



CIMAT

Centro de Investigación en Matemáticas, A.C.

---

# MATRICES ALEATORIAS Y PROBABILIDAD LIBRE PARA EXTENSIÓN DE MODELOS 2D-ARMA

**T E S I S**

Que para obtener el grado de

**Maestro en Ciencias**

con Orientación en

**Probabilidad y Estadística**

**Presenta**

Ricardo José López Dawn

**Director de Tesis:**

Dr. Víctor Pérez-Abreu Carrión

---

**Autorización de la versión final**



## **Dedicatoria**

*Dedico esta tesis a mis padres y hermanos, quienes me brindaron en todo momento su apoyo incondicional durante este trayecto*

# Agradecimientos

Es un honor para mí agradecer a través de estas líneas a todas aquellas personas y colegas, quienes me han brindado su apoyo durante el proceso de investigación y redacción de este trabajo. En primer lugar, quisiera agradecer a mis padres, quienes me han ayudado y respaldado en toda mi tesis. A mi asesor, el Dr. Víctor Pérez-Abreu Carrión por aceptar dirigir la tesis, por sus enseñanzas, por el apoyo, por el tiempo dedicado, por su comprensión y por haberme orientado en todos los momentos que necesité de sus consejos. Esto no sería posible sin su ayuda, experiencia y dirección.

Doy gracias a mis sinodales, el Dr. Enrique Villa Diharce y al Dr. Carlos Vargas Obieta quienes con sus aportaciones lograron que con éxito concluyera esta etapa de mi vida. También a las Dra. Eloísa Díaz Francés, Dra. Leticia Ramírez Ramírez y al Dr. Miguel Nakamura Savoy por su apoyo durante toda la maestría, así como su tutoría y cursos que impartieron. Por todo esto, muchas gracias por sus recomendaciones y orientación durante toda este proceso.

Agradezco a mis compañeros de la maestría, gracias a ustedes mi estancia en Cimat ha sido mucho más cómoda y agradable. También agradezco a mis compañeros de maestría graduados, entre ellos, Santiago Arenas ayudándome a dar orientación en diversas cuestiones de esta tesis. Muchas gracias por su amistad

y apoyo.

También agradezco a mis padres, Lourdes y Ricardo y a mis hermanos Aurora, Lillian y Jordi, que desde la distancia siempre me estuvieron apoyando y dándome ánimos y buenas energías. Reconozco su gran ayuda en especial a mi mamá, Lourdes y mi hermana Aurora por su contribución en la revisión de esta tesis.

Finalmente, quiero agradecer a CONACYT por la beca de maestría que me fue otorgada, al igual que al Centro de Investigación en Matemáticas A.C (CIMAT A.C) por todas las facilidades y apoyos para participar en eventos.

# Índice general

<b>Agradecimientos</b>	III
<b>1. Preliminares</b>	1
1.1. Modelos 2D-ARMA . . . . .	1
1.2. Elementos de probabilidad libre y matrices aleatorias . . . . .	4
<b>2. Equivalentes Deterministas Libres para 2D-ARMA</b>	11
2.1. Equivalentes Deterministas Libres (EDL) . . . . .	12
2.2. Aplicación a modelos 2D-ARMA . . . . .	18
<b>3. EDL con errores función de Poisson para 2D-ARMA</b>	23
3.1. 2D-ARMA con errores función de Poisson . . . . .	24
3.2. EDL con errores función de Poisson . . . . .	28
3.3. Aplicación a 2D-ARMA con errores función de Poisson . . . . .	33
3.4. Estudio de simulación . . . . .	37
3.4.1. Aplicado a EDL y EDL de Anscombe . . . . .	37
3.4.2. Aplicado a 2D-ARMA y 2D-ARMA de Anscombe . . . . .	46
<b>4. Discusión y Conclusiones</b>	56
<b>Bibliografía</b>	57

## Introducción

Desde su descubrimiento por Röntgen, los rayos X han sido objeto de investigación en una amplia variedad de campos. Uno de los primeros usos de los rayos X fue en el campo médico para obtener información sobre el interior del cuerpo sin disección. Además de la técnica básica de imágenes de rayos X de cortes fijos, se han desarrollado varios sistemas de imágenes, como lo son los ultrasonidos y por resonancia magnética. Un haz de rayos X es parcialmente absorbido por estructuras internas del cuerpo en un proceso conocido como atenuación, y el detector en el otro lado del cuerpo absorbe estos rayos X atenuados para producir una imagen de rayos X.

Varios de los algoritmos existentes en la literatura clásica de procesamiento de imágenes y señales del ruido de rayos X suponen una distribución de ruido Poisson (Cesarelli M. et al. [14]), con media representando la desviación media de fotones a través de la imagen obtenida por rayos X (con media generalmente mayores a 10). Sin embargo, Elbakri y Fessler [15] mostraron que el ruido real combina las distribuciones de Poisson y Gaussiana, con la distribución de Poisson derivada del ruido cuántico y la distribución Gaussiana proveniente del ruido térmico generado por los sensores y otros dispositivos electrónicos.

Debido a que el ruido de Poisson depende de la señal, es difícil diseñar un algoritmo general; para abordar esto, se han introducido transformaciones de estabilización de varianza. La transformada de Anscombe dirigida al ruido de Poisson fue propuesta por Anscombe [16] en 1948 y se extendió a la transformada de Anscombe generalizada y la transformada inversa para el ruido de Poisson-Gaussian. Entre otros aspectos, los modelos 2D-ARMA fueron introducidos por Zielinki et al. [8] en 2010, los cuales, a diferencia de los modelos ARMA de la estadística clásica definidos a mediados del siglo XX por Box y Jenkins radica en que los primeros son considerados en latices bi-dimensionales y enunciados con el objetivo

de clasificar y predecir distintos tipos de cáncer, detectar tuberculosis e incluso para inspección de calidad en relojes [8, 12, 13]. Es de mencionar que se han visto resultados exitosos en el uso de los modelos 2D-ARMA, específicamente para imágenes en rayos X [8].

A su vez, desde mediados del siglo XX e inicios del siglo XXI se ha desarrollado extensivamente la teoría de matrices aleatorias y de probabilidad libre, dando de este modo una nueva forma de pensar y abordar problemas. Es por ello que durante la última década se ha aplicado la teoría de matrices aleatorias y de probabilidad libre para el estudio de modelos ARMA en general y, recientemente en 2019 por Hayase [1] en modelos 2D-ARMA. Los intereses en el uso de estas herramientas son variados, pues consideran hipótesis desde las matrices de correlación de retardo [3, 10, 11] y más generalmente en los modelos Wishart compuestos introducidos por Speicher [6] y posteriormente estudiados por Bryc [7].

Con las herramientas de matrices aleatorias y probabilidad libre bien fundamentadas, los equivalentes deterministas libres (EDL) fueron definidos por primera vez por Speicher y Vargas [9] en 2012. En el 2014 Redelmeier [5] encuentra expresiones de los cumulantes clásicos de los modelos de Wishart compuestos usando libertad de segundo orden, que subsecuentemente son empleados por [1, 2, 3].

Es de notar que las matrices de Wishart compuestas y las teorías derivadas de ellas consideran genéricamente matrices cuyas entradas son v.a.i.i.d. con distribución normales estándar. Por ende, si estamos interesados en cambiar distribucionalmente las entradas de las matrices a considerar, entonces los resultados hallados que hagan uso de la hipótesis de normalidad no son necesariamente ciertos.

Hayase [1], emplea teoría de EDL y resultados de libertad de segundo orden para relacionar modelos 2D-ARMA con matrices de Wishart compuestas y encuentra



cantidades pivotaes de interés y potenciales para estudios estadísticos. Dado que [1] considera errores normales, y debido a que los modelos 2D-ARMA son usados en imágenes de rayos X, entonces es razonable usar errores con distribución función Poisson (centralización de transformada de Anscombe) y no distribuciones normales, pero preservando la independencia.

De este modo, esta tesis efectúa un puente al modificar la distribución de los errores en los modelos 2D-ARMA por distribuciones función Poisson y producir cantidades pivotaes aplicables a tales modelos similares a las de [1]. De esta forma, obtenemos pivotaes al modificar los EDL (el cual llamaremos EDL de Anscombe, pese a que nosotros lo definimos y estudiamos) mediante el cambio de la hipótesis distribucional de normales estándar a funciones Poisson. A su vez, conseguimos pivotaes derivados de un modelo 2D-ARMA (los cuales llamaremos modelos 2D-ARMA de Anscombe, pese a que nosotros los definimos y estudiamos) con distribución de error como las antes mencionadas y, finalmente efectuamos un estudio de simulación para analizar la velocidad de convergencia de las cantidades pivotaes halladas y compararlas con las de [1].

El objetivo principal de esta tesis es presentar conexiones entre los modelos 2D-ARMA definidos en [8], los resultados de [1] y el cambio distribucional a los errores función Poisson (transformada de Anscombe centralizada). Este enfoque permite obtener cantidades pivotaes asintóticas para los modelos 2D-ARMA usando herramientas de matrices aleatorias y probabilidad libre.

La organización y presentación de la tesis y nuestros resultados son los siguientes.

El Capítulo 1 contiene los preliminares de modelos 2D-ARMA y matrices aleatorias y probabilidad libre necesarios para el desarrollo de la tesis. En la Sección 1.1 se introducen y definen tales modelos, se mencionan propiedades y aplicaciones, así como la noción de dichos modelos y sus interpretaciones en latices y

los principales resultados acerca de sus propiedades. En la Sección 1.2 se resumen los principales elementos de matrices aleatorias y probabilidad libre, mencionando conceptos de importancia como los cumulantes y los cumulantes libres, así como lemas que son preparatorios. Dichos conceptos nos servirán para entender el desarrollo del Capítulo 2 y posteriormente, comprender la utilidad de la introducción de un parámetro en el Capítulo 3. Éste surge de modo natural en las imágenes de rayos X, dando lugar a una propuesta de modelo que considere errores función de Poisson con varianza no homogénea en lugar de errores normales estándar.

El Capítulo 2 trata sobre la obtención de cantidades pivotaes asintóticas mediante el uso de conceptos de EDL [9] y de libertad de segundo orden [5], relacionando tales pivotaes con modelos 2D-ARMA en el uso de imágenes de rayos X. En la Sección 2.1 se presentan definiciones y resultados de [1] sobre la distribución asintótica de la traza de matrices de Wishart compuestas, así como su media y varianza. Estos resultados nos permiten obtener un resultado en la siguiente sección con interpretación en imágenes de rayos X. En la Sección 2.2 presentamos los resultados de transformar los modelos 2D-ARMA a matrices de Wishart compuestas y poder usar resultados de la Sección 2.1 para nuevamente, enunciar y probar un resultado. La implicación de tal resultado nos permite obtener cantidades pivotaes sobre modelos 2D-ARMA cuando la definición (número de píxeles) de una imagen de rayos X sea mayor o igual a un número de muestras i.i.d. de dichos modelos y tienda al infinito el número de muestra, así como las definiciones de las imágenes.

El Capítulo 3 presenta una propuesta de modelaje diferente a los modelos 2D-ARMA introducidos por Zielinski et al. [8] en 2010. Nuestra propuesta surge de la

misma naturaleza dada por el problema inicial del modelado en imágenes de rayos X, pero a diferencia de los modelos 2D-ARMA considerados por [8], nosotros consideramos la distribución de los errores como una centralización de la transformación de estabilización de varianza de la distribución Poisson. En la Sección 3.1 abordamos el motivo que origina nuestro cambio de modelaje en los errores y el nuevo modelo propuesto. La Sección 3.2 contiene las modificaciones a tomar en los equivalentes deterministas libres, así como la prueba límite de considerar errores en la transformada de Anscombe centrada. La Sección 3.3 efectúa las adecuaciones para el caso 2D-ARMA, y la aplicación de los equivalentes deterministas libres bajo nuestra modificación. Finalmente, la Sección 3.4 expone los resultados de un estudio de simulación con la finalidad de analizar la velocidad de convergencia de las cantidades pivotaes halladas con nuestra propuesta de modelaje y compararlas con las del Capítulo 2.

Finalmente, el Capítulo 4 tiene como finalidad discutir líneas para seguir extendiendo este trabajo en la misma dirección o bien en otras direcciones, así como conclusiones recolectadas de este estudio.

# Capítulo 1

## Preliminares

Este capítulo contiene los elementos de modelos 2D-ARMA, probabilidad libre y matrices aleatorias necesarios para el desarrollo de esta tesis. La Sección 1.1 contiene una introducción a tales modelos, su origen y aplicaciones. La Sección 1.2 aborda el uso de probabilidad libre y matrices aleatorias, así como su relación con la Sección 1.1. Dichos conceptos nos servirán para entender el desarrollo del Capítulo 2 y posteriormente comprender la utilidad de la introducción de un parámetro en el Capítulo 3, el cual surge de modo natural en las imágenes de rayos X, dando lugar a una propuesta de modelo que considere errores función de Poisson con varianza no homogénea en lugar de errores normales estándar.

### 1.1. Modelos 2D-ARMA

Los modelos 2D-ARMA fueron introducidos en 2010 por Zielinski [8] para la detección y clasificación de cáncer de seno, observando propiedades similares a las que presentan los modelos ARMA introducidos por Box-Jenkins. Sin embargo, diferentes en naturaleza a estos últimos, debido a su natural tratamiento en imágenes, donde una imagen la define como un campo aleatorio 2D dado por va-

riables aleatorias reales  $\{x_{n,m}|n, m \in \mathbb{Z}\}$ . Estas imágenes son usadas para definir modelos 2D-ARMA como sigue a continuación.

**Definición 1.** Un modelo 2D-ARMA( $p_1, p_2, q_1, q_2$ ) ( $p_1, p_2, q_1, q_2 \in \mathbb{N}$ ) definido para un latís de dimensiones  $H \times W \in \mathbb{N} \times \mathbb{N}$  es una familia de variables aleatorias reales  $\{x_{n,m}|(n, m) \in \{1, \dots, H\} \times \{1, \dots, W\}\}$  que satisface para toda pareja  $(n, m) \in \{1, \dots, H\} \times \{1, \dots, W\}$ :

$$\sum_{i=1}^{p_1} \sum_{j=1}^{p_2} \phi_{i,j} x_{n-i+1, m-j+1} = \sum_{i=1}^{q_1} \sum_{j=1}^{q_2} \theta_{i,j} \epsilon_{n-i+1, m-j+1}$$

donde  $\phi = (\phi_{i,j}) \in M_{p_1 \times p_2}(\mathbb{R})$ ,  $\theta = (\theta_{i,j}) \in M_{q_1 \times q_2}(\mathbb{R})$  son matrices de parámetros con  $\phi_{1,1} = 1, \theta_{1,1} \neq 0$  y  $\{\epsilon_{n,m}|n, m \in \mathbb{Z}\}$  es una familia de v.a.i.i.d. con distribución normal con media cero y varianza  $\sigma^2$ .

Los estimadores de mínimos cuadrados de los parámetros del modelo anterior son conocidos como los estimadores de Yule-Walker. Estos son obtenidos al observar que la ecuación anterior puede ser escrita de modo matricial para todo  $n = L + 1, \dots, H - 1, m = M + 1, \dots, W - 1$  y para algún  $L > \max(p_1, q_1), M > \max(p_2, q_2)$  como:

$$\mathbf{x} + \Theta \Phi = \mathbf{E}$$

donde

$$\Phi = [\phi_{1,1}, \dots, \phi_{p_1, p_2}, \theta_{1,1}, \dots, \theta_{q_1, q_2}]^t$$

$$\mathbf{E} = [\epsilon[L + 1, M + 1], \dots, \epsilon[H, W]]^t.$$

$$\mathbf{x} = [x[L + 1, M + 1], \dots, x[H, W]]^t$$

$$\Theta = \begin{pmatrix} x_{L+1, M} & \cdots & x_{L+1-p_1, M+1-p_2} & -\epsilon_{L+1, M} & \cdots & -\epsilon_{L+1-q_1, M+1-q_2} \\ x_{L+2, M} & \cdots & x_{L+2-p_1, M+1-p_2} & -\epsilon_{L+2, M} & \cdots & -\epsilon_{L+2-q_1, M+1-q_2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{H, W-1} & \cdots & x_{H-p_1, W-p_2} & -\epsilon_{H, W-1} & \cdots & -\epsilon_{H-q_1, W-q_2} \end{pmatrix}$$

Es de observar que  $HW$  indica la calidad de la imagen, y  $(p_1, p_2)$ ,  $(q_1, q_2)$  los parámetros de ‘profundidad’ en la imagen de los elementos autoregresivos y de medias móviles. En aplicaciones, como es de esperar, mientras más grande sea el tamaño de muestra, la calidad de la imagen y la "profundidad", más precisión tendremos en nuestra tasa de clasificación correcta. Esta observación fue hecha por Zielinsky et al. [8], donde usando los estimadores de Yule-Walker logran clasificar correctamente cáncer de seno al 82 % en bases de datos privadas de imágenes, con lo cual mejora al 97.8 % mediante análisis secuencial e incluso generalizando sus resultados a la detección de microcalcificaciones en mamografías. Asimismo, se han encontrado buenas tasas de clasificación en aplicaciones como la detección de tuberculosis por medio de radiografía de tórax [12] e inspección de calidad de superficies de relojes [13].

Notemos que si  $q_2 = p_2 = 1$ , tenemos el modelo clásico de modelo ARMA, y similarmente, tenemos resultados análogos al modelo ARMA clásico extendido a los modelos 2D-ARMA, esto resulta por ser de diferente naturaleza y aplicaciones, ya que están íntimamente relacionados. En el caso de que  $p_1 = p_2 = 1$ , se le conoce como un modelo 2D-MA( $q_1, q_2$ ).

A partir de un modelo 2D-ARMA( $p_1, p_2, q_1, q_2$ ), es posible obtener los polinomios  $P(z_1, z_2) = \sum_{i=1}^{p_1} \sum_{j=1}^{p_2} \phi_{i,j} z_1^{i-1} z_2^{j-1}$ ,  $Q(z_1, z_2) = \sum_{i=1}^{q_1} \sum_{j=1}^{q_2} \theta_{i,j} z_1^{i-1} z_2^{j-1}$ .

Si el polinomio  $P$  no se anula en el disco unitario, entonces al modelo se le llamará modelo reversible. Una propiedad importante de los modelos reversibles es que todo 2D-ARMA( $p_1, p_2, q_1, q_2$ ) puede ser llevado a un modelo asintótico 2D-MA( $o_1, o_2$ ) ( $o_1, o_2 \rightarrow \infty$ ) reversible, es decir, todo modelo 2D-ARMA puede ser aproximado asintóticamente como un modelo 2D-MA, lo cual permite la optimización en tiempo y memoria de algoritmos que usen tales modelos.

Observemos que dado que nos interesa el modelo 2D-ARMA, podemos supo-

ner que cada modelo ARMA tiene un modelo autoregresivo reversible, esto es posible hacerlo al considerar la serie formal de potencias:

$$\frac{Q(z_1, z_2)}{P(z_1, z_2)} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g_{i,j} z_1^{i-1} z_2^{j-1}.$$

Es de observar que los coeficientes  $g_{i,j}$  pueden ser obtenidos recursivamente mediante la igualdad de:

$$g_{i,j} = \theta_{i,j} - \sum_{1 \leq k \leq i, 1 \leq l \leq j, (k,l) \neq (1,1)} g_{i-k+1, j-l+1} \phi_{k,l}.$$

Así, podemos truncar los coeficientes  $g_{i,j}$  a órdenes suficientemente grandes  $o_1, o_2$  para obtener una aproximación del modelo 2D-ARMA mediante un modelo 2D-MA. Entonces podemos tratar un modelo 2D-ARMA( $p_1, p_2, q_1, q_2$ ) como un modelo 2D-MA( $o_1, o_2$ ) de modo asintótico.

