

UNIVERSIDAD AUTONOMA METROPOLITANA

UAM-IZTAPALAPA

---

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

Análisis Bayesiano de Modelos Lineales para Datos Direccionales  
considerando la Distribución Normal bajo Proyección

T E S I S

QUE PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS

(MATEMÁTICAS)

PRESENTA

**GABRIEL NUÑEZ ANTONIO**

Dr. Gabriel Escarela Pérez

*UAM-I*

Asesor

Dr. Eduardo Gutiérrez-Peña

*IIMAS-UNAM*

Co-asesor

Junio 1, 2010.

## AGRADECIMIENTOS:

- A la UAM-I, particularmente al departamento de Matemáticas, por darme la oportunidad de concluir este proyecto.
- A mi asesor el Dr. Gabriel Escarela Pérez, por su tiempo, por todo su soporte durante el desarrollo de este trabajo y por su amistad.
- A la maestra Iseo González Christen ya que como asistente del posgrado siempre ha apoyado de manera eficiente a los alumnos, particularmente, su gran disposición me permitió concluir en tiempo este trámite.
- A mis sinodales: el Dr. Alberto Castillo, la Dra. Blanca Rosa Pérez Salvador, el Dr. Russell Bowater, el Dr. Manuel Mendoza Ramírez y el Dr. Eduardo Gutiérrez Peña. Por sus valiosas observaciones a este trabajo, pero sobre todo, por el apoyo que siempre me han ofrecido.
- A todos mis colegas y amigos de la UAM-I, del IIMAS-UNAM y del ITAM. Por ese ánimo que me sirvió para seguir adelante. Al Dr. Luis Felipe González, por compartir sus experiencias en el análisis gráfico de datos. Al Dr. Rubén Hernández Cid, por su gran calidad humana y la amistad que siempre me ha brindado. Al Dr. Luis Enrique Nieto por su legítimo interés de que este proyecto llegara a buen término. De manera particular, al Dr. José Ma. González Barrios, ya que sin saberlo su peculiar insistencia me hizo continuar en los momentos complicados. A Angélica Torres, porque siempre se puede contar con ella, especialmente en los momentos difíciles.

- Al Dr. Ernesto Barrios, a quien durante el desarrollo de este trabajo, tuve la oportunidad de conocer, de colaborar y de cultivar una gran amistad. Por la motivación que siempre me ha brindado, por su tiempo, por su brutal honestidad, pero sobre todo por ser un gran ser humano.
- Al Dr. Alberto Contreras, a Ernesto Barrios y a Eduardo Gutiérrez-Peña, quienes de manera desinteresada me permitieron usar los servidores Zamudio, Pascal y Nightwish para realizar mis últimas simulaciones.
- Al Dr. Manuel Mendoza, ya que a pesar de sus compromisos tuvo la amabilidad de darse un tiempo para participar en este proceso.
- De manera muy especial al Dr. Eduardo Gutiérrez Peña. No tengo palabras para agradecer su gran apoyo incondicional, por todo el tiempo invertido, por haber creído siempre en mi, pero sobre todo por su valiosa amistad. Su ejemplo como excelente profesionalista y excepcional ser humano siempre me impulsó a seguir adelante en este proyecto y a ser mejor cada día.

*A mi esposa  
y a mis hijos.*

# Indice

<b>Notación y Acrónimos</b>	<b>vii</b>
<b>Resumen</b>	<b>1</b>
<b>Introducción</b>	<b>5</b>
<b>1 Preliminares</b>	<b>10</b>
1.1 Naturaleza de los datos direccionales . . . . .	10
1.1.1 Medidas descriptivas . . . . .	11
1.2 El caso de datos circulares . . . . .	14
1.2.1 Medidas descriptivas muestrales . . . . .	17
1.2.2 La función característica . . . . .	23
1.3 Modelos para datos direccionales . . . . .	25
1.3.1 La distribución de von Mises-Fisher . . . . .	27
1.3.2 Relación entre algunas distribuciones . . . . .	28
1.4 Modelos lineales para una respuesta circular . . . . .	30
1.4.1 Modelos de regresión . . . . .	30
1.4.2 Modelos para datos longitudinales . . . . .	31
<b>2 La Distribución Normal Proyectada</b>	<b>33</b>

2.1	La distribución Normal proyectada $q$ -variada . . . . .	33
2.2	Derivación de la distribución Normal proyectada $PN(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ . . . . .	38
2.3	La distribución Normal proyectada $PN(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ . . . . .	41
<b>3</b>	<b>Modelos de Regresión para Datos Circulares</b>	<b>44</b>
3.1	Antecedentes . . . . .	44
3.2	El modelo . . . . .	47
3.2.1	El problema de datos faltantes . . . . .	49
3.2.2	El problema de selección de modelos . . . . .	51
3.3	Inferencias vía métodos MCCM . . . . .	52
3.3.1	Distribuciones condicionales . . . . .	53
3.3.2	Datos faltantes . . . . .	54
3.3.3	Selección de modelos . . . . .	55
3.4	Ejemplos . . . . .	56
3.5	Discusión . . . . .	65
<b>4</b>	<b>Un Modelo Longitudinal para Datos Circulares</b>	<b>67</b>
4.1	Introducción a los estudios longitudinales . . . . .	67
4.1.1	Diversos enfoques . . . . .	68
4.1.2	Modelos de regresión de efectos-mixtos . . . . .	70
4.2	Especificación de un modelo mixto . . . . .	74
4.3	Un modelo para datos longitudinales circulares . . . . .	76
4.3.1	Antecedentes . . . . .	76
4.3.2	El modelo longitudinal propuesto . . . . .	78

