



Universidad de Valladolid

Facultad de Ciencias

Trabajo Fin de Grado

Grado en Matemáticas

**Métricas probabilísticas en el Teorema
Central del Límite**

Autor: Kevin Muñoz Borondo

Tutor: Eustasio del Barrio Tellado

EUSTASIO DEL BARRIO TELLADO, Profesor Titular del Departamento de Estadística e Investigación Operativa de la Universidad de Valladolid,

CERTIFICA:

Que el presente trabajo, “Métricas Probabilísticas en el Teorema Central del Límite”, ha sido realizado bajo su dirección por Kevin Muñoz Bodondo y constituye su Trabajo de Fin de Grado para optar al Título de Graduado en Matemáticas.

Que le consta que el trabajo es original e inédito y que autoriza su presentación.

Y para que conste a los efectos oportunos firma la presente en Valladolid a 16 de julio de dos mil diecinueve.

Fdo: Eustasio del Barrio Tellado

Índice general

RESUMEN	V
INTRODUCCIÓN	1
1. EL MÉTODO DE STEIN PARA LA APROXIMACIÓN NORMAL.	7
1.1. El Lema de Stein	7
1.2. Métricas probabilísticas	11
1.3. El método de Stein para sumas de variables aleatorias inde- pendientes	15
1.3.1. Identidades de Stein	18
2. EL MÉTODO DE STEIN PARA MODELOS DE DEPENDENCIA LOCAL	25
2.1. Aproximación normal de sumas de variables aleatorias local- mente dependientes	25
2.2. Aplicación: un T.C.L. en el modelo de Erdős-Renyi	30
3. APROXIMACIÓN NORMAL DE FUNCIONALES DE POISSON MEDIANTE EL MÉTODO DE STEIN	43
3.1. Procesos de Poisson.	47
3.1.1. El desarrollo en caos de Poisson.	49
3.2. Cálculo de Malliavin	52
3.3. Aplicación: TCL para el volumen de conjuntos aleatorios. . .	59
3.4. Apéndice.	70

RESUMEN

Uno de los campos de estudio principales de la Teoría de la Probabilidad es el desarrollo de Teoremas Centrales del Límite generalizados. Dentro de este estudio, una de las técnicas más interesantes y con más desarrollo en los últimos años es el llamado método de Stein. Su ventaja frente a otras técnicas radica en que se desarrolla a partir de una idea muy sencilla, pues se basa en la comparación de esperanzas como método para determinar cómo de buena es una aproximación entre distribuciones probabilísticas. Tras desarrollar esta idea introduciendo los fundamentos del método de Stein y presentar el concepto de métrica probabilística, aplicamos el método a sumas de variables aleatorias independientes y damos una prueba del Teorema Central del Límite clásico y bajo la formulación de Lindeberg. Después, generalizamos el método de Stein a variables aleatorias con dependencia local y desarrollamos un Teorema Central del Límite en un caso más complejo. Concretamente estudiamos una situación combinatoria que genera un modelo especial de grafo. Finalmente, extendemos nuestro estudio sistematizándolo al estudio de Teoremas Centrales del Límite en espacios de funciones generalizados. En particular, estudiamos los procesos de Poisson y el cálculo de Malliavin, que combinado con el método de Stein nos permite realizar el análisis de un modelo clásico de la geometría estocástica, en el que estudiamos un funcional concreto del espacio de Poisson, para el cual seremos capaces de desarrollar un Teorema Central del Límite.

INTRODUCCIÓN

Uno de los temas principales en la Teoría de la Probabilidad es la determinación de Teoremas Centrales del Límite. En la práctica de la Estadística Inferencial, es poco común disponer de resultados distribucionales exactos que nos permitan realizar el cálculo efectivo de intervalos de confianza, valores críticos o p-valores. Por ello, es frecuente tener que recurrir a aproximaciones en distribución que subsanen la carencia de estos resultados exactos. Si consideramos que el estadístico de interés, llamémosle, T_n , es un elemento de la sucesión de variables aleatorias $\{T_n\}_{n \geq 1}$, se dice entonces que la sucesión converge en distribución a la variable aleatoria T si

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_n \leq t) = \mathbb{P}(T \leq t), \quad (0.1)$$

para cada $t \in \mathbb{R}$ en el que la función $F_T(x) = \mathbb{P}(T \leq x)$ es continua. Si F_T es continua en todo \mathbb{R} entonces (0.1) es equivalente a

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{R}} |\mathbb{P}(T_n \leq t) - \mathbb{P}(T \leq t)| = 0. \quad (0.2)$$

Este y otros resultados a los que hacemos referencia en esta sección, junto con sus respectivas demostraciones, pueden consultarse en las secciones 14, 25, 26 y 27 de [4].

La equivalencia entre (0.1) y (0.2) nos permite a menudo sustituir las probabilidades asociadas a T_n , que serán de difícil cálculo, por aquellas asociadas a T que serán más manejables.

Ahora bien, nuestro problema será el cómo probar la convergencia en distribución, que salvo en ciertos ejemplos notables (véase el caso de extremos tratado en la sección 14 de [4]), presentará el obstáculo que supone el no disponer de forma usual de una expresión sencilla de manejar de $F_n(t) = \mathbb{P}(T_n \leq t)$, lo que nos llevará a que en la mayoría de situaciones resultará más conveniente manejar transformadas de la distribución de probabilidad. El caso más importante y útil en este sentido será la conocida como *función característica*

$$\varphi_T(n) = \mathbb{E}(e^{inT}), \quad n \in \mathbb{R}.$$

Esta función está bien definida para cualquier variables aleatoria T que determina completamente su distribución. El interés en este contexto en manejar la función característica viene dado por el Teorema de Continuidad de Levy (véase [4], Teorema 26,3), por el cual podemos caracterizar la convergencia en distribución en términos de las funciones características, es decir, (1) es equivalente a que

$$\lim_{n \rightarrow \infty} \varphi_{T_n}(n) = \varphi_T(n)$$

para cada $n \in \mathbb{R}$, donde $\varphi_{T_n}(n) = \mathbb{E}(e^{inT_n})$.

Entre los estadísticos de uso más frecuente y de mayor utilidad se encuentra la media muestral, calculada a partir de una muestra de variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) y que representamos a través de la expresión $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$. Suponiendo que las variables están estandarizadas, es decir, para cada i se cumple $\mathbb{E}(X_i) = 0$ y $Var(X_i) = 1$ entonces si $T_n = \sqrt{n}\bar{X}_n$, propiedades elementales de las funciones características nos permitirán afirmar en primer lugar que

$$\varphi_{T_n}(n) = \varphi_{X_1} \left(\frac{n}{\sqrt{n}} \right)^n, \quad n \in \mathbb{R} \quad (0.3)$$

y, por otro lado,

$$\varphi_{X_1}(n) = 1 - \frac{u^2}{2} + \beta(u), \quad (0.4)$$

con $\beta(u) = o(u^2)$ en 0, lo cual, nos permite probar, a partir de (3) y (4) (véase [4], Teorema 27,1) que

$$\varphi_{T_n}(n) \longrightarrow e^{-u^2/2}, \quad n \in \mathbb{R}. \quad (0.5)$$

El lado derecho de (5) es la función característica de la distribución normal estándar $\mathcal{N}(0, 1)$, luego por el Teorema de Continuidad de Levy, esto demuestra que

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(\sqrt{n}\bar{X}_n \leq t) - \mathbb{P}(T \leq t)| \xrightarrow{n \rightarrow \infty} 0. \quad (0.6)$$

donde T es una variable aleatoria que sigue una distribución normal $\mathcal{N}(0, 1)$. Este es el Teorema Central del Límite de Lindeberg-Levy, que tiene una gran cantidad de importantes consecuencias en la práctica estadística.

Disponemos además de diversos resultados que nos permiten extender (6) a

otros estadísticos. Por ejemplo, si tomamos $T_n = g(\bar{X}_n)$ con g derivable en 0, entonces el Δ -método nos garantiza que el estadístico $\sqrt{n}(T_n - g(0))$ es aproximadamente $\mathcal{N}(0, (g'(0))^2)$.

Sin embargo, esta extensión que planteamos tiene sus limitaciones. Si T_n no es una función suave de un estadístico lineal, necesitaremos utilizar otro tipo de justificación. El problema que nos surge a la hora de adaptar el argumento que hemos utilizado antes, es que la factorización dada en (3) falla si T_n no es un estadístico expresado como suma de términos independientes. Cabe esperar que el Teorema Central del Límite siga siendo válido en situaciones que sean similares al caso de Lindeberg-Levy, sin embargo, necesitaremos introducir otro tipo de métodos para poder demostrarlo.

Otra de las limitaciones que encontramos en el método si nos basamos en el Teorema de continuidad de Levy y funciones características, va a ser que no nos da una cota para la distancia (más adelante nos ocuparemos de dar sentido a esto) entre la distribución de T_n y la distribución normal límite, es decir, no nos da una cota que muestre como de buena es la aproximación que realizamos entre la distribución de interés y la normal, lo cual sería deseable para poder valorar el error que estamos cometiendo si cambiamos percentiles o probabilidades de T_n por sus correspondientes de T .

Estas limitaciones que se producen sobre el método descrito, nos llevan a la necesidad de considerar otros métodos que nos permitan subsanar estas deficiencias y que nos permitan desarrollar Teoremas Centrales del Límite para estadísticos más generales. Uno de estos métodos será el llamado método de Stein, desarrollado por Stein en los años 70 y presentado en [19]. El método de Stein será un compendio de técnicas que nos permite la aproximación de distribuciones de interés a través de funciones probabilísticas modelo, de hecho, este método no se limita a distribución límite normal, pero nosotros sólo trataremos este caso. (véase [1] para encontrar una descripción detallada del método de Stein aplicado a otras distribuciones).

El método de Stein se basa en la siguiente caracterización de la distribución normal. Una variable aleatoria T sigue una distribución de tipo $\mathcal{N}(0, 1)$ si y sólo si

$$\mathbb{E}(f'(T) - Tf(T)) = 0,$$

para toda función f suave, en un sentido que se precisará más adelante.

La idea principal de Stein que es toda la base que sustenta el método, se basa en que T_n será aproximadamente $\mathcal{N}(0, 1) \Leftrightarrow$

$$\mathbb{E}(f'(T_n) - T_n f(T_n)) \simeq 0.$$

Para formalizar esta aproximación resulta conveniente manejar un cierto tipo de métricas probabilísticas, las llamadas *métricas integrales*, es decir, aquellas que toman la forma siguiente

$$d_{\mathcal{H}}(\mathcal{L}(T_n), \mathcal{L}(T)) = \sup_{h \in \mathcal{H}} |\mathbb{E}h(T_n) - \mathbb{E}h(T)|. \quad (0.7)$$

Esta categoría incluye métricas de uso frecuente como la distancia en variación total, la métrica de Kolmogorov y la métrica de Wasserstein.

Ahora bien, para calcular una cota del lado derecho de (0.7), nos valdremos de la ecuación diferencial de Stein.

$$f'(x) - xf(x) = h(x) - \mathbb{E}h(T). \quad (0.8)$$

Si tenemos que f_h es solución de (0.8) entonces

$$\mathbb{E}h(T_n) - \mathbb{E}h(T) = \mathbb{E}(f'_h(T_n) - T_n f_h(T_n)).$$

Así pues, el método de Stein se basa en la idea de a través de distintos tipos de técnicas, lograr estimar que el lado derecho es pequeño, y que por tanto podemos acotar la expresión

$$\sup_{h \in \mathcal{H}} |\mathbb{E}(f'_h(T_n) - T_n f_h(T_n))|, \quad (0.9)$$

lo cual nos dará una cota para $d_{\mathcal{H}}(\mathcal{L}(T_n), \mathcal{L}(T))$.

Esta idea simple, es sorprendentemente potente, pues elimina las limitaciones que se presentaban en otros métodos y es capaz de proporcionarnos tasas de convergencia en problemas de difícil tratamiento utilizando otras vías.

El método de Stein se basa en la comparación de esperanzas, tal y como plantea Stein en [20], problema para el cual disponemos de gran cantidad de técnicas para su estudio, no presenta requerimientos a priori sobre la estructura de la variable aleatoria a tratar, y además no exigirá independencia, lo que le hace particularmente útil en problemas de tipo geométrico.

En esta memoria realizamos una descripción del método de Stein aplicado a la aproximación normal para la distancia de Wasserstein, siguiendo principalmente [1], [2], [5], [15], [17] y [18]. En el primer capítulo comenzamos desarrollando la idea sobre la que se apoya, llamada Lema de Stein, que tras una breve introducción será presentada como la caracterización fundamental del método de Stein, tras esto introducimos las métricas probabilísticas, que

nos permitirá finalmente definir la noción de distancia entre distribuciones de probabilidad. Después aplicaremos el método de Stein para el cálculo de un Teorema Central del Límite en variables aleatorias independientes. En primer lugar para el caso en que estén idénticamente distribuidas mediante técnicas elementales, para más tarde introducir las identidades de Stein que nos permiten aplicar el enfoque del K -método. Esta técnica nos permitirá aplicar el método cuando la condición de distribución idéntica entre las variables aleatorias que conforman el funcional suma no se satisface. Por último, mejoramos este resultado introduciendo la condición de Lindeberg y probando el Teorema Central del Límite de Lindeberg a través del método de Stein.

En el segundo capítulo introducimos la aproximación normal para variables aleatorias localmente dependientes, para ello presentamos dos tipos especiales de dependencia local entre variables aleatorias, dando cotas que se ajustan a cada una de estas dos situaciones particulares junto con algunos ejemplos interesantes. Tras esto, tratamos en profundidad el método de los grafos de dependencia a través del estudio de una de las aplicaciones clásicas del método de Stein, tal y como puede verse en [3] y [18]. Nuestro objetivo será determinar un Teorema Central del Límite aplicado a grafos aleatorios que siguen un modelo generador de Erdos-Renyi, en concreto determinaremos que el número de triángulos que se forman en este tipo de grafo sigue una distribución normal. Este problema, que desde un punto de vista analítico resulta complicado, gracias al método de Stein se transforma en un problema combinatorio que será más sencillo de tratar a través de técnicas conocidas.

El último capítulo estará dedicado a dar una sistematización del método de Stein, para ello dedicamos este capítulo final al estudio de la aproximación normal de funcionales de Poisson en procesos de Poisson homogéneos. Esto nos permitirá mostrar como el método de Stein, basandonos en la teoría expuesta en [10], [11] y [14] puede combinarse de manera satisfactoria con otras teorías permitiéndonos obtener Teoremas Centrales del Límite en gran cantidad de situaciones. Introducimos un funcional de estudio con una sencilla interpretación geométrica, este funcional representará el volumen de la unión de bolas d -dimensionales en el espacio $[0, 1]^d$, estando estas bolas centradas en los puntos generados por un proceso de Poisson homogéneo η sobre el espacio \mathbb{R}^d , y vemos como este volumen sigue una distribución aproximadamente normal. Dedicaremos el capítulo a la introducción de la teoría relativa a procesos de Poisson junto con varios elementos del cálculo de Malliavin que serán necesarios. Finalmente vemos que el calculo de Malliavin combinado con el método de Stein nos permite obtener cotas para la distancia de Wasserstein entre funcionales de Poisson y la distribución normal y aplicamos esto a nuestro funcional de estudio.

De forma similar al capítulo anterior, vemos como el método de Stein lo que nos proporcionará será una simplificación del problema haciendo que finalmente los cálculos que necesitamos para fijar la cota se hagan a través de un análisis puramente geométrico.

Capítulo 1

EL MÉTODO DE STEIN PARA LA APROXIMACIÓN NORMAL.

1.1. El Lema de Stein

Comenzamos el capítulo mostrando los fundamentos del método de Stein, comenzando por la caracterización de la distribución normal dada por Stein en los 70 en [19], que se describirá en el llamado Lema de Stein, junto con la solución general de la llamada ecuación de Stein y algunas propiedades de la misma.

Sea X una variable aleatoria. Un operador de caracterización para X es un operador T_0 que actúa sobre una familia de funciones \mathcal{A} , de tal forma que para cualquier variable aleatoria Y se cumple la relación

$$\mathbb{E}T_0f(Y) = 0 \quad \forall f \in \mathcal{A} \quad \Leftrightarrow \quad Y \stackrel{d}{=} X.$$

Tal y como anunciamos en la Introducción, en el caso de la distribución normal estándar, un operador de caracterización es

$$T_0f(x) = f'(x) - xf(x),$$

con clase \mathcal{A} asociada dada por las funciones continuas, derivables a trozos en \mathbb{R} y tales que $\mathbb{E}|f'(z)| < \infty$ si $Z \sim \mathcal{N}(0, 1)$. Este resultado se conoce como Lema de Stein. Esta sección se dedica a presentar el resultado de forma precisa, así como una demostración del mismo.

Durante toda la memoria Z denota una variable aleatoria con distribución de probabilidad $\mathcal{N}(0, 1)$.

Damos a continuación la caracterización de Stein para la distribución normal, la cual conocemos como Lema de Stein. Denotamos \mathcal{C}_{bd} al conjunto

de todas las funciones $f : \mathbb{R} \rightarrow \mathbb{R}$ continuas y diferenciables a trozos que cumplen $\mathbb{E}|f'(Z)| < \infty$.

Lema 1.1. *Si W es una variable aleatoria que sigue una distribución normal estándar. Entonces*

$$\mathbb{E}f'(W) = \mathbb{E}\{Wf(W)\}, \quad (1.1)$$

para toda $f \in \mathcal{C}_{bd}$.

Recíprocamente, si (1.1) se cumple para toda función f acotada, continua y diferenciable a trozos con $\mathbb{E}|f'(Z)| < \infty$, entonces W sigue una distribución normal estándar.

Antes de probar el lema anterior, presentamos el siguiente resultado pues será necesario para la demostración.

Sea Φ la función de distribución acumulada de la distribución normal estándar $\Phi(x) = \mathbb{P}(Z \leq x)$ y sea $\Phi(h) = \mathbb{E}h(Z)$ donde Z es una variable aleatoria que sigue una distribución $\mathcal{N}(0, 1)$.

Nota: La notación anterior será de uso general durante toda la memoria, de ahora en adelante la suposición sobre la distribución de Z será omitida, ya que siempre que hablemos de la variable aleatoria Z está será $\mathcal{N}(0, 1)$.

Lema 1.2. *Sea $w \in \mathbb{R}$, y $h : \mathbb{R} \rightarrow \mathbb{R}$ tal que $\mathbb{E}|h'(Z)| < \infty$ entonces la única solución acotada $f(w) := f_h(w)$ de la ecuación diferencial*

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z) \quad (1.2)$$

viene dada por

$$\begin{aligned} f(w) &= e^{w^2/2} \int_{-\infty}^w [h(t) - \Phi(h)] e^{-t^2/2} dt \\ &= -e^{w^2/2} \int_w^{\infty} [h(t) - \Phi(h)] e^{-t^2/2} dt. \end{aligned} \quad (1.3)$$

Nota: A la ecuación diferencial de primer orden dada por (1.2) se la conoce como ecuación de Stein. (En este caso asociada a h y Z).

Demostración. Multiplicando ambos lados de (1.2) por el factor $e^{-w^2/2}$

$$\left(e^{-w^2/2} f(w) \right)' = -e^{-w^2/2} (h(w) - \Phi(h)).$$

Por tanto tendremos que f_h viene dado por

$$\begin{aligned} f_h(w) &= e^{w^2/2} \int_{-\infty}^w [h(t) - \Phi(h)] e^{-t^2/2} dt \\ &= -e^{w^2/2} \int_w^{\infty} [h(t) - \Phi(h)] e^{-t^2/2} dt. \end{aligned}$$

■

El Lema 1.3. nos muestra que la solución es acotada, sin embargo dejamos esto para más adelante en el capítulo.

En particular, podemos ver que si tenemos la situación particular en que $h(w) = \mathbf{1}_{\{w \leq x\}} = \mathbf{1}_{(-\infty, x]}(w)$ entonces podemos escribir la ecuación de Stein como

$$f'_x(w) - wf_x(w) = \mathbf{1}_{\{w \leq x\}} - \Phi(x),$$

y la única solución acotada viene dada por la función $f(w) = f_z(w)$ que viene dada por

$$\begin{aligned} f_z(w) &= e^{w^2/2} \int_{-\infty}^w [\mathbf{1}_{(-\infty, z]}(x) - \Phi(z)] e^{-x^2/2} dx = -e^{w^2/2} \int_w^{\infty} [\mathbf{1}_{(-\infty, z]}(x) - \Phi(z)] e^{-x^2/2} dx \\ &= \begin{cases} \sqrt{2\pi} e^{w^2/2} \Phi(w) [1 - \Phi(z)] & \text{si } w \leq z, \\ \sqrt{2\pi} e^{w^2/2} \Phi(z) [1 - \Phi(w)] & \text{si } w \geq z. \end{cases} \end{aligned}$$

La comprobación de esto es análoga al caso general, simplemente tendremos que integral tras multiplicar a ambos lados por el factor $e^{-w^2/2}$.

Probamos ahora el Lema 1.1.

Demostración. Comenzamos probando la primera parte del lema.

Si W sigue una distribución normal estándar, sea $f \in \mathcal{C}_{bd}$

$$\begin{aligned} \mathbb{E}(f'(W)) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(w) e^{-w^2/2} dw = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 f'(w) \left(\int_{-\infty}^w (-x) e^{-x^2/2} dx \right) dw \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f'(w) \left(\int_w^{\infty} x e^{-x^2/2} dx \right) dw \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \left(\int_x^0 f'(w) dw \right) (-x) e^{-x^2/2} dx \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \left(\int_0^x f'(w) dw \right) (x) e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [f(x) - f(0)] x e^{-x^2/2} dx = \mathbb{E}(W f(W)), \end{aligned}$$

donde la segunda igualdad es consecuencia del Teorema de Fubini.

Esto demuestra que la condición es necesaria. Veamos ahora que la condición es suficiente. Tenemos que f_z es una función acotada, y además cumple ser continua y diferenciable a trozos. Si suponemos que (1.1) se cumple para todo $f \in \mathcal{C}_{bd}$. Entonces se sigue por la definición de la solución f_z y (1.2)

$$0 = \mathbb{E}(f'_z(w) - wf_z(w)) = \mathbb{E}(\mathbf{1}_{\{w \leq x\}} - \Phi(z)) = \mathbb{P}(W \leq z) - \Phi(z).$$

Por lo que W seguirá una distribución normal estándar. ■

Como ya hemos mostrado, tenemos que $f'(w) - wf(w) = \mathbf{1}_{\{w \leq x\}} - \Phi(z)$ es un caso particular de la ecuación de Stein.

Ahora bien, si tenemos esta ecuación en su forma general dada como arriba, dada una función real h que es medible con valor esperada finito, entonces la solución $f = f_h$ dada para la ecuación de Stein será la introducida en el Lema 1.2 y tendremos el mismo resultado.

Pasamos ahora a estudiar las propiedades que cumple la solución general f_h (1.3) de la ecuación de Stein (1.2) a través del resultado siguiente.

Lema 1.3. *Para cualquier función real $h : \mathbb{R} \rightarrow \mathbb{R}$ absolutamente continua, sea f_h solución (1.3) de la ecuación de Stein (1.2). Si h es una función acotada*

$$\|f_h\| \leq \sqrt{\pi/2} \|h(\cdot) - \mathbb{E}h(Z)\| \quad y \quad \|f'_h\| \leq 2 \|h(\cdot) - \mathbb{E}h(Z)\|. \quad (1.4)$$

Si h es una función absolutamente continua entonces

$$\|f_h\| \leq 2 \|h'\|, \quad \|f'_h\| \leq \sqrt{\frac{2}{\pi}} \|h'\| \quad y \quad \|f''_h\| \leq 2 \|h'\|. \quad (1.5)$$

Demostración. Podemos ver la demostración en [2], Lema 2.4. ■

Tras la introducción del Lema de Stein junto con las propiedades sobre la solución de la ecuación de Stein que tenemos en el Lema 1.3., estamos en disposición de enunciar la siguiente proposición que es consecuencia directa de la teoría introducida.

Proposición 1.4. *Sea W una variable aleatoria y Z otra variable aleatoria que sigue una distribución normal estándar, entonces*

$$d_{\mathcal{H}}(W, Z) = \sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)| = \sup_{h \in \mathcal{H}} |\mathbb{E}(f'_h(W) - Wf_h(W))|. \quad (1.6)$$

Por lo que el método de Stein se basa en la comparación de esperanzas (tal y como nos muestra Stein con el título de [20]), y en definitiva, el objetivo a partir de ahora será buscar una cota para $\sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)|$ donde \mathcal{H} es una clase de funciones dada, lo curioso es que el lado derecho de la ecuación, que sólo incluye la variable aleatoria W será mucho más simple de comprobar que es pequeño y de acotar que el lado izquierdo. Así

la caracterización de la distribución normal se refleja en el hecho de que la cantidad $\mathbb{E}\{f'(W) - Wf(W)\}$ es relativamente fácil de demostrar que es pequeña a través de distintos métodos cuando la estructura de W es plausible de aproximar mediante una distribución normal.

Podemos notar como una de las principales diferencias y ventajas de el método de Stein respecto del método de funciones características es que el método de Stein es un método local tal y como hemos mostrado, ya que transformamos un problema global en uno de tipo local.

A continuación debemos dar sentido a la clase de funciones \mathcal{H} , para ello introducimos las métricas probabilísticas.

1.2. Métricas probabilísticas

Buscamos introducir el marco matemático sobre el que podremos estudiar los fundamentos del método de Stein. Necesitaremos introducir una noción de proximidad que nos permita medir distancias entre distribuciones asociadas a variables aleatorias, en esta primera parte de la memoria en particular, entre una distribución dada de interés y la distribución normal.

Dadas dos medidas de probabilidad \mathbb{P} y \mathbb{Q} , la métrica o distancia que asociamos será de un tipo particular que llamaremos integral, y por tanto será de la forma

$$d_{\mathcal{H}}(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathcal{H}} \left| \int h(x) d\mathbb{P}(x) - \int h(x) d\mathbb{Q}(x) \right|,$$

donde \mathcal{H} va a ser una familia de funciones test. En ocasiones se realizará un pequeño abuso de notación, expresando $d_{\mathcal{H}}(X, Y)$ en vez de $d_{\mathcal{H}}(\mathbb{P}, \mathbb{Q})$, siendo X, Y variables aleatorias cuyas leyes de probabilidad son respectivamente \mathbb{P}, \mathbb{Q} , tendremos entonces

$$d_{\mathcal{H}}(X, Y) = d_{\mathcal{H}}(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathcal{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|.$$

A continuación introducimos los principales ejemplos de métricas probabilísticas que siguen esta forma y que utilizaremos a lo largo de esta memoria, junto con algunas propiedades y relaciones entre ellas.