## 1.2. Elementos de probabilidad libre y matrices aleatorias

Las matrices de Wishart compuestas fueron introducidas por Speicher [6] en el año 1998 y posteriormente estudiados por Bryc [7] en 2008, estas pueden expresarse mediante una suma ponderada de matrices de Wishart y aparecen de modo natural como matrices de covarianza de muestreos correlacionados [3, 10, 11]. Asimismo, éstas son útiles en el análisis de algunos modelos estadísticos [2, 4] y pueden ser vistas como matrices de Wishart parametrizadas ([7]).

Así, en esta sección tendremos como objetivo dar los elementos para comprender el comportamiento de los cumulantes de trazas de modelos de Wishart compuestos a definir.

**Definición 2.** Definimos una matriz de Wishart compuesta como una familia de matrices aleatorias que satisface:

$$W_{(d/n, D)} = Z^t D Z, (d/n, D) \in \Theta \subset \Theta_{CW} = \cup_{n \in \mathbb{N}} \Theta_{CW}^n$$

donde  $\Theta_{CW}^n = \cup_{d \in \mathbb{N}, d \geq n} \{(d/n, D) \in \mathbb{Q} \times \mathbb{H}_d^+(\mathbb{R})\}$  y  $Z$  una matriz de dimensión  $d \times n$  con entradas independientes idénticamente distribuidas como  $\mathcal{N}(0, \frac{1}{\sqrt{n}})$ .

**Definición 3.** Para cada  $\nu = (\lambda, D) \in \Theta_{CW}$ , definimos su versión amplificada de orden  $k$  como  $\nu^k = (\lambda, D \otimes I_k) \in \Theta_{CW}, \forall k \in \mathbb{N}$ .

Como observaciones, destacamos que la diferencia que existe entre la matriz de Wishart y la matriz de Wishart compuesta, es que esta última generaliza el caso de una matriz hermitiana definida positiva diferente a la identidad, además que el espacio paramétrico carga información respecto a un racional desde la dimensión de la matriz hermitiana definida positiva hasta racionales más pequeños de su dimensión. Por otro lado, la versión amplificada de un parámetro toma bloques de la matriz  $D$  (al tomar tensor con la matriz identidad), fijando el racional asociado a la dimensión de  $D$ . Esta versión amplificada, servirá para definir en el Capítulo 2 una cantidad pivotal asintótica al tomar potencias de orden  $l$  sobre un matriz de Wishart compuesta asociada a una versión amplificada.

Así, para definir tal cantidad pivotal, es necesario fijar algunas definiciones a usar, para ello en nuestro estudio futuro consideramos trabajar sobre un espacio de probabilidad fijo  $(\Omega, \mathcal{F}, \mathbb{P})$ . Procedamos a fijar notación para una variable aleatoria  $X$ , denotamos por  $m_k(X) = \mathbb{E}(X^k)$  el  $k$ -ésimo momento centrado de  $X$ , por  $Tr$  la traza y por  $tr = \frac{Tr}{N}$  la traza normalizada de una matriz  $N \times N$ . Además denotemos a  $m_k(A) = tr(A^k)$  para una matriz cuadrada  $A$ . Procedamos a recordar algunos conceptos básicos sobre particiones y permutaciones tomados de Redelmeier [5].



**Definición 4.**

1. Definimos a  $[n] = \{1, 2, \dots, n\}$  para  $n \in \mathbb{N}$  y para cualquier conjunto  $S$  finito, denotamos a  $|S|$  como su cardinalidad.
2. Una partición  $\pi = \{V_1, \dots, V_k\} \in 2^S$  de un conjunto finito  $S$  es una familia de subconjuntos de  $S$  no vacíos, disjuntos por parejas y cuya unión es  $S$ . Definimos a  $P(S)$  como el conjunto de todas las particiones de  $S$ , y definimos a  $P(n) := P([n])$

Introduzcamos de forma combinatoria los cumulantes clásicos, así como elementos de particiones y permutaciones, los cuales servirán para examinar la traza de polinomios en matrices aleatorias mediante el estudio de Redelmeier [5].

**Definición 5.** Sea  $\mathcal{A}$  el álgebra de las variables aleatorias reales que tienen todos sus momentos finitos. Definamos funciones multi-lineales  $\mathcal{K}_\pi : \mathcal{A}^n \rightarrow \mathbb{R}$  ( $\pi \in P(n), n \in \mathbb{N}$ ) inductivamente por las relaciones cumulantes-momentos:

1. Para  $n \in \mathbb{N}$  y  $X_1, \dots, X_n \in \mathcal{A}$ ,  $\mathbb{E}(X_1 \dots X_n) = \sum_{\pi \in P(n)} \mathcal{K}_\pi(X_1, \dots, X_n)$ .
2. Para  $n \in \mathbb{N}$  y  $\pi \in P(n)$ ,  $\mathcal{K}_\pi(X_1, \dots, X_n) = \prod_{V \in \pi} \mathcal{K}_V(X_1, \dots, X_n)$ .
3. Para  $\pi \in P(n)$  y  $V \in \pi$ ,  $\mathcal{K}_V(X_1, \dots, X_n) = \mathcal{K}_{id_{|V|}}(X_{j_1}, \dots, X_{j_{|V|}})$  donde  $V = \{j_1, \dots, j_{|V|}\}$  con  $j_1 < j_2 < \dots < j_{|V|}$ .

A tales funciones multi-lineales  $\mathcal{K}_\pi$  les llamamos cumulantes clásicos. Con  $\mathcal{K}_n = \mathcal{K}_{id_n}$  y  $\mathcal{K}_n(X) = \mathcal{K}_n(X, \dots, X)$  ( $n$ -veces).

En teoría de matrices aleatorias y probabilidad libre, muchos resultados se basan en el cálculo de la distribución de valores propios asintóticos de matrices aleatorias grandes ([2, 3, 10, 11]). Una de las líneas de investigación hoy en día

es sobre los resultados más profundos sobre teoría de fluctuaciones de matrices aleatorias. El hecho fundamental en la teoría es que la fluctuación  $Tr(W_{\nu_n}^l) - \mathbb{E}(Tr(W_{\nu_n}^l))$  se distribuye asintóticamente normal si la matriz determinista  $D_n$  ( $\nu_n = (d_n/n, D_n) \in \Theta_{CW}^n$ ) tiene los momentos límite  $\lim_{n \rightarrow \infty} tr(D_n^l)$  ( $l \in \mathbb{N}$ ). Uno de los hechos más notables es que la varianza de la fluctuación límite se puede describir como un polinomio de los momentos límite de las partes deterministas (Redelmeier [5]). Hoy en día, este fenómeno se entiende como un resultado de la libertad de segundo orden. Esta contiene más información que la libertad, que es el concepto básico de la teoría de la probabilidad libre, la cual fue desarrollada extensivamente después por Voiculescu (mencionado en [6]) e incluye estrategias para entender el comportamiento asintótico colectivo de conjuntos matriciales aleatorias. De este modo, es necesario definir y desarrollar algunos elementos que nos serán de utilidad en las siguientes secciones.

**Definición 6.**

1. Para cualquier subconjunto  $S$  de números naturales, escribimos  $-S = \{-j \in \mathbb{Z} | j \in S\}$  y  $\pm S = S \cup (-S) \subset \mathbb{Z} - \{0\}$ .
2. Denotamos por  $S(I)$  el grupo de permutación del conjunto finito  $I$ . Escribimos  $S_l = S([l])$ .
3. Para cualquier permutación  $\pi$ , denotamos por  $\#\pi$  el número de ciclos de  $\pi$ . Usamos el mismo símbolo  $\pi$  para la partición generada por sus órbitas.

**Definición 7.** Para cualquier matriz auto adjunta  $A \in M_n(\mathbb{R})$  y permutación  $\pi \in S(\pm[n])$ , con la notación de ciclo  $\pi = \eta_1 \eta_2 \dots \eta_p$ , establecemos  $tr_\pi(A) = \prod_{j=1}^p m_{|\eta_j|}(A)$ .

Usamos la misma notación para el caso de matrices aleatorias. Recordemos algunas nociones de pre-mapeos:

**Definición 8.** Sea  $I \subset \mathbb{Z} - \{0\}$  un conjunto finito de enteros que no contenga  $k, -k$  para cualquier  $k \in \mathbb{N}$ .

1. Se dice que una permutación  $\pi \in S(\pm I)$  es un pre-mapeo si  $\pi(k) = -\pi^{-1}(-k)$  y ningún ciclo contiene tanto  $k$  como  $-k$  para cualquier  $k \in I$ . Denotamos por  $PM(\pm I)$  al conjunto de todos los pre-mapeos en  $S(\pm I)$ . Escribimos  $PM_l = PM(\pm[l])$ .
2. Para cualquier pre-mapeo  $\pi \in PM(\pm I)$ , un ciclo  $(c_1, \dots, c_m)$  de  $\pi$  es particular si  $c_{k_0} > 0$  donde  $k_0 = \operatorname{argmin}_{k=1, \dots, m} |c_k|$ .
3. Denotamos  $[\pi/2]$  al conjunto de todos los elementos que aparecen en ciclos particulares de  $\pi$ . Definamos  $\pi/2 \in S([\pi/2])$  a la permutación generada a partir de ciclos particulares de  $\pi$ .
4. Denotemos a  $\delta \in S(\pm[n])$  la permutación generada por  $\delta(-k) = k$ .
5. Para  $\pi \in S_l$ , definimos  $\pi_+, \pi_- \in PM_l$  por  $\pi_+$  la permutación  $\pi$  considerada como permutación en  $\pm I$  (donde esta actúa trivialmente en todo elemento fuera de  $\pm I$ ) y a  $\pi_- = \delta \pi_+ \delta$ .

Es de notar que por definición la inversa de un pre-mapeo es un pre-mapeo y  $\eta_-^{-1} \pi \eta_+$  es un pre-mapeo para  $\pi \in PM(\pm I), \eta \in S(I)$ .

**Definición 9.** Definimos a la característica de Euler de  $\pi \in PM_l$  con respecto a  $\eta \in S_{2l}$  por:

$$\chi(\eta, \pi) = \#((\eta_+ \eta_-^{-1})/2) + \#(\pi/2) + \#((\eta_-^{-1} \pi \eta_+)/2) - |I|.$$

**Definición 10.** Sea  $\pi, \eta \in S(I)$ , denotamos por  $\pi \vee \eta = id_I$  si para cualquier  $k, l \in I$  distintas, existen  $v_1, \dots, v_{2m} \in I$  y  $V_1, \dots, V_m \in \pi$  y  $W_1, \dots, W_m \in \eta$  con  $v_{2j}, v_{2j+1} \in V_j$ ,  $v_{2j-1}, v_j \in W_j$  para cualquier  $j \leq m$ , y  $(k, l) = (v_1, v_{2m+1})$  ó  $(k, l) = (v_{2m+1}, v_1)$ . Si  $\pi$  satisface esta condición para  $\eta$ , decimos que  $\pi$  conecta todos los bloques de  $\eta$ .

Los siguientes lemas son tomados de Redelmeier [5]. Usaremos tales lemas al explorar el comportamiento asintótico de la expansión del género de matrices compuestas de Wishart.

**Lema 1.** Sea  $\eta \in S_n$  y  $\{V_1, \dots, V_r\} \in P(n)$  las órbitas de  $\eta$ . Supongamos que  $\pi \in PM_n$  conecta la partición  $\eta_{\pm}$ , es decir,  $\pi \vee \eta_{\pm} = id_{\pm[n]}$ . Entonces  $\chi(\eta, \pi) \leq 2$ .

A continuación presentamos una consecuencia de la fórmula de inversión de Moebius dada por Redelmeier [5].

**Lema 2.** Sea  $f, g : \{I \subset [n]\} \rightarrow \mathbb{C}$ , entonces lo siguiente es equivalente:

$$g(I) = \sum_{I \subset J} f(J);$$

$$f(I) = \sum_{I \subset J} (-1)^{J-I} g(J).$$

**Lema 3.** Consideremos una matriz de Wishart compuesta  $\{W_{\lambda} \in M_N(\mathbb{R})\}_{\lambda \in \Lambda}$

con  $W_{\lambda_k} = Z^t D_{\lambda_k} Z$  para  $1 \leq k \leq n$ . Sea  $\eta \in S_n$  y  $Y_1, \dots, Y_n$  matrices aleatorias independientes de  $W_{\lambda_k}$  de dimensiones  $n \times n$ , entonces

$$\mathbb{E}(\text{tr}_\eta(W_{\lambda_1} Y_1, \dots, W_{\lambda_n} Y_n)) = \sum_{\pi \in PM(\pm[n])} N^{\chi(\eta, \pi) - 2\#(\eta)} \text{tr}_{\eta^{-1}}(D_{\lambda_1}, \dots, D_{\lambda_n}) \mathbb{E}(\text{tr}_{(\eta^{-1} \pi \eta_+)/2}(Y_1, \dots, Y_n)).$$

Así, los dos lemas anteriores y la relación cumulantes-momentos implica:

**Lema 4.** Para cualquier  $\nu = (d/n, D) \in \Theta$ , tenemos que

$$\mathcal{K}_r(\text{Tr}(W_\nu^l)) = \sum_{\pi \in PM_{lr, \pi \vee \eta_\pm = id_{\pm[lr]}}} n^{\chi(\eta, \pi) - r} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D)$$

donde  $\eta = (1, \dots, l)(l, l+1, \dots, 2l) \dots ((r-1)l, \dots, rl)$ .

## Capítulo 2

# Equivalentes Deterministas Libres para 2D-ARMA

En este capítulo introducimos los equivalentes deterministas libres (EDL) definidos por Speicher y Vargas [9] y la aplicación a los modelos 2D-ARMA considerados en [1]. Neu y Speicher [18] relacionan por primera vez el uso de operadores abstractos cercanos para el tratamiento de matrices aleatorias grandes, influenciados por los trabajos de Marchenko-Pastur [21] y Voiculescu [20]. En [9] se somete la heurística propuesta por [18, 19] y se señala que tales objetos eran conocidos como un método de aproximación de la transformada de Cauchy de matrices aleatorias en la literatura de comunicación inalámbrica. Más precisamente, mientras que la transformada de Cauchy del equivalente determinista libre se aproxima a la transformada de Cauchy de la matriz aleatoria. Es decir, la aproximación a nivel de operadores hace posible aproximar una amplia clase de matrices aleatorias como polinomios de gaussianas y deterministas.

El equivalente determinista libre también permite aproximar matrices aleatorias al nivel de la fluctuación de los momentos mediante sus cumulantes. Esto se realiza reemplazando cada variable aleatoria gaussiana en las entradas por una ma-

triz aleatoria gaussiana de “tamaño infinito”. Análogamente, se obtiene tal equivalente tomando el límite de los modelos amplificadas mediante la Definición 3 del Capítulo 1, es decir, las matrices se copian tomando el producto tensorial con identidad y las matrices aleatorias gaussianas se amplían simplemente aumentando el número de i.i.d. entradas.

## 2.1. Equivalentes Deterministas Libres (EDL)

El objetivo de esta sección es definir una cantidad pivotal, la cual sea función de  $Tr(W_\nu^l)$ , es decir, una cantidad pivotal en función de la traza de la versión amplificada de un modelo de Wishart compuesto. Para encontrar tal cantidad pivotal se definirá el Z-Score Equivalente Determinista Libre, y para probar que es una cantidad pivotal y con distribución límite normal estándar, se usarán la teoría de fluctuaciones de [5] mediante el Lema 4.

**Definición 11.** El Z-Score Equivalente Determinista Libre de orden  $l \in \mathbb{N}$  de una pareja  $\nu, \nu_0 \in \Theta_{CW}$  y  $\nu_0$  un parámetro conocido como la variable aleatoria dada por:

$$Z_l(\nu_0|\nu) = \frac{Tr(W_{\nu_0}^l) - \mu_l(\nu)}{\sqrt{Var_l(\nu)}}$$

donde  $\mu_l(\nu) = \lim_{k \rightarrow \infty} \mathbb{E}(Tr(W_{\nu^k}^l))/k$ ,  $Var_l(\nu) = \lim_{k \rightarrow \infty} Var(Tr(W_{\nu^k}^l))$ .

Es importante observar que si  $\nu = (\lambda, D) \in \Theta_{CW}$ , entonces los límites  $\mu_l(\nu) = \lim_{k \rightarrow \infty} \mathbb{E}(Tr(W_{\nu^k}^l))/k$  y  $Var_l(\nu) = \lim_{k \rightarrow \infty} Var(Tr(W_{\nu^k}^l))$  siempre existen y pueden ser escritos como polinomios en  $tr(D^k)$ . A partir de este concepto, es posible enunciar el siguiente teorema con implicaciones en modelos 2D-ARMA, tal teorema nos da hipótesis para la convergencia en distribución del Z-Score equivalente determinista libre.

**Teorema 1.** Si  $\nu_n = (d_n/n, D_n) \in \Theta_{CW}^n$  es una sucesión de parámetros y para algún  $l \in \mathbb{N}$  se tiene que:

$$\lim_{n \rightarrow \infty} \frac{\|D_n\|}{\sqrt{\text{tr}(D_n^2)}} = o(n^{1/3l}),$$

entonces  $Z_l(\nu_n|\nu_n)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $n \rightarrow \infty$ .

Por  $\|D_n\|$  denotamos la norma espectral de la matriz  $D_n$ , es decir, el máximo valor singular de una matriz. Intuitivamente, puede pensarse como la máxima “escala”, mediante la cual la matriz puede “extender” un vector.

Primeramente probaremos que el Z-Score Equivalente Determinista Libre esta bien definido, es decir, que  $\mu_l(\nu)$ ,  $\text{Var}_l(\nu)$  existen. Más aún, encontremos un comportamiento de ellos.

**Lema 5.** Sea:

$$\alpha_l = \sum_{\pi \in PM_l, \chi(\gamma_l, \pi)=2} \lambda^{\#(\pi/2)} \text{tr}_{\pi/2}(D),$$

$$\beta_l = \sum_{\pi \in PM_l, \chi(\gamma_l, \pi)=1} \lambda^{\#(\pi/2)} \text{tr}_{\pi/2}(D),$$

entonces  $\mu_l(\nu) = \lim_{k \rightarrow \infty} \mathbb{E}(\text{Tr}(W_{\nu^k}^l))/k = n\alpha_l + \beta_l$  donde  $\gamma_l = (1, 2, \dots, l)$  y  $\lambda = d/n$ . En particular  $\mu_l(\nu)$  esta bien definido, y desarrollando la sumas anteriores  $\mu_1(\nu) = nm_1$  y  $\mu_2(\nu) = n(m_2 + m_1^2) + m_2$ , con  $m_k = m_k(D)$  definido anteriormente para todo  $k \in \mathbb{N}$ .