4.3.3	Análisis del modelo . . . . .	81
4.4	Inferencias vía métodos MCCM . . . . .	83
4.5	Implementación . . . . .	85
4.6	Simulación por Bloques . . . . .	86
4.6.1	Ejemplos . . . . .	89
4.7	Discusión . . . . .	108
<b>5</b>	<b>Conclusiones y Perspectivas</b>	<b>111</b>
	<b>Referencias Bibliográficas</b>	<b>116</b>
	<b>Apéndice A</b>	<b>124</b>
	<b>Apéndice B</b>	<b>131</b>
	Programa: CircularRegression.R . . . . .	132
	Programa: RegressionFunctions.R . . . . .	136
	Programa: MissingData.R . . . . .	144
	Programa: Normal-proyectada.nb . . . . .	151
	Programa: LCP.R . . . . .	155
	<b>Apéndice C</b>	<b>165</b>
	Datos: Tortugas . . . . .	165
	Datos: Ejemplo 3.1 . . . . .	166
	Datos: Ejemplo 3.2 . . . . .	168
	Datos: Ejemplo 3.3 . . . . .	170
	Descripción de datos: Ejemplo 3.4 . . . . .	171
	Datos: Ejemplo 4.1 . . . . .	172
	Datos: Ejemplo 4.2 . . . . .	174

# Lista de Figuras

1.1	<i>Diagrama circular para los datos de tortugas.</i>	14
1.2	<i>Diagrama de rosa para los datos de tortugas.</i>	15
1.3	<i>Representación de la dirección <math>\mathbf{u}</math> por medio de un ángulo <math>\theta</math> y de un número complejo <math>z</math>.</i>	16
2.1	<i>Distribución Normal proyectada simétrica.</i>	36
2.2	<i>Distribución Normal proyectada asimétrica.</i>	36
2.3	<i>Distribución Normal proyectada bimodal asimétrica.</i>	37
2.4	<i>Distribución Normal proyectada bimodal simétrica.</i>	37
3.1	<i>Distribución final de <math>\mathbf{B}</math> (por componente) para el ejemplo 3.1. El verdadero valor de cada uno de los parámetros de regresión aparece entre paréntesis.</i>	57
3.2	<i>Conjunto de 100 observaciones (<math>\bullet</math>) de un modelo de regresión von Mises. Los valores <math>\theta + 360</math> están representados por (<math>\blacktriangle</math>). La línea representa la dirección media ajustada con base a un modelo Normal proyectado.</i>	58
3.3	<i>Densidades finales de <math>\beta^1</math> y <math>\beta^2</math> (por componente) para los datos de caracoles.</i>	59
3.4	<i>Intervalos predictivos finales al 95% para los datos de caracoles. La línea continua representa la dirección media ajustada.</i>	60



3.5	<i>Distribuciones finales de los valores faltantes <math>\theta_{13}</math> (en <math>x = 5</math>), <math>\theta_{23}</math> (en <math>x = 57</math>) y <math>\theta_1</math> (en <math>x = 107</math>), para los datos de caracoles.</i>	61
4.1	<i>Datos circulares longitudinales generados a partir del modelo LCP.</i>	91
4.2	<i>Distribuciones finales de los vectores <math>\beta^I</math> y <math>\beta^{II}</math> para el Ejemplo 4.1.</i>	92
4.3	<i>Distribuciones finales de los parámetros de la matriz de covarianzas <math>(\Omega^{II})^{-1}</math> para el Ejemplo 4.1.</i>	93
4.4	<i>Promedio ergódicos de los vectores <math>\beta^I</math> y <math>\beta^{II}</math>.</i>	94
4.5	<i>Promedios ergódicos para los parámetros de la matriz de covarianzas <math>(\Omega^{II})^{-1}</math>.</i>	94
4.6	<i>Autocorrelaciones para los elementos de los vectores <math>\beta^I</math> y <math>\beta^{II}</math>.</i>	95
4.7	<i>Autocorrelaciones para los parámetros de la matriz de varianzas y covarianzas <math>(\Omega^{II})^{-1}</math>.</i>	95
4.8	<i>Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.</i>	97
4.9	<i>Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.</i>	97
4.10	<i>Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.</i>	98
4.11	<i>Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.</i>	98
4.12	<i>Gráfica longitudinal de las direcciones de escape de 65 talitrus saltator en 5 ocasiones consecutivas.</i>	100
4.13	<i>Distribuciones finales de los parámetros en la componente I del modelo LCP para los datos de talitrus saltators.</i>	102

4.14	<i>Distribuciones finales de los parámetros en la componente II del modelo</i>	
	<i>LCP para los datos de talitrus saltators.</i>	103
4.15	<i>Promedios ergódicos para los parámetros en la componente I del modelo</i>	
	<i>LCP para los datos de talitrus saltators.</i>	104
4.16	<i>Promedios ergódicos para los parámetros en la componente II del modelo</i>	
	<i>LCP para los datos de talitrus saltators.</i>	105
4.17	<i>Autocorrelaciones de los parámetros en la componente I del modelo LCP</i>	
	<i>para los datos de talitrus saltators.</i>	106
4.18	<i>Autocorrelaciones de los parámetros en la componente II del modelo LCP</i>	
	<i>para los datos de talitrus saltators.</i>	107

