X,Y)} \mathbb{E}|X-Y|^r.$$
Esto es, la desigualdad en (13) es satisfecha.

Ahora demostraremos la desigualdad recíproca en (14). De hecho, dado $\pi \in \Pi(F,G)$, considere el espacio de probabilidad $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2), \pi)$ y las variables aleatorias (proyecciones) π_1 y π_2 definidas por $\pi_1(x,y) = x$ y $\pi_2(x,y) = y$, respectivamente. Desde que $\pi \in \Pi(F,G)$, las distribuciones marginales del vector (π_1, π_2) son F y G, respectivamente. Usando la definición de las variables π_1 y π_2, tenemos

$$\mathbb{E}|\pi_1 - \pi_2|^r = \int_{\mathbb{R}^2} |x-y|^r d\pi(x,y).$$

Así, para cada $\pi \in \Pi(F,G)$, se construyó un vector aleatorio (π_1, π_2) que tiene como distribuciones marginales a F y G, respectivamente, de modo que se cumple la igualdad anterior. Por lo tanto, hemos demostrado que

$$\inf_{(X,Y)} \mathbb{E}|X-Y|^r \le \int_{\mathbb{R}^2} |x-y|^r d\pi(x,y), \quad \pi \in \Pi(F,G).$$
Por lo tanto,
$$\inf_{(X,Y)} \mathbb{E}|X-Y|^r \le W_r^r(F,G).$$

Esto es, la desigualdad en (14) es válida.
De (13) y (14) la prueba del lema sigue inmediatamente.

APÉNDICE B. DEMOSTRACIÓN DE LA PROPOSICIÓN 21

Debemos verificar que, $\forall\, m_j > 1, k \in \mathbb{Z}$,
$$f_j\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right) \overset{d}{=} f_j\left(X_{i_1+k}, X_{i_2+k}, \dots, X_{i_{m_j}+k}\right),$$

De hecho, por cada $x \in \mathbb{R}$ tenemos

$$\mathbb{P}\left(f_i\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right) \le x\right) = \mathbb{P}\left(\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right) \in f_j^{-1}((-\infty, x])\right), \quad (16)$$

donde $f_j^{-1}(.)$ es la imagen de f_j y $f_j^{-1}\big((-\infty,x]\big)$ es un boreliano de $\mathbb{R}^n$. Una vez qué $\{X_j : j \in \mathbb{Z}\}$ es estacionario, vea que, $\forall\, m_j > 1, k \in \mathbb{Z}$,

$$\mathbb{P}\left(\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right) \in f_j^{-1}\big((-\infty,x]\big)\right)$$
$$= \mathbb{P}\left(\left(X_{i_1+k}, X_{i_2+k}, \dots, X_{i_{m_j}+k}\right) \in f_j^{-1}\big((-\infty,x]\big)\right). \quad (17)$$

De (16) y (17) obtenemos, $\forall\, m_j > 1, k \in \mathbb{Z}, x \in \mathbb{R}$,

$$\mathbb{P}\left(f_i\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right) \leq x\right)$$
$$= \mathbb{P}\left(\left(X_{i_1+k}, X_{i_2+k}, \dots, X_{i_{m_j}+k}\right) \in f_j^{-1}\big((-\infty,x]\big)\right)$$
$$= \mathbb{P}\left(f_j\left(X_{i_1+k}, X_{i_2+k}, \dots, X_{i_{m_j}+k}\right) \leq x\right).$$

En otras palabras, $\forall\, m_j \geq 1\, y\, k \in \mathbb{Z}$, las variables aleatorias $f_j\left(X_{i_1}, X_{i_2}, \dots, X_{i_{m_j}}\right)$ y $f_j\left(X_{i_1+k}, X_{i_2+k}, \dots, X_{i_{m_j}+k}\right)$ son iguales en distribución. Esto concluye la demostración.