*Demostración.* Sea  $\gamma_l = (1, \dots, l)$ , entonces usando el Lema 4. tenemos que:

$$\begin{aligned}
\mu_l(\nu) &= \lim_{k \rightarrow \infty} \mathbb{E}(\text{Tr}(W_{\nu, k}^l)) / k \\
&= \lim_{k \rightarrow \infty} k^{-1} \sum_{\pi \in PM_l, \pi \vee \gamma_l = id_{\pm[l]}} n^{\chi(\gamma_l, \pi) - 1} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D \otimes I_k) \\
&= \lim_{k \rightarrow \infty} k^{-1} \sum_{\pi \in PM_l, \pi \vee \gamma_l = id_{\pm[l]}} k n^{\chi(\gamma_l, \pi) - 1} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D) \\
&= \sum_{\pi \in PM_l, \pi \vee \gamma_l = id_{\pm[l]}} n^{\chi(\gamma_l, \pi) - 1} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D),
\end{aligned}$$

Finalmente por el Lema 1  $\chi(\gamma_l, \pi) \leq 2$ , y por tanto separando los sumandos como dos sumandos, tenemos que  $\mu_l(\nu) = n\alpha_l + \beta_l$ .

Por otro lado  $PM_2 = \{id, \pi_1 = (1, 2)(1, 2), \pi_2 = (1, 2)(1, 2)\}$ , así que para  $\gamma_2 = (1, 2)$ , tenemos que  $\chi(\gamma_2, id) = \chi(\gamma_2, \pi_1) = 2, \chi(\gamma_2, \pi_2) = 1$ , de donde  $\mu_2(\nu) = n\alpha_2 + \beta_2 = n(\text{tr}(D^2) + (\text{tr}(D))^2) + \text{tr}(D^2) = n(m_2 + m_1^2) + m_2$ , y para  $\gamma_1 = id$ , tenemos que  $\chi(\gamma_1, id) = 4, \chi(\gamma_2, \pi_1) = \chi(\gamma_2, \pi_2) = 2$ , por tanto  $\mu_1(\nu) = n\alpha_1 = n\text{tr}(D) = nm_1$ , tenemos lo deseado.  $\square$

Observemos que para  $n = 1, 2$ , tenemos que  $\mathbb{E}(\text{Tr}(W_{\nu}^l)) = \mu_1(\nu)$ . Prosigamos a enunciar y probar un lema que nos será de utilidad para encontrar cotas en  $\text{Var}_l(\nu)$ .

**Lema 6.** Se tiene que:

$$\text{Var}_l(\nu) = \lim_{k \rightarrow \infty} \text{Var}(\text{Tr}(W_{\nu, k}^l)) = \sum_{\pi \in PM_{2l}, \pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[2l]}, \chi(\gamma_l, \pi) = 2} \lambda^{\#(\pi/2)} \text{tr}_{\pi/2}(D)$$

donde  $\gamma_l = (1, 2, \dots, l)(l+1, l+2, \dots, 2l)$ ,  $V_1 = \{1, 2, \dots, l\}$ ,  $V_2 = \{l+1, l+2, \dots, 2l\}$  son particiones de  $\gamma_l$ , y  $\lambda = d/n$ . En particular,  $\text{Var}_1(\nu) = 2\lambda m_2$  y  $\text{Var}_2(\nu) = 2(4\lambda^3 m_1^2 m_2 + 2\lambda^2 m_2^2 + 8\lambda^2 m_1 m_3 + 4\lambda m_4)$  con  $m_k = m_k(D)$ .

*Demostración.* Sea  $\gamma_l, V_1, V_2$  y  $\lambda$  como en el enunciado. Notemos que usando el

Lema 3, tenemos que:

$$\begin{aligned}
Var_1(\nu) &= \lim_{k \rightarrow \infty} Var(Tr(W_{\nu^k}^l)) \\
&= \lim_{k \rightarrow \infty} \mathcal{K}_2(Tr(W_{\nu^k}^l)) \\
&= \lim_{k \rightarrow \infty} \sum_{\pi \in PM_{2l}, \pi \vee \gamma_l = id_{\pm[2l]}} n^{\chi(\gamma_l, \pi) - 2} \left(\frac{d}{n}\right)^{\#(\pi/2)} tr_{\pi/2}(D \otimes I_k).
\end{aligned}$$

Dado que por el Lema 1,  $\chi(\gamma_l, \pi) \leq 2$  y ya que  $\gamma_l = (V_1)(V_2)$ , entonces  $\chi(\gamma_l, \pi) \geq 2$ , y así  $\chi(\gamma_l, \pi) = 2$ . Por ende:

$$\begin{aligned}
Var_1(\nu) &= \lim_{k \rightarrow \infty} \sum_{\pi \in PM_{2l}, \pi \vee \gamma_l = id_{\pm[2l]}} n^{\chi(\gamma_l, \pi) - 2} \left(\frac{d}{n}\right)^{\#(\pi/2)} tr_{\pi/2}(D) \\
&= \lim_{k \rightarrow \infty} \sum_{\pi \in PM_{2l}, \pi \vee \gamma_l = id_{\pm[2l]}} n^{\chi(\gamma_l, \pi) - 2} \left(\frac{d}{n}\right)^{\#(\pi/2)} tr_{\pi/2}(D) \\
&= \lim_{k \rightarrow \infty} \sum_{\pi \in PM_{2l}, \pi \vee \gamma_l = id_{\pm[2l]}, \chi(\gamma_l, \pi) = 2} \lambda^{\#(\pi/2)} tr_{\pi/2}(D) \\
&= \sum_{\pi \in PM_{2l}, \pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[2l]}, \chi(\gamma_l, \pi) = 2} \lambda^{\#(\pi/2)} tr_{\pi/2}(D).
\end{aligned}$$

Para el caso  $Var_1(\nu)$ , como  $\pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[2]}$  si y solo si  $\pi/2 = (1, 2), (1, -2)$  y de donde  $Var_1(\nu) = \lambda m_2 + \lambda m_2 = 2\lambda m_2$ .

Para el caso  $Var_2(\nu)$ , como  $\pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[4]}$  si y solo si  $\pi/2$  pertenece al conjunto.

$$\begin{aligned}
&\{(1)(3)(2, 4), (1)(4)(2, 3), (2)(3)(1, 4), (2)(4)(1, 3), (1, 3)(2, 4), (1, 4)(2, 3), \\
&(1)(2, 3, 4), (1)(2, 4, 3), (2)(1, 3, 4), (2)(1, 4, 3), (3)(1, 2, 4), (3)(1, 4, 2), \\
&(4)(1, 2, 3), (4)(1, 3, 2), (1, 2, 3, 4), (1, 2, 4, 3), (1, 4, 3, 2), (1, 3, 4, 2), (1)(3)(2, 4), \\
&(1)(4)(2, 3), (2)(3)(1, 4), (2)(4)(1, 3), (1, 3)(2, 4), (1, 4)(2, 3), (1)(2, 3, 4), \\
&(1)(2, 4, 3), (2)(1, 3, 4), (2)(1, 4, 3), (3)(1, 2, 4), (3)(1, 4, 2), \\
&(4)(1, 2, 3), (4)(1, 3, 2), (1, 2, 3, 4), (1, 2, 4, 3), (1, 4, 3, 2), (1, 3, 4, 2)\}.
\end{aligned}$$

Al contar los pre-mapeos del mismo tipo de ciclo, tenemos que  $Var_2(\nu) = 2(4\lambda^3 m_1^2 m_2 + 2\lambda^2 m_2^2 + 8\lambda^2 m_1 m_3 + 4\lambda m_4)$ . Así, tenemos lo deseado.  $\square$

Efectuemos un cambio de notación de  $\nu = (d/n, D)$  a  $\nu = (d/N, D)$ , esto fue debido a que se le dará interpretación a  $N$  cuando se trabaje en un contexto de modelos 2D-ARMA.

Procedamos a dar una cota inferior para  $Var_l(\nu)$  mediante el siguiente lema:

**Lema 7.** Se tiene que  $Var_l(\nu) \geq (dN^{-1}m_2(D))^l$ .

*Demostración.* Por [5], tenemos que existe  $\pi \in PM_{2l}$  tal que  $\pi/2 = (1, l+1)(2, l+2)\dots(l, 2l)$ . Por tanto,  $\pi$  satisface  $\pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[2l]}$  y  $\chi(\gamma_2, \pi) = 2$ . Además por construcción de  $\pi/2$ , tenemos que  $tr_{\pi/2}(D) = m_2(D)^l$ , de donde al usar el Lema 6, obtenemos:

$$Var_l(\nu) = \sum_{\pi \in PM_{2l}, \pi \vee \{\pm V_1, \pm V_2\} = id_{\pm[2l]}, \chi(\gamma_l, \pi) = 2} \lambda^{\#(\pi/2)} tr_{\pi/2}(D) \geq (d/N)^l m_2(D)^l$$

así,  $Var_l(\nu) \geq (dN^{-1}m_2(D))^l$  y tenemos lo deseado.  $\square$

Sea  $X$  una matriz, definamos  $R(X)$  como el cociente del operador de la norma de  $X$  y su norma de Frobenious normalizada, es decir,  $R(X) = \|X\|/\sqrt{tr(X * X)}$ . El siguiente lema nos permite acotar los cumulantes clásicos de  $Z_l(\nu|\nu)$  con  $\nu = (d/n, D)$ , para posteriormente poder probar el Teorema 1 en la situación límite.

**Lema 8.** Sea  $Z_l(\nu|\nu)$  con  $\nu = (d/n, D)$  definido mediante la Definición 12. Entonces para todo  $r \geq 3$  se tiene:

$$\begin{aligned} |\mathcal{K}_1(Z_l(\nu|\nu))| &\leq R(D)^l \frac{|PM_l|}{n}, \\ |\mathcal{K}_2(Z_l(\nu|\nu)) - 1| &\leq R(D)^{2l} \frac{|PM_{2l}|}{n}, \\ |\mathcal{K}_r(Z_l(\nu|\nu))| &\leq R(D)^{rl} \frac{|PM_{rl}|}{n^{r-2}}. \end{aligned}$$

*Demostración.* Definamos  $\gamma_r = (1, 2, \dots, l)(l+1, l+2, \dots, 2)((r-1)l+1, (r1)l+2, \dots, rl)$ . Por el Lema 4, Lema 5 y Lema 6, al restar las igualdades respectivamente,

tenemos:

$$\mathcal{K}_1(\text{Tr}(W_\nu^l)) - \mu_l(\nu) = \sum_{\pi \in PM_l, \pi \vee \gamma_1 = id_{\pm[l]}, \chi(\gamma_1, \pi) \leq 0} n^{\chi(\gamma_1, \pi) - 1} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D),$$

$$\mathcal{K}_2(\text{Tr}(W_\nu^l)) - \text{Var}_l(\nu) = \sum_{\pi \in PM_{2l}, \pi \vee \gamma_2 = id_{\pm[2l]}, \chi(\gamma_2, \pi) \leq 1} n^{\chi(\gamma_2, \pi) - 2} \left(\frac{d}{n}\right)^{\#(\pi/2)} \text{tr}_{\pi/2}(D).$$

Por otro lado, por el Lema 1  $(d/n)^{\#(\pi/2)} \text{tr}_{\pi/2}[D] \leq (d/n)^l |D|^l$  y por el Lema 7  $\text{Var}_l(\nu) \geq (dN^{-1}m_2(D))^l$ . Por tanto, dividiendo las dos igualdades por las potencias respectivas de  $\text{Var}_l(\nu)$  y acotando los sumandos, tenemos las igualdades deseadas. La última igualdad puedes ser obtenida de modo análogo. □

A partir de ahora ya disponemos de las herramientas desarrolladas para probar el Teorema 1, la demostración puede ser también vista en [1] y la presentamos a continuación.

**Teorema 1.** Si  $\nu_n = (d_n/n, D_n) \in \Theta_{CW}^n$  es una sucesión de parámetros y para algún  $l \in \mathbb{N}$  se tiene que:

$$\lim_{n \rightarrow \infty} \frac{\|D_n\|}{\sqrt{\text{tr}(D_n^2)}} = o(n^{1/3l}),$$

entonces  $Z_l(\nu_n | \nu_n)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $n \rightarrow \infty$ .

*Demostración.* Observemos que por el Lema 8,  $\lim_{n \rightarrow \infty} \mathcal{K}_r(Z_l(\nu_n | \nu_n)) = o(1)$  si  $r \neq 2$  y si  $r = 2$ , tenemos que  $\lim_{n \rightarrow \infty} \mathcal{K}_2(Z_l(\nu_n | \nu_n)) = 1 + o(1)$ . Dado que la única distribución con cumulantes clásicos de orden mayor a dos nulos es la normal, con media el primer cumulante, (en este caso cero) y varianza (en este caso unitaria), concluimos que la distribución asintótica de  $Z_l(\nu_n | \nu_n)$  converge a una distribución normal estándar. □

## 2.2. Aplicación a modelos 2D-ARMA

El objetivo de esta sección es efectuar un puente entre los equivalentes deterministas libres y enfatizarlo en modelos 2D-ARMA introducidos en [8], la conexión es posible debido a la observación efectuada en la Sección 1.1, donde mencionamos que todo modelo 2D-ARMA puede ser llevado de modo asintótico a un modelo 2D-MA, y a través del mismo obtener una matriz de Wishart compuesta, de la cual usamos el Teorema 1 para la obtención de una cantidad pivotal asociada a los modelos 2D-ARMA introducidos en [8]. La interpretación del resultado y su relación con [8], radica en que cuando el tamaño de la imagen de rayos X (producto de las dimensiones) y la muestra tiendan al infinito con la restricción de que la cantidad de muestra menor o igual al tamaño de la imagen, tendremos una cantidad pivotal con convergencia en distribución a la normal estándar. Es decir, tendremos una cantidad pivotal con interpretación en imágenes de rayos X por [8].

De este modo, dado que es posible convertir todo modelo 2D-ARMA a un modelo 2D-MA de modo asintótico, nuestro problema se reduce a considerar modelos 2D-MA. Por otro lado, debido a que los modelos 2D-ARMA provienen de imágenes obtenidas por rayos X, es razonable considerar muestras de rayos X independientes e idénticamente distribuidas, es decir, consideremos  $N$  muestras independientes e idénticamente distribuidas  $y_{h,w}(n)$  con distribución 2D-MA( $q_1, q_2$ ) donde  $(h, w, n) \in \{1, \dots, H\} \times \{1, \dots, W\} \times \{1, \dots, N\}$ . Recordemos que la interpretación de  $H$  y  $W$  representan las dimensiones del modelo 2D-MA (o bien el tamaño de las imágenes de rayos X).

Formalizando lo anterior consideramos:

$$y_{h,w}(n) = \sum_{i=1}^{q_1} \sum_{j=1}^{q_2} \theta_{i,j} \epsilon_{h-i+1, w-j+1}(n)$$

donde  $\theta = (\theta_{i,j}) \in M_{q_1 \times q_2}(\mathbb{R})$  son matrices de parámetros con  $\theta_{1,1} \neq 0$  y  $\{\epsilon_{n,m} | n, m \in \mathbb{Z}\}$  es una familia de v.a.i.i.d. con distribución normal estándar.

Notemos que podemos inmediatamente considerar la matriz  $Y \in M_{HH \times N}(\mathbb{R})$  definida como  $Y_{h(W-1)+w,n} = y_{h,w}(n)/\sqrt{N}$ . A su vez, observemos que del mismo modo podemos definir una matriz de Wishart compuesta del modelo 2D-MA subyacente, para ello usemos la matriz de coeficientes  $\theta$  y  $H, W$  para su construcción, definiendo primeramente a  $H_e = H - 1 + q_1, W_e = W + 1 - q_2$  y procediendo a definir  $B \in M_{HW \times H_e W_e}(\mathbb{R})$  como:

$$B_{(h-1)W+w, (h+i-2)W_e+j+w-1} = \theta_{i,j}$$

donde  $1 \leq h \leq H, 1 \leq w \leq W, 1 \leq i \leq q_1, 1 \leq j \leq q_2$  y las otras entradas cero.

A partir de  $Y$  y  $B$  es conocido que  $Y^t Y$  converge en distribución a la distribución de  $W_{(H_e W_e / N, B^t B)}$ , como variables aleatorias en  $M_N(R)$ . Más aún, podemos probar un resultado muy similar al presentado en el Teorema 1 al cual involucra la traza de  $Y^t Y$ . Por ello, procedemos a las siguientes definiciones, lemas y teorema cuyo objetivo radica en probar tal hecho. La interpretación obtenida para el caso de rayos X se encuentra en que cuando la calidad de la imagen y la muestra tiendan al infinito con la restricción de que la calidad de la misma sea mayor o igual a la muestra, tendremos una cantidad pivotal (el Z-Score equivalente determinista libre) que converge en distribución a una normal estándar.

**Definición 12.** Definimos una matriz de Wishart compuesta para un modelo 2D-MA( $q_1, q_2$ ) como una familia  $(W_\nu)_{\nu \in \Theta_{MA}}$  donde:

$$\Theta_{MA} = \cup_{H_e, W_e, N \in \mathbb{N}, H_e W_e \geq N} \{(H_e W_e / N, B^t B) | q_1, q_2 \in \mathbb{N}, \theta \in M_{q_1 \times q_2}(\mathbb{R}), \theta_{1,1} \neq 0\}.$$

Nos interesa usar el Teorema 1 para probar que  $Z_l$  converge en distribución a una normal estándar en el caso de matrices de Wishart compuestas de un modelo 2D-MA. Por ello nos interesa encontrar cotas asintóticas de  $B^t B$  mediante el uso de normas matriciales y de sucesiones, por tanto, recordemos algunas propiedades.

Consideremos una sucesión  $c = (c_n)_{n \in \mathbb{Z}}$  no cero, excepto en un número finito de índices. Una sucesión  $c = (c_n)_{n=0}^{\infty}$  lo tratamos como una sucesión  $c = (c_n)_{n \in \mathbb{Z}}$  con  $c_n = 0$  para cualquier  $n < 0$ . Definamos la transformada de Fourier de  $c$  como:

$$\hat{c}(\xi) = \sum_{n \in \mathbb{Z}} c_n \exp(-in\xi), \xi \in [0, 2\pi].$$

Sea  $t_c$  un operador en un subespacio denso del espacio de Hilbert de sucesiones con norma acotada  $l^2(\mathbb{Z}) = \{(a_n)_{n \in \mathbb{Z}} \mid \|a_n\|_2 = \sqrt{\sum_{n \in \mathbb{Z}} |a_n|^2} < \infty\}$  definido por  $(t_c)_{i,j} = c_{j-i}$ . Entonces es conocido que  $t_c$  es unitario y equivalente al operador de multiplicación de  $\hat{c}$  en  $L^2([0, 2\pi])$ , por tanto  $\|t_c\| = \|\hat{c}\|_{\infty} = \sup_{\xi \in [0, 2\pi]} |\hat{c}(\xi)| \leq \|c\|_1 = \sum_{n \in \mathbb{Z}} |c_n| < \infty$ . En particular,  $t_c$  es acotado y bien definido en  $l^2(\mathbb{Z})$ .

Por otro lado, definimos la matriz de Toeplitz  $T_c \in M_H(\mathbb{R})$  de la sucesión  $c$  definida como  $(T_c)_{i,j} = c_{j-i}$  para  $i, j = 1, 2, \dots, H$ . Entonces  $T_c$  satisface por construcción que  $T_c \leq t_c \leq \|c\|_1$ .