Métrica de variación total. Sean \mathbb{P}, \mathbb{Q} funciones de distribución asociadas a variables aleatorias W y Z , para el caso en que $\mathcal{H} = \mathcal{A}$, el conjunto formado por la familia de subconjuntos medibles A de \mathbb{R} y de tal forma que para cada A le asociamos la función test $1_A = h$ y tendremos que

$$\begin{aligned}
d_{VT}(\mathbb{P}, \mathbb{Q}) &= \sup_A |\mathbb{P}(A) - \mathbb{Q}(A)| = \sup_{h \in \mathcal{A}} \left| \int_A h d\mathbb{P} - \int_A h d\mathbb{Q} \right| \\
&= \sup_{A \in \mathcal{A}} |\mathbb{E}(h(w)) - \mathbb{E}(h(z))|.
\end{aligned}$$

Esta métrica es muy fuerte y no nos permite manejar aproximaciones de distribuciones continuas por discretas, esto se ve fácilmente si imaginamos una distribución binomial de $n+1$ elementos $\{0, 1, \dots, n\}$, que presenta $\mathbb{P}(\{0, 1, \dots, n\}) = 1$. Ahora bien, si consideramos la distribución normal, que es una distribución continua y por tanto con función de densidad, tenemos que $\mathbb{P}(\{0, 1, \dots, n\}) = 0$, luego la distancia entre ambas distribuciones será 1, que es la máxima distancia posible, lo cual muestra lo anteriormente señalado.

Esta distancia funciona bien en distribuciones de tipo discreto para medir la calidad de la aproximación. Representa la mayor distancia posible entre las probabilidades de que dos distribuciones de probabilidad sean asignadas al mismo evento.

Métrica de Kolmogorov. Tomamos ahora $\mathcal{H} = \{\mathbf{1}_{(-\infty, z]} : z \in \mathbb{R}\}$ y $h = \mathbf{1}_{(-\infty, z]}$ tal que $z \in \mathbb{R}$. Esta métrica será utilizada únicamente en medidas de probabilidad en \mathbb{R} .

$$d_K(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathcal{H}} \left| \int h d\mathbb{P} - \int h d\mathbb{Q} \right| = \sup_{z \in \mathbb{R}} |\mathbb{P}(-\infty, z] - \mathbb{Q}(-\infty, z]|.$$

Esta métrica nos da la distancia máxima entre funciones de distribución. Es por tanto la métrica de mayor interés para distribuciones continuas pues es la que se ajusta a la mayoría de modelos de estudio ajustados a casos reales, además de que una sucesión de distribuciones que converge en esta métrica a una distribución fijada implicará la existencia de convergencia débil o en distribución.

A las cotas calculadas por el método de Stein directamente bajo esta métrica las llamamos cotas de Berry-Essen, no las estudiamos aquí pues excede los objetivos de esta memoria. (Véase por ejemplo [1] y [2], donde se estudian en profundidad este tipo de cotas). Sin embargo, presenta el problema de tener un mal comportamiento para funciones no suaves, propiamente, las soluciones de Stein para la función h asociada son más difíciles de tratar.

Métrica de Wasserstein. Tomando $\mathcal{H} = \{h : \mathbb{R} \rightarrow \mathbb{R} : |h(x) - h(y)| \leq |x - y|\} := \text{Lip}(1)$ entonces

$$d_W(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathcal{H}} \left| \int h d\mathbb{P} - \int h d\mathbb{Q} \right|.$$

Esta métrica, también conocida como métrica de Kantorovich, presenta un gran interés ya que permite calcular cotas con relativa facilidad lo que hace que el método de Stein para esta distancia sea idóneo, pues tiene un buen comportamiento solo teniendo que exigir que la esperanza sea finita, es decir, la distancia de Wasserstein es una métrica en el conjunto de probabilidades en \mathbb{R} con esperanza finita.

Una de las características más importantes es que presenta relaciones directas con otras métricas, entre ellas para la distancia de Kolmogorov, esta relación la convierte en la métrica más utilizada en el contexto general de la aproximación para distribuciones continuas, pues nos permitirá conseguir cotas, si bien estas no serán óptima, para otras métricas de tratamiento más complicado como la de Kolmogorov, otorgándonos resultados que en ocasiones no serían posibles de obtener directamente utilizando esta métrica.

El siguiente lema clarifica la relación entre las tres métricas que acabamos de introducir.

Lema 1.5. *Tenemos las siguientes relaciones existentes entre las métricas, no demostraremos todas por el momento, pero estas relaciones serán una herramienta fundamental en capítulos posteriores.*

1. *Para variables aleatorias W y Z que toman valores en un espacio discreto Ω*

$$d_{VT}(W, Z) = \frac{1}{2} \sum_{w \in \Omega} |P(W = w) - P(Z = w)|.$$

2. *Sean X, Y variables aleatorias. $d_K(X, Y) \leq d_{VT}(X, Y)$.*

3. *Si la variable aleatoria Z tiene densidad con respecto a la medida de Lebesgue acotada por un cierto valor δ , entonces para cualquier variable aleatoria W tendremos*

$$d_K(W, Z) \leq \sqrt{2\delta d_W(W, Z)}.$$

Nota: Este resultado tendrá especial importancia cuando comencemos a desarrollar el cálculo de cotas para funciones suaves.

Demostración. El segundo punto es obvio, pues en la distancia de variación total los conjuntos de estudio son más grandes.

Para el primer punto, suponemos que tanto W como Z toman valores sobre un espacio discreto Ω . Definimos entonces un conjunto de puntos discreto de Ω de la forma siguiente

$$A := \{\omega \in \Omega : \mathbb{P}(W = \omega) \geq \mathbb{P}(Z = \omega)\},$$

y sea B un cierto subconjunto de Ω , tenemos entonces que basándonos en argumentos básicos de teoría de conjuntos y probabilística

$$\mathbb{P}(W \in B) - \mathbb{P}(Z \in B) \leq \mathbb{P}(W \in A \cap B) - \mathbb{P}(Z \in A \cap B) \leq \mathbb{P}(W \in A) - \mathbb{P}(Z \in A).$$

Razonando de forma análoga para el otro sentido

$$\mathbb{P}(Z \in B) - \mathbb{P}(W \in B) \leq \mathbb{P}(Z \in A \cap B) - \mathbb{P}(W \in A \cap B) \leq \mathbb{P}(Z \in A^c) - \mathbb{P}(W \in A^c),$$

por definición de el conjunto A .

Ahora bien, es obvio que si juntamos ambos términos tendremos

$$\mathbb{P}(W \in A) - \mathbb{P}(Z \in A) + \mathbb{P}(Z \in A) - \mathbb{P}(W \in A^c) = 1,$$

$$\mathbb{P}(Z \in A^c) - \mathbb{P}(W \in A^c) + \mathbb{P}(W \in A^c) - \mathbb{P}(Z \in A) = 1.$$

Luego los términos $\mathbb{P}(Z \in A^c) - \mathbb{P}(W \in A^c)$ y $\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)$ son el mismo y tomando $A = B$ se tiene que $|\mathbb{P}(W \in B) - \mathbb{P}(Z \in B)|$ es igual a su valor máximo, es decir, a su cota superior, por lo que llegamos a lo buscado ya que

$$d_{VT}(W, Z) = \frac{1}{2} (\mathbb{P}(W \in A) - \mathbb{P}(Z \in A) + \mathbb{P}(Z \in A^c) - \mathbb{P}(W \in A^c)).$$

Y podemos ver finalmente que esto es

$$d_{TV}(W, Z) = \frac{1}{2} \sum_{w \in \Omega} |P(W = w) - P(Z = w)|.$$

Lo que finaliza la demostración del segundo punto.

Demostramos ahora el tercer aserto, consideramos las funciones $h(w) = \mathbf{1}_{\{w \leq z\}}$ y definimos una función de suavizamiento $h_\alpha(w)$ que vendrá definida como sigue

$$h_\alpha(w) = \begin{cases} 1 & \text{si } w \leq z, \\ 0 & \text{si } w \geq z + \alpha, \\ \text{Lineal} & \text{si } z \leq w \leq z + \alpha. \end{cases}$$

Tenemos entonces

$$\begin{aligned}
\mathbb{E}(h(W)) - \mathbb{E}(h(Z)) &= \mathbb{E}(h(W)) - \mathbb{E}(h_\alpha(Z)) + \mathbb{E}(h_\alpha(Z)) - \mathbb{E}(h(Z)) \\
&\leq \mathbb{E}(h(W)) - \mathbb{E}(h_\alpha(Z)) + \delta\alpha/2 \\
&\leq \mathbb{E}(h_\alpha(W)) - \mathbb{E}(h_\alpha(Z)) + \delta\alpha/2 \\
&\leq d_W(W, Z)/\alpha + \delta\alpha/2
\end{aligned}$$

ya que $\alpha h_\alpha(x)$ es una función test relativa a la métrica de Wasserstein. Basta por tanto tomar $\alpha = \sqrt{2d_W(W, Z)}/\delta$ para demostrar la mitad de la desigualdad, la otra mitad será de prueba similar a través de un argumento equivalente. ■

El lema anterior nos muestra la idea de relación entre las métricas de Wasserstein y Kolmogorov, propiamente nos dice que la convergencia en distancia de Wasserstein implica convergencia en distancia de Kolmogorov, y en consecuencia, un resultado de velocidad de convergencia en distancia de Wasserstein nos dará un mismo resultado, aunque no óptimo, en velocidad de convergencia de Kolmogorov.

Pasamos ahora a la parte más técnica del capítulo, donde veremos como el método se aplica a distintas situaciones en las que nuestro estadístico de interés es una suma de variables aleatorias independientes.

1.3. El método de Stein para sumas de variables aleatorias independientes

Hemos visto que la idea principal en la que se basa el método de Stein para la obtención de cotas en el caso de la aproximación normal se basa en tratar de acotar el lado derecho de (1.6), para lo cual usamos la estructura de la variable aleatoria W y las propiedades de la solución f_h . Volviendo sobre la cuestión de el porqué utilizar la métrica de Wasserstein es tan adecuado, podemos fijarnos en que si tomamos como familia de funciones test \mathcal{H} , el conjunto de funciones con constante de Lipschitz uno, f_h , solución de la ecuación de Stein, es acotada con su segunda derivada acotada, mientras que para la familia de funciones de la métrica de Kolmogorov, sólo tendremos que la solución f_z tendrá primera derivada acotada, pero no doblemente diferenciable.

Comenzamos dando el resultado siguiente, que será un corolario del Lema 1.1 y el Lema 1.3. Este teorema resulta ser la base principal de la aproximación normal a través del método de Stein.

Teorema 1.6. *Si W es una variable aleatoria tal que $\mathbb{E}W^2 < \infty$ y Z sigue una distribución normal estándar, si definimos la familia de funciones $\mathcal{F} = \{f :$*

$\mathbb{R} \rightarrow \mathbb{R}/f$ es dos veces derivable con $\|f\|_\infty \leq 2$, $\|f'\|_\infty \leq \sqrt{2/\pi}$, $\|f''\|_\infty \leq 2$ entonces

$$d_W(W, Z) \leq \sup_{f \in \mathcal{F}} |\mathbb{E}[f'(W) - Wf(W)]|. \quad (1.7)$$

El hecho de que la variable aleatoria W es de cuadrado integrable, entonces la cantidad $\mathbb{E}[Wf(W)]$ está bien definida, (recordemos que f es Lipschitz). Los siguientes pasos que daremos en esta sección se dirigen a mostrar que tenemos las herramientas necesarias para aplicar este método a diversas situaciones y empezar a obtener cotas del error en la aproximación de ciertas leyes de distribución respecto de la normal. Nuestro objetivo por tanto va a ser mostrar que uno puede efectivamente acotar la cantidad que aparece en el lado derecho de (1.7). Por tanto estimar el valor de $\mathbb{E}h(W) - \mathbb{E}h(Z)$ para una gran variedad de situaciones según la forma de la variable aleatoria W , cuando h es una función test de la métrica de Wasserstein.

Comenzamos desarrollando un T.C.L. en su formulación clásica. Podemos ver cómo con la teoría introducida podemos dar ya de manera sencilla una primera cota para este caso, sin necesidad de introducir técnicas suplementarias en las que apoyarnos dentro del método de Stein.

Ilustramos a continuación qué ocurre en el caso de que la variable aleatoria W es suma de variables aleatorias independientes idénticamente distribuidas, por lo que nos centraremos en el estudio del Teorema Central del Límite en la formulación de Lindeberg-Levy, y que podemos reformular en términos matemáticos como presentamos a continuación.

Supongamos que X_1, X_2, \dots, X_n son una colección de variables aleatorias independientes idénticamente distribuidas, con media cero y varianza unitaria, y supongamos que se cumple $\mathbb{E}|X_1|^3 < \infty$; Definimos de la manera siguiente la variable aleatoria suma $W = n^{-1/2} \sum_{j=1}^n X_j$, nuestro objetivo es estimar el valor de $\mathbb{E}\{f'(W) - Wf(W)\}$ para una función f de la clase \mathcal{F} definida en el teorema 1.4.

Puesto que W es suma de variables aleatorias idénticamente distribuidas, podemos escribir

$$\mathbb{E}\{Wf(W)\} = n\mathbb{E}\{n^{-1/2}X_1f(W)\},$$

ya que $Wf(W) = \frac{X_1}{\sqrt{n}}f(W) + \dots + \frac{X_n}{\sqrt{n}}f(W)$, $(X_1, W) = (X_2, W)$ y tendremos por lo tanto $\mathbb{E}\{Wf(W)\} = n\mathbb{E}\{n^{-1/2}X_1f(W)\} = \sqrt{n}\mathbb{E}\{X_1f(W)\}$.

Podemos entonces escribir $W = X_1/\sqrt{n} + W_1$ con $W_1 = \sum_{j=2}^n X_j/\sqrt{n}$, siendo W_1 y X_1 independientes pues W_1 estará formado por X_2, X_3, \dots, X_n .

Mediante un desarrollo de Taylor llegamos a lo siguiente

$$\begin{aligned}\mathbb{E}\{Wf(W)\} &= n^{1/2}\mathbb{E}\{X_1f(W_1 + n^{-1/2}X_1)\} \\ &= n^{1/2}\mathbb{E}\{X_1(f(W_1) + n^{-1/2}X_1f'(W_1))\} + \eta_1.\end{aligned}$$

Tomaremos como máximo del valor absoluto de η_1 , el valor que se obtiene como resto en el desarrollo de Taylor de forma usual.

Así pues se tiene que $\eta_1 \leq \frac{1}{2} \frac{1}{\sqrt{n}} \mathbb{E}(X_1^3 f''(\tilde{W}_1))$. Con \tilde{W}_1 entre W_1 y $W_1 + \frac{X_1}{\sqrt{n}}$, y con $|\frac{f''(\tilde{W}_1)}{2}| \leq 1$.

Ahora bien, de la independencia de las variables X_1 y W_1 se sigue

$$\mathbb{E}(X_1 f(W_1)) = \mathbb{E}(X_1) \mathbb{E}(f(W_1)) = 0.$$

Usando ahora que $\mathbb{E}(X_1) = 0$ y puesto que $\text{Var}(X_1) = \mathbb{E}(X_1^2) = 1$, entonces

$$\mathbb{E}(X_1^2 f'(W_1)) = \mathbb{E}(X_1^2) \mathbb{E}(f'(W_1)) = \mathbb{E}(f'(W_1)).$$

Luego $\mathbb{E}(Wf(W)) = \mathbb{E}f'_h(W_1) + \eta_1$.

Aplicando otro desarrollo de Taylor esta vez tenemos

$$\mathbb{E}\{f'(W)\} = \mathbb{E}f'(W_1 + n^{-1/2}X_1) = \mathbb{E}f'(W_1) + \eta_2,$$

donde $|\eta_2| \leq \frac{1}{\sqrt{n}} \mathbb{E}(X_1 f''_h(\tilde{W}_1))$, como antes con \tilde{W}_1 entre W_1 y $W_1 + \frac{X_1}{\sqrt{n}}$.

Ahora bien puesto que $\mathbb{E}X_1 = 0$ y $\mathbb{E}X_1^2 = 1$ y como X_1 y W_1 son independientes llegaremos a que

$$|\mathbb{E}\{f'(W) - Wf(W)\}| \leq n^{-1/2} \{1 + \frac{1}{2} \mathbb{E}|X_1|^3\} \cdot 2 \quad f \in \mathcal{F}.$$

puesto que la familia de funciones test que estamos considerando es $\mathcal{H} := \text{Lip}(1)$ y se tiene de forma directa que $\|f''\|_\infty \leq 2$, y por tanto llegamos a que

$$d_W(W, Z) = \sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \frac{2 + \mathbb{E}|X_1|^3}{\sqrt{n}}.$$

Hemos conseguido por lo tanto una demostración alternativa aplicando el método de Stein para el Teorema Central del Límite clásico de Lindeberg-Levy, con unas hipótesis algo más restrictivas pues hemos supuesto momentos de tercer orden finito en vez de segundo orden, pero a cambio hemos obtenido una velocidad de convergencia, la cual nos dice que el error en la aproximación tiene a 0 a velocidad como mucho $\frac{1}{\sqrt{n}}$.

Así pues ya tenemos hemos conseguido una cota y por ende una velocidad de convergencia para el Teorema Central del Límite clásico, sin embargo este tipo de argumentos no será suficiente en casos más complejos y en situaciones

más generales que la estudiada, por ello necesitamos introducir las llamadas identidades de Stein, que nos permitirán desarrollar el llamado K-método y nos pondrán en disposición de estudiar un T.C.L. para sumas de variables aleatorias independientes que siguen una forma más general.

1.3.1. Identidades de Stein

Existen así varios enfoques o variaciones del argumento que acabamos de usar. Aborda primero el método más simple, pero que a la vez resulta ser una herramienta de gran utilidad en el método de Stein, llamado el método de la función K o K -método, que es la técnica más sencilla del método de Stein y se utiliza cuando W es un estadístico de interés expresado como suma de variables aleatorias independientes. En muchos textos podemos encontrar este método como una primera forma de aplicación del método de Stein, utilizándolo de modo introductorio sin nombrarlo de forma explícita.

El método se basa en argumentaciones similares a la vista en el apartado anterior, solo que trata de limitar los restos η_1 y η_2 que encontrábamos en el apartado anterior de una forma más sofisticada, en busca de mejores resultados, que se mostrarán en el apartado siguiente.

Tomamos $\xi_1, \xi_2, \dots, \xi_n$ variables aleatorias independientes con $\mathbb{E}\xi_i = 0$, $1 \leq i \leq n$ y tal que $\sum_{i=1}^n \mathbb{E}\xi_i^2 = 1$. Haciendo la correspondencia $\xi_i = n^{-1/2}X_i$. Aquí vendrá la primera mejora, ya que podemos obviar la condición de que las variables sean idénticamente distribuidas, sea

$$W = \sum_{i=1}^n \xi_i \quad y \quad W^{(i)} = W - \xi_i,$$

Introducimos la función en la que se basa el método, la llamada función K que vendrá dada como

$$K_i(t) = \mathbb{E}\{\xi_i(\mathbf{1}_{\{0 \leq t \leq \xi_i\}} - \mathbf{1}_{\{\xi_i \leq t < 0\}})\}, \quad (1.8)$$

y que cumple las siguientes propiedades de gran importancia

$$\int_{-\infty}^{\infty} K_i(t) dt = \mathbb{E}\xi_i^2 \quad y \quad \int_{-\infty}^{\infty} |t|K_i(t) dt = \frac{1}{2}\mathbb{E}|\xi_i|^3. \quad (1.9)$$

Sea $f \in \mathcal{F}$ la solución general de la ecuación de Stein, buscamos estimar el valor de

$$\mathbb{E}\{f'(W) - Wf(W)\},$$

donde seguimos el argumento que vemos a continuación, que será de uso usual en los resultados posteriores.

Por construcción tendremos que la variable aleatoria ξ_i es independiente de su variable aleatoria asociada W_i para cada caso $1 \leq i \leq n$.

Tenemos entonces

$$\mathbb{E}(Wf(W)) = \sum_{i=1}^n \mathbb{E}\{\xi_i f(W)\} = \sum_{i=1}^n \mathbb{E}\{\xi_i [f(W) - f(W^{(i)})]\},$$

por lo que transcribiendo esto a su forma integral lo que nos queda es

$$\begin{aligned} \mathbb{E}(Wf(W)) &= \sum_{i=1}^n \mathbb{E} \left(\xi_i \int_0^{\xi_i} f'(w_i + t) dt \right) \\ &= \sum_{i=1}^n \mathbb{E} \left\{ \int_{-\infty}^{\infty} f'(w_i + t) \xi_i (\mathbf{1}_{\{0 \leq t \leq \xi_i\}} - \mathbf{1}_{\{\xi_i \leq t < 0\}}) dt \right\}, \end{aligned}$$

luego en definitiva llegamos a la identidad

$$\mathbb{E}(Wf(W)) = \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E} \left\{ (f'(W^{(i)} + t)) \right\} K_i(t) dt. \quad (1.10)$$

Aplicando (1.10) y usando la independencia de los términos, como tenemos $\mathbb{E}f'(W) = \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}(f'(W)) K_i(t) dt$ llegamos al resultado siguiente que nos da la igualdad

$$\mathbb{E}(f'(W) - Wf(W)) = \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}(f'(w) - f'(w^{(i)} + t)) K_i(t) dt, \quad (1.11)$$

la cual será junto con (1.10) de vital importancia en la demostración de buenas aproximaciones normales.

En el apartado siguiente veremos una aplicación de el método de la K -función, que funcionará bien cuando W sea suma de variables aleatorias independientes, esto será ampliado en el siguiente capítulo, donde presentaremos métodos que aporten solución al problema de la independencia, así tendremos el método de los grafos aleatorios que nos dará una forma de encarar problemas en los que se presenta un tipo especial de dependencia, que sin embargo será de presencia común en una gran clase de problemas.

Utilizamos ahora el K -método que hemos introducido para presentar dos resultados que nos permitirán desarrollar Teoremas Centrales del Límite, en primer lugar tenemos el teorema siguiente.

Teorema 1.7. Sean $\xi_1, \xi_2, \dots, \xi_n$ variables aleatorias independientes tales que $\mathbb{E}\xi_i = 0$ y $\mathbb{E}|\xi_i|^3 < \infty$ para cada $1 \leq i \leq n$, tal que $\sum_{i=1}^n \mathbb{E}\xi_i^2 = 1$ y $W = \sum_{i=1}^n \xi_i$. Entonces tenemos

$$d_W(W, Z) \leq 3 \sum_{i=1}^n \mathbb{E}|\xi_i|^3. \quad (1.12)$$

Demostración. Para $f \in \mathcal{F}$ por (1.14) y el teorema del valor medio tendremos

$$\begin{aligned} |\mathbb{E}\{f'(W) - Wf(W)\}| &\leq \sum_{i=1}^n \int_{-\infty}^{\infty} |f'(W) - f'_h(W^{(i)} + t)| K_i(t) dt \\ &\leq 2 \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}(|t| + |\xi_i|) K_i(t) dt. \end{aligned}$$

Usando (1.12) y el K-método se sigue que como $\int_{-\infty}^{\infty} K_i(t) dt = \mathbb{E}\xi_i^2$ y $\int_{-\infty}^{\infty} |t| K_i(t) dt = 1/2 \mathbb{E}|\xi_i|^3$ llegamos finalmente tras sustituir a que

$$|\mathbb{E}\{f'(W) - Wf(W)\}| \leq 2 \sum_{i=1}^n (\mathbb{E}|\xi_i|^2/2 + \mathbb{E}|\xi_i| \mathbb{E}\xi_i^2) \leq 3 \sum_{i=1}^n \mathbb{E}|\xi_i|^3.$$

Lo que finaliza la prueba. ■

Sin embargo este resultado sigue suponiendo el buen comportamiento del tercer momento de la suma de variables, lo cual es un impedimento para la amplitud de aplicación que pretendemos transmitir mediante el método de Stein, podemos por tanto mejorar este resultado definiendo la cota en términos de elementos del Teorema Central del Límite de Lindeberg.

Teorema 1.8. Sean $\xi_1, \xi_2, \dots, \xi_n$ variables aleatorias independientes que cumplen $\mathbb{E}\xi_i = 0$ para cada $1 \leq i \leq n$, tal que $\sum_{i=1}^n \mathbb{E}\xi_i^2 = 1$ y $W = \sum_{i=1}^n \xi_i$. Entonces tenemos

$$d_W(W, Z) \leq 4(4\beta_2 + 3\beta_3), \quad (1.13)$$

donde

$$\beta_2 = \sum_{i=1}^n \mathbb{E}\xi_i^2 I_{\{|\xi_i| > 1\}} \quad y \quad \beta_3 = \sum_{i=1}^n \mathbb{E}|\xi_i|^3 I_{\{|\xi_i| \leq 1\}}. \quad (1.14)$$

Demostración. Utilizamos las desigualdades mostradas en el Lema 1.3. para mostrar que

$$|f'(W) - f'(W^{(i)} + t)| \leq \min(8, 2(|t| + |\xi_i|)) \leq 8(|t| \wedge 1 + |\xi_i| \wedge 1),$$

donde utilizamos la notación \wedge para denotar el mínimo de los números sobre los que se aplica. Sustituyendo la cota obtenemos

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq 8 \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}(|t| \wedge 1 + |\xi_i| \wedge 1) K_i(t) dt,$$

ahora

$$\int_{-\infty}^{\infty} (|t| \wedge 1) \{ \mathbf{1}_{[0,x]}(t) - \mathbf{1}_{[-x,0]}(t) \} dt = \begin{cases} \frac{1}{2}|x|t^2 + |x|(|x| - 1) & \text{si } |x| > 1; \\ \frac{1}{2}|x|^3, & \text{si } |x| \leq 1, \end{cases}$$

y tendremos por lo tanto

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbb{E}(|t| \wedge 1 + |\xi_i| \wedge 1) K_i(t) dt &= \mathbb{E}\{|\xi_i|(|\xi_i| - 1)\mathbf{1}_{\{|\xi_i| > 1\}}\} + \frac{1}{2}\mathbb{E}\{|\xi_i|(|\xi_i| \wedge 1)^2\} \\ &+ \mathbb{E}\{\xi_i^2 \mathbb{E}(|\xi_i| \wedge 1)\} = \mathbb{E}\{\xi_i^2 \mathbf{1}_{\{|\xi_i| > 1\}}\} - \frac{1}{2}\mathbb{E}\{|\xi_i| \mathbf{1}_{\{|\xi_i| > 1\}}\} \\ &+ \frac{1}{2}\mathbb{E}\{|\xi_i|^3 \mathbf{1}_{\{|\xi_i| \leq 1\}}\} + \mathbb{E}\{\xi_i^2 \mathbb{E}(|\xi_i| \wedge 1)\}, \end{aligned}$$

de lo cual se deduce que

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq 8 \left\{ \beta_2 + \frac{1}{2}\beta_3 + \sum_{i=1}^n \mathbb{E}\xi_i^2 \mathbb{E}(|t| \wedge 1) \right\}.$$

Ahora bien puesto que x^2 y $(x \wedge 1)$ son funciones crecientes de $x \geq 0$ se sigue que para cualquier variable aleatorias ξ tendremos

$$\mathbb{E}\xi_i^2 \mathbb{E}(|\xi_i| \wedge 1) \leq \mathbb{E}\{\xi^2(|\xi^2| \wedge 1)\} = \mathbb{E}\xi^2 \mathbf{1}_{\{|\xi_i| > 1\}} + \mathbb{E}|\xi|^3 \mathbf{1}_{\{|\xi_i| \leq 1\}}.$$

Ergo la suma final no es mayor que la suma de $\beta_2 + \beta_3$ lo cual completa la cota y nuestra demostración. ■

Aplicamos el Teorema anterior de tal forma que podremos demostrar el Teorema Central del Límite bajo las condiciones de Lindeberg.