APÉNDICE C. UN RESULTADO ADICIONAL

Lema 28. *Sean X e Y dos variables aleatorias tal que Y es absolutamente continua con densidad f_Y tal que $|f_Y(y)| \leq C$, $\forall y \in Supp(Y)$ y para alguna constante $C > 0$. Entonces,*

$$d_K(F_X, F_Y) \leq 2\sqrt{CW_1(F_X, F_Y)},$$

donde $d_K(F_X, F_Y) = sup_{x\in\mathbb{R}} |F_X(x) - F_Y(x)|$es la distancia de Kolmogorov entre las distribuciones F_X y F_Y.

Demostración. Dado $\varepsilon > 0$ y $t \in \mathbb{R}$ fijo, definimos

$$g_t(x) \equiv \mathbb{1}_{(-\infty,t]}(x) + \left[1 - \frac{1}{\varepsilon}(x-t)\right]\mathbb{1}_{(t,t+\varepsilon)}(x),$$
$$h_t(x) \equiv g_{t-\varepsilon}(x).$$

Para cada $x, y \in \mathbb{R}$ note que

$$g_t(x) - g_t(y) = \frac{1}{\varepsilon}(y - x)\mathbb{1}_{(t,\infty)}(x)\mathbb{1}_{(-\infty,t+\varepsilon)}(y),$$
$$h_t(x) - h_t(y) = \frac{1}{\varepsilon}(y - x)\mathbb{1}_{(t-\varepsilon,\infty)}(x)\mathbb{1}_{(-\infty,t)}(y),$$

lo cual implica que g_t y h_t son funciones Lipschitz con constante Lipschiptz $\frac{1}{\varepsilon}$. Dado que

$$\mathbb{E}[g_t(X)] \geq F_X(t), \quad \mathbb{E}[h_t(Y)] \leq F_Y(t) + C_{\varepsilon,}$$

Obtenemos

$$\begin{aligned} F_X(t) - F_Y(t) &\leq \mathbb{E}[g_t(X) - g_t(Y)] + \{\mathbb{E}[g_t(Y)] - F_Y(t)\} \\ &\leq \frac{1}{\varepsilon}\mathbb{E}|X - Y| + C_{\varepsilon,} \end{aligned}$$

donde hemos usado el hecho de que g_t es Lipschitz. Por la desigualdad anterior y de la definición de la distancia de Wasserstein se deduce que

$$F_X(t) - F_Y(t) \leq \frac{1}{\varepsilon} W_1(F_X, F_Y) + C_\varepsilon.$$

De forma análoga, usando que h_t es Lipschitz, encontramos que

$$F_Y(t) - F_X(t) \leq \frac{1}{\varepsilon} W_1(F_X, F_Y) + C_\varepsilon.$$

Por lo tanto,

$$|F_X(t) - F_Y(t)| \leq \frac{1}{\varepsilon} W_1(F_X, F_Y) + C_\varepsilon, \qquad t \in \mathbb{R}.$$

Una vez que la función $\varepsilon \to \frac{W_1(F_X,F_Y)}{\varepsilon} + C_\varepsilon$ tiene un valor máximo $2\sqrt{CW_1(F_X,F_Y)}$ en $\varepsilon = \varepsilon = \sqrt{W_1(F_X,F_Y)/C}$, la demostración del corolário sigue rápidamente.

4. REFERENCIAS

Ambrosio, L. (2003). Lecture Notes on Optimal Transport Problems. In: Ambrosio, L., Deckelnick, K., Dziuk, G., Mimura, M., Solonnikov, V. A., Soner, H. M., & Ambrosio, L. (Eds.), *Mathematical Aspects of Evolving Interfaces* (pp. 1-52). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-39189-0_1

Barlow, R. E., & Proschan, F. (1975). *Statistical theory of reliability and life testing: probability models*. Holt, Rinehart and Winston. https://apps.dtic.mil/sti/citations/ADA006399

Bickel, P. J., & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The annals of statistics, 9*(6), 1196-1217. https://doi.org/10.1214/aos/1176345637