El siguiente Lema encuentra cotas del cociente  $R(B^t B)$ , el cual recordemos es el cociente de la norma y su norma de Frobenius normalizada. A través de tal cota podemos usar el Teorema 1 y obtener el resultado de asintoticidad deseado para modelos 2D-MA.

**Lema 9.** Sea  $D = B^t B$  y  $\zeta = H_e W_e / HW$ . Entonces se satisface:

$$\begin{aligned} \text{tr}(D^2) &\geq \zeta^{-1} \|\theta\|_2^2, \\ \|D\| &\leq \|\theta\|_1^2, \\ R(D) &\leq \sqrt{\zeta} \|\theta\|_1^2 / \|\theta\|_2. \end{aligned}$$

*Demostración.* Observemos que por definición:

$$\text{Tr}(B^t B B^y B) = \sum_{(i, i_1, i_2, i_3, j, j_1, j_2, j_3) \in S} \theta_{i,j} \theta_{i_1, j_1} \theta_{i_2, j_2} \theta_{i_3, j_3}$$

donde  $S \subset [q_1]^4 \times [q_2]^4$  definido como  $(i, i_1, i_2, i_3, j, j_1, j_2, j_3) \in S$  si existen  $h, h' \in [H]$ ,  $w, w' \in [W]$  tal que:

$$h + i = h + i_3, w + j = w + j_3, h + i_1 = h + i_2, w + j_1 = w + j_2.$$

Sumando solamente los términos cuyos índices corresponden al caso  $h = h'$ ,  $w = w'$ , tenemos que  $\text{tr}(D^2)$  es mayor o igual a tales sumandos que son  $HW \|\theta\|_2^2 = \zeta^{-1} \|\theta\|_2^2$ . Así, tenemos la primera desigualdad.

Definamos la matriz  $V$  de  $HW \times H_e W_e$  por  $V_{(h-1)W+1, (h-1)W_e+w} = 1$  donde  $1 \leq w \leq W$  y  $1 \leq h \leq H$ , y las otras entradas cero.

Definamos la matriz  $T$  de tamaño  $H_e W_e \times H_e W_e$  para  $1 \leq w \leq W_e$  y  $1 \leq h \leq H_e$ ,  $T_{(h-1)W_e+w, (h+i-2)W_e+j+w-1} = \theta_{i,j}$ , donde  $1 \leq i \leq q_1$ ,  $1 \leq j \leq q_2$  con  $(h+i-2)W_e + j + w - 1 \leq H_e W_e$  y las otras entradas nulas. Entonces por construcción  $B = VT$ . Más aún, definamos la sucesión  $c = (c_n)_{0 \leq n \leq H_e W_e}$  por  $c_{(i-1)W_e+j-1} = \theta_{i,j}$  para  $1 \leq i \leq q_1$ ,  $1 \leq j \leq q_2$  y cero de otra modo. Entonces  $T$  es igual a la matriz triangular superior de Toeplitz  $T_c$  por definición. Denotemos  $c_i(n) = \theta_{i,n}$ , y sea  $S$  una matriz  $H_e W_e \times H_e W_e$  nilpotente con  $S_{i, i+1} = 1, 1 \leq i \leq H_e W_e$ , y las otras entradas son cero. Por tanto, hay  $m \in \mathbb{N}$  tal que:

$$T_c = \sum_{i=1}^{q_1} S^{(i-1)m} T_{c_i}.$$



Así, tenemos que  $T_c \leq \sum_{i=1}^{q_1} \|T_{c_i}\|$ , y  $D \leq \|T_c\|^2 \|V^t V\| \leq \|T_c\|^2 \leq (\sum_{i=1}^{q_1} \|T_{c_i}\|)^2$ .

Dado que  $\|T_{c_i}\| \leq \|c_i\|_1$ , entonces  $\|D\| \leq \|\theta\|_1^2$ .

Finalmente, dado que  $R(D) = \|D\|/\sqrt{\text{tr}(D^2)} \leq \frac{\|\theta\|_1^2}{\sqrt{\xi^{-1}\|\theta\|_2^2}} = \sqrt{\xi}\|\theta\|_1^2/\|\theta\|_2$ .

□

Finalizamos la sección y el capítulo enunciando y demostrando el teorema principal de esta sección, en el cual encontramos que el Z-Score equivalente determinista libre de cualquier orden converge en distribución a una normal estándar (tal demostración puede ser también vista en [1]). La interpretación de este hecho en imágenes de rayos X como mencionamos se encuentra en que, cuando la muestra y la calidad de las imágenes obtenidas tiendan al infinito con la muestra creciendo a un ritmo menor o igual al de la calidad de las imágenes, entonces el Z-Score equivalente determinista libre de cualquier orden es una cantidad pivotal.

**Teorema 2.** Para un modelo 2D-MA( $q_1, q_2$ ) fijo con matriz de parámetros  $\theta \in M_{q_1, q_2}(\mathbb{R})$ , sea  $\nu_N = (H_e W_e / N, B^t B)$  construido como anteriormente. Entonces para todo  $l \in \mathbb{N}$  la distribución del Z-Score equivalente determinista libre de orden  $l$ ,  $Z_l(\nu_N | \nu_N)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $N, HW \rightarrow \infty$  con  $HW \geq N$ .

*Demostración.* Por el Lema 9,  $R(D_N) \leq \sqrt{H_e W_e / HW} \|\theta\|_1^2 / \|\theta\|_2$ , entonces:

$$R(D_N) \leq \sqrt{H_e W_e / HW} \|\theta\|_1^2 / \|\theta\|_2 \rightarrow \|\theta\|_1^2 / \|\theta\|_2$$

cuando  $N, HW \rightarrow \infty$  con  $HW \geq N$ .

Finalmente, por el Teorema 1 y ya que  $\theta$  es fijo, tenemos que para todo  $l \in \mathbb{N}$ ,  $Z_l(\nu_N | \nu_N)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $N, HW \rightarrow \infty$  con  $HW \geq N$ .

□

## Capítulo 3

# EDL con errores función de Poisson para 2D-ARMA

El objetivo de este capítulo es presentar una propuesta de modelaje diferente a los modelos 2D-ARMA introducidos por Zielinski et al. [8] en 2010. Nuestra propuesta surge de la misma naturaleza dada por el problema inicial del modelado en imágenes de radiografías y mamografías, pero a diferencia de los modelos 2D-ARMA considerados por [8], nosotros consideramos la distribución de los errores como una centralización de la transformación de estabilización de varianza de la distribución Poisson. En la Sección 3.1 abordaremos el motivo que origina nuestro cambio de modelaje en los errores y el nuevo modelo propuesto. La Sección 3.2 contiene las modificaciones a tomar en los equivalentes deterministas libres, así como la prueba límite de considerar errores la transformada de Anscombe centrada. La Sección 3.3 efectúa las adecuaciones para el caso 2D-ARMA, y la aplicación de los equivalentes deterministas libres bajo nuestra modificación. Finalmente, la Sección 3.4 expone un estudio de simulación con la finalidad de analizar la velocidad de convergencia de las cantidades pivotaes halladas con nuestra propuesta de modelaje y compararlas con las del Capítulo 2.

### **3.1. 2D-ARMA con errores función de Poisson**

Los modelos 2D-ARMA en [8] fueron introducidos para abordar la problemática natural de trabajar en un contexto de imágenes de rayos X, llegando a ser objeto de estudio en una amplia variedad de campos. Uno de los primeros usos de los rayos X fue en el campo médico para obtener información sobre el interior del cuerpo sin disección. Además de la técnica básica de imágenes de rayos X de cortes fijos, se han desarrollado varios sistemas de imágenes, como lo son los ultrasonidos y las imágenes por resonancia magnética. Un haz de rayos X es parcialmente absorbido por estructuras internas del cuerpo en un proceso conocido como atenuación, y el detector en el otro lado del cuerpo absorbe estos rayos X atenuados para producir una imagen de rayos X. Muchos de los algoritmos existentes en la literatura clásica de procesamiento de imágenes y señales considera que el ruido de una imagen de rayos X se distribuye Poisson (Cesarelli M. et al. [14]) con media representando la desviación media de fotones a través de la imagen obtenida por rayos X (con media generalmente mayores a 10). Sin embargo, tal desviación media de fotones varía de imagen a imagen en términos reales y, por ende, la varianza del ruido de las imágenes no es necesariamente igual. Debido a que el parámetro de la distribución Poisson depende de la imagen, se han introducido transformaciones de estabilización de varianza para generar tal independencia y así, tener una distribución de error independiente de la imagen. La transformada de Anscombe dirigida al ruido de Poisson fue propuesta por Anscombe [16] en 1948, y se extendió a la transformada de Anscombe generalizada. Utilizando esta transformación de estabilización de varianza y posteriormente efectuar una centralización, el ruido función Poisson se convierte en un ruido con distribución límite normal estándar.

En términos formales, supongamos que tenemos  $N$  imágenes tomadas por

rayos  $X, Im_1, \dots, Im_N$  cuyas variables aleatorias reales de los errores originados por la desviación media de fotones a través de la  $i$ -ésima imagen están dados por  $\{\epsilon_{n,m}^i | n, m \in \mathbb{Z}\}$ . Es razonable suponer que la familia de los errores es independiente, es decir,  $\{\epsilon_{n,m}^i | i = \{1, \dots, N\}, n, m \in \mathbb{Z}\}$  es una familia de variables aleatorias reales independientes. Cesarelli M. et al. [14] postulan que la distribución de los errores es Poisson, es decir, existen  $\lambda_1, \dots, \lambda_N$  positivas, no necesariamente iguales, tales que  $\epsilon_{n,m}^i \sim Poisson(\lambda_i)$ .

Dado que  $Var(Poisson(\lambda_i)) = \lambda_i$  y las varianzas no son necesariamente iguales, nos vemos en la necesidad de introducir una transformación de estabilización de varianza para los errores, debido a que es poco realista suponer que la desviación media de fotones siempre es la misma y no varía de imagen a imagen. De este modo, la distribución de los errores será independiente de la imagen específica, por lo tanto, tendremos una distribución de errores que podemos usar para toda imagen obtenida por rayos X.

Observemos que tal transformación debe ser de una distribución Poisson, así consideramos la transformada de Anscombe dirigida al ruido de Poisson, es decir, consideramos la distribución de los errores como la transformada de Anscombe de variables aleatorias independientes Poisson. Dado que nuestro interés es en la distribución de errores, deseamos que su media sea nula, y por ello tomamos como tal distribución la centralización de la transformada de Anscombe, y por lo tanto, consideramos que  $\epsilon_{n,m}^i \sim 2\sqrt{Poisson(\lambda) + \frac{3}{8}} - \mathbb{E}(2\sqrt{Poisson(\lambda) + \frac{3}{8}})$  para algún  $\lambda$  positivo, el parámetro  $\lambda$  puede ser interpretado en función de la varianza de la distribución de los errores subyacentes, es decir, la variabilidad de cada imagen pixel a pixel. Es de notar que al tomar la centralización de la transformada de Anscombe, forzamos a que la varianza sea aproximadamente unitaria, si estamos interesados en varianzas distintas a uno, consideramos múltiples de la centralización de la transformada de Anscombe.

De esta forma, estaremos interesados en suponer errores tratados mediante la centralización de la transformada de Anscombe o múltiplos de ella, con intención de la obtención de cantidades pivotaes asintóticas, para posteriormente poder efectuar estudios de simulación para comprobar orden de convergencia de modo empírico.

Definamos la transformación de Anscombe a trabajar en el resto del documento, la cual definimos por medio de una variable aleatoria  $X \sim Poisson(\lambda)$  como la variable aleatoria  $Y = 2\sqrt{X + \frac{3}{8}}$ . Una observación importante de tal transformada es que ella es la "mejor" transformación de la clase de funciones de la forma  $\sqrt{X + c}$  que estabiliza la varianza, y en la práctica la transformada de Anscombe estabiliza la varianza para  $\lambda \geq 5$  por las siguientes propiedades de [16].

**Lema 10.** La transformada de Anscombe de  $Y = 2\sqrt{X + \frac{3}{8}}$  de una variable aleatoria  $X \sim Poisson(\lambda)$  tiene las siguientes propiedades:

1.  $\mathbb{E}(Y) \sim 2\sqrt{\lambda + \frac{3}{8}} - \frac{1}{4\lambda^{1/2}} + \frac{1}{32\lambda^{3/5}}$ .
2.  $Var(Y) \sim 1 + \frac{1}{16\lambda^2}$ .

A través de estas dos propiedades, observemos que la transformada de Anscombe cuando  $\lambda \geq 5$  se distribuye aproximadamente como una variable aleatoria normal por el siguiente lema.

**Lema 11.** Sea  $X$  una variable aleatoria tal que  $X \sim Poisson(\lambda)$ , entonces  $2\sqrt{X} - 2\sqrt{\lambda} \xrightarrow{d} \mathcal{N}(0, 1)$  cuando  $\lambda \rightarrow \infty$ .

*Demostración.* Consideremos  $n$  variables aleatorias i.i.d.  $\{X_i\}_{i=1, \dots, n}$  tales que  $X_i \sim Poisson(\frac{\lambda}{n}), \forall i = 1, \dots, n$ . Sea  $g$  la función definida sobre  $[0, \infty)$  tal que:  $g(x) = 2\sqrt{x}$ . Notemos que  $g$  es diferenciable en  $(0, \infty)$  con derivada  $g'(x) =$

$$\frac{1}{\sqrt{x}} > 0.$$

Por otro lado, por el teorema central del límite, tenemos que:

$$\frac{\sum_{i=1}^n X_i}{n} \sim \mathcal{AN}\left(\frac{\lambda}{n}, \frac{1}{n} \frac{\lambda}{n}\right).$$

Como  $g$  es diferenciable en  $\lambda > 0$ , las hipótesis del método delta se satisfacen de donde:

$$\begin{aligned} g\left(\frac{\sum_{i=1}^n X_i}{n}\right) &\sim \mathcal{AN}\left(g\left(\frac{\lambda}{n}\right), \frac{\lambda}{n^2} [g'(\lambda/n)]^2\right), \\ 2\sqrt{\frac{\sum_{i=1}^n X_i}{n}} &\sim \mathcal{AN}\left(2\sqrt{\frac{\lambda}{n}}, \frac{1}{n}\right), \\ \sqrt{n}\left(2\sqrt{\frac{\sum_{i=1}^n X_i}{n}} - 2\sqrt{\frac{\lambda}{n}}\right) &\xrightarrow{d} \mathcal{N}(0, 1). \end{aligned}$$

Finalmente, dado que  $\sum_{i=1}^n X_i \sim Poisson(\lambda)$  y que la suma de variables aleatorias independientes Poisson es Poisson del parámetro de la suma de los parámetros, entonces  $\sum_{i=1}^n X_i \sim X$ .

Así, obtenemos que:

$$2\sqrt{X} - 2\sqrt{\lambda} \xrightarrow{d} \mathcal{N}(0, 1)$$

cuando  $\lambda \rightarrow \infty$ . □

Notemos que  $2\sqrt{X + \frac{3}{8}} \sim 2\sqrt{X}$  y por ende,  $2\sqrt{X + \frac{3}{8}}$  se distribuye aproximadamente normal. Es de comentar, que la prueba del lema anterior difiere de la dada en [16] mediante el uso del método delta y del teorema central del límite y no de aproximaciones en series de modo explícito.

Así, nuestras modificaciones de interés son considerar modelos 2D-ARMA explicados en Definición 1, y con distribución de los errores  $\{\epsilon_{n,m} | n, m \in \mathbb{Z}\}$  una familia de v.a.i.i.d. con distribución  $\sigma\left(2\sqrt{Poisson(\lambda) + \frac{3}{8}} - \mathbb{E}\left(2\sqrt{Poisson(\lambda) + \frac{3}{8}}\right)\right)$ , es decir, múltiplos de la centralización de la transformada de Anscombe (a tales modelos les llamaremos 2D-ARMA de Anscombe, pese a que nosotros los proponemos).

En la siguiente sección efectuaremos modificaciones a las matrices de Wishart compuestas considerando tal distribución de errores, probamos el caso asintótico de nuestra modificación al Z-Score equivalente determinista libre y efectuamos un estudio de simulación para observar velocidad de convergencia.

### 3.2. EDL con errores función de Poisson

En esta sección efectuamos adaptaciones a la teoría de equivalentes deterministas libres para nuestro caso de errores dados por la transformada de Anscombe centrada, probamos el caso asintótico del Z-Score equivalente determinista libre adaptado a nuestra distribución de errores y efectuamos en las secciones subsecuentes un estudio de simulación para observar empíricamente la velocidad de convergencia. Este estudio de simulación tiene una fuerte motivación, debido a que el Z-Score equivalente determinista libre considera matrices con entradas v.a.i.i.d. normales centradas de varianza  $1/\sqrt{n}$ , y nosotros estamos interesados en modificaciones de tal matriz entrada por entrada, así como su velocidad de convergencia al caso del Z-Score equivalente determinista libre con errores normales. Debido a que esta última modela la distribución de los errores en los modelos 2D-ARMA,

La modificación a efectuar en las matrices de Wishart compuestas es la de considerar matrices de la forma  $P_{(d/n,D)} = P^t D P$  con la matriz  $P$  teniendo entradas obtenidas por medio en la centralización de la transformación de Anscombe de variables aleatorias independientes idénticamente distribuidas como Poisson de un mismo parámetro.

**Definición 13.** Definimos una matriz de Anscombe compuesta de parámetro

$\lambda > 0$  como una familia de matrices aleatorias que satisface

$$P_{\lambda,(d/n,D)} = P_{\lambda}^t D P_{\lambda}, (d/n, D) \in \Theta \subset \Theta_{CW} = \cup_{n \in \mathbb{N}} \Theta_{CW}^n$$

donde  $\Theta_{CW}^n = \cup_{d \in \mathbb{N}, d \geq n} \{(d/n, D) \in \mathbb{Q} \times \mathbb{H}_d^+(\mathbb{R})\}$  y  $P_{\lambda}$  una matriz de dimensión  $d \times n$  con entradas independientes idénticamente distribuidas como  $\frac{2}{n^{1/4}}(\sqrt{Poisson(\lambda) + \frac{3}{8}} - \mathbb{E}(\sqrt{Poisson(\lambda) + \frac{3}{8}}))$ .

Notemos que no es necesario efectuar modificaciones a las definiciones de  $\Theta_{CW}$ , así como la versión amplificada de un parámetro  $\nu \in \Theta_{CW}$ , pues son independientes de la matriz  $Z$ . Definamos ahora nuestra adaptación del Z-Score equivalente determinista libre.

**Definición 14.** El Z-Score Equivalente Determinista Libre de Anscombe de parámetro  $\lambda > 0$  y orden  $l \in \mathbb{N}$  de una pareja  $\nu, \nu_0 \in \Theta_{CW}$  y  $\nu_0$  un parámetro conocido como la variable aleatoria dada por:

$$Z_{\lambda,l}^A(\nu_0|\nu) = \frac{Tr(P_{\lambda,\nu_0}^l) - \mu_{\lambda,l}(\nu)}{\sqrt{Var_{\lambda,l}(\nu)}}$$

donde  $\mu_{\lambda,l}(\nu) = \lim_{k \rightarrow \infty} \mathbb{E}(Tr(P_{\lambda,\nu^k}^l))/k$ ,  $Var_{\lambda,l}(\nu) = \lim_{k \rightarrow \infty} Var(Tr(P_{\lambda,\nu^k}^l))$ .