# Lista de Tablas

3.1	<i>Medidas predictivas para selección de covariables para los datos de caracoles.</i>	62
3.2	<i>Intervalos finales de credibilidad al 95% para cada uno de los componentes de <math>\beta^1</math> y <math>\beta^2</math>.</i>	64
3.3	<i>Medidas predictivas para la selección de covariables, datos REDMET.</i>	65
4.1	<i>Intervalos finales de credibilidad al 95% para cada uno de los componentes de los vectores <math>\beta^I</math> y <math>\beta^{II}</math>, así como para los elementos relevantes de la matriz <math>(\Omega^{II})^{-1}</math>.</i>	92
4.2	<i>Intervalos finales de credibilidad al 95% para cada uno de los componentes del modelo LCP para los datos de talitrus saltators.</i>	108

# Notación y Acrónimos

<i>MCCM</i>	Monte Carlo vía Cadenas de Markov
<i>GEE</i>	Ecuaciones Estimadoras Generalizadas
<i>f.d.p.</i>	función de densidad de probabilidad
<i>v.a.</i>	variable aleatoria
<i>vs.as.</i>	variables aleatorias
<i>m.a.</i>	muestra aleatoria

$N_q(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \equiv N_q(\cdot   \boldsymbol{\mu}, \boldsymbol{\Lambda})$	Normal $q$ -variada con vector de medias $\boldsymbol{\mu}$ y matriz de precisión $\boldsymbol{\Lambda}$
$Wi(t, T)$	Wishart con parámetros $t$ y $T$
$N_2(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \equiv N_2(\cdot   \boldsymbol{\mu}, \boldsymbol{\Lambda})$	Normal bivariada con vector de medias $\boldsymbol{\mu}$ y matriz de precisión $\boldsymbol{\Lambda}$
$PN_q(\boldsymbol{\mu}, \boldsymbol{\Lambda})$	Normal proyectada $q$ -variada con parámetros $\boldsymbol{\mu}$ y $\boldsymbol{\Lambda}$
$PN(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \equiv PN(\cdot   \boldsymbol{\mu}, \boldsymbol{\Lambda})$	Normal proyectada con parámetros $\boldsymbol{\mu}$ y $\boldsymbol{\Lambda}$

# Resumen

En el área de análisis de datos, el investigador se puede encontrar con observaciones de tipo direccional. Particularmente, este tipo de datos puede aparecer en modelos lineales como la variable de respuesta. Por lo anterior, se tiene la necesidad de contar con modelos que describan la relación de dependencia entre una variable de tipo direccional y un conjunto de variables explicativas o covariables.

El estudio de datos con estructuras de dependencia es un reto ya sea en términos técnicos o desde el punto de vista aplicado. Aún más, cuando se trabaja con datos donde la variable de respuesta es de tipo direccional, proponer modelos que permitan analizar adecuadamente relaciones de dependencia, requiere de un esfuerzo adicional.

En el contexto de regresión se han propuesto varios modelos en la literatura que consideran una respuesta de tipo circular - direcciones en dos dimensiones -. Sin embargo, la mayoría de ellos presentan limitaciones para realizar inferencias adecuadas para todos los parámetros involucrados en el modelo. Adicionalmente, algunos presentan diversos problemas numéricos que son relevantes en su implementación práctica.

Por su parte, cuando se trabaja con conjuntos de datos circulares de tipo longitudinal el problema se complica. Por un lado, hay que tomar en cuenta la dependencia natural que se tiene cuando se trabaja con medidas repetidas para un mismo conjunto de individuos. Por otro lado, el tratamiento de la naturaleza direccional de los datos implica trabajar con modelos probabilísticos especiales; para los cuales los métodos inferenciales asociados aún se encuentran en una etapa de continuo desarrollo. Lo anterior, da como resultado un vacío tanto en el planteamiento de modelos, que describan adecuadamente estructuras longitudinales de tipo circular, así como en propuestas para realizar inferencias en esta clase de modelos.

Este trabajo de investigación pretende contribuir al estudio de modelos lineales para datos direccionales. Particularmente, se implementa un análisis Bayesiano de un modelo de regresión donde la variable de respuesta es circular. Este modelo de regresión se basa en la distribución de probabilidad que se obtiene al proyectar radialmente una distribución normal bivariada específica. Adicionalmente, en este contexto, se analizan y discuten los problemas de datos faltantes y selección de modelos.

Para el tratamiento de datos longitudinales donde la variable de respuesta es de tipo circular, en este trabajo se introduce un modelo basado en una versión de la distribución normal bivariada proyectada, donde cada uno de sus componentes se define a través de un modelo lineal de efectos mixtos. Por último, se propone una metodología Bayesiana para realizar inferencias para todos los parámetros involucrados en el modelo.