Sean X_1, X_2, \dots, X_n variables aleatorias independientes con $\mathbb{E}X_i = 0$ y $\mathbb{E}X_i^2 = \sigma_i^2 < \infty$ para cada $1 \leq i \leq n$. Definiendo entonces

$$S_n = \sum_{i=1}^n X_i \quad \text{y} \quad B_n^2 = \sum_{i=1}^n \mathbb{E}X_i^2. \quad (1.15)$$

En vistas de poder aplicar el Teorema a través del método de Stein tendremos que definir

$$\xi_i = X_i/B_n \quad \text{y} \quad W = S_n/B_n. \quad (1.16)$$

Sean β_2 y β_3 variables definidas como en el teorema anterior en la ecuación (1.14), podemos notar que estos valores no son escogidos al azar, se basan en la expresión de la condición de Lindeberg que más tarde definiremos.

Es fácil observar que para cualquier ε de tal forma que $0 < \varepsilon < 1$ se tendrá que si sumamos ambas cantidades estas pueden ser acotadas de la forma siguiente

$$\begin{aligned}
\beta_2 + \beta_3 &= \frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| > B_n\}}\} + \frac{1}{B_n^3} \sum_{i=1}^n \mathbb{E}\{|X_i|^3 \mathbf{1}_{\{|X_i| \leq B_n\}}\} \\
&\leq \frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| > B_n\}}\} + \frac{1}{B_n^3} \sum_{i=1}^n B_n \mathbb{E}\{X_i^2 \mathbf{1}_{\{\varepsilon B_n \leq |X_i| \leq B_n\}}\} \\
&\quad + \frac{1}{B_n^3} \sum_{i=1}^n \varepsilon B_n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| < \varepsilon B_n\}}\} \\
&\leq \frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| > \varepsilon B_n\}}\} + \varepsilon.
\end{aligned}$$

Si se cumple la condición de Lindeberg que en nuestros términos vendrá dada como

$$\frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| > \varepsilon B_n\}}\} \rightarrow 0 \quad \text{si } n \rightarrow \infty \quad \forall \varepsilon > 0,$$

entonces puesto que hemos escogido un ε arbitrario, dado que por la desigualdad anterior sabemos que $\beta_2 + \beta_3 \leq \frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}\{X_i^2 \mathbf{1}_{\{|X_i| > \varepsilon B_n\}}\} + \varepsilon$ se dará que $\beta_2 + \beta_3 \rightarrow 0$ si $n \rightarrow \infty$.

Y consecuentemente, aplicando $d_K(W, Z) \leq \sqrt{2\delta d_W(W, Z)}$, como $d_W(W, Z) \leq 4(4\beta_2 + 3\beta_3)$ llegamos a que

$$d_K(W, Z) = \sup_z |\mathbb{P}(S_n/B_n \leq z) - \Phi(z)| \leq 8\sqrt{(\beta_2 + \beta_3)} \rightarrow 0 \quad \text{si } n \rightarrow \infty.$$

En particular tenemos el Teorema Central del Límite de Lindeberg:

Teorema 1.9. Sean X_1, \dots, X_n variables aleatorias independientes tales que $\mathbb{E}X_i = 1$ y $\text{Var}(X_i) = \sigma_i^2$ existe y es finita. Sea $B_n^2 = \sum_{i=1}^n \sigma_i^2$ y $W_n := \frac{\sum_{i=1}^n X_i}{B_n}$. Si se cumple la condición de Lindeberg siguiente

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{i=1}^n \mathbb{E}|X_i^2 \mathbf{1}_{\{|X_i| > \varepsilon B_n\}}| = 0 \quad \forall \varepsilon > 0,$$

entonces tenemos

$$W_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1).$$

La condición de Lindeberg es una condición de suficiencia (y bajo ciertas condiciones necesaria) para el Teorema Central del Límite, que nos asegura que una secuencia de variables aleatorias independientes que cumplen dicha condición cumplen un T.C.L.. Esto supone una mejora respecto a la versión clásica de Lindeberg-Levy ya que no necesita de la condición sobre las variables aleatorias de distribución idéntica.

En conclusión, hemos visto cómo el método de Stein nos permite desarrollar Teoremas Centrales del Límite para sumas de variables aleatorias independientes, sin necesidad de que estas estén idénticamente distribuidas. Además hemos conseguido cotas que muestran la velocidad de convergencia en los casos estudiados. Sin embargo, tenemos un funcional muy limitado, así pues, nuestro objetivo en los capítulos posteriores será explorar la posibilidad de ampliar el uso del método de Stein fuera del modelo rígido dado por la suma de variables aleatorias independientes.

Capítulo 2

EL MÉTODO DE STEIN PARA MODELOS DE DEPENDENCIA LOCAL

2.1. Aproximación normal de sumas de variables aleatorias localmente dependientes

En esta sección veremos una de las características que definen al método de Stein como una herramienta muy valiosa dentro de la Teoría de la Probabilidad, y es el hecho de que nos permite obtener cotas para el caso en que tengamos sumas de variables aleatorias no necesariamente independientes. Podemos ver que si relajamos la independencia a un tipo limitado de dependencia, podremos obtener cotas de forma similar al último apartado del capítulo anterior, a este tipo limitado de dependencia la denominaremos dependencia local, el concepto de dependencia local será dado más adelante. Sin embargo, debemos prestar especial atención a que este tipo de relación entre variables aleatorias puede ser definida de forma más o menos restrictiva, obteniendo así distintos resultados de cotas y convergencias, pero mostrando resultados aún en el caso más restrictivo de dependencia local, tal y como se mostrará a lo largo del capítulo (Nótese que veremos la versión más elemental, para ampliar resultados y ver más nociones de dependencia local véase [6]).

Para introducir la noción de dependencia local que vamos a utilizar, que aún puede ser ampliado pero que para la teoría que buscamos tratar en esta memoria será suficiente, comenzamos definiendo la noción de m -dependencia. Diremos que una colección de variables aleatorias $\xi_i, i \in \mathbb{Z}$ es m -dependiente si tiene la propiedad de que para cada i , el conjunto de variables aleatorias $\{\xi_j, j \leq i\}$ y $\{\xi_j, j > i + m\}$ son independientes. Obviamente, podemos considerar el caso especial en que tenemos 0-dependencia, pero esto no es más que la situación en que las variables aleatorias de nuestra

secuencia son independientes.

La dependencia local generaliza este concepto a variables aleatorias que están indexadas de forma arbitraria. Sin embargo, debido a sus múltiples aplicaciones y el interés en su estudio nos interesará sobre todo aplicar la noción de dependencia local al caso en que las variables aleatorias a estudiar vienen indexadas por los vértices de un grafo, y se cumple que las colecciones $\{\xi_i, i \in I\}$ y $\{\xi_j, j \in J\}$ son independientes cuando $I \cap J = \emptyset$ y el grafo no tiene aristas $\{i, j\}$ con $i \in I$ y $j \in J$, que como veremos se podrá generalizar en el método de los grafos de dependencia.

Sea \mathcal{J} un conjunto finito que utilizaremos para indexar, formado por n elementos. sea $\{\xi_i, i \in \mathcal{J}\}$ una colección de variables aleatorias de media cero y varianza $\sigma^2 < \infty$. Definimos $W = \sum_{i \in \mathcal{J}} \xi_i$, y asumimos que $\text{Var}(W) = 1$. Para $A \subset \mathcal{J}$ denotamos como $\xi_A = \{\xi_i : i \in A\}$ y sea $A^c = \{j \in \mathcal{J} : j \notin A\}$. Introduciremos los siguientes tipos de dependencia local.

(DL1) Para cada $i \in \mathcal{J}$ existe $A_i \subset \mathcal{J}$ de tal forma que ξ_i y $\xi_{A_i^c}$ son independientes.

(DL2) Para cada $i \in \mathcal{J}$ existe $A_i \subset B_i \subset \mathcal{J}$ de tal forma que ξ_i es independiente de $\xi_{A_i^c}$ y ξ_{A_i} lo es de $\xi_{B_i^c}$.

Podemos entonces denotar $\eta_i = \sum_{j \in A_i} \xi_j$ y $\tau_i = \sum_{j \in B_i} \xi_j$. Nos podemos fijar además, que para el caso de variables aleatorias independientes ξ_i podemos tomar $A_i = B_i = \{i\}$ en cuyo caso tendremos $\eta_i = \tau_i = \xi_i$.

Teorema 2.1. *Sea Z una variable aleatoria con distribución probabilística $\mathcal{N}(0, 1)$, y sea W una suma de variables aleatorias de media cero y varianza finita que siguen un cierto tipo de dependencia local, asumiendo que $\text{Var}(W) = 1$, entonces tenemos las cotas siguientes según el tipo de dependencia local que sigan las variables aleatorias*

Cota de error para la dependencia (DL1)

$$d_W(W, Z) \leq 4\mathbb{E} \left| \sum_{i \in \mathcal{J}} \{\xi_i \eta_i - \mathbb{E}(\xi_i \eta_i)\} \right| + \sum_{i \in \mathcal{J}} \mathbb{E} |\xi_i \eta_i^2|. \quad (2.1)$$

Cota de error para la dependencia (DL2)

$$d_W(W, Z) \leq 2 \sum_{i \in \mathcal{J}} (\mathbb{E} |\xi_i \eta_i \tau_i| + |\mathbb{E}(\xi_i \eta_i)| \mathbb{E} |\tau_i|) + \sum_{i \in \mathcal{J}} \mathbb{E} |\xi_i \eta_i^2|. \quad (2.2)$$

Nota: Esta cota sería también válida para variables aleatorias independientes, descrito como en el caso mostrado arriba, cuyo valor es $5 \sum_{i \in \mathcal{J}} \mathbb{E} |\xi_i|^3$, el cual resulta mayor que la cota directa que obtuvimos en el Teorema 1.5. .

Demostración. En primer lugar comenzamos obteniendo las identidades de Stein de forma semejante a como hacíamos en (1.10) y (1.11). Sea $f \in \mathcal{F}$ entonces

$$\mathbb{E}\{Wf(W)\} = \sum_{i \in \mathcal{J}} \mathbb{E}\xi_i f(W) = \sum_{i \in \mathcal{J}} \mathbb{E}\xi_i [f(W) - f(W - \eta_i)], \quad (2.3)$$

por la independencia de ξ_i y $W - \eta_i$ entonces

$$\mathbb{E}\{Wf(W)\} = \sum_{i \in \mathcal{J}} \mathbb{E}\{\xi_i [f(W) - f(W - \eta_i) - \eta_i f'(W)]\} + \mathbb{E}\left\{\left(\sum_{i \in \mathcal{J}} \xi_i \eta_i\right) f'(W)\right\}. \quad (2.4)$$

Ahora como $\mathbb{E}\xi_i = 0$ para todo i , ahora bien para (DL1) se sigue que

$$1 = \mathbb{E}W^2 = \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} \mathbb{E}\{\xi_i \xi_j\} = \sum_{i \in \mathcal{J}} \mathbb{E}\{\xi_i \eta_i\},$$

lo que nos da

$$\begin{aligned} \mathbb{E}\{f'(W) - Wf(W)\} &= -\mathbb{E}\left(\sum_{i \in \mathcal{J}} \{\xi_i \eta_i - \mathbb{E}(\xi_i \eta_i)\} f'(W)\right) \\ &\quad - \sum_{i \in \mathcal{J}} \mathbb{E}\{\xi_i [f(W) - f(W - \eta_i) - \eta_i f'(W)]\}. \end{aligned} \quad (2.5)$$

Aplicando entonces (1.5), puesto que $\|f'\| \leq \sqrt{\frac{2}{\pi}} \|h'\| \leq 4\|h'\|$ (Usamos esto para un resultado más limpio ya que las constantes no suponen grandes problemas) y $\|f''\| \leq 2\|h'\|$. Como hemos considerado $h \in Lip(1)$ podemos tomar directamente como en ocasiones anteriores $\|h'\| = 1$ y entonces se sigue de (2.5) y del desarrollo de Taylor que

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \left\{ 4\mathbb{E}\left|\sum_{i \in \mathcal{J}} \{\xi_i \eta_i - \mathbb{E}(\xi_i \eta_i)\}\right| + \sum_{i \in \mathcal{J}} \mathbb{E}|\xi_i \eta_i^2| \right\},$$

y con esto tenemos probada la cota para la dependencia de tipo (DL1).

Ahora bien en el caso de (DL2) tendremos que $f'(W - \tau_i)$ y $\xi_i \eta_i$ son independientes para cada $i \in \mathcal{J}$. Usando la expresión (2.5) podemos escribir

$$\begin{aligned}
& |\mathbb{E}h(W) - \mathbb{E}h(Z)| \\
& \leq \left| \mathbb{E} \sum_{i \in \mathcal{J}} \{\xi_i \eta_i - \mathbb{E}(\xi_i \eta_i)\} (f'(W) - f'(W - \tau_i)) \right| + \|h'\| \sum_{i \in \mathcal{J}} \mathbb{E}|\xi_i \eta_i^2| \\
& \leq \left\{ 2 \sum_{i \in \mathcal{J}} (\mathbb{E}|\xi_i \eta_i \tau_i| + |\mathbb{E}(\xi_i \eta_i)| \mathbb{E}|\tau_i|) + \sum_{i \in \mathcal{J}} \mathbb{E}|\xi_i \eta_i^2| \right\}
\end{aligned}$$

tal y como buscábamos probar. ■

El Teorema 2.1 nos proporciona un método útil para probar Teoremas Centrales del Límite para muchos funcionales definidos sobre grafos aleatorio (en un sentido amplio). Exponemos algunos ejemplos a continuación.

1. Grafos de dependencia.

Describimos el método de Grafos de dependencia, en este caso consideramos un conjunto de variables aleatorias $\{X_i, i \in \nu\}$ que van a encontrarse indexadas por los vértices de un cierto grafo $G = (\nu, \varepsilon)$, diremos entonces que G es un grafo de dependencia si cumple que para cualquier par de subconjuntos disjuntos Γ_1, Γ_2 de ν , es decir, conjuntos de vértices de tal forma que no haya ninguna arista en ε que una Γ_1 y Γ_2 , los conjuntos de variables aleatorias $\{X_i, i \in \Gamma_1\}$ y $\{X_i, i \in \Gamma_2\}$ son independientes.

Sea D el grado máximo de G , es decir, el número máximo de aristas que tienen incidencia en un vértice. Denotaremos

$$A_i = \{i\} \cup \{j \in \nu : \text{existe una arista que conecta } j \text{ e } i\}$$

y $B_i = \cup_{j \in A_i} A_j$. Entonces el conjunto de variables aleatorias $\{X_i, i \in \nu\}$ cumple dependencia local del tipo (DL2) descrita anteriormente y por tanto podemos aplicar la cota de error que hemos planteado para la aproximación normal.

2. Número de extremos locales en un grafo.

Consideramos un grafo $G = (\nu, \varepsilon)$ que puede no ser un grafo de dependencia, sean $\{Y_i, i \in \nu\}$ variables aleatorias continuas independientes e idénticamente distribuidas. Para $i \in \nu$ definiremos la variable como función indicadora

$$X_i = \begin{cases} 1 & \text{si } Y_i > Y_j \forall j \in N_i \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Donde $N_i = \{j \in \nu : \{i, j\} \in \varepsilon\}$ de tal forma que $X_i = 1$ indicaba que Y_i es un máximo local. Sea $W = \sum_{i \in \nu} X_i$ es el número de máximos locales. Definimos

$$A_i = \{i\} \cup \left\{ \bigcup_{j \in N_i} N_j \right\} \quad \text{y} \quad B_i = \bigcup_{j \in A_i} A_j$$

Entonces $\{X_i, i \in \nu\}$ cumple (DL2) y se cumple la cota para variables aleatorias que se encuentran ligadas bajo la dependencia local de tipo (DL2). Esto nos permite concluir que

$$d_W \left(\frac{W - \mathbb{E}W}{\sqrt{\text{Var}(W)}}, Z \right) \leq 2 \sum_{i \in \mathcal{J}} (\mathbb{E}|\xi_i \eta_i \tau_i| + |\mathbb{E}(\xi_i \eta_i)| |\mathbb{E}|\tau_i|) + \sum_{i \in \mathcal{J}} \mathbb{E}|\xi_i \eta_i^2|.$$

Esto permitiría probar un TCL para W si la estructura del grafo es adecuada para acotar la lado derecho en la desigualdad anterior. Para más ejemplos véase [16].

Hemos introducido dos ejemplos básicos de situaciones en las que el método de Stein nos permite calcular Teoremas Centrales del Límite para funcionales definidos sobre grafos aleatorios. A continuación desarrollamos más en profundidad una situación descrita en el ejemplo 1, y que propiamente en la literatura se conoce como método de los grafos de dependencia. El interés principal viene dado porque tratamos un problema que a priori es de tipo probabilístico, pero vemos como el método de Stein transforma este problema en un problema combinatorio que es fácil de analizar con técnicas elementales.

2.2. Aplicación: un T.C.L. en el modelo de Erdős-Renyi

Una de las situaciones de estudio clásicas del método de Stein en el caso de variables aleatorias localmente dependientes para grafos de dependencia (véase [3] ó [16]) será el caso en que nos encontremos con un cierto grafo aleatorio $K(n, p)$ en el conjunto de vértices $N := \{1, 2, \dots, n\}$, donde la aleatoriedad viene dada por la aparición o no de aristas para unos nodos fijos, viniendo la aparición de aristas determinada por una cierta probabilidad p .

En esta situación descrita, llegamos a una de las aplicaciones más estudiadas dentro del método de Stein, que será el estudio del número de subgrafos inducidos por el grafo aleatorio $K(n, p)$ isomorfos a un cierto grafo dado G con k vértices. Sea $\mathcal{J} := \{\{i_1, i_2, \dots, i_k : 1 \leq i_1 < i_2 < \dots < i_k \leq n\}\}$ el conjunto que indexa los subconjuntos de k vértices de N , podremos demostrar cómo el número de subconjuntos de k vértices unidos por aristas entre sí sigue una distribución normal en ciertas situaciones.

Uno de estos grafos aleatorios será el llamado Grafo de Erdős-Renyi, que denotaremos de forma usual como $G(n, p)$ que será un modelo de generación de grafos aleatorio. n representa el número de vértices en el grafo y p es la probabilidad de arista. Para cada par de vértices distintos v y w , p va a representar la probabilidad de que el grafo presente la arista (v, w) , al ser un grafo no dirigido, nuestra notación será de ahora en adelante $\{v, w\}$, representando esta notación que $(v, w) = (w, v)$. La presencia de cada arista es estadísticamente independiente de todas las demás aristas.

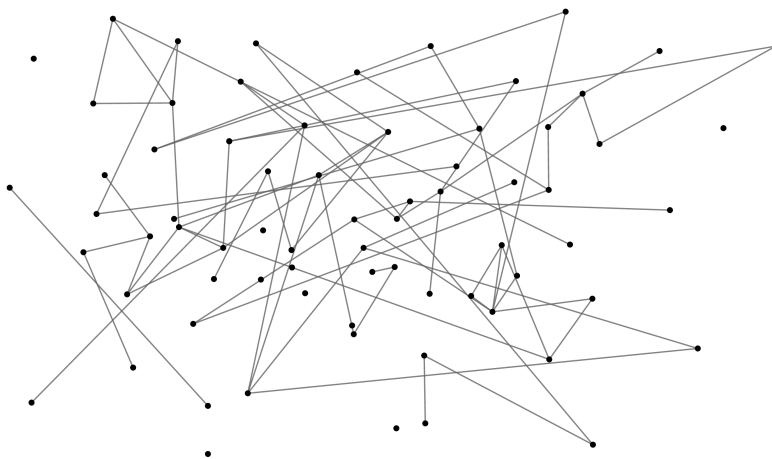


Figura 2.1: Simulación de un grafo de Erdos-Renyi aleatorio.

La construcción de un grafo de Erdős-Renyi es por tanto sencilla, consideramos una red de n nodos sin conectar, distribuidos de forma aleatoria, y

vamos enlazando de forma aleatoria en pasos sucesivos dos vértices a través de una arista, descartando vértices ya enlazados de tal forma que no puede ocurrir que vuelvan a enlazarse entre ellos. Así pues tras M pasos habrá M conexiones establecidas y según este número tendremos una mayor o menor conexión entre vértices, La particularidad e interés de este grafo es que unir los vértices a través de este proceso descrito genero propiedades específicas en cuanto a la distribución de las aristas.

El objeto de nuestro estudio será el descrito anteriormente, bajo la situación en que tenemos $G(n, p)$ un grafo generado por el modelo de Erdős-Renyi, buscaremos ver que el número de triángulos, entendidos como el número de subconjuntos de 3 vértices que presentan una arista que les une dos a dos, sigue una distribución normal cuando el número de vértices es suficientemente alto, es decir, determinaremos la existencia de un Teorema Central del límite para el número de triángulos en un grafo de Erdős-Renyi.

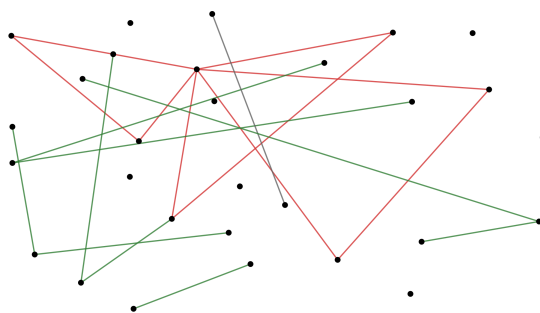


Figura 2.2: Grafo según el modelo de Erdős-Renyi. En rojo; Triángulos generados. En verde; Aristas que no pertenecen a ningún triángulo.

Buscamos determinar un Teorema Central del Límite para grafos aleatorios que siguen el modelo de Erdős-Renyi a través de la aplicación del método de Stein. En concreto, vamos a determinar que el número de triángulos que podemos encontrar en un grafo aleatorio que sigue el modelo generador indicado se puede aproximar por la distribución normal. Para demostrar esto aplicaremos el método de los grafos de dependencia, encuadrado dentro del estudio de variables aleatorias con dependencia local a través del método de Stein.

Sea $G = G(n, p)$ un grafo aleatorio generado por el modelo de Erdős-Renyi, si denominamos W_n al número de triángulos en el grafo total, vamos a ver que el número de triángulos en dicho grafo cumplirá un Teorema Central del Límite de la forma siguiente

$$W = \frac{W_n - \mathbb{E}W_n}{\sqrt{\text{Var}(W_n)}} \xrightarrow{d} \mathcal{N}(0, 1). \quad (2.6)$$

Denotamos por n el número de vértices o nodos asociados al grafo G y denotemos $\{i, j\}$ la arista que une el vértice i con el vértice j y tendremos entonces que la probabilidad de que en el grafo encontremos dicha arista vendrá dada por

$$p = \mathbb{P}(\{i, j\} \in G) \quad 0 \leq p \leq 1. \quad (2.7)$$

Vamos a definir la variable aleatoria W_n como el número de triángulos que realmente se encuentran en el grafo, que en términos de elementos del grafo será el conjunto formado por aquellos tríos de vértices que verifican estar todos ellos en conexión por alguna arista en $G(n, p)$. Podemos escribir la variable aleatoria a estudiar de la forma siguiente

$$W_n = \sum_{\{i,j,k\} \subset \{1,\dots,n\}} \mathbf{1}_{\{B_{\{i,j\}}=1, B_{\{i,k\}}=1, B_{\{k,j\}}=1\}}, \quad (2.8)$$

siendo $B_{\{i,k\}} = 1$ una variable tipo bernoulli de parámetro p de tal forma que si tenemos una arista entre los vértices i y j y 0 si no.

Ahora bien, para el análisis posterior será conveniente reescribir W_n como una suma indexada por los elementos de \mathcal{T}

Sea \mathcal{J} el conjunto de triángulos posibles en el grafo completo. Vamos a reformular este conjunto de tal forma que será un conjunto indexador, que vendrá definido como

$$\mathcal{J} = \{i = \{i_1, i_2, i_3\} : 1 \leq i_1 < i_2 < i_3 \leq n\}, \quad (2.9)$$

el cual podemos entender como el conjunto formado por todos los subconjuntos de 3 vértices del grafo G , de tal manera que a cada conjunto de 3 vértices le asigna de forma arbitraria pero predeterminada un cierto índice.

Así, si denotamos los vértices en forma numérica, estará formado por todos los conjuntos de 3 elementos de $\{1, 2, \dots, n\}$ sin importancia del orden, es decir, $\{1, 2, 3\} = \{1, 3, 2\} = \dots = \{3, 2, 1\}$

Podemos entonces reescribir la variable aleatoria antes definida en la forma siguiente

$$W_n = \sum_{i \in \mathcal{J}} \mathbf{1}_{\{i = \{i_1, i_2, i_3\} \in G(n, p)\}}. \quad (2.10)$$

Ahora bien, esto nos lleva a que $|\mathcal{J}| = N = \binom{n}{3}$ define el número de conjuntos de 3 elementos que se pueden formar en el grafo.

Definiendo entonces la variable aleatoria $\xi_i = \mathbf{1}_{\{i=\{i_1, i_2, i_3\} \in \mathcal{T}\}}$ tendremos como $p = \mathbb{P}(\{i, j\} \in G) = \mathbb{P}(B_{\{i, j\}} = 1)$ se tiene por independencia que $\mathbb{E}\xi_i = p^3$.