Es de observar que el Z-Score Equivalente Determinista Libre de Anscombe nosotros lo proponemos y estudiamos por primera vez (lo llamamos de Anscombe debido al uso de tal transformación).

La importancia del teorema siguiente y subsecuentes lemas tienen como finalidad efectuar una conexión entre nuestro modelo propuesto y los resultados de [1]. Resaltamos que en nuestro caso, hemos introducido el parámetro  $\lambda$ , el cual surge del origen de los modelos 2D-ARMA y cuya interpretación es asociada a la variabilidad de cada imagen obtenida por rayos-X pixel a pixel, de tal modo, que



si nuestro parámetro es grande nuestros errores función Poisson convergerán en distribución a una normal estándar, sin embargo, si tal parámetro es pequeño no tendremos una convergencia rápida como la mencionada.

**Teorema 3.** Si  $\nu_n = (d_n/n, D_n) \in \Theta_{CW}^n$  es una sucesión de parámetros y para algún  $l \in \mathbb{N}$  se tiene que:

$$\lim_{n \rightarrow \infty} \frac{\|D_n\|}{\sqrt{\text{tr}(D_n^2)}} = o(n^{1/3l}),$$

entonces  $Z_{\lambda,l}^A(\nu_n|\nu_n)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $\lambda, n \rightarrow \infty$ .

Para probar este teorema, requerimos los siguientes lemas.

**Lema 12.** Sea  $\nu = (d/n, D) \in \Theta_{CW}$ , entonces  $\text{Tr}(P_{\lambda,\nu}^l)$  converge en distribución a  $\text{Tr}(W_\nu^l)$  cuando  $\lambda \rightarrow \infty$ .

*Demostración.* Consideremos la disposición en vector de la matriz  $P_\lambda$ ,

$$((P_\lambda)_{1,1}, \dots, (P_\lambda)_{1,n}, \dots, ((P_\lambda)_{d,1}, \dots, (P_\lambda)_{d,n}),$$

observemos que la familia de variables aleatorias  $\{(P_\lambda)_{i,j}\}$  es independiente por construcción, entonces:

$$((P_\lambda)_{1,1}, \dots, (P_\lambda)_{1,n}, \dots, ((P_\lambda)_{d,1}, \dots, (P_\lambda)_{d,n}),$$

converge en distribución a:

$$(W_{1,1}, \dots, W_{1,n}, \dots, W_{d,1}, \dots, W_{d,n}),$$

cuando  $\lambda \rightarrow \infty$ .

Consideremos la función  $g : \mathbb{R}^{dn} \rightarrow \mathbb{R}$ , definida por:

$$g(x_{1,1}, \dots, x_{1,n}, \dots, x_{d,1}, \dots, x_{d,n}) = \sum_{i,j_3,\dots,j_{3l-3}=1}^n \left( \sum_{j_1,\dots,j_{3l-1}=1}^d \omega_{i,j_1,\dots,j_{3l-1}} \right),$$

donde

$$\omega_{i,j_1,\dots,j_{3l-1}} = x_{j_1,i} D_{j_1,j_2} x_{j_2,j_3} \dots x_{j_{3l-2},j_{3l-3}} D_{j_{3l-2},j_{3l-1}} x_{j_{3l-1},i}.$$

Por construcción  $g$  es una función continua, pues es suma de funciones continuas.

Por ende, por el teorema de continuidad, tenemos que:

$$g((P_\lambda)_{1,1}, \dots, (P_\lambda)_{1,n}, \dots, ((P_\lambda)_{d,1}, \dots, (P_\lambda)_{d,n}),$$

converge en distribución a:

$$g(W_{1,1}, \dots, W_{1,n}, \dots, W_{d,1}, \dots, W_{d,n}),$$

cuando  $\lambda \rightarrow \infty$ .

Finalmente, observemos que:

$$Tr(P_{\lambda,\nu}^l) = \sum_{i,j_3,\dots,j_{3l-3}=1}^n \left( \sum_{j_1,\dots,j_{3l-1}=1}^d \omega_{i,j_1,\dots,j_{3l-1}}^* \right),$$

donde

$$\omega_{i,j_1,\dots,j_{3l-1}}^* = (P_\lambda^t)_{i,j_1} D_{j_1,j_2} (P_\lambda)_{j_2,j_3} \dots (P_\lambda^t)_{j_{3l-3},j_{3l-2}} D_{j_{3l-2},j_{3l-1}} (P_\lambda)_{j_{3l-1},i}$$

así,

$$Tr(P_{\lambda,\nu}^l) = g((P_\lambda)_{1,1}, \dots, (P_\lambda)_{1,n}, \dots, ((P_\lambda)_{d,1}, \dots, (P_\lambda)_{d,n}),$$

converge en distribución a:

$$Tr(W_\nu^l) = g(W_{1,1}, \dots, W_{1,n}, \dots, W_{d,1}, \dots, W_{d,n}),$$

cuando  $\lambda \rightarrow \infty$ .

□

Es inmediato de este lema que tenemos que:

$$\lim_{\lambda \rightarrow \infty} \mu_{\lambda,l}(\nu) = \lim_{k,\lambda \rightarrow \infty} \mathbb{E}(Tr(P_{\lambda,\nu^k}^l))/k = \mu_l(\nu),$$

$$\lim_{\lambda \rightarrow \infty} \text{Var}_{\lambda,l}(\nu) = \lim_{\lambda,k \rightarrow \infty} \text{Var}(\text{Tr}(P_{\lambda,\nu^k}^l)) = \text{Var}_l(\nu).$$

Por tanto, cuando  $\lambda \rightarrow \infty$ , podemos calcular  $\mu_{\lambda,l}(\nu)$  y  $\text{Var}_{\lambda,l}(\nu)$  como  $\mu_l(\nu)$  y  $\text{Var}_l(\nu)$  respectivamente, y estos últimos fueron calculados en la Sección 2.2.

**Teorema 3.** Si  $\nu_n = (d_n/n, D_n) \in \Theta_{CW}^n$  es una sucesión de parámetros y para algún  $l \in \mathbb{N}$  se tiene que:

$$\lim_{n \rightarrow \infty} \frac{\|D_n\|}{\sqrt{\text{tr}(D_n^2)}} = o(n^{1/3l}),$$

entonces  $Z_{\lambda,l}^A(\nu_n|\nu_n)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $\lambda, n \rightarrow \infty$ .

*Demostración.* Notemos que es suficiente probar que  $\lim_{n,\lambda \rightarrow \infty} \mathcal{K}_r(Z_{\lambda,l}^A(\nu_n|\nu_n)) = o(1)$  si  $r \neq 2$  y  $\lim_{n,\lambda \rightarrow \infty} \mathcal{K}_2(Z_{\lambda,l}^A(\nu_n|\nu_n)) = 1 + o(1)$ . Ya que la única distribución con cumulantes clásicos de orden mayor a dos igual a cero es la normal, con media el primer cumulante, en este caso cero y varianza en este caso unitaria, y concluiríamos que la distribución asintótica de  $Z_{\lambda,l}^A(\nu_n|\nu_n)$  converge a una distribución normal estándar cuando  $\lambda, n \rightarrow \infty$ .

Por el Lema 12, tenemos que para todo  $\nu_n \in \Theta_{CW}$ ,  $\text{Tr}(P_{\lambda,\nu_n}^l)$  converge en distribución a  $\text{Tr}(W_{\nu_n}^l)$  cuando  $\lambda \rightarrow \infty$ , y de donde tenemos convergencia en momentos (pues todos los momentos y cumulantes de  $\text{Tr}(W_{\nu_n}^l)$  son finitos), cumulantes y función generadora de momentos.

Por tanto:

$$\lim_{\lambda \rightarrow \infty} \mathcal{K}_r(\text{Tr}(P_{\lambda,\nu_n}^l)) = \mathcal{K}_r(\text{Tr}(W_{\nu_n}^l)),$$

de donde:

$$\lim_{\lambda,n \rightarrow \infty} \mathcal{K}_r(\text{Tr}(P_{\lambda,\nu_n}^l)) = \lim_{n \rightarrow \infty} \mathcal{K}_r(\text{Tr}(W_{\nu_n}^l)).$$

Finalmente, como:

$$\lim_{\lambda,n \rightarrow \infty} \mathcal{K}_r(\text{Tr}(P_{\lambda,\nu_n}^l)) = \lim_{n \rightarrow \infty} \mathcal{K}_r(\text{Tr}(W_{\nu_n}^l)),$$

entonces:

$$\lim_{n, \lambda \rightarrow \infty} \mathcal{K}_r(Z_{\lambda, l}^A(\nu_n | \nu_n)) = \lim_{n \rightarrow \infty} \mathcal{K}_r(Z_l(\nu_n | \nu_n)),$$

y así, tenemos lo deseado.  $\square$

La Subsección 3.4.1. es dedicada a un estudio de simulación para analizar la velocidad de convergencia al introducir el parámetro  $\lambda$  de nuestra cantidad pivotal  $Z_{\lambda, l}^A(\nu_n | \nu_n)$ , así como compararla con la cantidad pivotal de la Sección 2.1.

### 3.3. Aplicación a 2D-ARMA con errores función de Poisson

En esta sección se efectúa la relación entre modelos 2D-ARMA [8] y los equivalentes deterministas libres con errores función de Poisson y varianza no homogénea de la Sección 3.2, con el fin de obtener un resultado asintótico para modelos 2D-ARMA similar al Teorema 2 de la Sección 2.2, pero a diferencia de la Sección 2.2 consideramos errores función de Poisson y varianza no homogénea para los modelos 2D-ARMA. En la Sección 3.2 introducimos un parámetro  $\lambda$ , derivado del origen de los modelos 2D-ARMA descritos en [8], cuyo objetivo inicial es introducir modelos para el problema de clasificación y predicción de imágenes obtenidas de rayos X de cáncer de seno con tasas de correcta clasificación cercanas al 80 % y, posteriormente, mejorada tal tasa al orden del 98 % mediante calibración de los estimadores de mínimos cuadrados mediante análisis secuencial.

En nuestro caso, nosotros consideramos la introducción del parámetro  $\lambda$  (como mencionamos en la Sección 3.1) con la finalidad del modelado de los errores, mediante la transformación de distribuciones Poisson postuladas en Elbakri et al.

[15] y Sangyoon et al. [17].

Finalmente, en la siguiente sección efectuaremos un estudio de simulación para analizar el orden de convergencia de un resultado de naturaleza semejante al del Teorema 2 de la Sección 2.2 donde modificaremos el modelo 2D-ARMA a trabajar, mediante el cambio distribucional de los errores. Para ello definamos primeramente nuestro modelo 2D-ARMA a trabajar con nuestro cambio en la distribución de errores postulado en la Sección 3.1.

**Definición 14.** Un modelo 2D-ARMA( $p_1, p_2, q_1, q_2$ ) de Anscombe de parámetro  $\lambda > 0$  ( $p_1, p_2, q_1, q_2 \in \mathbb{N}$ ) definido para un latís de dimensiones  $H \times W \in \mathbb{N} \times \mathbb{N}$  es una familia de variables aleatorias reales  $\{y_{n,m} | (n, m) \in \{1, \dots, H\} \times \{1, \dots, W\}\}$  que satisface para toda pareja  $(n, m) \in \{1, \dots, H\} \times \{1, \dots, W\}$ :

$$\sum_{i=1}^{p_1} \sum_{j=1}^{p_2} \phi_{i,j} y_{n-i+1, m-j+1} = \sum_{i=1}^{q_1} \sum_{j=1}^{q_2} \theta_{i,j} \epsilon_{n-i+1, m-j+1}$$

donde  $\phi = (\phi_{i,j}) \in M_{p_1 \times p_2}(\mathbb{R})$ ,  $\theta = (\theta_{i,j}) \in M_{q_1 \times q_2}(\mathbb{R})$  son matrices de parámetros con  $\phi_{1,1} = 1, \theta_{1,1} \neq 0$  y  $\{\epsilon_{n,m} | n, m \in \mathbb{Z}\}$  es una familia de v.a.i.i.d. con distribución  $2(\sqrt{Poisson(\lambda) + \frac{3}{8}} - \mathbb{E}(\sqrt{Poisson(\lambda) + \frac{3}{8}}))$ .

Es de observar que los modelos 2D-ARMA de Anscombe nosotros los proponemos y estudiamos por primera vez (lo llamamos de Anscombe debido al uso de tal transformación).

Similarmente como en la Sección 1.1, es posible hallar estimadores de mínimos cuadrados de los parámetros del modelo anterior. Estos son obtenidos debido, a que la ecuación anterior puede ser escrita de modo matricial semejante a la Sección 1.1. De igual modo que en el Capítulo 1, es posible definir cuando un modelo 2D-ARMA de Anscombe es reversible y tendremos la observación asintóticamente, es decir, cuando  $\lambda \rightarrow \infty$ , tendremos que es posible convertir todo

modelo 2D-ARMA de Anscombe a un modelo 2D-MA de Anscombe de modo nuevamente asintótico. De esta forma, como en la Sección 2.2 nuestro problema se reduce a considerar modelos 2D-MA de Anscombe.

Por otro lado, debido a que los modelos 2D-ARMA de Anscombe provienen de imágenes obtenidas por rayos X, es razonable considerar muestras de rayos X independientes e idénticamente distribuidas, es decir, consideremos  $N$  muestras independientes e idénticamente distribuidas  $y_{h,w}(n)$  con distribución 2D-MA( $q_1, q_2$ ) de Anscombe de parámetro  $\lambda$  donde  $(h, w, n) \in \{1, \dots, H\} \times \{1, \dots, W\} \times \{1, \dots, N\}$ , y recordemos que la interpretación de  $H$  y  $W$  representan las dimensiones del modelo 2D-MA de Anscombe (o bien el tamaño de las imágenes de rayos X).

Formalizando lo anterior consideramos:

$$y_{h,w}(n) = \sum_{i=1}^{q_1} \sum_{j=1}^{q_2} \theta_{i,j} \epsilon_{h-i+1, w-j+1}(n)$$

donde  $\theta = (\theta_{i,j}) \in M_{q_1 \times q_2}(\mathbb{R})$  son matrices de parámetros con  $\theta_{1,1} \neq 0$  y  $\{\epsilon_{n,m} | n, m \in \mathbb{Z}\}$  es una familia de v.a.i.i.d. con distribución  $2(\sqrt{Poisson(\lambda) + \frac{3}{8}} - \mathbb{E}(\sqrt{Poisson(\lambda) + \frac{3}{8}}))$ .

Notemos que podemos considerar similarmente la matriz  $Y_\lambda \in M_{HH \times N}(\mathbb{R})$  definida como  $Y_{h(W-1)+w, n} = y_{h,w}(n) / \sqrt{N}$ . A su vez, observemos que del mismo modo podemos definir una matriz de Wishart compuesta del modelo 2D-MA de Anscombe subyacente, para ello usemos la matriz de coeficientes  $\theta$  y  $H, W$  para su construcción, definiendo primeramente a  $H_e = H - 1 + q_1, W_e = W + 1 - q_2$  y procediendo a definir  $B \in M_{HW \times H_e W_e}(\mathbb{R})$  como:

$$B_{(h-1)W+w, (h+i-2)W_e+j+w-1} = \theta_{i,j}$$

donde  $1 \leq h \leq H, 1 \leq w \leq W, 1 \leq i \leq q_1, 1 \leq j \leq q_2$  y las otras entradas cero.

Es de observar que tal matriz construida es independiente de  $\lambda$  como en la sección anterior, así podemos efectuar la modificación sobre  $Z_{\lambda,l}^A(\nu_n|\nu_n)$  y enunciar el siguiente teorema cuya interpretación es similar a la enunciada en el Teorema 2, es decir, en el cual encontramos que el Z-Score equivalente determinista libre de Anscombe de cualquier orden converge (con la adición  $\lambda \rightarrow \infty$ ) en distribución a una normal estándar. La interpretación de este hecho en imágenes de rayos X se encuentra en que cuando la postulación de nuestros errores sigue una distribución que corresponde más razonablemente por [15, 17] a la motivación del problema original de [8]. Así, la muestra y la calidad de las imágenes obtenidas tiendan al infinito con la muestra creciendo a un ritmo menor o igual al de la calidad de las imágenes, adicionado con que la tasa de nuestro modelo 2D-ARMA de Anscombe tienda a infinito, entonces el Z-Score equivalente determinista libre de Anscombe de cualquier orden es una cantidad pivotal.

**Teorema 4.** Para un modelo 2D-MA( $q_1, q_2$ ) de Anscombe de parámetro  $\lambda > 0$  fijo con matriz de parámetros  $\theta \in M_{q_1, q_2}(\mathbb{R})$ , sea  $\nu_N = (H_e W_e / N, B^t B)$  construido como anteriormente. Entonces para todo  $l \in \mathbb{N}$  la distribución del Z-Score equivalente determinista libre de Anscombe con parámetro  $\lambda$  y de orden  $l$ ,  $Z_{\lambda,l}^A(\nu_N|\nu_N)$  converge distribución a  $\mathcal{N}(0, 1)$  cuando  $N, HW, \lambda \rightarrow \infty$  con  $HW \geq N$ .

*Demostración.* Por la demostración del Teorema 3, tenemos que  $Z_{\lambda,l}^A(\nu_N|\nu_N)$  converge en distribución a  $Z_l(\nu_N|\nu_N)$  cuando  $\lambda \rightarrow \infty$ , finalmente dado que  $Z_l(\nu_N|\nu_N)$  converge en distribución a  $\mathcal{N}(0, 1)$  cuando  $\lambda \rightarrow \infty$ , tenemos que por la transitividad de la convergencia distribucional,  $Z_{\lambda,l}^A(\nu_N|\nu_N)$  converge distribución a  $\mathcal{N}(0, 1)$  cuando  $N, HW, \lambda \rightarrow \infty$  con  $HW \geq N$ .

Así tenemos lo deseado.  $\square$

La Subsección 3.4.2. es dedicada a un estudio de simulación para analizar la velocidad de convergencia al introducir el parámetro  $\lambda$  de la cantidad pivotal  $Z_{\lambda,l}^A(\nu_N|\nu_N)$ , obtenida por medio de modelos 2D-ARMA de Anscombe de parámetro  $\lambda$ . Así mismo, compararla con la cantidad pivotal de la Sección 2.2.

### 3.4. Estudio de simulación

Esta sección es dedicada a efectuar estudios de simulación para el análisis de la velocidad de convergencia con la introducción del parámetro  $\lambda$ , del cual derivan las cantidades pivotaes halladas en la Secciones 3.2 y 3.3. A su vez, otro de nuestros intereses radica en comparar tales pivotaes a los obtenidas en las Secciones 2.1 y 2.2 respectivamente.

#### 3.4.1. Aplicado a EDL y EDL de Anscombe

Esta subsección es dedicada a un estudio de simulación para analizar la velocidad de convergencia de nuestra cantidad pivotal  $Z_{\lambda,l}^A(\nu_n|\nu_n)$ .