En la literatura los modelos de probabilidad más utilizados para describir datos direccionales son los modelos tipo von Mises y los modelos *wrapped* o “*envueltos*” (descritos en el siguiente capítulo). Aunque los modelos anteriores son útiles para describir comportamientos probabilísticos para datos direccionales, los procedimientos para llevar a cabo inferencias adecuadas sobre todos los parámetros involucrados en los respectivos modelos, no han sido del todo desarrollados. Aún más, cuando el objetivo es modelar datos con una estructura de dependencia, relativamente poco trabajo se ha realizado y, los modelos más comúnmente empleados en este caso son los modelos tipo von Mises y los modelos *wrapped*. Sin embargo, el análisis y la aplicabilidad de este tipo de modelos no ha sido directa ni desarrollada del todo de manera satisfactoria. Lo anterior, por un lado, debido a la complejidad del análisis que implica trabajar con este tipo de modelos y por otro lado, la particular estructura de dependencia de los datos.

En ese proyecto de investigación se decidió trabajar con un modelo generado al proyectar radialmente una distribución bivariada, la distribución normal proyectada. Este tipo de modelos proyectados, al igual que los modelos tipo von Mises y los modelos *wrapped*, no son la excepción en cuanto a la dificultad de su análisis, y han resultado históricamente quizá los más complejos de estudiar. Sin embargo, con los recursos computacionales de que se disponen hoy en día y bajo un enfoque inferencial Bayesiano, como se muestra en este trabajo el modelo normal proyectado presenta ciertas características que lo vuelven atractivo para describir comportamientos probabilísticos de datos direccionales y, en particular para modelar relaciones de dependencia.

Los procedimientos para llevar a cabo las inferencias para los modelos analizados en este trabajo se basan fundamentalmente en métodos de Monte Carlo vía Cadenas

de Markov (MCCM). Finalmente, las metodologías desarrolladas en esta tesis se ilustran empleando conjuntos de datos tanto simulados como reales y datos previamente analizados en la literatura.



# Introducción

En varias campos del conocimiento se pueden encontrar mediciones que representan direcciones. En los últimos años se ha vuelto a tener un desarrollo importante en las propuestas de métodos estadísticos para analizar tales datos. Lo anterior, con énfasis en problemas que van desde métodos gráficos adecuados para desplegar observaciones de esta clase, hasta el planteamiento de modelos estadísticos que expliquen relaciones de dependencia. No sin mencionar, los trabajos enfocados en ofrecer metodologías para realizar inferencias sobre los parámetros involucrados en los modelos probabilísticos utilizados para describir este tipo de datos direccionales. Ver, por ejemplo, Fisher (1993), Mardia y Jupp (2000), Artes y Jørgensen (2000), Artes *et al.* (2000), D'Elia *et al.* (2001), Jammalamadaka y SenGupta (2001), Downs y Mardia (2002), Nuñez-Antonio y Gutiérrez-Peña (2005a y 2005b), George y Ghosh (2006), Arnold y SenGupta (2006), Jammalamadaka y Lund (2006), Kume y Walker (2006), Song (2007) y Kume y Walker (2009).

El interés en desarrollar técnicas para analizar datos direccionales se remonta a la época en la que Gauss desarrolló la teoría de errores para analizar ciertas medidas direccionales en astronomía. Es un accidente histórico que los errores involucrados fueran suficientemente pequeños para que Gauss empleara una aproximación lineal y, como una consecuencia, que desarrollara una teoría lineal en lugar de una teoría direccional de los errores.

Los datos direccionales tienen que ver con observaciones que son vectores unitarios en el espacio  $q$ -dimensional. Los datos direccionales en el plano 2-dimensional se denominan *datos circulares* y, las direcciones en el plano 3-dimensional se denominan *datos esféricos*. Así, los espacios muestrales más comunes son el círculo unitario o la esfera unitaria. Por lo anterior, el empleo de métodos lineales univariados o multivariados estándar para el análisis de datos direccionales puede no ser adecuado, por lo que se requiere de métodos estadísticos especiales que tomen en cuenta la estructura topológica de dichos espacios muestrales.

Históricamente, se han propuesto modelos para describir el comportamiento probabilístico de datos de tipo direccional. Sin pérdida de generalidad, estos modelos se pueden agrupar en tres grandes categorías: *modelos generados por proyecciones*, dentro de los cuales la distribución más representativa es la distribución Normal proyectada; *modelos wrapped o “envueltos”*, que incluyen a la Normal envuelta, la Cauchy envuelta y la Poisson envuelta, por citar algunos, y *modelos tipo von Mises-Fisher* cuya distribución principal es la distribución von Mises-Fisher. La distribución de von Mises-Fisher para datos circulares –direcciones sobre el círculo unitario– es conocida como distribución von Mises y es una de las distribuciones más relevantes en el análisis inferencial para

datos circulares. En el caso de datos esféricos –direcciones en la esfera unitaria– la distribución von Mises-Fisher es conocida como distribución Fisher.

Los datos direccionales pueden aparecer en modelos estadísticos como la variable de respuesta, por lo que es de interés el estudio de este tipo de modelos. Sin embargo, los modelos lineales propuestos en la literatura para una respuesta direccional sufren de ciertos problemas que los vuelven difíciles de aplicar en el análisis de tales datos. Esta dificultad para implementar los modelos anteriores es relevante cuando se contrasta con la metodología de los modelos lineales generalizados para el análisis de una respuesta escalar.