Y por lo tanto

$$\mathbb{E}W_n = \binom{n}{3} p^3. \quad (2.11)$$

Realizamos a continuación el cálculo de la varianza.

$$\text{Var}(W_n) = \text{Cov}(\xi_n, \xi_n) = \sum_{S \in \mathcal{J}, T \in \mathcal{J}} \text{Cov}(\mathbf{1}(S \in G(n, p)), \mathbf{1}(T \in G(n, p))).$$

- a) $S \cap T = \emptyset \Rightarrow \text{Cov}(I_S, I_T) = 0,$
- b) $\text{Card}(S \cap T) = 1 \Rightarrow \text{Cov}(I_S, I_T) = 0,$
- c) $\text{Card}(S \cap T) = 2 \Rightarrow \text{Cov}(I_S, I_T) = p^5(1 - p),$
- d) $\text{Card}(S \cap T) = 3 \Rightarrow \text{Cov}(I_S, I_T) = p^3(1 - p^3).$

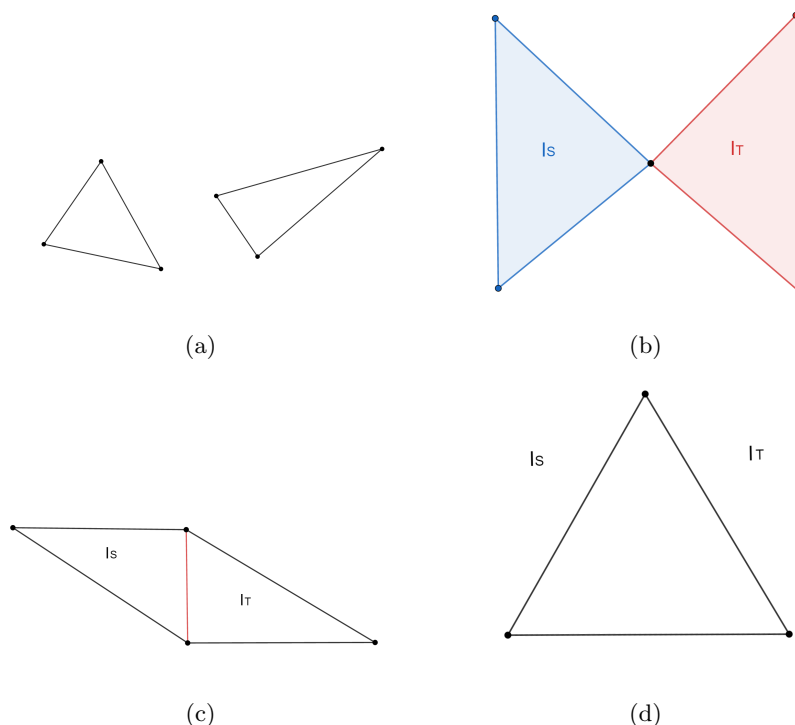


Figura 2.3: Esquema para el cálculo de la varianza.

El cálculo es sencillo, obviamente en el caso a) y b) tenemos que los triángulos que se forman son independientes, puesto que no hay ninguna arista en común luego su covarianza es 0.

En los otros casos el valor de la covarianza vendrá dado por la expresión $p^{6-e(T)}(1-p^{e(T)})$ con $e(T)$ el número de aristas en común de los triángulos.

Tratamos ahora de contar cuantos sumandos habrá de los tipos c) y d).

En el caso c) tendremos que hay $\binom{n}{2}$ posibles aristas que unan 2 vértices, y luego habrá $(n-2)$ posibles elecciones para el vértice que uniremos a esos 2 vértices escogidos para formar un triángulo, y acto seguido $(n-3)$ posibles vértices para formar el otro triángulo, lo que nos deja que el número total de elementos representados en c) y cuya covarianza es $p^5(1-p)$ será $\binom{n}{2}(n-2)(n-3)$.

El último caso resulta inmediato pues tenemos $\binom{n}{3}$ posibles triángulos en $G(n, p)$.

De ahora en adelante supongamos que se cumple $n \geq 3$ (Nótese que $n=3$ no tiene interés).

Agrupando las cantidades halladas en el cálculo de la varianza tendremos

$$\begin{aligned} \sigma_n^2 = Var(\xi_n) &= \binom{n}{3} p^3(1-p^3) + \binom{n}{2} (n-2)(n-3) p^5(1-p) \\ &= \frac{n(n-1)(n-2)}{6} p^3(1-p^3) + \frac{n(n-1)(n-2)(n-3)}{2} p^5(1-p) \\ &= \frac{n(n-1)(n-2)}{2} p^3 \left(\frac{(1-p^3)}{3} + (n-3)p^2(1-p) \right). \end{aligned}$$

Queremos aplicar el método de los grafos de dependencia a través de la expresión (2.10) y necesitamos un grafo cuya estructura se ajuste a la de nuestro grafo aleatorio y que sea un grafo de dependencia. Para ello en primer lugar necesitamos fijar un conjunto de vértices y aristas, que representen las de nuestro grafo aleatorio y que dicho grafo resultante sea un grafo de dependencia.

Recuperamos nuestro conjunto indexador, que habíamos utilizado para definir la variable aleatoria a estudiar, y que venía dado como

$$\mathcal{I} = \{i = (i_1, i_2, i_3) : 1 \leq i_1 < i_2 < i_3 \leq 1\},$$

tendremos que el grafo de dependencia que denominamos $\mathcal{G} = \mathcal{G}(\nu, \varepsilon)$ tendrá como conjunto de vértices el indexado por \mathcal{I} , $\nu(\mathcal{G}) = \mathcal{I}$ y el de aristas vendrá dado por

$$\varepsilon(\mathcal{G}) = \{(i, j) : |\{i_1, i_2, i_3\} \cap \{j_1, j_2, j_3\}| \geq 2\},$$

Para aplicar la teoría anterior bastará comprobar que efectivamente \mathcal{G} cumple la propiedad fundamental para ser un grafo de dependencia, que era

la siguiente

Propiedad: Sean $\{X_i : i \in \nu\}$ variables aleatorias indexadas por los vértices del grafo $\mathcal{G} = \mathcal{G}(\nu, \varepsilon)$, entonces \mathcal{G} es un grafo de dependencia si para cualquier par de conjuntos disjuntos Γ_1, Γ_2 en ν de tal forma que no haya aristas en ε que tenga un punto inicial en Γ_1 y otro en Γ_2 , el conjunto $\{X_i : i \in \Gamma_1\}$ y $\{X_j : j \in \Gamma_2\}$ son independientes.

Esto es sencillo de comprobar gracias a la construcción que hemos hecho, pues si tenemos Γ_1 y Γ_2 disjuntos como conjuntos de tres vértices en ν , si no tenemos una arista que les una, esto implica que los $i, j \in \Gamma_1, \Gamma_2$ cumplen

$$|\{i_1, i_2, i_3\} \cap \{j_1, j_2, j_3\}| < 2,$$

luego por tanto tendremos que como máximo tiene uno de los vértices en común, tal y como vemos en el dibujo (b) de la figura 2.3, y por tanto los triángulos serán independientes por construcción.

Podemos entonces aplicar la cota (2.2) y desarrollar el Teorema Central del Límite que estamos buscando.

Antes de eso, definiremos los llamados vecindarios o entornos, ya que los necesitamos a la hora del estudio de la cota.

Definimos para todo $i = \{i_1, i_2, i_3\} \in \mathcal{J}$

$$A_i = \{k = \{k_1, k_2, k_3\} \in \mathcal{J} : |\{i_1, i_2, i_3\} \cap \{j_1, j_2, j_3\}| \geq 2\}.$$

Siendo sencillo ver que $|A_i| = 1 + 3(n - 3) \forall i \in \mathcal{J}$ pues basta para determinar este calculo incluir el índice $j = i$ caso en que los 3 vértices coinciden, y después basta con componer los demás índices j dejando fijos 2 vértices y variando el que resta, lo que nos da $(n - 3)$ índices, que son $3(n - 3)$ teniendo en cuenta que cada vez fijaremos 2 vértices distintos y el otro variará.

De igual forma podemos construir los B_i , que serán aquellos elementos que cumplen la condición de pertenecer al entorno, o al menos tener en común uno de los vértices que forman el índice, luego definiremos

$$B_i = \{k = \{k_1, k_2, k_3\} \in \mathcal{J} : |\{i_1, i_2, i_3\} \cap \{j_1, j_2, j_3\}| \geq 1\}.$$

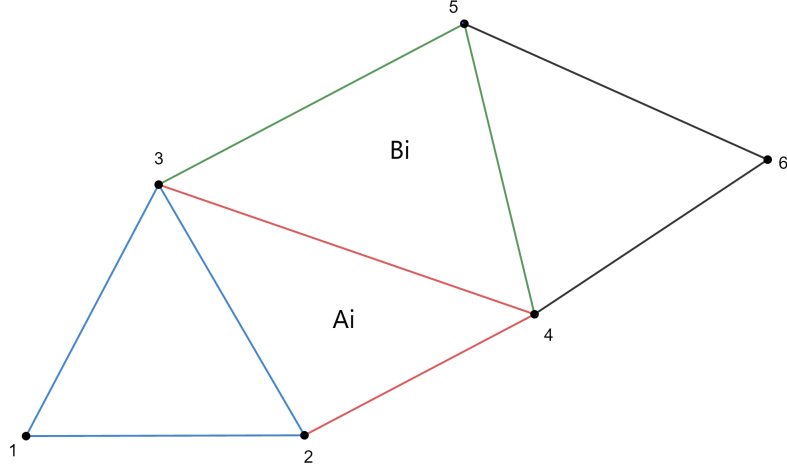


Figura 2.4: Esquema de vecindarios.

Resta estudiar qué ocurre en los distintos casos que se nos presentan para determinar bajo qué condiciones podremos asegurar que se cumple un teorema central del límite para la suma de los triángulos, es decir.

$$W = \frac{W_n - \mathbb{E}W_n}{\sqrt{\text{Var}(W_n)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Consideramos ahora para realizar nuestro estudio que p es fijo y cumple $p \leq \frac{1}{2}$ ya que en cualquier otro caso basta tomar $(1 - p)$ y que tenemos $np \rightarrow \infty$ luego $\mathbb{E}W_n \rightarrow \infty$.

Vamos a continuación a realizar el cálculo de acotaciones para cada uno de los términos que tienen influencia en la cota (2.2) que recordamos venía expresada como

$$d_W(\mathcal{L}(W), \mathcal{N}(0, 1)) \leq 2 \sum_{i \in \mathcal{J}} (\mathbb{E}|\xi_i \eta_i \tau_i| + |\mathbb{E}(\xi_i \eta_i)| \mathbb{E}|\tau_i|) + \sum_{i \in \mathcal{J}} \mathbb{E}|\xi_i \eta_i^2|.$$

Esta cota viene expresada para variables centradas, luego tomamos $\bar{\xi}_i = \xi_i - \mathbb{E}\xi_i$ y realizamos los cálculos con esta variable.

Tenemos en cuenta en todo momento que $\eta_i = \sum_{j \in A_i} \bar{\xi}_j$ y $\tau_i = \sum_{j \in B_i} \bar{\xi}_j$.

Necesitamos así calcular cada uno de los términos y sus órdenes, comenzaremos entonces por el que a todas luces será más sencillo, que será el dado por $\delta_2 = \sum_{i \in \mathcal{J}} |\mathbb{E}(\bar{\xi}_i \eta_i)| \mathbb{E}|\tau_i|$.

Así pues comenzamos con el cálculo de

$$|\mathbb{E}(\bar{\xi}_i \eta_i)| = \mathbb{E}(\bar{\xi}_i \eta_i) = \mathbb{E}(\bar{\xi}_j \sum_{j \in A_i} \bar{\xi}_i) = \mathbb{E}(\bar{\xi}_i^2) + \sum_{\substack{j \in A_i \\ i \neq j}} \mathbb{E}(\bar{\xi}_i \bar{\xi}_j),$$

cuyo cálculo va a ser inmediato pues tendremos que si tomamos $\bar{\xi}_i = I_{ABC} - p^3$ y $\bar{\xi}_i = I_{BDE} - p^3$ como variables centradas entonces cálculos sencillos nos llevan a

$$\begin{aligned} \mathbb{E}(\bar{\xi}_i^2) &= \mathbb{E}((I_{ABC} - p^3)(I_{ABC} - p^3)) \\ &= \mathbb{E}(I_{ABC} - p^3 I_{ABC} - p^3 I_{ABC} + p^6) = p^3(1 - p^3), \end{aligned}$$

$$\mathbb{E}(\bar{\xi}_i \bar{\xi}_j) = \mathbb{E}(I_{ABCDE} - p^3 I_{ABC} - p^3 I_{BDE} + p^6) = p^5(1 - p),$$

y por lo tanto

$$\mathbb{E}(\bar{\xi}_i \eta_i) = p^3(1 - p^3) + 3(n - 3)p^5(1 - p).$$

Nos falta calcular el término restante, para el cual haremos un análisis combinatorio.

$$\mathbb{E}|\tau_i| = \mathbb{E} \left| \sum_{k \in B_i} \bar{\xi}_k \right|,$$

$$\mathbb{E}|\tau_i| \leq (\mathbb{E}(\tau_i)^2)^{1/2},$$

$$\mathbb{E}(\tau_i)^2 \leq \sum_{k \in B_i} \sum_{m \in B_i} (\mathbb{E}(\bar{\xi}_k \bar{\xi}_m)).$$

Ahora bien, la aplicación de una propiedad como la del valor medio no nos dará los resultados que buscamos. Así pues necesitamos hacer un análisis combinatorio del problema. Para ello veremos el número de términos que son nulos. Serán todos aquellos en que ambos triángulos no compartan al menos 1 arista común.

Es inmediato que el número de términos no nulos sigue un orden de n^2 . Tendremos por tanto un orden de n^2 términos iguales, y por tanto de la forma $p^3(1 - p^3)$ y a lo sumo n^3 de la forma $p^5(1 - p)$.

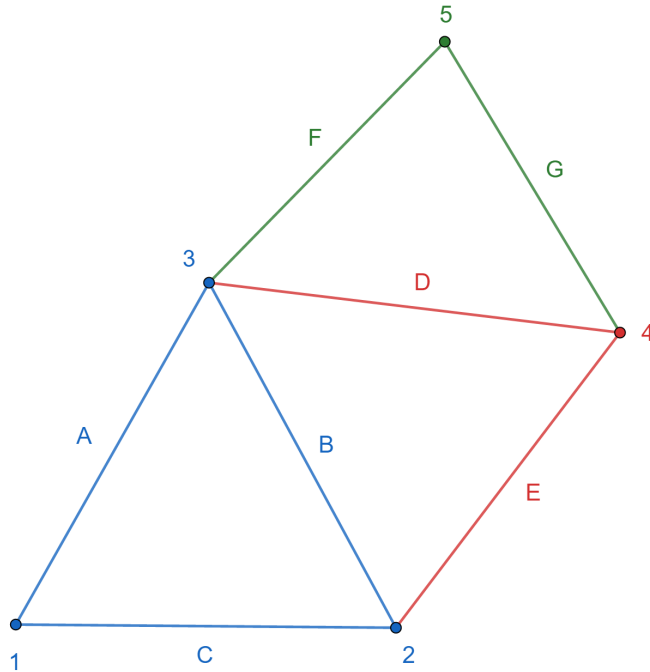


Figura 2.5: Esquema del análisis combinatorio sobre los triángulos.

Por tanto

$$\sum_{k \in B_i} \sum_{m \in B_i} (\mathbb{E}(\bar{\xi}_k \bar{\xi}_m) \leq C(n^3 p^5 (1-p) + n^2 p^3 (1-p^3)))$$

$$\Rightarrow \mathbb{E}|\tau_i| \leq (\mathbb{E}(\tau_i)^2)^{1/2} \leq C_1 n^{3/2}.$$

Hemos hecho el análisis con las variables centradas, sin embargo debemos estandarizar. Para ello necesitamos el valor de σ^3 . Hacemos el cálculo del orden

$$\sigma^3 = n^{9/2} p^{9/2} + n^5 p^{11/2} + n^{11/2} p^{13/2} + n^6 p^{15/2} \Rightarrow \sigma^3 \sim n^6,$$

luego tenemos que la cota para este término será la siguiente

$$\delta_2 \leq \frac{1}{\sigma^3} \binom{n}{3} (p^3 (1-p^3) + 3(n-3)p^5 (1-p)) (np^3 (1-p^3) + n^2 p^5 (1-p))^{1/2}$$

$$\leq \frac{C_1}{\sqrt{n}},$$

Y estará determinado el orden final del término segundo de la cota.

$$\delta_2 \sim O\left(\frac{C_1}{\sqrt{n}}\right).$$

En segundo lugar calcularemos el valor de $\delta_3 = \sum_{i \in \mathcal{J}} \mathbb{E}|\bar{\xi}_i \eta_i^2|$. Es fácil comprobar cómo en este caso el análisis combinatorio no será necesario. Esto viene dado por la ausencia de términos nulos y por tanto estamos en disposición de aplicar directamente la desigualdad aritmético geométrica de la media, la cual nos conduce a un orden de convergencia satisfactorio.

Tenemos

$$\sum_{i \in \mathcal{J}} \mathbb{E}|\bar{\xi}_i \eta_i^2| = \sum_{i \in \mathcal{J}} \mathbb{E}|\bar{\xi}_i| \sum_{j \in A_i} \bar{\xi}_j \sum_{l \in A_i} \bar{\xi}_l \leq \sum_{i \in \mathcal{J}} \sum_{j, l \in A_i} \mathbb{E}|\bar{\xi}_i \bar{\xi}_j \bar{\xi}_l|,$$

$$\mathbb{E}|\bar{\xi}_i \bar{\xi}_j \bar{\xi}_l| \leq \frac{1}{3}(\mathbb{E}|\bar{\xi}_i|^3 + \mathbb{E}|\bar{\xi}_j|^3 + \mathbb{E}|\bar{\xi}_l|^3) \leq \mathbb{E}|\bar{\xi}_i|^3.$$

Ahora bien podemos calcular

$$\mathbb{E}|\bar{\xi}_i|^3 = p^3(1-p^3)^3 + (1-p^3)p^9,$$

y finalmente

$$\sum_{j, l \in A_i} \mathbb{E}|\bar{\xi}_i \bar{\xi}_j \bar{\xi}_l| \leq |A_i|^2 \mathbb{E}|\bar{\xi}_i|^3 = |A_i|^2 (p^3(1-p^3)^3 + (1-p^3)p^9),$$

por lo que el valor de la cota es el que sigue

$$\delta_3 = \frac{1}{\sigma^3} \binom{n}{3} |A_i|^2 (p^3(1-p^3)^3 + (1-p^3)p^9) \sim O\left(\frac{n^5 p^3}{n^6 p^{15/2}}\right),$$

lo cual prueba el orden del tercer término de la cota, δ_3 .

$$\delta_3 \sim O\left(\frac{C_2}{n}\right).$$

Para el término de la cota que resta por hallar tendremos que calcular $\delta_1 = \sum_{i \in \mathcal{J}} (\mathbb{E}|\bar{\xi}_i \eta_i \tau_i|)$, este será a todas luces el cálculo más complicado, un análisis combinatorio puro será extremadamente difícil tal y como podemos ver en el dibujo que da una pequeña representación esquemática de la situación que tenemos (figura 2.6).

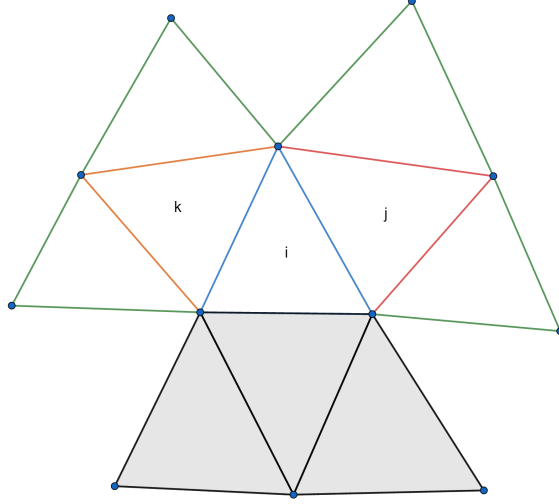


Figura 2.6: Esquema de situación para análisis combinatorio sobre δ_1 . Podemos ver como a la hora de contar términos nulos, si tomamos como situación ideal de un sumando particular en que los entornos del triángulo azul i , son los dados por k y j , entonces todos los B_i serán de la forma de los triángulos verdes y por tanto, aquellos representados por los triángulos negros, formarán parte de B_i^c y serán términos nulos, aunque no serán todos los que hay, luego a todas luces el análisis de ordenes de términos nulos es extremadamente complicado.

Luego para facilitar los cálculos aplicamos el siguiente argumento

$$\mathbb{E}|\bar{\xi}_i \eta_i \tau_i| \leq (\mathbb{E}|\bar{\xi}_i \eta_i^2|)^{1/2} (\mathbb{E}|\bar{\xi}_i \tau_i^2|)^{1/2},$$

donde estamos aplicando la desigualdad de uso común

$$\mathbb{E}|XY| \leq (\mathbb{E}X^2)^{1/2} (\mathbb{E}Y^2)^{1/2}.$$

Para nuestros propósitos, tendremos que tomar $X = |\bar{\xi}_i|^{1/2} |\eta_i|$ e $Y = |\bar{\xi}_i|^{1/2} |\tau_i|$.

Ahora bien, por la definición de las variables aleatorias que tienen influencia sobre este término de la cota, podemos escribir $\tau_i = \eta_i + \sum_{k \in B_i \cap A_i^c} \bar{\xi}_k$ y ya que obviamente $(a + b)^2 \leq 2a^2 + 2b^2$ entonces

$$\tau_i^2 \leq 2\eta_i^2 + 2 \left(\sum_{k \in B_i \cap A_i^c} \bar{\xi}_k \right)^2.$$

Una vez tenemos esto, combinándolo podemos llegar a acotar cada uno de los dos términos de la ecuación resultante

$$\mathbb{E}|\bar{\xi}_i \tau_i^2| \leq 2\mathbb{E}|\bar{\xi}_i \eta_i^2| + 2\mathbb{E}|\bar{\xi}_i| \mathbb{E}\left(\sum_{k \in B_i \cap A_i^c} \bar{\xi}_k^2\right),$$

pero podemos interpretar este último término como $\sum_{k \in B_i \cap A_i^c} \bar{\xi}_k = \tau_i - \eta_i$. Teniendo esto en cuenta

$$\mathbb{E}(\tau_i - \eta_i)^2 \leq 2\mathbb{E}\tau_i^2 + 2\mathbb{E}\eta_i^2 \leq C_{3,1}n^3,$$

y puesto que ya conocemos la desigualdad $\mathbb{E}|\bar{\xi}_i \eta_i^2|^{1/2} \leq C_2n$ ya que que ha sido calculado en el apartado anterior, juntando la aportación de este término y del calculado ahora, llegamos finalmente a acotar el término restante

$$\mathbb{E}|\bar{\xi}_i \eta_i \tau_i| \leq C_2n + C_{3,1}n^3 \leq C_3n^{5/2},$$

con lo que

$$\sum_{i \in \mathcal{J}} \mathbb{E}|\bar{\xi}_i \eta_i \tau_i| \leq Cn^{11/2} \Rightarrow \frac{1}{\sigma_n^3} \sum_{i \in \mathcal{J}} \mathbb{E}|\bar{\xi}_i \eta_i \tau_i| \leq \frac{C_3}{\sqrt{n}}.$$

Lo que nos deja el orden de convergencia siguiente y nos da una acotación adecuada para el término de la cota que restaba por determinar, que será del orden de

$$\delta_1 \sim O\left(\frac{C_3}{\sqrt{n}}\right).$$

Combinando las acotaciones anteriores con la desigualdad (2.2) del Teorema 2.1, habremos demostrado el siguiente teorema

Teorema 2.2. *Si la variable aleatoria W_n denota el número de triángulos en el grafo de Erdős-Renyi $G(n, p)$ entonces*

$$d_W\left(\frac{W_n - \mathbb{E}W_n}{\sqrt{\text{Var}(W_n)}}, Z\right) \leq \frac{C}{\sqrt{n}}$$

para una cierta constante $C > 0$. En particular

$$\frac{W_n - \mathbb{E}W_n}{\sqrt{\text{Var}(W_n)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Y por lo tanto tendremos un Teorema Central del Límite para la ley de distribución del número de triángulos en un grafo que sigue un modelo generador de Erdős-Renyi.

Capítulo 3

APROXIMACIÓN NORMAL DE FUNCIONALES DE POISSON MEDIANTE EL MÉTODO DE STEIN

Hemos mostrado cómo el método de Stein para la aproximación normal es una gran herramienta en la demostración de Teoremas Centrales del Límite para funcionales generales de variables aleatorias, pudiendo aplicarse a una gran cantidad de casos. Sin embargo cada ejemplo tratado ha requerido una idea diferente para conseguir acotar el término principal

$$\sup_{f \in \mathcal{F}} |E(f'(W) - Wf(W))|.$$

La combinación del método de Stein con el cálculo de Malliavin sí que proporciona una forma sistemática de control del término anterior.

Su aplicación no se limita a los casos expuestos en los capítulos anteriores y su desarrollo reciente ha conseguido ampliar su campo de acción. Como muestra de esto, veremos cómo la aplicación de las ideas desarrolladas por Stein pueden ser combinadas con el Cálculo de Malliavin para obtener cotas explícitas para la aproximación normal de funcionales de Poisson aleatorios para la distancia de Wasserstein.

El cálculo de Malliavin es el resultado de una teoría para dotar de un formalismo de cálculo diferencial e integral a ciertos espacios de variables aleatorias.

Originalmente se desarrolló para funcionales de un ruido Gaussiano” (véase [13]). La exposición de esta construcción resulta técnicamente más compleja. Aquí optamos por estudiar el caso más simple de funcionales de un ruido de Poisson”.

Con el término "ruido de Poisson" nos referimos a una medida aleatoria de Poisson centrada. Una medida aleatoria en un espacio medible $(\mathbb{X}, \mathcal{A})$ es una aplicación

$$\begin{aligned} \eta: \quad \Omega \times \mathcal{A} &\longrightarrow [0, \infty] \\ (w, A) &\longmapsto \eta(w, A) \end{aligned}$$

de forma que $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio probabilístico y se cumple:

- a) Para cada $A \in \mathcal{A}$ fijo, $w \mapsto \eta(w, A)$ es una variable aleatoria ($\mathcal{F}|_{\mathcal{A}}$ -medible).
- b) Para cada w fijo $\eta(w, \cdot)$ es una medida positiva en $(\mathbb{X}, \mathcal{A})$.

En lo sucesivo suprimiremos la mención a w en la notación y escribiremos de forma general $\eta(A)$.