Efectuemos un estudio de simulación considerando a la matriz  $D_n$  como la matriz identidad de  $n \times n$ , ciertamente se satisface las hipótesis del Teorema 1 y el Teorema 3, esto debido a que:

$$\lim_{n \rightarrow \infty} \frac{\|D_n\|}{\sqrt{\text{tr}(D_n^2)}} = o(n^{1/3l}).$$

Por tanto, para realizar el estudio de simulación, se efectuó un estudio de Montecarlo con la siguiente mecánica:



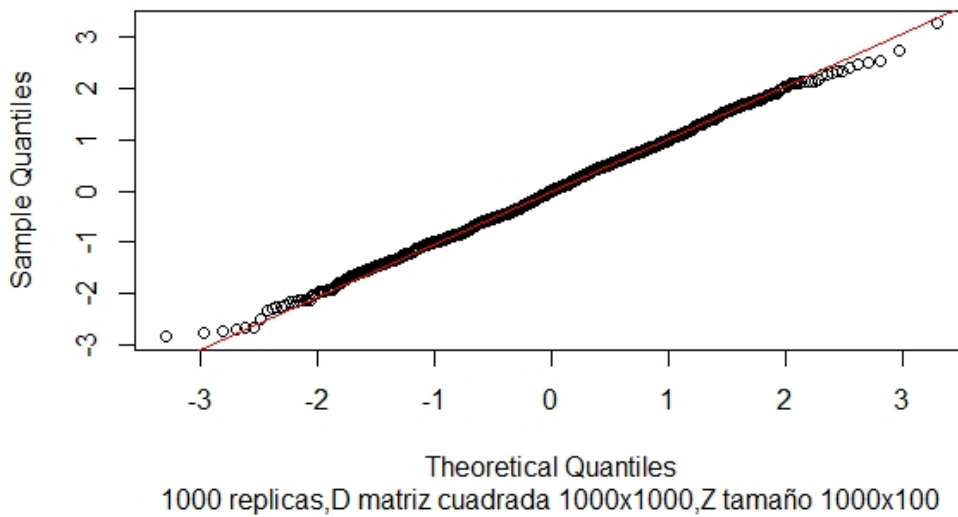
1. Se toma un entero  $n$  lo suficientemente grande que representa la dimensión de la matriz cuadrada  $D_n$ , un entero  $d$ , un número  $M$  representando el número de matrices a generar con nuestro estudio de Montecarlo y se fijan enteros  $\lambda$  representando la tasa del Z-Score equivalente determinista libre de Anscombe a estudiar.
2. Se generan  $M$  matrices aleatoria  $Z$  de dimensiones  $d \times n$  con entradas independientes y distribución deseada  $\mathcal{N}(0, 1/\sqrt{n})$  ó  $\frac{2}{n^{1/4}}(\sqrt{Poisson(\lambda) + \frac{3}{8}} - \sqrt{\lambda + \frac{3}{8}})$ .
3. Se fija un entero  $l$ , representando el orden del Z-Score equivalente determinista libre ó Z-Score equivalente determinista libre de Anscombe y se calcula  $\mu_l(\nu)$ ,  $\sqrt{Var_l(\nu)}$  y, para simplificar el cómputo suponemos que  $\mu_{\lambda,l}(\nu) \approx \mu_l(\nu)$ ,  $\sqrt{Var_{\lambda,l}(\nu)} \approx \sqrt{Var_l(\nu)}$ .
4. Se calcula y almacena en un vector de longitud  $M$  para cada una de las matrices generadas:
 
$$\frac{Tr(W_{\nu_0}^l) - \mu_l(\nu)}{\sqrt{Var_l(\nu)}}$$
 ó
 
$$\frac{Tr(P_{\lambda,\nu_0}^l) - \mu_{\lambda,l}(\nu)}{\sqrt{Var_{\lambda,l}(\nu)}}.$$
5. Dado que nos interesa comprobar normalidad, se efectúa la Q-Q plot del vector de longitud  $M$  y se discute.

Primeramente fijemos algunos elementos que serán comunes en las gráficas Q-Q a efectuar entre ellos que  $n = 1000$ ,  $d = 100$ ,  $l = 1$  por simplicidad, sin embargo, podemos calcularlos para  $l = 2$  y estimarlo mediante un límite para  $l > 2$ ,  $\lambda \in \{\frac{1}{2}, 5, 50, 1000\}$  y el número de simulaciones  $M$  las tomaremos de 1000 ó 2000 para observar de modo empírica la velocidad de convergencia.

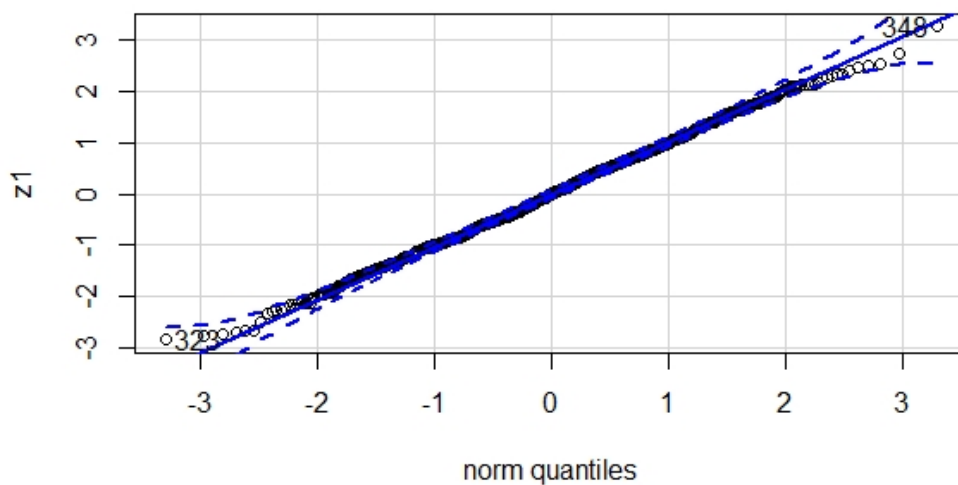
De este modo, veamos las gráficas Q-Q obtenidas al efectuar el estudio de simulación mediante el Z-Score equivalente determinista libre, con los parámetros ya mencionados.

Para el caso de que el número de simulaciones es 1000, tenemos las gráficas Q-Q con las franjas de color azul bandas de confianza al 95%:

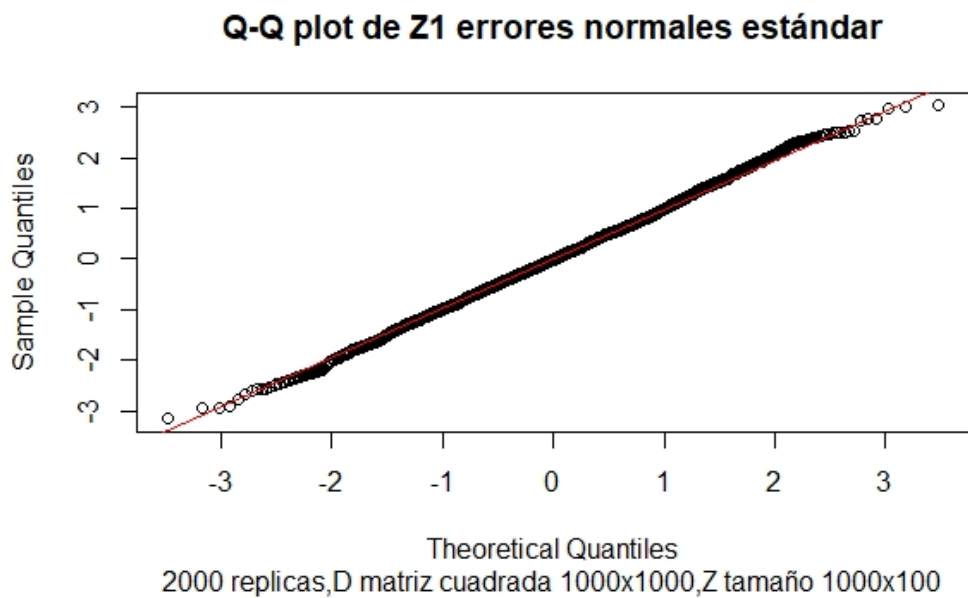
**Q-Q plot de Z1 errores normales estándar**

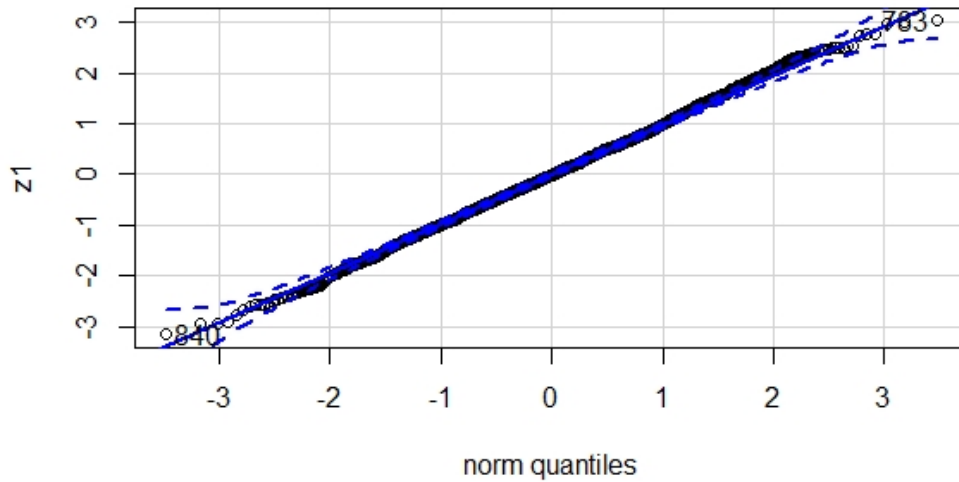


**Q-Q plot de Z1 errores normales estándar con bandas al 95%**



De aquí podemos notar diversas anotaciones, entre ellas, como es un resultado asintótico, percibimos que para  $n = 1000$  y para  $M = 1000$ , tenemos que los cuantiles empíricos del el Z-Score equivalente determinista libre calculados y los teóricos de una distribución normal se ajustan razonablemente bien a la recta identidad, incluso vemos que tanto mejora el resultado si efectuamos el estudio de Montecarlo al aumentar la cantidad de matrices generadas  $M$  a 2000.

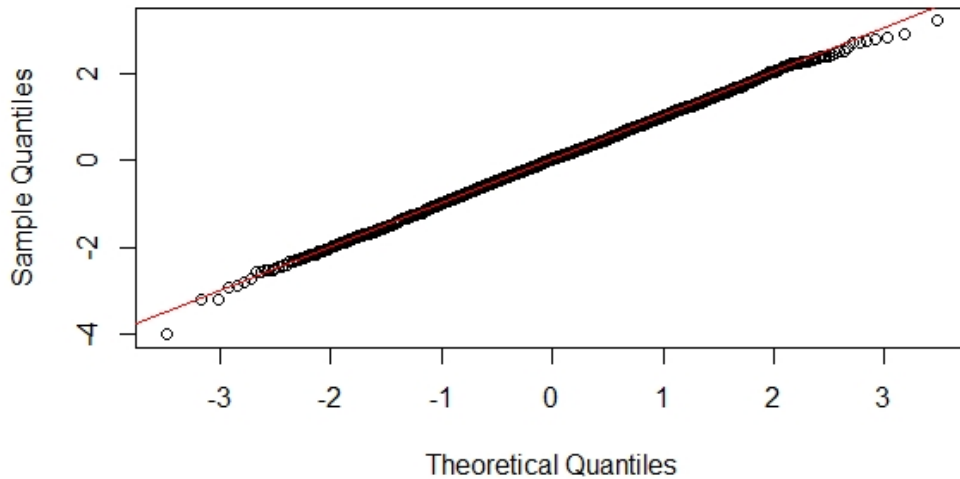


**Q-Q plot de Z1 errores normales estándar con bandas al 95%**

Notamos que el ajuste a la recta identidad se mejora razonablemente (es posible verlo en el eje x para valores cercanos a -2), sin embargo, se incrementa el costo computacional temporal y de memoria debido al incremento en el número de simulaciones. En este caso, se duplica el tiempo de cómputo y de memoria requerida, pero sigue siendo relativamente rápido en un computador de 2Gb de procesador y 4Gb de memoria RAM.

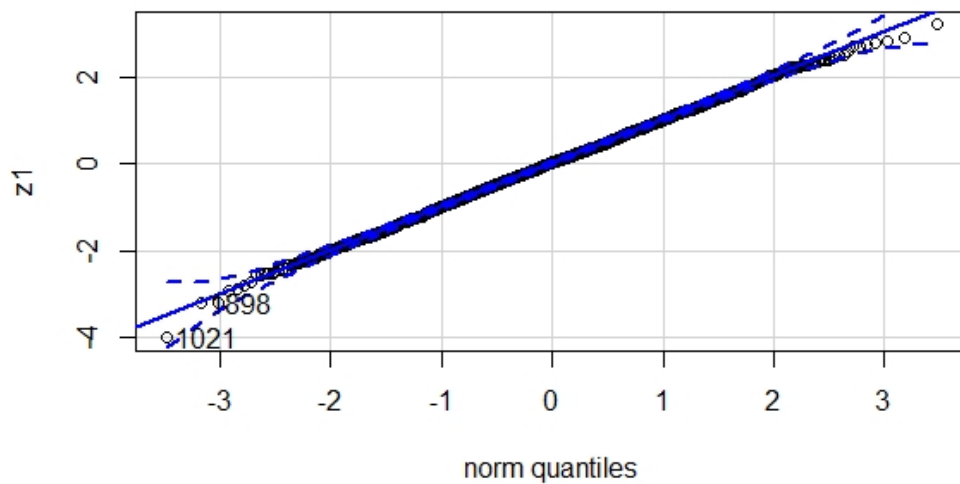
Procedamos al análisis del Z-Score equivalente determinista libre de Anscombe para los casos de  $\lambda \in \{\frac{1}{2}, 5, 50, 1000\}$ . Empecemos con  $\lambda$  decreciendo desde 1000 hasta  $\frac{1}{2}$  en nuestro conjunto objetivo, el cual genera las siguientes gráficas Q-Q con 2000 simulaciones efectuadas:

**Q-Q plot de Z1 errores función Poisson**



2000 replicas, D matriz cuadrada 1000x1000, Z tamaño 1000x100, lambda=1000

**Q-Q plot de Z1 errores función Poisson con bandas al 95%**

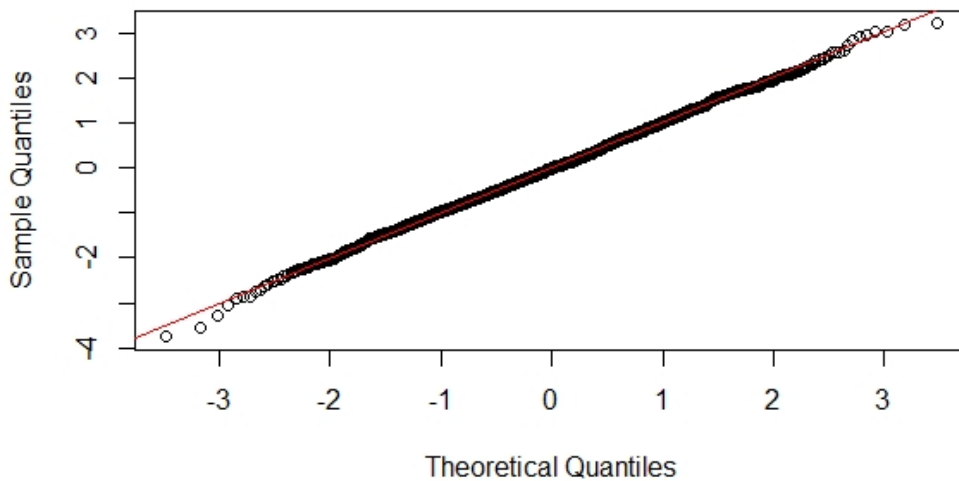


Observemos que para  $\lambda = 1000$ , las gráficas Q-Q son muy parecidas a las anteriormente presentadas, esto es razonable, debido por el Teorema 3 cuando  $\lambda, n \rightarrow \infty$  tendremos las mismas gráficas Q-Q generadas. Es importante destacar que no hay incremento significativo en el costo computacional de memoria y de tiempo.

Esto sucede en general, independientemente del valor de  $\lambda$ .

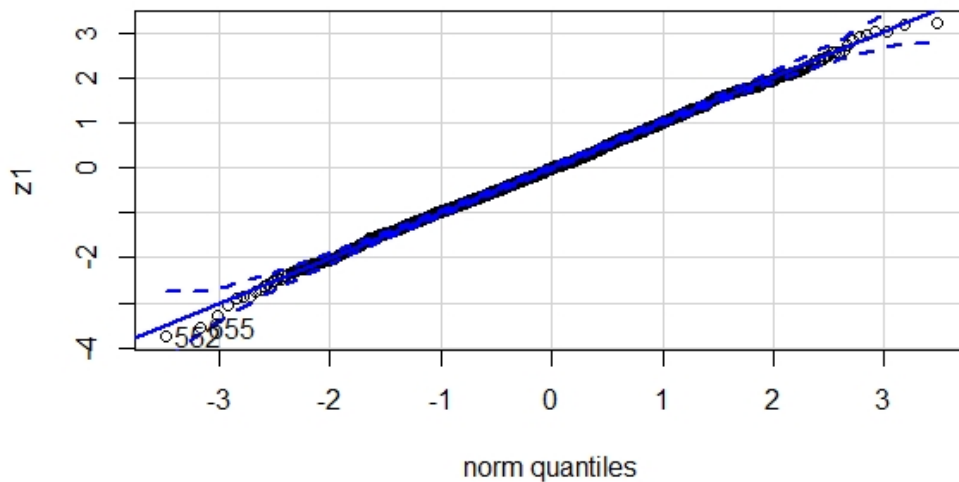
En las gráficas es esperado que el comportamiento de los puntos se empiecen a aislar de la recta identidad, y esto es, debido al decremento del parámetro  $\lambda$  y la consecuencia de la mala aproximación a la normal estándar de nuestros errores función Poisson, observémoslo de modo empírico a continuación:

**Q-Q plot de Z1 errores función Poisson**

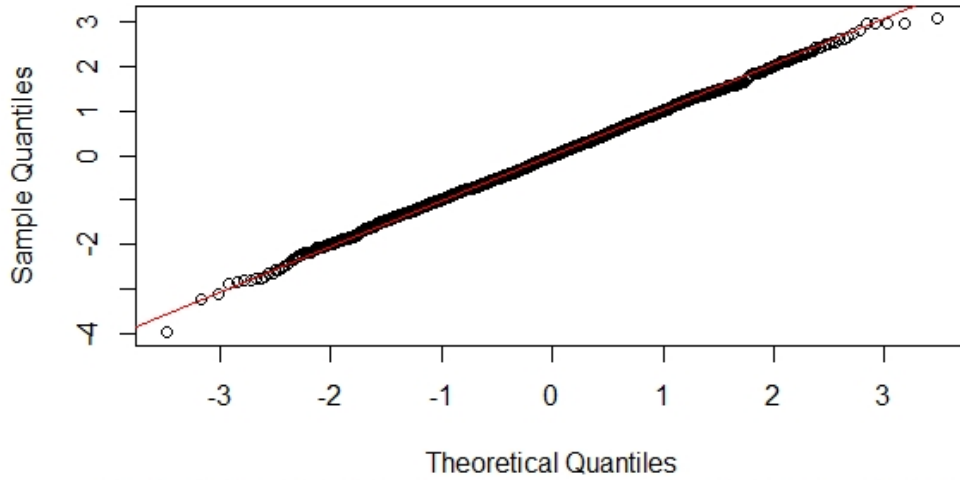


2000 replicas, D matriz cuadrada 1000x1000, Z tamaño 1000x100, lambda=50

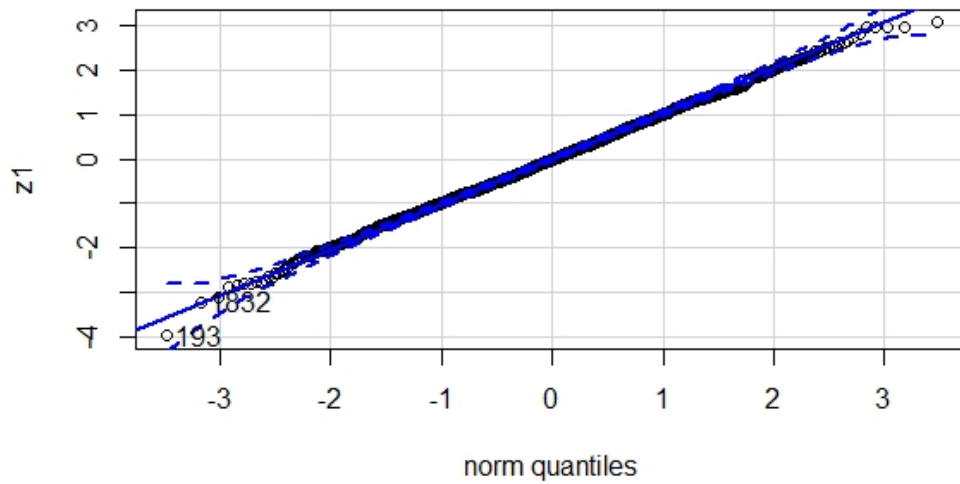
**Q-Q plot de Z1 errores función Poisson con bandas al 95%**

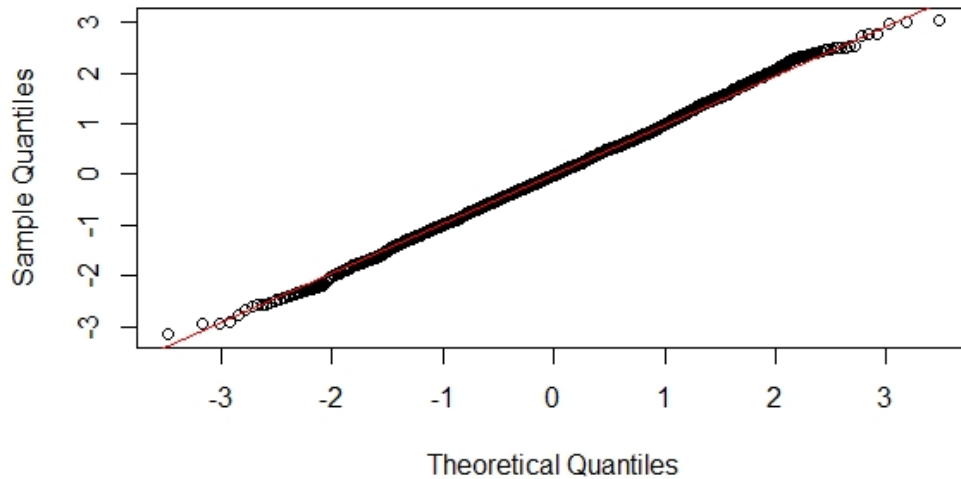
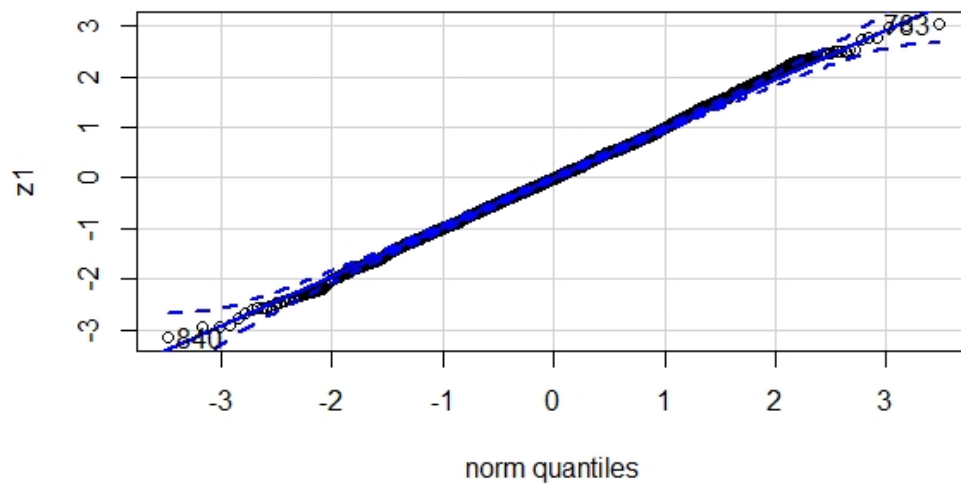


**Q-Q plot de Z1 errores función Poisson**



**Q-Q plot de Z1 errores función Poisson con bandas al 95%**



**Q-Q plot de Z1 errores función Poisson****Q-Q plot de Z1 errores función Poisson con bandas al 95%**

Notamos, que el ajuste a la recta identidad empeora significativamente para el caso de  $\lambda = \frac{1}{2}$ , sin embargo, en los casos de 5, 50 tenemos un buen comportamiento, recordemos que en nuestro caso hemos introducido el parámetro  $\lambda$ , debido a que el ruido de una imagen de rayos X se distribuye Poisson con me-



dia representando la desviación media de fotones a través de la imagen obtenida por rayos X (con media generalmente mayores a 10). Sin embargo, tal desviación media de fotones varía de imagen a imagen en términos reales y por ende, la varianza del ruido de las imágenes no es necesariamente igual. Debido a que el parámetro de la distribución Poisson depende de la imagen, se han introducido transformaciones de estabilización de varianza para generar tal independencia y así, tener una distribución de error independiente de la imagen. De este modo, surge el origen de los modelos 2D-ARMA y en la aplicación se ha visto que el parámetro  $\lambda$  de nuestros errores función Poisson tiende a estar en valores  $\lambda < 32$ . En conclusión, respecto a nuestro estudio de simulación vemos que coinciden los resultados empíricos encontrados y teóricos del Teorema 1 y el Teorema 3.

### 3.4.2. Aplicado a 2D-ARMA y 2D-ARMA de Anscombe

Esta subsección es dedicada a un estudio de simulación para analizar velocidad de convergencia de nuestra cantidad pivotal  $Z_{\lambda,l}^A(\nu_n|\nu_n)$ . Ponemos especial énfasis sobre el parámetro  $\lambda$  y haciéndolo variar para el caso de la aplicación de los modelos 2D-ARMA de Anscombe, los cuales fueron introducidos en la sección anterior. Además compararlos con la aplicación a los modelos 2D-ARMA dada en la Sección 2.2.

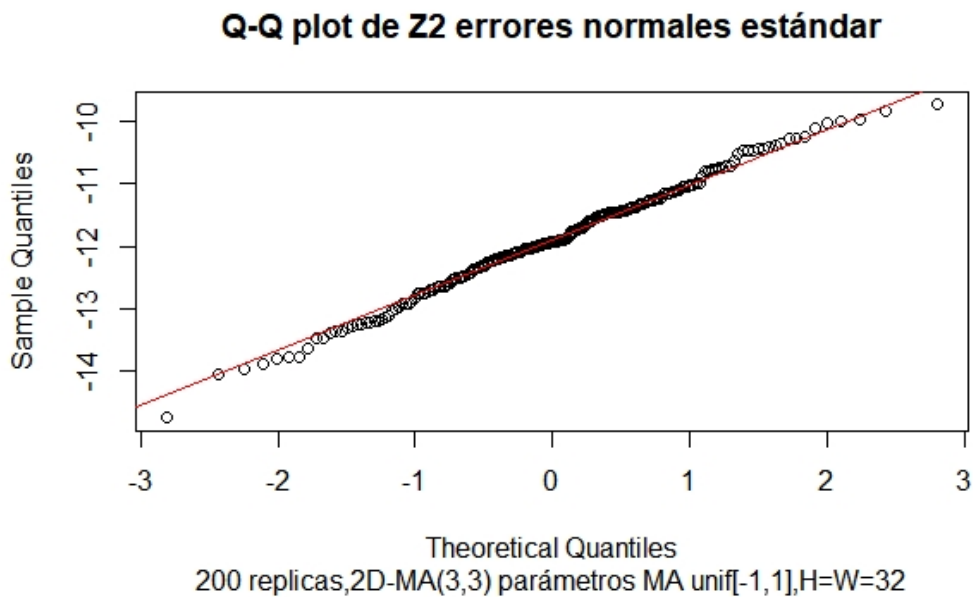
Para realizar el estudio de simulación, se efectuaron simulaciones vía el método de Montecarlo con la siguiente mecánica:

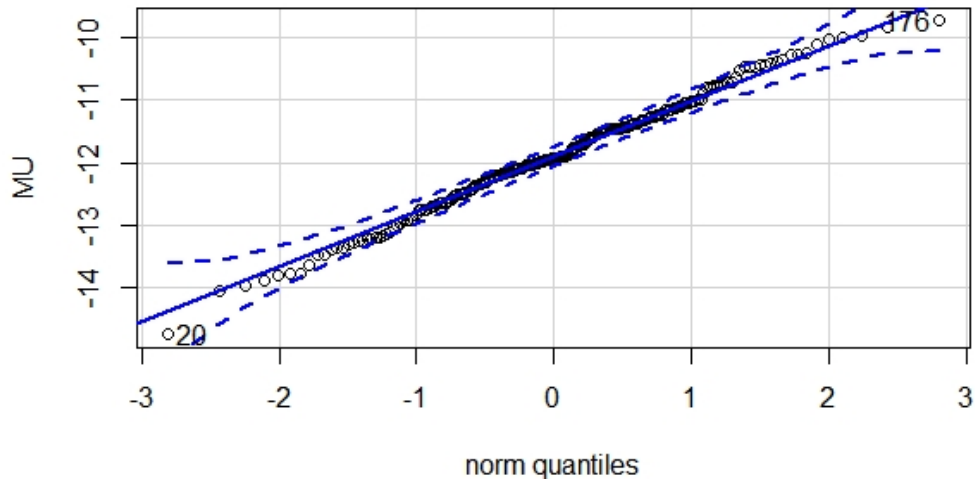
1. Se fija un modelo 2D-ARMA( $p_1, p_2, q_1, q_2$ ) o un 2D-ARMA( $p_1, p_2, q_1, q_2$ ) de Anscombe de parámetro  $\lambda$ . Es decir, se fijan,  $p_1, p_2, q_1, q_2, H, W$  y el parámetro  $\lambda$  (si es el caso), así como las matrices de parámetros, considerando para el primer caso errores normales estándar y el segundo funciones Poisson.

2. Consideramos un natural  $M$  fijo que representa el número de simulaciones a efectuar del mismo modelo fijado en el paso 1.
3. Tomamos  $N$  muestras  $\{y(i) \in \mathbb{R}^{HW}\}_{i=1,2,\dots,N}$  i.i.d. del modelo 2D-ARMA o 2D-ARMA de Anscombe fijada en el paso 1. De modo posterior, obtenemos la matriz  $X \in M_{HW,N}(\mathbb{R})$  de expresar las  $N$  muestras en forma matricial, es decir,  $X_{hw,n} = y(n)_{h,w}$ .
4. Fijamos un natural  $l$ , que representa el orden del Z-Score equivalente determinista libre o del Z-Score equivalente determinista libre de Anscombe. En nuestro caso nos restringimos al caso  $l = 1, 2$ , debido a que podemos calcularlos de modo explícito por el Lema 5 y el Lema 6. De hecho, podemos calcularlos para  $l > 2$ , sin embargo, se requiere más recursos computacionales en tiempo y memoria.
5. Procedemos a calcular  $\mu_l^* = Tr[(X^t X/N)^l]$ .
6. Se calcula la matriz  $B$ , los naturales  $H_e, W_e$  del paso 1 mediante la Sección 2.2. o la Sección 3.3. Además, computamos el parámetro  $\nu = (H_e W_e/N, B^t B)$ .
7. Se procede a usar el Lema 5 y el Lema 6 mediante  $tr(B^t B)$  para hallar  $\mu_l(\nu)$  y  $Var_l(\nu)$ , ó bien, estimar (burdamente por restricciones de recursos computacionales de memoria) a  $\mu_{\lambda,l}(\nu)$  y  $Var_{\lambda,l}(\nu)$ .
8. Se almacena  $z_l = \frac{\mu_l^* - \mu_l(\nu)}{\sqrt{Var_l(\nu)}}$  o  $z_l = \frac{\mu_l^* - \mu_{\lambda,l}(\nu)}{\sqrt{Var_{\lambda,l}(\nu)}}$  en un vector de longitud  $M$ .
9. Se repite  $M - 1$  veces del paso 3 al paso 8, almacenando los elementos  $z_l$  dejados en cada ciclo.
10. Dado que nos interesa comprobar normalidad, se efectúa la Q-Q plot del vector de longitud  $M$  y se discute.

Primeramente, fijemos algunos elementos que serán comunes en las gráficas Q-Q a efectuar entre ellos que  $M = N = 200$ ,  $l = 2$  y consideraremos un 2D-MA(3, 3). Es decir, nuestro estudio de Montecarlo será sobre 200 realizaciones del Z-Score equivalente determinista libre o de Anscombe de segundo orden, donde cada 2D-MA(3, 3) es simulado 200 veces. Como mencionamos, para el caso  $l > 2$  podemos estimarlo mediante límites, sin embargo, es pesado computacionalmente en memoria. También tomamos nuevamente  $\lambda \in \{\frac{1}{2}, 5, 50, 1000\}$ . Además, para el modelo 2D-MA(3, 3) consideraremos que cada coeficiente de la matriz de media móvil del modelo sigue una distribución uniforme en el intervalo  $[-1, 1]$ , y para nuestro modelo fijamos  $H = W = 32$ . Con estos elementos fijados, podremos observar de modo empírico la velocidad de convergencia al variar el parámetro  $\lambda$ .

De este modo, veamos las gráficas Q-Q y de color azul las bandas de confianza al 95 % obtenidas al efectuar el estudio de simulación con los parámetros ya mencionados y para el caso de 2D-MA(3, 3).

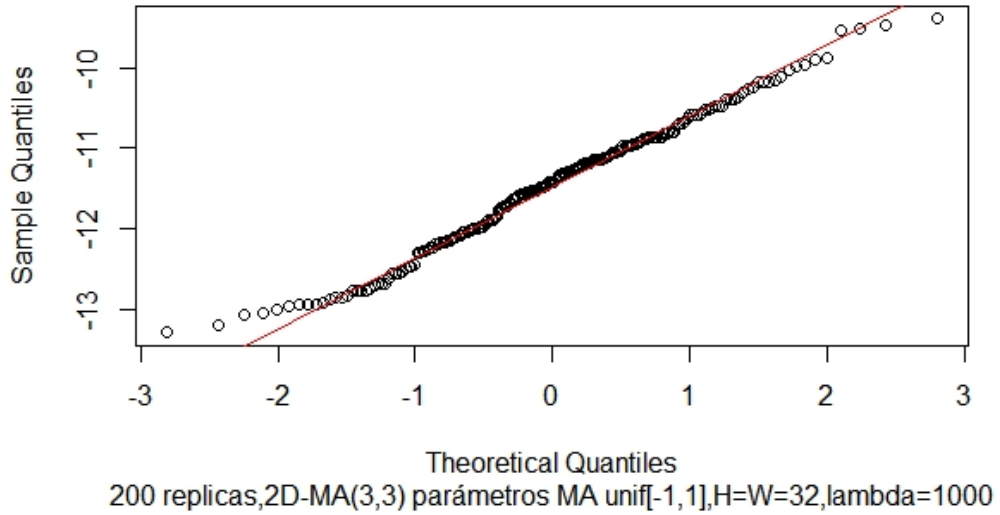


**Q-Q plot de Z2 errores normales estándar con bandas al 95%**

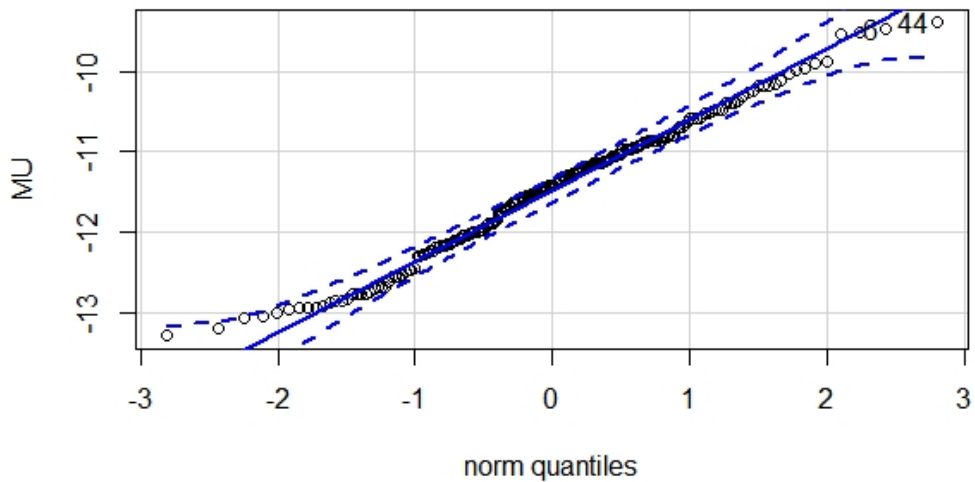
De aquí podemos notar diversas observaciones, entre ellas, como es un resultado asintótico, percibimos que para  $1024 = HW > N = 200$ , y  $M = 200$  simulaciones tenemos que los cuantiles empíricos del modelo 2D-MA calculados y los teóricos de una distribución normal se ajustan razonablemente bien a la recta identidad para un tamaño de muestra de 200 elementos por modelo 2D-MA. Además recordemos que  $HW$  representa la definición de la imagen y  $N$ , el número de imágenes que tenemos con cada uno de los modelos.