Aunque algunas definiciones y resultados mostrados en este trabajo de investigación están asociados a datos direccionales en general, las propuestas y resultados metodológicos obtenidos son dirigidos básicamente al caso de datos circulares.

En este proyecto se pretende implementar el análisis Bayesiano de un modelo de regresión para datos circulares basado en la distribución Normal proyectada, y proponer una metodología que permita modelar estudios longitudinales de series cortas de observaciones donde la variable de respuesta sea de tipo angular.

La estructura de esta tesis es la siguiente. Debido a que la mayoría de estadísticos y matemáticos no están familiarizados con datos de tipo direccional, en el Capítulo 1 se ofrece una introducción al análisis de datos direccionales en general y en particular a los datos circulares. Se presentan algunos modelos de probabilidad relevantes para el estudio de datos direccionales y específicamente se introduce el modelo Normal proyectado. Finalmente, se describe un panorama general sobre los modelos lineales propuestos en la literatura para una respuesta circular.

Los modelos lineales propuestos en este trabajo se basan, desde el punto de vista probabilístico, fundamentalmente en una familia particular que se obtiene al proyectar radialmente una distribución normal bivariada. Por la relevancia que tiene esta distribución en los trabajos realizados en esa tesis, en el Capítulo 2, se define ampliamente la distribución Normal proyectada  $q$ -variada. Se muestra una forma de derivar esta distribución y se analiza el caso bivariado. Adicionalmente, se obtienen a detalle algunos resultados asociados a la distribución normal bivariada bajo proyección, los cuales son relevantes en todo el análisis posterior.

En el Capítulo 3, se desarrolla una propuesta de análisis Bayesiano de un modelo de regresión para datos circulares, basado en la distribución Normal proyectada. Además, se analiza el problema de datos faltantes en la variable respuesta, así como la implementación de un criterio predictivo para la selección de modelos. Se exhibe la forma de llevar a cabo inferencias basadas en métodos de Monte Carlo vía Cadenas de Markov. Las metodologías propuestas se ejemplifican a través de la implementación de diversos ejemplos asociados a conjuntos de datos tanto simulados como reales y datos previamente analizados en la literatura. El Capítulo 3 concluye con una breve discusión sobre los resultados más relevantes obtenidos.

En el Capítulo 4 se propone y analiza un modelo para datos longitudinales circulares. La propuesta se basa en un modelo normal bivariado proyectado donde cada una de sus componentes se especifica a través de modelos lineales de efectos mixtos. Debido a que los modelos mixtos juegan un papel relevante en la estructura del modelo longitudinal propuesto, en este capítulo se presenta un panorama general sobre los modelos de efectos mixtos, resaltando sus ventajas en la modelación de datos longitudinales para una

respuesta escalar. Se ilustra la metodología propuesta con la implementación de algunos ejemplos y se da una discusión asociada a los resultados más importantes conseguidos.

Aunque los Capítulos 3 y 4 finalizan con sus discusiones respectivas, el Capítulo 5 ofrece las conclusiones generales sobre este trabajo de investigación, así como algunas reflexiones finales sobre los retos que implica el tratar de modelar relaciones de dependencia, como son los modelos de regresión y de datos longitudinales, cuando se tiene una variable respuesta de tipo direccional.

Debido a su extensión, el Apéndice *A* contiene las demostraciones de algunas proposiciones incluidas en los diversos capítulos de este trabajo. En el Apéndice *B* se muestra el código de los distintos programas utilizados para la implementación de la metodología propuesta. Finalmente, el Apéndice *C* contiene las bases, y en su caso la estructura, de los datos asociadas a los ejemplos desarrollados en esta tesis.

# Capítulo 1

## Preliminares

### 1.1 Naturaleza de los datos direccionales

Los datos direccionales aparecen en varias áreas de manera natural y son especialmente comunes en las ciencias biológicas, meteorológicas y ecológicas. Algunas aplicaciones se pueden encontrar, por ejemplo, en el análisis de direcciones de viento, direcciones de migración de aves, propagación de fisuras en concreto y otros materiales, orientación de depósitos geológicos, análisis de datos composicionales, análisis de datos axiales, etc. Para una revisión detallada sobre el tema, el lector se puede dirigir a Fisher *et al.* (1987), Mardia (1972, 1975), Fisher (1993) y más recientemente a Mardia y Jupp (2000), y Jammalamadaka y SenGupta (2001). Arnold y SenGupta (2006) presentan una revisión de las aplicaciones del análisis de datos direccionales en las ciencias ecológicas y del medio ambiente.

La literatura Bayesiana es poco extensa y en cierto sentido, al igual que la perspectiva clásica, no es del todo adecuada. Lo anterior debido a las dificultades que representa trabajar con distribuciones de probabilidad asociadas con el análisis de datos direccionales.

En el caso general, los datos direccionales se pueden representar como puntos sobre la esfera unitaria  $\mathbb{S}^q = \{\mathbf{u} \in \mathbb{R}^q : \mathbf{u}'\mathbf{u} = 1\}$ . Cuando  $q = 2$  se pueden emplear coordenadas polares definidas por

$$\mathbf{u} = (\cos \theta, \text{sen} \theta)'$$

para especificar direcciones en el plano. En el caso  $q = 3$  se pueden usar coordenadas polares esféricas definidas por

$$\mathbf{u} = (\cos \theta, \text{sen} \theta \cos \phi, \text{sen} \theta \text{sen} \phi)'$$

para especificar datos esféricos.