Además, las medidas aleatorias de Poisson se caracterizan por presentar las propiedades adicionales siguientes:

- (i) Para todo $A \in \mathcal{A}$ la distribución de $\eta(A)$ sigue una distribución de Poisson de parámetro $\mu(A) = \mathbb{E}(\eta(A))$ (si $\mathbb{E}(\eta(A)) < \infty$), es decir

$$\mathbb{P}(\eta(A) = k) = Po(\mu(A); k) = e^{-\mu(A)} \frac{\mu(A)^k}{k!}. \quad (3.1)$$

- (ii) Para todo $n \in \mathbb{N}$ y todos los conjuntos disjuntos dos a dos A_1, A_2, \dots, A_n las variables aleatorias $\eta(A_1), \eta(A_2), \dots, \eta(A_n)$ son independientes.

La medida μ es conocida como "intensidad" del proceso de Poisson.

Desarrollar la teoría relativa a las medidas de Poisson aleatorias de forma general sería una tarea larga, que excede los objetivos de esta memoria. Se pueden consultar los aspectos fundamentales en [9].

Aquí nos limitamos a manejar el caso en el que $\mathbb{X} = \mathbb{R}^d$, o un subconjunto de Borel de \mathbb{R}^d , $\mathcal{A} = \mathcal{B}(\mathbb{X})$ y $\mu = \lambda \cdot \ell_d$ (ℓ_d es la medida de Lebesgue d -dimensional) con $\lambda > 0$. Al proceso η correspondiente se le conoce como proceso de Poisson homogéneo de tasa λ en \mathbb{X} .

El conjunto de funcionales de cuadrado integrable que están generados por un proceso de Poisson va a admitir una descomposición ortogonal que vendrá dada en términos de ciertas integrales estocásticas. Esta propiedad, que será fundamental, es la conocida como descomposición en caos de Poisson. Esta descomposición es la base sobre la que se apoya todo el cálculo de Malliavin, ya que nos va a permitir definir ciertos operadores (operador derivada,

Integral de Skorohod , operador de Ornstein-Uhlenbeck, etc...)) que proporcionan una cota sistemática en la aproximación normal mediante el método de Stein, tal y como expondremos en este capítulo.

Como ejemplo de aplicación de la teoría anterior, estudiaremos un ejemplo clásico de la Geometría Estocástica: el volumen de la unión de bolas centradas en puntos generados por un proceso de Poisson homogéneo de tasa λ en \mathbb{R}^d , η_λ .

Precisando, tal y como se verá en el desarrollo del capítulo, podemos entender η_λ como la medida de conteo asociada a un conjunto numerable aleatorio de puntos en \mathbb{R}^d .

Consideramos entonces el funcional

$$F_\lambda = F(\eta_\lambda) = \ell_d \left(\left(\bigcup_{z \in \eta_\lambda} B(z, r) \right) \cap [0, 1]^d \right). \quad (3.2)$$

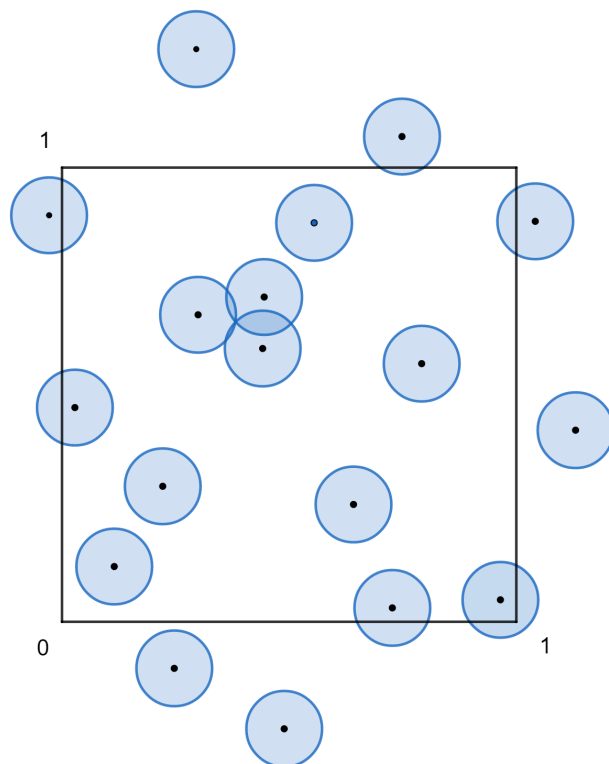


Figura 3.1: El caso de estudio en el caso $d = 2$.

La descripción del funcional es la que habíamos dado antes de manera informal, sobre cada punto de η construimos una bola de radio r aún por determinar. Limitamos el espacio al subespacio $[0, 1]^d$, entonces el funcional que actúa como nuestra variable aleatoria es la medida d -dimensional dada por la unión de todas estas bolas, y nuestro objetivo es probar que este volumen total puede ser aproximado utilizando la distribución normal.

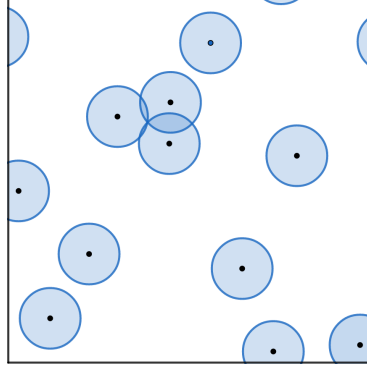


Figura 3.2: El funcional de estudio limitado a $[0, 1]^d$; $d = 2$.

Nuestro objetivo será aplicar el método de Stein para calcular una cota para la distancia de Wasserstein que nos permita encontrar un resultado del tipo

$$d_W \left(\mathcal{L} \left(\frac{F(\eta_\lambda) - \mathbb{E}F(\eta_\lambda)}{\sqrt{\text{Var}F(\eta_\lambda)}} \right), \mathcal{N}(0, 1) \right) \leq \delta_\lambda, \quad \text{con } \delta_\lambda \rightarrow 0, \lambda \rightarrow \infty,$$

para un cierto valor de δ_λ que debemos determinar.

Y como consecuencia demostrar que se cumple el siguiente Teorema Central del Límite

$$\frac{F(\eta_\lambda) - \mathbb{E}F(\eta_\lambda)}{\sqrt{\text{Var}F(\eta_\lambda)}} \xrightarrow[\lambda \rightarrow \infty]{d} \mathcal{N}(0, 1),$$

Sin embargo, para poder presentar esta cota de forma adecuada, debemos en primer lugar introducir una serie de conceptos y resultados teóricos del cálculo de Malliavin. En concreto nos interesa introducir los operadores de tipo Malliavin que actuarán sobre el funcional de Poisson. Para este cometido nos basaremos en la teoría expuesta en los artículos [10] y [14] en los que estos conceptos teóricos se presentan de forma generalizada.

El resto del capítulo está estructurado de la forma siguiente. Dedicamos la sección 1.1 a desarrollar los elementos necesarios de la teoría de procesos de Poisson, incluyendo la construcción de nuestro proceso de Poisson, la teoría necesaria sobre la integral estocástica referente a los procesos de Poisson,

llamada integral de Wiener-Ito, y una presentación del desarrollo en caos de Poisson. La sección 1.2 introduce los operadores del cálculo de Malliavin y cómo emplearlos en relación con el método de Stein. Finalmente la sección 1.3 presenta la aplicación que hemos descrito con anterioridad.

3.1. Procesos de Poisson.

En esta sección se presentan algunos elementos de la teoría de Procesos de Poisson que serán necesarios para el desarrollo de este trabajo. En la introducción a este capítulo hemos hablado de los procesos de Poisson como medidas aleatorias. Sin embargo, podemos entenderlos también como procesos puntuales. Un proceso puntual es una colección aleatoria Z , a lo sumo numerable, de puntos sobre un cierto espacio \mathbb{X} (en nuestro caso $\mathbb{X} = \mathbb{R}^d$ o $B \subset \mathbb{R}^d$ de Borel). Un proceso puntual induce de forma natural una medida aleatoria mediante la igualdad

$$(\omega, A) \rightarrow \eta(\omega, A) := \text{Card}(Z(\omega) \cap A) \quad A \in \mathcal{A}.$$

Si bien se pueden probar resultados más generales de existencia de procesos de Poisson. Para los objetivos de esta memoria basta con presentar la siguiente construcción.

En primer lugar consideramos una colección $\{X_n\}_{n \geq 1}$ de v.a.i.i.d. en $[0, 1]^d$ y N v.a. de Poisson de parámetro $\lambda > 0$, la cual es independiente de las X_n . Definimos

$$\eta(A) = \sum_{n \leq N} \mathbf{1}(X_n \in A)$$

($\eta(A) = 0$ si $N = 0$).

Es un ejercicio elemental comprobar que η es un Proceso de Poisson con intensidad $\mu = \lambda \ell_d|_{[0,1]^d}$ en el sentido definido en la introducción al capítulo. Claramente, η es también un proceso puntual, asociado al conjunto de puntos $\{X_n\}_{n \leq N}$.

Para extender la construcción a un proceso de Poisson homogéneo de intensidad $\lambda > 0$ en todo \mathbb{R}^d basta considerar una familia de procesos η_Q de Poisson de tasa λ en cada hipercubo de lado uno en \mathbb{R}^d , mutuamente independientes y definir

$$\eta(A) = \sum_Q \eta_Q(A).$$

Es fácil comprobar que el proceso η es un proceso de Poisson homogéneo de tasa λ en \mathbb{R}^d (y también un proceso puntual). Si ahora tomamos $\tilde{\eta}(A) = \eta(A \cup B)$ (con B de Borel) entonces $\tilde{\eta}$ es un proceso de Poisson en \mathbb{R}^d con medida de intensidad $\mu(A) = \lambda \ell_d(A \cap B)$ (o un proceso de Poisson homogéneo de tasa λ en B). Todos los procesos de Poisson que tratemos en este capítulo serán de este tipo.

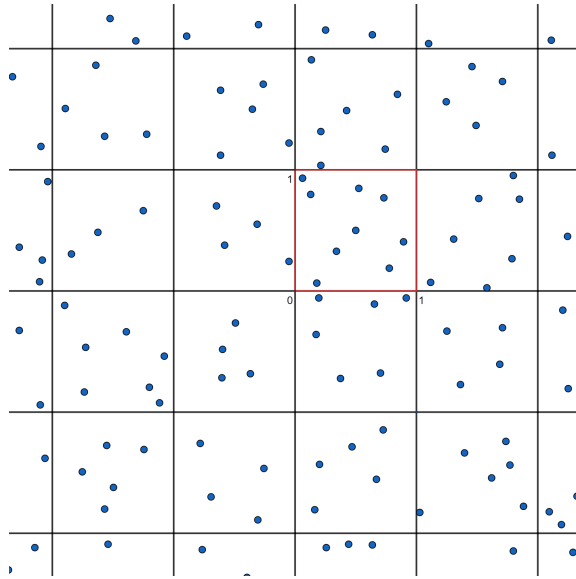


Figura 3.3: Esquema de la construcción del proceso sobre \mathbb{R}^d mediante el método de "pegar" hipercubos para $d=2$.

En definitiva, puesto que vamos a tomar uno de los hipercubos $[0, 1]^d$, y definir sobre el proceso η generado en \mathbb{R}^d bolas centradas sobre los puntos de η con un cierto radio r , bastará con tomar un engrosamiento de nuestro hipercubo para el radio fijado de tal forma que los puntos que aportan volumen al funcional aunque queden fuera del hipercubo se tengan en cuenta.

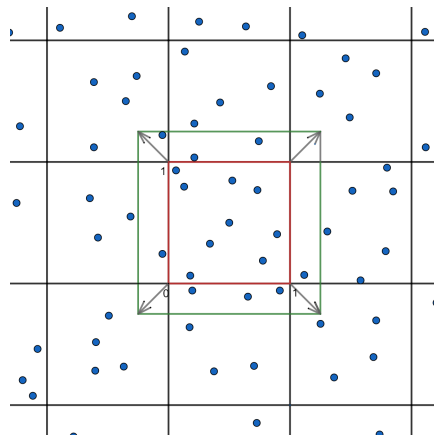


Figura 3.4: Esquema del engrosamiento para $d=2$ en la situación descrita por la figura 3.3. Nótese que esto encaja con el esquema presentado en la figura 3.1.

3.1.1. El desarrollo en caos de Poisson.

Cualquier funcional de cuadrado integrable de un proceso de Poisson puede ser expresado en términos de su desarrollo en caos de Poisson. Presentamos los detalles a continuación. Aunque la teoría expuesta puede ser presentada de forma generalizada (véase [11] o [8]), para nuestros propósitos bastará con considerar, tal como hemos anunciado, el caso η_λ medida de Poisson aleatoria sobre \mathbb{R}^d con medida de intensidad μ finita (en realidad, de la forma $\mu(A) = \lambda \ell_d(A \cap B)$ con B medible y acotado).

La integral de Wiener-Ito es la integral estocástica central de los procesos de Poisson. Esta integral puede ser definida para cualquier función de cuadrado integrable con respecto a la medida de intensidad. Definimos lo que llamaremos la *medida de Poisson compensada* o centrada que denotaremos como $\hat{\eta} = \eta - \mu$. Entonces, si $f \in L^2(\mathbb{X}, \mu)$ entonces denotaremos a la integral estocástica de f respecto del proceso de Poisson η , también llamada Integral de Wiener-Ito, como

$$I_1(f) = \int_{\mathbb{R}^d} f(x) d\hat{\eta} = \int_{\mathbb{R}^d} f(x) d(\eta - \lambda \ell_d) = \sum_{z \in \eta} f(z) - \int_{\mathbb{R}^d} f(x) d\mu(x).$$

Observamos que la integral en la última expresión es convergente (porque $f \in L^2(\mathbb{X}, \mu)$ y la medida μ es finita) y que la suma es sobre una cantidad finita (c.s.) de términos.

De forma más general, si $f \in L^2(\mathbb{X}^n, \mu^n)$ entonces la integral estocástica múltiple de f respecto del proceso de Poisson η se define como

$$\begin{aligned} I_n(f) &= \int_{\mathbb{X}_*^n} f(x_1, \dots, x_n) (d\hat{\eta}(x_1) \dots d\hat{\eta}(x_n)) \\ &= \int_{\mathbb{X}_*^n} f(x_1, \dots, x_n) (\eta - \mu)(dx_1) \dots (\eta - \mu)(dx_n), \end{aligned} \tag{3.3}$$

siendo el espacio de integración de esta integral el espacio total menos los puntos que tengan coordenadas coincidentes, es decir,

$$\mathbb{X}_*^n = \{(x_1, \dots, x_n) \in \mathbb{X}^n : x_i \neq x_j \forall i \neq j\}.$$

En el caso $n = 0$ escribimos

$$I_0(f) = \int_{\mathbb{R}^d} f(x) d\mu(x).$$

Para entender mejor el sentido de la definición (3.3) nos fijamos en el

caso $n = 2$. Entonces

$$\begin{aligned} I_2(f) &= \int_{(B \times B) \setminus D} f(x, y) d\widehat{\eta}(x) d\widehat{\eta}(y) = \\ &= \sum_{\substack{i=1 \\ i \neq j}}^{N_\lambda} \sum_{j=1}^{N_\lambda} f(x_i, x_j) - \lambda \left(\sum_{i=1}^{N_\lambda} \int_B f(x_i, y) dy + \sum_{j=1}^{N_\lambda} \int_B f(x, x_j) dy \right) \\ &\quad + \lambda^2 \int_{B \times B} f(x, y) dx dy, \end{aligned}$$

donde $D = \mathbb{X}_*^2$.

De forma más general, si f es una función simétrica en $L_2(\mathbb{X}^n, \mu^n)$ entonces

$$I_n(f) = \sum_{I \subset [n]} (-1)^{n-|I|} \int_{(\mathbb{R}^d)^n} f(x_1, \dots, x_n) d(\eta^{(I)} \otimes \mu^{([n]-I)})(x_1, \dots, x_n), \quad (3.4)$$

donde $[n] = \{1, \dots, n\}$ elementos y la integración sobre el producto tensorial dado representa la suma de evaluar f en los x_i con $i \in I$ integrando respecto μ las demás variables.

Las integrales de Wiener-Ito respecto el proceso de Poisson satisfacen las siguientes propiedades.

Lema 3.1. *Si $f \in L^2(\mathbb{X}^n, \mu^n)$ entonces*

1. $I_n(f) = I_n(\widehat{f})$, donde \widehat{f} es la forma simétrica de f , es decir,

$$\widehat{f}(x_1, \dots, x_n) = \frac{1}{n!} \sum_{(\pi)} f(x_{\pi_1}, \dots, x_{\pi_n}),$$

siendo $(\pi) = (\pi_1, \dots, \pi_n)$ todas las permutaciones de $(1, 2, \dots, n)$

2. $\mathbb{E}(I_n(f)) = 0$.
3. $\mathbb{E}(I_n(f)I_n(g)) = n! \langle f, g \rangle_{L^2(\mathbb{X}^n, \mu^n)} \mathbf{1}(m = n)$ para $f \in L_s^2(\mathbb{X}^n, \mu^n)$, $g \in L_s^2(\mathbb{X}^m, \mu^m)$.
4. $I_n(f) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$.

La demostración de este resultado se puede consultar en el Apéndice. El Lema 3.1 es fundamental en el desarrollo en caos de Poisson. Para llegar a esta descomposición necesitamos manejar el espacio de Fock asociado a un espacio $(\Omega, \mathcal{F}, \mathbb{P})$. La construcción se puede hacer de forma más general, pero esto no es necesario en esta memoria, (para una descripción detallada véase [13]). El espacio al que nos referimos consta de sucesiones de la forma

$f = (f_n)_{n \geq 0}$ donde f_0 es un número real y f_n es una función simétrica en $L_2(\mathbb{X}^n, \mu^n)$. En \mathbb{H} se define el producto escalar

$$\langle f, g \rangle_{\mathbb{H}} := \sum_{n \geq 0} \frac{1}{n!} \langle f_n, g_n \rangle_{L_2(\mathbb{X}^n, \mu^n)}. \quad (3.5)$$

El espacio de Fock \mathbb{H} es el espacio de sucesiones de funciones de la forma anterior tales que

$$\|f\|_{\mathbb{H}} := \langle f, f \rangle_{\mathbb{H}} < \infty.$$

Se puede comprobar que el espacio \mathbb{H} así definido es un espacio de Hilbert con producto escalar $\langle \cdot, \cdot \rangle_{\mathbb{H}}$.

El resultado principal al que nos referimos como descomposición en caos de Poisson es que el espacio de funciones de cuadrado integrable medibles en la σ -álgebra generada por el proceso de Poisson (por las variables $\eta(A)$, $A \in \mathbb{R}^d$) es isométrico al espacio de Fock anterior. Para llegar a este resultado necesitamos introducir el operador de diferencias D_z . De forma general denotaremos por $F = F(\eta)$ un funcional de los anteriores (de cuadrado integrable medibles en la σ -álgebra generada por el proceso de Poisson). Entonces la función $D_z f$ en \mathbf{N} se define como

$$D_z F(\eta) := F(\eta + \delta_z) - F(\eta), \quad (3.6)$$

siendo δ_x la medida de Dirac en un punto $x \in \Omega$.

Iterando la definición para $n \geq 2$ y (x_1, x_2, \dots, x_n) podemos definir de forma inductiva

$$D_{x_1, \dots, x_n}^n F(\eta) = D_{x_1}^1 D_{x_2, \dots, x_n}^{n-1} F(\eta). \quad (3.7)$$

Es conveniente además denotar $D^0 F(\eta) := F(\eta)$.

Por ejemplo, en el caso $n = 2$ podemos ver

$$\begin{aligned} D_{x_1, x_2}^2 F(\eta) &= D_{x_1}(D_{x_2} F(\eta)) = D_{x_1}[F(\eta + \delta_{x_2}) - F(\eta)] \\ &= F(\eta + \delta_{x_1} + \delta_{x_2}) - F(\eta + \delta_{x_2}) - F(\eta + \delta_{x_1}) + F(\eta). \end{aligned}$$

A partir del operador de diferencias definimos la función

$$T_n F(y_1, \dots, y_n) := \mathbb{E} D_{y_1, \dots, y_n}^n F(\eta). \quad (3.8)$$

Observamos que la función $T_n F(x_1, \dots, x_n)$ es una función simétrica de x_1, \dots, x_n .

Estamos en disposición de introducir el teorema principal de esta sección y en el que nos apoyamos para el desarrollo de toda la teoría del cálculo estocástico de Malliavin.

Teorema 3.2. *Sea $F = F(\eta)$ un funcional de cuadrado integrable medible en la σ -álgebra generada por el proceso de Poisson η . Entonces*

$$\mathbb{E} F(\eta)^2 = (\mathbb{E} F(\eta))^2 + \sum_{n=1}^{\infty} \frac{1}{n!} \|T_n F\|_n^2. \quad (3.9)$$

Si además $G \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, más generalmente

$$\mathbb{E}F(\eta)F(\eta) = (\mathbb{E}F(\eta))(\mathbb{E}G(\eta)) + \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n F, T_n G \rangle_n, \quad (3.10)$$

de lo cual podemos derivar además que si $f \in L_s^2(\mathbb{X}^n, \mu^n), g \in L_s^2(\mathbb{X}^m, \mu^m)$ entonces

$$\mathbb{E}(I_n(f)I_m(g)) = \delta_{n,m} n! \langle f, g \rangle_{L^2(\mathbb{X}^n, \mu^n)}, \quad (3.11)$$

considerando $\delta_{n,m} = \mathbf{1}(n = m)$.

La conclusión del Teorema 3.2 es que $TF \in \mathbb{H}$ para $F \in L^2(\mathbb{P}_\eta)$ y que, en realidad, T es una isometría entre el espacio de funcionales de Poisson de cuadrado integrable y el espacio de Fock \mathbb{H} . La demostración del resultado es larga y se desvía de los objetivos principales de este capítulo, por lo que se ha llevado al Apéndice.

3.2. Cálculo de Malliavin

El desarrollo en caos de Poisson dado por el Teorema 3.2 nos permite introducir una serie de operadores que, combinados con el método de Stein, conducirán a un tratamiento sistemático del Teorema Central del Límite para funcionales de Poisson. Estamos ya en disposición de introducir los verdaderos protagonistas de la cota que buscamos presentar, los distintos operadores que encontraremos en la expresión de la cota y que vendrán dados a través de una representación en caos de Poisson. Será necesario además conocer sus propiedades pues nuestro objetivo una vez presentada la cota será poder acotar sus términos para nuestro caso de estudio, y como ya hemos dicho estos términos vendrán en función de dichos operadores.

Notación: Denotamos $L^2(\mathbb{P}_\eta, L^2(\mathbb{X}, \mu)) \simeq L^2(\Omega \times \mathbb{X}, \mathcal{F} \times \mathcal{A}, \mathbb{P}_\eta \times \mu)$ el espacio de funciones medibles $u : \Omega \times \mathbb{X} \rightarrow \mathbb{R}$ que verifican

$$\mathbb{E} \left[\int_B u_z^2 \mu(dz) \right] < \infty.$$

El operador derivada D

El primer operador del cálculo de Malliavin es el operador derivada, D . En realidad este operador ya ha sido manejado, porque lo hemos necesitado para presentar el desarrollo en caos de Poisson. Este operador transforma variables aleatorias en funciones aleatorias, tal y como habíamos visto. Lo que vamos a ver aquí es que este operador se caracteriza de forma equivalente a través de la representación en caos, en la forma siguiente

$$D_z F(\eta) = \sum_{n \geq 1} n I_{n-1}(f_n(z, \cdot)), \quad z \in B, \quad (3.12)$$

donde f_n son las funciones del desarrollo en caos de Poisson de η . Por tanto tendremos que, explotando las propiedades isométricas de la integral múltiple de Wiener-Ito, $D_z F(\eta) \in L^2(\mathbb{P}_\eta, L^2(\mathbb{X}, \mu)) \forall F \in \text{dom} D$. Aquí, el *dominio* de D , que denotamos $\text{dom} D$ es el conjunto de todas las variables aleatorias $F \in L^2(\mathbb{P}_\eta)$ que admiten una descomposición en caos de Poisson tal que

$$\mathbb{E}F(\eta)^2 = \sum_{n \geq 1} n n! \|f_n\|_{L^2(\mathbb{X}^n, \mu^n)}^2 < \infty.$$

Demostremos ahora que la definición básica del operador y la definición como expresión en caos de Poisson son equivalentes, es decir, que

$$D_z F(\eta) = F(\eta + \delta_z) - F(\eta) = \sum_{n \geq 1} n I_{n-1}(\tilde{f}_n(z, \cdot)). \quad (3.13)$$

Si denotamos $z = x_{n+1}$ entonces

$$\begin{aligned} D_{x_1, \dots, x_n, x_{n+1}}^{n+1} F(\eta) &= D_{x_1, \dots, x_n, z}^{n+1} F(\eta) = \sum_{J \subset \{1, \dots, n+1\}} (-1)^{n+1-|J|} F(\eta + \sum_{j \in J} \delta_{x_j}) \\ &= - \sum_{J \subset \{1, \dots, n\}} (-1)^{n-|J|} F(\eta + \sum_{j \in J} \delta_{x_j}) \\ &\quad + \sum_{J \subset \{1, \dots, n\}} (-1)^n F(\eta + \delta_z + \sum_{j \in J} \delta_{x_j}), \end{aligned}$$

cuya última igualdad viene dada por dividir los conjuntos J que contienen el elemento $n+1$ y los que no. Por otro lado

$$\begin{aligned} D_{x_1, \dots, x_n, z}^{n+1} F(\eta) &= D_{x_1, \dots, x_n}^n F(\eta + \delta_z) - D_{x_1, \dots, x_n}^n F(\eta) \\ &\Rightarrow D_{x_1, \dots, x_n}^n F(\eta + \delta_z) = D_{x_1, \dots, x_n, z}^{n+1} F(\eta) + D_{x_1, \dots, x_n}^n F(\eta) \end{aligned}$$

Combinando lo anterior obtenemos que

$$\begin{aligned} ED_{x_1, \dots, x_n}^n F(\eta + \delta_z) &= f_n(x_1, \dots, x_n) + f_{n+1}(z, x_1, \dots, x_n) \\ &= \tilde{f}_n(x_1, \dots, x_n) + (n+1)\tilde{f}_{n+1}(z, x_1, \dots, x_n), \end{aligned}$$

lo que nos lleva a la siguiente expresión:

$$\begin{aligned} F(\eta + \delta_z) &= \sum_{n=0}^{\infty} I_n(\tilde{f}_n + (n+1)\tilde{f}_{n+1}(z, \cdot)) \\ &= \sum_{n=0}^{\infty} I_n(\tilde{f}_n) + (n+1)I_n(\tilde{f}_{n+1}(z, \cdot)). \end{aligned}$$

Como

$$F(\eta) = \sum_{n=0}^{\infty} I_n(\tilde{f}_n),$$

concluimos que

$$D_z F(\eta) = F(\eta + \delta_z) - F(\eta) = \sum_{n=0}^{\infty} (n+1) I_n(\tilde{f}_{n+1}(z, \cdot)),$$

y con un cambio de índice de suma $n+1 = m$ obtenemos (3.13).