Procedamos al análisis del modelo 2D-MA(3, 3) de Anscombe para los casos de  $\lambda \in \{\frac{1}{2}, 5, 50, 1000\}$ . Empecemos con  $\lambda$  decreyendo desde 1000 hasta  $\frac{1}{2}$  en nuestro conjunto objetivo, el cual genera las siguientes gráficas Q-Q con 200 simulaciones efectuadas:

**Q-Q plot de Z2 errores función Poisson**



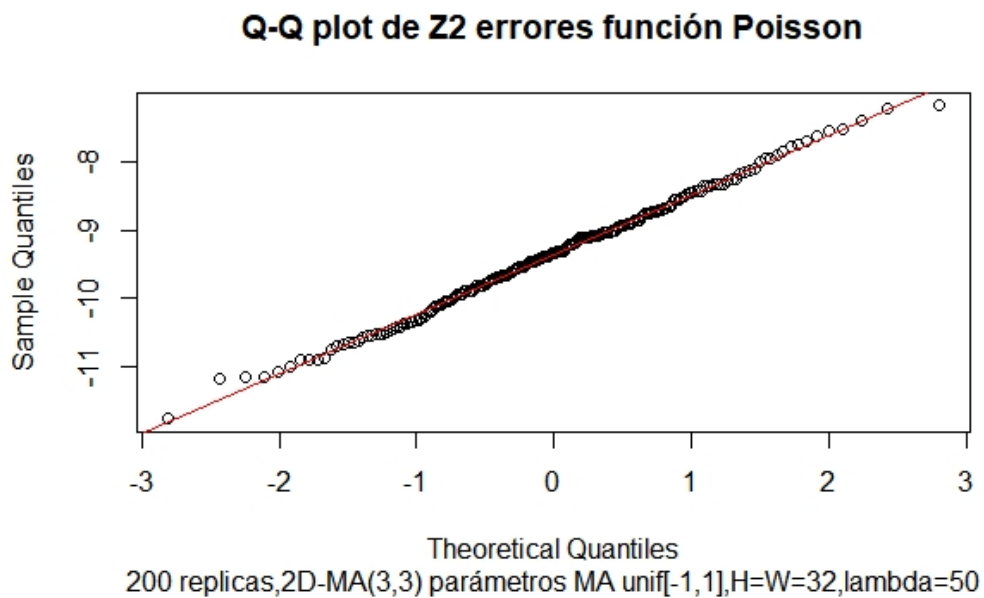
**Q-Q plot de Z2 errores función Poisson con bandas al 95%**



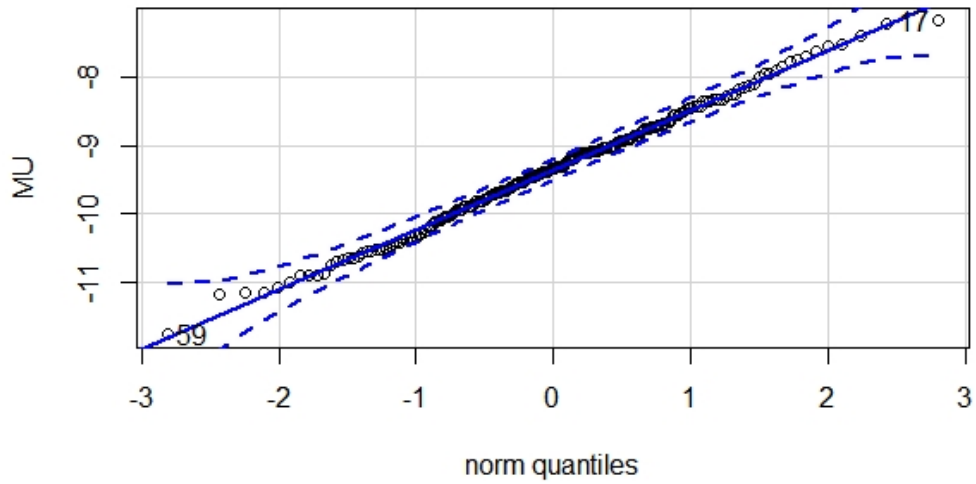
Observemos que para  $\lambda = 1000$ , las gráficas Q-Q son muy parecidas a las anteriormente presentadas, esto es razonable, debido por el Teorema 3 cuando  $\lambda, n \rightarrow \infty$  tendremos gráficas Q-Q generadas muy similares a las ya presentadas. Es importante destacar que no hay incremento significativo en el costo

computacional de memoria y de tiempo. Esto sucede en general, independientemente del valor de  $\lambda$ .

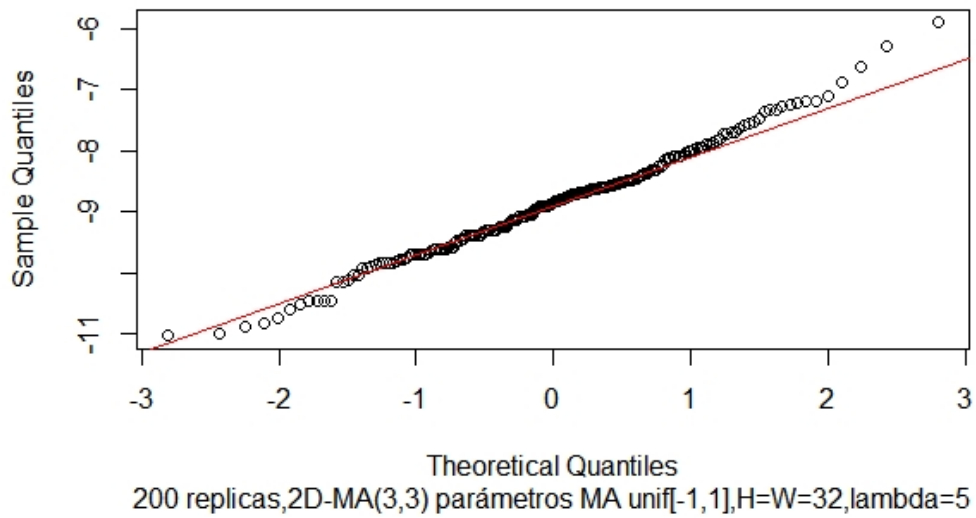
En las gráficas es esperado que el comportamiento de los puntos se empiecen a aislar de la recta identidad, y esto es debido al decremento del parámetro  $\lambda$  y la consecuencia de la mala aproximación a la normal estándar de nuestros errores función Poisson, observémoslo de modo empírico a continuación:



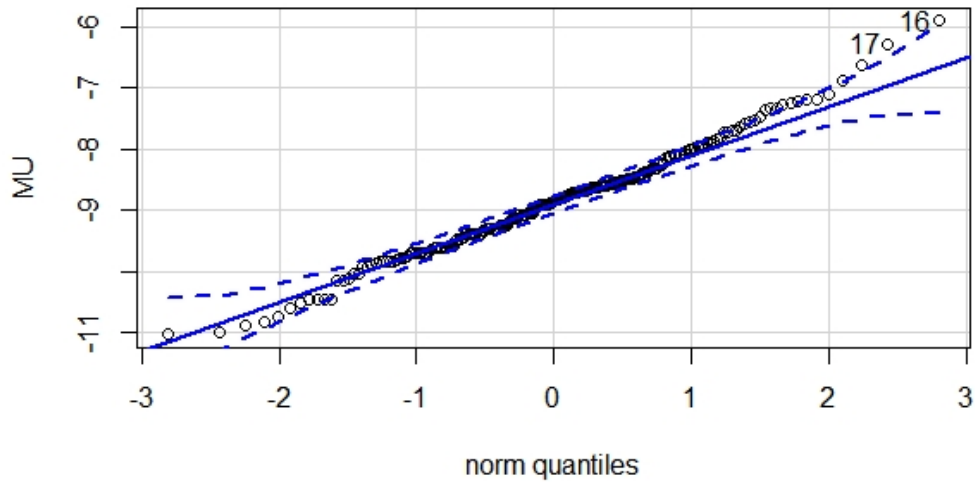
**Q-Q plot de Z2 errores función Poisson con bandas al 95%**



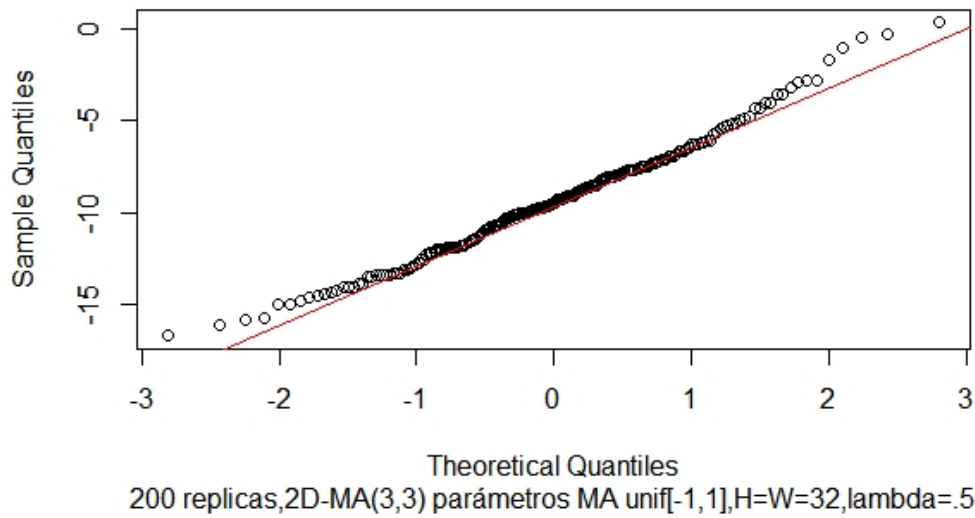
**Q-Q plot de Z2 errores función Poisson**



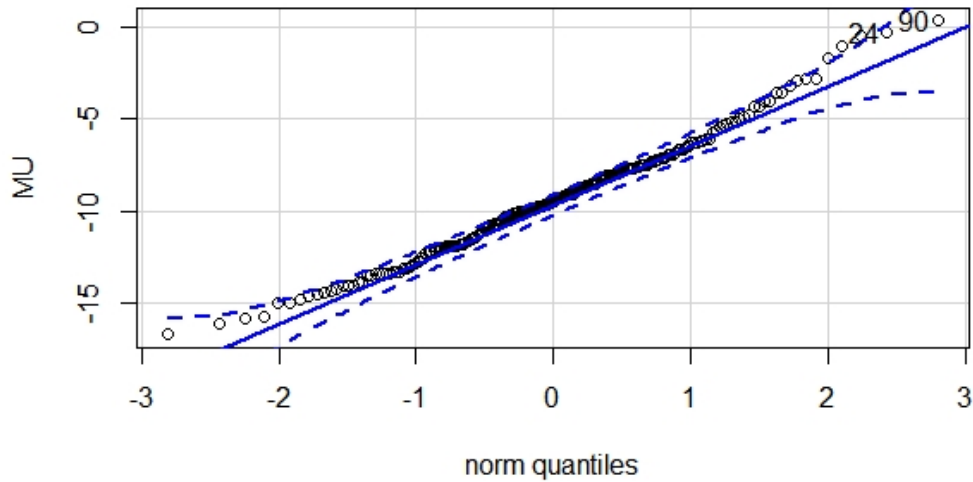
**Q-Q plot de Z2 errores función Poisson con bandas al 95%**



**Q-Q plot de Z2 errores función Poisson**





**Q-Q plot de Z2 errores función Poisson con bandas al 95%**

Observemos que el ajuste a la recta identidad empeora significativamente para los casos  $\lambda = \frac{1}{2}$  y 5. Más aún, podemos ver que hay una cantidad considerable de puntos que se salen de las bandas al 95 %. Sin embargo, en los casos de 50, 1000 tenemos una alineación más notable y los puntos caen en las bandas para 200 realizaciones de 200 modelos 2D-MA establecidos en la mecánica anterior. Esto es razonable, ya que cuando  $\lambda \rightarrow \infty$  esperamos ver gráficas Q-Q parecidas a las del caso del modelo 2D-MA presentadas al inicio de esta subsección, incluso estamos estimando para  $\lambda$  pequeños a  $\mu_{\lambda,l}(\nu)$  y  $Var_{\lambda,l}(\nu)$  como  $\mu_l(\nu)$  y  $Var_l(\nu)$  respectivamente. Es decir, introducimos más margen error mediante la forma de calcularlo, debido a que es computacionalmente costoso para cada  $M$  de 1 a 200 estimar  $\mu_{\lambda,l}(\nu)$  y  $Var_{\lambda,l}(\nu)$ . Es de notar, que estos procesos son significativamente más exigentes computacionalmente en recursos que los de la pasada subsección.

Como mencionamos en la Subsección 3.4.1. el parámetro  $\lambda$  tiene intrínse-

ca relación al ruido de una imagen de rayos X y, por ello, se introducen transformaciones de estabilización de varianza. Así, surge el origen de los modelos 2D-ARMA y en la aplicación se ha visto que el parámetro  $\lambda$  de nuestros errores función Poisson tiende a estar en valores  $\lambda < 32$ .

En conclusión, respecto a nuestro estudio de simulación vemos que coinciden los resultados empíricos encontrados y teóricos del Teorema 2 y el Teorema 4. Sin embargo, con lenta convergencia *ceteris paribus* en función del parámetro  $\lambda$ .

## Capítulo 4

# Discusión y Conclusiones

En esta tesis presentamos una propuesta de modelaje para el tratamiento de imágenes de rayos X, que surge de modo natural mediante los modelos 2D-ARMA introducidos por [8]. Con la principal diferencia en la distribución de errores subyacentes y mediante la introducción de un parámetro (asociado a la tasa media de fotones desviados que difiere de equipo a equipo de rayos X), es a través de esta necesidad que surge el uso natural de la transformada de Anscombe centralizada. Similarmente a como lo hace [1], es decir, con el punto clave que le permite aproximar eficientemente las desviaciones del modelos de Wishart compuesto (por lo que no es necesario determinar el límite del valor propio de los parámetros de tales matrices). Podemos definir el Z-Score equivalente determinista libre de Anscombe, el cual probamos que es una cantidad pivotal.

Mediante este resultado obtenemos una cantidad pivotal que es aplicada a modelos 2D-ARMA de Anscombe con interpretación de este hecho en imágenes de rayos X, que se encuentra cuando la postulación de nuestros errores sigue una distribución más razonablemente a [15, 17], resulta de la motivación del problema original de [8]. Además de que la muestra y la calidad de las imágenes obtenidas tiendan al infinito con la muestra creciendo a un ritmo menor o igual al de la cali-

dad de las imágenes, adicionado con que la tasa de nuestro modelo 2D-ARMA de Anscombe tienda al infinito, entonces el Z-Score equivalente determinista libre de Anscombe de cualquier orden es una cantidad pivotal con distribución normal estándar.

Mostramos mediante estudios de simulación que nuestras cantidades pivotaes se aproximan a normales para  $\lambda > 5$ . Incluso resultó que no requerimos que los modelos sean de un tamaño muy grande, y funciona relativamente razonable para los modelos 2D-ARMA de  $32 \times 32$ .

Es posible extender este trabajo en varias líneas. Una de ellas, es dado que nosotros estimamos a  $\mu_{\lambda,l}(\nu)$  y  $Var_{\lambda,l}(\nu)$  como  $\mu_l(\nu)$  y  $Var_l(\nu)$  respectivamente, entonces para  $\lambda$  pequeños (menores o iguales a 5) tenemos que nuestra estimación es muy burda, y por tanto una línea es estimarlos de mejor modo (pues el costo computacional en tiempo y memoria es alto) o hallarlos de modo explícito para que el error observado de modo empírico de la Sección 3.4.2 sea de menor medida ocasionado por el método de estimación para casos de  $\lambda$  pequeños. Otra línea, es usar estos modelos 2D-ARMA de Anscombe para emplearlos con bases de datos reales de radiografías y compararlos con los resultados de [8] para el caso de clasificación y predicción de cáncer de seno. Dado que estamos usando en principio una distribución de error más realista, es esperado que la tasa de clasificación sea mejor que las obtenidas del trabajo de [8]. Sin embargo, no pudimos efectuar este estudio debido a que son bases de datos privadas las usadas en [8]. Finalmente, una dirección futura es usar métodos de [5] para encontrar cantidades pivotaes que difieren de las aquí presentadas, conjeturamos que esta línea puede ser muy interesante por la aplicación de tales cantidades pivotaes a hamiltonianos en un contexto de física atómica.

# Bibliografía

- [1] T. Hayase. Free deterministic equivalent Z-scores of compound Wishart models: A goodness of fit test of 2D-ARMA models. *Random Matrices: Theory and Applications*, pages 19, Vol. **8**, No. 2, 2019.
- [2] A. Hasegawa, N. Sakuma and H. Yoshida. Fluctuations of Marchenko–Pastur limit of random matrices with dependent entries. *Statistics Probability Letters*, **127**, 85-96, August 2017.
- [3] M. A. Nowak and W. Tarnowski. Spectra of large time-lagged correlation matrices from random matrix theory. *Journal of Statistical Mechanics: Theory and Experiment*, 063405, 2017.
- [4] A. Hasegawa, N. Sakuma and H. Yoshida. Random matrices by MA models and compound free poisson laws. *Probability and Mathematical Statistics*, **33**(2), 243-254, 2013.
- [5] E. Redelmeier. Real second-order freeness and the asymptotic real second-order freeness of several real matrix models. *International Mathematics Research Notices*, **3353-3395**, 2014.
- [6] R Speicher. *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory* . (Vol. **627**). American Mathematical Soc. 1998.

- [7] W. Bryc. Compound real Wishart and q-Wishart matrices. *International Mathematics Research Notices*, **2008**(9), 2008.
- [8] N. Bouaynaya, J. Zielinski and D. Schonfeld. Two-Dimensional ARMA modeling for breast cancer detection and classification. In *2010 International Conference on Signal Processing and Communications (SPCOM)* (pp. 1-4). IEEE.
- [9] R. Speicher and C. Vargas with an appendix by T. Mai. Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory. *Random Matrices: Theory and Applications*, **1**(02), 1150008, 2012.
- [10] S. Thurner and C. Biely. Random matrix ensembles of time-lagged correlation matrices: derivation of eigenvalue spectra and analysis of financial time-series *Quantitative Finance*, **8**(7), 705-722, 2006.
- [11] M. A. Nowak, Z. Burda, A. Jarosz, J. Jurkiewicz, G. Papp and I. Zahed. Applying free random variables to random matrix analysis of financial data Part I: A Gaussian case *Quantitative Finance*, **11**(7), 1103-1124, 2011.
- [12] M. J. Russell, A. Nel and T. Marwala. ARMA Analysis of Chest X-rays for Computer Assisted Detection of Tuberculosis. In *World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China* (pp. 896-899). Springer, Berlin, Heidelberg.
- [13] G. Zamuner and J. Jacot. A System for the Quality Inspection of Surfaces of Watch Parts. In *International Precision Assembly Seminar* (pp. 134-143). Springer, Berlin, Heidelberg. February, 2012.
- [14] M. Cesarelli, P. Bifulco, T. Cerciello, M. Romano, L. Paura. X-ray fluoroscopy noise modeling for filter design *International journal of computer assisted radiology and surgery*, **8**(2), 269-278. 2013.

- [15] I.A. Elbakri, J.A. Fessler. Poisson–Gaussian Noise Analysis and Estimation for Low-Dose X-ray imágenes in the NSCT Domain *Sensors*, **18**(4), 1019. 2002.
- [16] F.J. Anscombe. The transformation of Poisson, binomial and negative-binomial data *Biometrika*, **35**(3/4), 246-254. 1948.
- [17] L. Sangyoon, S. Min and G. K. Moon. Poisson–Gaussian Noise Analysis and Estimation for Low-Dose X-ray imágenes in the NSCT Domain. *Sensors (Basel)*, Mar. 2018.
- [18] P. Neu and R. Speicher, Rigorous mean-field theory for Coherent Potential Approximation: Anderson model with free random variables, *J. Stat. Phys.* **80**, 1279–1308. 1995.
- [19] Vyacheslav Girko, Theory of Stochastic Canonical Equations, *Mathematics and Its Applications*, vol. **1200**, Kluwer Academic Publications. 2001.
- [20] D. Voiculescu , Limit laws for random matrices and free products, *Invent. Math.* **104**, 201–220. 1991.
- [21] Vladimir Marchenko and Leonid Pastur, Distribution of eigenvalues for some sets of random matrices, *Math. USSR-Sbornik* **1**, 457–483. 1967.