Desde un punto de vista teórico existen tres enfoques para el análisis de datos direccionales, los cuales se denominan enfoque *embedding*, enfoque *wrapping* y enfoque *intrínseco*. En el enfoque *embedding*, la esfera  $\mathbb{S}^q$  es reconocida como un subconjunto de  $\mathbb{R}^q$ . En el enfoque *wrapping*, los vectores tangentes  $\mathbf{u}$  a la esfera en  $\boldsymbol{\mu}$  son “envueltos” sobre la esfera por medio del mapeo

$$\mathbf{u} \mapsto (\text{sen} \|\mathbf{u}\|)\boldsymbol{\mu} + (\cos \|\mathbf{u}\|)\mathbf{u},$$

donde  $\mathbf{u}'\boldsymbol{\mu} = 0$ . En el caso circular,  $q = 2$ , el mapeo anterior equivale esencialmente a considerar los valores módulo  $2\pi$ . En el enfoque *intrínseco*, la esfera es reconocida propiamente como un subespacio en su propio contexto.

### 1.1.1 Medidas descriptivas

Sean  $\mathbf{u}_1, \dots, \mathbf{u}_n$  puntos sobre la esfera unitaria  $\mathbb{S}^q$ . Entonces la localización de estos datos puede ser resumida a través de su media muestral en  $\mathbb{R}^q$ , la cual está dada por

$$\bar{\mathbf{u}} = \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i$$

Una forma útil de expresar el vector  $\bar{\mathbf{u}}$  es a través de su forma polar, la cual está dada por

$$\bar{\mathbf{u}} = \bar{R}\bar{\mathbf{u}}_0, \quad (1.1)$$

donde  $\bar{\mathbf{u}}_0$  es un vector unitario y  $\bar{R} \geq 0$ , por lo que

$$\bar{R} = \|\bar{\mathbf{u}}\|$$

y

$$\bar{\mathbf{u}}_0 = \|\bar{\mathbf{u}}\|^{-1}\bar{\mathbf{u}}.$$

El vector unitario  $\bar{\mathbf{u}}_0$  es denominado la *dirección media muestral* y  $\bar{R}$  es denominada la *longitud de la resultante promedio*. Si los puntos  $\mathbf{u}_1, \dots, \mathbf{u}_n$  tienen la misma masa, entonces su centro de masa es  $\bar{\mathbf{u}}$ , la cual tiene dirección  $\bar{\mathbf{u}}_0$  y distancia al origen  $\bar{R}$ .

De manera análoga, para un vector aleatorio unitario  $\mathbf{u} = (u_1, \dots, u_q)'$  se define la *longitud de la resultante media poblacional*  $\rho$  como

$$\rho = \sum_{j=1}^q [E(u_j^2)]^{1/2} = [E(\mathbf{u})'E(\mathbf{u})]^{1/2}.$$

Cuando  $\rho > 0$  la *dirección media poblacional* se define como

$$\boldsymbol{\mu} = \rho^{-1}E(\mathbf{u}).$$

La dirección media posee la siguiente propiedad, análoga a aquella sobre la recta real.

**Proposición 1.1.1** *Sea  $\mathbf{A}$  una transformación ortogonal, es decir, una rotación o una reflexión. Entonces, la dirección media de  $\mathbf{A}\mathbf{u}_1, \dots, \mathbf{A}\mathbf{u}_n$  es  $\mathbf{A}\bar{\mathbf{u}}_0$  y la longitud media resultante está dada por  $\bar{R}$ . Así,  $\bar{R}$  es invariante bajo rotaciones y reflexiones.*

Una medida importante de dispersión es la matriz  $\bar{\mathbf{T}}$  definida como

$$\bar{\mathbf{T}} = \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i \mathbf{u}_i' \quad (1.2)$$



Los lectores familiarizados con mecánica pueden encontrar útil la interpretación de  $\bar{\mathbf{T}}$  como el tensor de inercia alrededor del origen de partículas con pesos  $n^{-1}$  en cada uno de los puntos  $\mathbf{u}_1, \dots, \mathbf{u}_n$ .

Así, si  $\mathbf{S}$  denota la varianza muestral de los vectores  $\mathbf{u}_1, \dots, \mathbf{u}_n$ , dada por

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{u}_i - \bar{\mathbf{u}}_0)(\mathbf{u}_i - \bar{\mathbf{u}}_0)',$$

entonces

$$\bar{\mathbf{T}} = \frac{n-1}{n} \mathbf{S} + \bar{\mathbf{u}}_0 \bar{\mathbf{u}}_0'.$$

De (1.2) se puede notar que la restricción  $\mathbf{u}_i' \mathbf{u}_i = 1$  produce  $\text{tr}(\bar{\mathbf{T}}) = 1$  y así

$$\frac{n-1}{n} \text{tr}(\mathbf{S}) + \bar{R}^2 = 1. \quad (1.3)$$

En vista de (1.3)  $\bar{R}$  juega un papel importante como medida de variación. Se puede notar que  $\bar{R} \simeq 0$  cuando  $\mathbf{u}_1, \dots, \mathbf{u}_n$  estén ampliamente dispersos y  $\bar{R} \simeq 1$  cuando  $\mathbf{u}_1, \dots, \mathbf{u}_n$  estén fuertemente concentrados. Así,  $\bar{R}$  resulta ser una medida de concentración alrededor de la dirección media.