La Integral de Skorohod δ

El operador δ se ocupa de transformar funciones aleatorias $u \in L^2(\mathbb{P}_\eta, L^2(\mathbb{X}, \mu))$ que admiten una expresión única en caos de Poisson de la forma

$$u_z = \sum_{n=0}^{\infty} I_n(f_n(z, \cdot)), \quad z \in Z, \quad (3.14)$$

donde $f_n(z, \cdot) \in L_s^2(\mathbb{X}^n, \mu^n)$. Tales que su expansión en caos de la forma anterior cumple

$$\sum_{n=0}^{\infty} (n+1)! \|f_n\|_{L^2(\mu^{n+1})}^2 < \infty. \quad (3.15)$$

Así definimos la variable aleatoria $\delta(u)$ de la forma siguiente

$$\delta(u) = \sum_{n \geq 0} I_{n+1}(\tilde{f}_n). \quad (3.16)$$

Podemos definir este operador además como el adjunto del operador D , lo cual manifestamos en el lema siguiente.

Lema 3.3. (*Formula de integración por partes*). Para todo $G \in \text{dom } D$ y $u \in \text{dom } \delta$ se cumple la igualdad

$$\mathbb{E}[G \cdot \delta(u)] = \mathbb{E}[\langle DG, u \rangle_{L^2(\mathbb{X}, \mu)}], \quad (3.17)$$

Demostración. Comenzamos desarrollando el lado derecho de la igualdad

$$\langle D_z G, u \rangle_{L^2(\mathbb{X}, \mu)} = \int_Z D_z G \times u(z) \mu(dz) = \int_B (G(\eta + \delta_z) - G(\eta)) u(z) dz,$$

por tanto por linealidad su esperanza será

$$\mathbb{E} \langle DG, u \rangle_{L^2(\mathbb{X}, \mu)} = \int_B \mathbb{E}(D_z G \cdot u(z)) dz.$$

Necesitamos calcular la esperanza del interior de la integral y expresándolas según sus desarrollos ortogonales estas vendrán dadas como

$$\begin{aligned}\mathbb{E}\langle DG, u \rangle_{L^2(\mathbb{X}, \mu)} &= \mathbb{E} \left(\sum_{n=0}^{\infty} (n+1) I_n(\tilde{f}_{n+1}(z, \cdot)) \sum_{n=0}^{\infty} I_n(u_n(z, \cdot)) \right) \\ &= \sum_{n=0}^{\infty} (n+1) \mathbb{E} \left(I_n(\tilde{f}_{n+1}(z, \cdot)) I_n(u_n(z, \cdot)) \right),\end{aligned}$$

por lo que aplicando a esto la relación isométrica (3.14) podemos escribir esto finalmente en la forma

$$\mathbb{E}(\langle DG, u \rangle_{L^2(\mathbb{X}, \mu)}) = \sum_{n=0}^{\infty} (n+1)! \int_{B^{n+1}} \tilde{f}_{n+1}(z, \cdot) u_n(z, \cdot) d(x_1, \dots, x_n) dz,$$

y finalmente utilizando la simetría de la función \tilde{f}_{n+1} ,

$$= \int_{B^{n+1}} \sum_{n=0}^{\infty} (n+1)! u_n(z, \cdot) \tilde{f}_{n+1}(z, \cdot) d(x_1, \dots, x_n) dz = \mathbb{E}[G \cdot \delta(u)].$$

■

El operador de Ornstein-Uhlenbeck L

El dominio del operador de Ornstein-Uhlenbeck denominado $\text{dom } L$, será el formado por todos los $F \in L^2(\mathbb{P}_\eta)$ tal que su expansión en caos de Poisson verifica

$$\sum_{n \geq 1} n^2 n! \|f_n\|_{L^2(\mathbb{X}^n, \mu^n)}^2 < \infty. \quad (3.18)$$

Si el funcional de Poisson F cumple esta condición y por tanto $F \in \text{dom } L$

$$LF = - \sum_{n \geq 1} n I_n(f_n). \quad (3.19)$$

Luego $\mathbb{E}(LF) = 0$ por definición y tendremos el siguiente resultado que pone en relación los tres operadores de Malliavin que hemos presentado.

Lema 3.4. *Para todo $F \in \text{dom } L$, tenemos que $F \in \text{dom } D$ y $DF \in \text{dom } (\delta)$. Y además tenemos entonces*

$$\delta DF = -LF. \quad (3.20)$$

Demostración. La primera parte es obvia pues si cumple $\sum_{n \geq 1} n^2 n! \|f_n\|_{L^2(\mathbb{X}^n, \mu^n)}^2 < \infty$ entonces se cumplirá $\sum_{n \geq 1} n n! \|f_n\|_{L^2(\mathbb{X}^n, \mu^n)}^2 < \infty$ por definición del dominio de D y $\sum_{n=0}^{\infty} (n+1)! \|f_n\|_{L^2(\mathbb{X}^{n+1}, \mu^{n+1})}^2 < \infty$ por definición del dominio de δ .

Ahora bien, probamos la igualdad para un cierto funcional F de la forma $F = I_q(f)$, $q \geq 1$. Ya que esto se puede extender a cualquier funcional general $F \in L^2(\mathbb{P}_\eta)$ que verifique $F \in \text{dom } D$. Tendremos entonces $D_z(F) = qI_{q-1}(f(z, \cdot))$, de tal forma que tendremos $\delta D_z F = qI_q(f) = -LF$, lo cual finaliza la prueba. ■

El operador inverso de Ornstein-Uhlenbeck L^{-1}

El dominio del operador L^{-1} que denotaremos por $L_0^2(\mathbb{P}_\eta)$ en el espacio de variables centradas en $L^2(\mathbb{P}_\eta)$. Si $F \in L_0^2(\mathbb{P}_\eta)$ y $F = \sum_{n \geq 1} I_n(f_n)$ entonces

$$L^{-1}F = - \sum_{n \geq 1} \frac{1}{n} I_n(f_n). \quad (3.21)$$

Comprobamos que efectivamente esta definición es el operador inverso del operador presentado como operador de Ornstein-Uhlenbeck L .

Lema 3.5. *Sea L el operador de Ornstein-Uhlenbeck, si definimos el operador $L^{-1}F = - \sum_{n \geq 1} \frac{1}{n} I_n(f_n)$ entonces*

$$L^{-1}LF = F \quad \text{y} \quad LL^{-1}F = F.$$

Demostración. Cálculos directos sobre las representaciones en caos de Poisson nos llevarán al resultado, cuyos cálculos estarán justificado por la linealidad de los operadores considerados. Sea $F = \sum_{n \geq 1} I_n(f_n)$,

$$LF = - \sum_{n \geq 1} n I_n(f_n),$$

por lo que aplicando ahora el operador L^{-1} sobre este desarrollo

$$L^{-1}LF = - \left(- \sum_{n \geq 1} \frac{1}{n} n I_n(f_n) \right) = \sum_{n \geq 1} I_n(f_n).$$

Estos cálculos tal y como hemos nombrado son válidos por linealidad, lo cual prueba la igualdad. El caso $LL^{-1}F = F$ será de similar comprobación. ■

Finalmente probamos un resultado que será útil más adelante sobre la representación en caos de la expresión $D_z L^{-1}F$.

Lema 3.6. *La descomposición en caos de $D_z L^{-1}F$ para un cierto funcional de Poisson F en el dominio de L^{-1} será*

$$D_z L^{-1}F = - \sum_{n=1}^{\infty} I_{n-1}(f_n(z, \cdot)).$$

Demostración.

$$D_z L^{-1}F = D_z \left(- \sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n) \right).$$

Puesto que recordamos que la expresión en caos de Poisson del operador derivada venía dado por

$$D_z F(\eta) = \sum_{n=1}^{\infty} n I_{n-1}(f_n(z, \cdot)),$$

aplicando esta representación sobre la expresión de $L^{-1}F(\eta)$ nos lleva finalmente a su representación en caos de Poisson

$$D_z L^{-1}F = - \sum_{n=1}^{\infty} I_{n-1}(f_n(z, \cdot)).$$

■

Ahora que hemos presentado los aspectos básicos del Cálculo de Malliavin sobre medidas de Poisson aleatorias estamos en condiciones de presentar el resultado principal en relación con el método de Stein (este resultado se presenta como Teorema 3.1. en [14]).

Teorema 3.7. *Sea $F \in \text{dom } D$, y tal que $\mathbb{E}(F_\lambda) = 0$. Si $Z \sim \mathcal{N}(0, 1)$. Entonces*

$$d_W(F, Z) \leq \mathbb{E}[|1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mu)}|] + \int_B \mathbb{E}[|D_z F|^2 |D_z L^{-1}F|] \mu(dz) \quad (3.22)$$

$$\leq \sqrt{\mathbb{E}[(1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mu)})^2]} + \int_B \mathbb{E}[|D_z F|^2 |D_z L^{-1}F|] \mu(dz). \quad (3.23)$$

Nota: Con la notación que estamos usando

$$\langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)} = - \int_B [D_z F \times D_z L^{-1}F] \mu(dz).$$

A partir de los desarrollos ortogonales para $-DL^{-1}F$ y DF es inmediato comprobar que

$$E(\langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}) = \text{Var}(F(\eta)). \quad (3.24)$$

Demostración. Aplicamos el Teorema 1.6 del primer capítulo. En virtud de una acotación del tipo

$$d_W(W, Z) \leq \sup_{f \in \mathcal{F}} |\mathbb{E}(f'(W) - Wf(W))|.$$

es suficiente que probemos que para toda función f tal que $\|f'\|_\infty \leq \sqrt{\frac{2}{\pi}} \leq 1$ y $\|f''\|_\infty \leq 2$, es decir, para toda $f \in \mathcal{F}$ siguiendo la notación introducida

en el Teorema 1.6, se cumple que el lado derecho de la ecuación anterior es más pequeño que el lado derecho de la cota que nos da el teorema.

Fijamos para ello un cierto $f \in \mathcal{F}$ y podemos ver que si usamos la expresión general de el operador derivada, tendremos que para todo $\omega, z \in \Omega$ se tiene que $D_z f(F)(\omega) = f(F)_z(\omega) - f(F)(\omega) = f(F_z(\omega) - f(F)(\omega))$.

Ahora, usando dos veces la representación $D_z F = F_z - F$ junto con un desarrollo de Taylor, podremos escribir

$$D_z f(F) = f(F)_z - f(F) = f(F_z) - f(F) = f'(F)(F_z - F) + \left[\frac{f''(\xi)}{2} \right] (F_z - F)^2,$$

y por lo tanto denotaremos

$$D_z f(F) = f'(F)(D_z F) + R(D_z F),$$

Puesto que $\left| \frac{f''(\xi)}{2} \right| \leq 1$ ya que $\|f''\|_\infty \leq 2$ llegamos a que $|R(x)| \leq x^2$ $\forall x \in \mathbb{R}$ entonces aplicaremos los Lemas 3.6 y 3.7 en ese mismo orden, siendo $u = DL^{-1}F$ y $G = f(F)$

$$\begin{aligned} \mathbb{E}[Ff(F)] &= \mathbb{E}[LL^{-1}Ff(F)] = \mathbb{E}[-\delta(DL^{-1}F)f(F)] \\ &= \mathbb{E}[\langle Df(F), -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}]. \end{aligned}$$

Ahora bien, por la expresión en expansión de Taylor de $D_z f(F)$ tendremos que

$$\begin{aligned} \mathbb{E}[\langle Df(F), -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}] &= \mathbb{E}[f'(F)\langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}] \\ &\quad + \mathbb{E}[\langle R(DF), -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}], \end{aligned}$$

de tal forma que tenemos

$$\begin{aligned} |\mathbb{E}[f'(F) - Ff(F)]| &\leq |\mathbb{E}[f'(F)(1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)})]| \\ &\quad + |\mathbb{E}[\langle R(DF), -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}]|. \end{aligned}$$

Utilizando ahora el hecho de que $\|f'\| \leq 1$ y la desigualdad de Cauchy-Schwarz,

$$\begin{aligned} |\mathbb{E}[f'(F)(1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)})]| &\leq \mathbb{E}[|1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}|] \\ &\leq \sqrt{\mathbb{E}[(1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)})^2]}. \end{aligned}$$

Por otro lado tenemos de inmediato que

$$\begin{aligned} |\mathbb{E}[\langle R(DF), -DL^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}]| &\leq \int_Z |\mathbb{E}[\langle R(D_z F)D_z L^{-1}F \rangle_{L^2(\mathbb{X}, \mu)}]| \mu(dz) \\ &\leq \int_Z \mathbb{E}[|D_z F|^2 |D_z L^{-1}F|] \mu(dz), \end{aligned}$$

lo cual completa la demostración. ■

Generalmente en la aplicación del Teorema 3.7 la variable para la que queremos probar la aproximación normal estará estandarizada, por lo que resultará más conveniente la siguiente versión simplificada.

Teorema 3.8. *Sea $F = F(\eta) \in \text{dom } D$, satisfaciendo $0 < \text{Var}(F) < \infty$. Si $Z \sim \mathcal{N}(0, 1)$*

$$d_W \left(\frac{F - \mathbb{E}(F)}{\sqrt{\text{Var}(F)}}, Z \right) \leq \frac{(\text{Var}(\int (D_z F)(-D_z L^{-1} F) d\mu(z)))^{1/2}}{\text{Var}(F)} + \frac{\mathbb{E}(\int (D_z F)^2 | -D_z L^{-1} F | d\mu(z))}{(\text{Var}(F))^{3/2}} \quad (3.25)$$

Demostración. Basta con aplicar el teorema anterior para $\tilde{F} = \frac{F - \mathbb{E}F}{\sqrt{\text{Var}(F)}}$, puesto que $\mathbb{E}(\int (D_z F)(-D_z L^{-1} F) d\mu(z)) = \text{Var}(F)$. ■

Podemos fijarnos, que en general, las cotas presentadas pueden no ser finitas. Nuestro trabajo por tanto en el estudio del caso planteado, será ver que efectivamente para la variable aleatoria que actúa como funcional de Poisson estas cotas son finitas. Por otra parte, las cotas se pueden presentar como dificultosas de calcular pese a que el funcional sea de sencillo, por ello debemos encontrar a su vez acotaciones sencillas para los términos que aparecen en las cotas que sean fácilmente controlables para nuestro caso de estudio, dedicaremos el apartado siguiente a este cometido y al cálculo explícito final de una cota válida para el funcional que estamos analizando.

3.3. Aplicación: TCL para el volumen de conjuntos aleatorios.

Toda la teoría que hemos introducido, tiene como objetivo final desarrollar un Teorema Central del Límite para el caso introducido en la primera sección.

Recordemos que habíamos presentado la situación en que teníamos $\eta_\lambda \sim \text{PPH}(\lambda)$ definido sobre el conjunto \mathbb{R}^d . El funcional objeto de nuestro estudio es

$$\tilde{F}(\eta_\lambda) = \ell_d \left(\left(\bigcup_{z \in \eta_\lambda} B(z, r) \right) \cap [0, 1]^d \right).$$

Resultará conveniente observar que

$$\begin{aligned} \tilde{F}(\eta_\lambda) &= \int_{[0,1]^d} \mathbf{1}(x \in \bigcup_{x \in \eta} B(x, r)) dx = 1 - \{x \in [0, 1]^d : \eta(B(x, r)) = 0\} \\ &= 1 - \int_{[0,1]^d} \mathbf{1}(x \notin \bigcup_{x \in \eta} B(x, r)) dx = 1 - \int_{[0,1]^d} \mathbf{1}(B(x, r) \cap \eta = \emptyset) dx. \end{aligned}$$

y que podemos, por lo tanto, centrarnos en el siguiente funcional, que consideramos de ahora en adelante para todos nuestros cálculos.

$$F(\eta_\lambda) = \int_{[0,1]^d} \mathbf{1}(B(x,r) \cap \eta = \emptyset) dx = \int_{[0,1]^d} \mathbf{1}(\eta(B(x,r)) = 0) dx. \quad (3.26)$$

Para calcular la esperanza de $F(\eta_\lambda)$ podemos intercambiar esperanza e integral y tenemos la integral de la esperanza de una función indicadora, que será su probabilidad. En consecuencia

$$\begin{aligned} \mathbb{E}F(\eta_\lambda) &= \mathbb{E} \int_{[0,1]^d} \mathbf{1}(B(x,r) \cap \eta = \emptyset) dx = \int_{[0,1]^d} \mathbb{E}\{\mathbf{1}(B(y,r) \cap \eta = \emptyset)\} dx \\ &= \int_{[0,1]^d} \mathbb{P}(\eta(B(x,r)) = 0) dx, \end{aligned} \quad (3.27)$$

donde hemos hecho el cambio $B(x,r) \cap \eta = \emptyset \Leftrightarrow \eta(B(x,r)) = 0$ para poder hablar correctamente de una probabilidad, su justificación sin más que pensar en lo que describen ambas expresiones geoméricamente tiene una fácil lectura, puesto que ambas nos dicen que el número de puntos del proceso de Poisson que encontramos dentro de la bola centrada en x y de radio r es nulo. Ahora bien, como sabemos al ser un proceso de Poisson tendremos que la probabilidad será la siguiente.

$$\mathbb{P}(\eta(B(x,r)) = 0) = e^{-\lambda C_d r^d}, \quad (3.28)$$

donde C_d denota el volumen de la bola unitaria d -dimensional.

Una vez tenemos esto, ya estamos en disposición de definir el valor del radio que habíamos pospuesto hasta este momento de forma correcta, pues ahora podemos ver cómo la selección adecuada del radio para que todo cuadre, será dar el valor $r = \frac{1}{\lambda^{1/d}}$ luego $\mathbb{P}(\eta(B(x,r)) = 0) = e^{-C_d}$ luego tendremos que cuando $\lambda \rightarrow \infty$ entonces $\mathbb{E}F(\eta_\lambda) \rightarrow e^{-C_d}$.

Tenemos por tanto, que una vez fijado el radio que hace que la situación funcione correctamente, tomaremos el funcional en la forma siguiente para el radio ya definido

$$F(\eta_\lambda) = \int_{[0,1]^d} \mathbf{1}(\eta(B(z, \lambda^{-1/d})) = 0) dz, \quad (3.29)$$

funcional que tiene como esperanza tal y como hemos calculado

$$\mathbb{E}(F(\eta_\lambda)) = e^{-C_d}. \quad (3.30)$$

Ya tenemos por tanto calculado el valor de la esperanza, y además hemos fijado el valor del radio que hace que el problema funcione correctamente, el siguiente paso será por tanto realizar el cálculo de la varianza.

Utilizando la expresión conocida de la varianza tenemos que

$$Var(F(\eta)) = \mathbb{E}(F(\eta)^2) - \mathbb{E}(F(\eta))^2,$$

y puesto que ya tenemos el valor de la esperanza, para desarrollar esta expresión comenzamos calculando

$$\mathbb{E}(F(\eta)^2) = \int_{[0,1]^d \times [0,1]^d} \mathbf{1}(B(x, \lambda^{-1/d}) \cup B(y, \lambda^{-1/d})) dx dy.$$

Teniendo en cuenta que $\ell_d(A \cup B) = \ell_d(A) + \ell_d(B) - \ell_d(A \cap B)$ y sustituyendo esto dentro de la integral tendremos

$$\begin{aligned} \mathbb{E}(F(\eta)^2) &= \int_{[0,1]^d \times [0,1]^d} \mathbf{1}(B(x, \lambda^{-1/d}) + B(y, \lambda^{-1/d}) - \\ &\quad B(x, \lambda^{-1/d}) \cap B(y, \lambda^{-1/d})) dx dy \\ &= e^{-2C_d} \int_{[0,1]^d \times [0,1]^d} e^{-\lambda \ell_d(B(x, \lambda^{-1/d}) \cap B(y, \lambda^{-1/d}))} dx dy, \end{aligned}$$

por lo que podemos llegar a que dado que $\mathbb{E}(F(\eta))^2 = e^{-2C_d}$ la varianza vendrá representada por la expresión siguiente

$$Var(F(\eta)) = e^{-2C_d} \int_{[0,1]^d \times [0,1]^d} (e^{-\lambda \ell_d((B(x, \lambda^{-1/d}) \cap B(y, \lambda^{-1/d})))} - 1) dx dy.$$

Ahora bien, para calcular el interior de la integral, necesitaremos hacer el siguiente cambio de variable, que vendrá dado por $y = x + z\lambda^{-1/d}$ y por tanto tendremos la transformación siguiente

$$\lambda \ell_d((B(x, \lambda^{-1/d}) \cap B(y, \lambda^{-1/d}))) = \lambda \ell_d((B(0, \lambda^{-1/d}) \cap B(y - x, \lambda^{-1/d}))),$$

ya que se cumple $\lambda \ell(B(0, 1) \cap B(\lambda^{-1/d}z, 1)) = \ell(B(0, 1) \cap B(z, 1))$

Y llegamos a que $z = (y - x)\lambda^{1/d}$ y $[0, 1]^d$ se transforma en el conjunto $V_\lambda(x) = \{z \in \mathbb{R}^d : z = (y - x)\lambda^{1/d}, y \in (0, 1)^d\}$ cumpliéndose que $V_\lambda(x) \uparrow \mathbb{R}^d$.

$$Var(F(\eta)) = \frac{e^{-2C_d}}{\lambda} \int_{(0,1)^d} \left[\int_{V_\lambda(x)} (e^{-\ell_d((B(0,1) \cap B(z,1)))} - 1) dz \right] dx,$$

y por convergencia tendremos cuando $\lambda \rightarrow \infty$

$$\begin{aligned} \lambda Var(F(\eta)) &= e^{-2C_d} \int_{(0,1)^d} \left[\int_{V_\lambda(x)} (e^{-\ell_d((B(0,1) \cap B(z,1)))} - 1) dz \right] dx = C(\lambda) \\ &\downarrow e^{-2C_d} \int_{(0,1)^d} \left[\int_{\mathbb{R}^d} (e^{-\ell_d((B(0,1) \cap B(z,1)))} - 1) dz \right] dx. \end{aligned}$$

Ahora bien, fijándonos en la geometría tenemos que la integral se puede limitar a la bola $B(0, 2)$ ya que fuera de ese conjunto todos los valores serán nulos, y por tanto llegamos a que

$$\lambda \text{Var}(F(\eta)) \nearrow e^{-2C_d} \int_{(0,1)^d} \left[\int_{B(0,2)} (e^{-\ell_d((B(0,1) \cap (B(z,1)))} - 1) dz \right] dx = C < \infty.$$

Para probar un TCL para $F(\eta_\lambda)$ con el método de Stein-Malliavin necesitamos encontrar su desarrollo en caos de Poisson. La clave será encontrar este desarrollo para un tipo muy simple de funcional, de la forma siguiente. Sea $A \subset \mathbb{R}^d$ un conjunto de Borel definido sobre un proceso puntual de Poisson en \mathbb{R}^d con intensidad de medida μ y $\mu(A) < \infty$. Definimos el funcional centrado

$$F(\eta) = \mathbf{1}(\eta(A) = 0) - \mathbb{P}(\eta(A) = 0) = \mathbf{1}(\eta(A) = 0) - e^{-\lambda \ell_d(A)}. \quad (3.31)$$

Tenemos, por definición,

$$f_n(x_1, \dots, x_n) = \mathbb{E} \frac{D_{x_1, \dots, x_n}^n F(\eta)}{n!}. \quad (3.32)$$

Vamos desarrollando esta expresión en busca de conseguir una forma general para su representación. Comenzamos viendo su forma para $n = 1$ que vendrá dada por la siguiente expresión

$$f_1(x_1) = \mathbb{E} F((\eta + \delta_{x_1}) - G(\eta)) = \mathbb{E}(\mathbf{1}(\eta(A) = 0, x_1 \notin A) - \mathbf{1}(\eta(A) = 0))$$

Puesto que la función indicadora de la intersección es el producto de las mismas y sacando factor común a $\mathbb{E}(\mathbf{1}(\eta(A) = 0))$ llegamos a lo siguiente

$$\begin{aligned} f_1(x_1) &= \mathbb{E}(\mathbf{1}(\eta(A) = 0) \mathbf{1}(x_1 \notin A) - \mathbf{1}(\eta(A) = 0)) \\ &= \mathbb{E}(\mathbf{1}(\eta(A) = 0)) (\mathbf{1}(x_1 \notin A) - 1) \\ &= -\mathbb{E}(\mathbf{1}(\eta(A) = 0)) \mathbf{1}(x \in A) = -e^{-\lambda \ell_d(A)} \mathbf{1}_A(x) \end{aligned}$$

Ahora bien, para el caso $n = 2$ se tiene

$$\begin{aligned} 2f_2(x_1, x_2) &= \mathbb{E} F((\eta + \delta_{x_1} + \delta_{x_2}) - F((\eta + \delta_{x_1}) - F((\eta + \delta_{x_2}) + F(\eta))) \\ &= \mathbb{E}(\mathbf{1}(\eta(A) = 0)) (\mathbf{1}(x_1 \notin A, x_2 \notin A) - \mathbf{1}(x_1 \notin A) - \mathbf{1}(x_2 \notin A) + 1) \end{aligned}$$

Por las reglas de complementarios para la función indicadora tenemos que $\mathbf{1}(x_1 \notin A) = 1 - \mathbf{1}_A(x)$ y sustituyendo esto en la ecuación tendremos

$$\begin{aligned} 2f_2(x_1, x_2) &= \mathbb{E}(\mathbf{1}(\eta(A) = 0)) = (1 - \mathbf{1}_A(x_1))(1 - \mathbf{1}_A(x_2)) - \\ &\quad (1 - \mathbf{1}_A(x_1)) - (1 - \mathbf{1}_A(x_2)) + 1 \\ &= -\mathbf{1}_A(x_1)(1 - \mathbf{1}_A(x_2) - 1 + \mathbf{1}_A(x_1)) + 1 = \mathbf{1}(x_1 \in A, x_2 \in A). \end{aligned}$$

Así pues, razonando de forma similar para el caso n -ésimo es fácil ver que la fórmula general será de la forma siguiente

$$f_n(x_1, \dots, x_n) = \frac{(-1)^n}{n!} e^{-\lambda \ell_d(A)} \mathbf{1}_{(A \times \dots \times A)}(x_1, \dots, x_n). \quad (3.33)$$

Buscamos ahora dar una representación cerrada general para $I_n(f_n)$. Desarrollamos término a término para poder ver la forma general que nos proporciona la expresión buscada

$$\begin{aligned} I_1(f_1) &= \int_{[0,1]^d} f_1(x_1) d\hat{\eta}(x_1) = \int_{[0,1]^d} f_1(x_1) d(\eta(x_1) - \lambda \int_{[0,1]^d} f_1(x_1) dx_1) \\ &= \sum_{z \in \eta} \mathbf{1}_A(z) - \lambda \int_{[0,1]^d} \mathbf{1}_A(x_1) dx_1 = \eta(A) - \lambda \ell_d(A). \end{aligned}$$

Para $n = 2$

$$\begin{aligned} I_2(f_2) &= \int_{[0,1]^d} f_2(x_1, x_2) d\hat{\eta}(x_1) d\hat{\eta}(x_2) = \sum_{z_1, z_2} I_A(z_1) I_A(z_2) - \sum_{z_1 \in \eta} I_A(z_1) \lambda \ell_d(A) \\ &\quad - \sum_{z_2 \in \eta} I_A(z_2) \lambda \ell_d(A) + (\lambda \ell_d(A))^2 \\ &= \left(\sum_{z_1 \in \eta} I_A(z_1) \right) \left(\sum_{z_2 \in \eta} I_A(z_2) \right) - \sum_{z \in \eta} I_A(z) - \sum_{z_1 \in \eta} I_A(z_1) \lambda \ell_d(A) \\ &\quad - \sum_{z_2 \in \eta} I_A(z_2) \lambda \ell_d(A) + (\lambda \ell_d(A))^2 \\ &= \eta(A)^2 - \eta(A) - 2\lambda \ell_d(A) \eta(A) + (\lambda \ell_d(A))^2. \end{aligned}$$

Generalizando estos resultados, fijando un cierto $I \subset [n]$ con $|I| = k$ entonces, podemos escribir de forma general

$$\int_{B^n} f(x_1, \dots, x_n) d(\eta^{(I)} \otimes \mu^{[n]-I})(x_1, \dots, x_n) = \frac{(-1)^n}{n!} e^{-\mu(A)} \mu(A)^{n-k} k! \binom{\eta(A)}{k},$$

y llegamos por tanto a que la fórmula general vendrá dada por la siguiente expresión

$$I_n(f_n) = \frac{e^{-\lambda \ell_d(A)}}{n!} \sum_{k=0}^n (-1)^k \binom{n}{k} \binom{\eta(A)}{k} k! (\lambda \ell_d(A))^{n-k}. \quad (3.34)$$

Como consecuencia de la representación (3.34) obtenemos la siguiente expresión explícita para la acción de los operadores D y DL^{-1} sobre F .