Birkel, T. (1988). Moment bounds for associated sequences. *The annals of Probability, 16*(3), 1184-1193. https://www.jstor.org/stable/2244116

Cioletti, L., Dorea, C. C. Y., & Vila, R. (2017). *Limit Theorems in Mallows Distance for Processes with Gibssian Dependence*. arXiv. https://doi.org/10.48550/arXiv.1701.03747

Dorea, C. C., & Ferreira, D. B. (2012). Conditions for equivalence between Mallows distance and convergence to stable laws. *Acta Mathematica Hungarica, 134*(1-2), 1-11. https://doi.org/10.1007/s10474-011-0101-7

Esary, J. D., Proschan, F., & Walkup, D. W. (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics, 38*(5), 1466-1474. https://doi.org/10.1214/aoms/1177698701

Gabriel, R. V. (2017). Representações gráficas para sistemas de spins com presença de campo externo: algumas relações em teoria de probabilidades [Tese para obtenção do grau de Doutor em Matemática]. Universidade de Brasília. Instituto de Ciências Exatas. Departamento de Matemática. http://icts.unb.br/jspui/handle/10482/22471

Gray, R. M. (2009). *Probability, random processes, and ergodic properties*. Springer. https://doi.org/10.1007/978-1-4419-1090-5

Jordan, R., Kinderlehrer, D., & Otto, F. (1998). The variational formulation of the Fokker--Planck equation. *SIAM journal on mathematical analysis, 29*(1), 1-17. https://doi.org/10.1137/S0036141096303359

Kantorovič, L. V., Rubinšteĭn, G. Š. (1958). On a space of completely additive functions. *Vestnik Leningrad University, 13*, 52-59.

Mallows, C. L. (1972). A note on asymptotic joint normality. *The Annals of Mathematical Statistics, 43*(2), 508-515. https://www.jstor.org/stable/2239988

Newman, C. M. (1980). Normal fluctuations and the FKG inequalities. *Communications in Mathematical Physics, 74*(2), 119-128. https://doi.org/10.1007/BF01197754

Newman, C. M., & Wright, A. L. (1981). An invariance principle for certain dependent sequences. *The Annals of Probability, 9*(4), 671-675. https://doi.org/10.1214/aop/1176994374

Oliveira, P. E. (2012). *Asymptotics for associated random variables*. Springer Science & Business Media.

Otto, F. (2001). The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations, 26*(1-2), 101-174. https://doi.org/10.1081/PDE-100002243

Pitt, L. D. (1982). Positively correlated normal variables are associated. *The Annals of Probability, 10*(2), 496-499. https://www.jstor.org/stable/2243445

Rachev, S. T., & Rüschendorf, L. (1998). *Mass Transportation Problems: Volume I: Theory*. Springer Science & Business Media.

Shorack, G. R., & Wellner, J. A. (2009). *Empirical processes with applications to statistics*. Society for Industrial and Applied Mathematics. https://epubs.siam.org/doi/pdf/10.1137/1.9780898719017.bm

Sommerfeld, M., & Munk, A. (2018). Inference for empirical Wasserstein distances on finite spaces. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 80*(1), 219-238. https://doi.org/10.1111/rssb.12236

Vaserstein, L. N. (1969). Markov processes over denumerable products of spaces, describing large systems of automata. *Problemy Peredachi Informatsii, 5*(3), 64-72. https://www.mathnet.ru/eng/ppi1811

Villani, C. (2003). Topics in optimal transportation. *OR/MS Today, 30*(3), 66-67. https://link.gale.com/apps/doc/A104669453/AONE?u=anon~6226aa1c&sid=googleScholar&xid=2585334e

Villani, C. (2009). *Optimal transport: old and new*. Springer. https://doi.org/10.1007/978-3-540-71050-9

Yeh, J. (2006). *Real Analysis: Theory of Measure and Integration*. World Scientific.