En una manera similar a la deducción de (1.3), la restricción de que un vector aleatorio pertenezca a  $\mathbb{S}^q$  produce la relación

$$\text{tr}(\boldsymbol{\Sigma}) + \rho^2 = 1,$$

entre  $\rho$  y la matriz de varianzas-covarianzas  $\boldsymbol{\Sigma}$  de  $\mathbf{u}$ . Así, la naturaleza *singular* del espacio muestral para datos direccionales significa que existe una conexión entre la media  $E(\mathbf{u})$  y la matriz de varianzas-covarianzas del vector aleatorio  $\mathbf{u}$ . Lo anterior, en contraste a los usuales casos irrestrictos de los espacios muestrales en  $\mathbb{R}^q$ .

## 1.2 El caso de datos circulares

Los datos circulares son una clase particular de datos direccionales. Específicamente, los datos circulares son direcciones en dos dimensiones. Dos de las formas más relevantes en las que surgen los datos circulares están asociadas a dos de los instrumentos más importantes de medición circular: *la brújula* y *el reloj*. Observaciones típicas medidas por una brújula incluyen direcciones de viento y direcciones de migración de aves. Como ejemplo de observaciones típicas medidas con un reloj se pueden mencionar datos asociados a tiempos de arribo a una unidad de terapia intensiva de un hospital, sobre un reloj de 24 hrs. Datos similares aparecen como tiempos a lo largo del año (o tiempos en meses) en los que se presenta algún evento de interés.

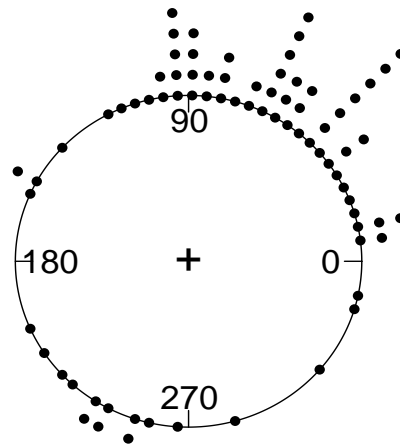


Figura 1.1: *Diagrama circular para los datos de tortugas.*

La representación gráfica de este tipo de datos es a través de puntos sobre la cir-

cunferencia del círculo unitario. La Figura 1.1 muestra las direcciones que tomaron 76 tortugas hembras después de depositar sus huevos sobre la playa. Los datos fueron tomados de la Tabla 1.5 de Mardia y Jupp (2000) y se pueden revisar en el Apéndice *C* de este trabajo. Existe toda una discusión sobre la construcción de métodos adecuados de análisis exploratorio para datos circulares (ver, por ejemplo, Mardia y Jupp, 2000 y Fisher, 1993). Dentro de estos métodos gráficos se encuentra el *diagrama de rosa*, el cual se puede pensar como la representación análoga del histograma para datos en la recta real. En la Figura 1.2 se muestra el correspondiente diagrama de rosa para los datos de tortugas. A partir de las Figuras 1.1 y 1.2 se podría considerar, por ejemplo, que una distribución bimodal sería adecuada para describir estos datos.

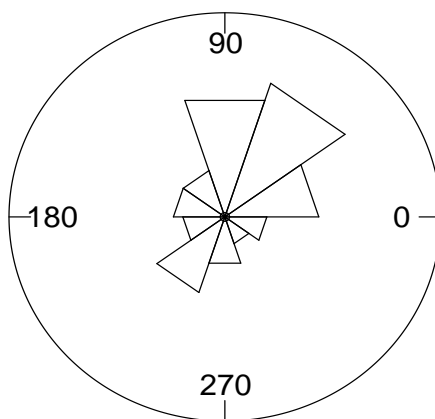


Figura 1.2: *Diagrama de rosa para los datos de tortugas.*

La aplicación de técnicas lineales convencionales puede producir paradojas en el análisis de este tipo de datos, dada la periodicidad inherente del círculo y la diferente

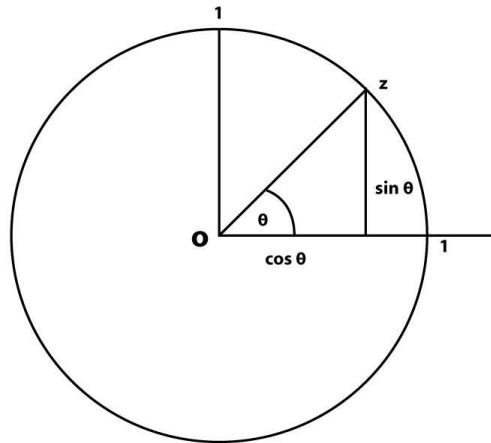


Figura 1.3: Representación de la dirección  $u$  por medio de un ángulo  $\theta$  y de un número complejo  $z$ .

topología del círculo y la línea recta. Un ejemplo ilustrativo es el siguiente: Si se tiene un conjunto de datos conformado por los ángulos 1 y 359, entonces su media aritmética resulta ser 180, la cual como medida descriptiva de este conjunto de datos es engañosa. Sin embargo, resulta más adecuado, y geoméricamente más intuitivo, considerar a la dirección 0 como un mejor valor para representar la localización de este conjunto de datos.