Lema 3.9. *Sea η un proceso puntual de Poisson en $[0, 1]^d$ con intensidad de medida μ , $A \subset [0, 1]^d$ un conjunto de Borel tal que $\mu(A) < \infty$ y $\mathbf{1}(\eta(A) = 0) - \mathbb{P}(\eta(A) = 0) = \mathbf{1}(\eta(A) = 0) - e^{-\lambda \ell_d(A)}$ entonces*

$$\begin{aligned} D_z F &= -\mathbf{1}_A(z) F(\eta). \\ D_z L^{-1} F &= \mathbf{1}_A(z) \int_0^1 t^{\eta(A)} e^{-\mu(A)t} dt. \end{aligned}$$

Demostración. Probamos la representación para el operador $D_z L^{-1} F$ pues la otra expresión es obvia, para ello nos valdremos de la representación en caos de Poisson del funcional antes presentada y la linealidad de la integral estocástica.

$$\begin{aligned}
D_z L^{-1} F(\eta) &= \mathbf{1}_A(z) e^{-\mu(A)} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \left[\sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{m-k} \mu(A)^{m-k} k! \binom{\eta(A)}{k} \right] \\
&= \mathbf{1}_A(z) e^{-\mu(A)} \sum_{m=0}^{\infty} \frac{(-1)^m}{(m+1)!} \left[\sum_{k=0}^m \binom{m}{k} (-1)^{m-k} \mu(A)^{m-k} k! \binom{\eta(A)}{k} \right] \\
&= \mathbf{1}_A(z) e^{-\mu(A)} \sum_{k=0}^{\infty} (-1)^k \binom{\eta(A)}{k} \left[\sum_{m=k}^{\infty} \frac{1}{m+1} \frac{\mu(A)^{m-k}}{(m-k)!} \right] \\
&= \mathbf{1}_A(z) e^{-\mu(A)} \sum_{k=0}^{\infty} (-1)^k \binom{\eta(A)}{k} \left[\sum_{n=0}^{\infty} \frac{1}{n+k+1} \frac{\mu(A)^n}{(n)!} \right] \\
&= \mathbf{1}_A(z) e^{-\mu(A)} \int_0^1 \left[\sum_{k=0}^{\infty} \binom{\eta(A)}{k} (-x)^k \left[\sum_{n=0}^{\infty} \frac{(x\mu(A))^n}{(n)!} \right] \right] dx \\
&= \mathbf{1}_A(z) e^{-\mu(A)} \int_0^1 \left[\sum_{k=0}^{\infty} \binom{\eta(A)}{k} (-x)^k \right] e^{\mu(A)x} dx \\
&= \mathbf{1}_A(z) \int_0^1 (1-x)^{\mu(A)} e^{-\mu(A)(1-x)} dx,
\end{aligned}$$

donde hemos usado $\frac{1}{n+k+1} = \int_0^1 x^{k+n} dx$. El intercambio entre integrales y sumatorios vendrá justificado porque $\mathbb{E}((1-x)^{\eta(A)}) = e^{-\mu(A)x}$ luego la integral es casi siempre finita. Por el cambio de variable $t = 1-x$ tendremos completada la prueba. ■

Aunque el Lema 3.9 se refiere a la acción de operadores de Malliavin sobre funcionales de la forma $\mathbf{1}(\eta(A) = 0)$ (centrados), el función $F(\eta_\lambda)$ definido en (3.26) se puede manejar de forma simple a partir de este resultado, porque se trata de un promedio (una integral) de funcionales del tipo del Lema 3.9. La linealidad de los operadores de Malliavin permite extender el resultado. La acción de los operadores de Malliavin sobre el funcional (3.26) será el promedio correspondiente de las acciones sobre los operadores del tipo de Lema 3.9.

Ahora estamos en condiciones de probar un TCL en nuestro caso de estudio. Para ello necesitamos en primer lugar dar una acotación útil para cada uno de los términos que aparecen en la cota del Teorema 3.8, de tal forma que podamos trabajar con ellos asegurando su acotación y su buen comportamiento operativo. Nuestro objetivo por tanto será acotar cada uno

de los dos términos que aparecen mediante expresiones sencillas de calcular de tal manera que podamos razonar de forma geométrica como con la varianza y convertir este problema de un problema analíticamente complicado en uno geométrico de fácil resolución.

Comenzamos acotando el término $\int \mathbb{E}[|D_z F|^2 |D_z L^{-1} F|] d\mu(z)$, para ello aplicamos la desigualdad de Schwarz, obteniendo

$$\int \mathbb{E}[|D_z F|^2 |D_z L^{-1} F|] \leq \int (\mathbb{E}|D_z F|^4)^{1/2} (\mathbb{E}|D_z L^{-1} F|^2)^{1/2} \mu(dz).$$

Buscamos una expresión de la cota explícita para el primer elemento,

$$\begin{aligned} & \int_{[0,1]^d} \mathbb{E}(D_z F)^4 \lambda dz \\ &= \lambda \int_{[0,1]^d} \int_{([0,1]^d)^4} e^{-\lambda \ell_d(\cup_{i=0}^3 B(x_i, \lambda^{-1/d}))} \mathbf{1}(\cap_{i=0}^3 B(x_i, \lambda^{-1/d})(z)) dx_0 \dots dx_3 dz. \end{aligned}$$

Hacemos el cambio de variable $x_i = x_0 + \lambda^{-1/d} z_i$ $i = 1, 2, 3$ denotando $V_\lambda = \{(x_0, z_1, z_2, z_3) : x_0 \in (0, 1)^d, z_i \in \lambda^{1/d}((0, 1)^d - x_0), i = 1, 2, 3\} \uparrow (\mathbb{R}^d)^4$. Notemos que por el Teorema del Jacobiano, saldrá fuera con valor $\frac{1}{\lambda^4}$ que al multiplicarle con el λ ya presente en la expresión nos lleva a que tomando $z_0 = 0$ posible pues por traslaciones no generamos problemas en la acotación tendremos

$$\begin{aligned} & \int_{[0,1]^d} \mathbb{E}(D_z F)^4 \lambda dz \\ &= \frac{1}{\lambda^3} \int_{[0,1]^d} \int_{V_\lambda} e^{-\ell_d(\cup_{i=0}^3 B(z_i, 1))} \ell_d(\cap_{i=0}^3 B(z_i, 1)(z)) dz_0 \dots dz_3 dz \\ & \downarrow \frac{1}{\lambda^3} \int_{[0,1]^d} \int_{(\mathbb{R}^d)^3} e^{-\ell_d(B(0,1) \cup \cup_{i=1}^3 B(z_i, 1))} \ell_d((B(0, 1) \cap (\cap_{i=1}^3 B(z_i, 1)(z)) dz_1 \dots dz_3 dz \\ &= \frac{1}{\lambda^3} \int_{[0,1]^d} \int_{(B(0,6))^3} e^{-\ell_d(B(0,1) \cup \cup_{i=1}^3 B(z_i, 1))} \ell_d((B(0, 1) \cap (\cap_{i=1}^3 B(z_i, 1)(z)) dz_1 \dots dz_3 dz \\ &= \frac{1}{\lambda^3} \ell_d([0, 1]^d) \int_{(B(0,6))^3} e^{-\ell_d(B(0,1) \cup \cup_{i=1}^3 B(z_i, 1))} \ell_d((B(0, 1) \cap (\cap_{i=1}^3 B(z_i, 1)(z)) dz_1 \dots dz_3 dz \\ & \leq \frac{C_1}{\lambda^3} < \infty. \end{aligned}$$

Puesto que como en el cálculo de la varianza, si $\|z_i\| > 2$ entonces $B(0, 1) \cap B(z_i, 1) = \emptyset$.

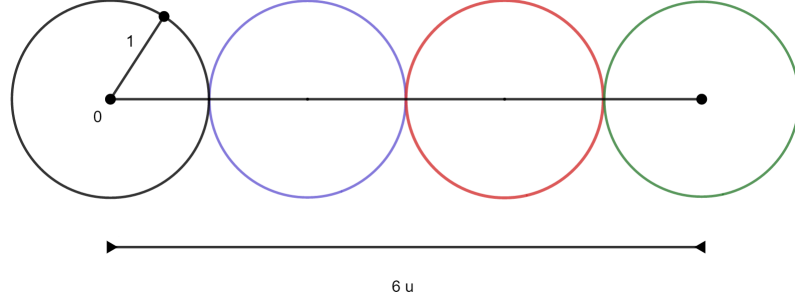


Figura 3.5: Esquema de máxima distancia entre las 4 bolas.

En consecuencia, hemos llegado a una acotación de la forma

$$\int_{[0,1]^d} \mathbb{E}(D_z F)^4 \lambda dz \leq \frac{C_1}{\lambda^3}.$$

Ahora bien, necesitamos el cálculo de un acotación adecuada para el otro término que aparece en la primera parte de la cota, comprobamos para ello que $\int \mathbb{E}|D_z L^{-1} F|^2 d\mu(z) \leq \text{Var}(F)$, lo que ya nos arroja una acotación válida.

Esta comprobación es inmediata utilizando la fórmula de representación cerrada para la expresión del operador combinado $D_z L^{-1} F$, ya que tenemos

$$D_z L^{-1} F(\eta) = \int_0^1 \int_{[0,1]^d} t^\eta (B(x, \lambda^{-1/d})) e^{-\lambda C_d \lambda^{-1} t} \mathbf{1}(z \in B(x, \lambda^{-1/d})) dx dt,$$

luego tomando esperanza sobre el cuadrado de la expresión

$$\begin{aligned} \mathbb{E}|D_z L^{-1} F(\eta)|^2 &= \\ \mathbb{E} \left[\int_{[0,1]^d \times [0,1]^d} \left[\int_{[0,1] \times [0,1]} t^\eta (A)_s \eta (B) e^{-C_d t} e^{-C_d s} dt ds \right] \right. \\ &\quad \left. \mathbf{1}(z \in B(x, \lambda^{-1/d} \cap B(y, \lambda^{-1/d}))) dx dy \right], \end{aligned}$$

que sometándolo a la acción de la integral sobre la expresión general

$$\begin{aligned} \int \mathbb{E}|D_z L^{-1} F(\eta)|^2 \mu(dz) &= \int \sum_{n=1}^{\infty} \mathbb{E}(I_{n-1}(f_n(z, \cdot)))^2 \mu(dz) \\ &= \sum_{n=1}^{\infty} \int \mathbb{E}(I_{n-1}(f_n(z, \cdot)))^2 \mu(dz) \\ &= \sum_{n=1}^{\infty} (n-1)! \|f_n(z, \cdot)\|_n^2, \end{aligned}$$

con lo que llegamos a que $\int \mathbb{E}|D_z L^{-1}F|^2 d\mu(z) = \sum_{n=1}^{\infty} (n-1)! \|f_n\|_n^2 \leq \sum_{n=1}^{\infty} (n)! \|f_n\|_n^2 = \text{Var}(F)$ lo que nos da la acotación deseada.

Aplicando la cota del Teorema 3.2 tendremos finalmente la acotación siguiente para el primer término de la cota.

$$\frac{\mathbb{E} \int [|D_z F|^2 |D_z L^{-1}F|] \mu(dz)}{(\text{Var}(F))^{3/2}} \leq \frac{C_1^{1/2}}{\lambda^{1/2} C(\lambda)}.$$

Resta por calcular el otro término de la cota, tenemos mediante cálculos elementales que

$$\begin{aligned} \text{Var}\left(\int X_s ds\right) &= \text{Cov}\left(\int X_s ds, \int X_t dt\right) \\ &= \mathbb{E}\left(\int X_s ds \cdot \int X_t dt\right) - \mathbb{E}\int X_s ds \cdot \mathbb{E}\int X_t dt \\ &= \mathbb{E}\left[\int \int X_s X_t ds dt\right] - \int \mathbb{E}X_s ds \int \mathbb{E}X_t dt \\ &= \int \int \mathbb{E}(X_s X_t) ds dt - \int \mathbb{E}X_s \mathbb{E}X_t ds dt, \end{aligned}$$

y por tanto

$$\text{Var}\left(\int X_s ds\right) = \text{Cov}\left(\int X_s ds, \int X_t dt\right) = \int [\mathbb{E}(X_s X_t) - (\mathbb{E}X_s \mathbb{E}X_t ds dt)],$$

luego a todas luces necesitaremos una acotación en forma de desigualdad de covarianza, la cual podemos dar gracias al lema siguiente, cuya prueba podemos encontrar en el Lema 3.2 de [7].

Lema 3.10. *Sea η una medida de Poisson con intensidad de medida μ y A, B, C y D son conjuntos medibles de medida μ -finita entonces*

$$\mathbb{E}\left(\mathbf{1}(\eta(A) = 0) \int_0^1 u^{\eta(B)} e^{-\mu(B)u} du\right) = e^{-\mu(A \cup B)} \frac{1 - e^{-\mu(A \cap B)}}{\mu(A \cap B)},$$

$$= e^{-\mu(A \cap B)} \text{ si } \mu(A \cap B) = 0.$$

Luego tenemos

$$\begin{aligned} &\text{Cov}\left(\mathbf{1}(\eta(A) = 0) \int_0^1 u^{\eta(B)} e^{-\mu(B)u} du, \mathbf{1}(\eta(C) = 0) \int_0^1 v^{\eta(D)} e^{-\mu(D)v} dv\right) \\ &\leq e^{-\mu(A \cup B \cup C \cup D)} \mathbf{1}((A \cup B) \cap (C \cup D) \neq \emptyset), \end{aligned}$$

luego por tanto tendremos que podemos acotar la expresión gracias al lema anterior ya que tendremos que la varianza a calcular es igual a la

varianza de la integral siguiente, es decir

$$\begin{aligned}
& \text{Var}\left(\int (D_z F)(-D_z L^{-1} F) d\mu(z)\right) \\
& \leq \int_{[0,1]^d \times [0,1]^d} \mathbf{1}(\eta_\lambda(B(x_0, \lambda^{-1/d}) = 0)) \left(\int_0^1 t^{\eta_\lambda(B(x_1, \lambda^{-1/d}))} e^{-\lambda \eta_\lambda(B(x_0, \lambda^{-1/d}))t} dt\right) \\
& \quad \cdot \lambda \ell(B(x_0, \lambda^{-1/d}) \cap B(x_1, \lambda^{-1/d})) dx_0 dx_1,
\end{aligned}$$

y usando la desigualdad de covarianza vista en el Lema 3.3 anterior, tendremos que estará acotada por

$$\begin{aligned}
& \leq \int_{([0,1]^d)^2 \times ([0,1]^d)^2} e^{-\lambda \ell_d(\cup_{i=0}^3 B(x_i, \lambda^{-1/d}))} \mathbf{1}(B(x_0, \lambda^{-1/d}) \cap B(x_1, \lambda^{-1/d})) \\
& \quad \mathbf{1}(B(x_2, \lambda^{-1/d}) \cap B(x_3, \lambda^{-1/d})) dx_0 dx_1 dx_2 dx_3.
\end{aligned}$$

Ahora bien, podemos hacer el cambio de variable ya utilizado $x_i = x_0 + \lambda^{-1/d} z_i$ $i = 1, 2, 3$ denotando $U_\lambda = \{(x_0, z_1, z_2, z_3) : x_0 \in (0, 1)^d, z_i \in \lambda^{1/d}((0, 1)^d - x_0), i = 1, 2, 3\}$ tomando $x_0 = z_0 = 0$. Además, podemos fijarnos en que la integral sólo será no nula en el conjunto $U = \{(z_1, z_2, z_3) : (B(0, 1) \cup B(z_1, 1)) \cap (B(z_2, 1) \cup B(z_3, 1)) \neq \emptyset\}$ lo que nos lleva a

$$\begin{aligned}
\frac{C_2}{\lambda^3} &= \int_U e^{-\ell_d(B(0,1) \cup_{i=1}^3 B(z_i, 1))} \ell_d(B(0, 1) \cap B(z_1, 1)) \ell_d(B(z_2, 1) \cap B(z_3, 1)) dz_1 dz_2 dz_3 \\
&\leq \int_{B(0,2) \times B(0,6)^2} e^{-\ell_d(B(0,1) \cup_{i=1}^3 B(z_i, 1))} \ell_d(B(0, 1) \cap B(z_1, 1)) \\
&\quad \cdot \ell_d(B(z_2, 1) \cap B(z_3, 1)) dz_1 dz_2 dz_3 \\
&< \infty.
\end{aligned}$$

Ya que obviamente, entre las dos primeras bolas la distancia máxima que puede existir será una bola de radio 2, mientras que con las otras, al haber un total de cuatro bolas, la distancia máxima será la misma que la planteada en el dibujo anterior, donde teníamos que la distancia máxima era de 6 unidades.

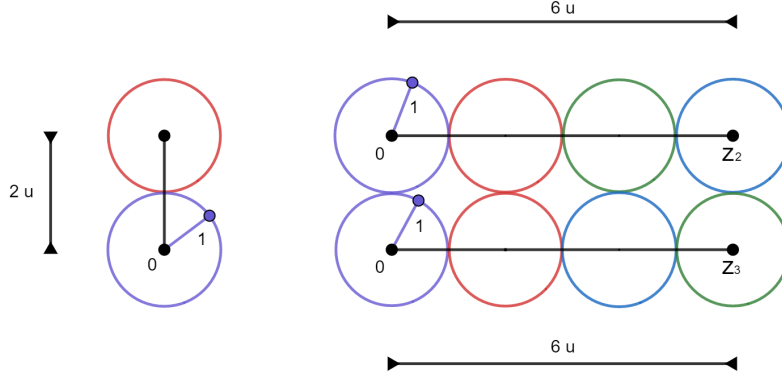


Figura 3.6: Esquema de máxima distancia entre bolas para la cota anterior.

Como antes, juntando las aportaciones llegaremos a que

$$\text{Var}\left(\int (D_z F)(-D_z L^{-1} F) d\mu(z)\right) \leq \frac{C_2}{\lambda^3}$$

Finalmente juntando las aportaciones de todos los elementos el valor de la cota vendrá dado por

$$\frac{\text{Var}\left(\int (D_z F)(-D_z L^{-1} F) d\mu(z)\right)^{1/2}}{\text{Var}(F)} \leq \frac{C_2^{1/2}}{\lambda^{1/2} C(\lambda)},$$

$$d_W\left(\mathcal{L}\left(\frac{F(\eta_\lambda) - \mathbb{E}F(\eta_\lambda)}{\sqrt{\text{Var}F(\eta_\lambda)}}\right), N(0, 1)\right) \leq \frac{1}{\sqrt{\lambda}} \frac{C_1^{1/2} + C_2^{1/2}}{C(\lambda)} \sim \frac{1}{\sqrt{\lambda}}.$$

Todos los cálculos anteriores han demostrado el teorema siguiente.

Teorema 3.11. *Si $F(\eta_\lambda)$ es el funcional definido en (3.26) entonces existe una constante $C > 0$ tal que*

$$d_W\left(\mathcal{L}\left(\frac{F(\eta_\lambda) - \mathbb{E}F(\eta_\lambda)}{\sqrt{\text{Var}F(\eta_\lambda)}}\right), N(0, 1)\right) \leq \frac{C}{\sqrt{\lambda}}, \quad \lambda \geq 1.$$

En particular

$$\frac{F(\eta_\lambda) - \mathbb{E}F(\eta_\lambda)}{\sqrt{\text{Var}F(\eta_\lambda)}} \xrightarrow[\lambda \rightarrow \infty]{d} \mathcal{N}(0, 1),$$

Y por lo tanto tendremos un Teorema Central del Límite para la ley de distribución del volumen de bolas de radio $\lambda^{-1/d}$ sobre un proceso de Poisson definido en todo \mathbb{R}^d y limitado a un espacio $[0, 1]^d$

Esto nos muestra cómo el método de Stein puede ser aplicado de forma satisfactoria en funcionales más generales definidos sobre espacios de variables aleatorias como el de Poisson.

3.4. Apéndice.

Reunimos en este apéndice algunas demostraciones relativas al desarrollo en caos de Poisson.

Demostración del Lema 3.1. Supongamos f simple, es decir, de la forma

$$f(x) = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}(x)$$

con los A_i medibles y disjuntos dos a dos. Entonces

$$I_1(f) = \sum_{i=1}^k \alpha_i (\eta(A_i) - \mu(A_i))$$

y, trivialmente, $E(I_1(f)) = 0$. Además

$$(I_1(f))^2 = \sum_{i=1}^k \alpha_i^2 (\eta(A_i) - \mu(A_i))^2 + \sum_{j \neq i} \alpha_i \alpha_j (\eta(A_i) - \mu(A_i)) (\eta(A_j) - \mu(A_j)).$$

Por la independencia el segundo sumando es centrado y comprobamos que

$$E[(I_1(f))^2] = \sum_{i=1}^k \alpha_i^2 \mu(A_i) = \int f^2 d\mu.$$

De forma más general, si f y g son simples, con

$$g(x) = \sum_{i=1}^k \beta_i \mathbf{1}_{A_i}(x)$$

(esto no supone pérdida de generalidad) tendremos que

$$I_1(f)I_1(g) = \sum_{i=1}^k \alpha_i \beta_i (\eta(A_i) - \mu(A_i))^2 + \sum_{j \neq i} \alpha_i \beta_j (\eta(A_i) - \mu(A_i)) (\eta(A_j) - \mu(A_j)),$$

de donde

$$E[I_1(f)I_1(g)] = \sum_{i=1}^k \alpha_i \beta_i \mu(A_i) = \int fg d\mu.$$

Para entender mejor el caso de funciones de varias variables tratamos primero el caso $n = 2$ y $f(x_1, x_2) = \mathbf{1}_{A_1}(x_1)\mathbf{1}_{A_2}(x_2) + \mathbf{1}_{A_1}(x_2)\mathbf{1}_{A_2}(x_1)$ con A_1

y A_2 disjuntos. Entonces, teniendo en cuenta que $f(x_1, x_2) = 0$ si $x_1 = x_2$ obtenemos

$$\begin{aligned}
I_2(f) &= \sum_{z_1 \in \eta, z_2 \in \eta} f(z_1, z_2) - \sum_{z_1 \in \eta} \int f(z_1, z_2) d\mu(z_2) \\
&\quad - \sum_{z_1 \in \eta} \int f(z_1, z_2) d\mu(z_2) + \int f(z_1, z_2) d\mu(z_1) d\mu(z_2) \\
&= 2(\eta(A_1)\eta(A_2) - \mu(A_1)\eta(A_2) - \eta(A_1)\mu(A_2) + \mu(A_1)\mu(A_2)) \\
&= 2(\eta(A_1) - \mu(A_1))(\eta(A_2) - \mu(A_2)).
\end{aligned}$$

En general, si $n \geq 2$ consideramos la función

$$f(x_1, \dots, x_n) = \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \mathbf{1}_{A_1}(x_{i_1}) \cdots \mathbf{1}_{A_n}(x_{i_n})$$

con A_1, \dots, A_n disjuntos dos a dos. De la igualdad

$$I_n(f) = \sum_{I \subset [n]} (-1)^{n-|I|} \int_{(\mathbb{R}^d)^n} f(x_1, \dots, x_n) d(\eta^{(I)} \otimes \mu^{([n]-I)})(x_1, \dots, x_n)$$

obtenemos que, en este caso,

$$I_n(f) = n!(\eta(A_1) - \mu(A_1)) \cdots (\eta(A_n) - \mu(A_n)). \quad (3.35)$$

Por la independencia, deducimos

$$E(I_n(f)) = 0, \quad \text{Var}(I_n(f)) = (n!)^2 \mu(A_1) \cdots \mu(A_n) = n! \int f^2 d\mu^n. \quad (3.36)$$

Manejamos ahora una función simple de n variables de varias variables, de la forma

$$f(x_1, \dots, x_n) = \sum_{i_1=1}^k \cdots \sum_{i_n=1}^k \alpha_{i_1, \dots, i_n} \mathbf{1}_{A_{i_1}}(x_1) \cdots \mathbf{1}_{A_{i_n}}(x_n), \quad (3.37)$$

donde asumimos $k \geq n$, A_1, \dots, A_k son disjuntos dos a dos, $\alpha_{i_1, \dots, i_k} = 0$ si en (i_1, \dots, i_k) hay algún índice repetido y $\alpha_{i_1, \dots, i_n} = \alpha_{\sigma(i_1), \dots, \sigma(i_n)}$ para cualquier permutación de los índices i_1, \dots, i_k distintos, para garantizar que f es simétrica. Escribiendo α_I en lugar de α_{i_1, \dots, i_k} si $I = \{i_1, \dots, i_n\} \subset \{1, \dots, k\}$ tiene cardinal n podemos reescribir f en la forma

$$f(x_1, \dots, x_n) = \sum_{I \subset \{1, \dots, k\}: |I|=n} \alpha_I f_I(x_1, \dots, x_n)$$

con

$$f_I(x_1, \dots, x_n) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \mathbf{1}_{A_{i_1}}(x_{j_1}) \cdots \mathbf{1}_{A_{i_n}}(x_{j_n})$$

si $I = \{i_1, \dots, i_n\} \subset \{1, \dots, k\}$. Aplicando los cálculos anteriores obtenemos $I_n(f_I) = n! \prod_{i \in I} (\eta(A_i) - \mu(A_i))$ y, por linealidad,

$$I_n(f) = \sum_{I \subset \{1, \dots, k\}: |I|=n} \alpha_I I_n(f_I).$$

Concluimos que $E(I_n(f)) = 0$. Además, si $I \neq J$ claramente $E(I_n(f_I)I_n(f_J)) = 0$ (en el producto $\prod_{i \in I} (\eta(A_i) - \mu(A_i)) \prod_{j \in J} (\eta(A_j) - \mu(A_j))$ aparece, necesariamente, algún factor $(\eta(A_i) - \mu(A_i))$ una sola vez). Esto implica que

$$\begin{aligned} \text{Var}(I_n(f)) &= \sum_{I \subset \{1, \dots, k\}: |I|=n} \alpha_I^2 \text{Var}(I_n(f_I)) \\ &= (n!)^2 \sum_{I \subset \{1, \dots, k\}: |I|=n} \alpha_I^2 \prod_{i \in I} \mu(A_i) = n! \int f^2 d\mu^n. \end{aligned} \quad (3.38)$$

Como resultado de los argumentos anteriores hemos probado que si f es una función simple, de la forma (3.37) entonces

$$E(I_n(f)) = 0, \quad \text{Var}(I_n(f)) = n! \int f^2 d\mu^n. \quad (3.39)$$

Un argumento de aproximación permite extender (3.39) a cualquier f simétrica de cuadrado integrable. A partir de (3.39) y de la igualdad $\text{Cov}(X, Y) = (\text{Var}(X + Y) - \text{Var}(X) - \text{Var}(Y))/2$ se concluye

$$\text{Cov}(I_n(f)I_n(g)) = n! \int fg d\mu^n. \quad (3.40)$$

Finalmente, para probar que

$$\text{Cov}(I_n(f)I_m(g)) = 0$$

si $n \neq m$ es suficiente considerar el caso $m > n$,

$$\begin{aligned} f(x_1, \dots, x_n) &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \mathbf{1}_{A_1}(x_{i_1}) \cdots \mathbf{1}_{A_n}(x_{i_n}), \\ g(x_1, \dots, x_m) &= \sum_{j_1=1}^m \cdots \sum_{j_m=1}^m \mathbf{1}_{A_1}(x_{j_1}) \cdots \mathbf{1}_{A_m}(x_{j_m}), \end{aligned}$$

con $A_1, \dots, A_n, \dots, A_m$ disjuntos dos a dos. Ahora (3.40) es obvio a partir de (3.35). Por linealidad y densidad se concluye (3.40) en el caso general. Esto completa la demostración. \square

Demostración del Teorema 3.2.

Para demostrar el Teorema 3.2 debemos completar una serie de etapas. Para abreviar, si $F = F(\eta)$ es un funcional de cuadrado integrable medible en la σ -álgebra generada por el proceso de Poisson η , entonces escribiremos $F \in L^2(\mathbb{P}_\eta)$. Definimos $TF := (T_n F)_{n \geq 0}$. La conclusión del Teorema 3.2 es que $TF \in \mathbf{H}$ para $F \in L^2(\mathbb{P}_\eta)$ y

$$\mathbb{E}F(\eta)G(\eta) = \langle TF, TG \rangle_{\mathbf{H}}, \quad F, G \in L^2(\mathbb{P}_\eta). \quad (3.41)$$

La idea de la demostración será la siguiente. En primer lugar comprobamos el resultado para funciones de la forma $\mu \rightarrow \exp[-\int v(y)\mu(dy)]$. Tras esto vemos que el conjunto G de todas las combinaciones lineales de funciones del tipo anterior es denso sobre el espacio $L_2(\Omega, \mathcal{F}, \mathbb{P}) = L_2(\mathbb{P}_\eta)$. Finalmente usando que el espacio es de Hilbert junto con argumentos de completitud extendemos el resultado desde G a $L_2(\mathbb{P}_\eta)$.

Comenzamos entonces considerando el espacio \mathbf{F}^+ de funciones medibles acotadas.

$$\mathbf{F}^+ = \{v : B \rightarrow \mathbb{R}_+ : v \text{ medible y acotada}\},$$

Consideramos el *funcional de Laplace* del proceso η , dado por

$$\mathbb{E}(\exp(-\eta(v))), v \in \mathbf{F}^+.$$

Si la función $v \in \mathbf{F}^+$ es simple, es decir, si puede escribirse en la forma siguiente

$$v = \sum_{j=1}^k v_j \mathbf{1}_{A_j} \quad A_1, \dots, A_k \subset B \text{ medibles y disjuntos dos a dos,}$$

entonces

$$\eta(v) = \int_B v d\eta = \sum_{j=1}^k v_j \eta(A_j),$$

En este caso

$$\mathbb{E}(\exp(-\eta(v))) = \mathbb{E}(e^{-\sum_{j=1}^k v_j \eta(A_j)}) = \prod_{j=1}^k \mathbb{E}(e^{-v_j \eta(A_j)}).$$

Un cálculo elemental muestra que si $Z \sim Po(\mu)$ entonces

$$\mathbb{E}(e^{-xZ}) = \sum_{j=0}^{\infty} (e^{-x})^j \frac{\mu^j e^{-\mu}}{j!} = e^{-\mu} \sum_{j=0}^{\infty} \frac{(\mu e^{-x})^j}{j!} = e^{-\mu} e^{\mu e^{-x}} = e^{-\mu(1-e^{-x})}.$$

Aplicando esto a la expresión anterior se deduce que

$$\begin{aligned} \prod_{j=1}^k \mathbb{E}(e^{-v_j \eta(A_j)}) &= \prod_{j=1}^k \exp(-\lambda \ell_d(A_j)(1 - e^{-v_j})) \\ &= \exp\left(-\lambda \sum_{j=1}^k (1 - e^{-v_j}) \ell_d(A_j)\right) \\ &= \exp\left(-\lambda \int_B (1 - e^{-v(x)}) dx\right). \end{aligned}$$

Concluimos que si $v \in \mathbf{F}^+$ es simple entonces

$$\mathbb{E}(\exp(-\eta(v))) = \exp\left(-\lambda \int_B (1 - e^{-v(x)}) dx\right). \quad (3.42)$$

Para $v \in \mathbf{F}^+$ general se comprueba (por el Teorema de Convergencia Monótona) que la igualdad anterior sigue siendo válida.

Denotamos como \mathbf{G} el espacio de funciones medibles y acotadas $G : \mathbf{N} \rightarrow \mathbb{R}$ que se expresan de la forma $G(\mu) = a_1 e^{-\mu(v_1)} + \dots + a_n e^{-\mu(v_n)}$ con $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$ y $v_1, \dots, v_n \in \mathbf{F}^+$. Demostramos a continuación que se cumple la ecuación principal de la representación en caos de Poisson para cualquier funcional de la forma anterior.

Lema 3.12. *Para cualesquiera $F, G \in \mathbf{G}$ se cumple la ecuación*

$$\mathbb{E}F(\eta)G(\eta) = \langle TF, TG \rangle_H, \quad F, G \in L^2(\mathbb{P}_\eta).$$

Demostración. Sea $|J|$ el cardinal de J , entonces podemos expresar el operador en diferencias múltiple a través de la ecuación siguiente

$$D_{y_1, \dots, y_n}^n F(\eta) = \sum_{J \subset \{1, 2, \dots, n\}} (-1)^{n-|J|} F\left(\eta + \sum_{j \in J} \delta_{y_j}\right), \quad (3.43)$$

es fácil ver que por linealidad, será suficiente con considerar la función F en la forma

$$F(\eta) = \exp(-\eta(v)).$$

Tomamos el funcional de la forma convenida y aplicando el operador de diferencias sobre él

$$D_y F(\eta) = \exp(-(\eta(v) + \delta(y)))(-\exp(\eta(v))) = \exp(-\eta(v))(e^{-v(y)} - 1).$$

Expresando

$$\eta(v) = \int_B v d\eta \quad \Rightarrow \quad (\eta + \delta_y)(v) = \eta(v) + v(y),$$

podemos fijarnos en cómo se comporta para dos variables y tendremos

$$\begin{aligned} D_{y_1, y_2} F(\eta) &= F(\eta + \delta_{y_1} + \delta_{y_2}) - F(\eta + \delta_{y_1}) - F(\eta + \delta_{y_2}) + F(\eta) \\ &= \exp(-\eta(v) - v(y_1) - v(y_2)) \\ &= \exp(-\eta(v)) [\exp(-(v(y_1) + v(y_2))) - \exp(-v(y_1)) - \exp(-v(y_2)) + 1] \\ &= \exp(-\eta(v)) (-1 + \exp(-v(y_1))) (-1 + \exp(-v(y_2))). \end{aligned}$$

Y de forma general

$$D^n F(\eta) = \exp[-\eta(v)] (e^{-v} - 1)^{\otimes n},$$

por lo que podemos llegar a que

$$T_n F(\eta) = \exp \left(-\lambda \int_B v(x) dx \right) (e^{-v} - 1)^{\otimes n}.$$

Si ahora consideramos otra función $G = \exp[-\eta(w)]$ podemos entonces desarrollar cada una de las expresiones buscadas. Comenzamos desarrollando la expresión $\langle T_n F, T_n G \rangle$.

$$\begin{aligned} \langle T_n F, T_n G \rangle &= \exp \left(-\lambda \int_B v(x) dx \right) \exp \left(-\lambda \int_B w(x) dx \right) \\ &\quad \cdot \int_{B^n} \prod_{j=1}^d (e^{-v(y_j)} - 1) (e^{-w(y_j)} - 1) dy, \end{aligned}$$

y por lo tanto se tiene

$$\begin{aligned} \langle TF, TG \rangle &= \sum_{n=0}^{\infty} \frac{\langle T_n F, T_n G \rangle}{n!} = \exp \left(-\lambda \int_B v(x) dx \right) \exp \left(-\lambda \int_B w(x) dx \right) \\ &\quad \exp \left(\lambda \int_B (e^{-v(x)} - 1) (e^{-w(x)} - 1) dx \right). \end{aligned}$$

Desarrollando por otro lado la expresión $\mathbb{E}(F(\eta)G(\eta))$,

$$\begin{aligned} \mathbb{E}(F(\eta)G(\eta)) &= \mathbb{E}(\exp(-\eta(v+w))) \\ &= \exp \left(-\lambda \int_B (1 - e^{-(v(x)+w(x))}) dx \right), \end{aligned}$$

por lo que agrupando los términos de $\langle TF, TG \rangle$ llegamos conseguiremos la igualdad buscada

$$\begin{aligned}
\langle TF, TG \rangle &= \exp \left(-\lambda \int_B v(x) dx - \lambda \int_B w(x) dx \right) \\
&\quad \exp \left(\lambda \int_B (e^{-v(x)} - 1)(e^{-w(x)} - 1) dx \right) \\
&= \exp \left(-\lambda \int_B 2 - e^{-v(x)} - e^{-w(x)} - e^{-(v(x)+w(x))} - 1 + e^{-v(x)} + e^{-w(x)} dx \right) \\
&= \exp \left(-\lambda \int_B 1 - e^{-(v(x)+w(x))} dx \right) = \mathbb{E}(F(\eta)G(\eta)).
\end{aligned}$$

Lo cual finaliza la prueba. ■

Una vez tenemos esta igualdad probada, nuestro siguiente paso para llegar a nuestro objetivo será extender el resultado a funcionales generales. Con este propósito comenzamos probando el lema siguiente.

Lema 3.13. *El conjunto \mathbf{G} es denso en $L^2(P_\eta)$*

Demostración. Nos basamos en el siguiente resultado que encontramos en [12] (Th 21 cap. 1). *Sea \mathcal{H} es un subespacio de $L^0(\Omega, \mathcal{F}, \mathbb{P})$ que contiene a las constantes, es cerrado para convergencia uniforme y tal que si $f_n \in \mathcal{H}, 0 \leq f_n \leq k; y f_n \uparrow f$ entonces $f \in \mathcal{H}$. Sea $\mathcal{C} \subset \mathcal{H}$ estable para la multiplicación. Entonces \mathcal{H} contiene todas las funciones acotadas y medibles en $\sigma(\mathcal{C})$.*

En el caso que nos concierne $\mathcal{H} = L^2(P_\eta)$ y $\mathcal{C} = \mathbf{G}$. Claramente \mathbf{G} es cerrado para multiplicaciones. y por propia construcción $\sigma(\mathbf{G}) = \mathcal{F}$. Además \mathcal{H} contiene a las constantes (si $v = 0, a = cte$ entonces $\eta(v) = 0$ y por tanto $a \in \mathbf{G} \subset \mathcal{H}$). Para ver que \mathcal{H} es cerrado para límites monótonos acotados sea g_n sucesión de funciones en \mathcal{H} que cumple $0 \leq g_n \uparrow g$ y $0 \leq g_n \leq M$. Sea $g_{n,m} \in \mathbf{G}$. Entonces

$$\begin{aligned}
\mathbb{E}(g_n(\eta) - g_{n,m}(\eta))^2 &\xrightarrow{m \rightarrow \infty} 0 \\
\mathbb{E}(g_n(\eta) - g_{n,m_n}(\eta))^2 &\leq \frac{1}{2^n} \\
|g_n - g|^2 \leq M^2 &\Rightarrow \mathbb{E}(g_n(\eta) - g(\eta))^2 = \delta_n \rightarrow 0 \\
\Rightarrow \sqrt{\mathbb{E}(g_{n,m_n}(\eta) - g(\eta))^2} &\leq \sqrt{\frac{1}{2^n}} + \sqrt{\delta_n} \rightarrow 0
\end{aligned}$$

Tenemos entonces que \mathcal{H} contiene todas las funciones \mathcal{F} -medibles y acotadas, por tanto tendremos que \mathcal{H} es denso en $L^2(P_\eta) = L^2(\Omega, \mathcal{F}, \mathbb{P})$ y por tanto, llegamos a que \mathbf{G} es denso en $L^2(\Omega, \mathcal{F}, \mathbb{P})$ tal y como queríamos probar. ■

Para $F, G \in L^2(P_\eta)$ generales, el Lema 3.13 nos permite tomar $F_n, G_n \in \mathbf{G}$ tales que $\|F_n - F\|_{L^2(P_\eta)} \rightarrow 0$, $\|G_n - G\|_{L^2(P_\eta)} \rightarrow 0$. Tenemos ya probado que $\mathbb{E}F_n(\eta)G_n(\eta) = \langle TF_n, TG_n \rangle$ y, obviamente,

$$\mathbb{E}F_n(\eta)G_n(\eta) \rightarrow \mathbb{E}(F(\eta)G(\eta)).$$

Solo nos falta comprobar la convergencia $\langle TF_n, TG_n \rangle \rightarrow \langle TF, TG \rangle$ lo cual demostrará finalmente que se cumple el Teorema 3.2. Para este proposito necesitaremos demostrar la siguiente igualdad que será fundamental en la demostración y que es una herramienta básica de sistemas estocásticos de tipo Poisson (esta fórmula fue desarrollada por Mecke en 1967):

$$\mathbb{E} \int h(\eta, y) \eta(dy) = \mathbb{E} \int_B h(\eta + \delta_y, y) \lambda dy. \quad (3.44)$$

Para comprobar (3.44) comenzamos por el lado izquierdo de la ecuación para llegar a que

$$\int h(\eta, y) \eta(dy) = \begin{cases} \sum_{i=1}^N h(\eta, x_i) = \sum_{i=1}^N h(\sum_{j=1}^N \delta_{x_j}, x_i) & \text{si } N \geq 1 \\ 0 & \text{si } N = 0 \end{cases}$$

Por lo tanto

$$\mathbb{E} \left(\int h(\eta, y) \eta(dy) \mid N = n \right) = \begin{cases} n \mathbb{E} h(\sum_{j=1}^n \delta_{x_j}, x_i) & \text{si } n \geq 1 \\ 0 & \text{si } n = 0 \end{cases}$$

$$\begin{aligned} \text{de donde } \mathbb{E} \left(\int h(\eta, y) \eta(dy) \right) &= \sum_{n=1}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} n \mathbb{E} h \left(\sum_{j=1}^n \delta_{x_j}, x_i \right) \\ &= \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \mathbb{E} h \left(\sum_{j=1}^{k+1} \delta_{x_j}, x_i \right). \end{aligned}$$

Centrándonos ahora en el lado derecho de la ecuación

$$\begin{aligned} &\mathbb{E} \left(\int_B h(\eta + \delta_y, y) \lambda dy \mid N = n \right) \\ &= \begin{cases} \lambda \int_B h(\delta_y, y) dy & \text{si } n = 0 \\ \lambda \mathbb{E} \int_B h(\sum_{j=1}^n \delta_{x_j} + \delta_y, y) dy & \text{si } n \geq 1 \end{cases} \end{aligned}$$

Ahora bien,

$$\begin{aligned} \mathbb{E} h \left(\sum_{j=1}^{k+1} \delta_{x_j}, x_i \right) &= \mathbb{E} h \left(\sum_{j=1}^{k+1} \delta_{x_j}, x_{k+1} \right) \\ &= \int_{B^{k+1}} h \left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y \right) dx_1, \dots, dx_n dy \end{aligned}$$

$$\mathbb{E}h\left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y\right) dy = \int_{B^{k+1}} h\left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y\right) dx_1, \dots, dx_n dy$$

Llegamos entonces a que el lado derecho puede expresarse como

$$\begin{aligned} \mathbb{E}\left(\int_B h(\eta + \delta_y, y) \lambda dy \mid N = n\right) &= \lambda \int_{B^{k+1}} h\left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y\right) dx_1, \dots, dx_n dy \\ &\Rightarrow \mathbb{E} \int_B h(\eta + \delta_y, y) \lambda dy \\ &= \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} \lambda \int_{B^{k+1}} h\left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y\right) dx_1, \dots, dx_n dy \end{aligned}$$

De la misma forma tenemos que

$$\mathbb{E} \int h(\eta, y) \eta(dy) = \lambda \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} \int_{B^{k+1}} h\left(\sum_{j=1}^n \delta_{x_j} + \delta_y, y\right) dx_1, \dots, dx_n dy$$

Y por tanto hemos demostrado (3.44).

Para finalizar, nos queda demostrar el siguiente lema.

Lema 3.14. Sean $F, F^1, F^2, \dots \in L^2(\mathbb{P}_\eta)$ satisfaciendo $F^k \rightarrow F$ en $L^2(\mathbb{P}_\eta)$ cuando $k \rightarrow \infty$ y de tal forma que $h \rightarrow [0, 1]$ es medible. Entonces

$$\lim_{k \rightarrow \infty} \int_{B^n} \mathbb{E}[|D_{x_1, \dots, x_n}^n F(\eta) - D_{x_1, \dots, x_n}^n F^k(\eta)| h(\eta)] \mu^n(dx_1 \dots dx_n) = 0. \quad (3.45)$$

Demostraci3n. Necesitamos probar la convergencia de esta expresi3n, por (3.17) la convergencia de (3.19) estar3a asegurada si comprobamos la convergencia de

$$\lim_{n \rightarrow \infty} \int_{B^n} \mathbb{E} \left[\left| F\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) - F^k\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) \right| h(\eta) \right] \mu^n(dx_1 \dots dx_n) = 0.$$

Sea $m \in \{0, \dots, n\}$ entonces tendremos la igualdad siguiente para la integral

$$\begin{aligned} &\int_{B^n} \mathbb{E} \left| F\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) - F^k\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) \right| h(\eta) \mu^m(dx_1 \dots dx_m) \\ &= \mu(B)^{n-m} \mathbb{E} \int_{B^m} \left| F\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) - F^k\left(\eta + \sum_{i=1}^m \delta_{x_i}\right) \right| h(\eta) \mu^m(dx_1 \dots dx_m), \end{aligned}$$

y utilizando la ecuación de Mecke (3.18) tenemos la igualdad

$$\begin{aligned} & \mu(B)^{n-m} \mathbb{E} \int_{B^m} \left| F \left(\eta + \sum_{i=1}^m \delta_{x_i} \right) - F^k \left(\eta + \sum_{i=1}^m \delta_{x_i} \right) \right| h(\eta) \mu^m(dx_1 \dots dx_n) \\ &= \mu(B)^{n-m} \mathbb{E} \int_{B^m} |F(\eta) - F^k(\eta)| h \left(\eta - \sum_{i=1}^n \delta_{x_i} \right) \eta^{(m)}(dx_1 \dots dx_n). \end{aligned}$$

Esto es sencillo de acotar pues tendremos que la ecuación anterior tendrá un valor a lo sumo como el de la expresión siguiente.

$$\leq \mu(B)^{n-m} \mathbb{E} |F(\eta) - F^k(\eta)| \eta^m(B^m)$$

Utilizando entonces la desigualdad de Cauchy-Schwarz podemos acotar esta última expresión por

$$\mu(B)^{n-m} (\mathbb{E}(F(\eta) - F^k(\eta))^2)^{1/2} (\mathbb{E}[(\eta^{(m)}(B^m))^2])^{1/2}.$$

Ahora bien, es obvio que tomando el límite tenemos que este converge a 0 para la ecuación (2.29) lo cual prueba el lema. \blacksquare

Esto prueba finalmente el aserto que nos asegura que

$$\langle T f_n, T g_n \rangle \rightarrow \langle T f, T g \rangle,$$

lo cual nos permite asegurar finalmente la validez de la representación en caos de Poisson que vendrá dada por

$$F(\eta) = \mathbb{E}F(\eta) + \sum_{n=1}^{\infty} \frac{1}{n!} I_n(f_n) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(f_n) = \sum_{n=0}^{\infty} I_n(\tilde{f}_n), \quad (3.46)$$

siendo $f_n(x_1, \dots, x_n) = \mathbb{E}D_{(x_1, \dots, x_n)}F(\eta)$ y $\tilde{f}_n = \frac{f_n}{n!}$.

Bibliografía

- [1] BARBOUR, A.D. & CHEN, L. H. Y. (2005). *An introduction to Stein's Method*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Vol 4. (Cited on pp. 3, 4 y 12)
- [2] BARBOUR, A.D. & CHEN, L. H. Y. & SHAO, Q.M. (2011). *Normal approximation by Stein's Method*, Probability and its applications. (Cited on pp. 4, 10 y 12)
- [3] BARBOUR, A.D. & KARONSKI, M. & RUCINSKI, A. (1989). *A central limit theorem for descomposable random variable with applications to random graphs*, J.Combin. Theory Ser. B 47. **2** 125–145. (Cited on pp. 5 y 30)
- [4] BILLINGSLEY, P. (1986). *Probability and Measure*, Wiley, New York. (Cited on pp. 1 y 2)
- [5] CHATTERJEE, S. (2014). *A short survey of Stein's method.*, Stanford University. (Cited on p. 4)
- [6] CHEN, L. H. Y. & SHAO, Q.M. (2004). *Normal approximation under local dependence*, Ann. Probab. **32**, págs. 1985–2028. (Cited on p. 25)
- [7] DEL BARRIO, EUSTASIO (2015). *Berry-Esseen bounds for weighted averages of Poisson avoidance functionals*, Universidad de Valladolid, IMUVA. (Cited on p. 67)
- [8] ITO, K. (1951). *Multiple Wiener integral*, Journal of the Mathematical Society of Japan. **3**, No.1, págs. 157–169. (Cited on p. 49)
- [9] KINGMAN, J.F.C. (1993). *Poisson Processes*, Oxford University Press. (Cited on p. 44)
- [10] LAST, G. & PENROSE, M.D. (2011). *Poisson process Fock space representation, chaos expansion and covariance inequalities*, Probab. Theory Relat. Fields, **150**, págs. 663–690. (Cited on pp. 5 y 46)
- [11] LAST, G. & PENROSE, M.D. (2017). *Lectures on the Poisson Process*, IMS Textbook by Cambridge University Press. (Cited on pp. 5 y 49)

- [12] MEYER, P. & DELLACHERIE, C. (1997). *Probabilités et Potentiel*, (Cited on p. 76)
- [13] NUALART, D. & VIVES, J. (1990). *Anticipative calculus for the Poisson process based on the fock space*, Séminaire de probabilités (Strasbourg), **24**, págs. 154—165. (Cited on pp. 43 y 50)
- [14] PECCATI, G. & SOLÉ, J.L. & TAQQU, M.S. & UTZET, F. (2010). *Stein's method and Normal approximation of Poisson functionals*, The Annals of Probability, **38**, No.2, págs. 443–478. (Cited on pp. 5, 46 y 57)
- [15] RAIC, M. (2014). *Normal approximations by Steins method*, Proceedings of the Seventh Young Statiscians Meeting. Metodoloki zvezki, **21**, 71–97 University of Ljubljana, Slovenia. (Cited on p. 4)
- [16] RINOTT, Y. & ROTAR, V. (1996). *A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate and applications to multivariate graph related statistics*, J. Multivariate Analysis, **56**, págs. 333 –350. (Cited on pp. 29 y 30)
- [17] RINOTT, Y. & ROTAR, V. (2000). *Normal Approximations by Stein's method*, Decis. Econ. Finance, **23**, págs. 15 –29. (Cited on p. 4)
- [18] ROSS, N (2011). *Fundamentals of Stein's method*, Probability Surveys, **8**, págs. 210—293. (Cited on pp. 4 y 5)
- [19] STEIN, C. (1972). *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Vol **II**. Probab. theory, 583–602. Berkeley, California. (Cited on pp. 3 y 7)
- [20] STEIN, C. (1986). *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes-Monograph Series, **7**, Institute of Mathematical Statistics, Hayward, California. (Cited on pp. 4 y 10)