Como una observación final para apreciar la diferente naturaleza de los datos circulares con respecto a los datos sobre la línea real, se puede ver que el círculo es una curva cerrada pero la línea recta no, por lo que se pueden anticipar diferencias entre la teoría estadística sobre la línea y sobre el círculo. Así, es necesario definir, por ejemplo, funciones de distribución de probabilidad y medidas descriptivas numéricas de tal manera que tomen en cuenta la particular topología de estos espacios muestrales.

Los datos circulares son direcciones en el plano y se pueden considerar como vectores

unitarios  $\mathbf{u}$ , o de manera equivalente, como puntos sobre el círculo unitario. Sin embargo, existen otras dos maneras útiles de considerar estos datos: como ángulos o como números complejos unitarios. Una vez seleccionado un sistema de coordenadas ortogonales en el plano (lo cual es equivalente a seleccionar una dirección y una orientación inicial), cada punto  $\mathbf{u}$  sobre el círculo unitario se puede representar por un ángulo  $\theta$  o, en forma equivalente, por un número complejo  $z$ . Esto se relaciona con  $\mathbf{u}$  mediante

$$\mathbf{u} = (\cos \theta, \sin \theta)' \quad \text{y} \quad z = e^{i\theta} = \cos \theta + i \sin \theta,$$

ver Figura 1.3

La representación de datos circulares por medio de ángulos se puede ver como un enfoque *intrínseco*, dado que las direcciones son consideradas como puntos sobre el círculo mismo. Por otro lado, la representación por un número complejo se puede pensar como un enfoque *embedding*, ya que las direcciones se consideran como puntos especiales en el plano.

## 1.2.1 Medidas descriptivas muestrales

### Medidas de localización

**Definición 1.2.1** Sean  $\theta_1, \dots, \theta_n$  los ángulos asociados a los vectores aleatorios unitarios  $\mathbf{u}_1, \dots, \mathbf{u}_n$ . La **dirección media**  $\bar{\theta}$  de  $\theta_1, \dots, \theta_n$  se define como la dirección de la resultante  $\mathbf{u}_1 + \dots + \mathbf{u}_n$  de los vectores  $\mathbf{u}_1, \dots, \mathbf{u}_n$ . La cual, es también la dirección del centro de masa,  $\bar{\mathbf{u}}$ , de  $\mathbf{u}_1, \dots, \mathbf{u}_n$ .

En la Proposición (1.2.1) se muestra que  $\bar{\theta}$  así definida minimiza cierta medida de dispersión que da lugar a la denominada varianza circular muestral.

Como las coordenadas Cartesianas de  $\mathbf{u}_i$  son  $(\cos \theta_i, \text{sen} \theta_i)$  para  $i = 1, \dots, n$ , entonces las coordenadas Cartesianas del centro de masa son  $(\bar{C}, \bar{S})$ , donde

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i \quad \bar{S} = \frac{1}{n} \sum_{i=1}^n \text{sen} \theta_i. \quad (1.4)$$

Así,  $\bar{\theta}$  es la solución de las ecuaciones

$$\bar{C} = \bar{R} \cos \bar{\theta}, \quad \bar{S} = \bar{R} \text{sen} \bar{\theta}, \quad (1.5)$$

donde

$$\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2}.$$

Se puede notar que  $\bar{\theta}$  no está definida cuando  $\bar{R} = 0$ . Cuando  $\bar{R} > 0$ ,  $\bar{\theta}$  está dada explícitamente por

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S} / \bar{C}) & \text{si } \bar{C} \geq 0 \\ \tan^{-1}(\bar{S} / \bar{C}) + \pi & \text{si } \bar{C} < 0, \end{cases}$$

donde  $\tan^{-1}(\cdot)$  toma valores en  $[-\pi/2, \pi/2]$ . Hay que notar que en el contexto de estadística circular  $\bar{\theta}$  no significa  $\sum \theta_i/n$ , la cual no está bien definida. En el caso general,  $q$ -dimensional, (1.4) y (1.5) son equivalentes a (1.1).

**Definición 1.2.2** Sea  $\mathbf{u}_1, \dots, \mathbf{u}_n$  vectores unitarios correspondientes a los ángulos  $\theta_1, \dots, \theta_n$ . La **longitud de la resultante promedio**,  $\bar{R}$ , de los vectores  $\mathbf{u}_1, \dots, \mathbf{u}_n$ , es decir, la longitud del centro de masa  $\bar{\mathbf{u}}$ , está dada por

$$\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2},$$

donde  $\bar{C}$  y  $\bar{S}$  están definidas por (1.4).

## Medidas de concentración y dispersión

La longitud de la resultante promedio  $\bar{R}$  fue introducida anteriormente como la longitud del centro de masa  $\bar{\mathbf{u}}$ . Dado que los vectores  $\mathbf{u}_1, \dots, \mathbf{u}_n$  son vectores unitarios, entonces  $0 \leq \bar{R} \leq 1$ . Si las direcciones  $\theta_1, \dots, \theta_n$  están muy agrupadas,  $\bar{R}$  será cercana a 1 y si están muy dispersas entonces será cercana a 0. Así,  $\bar{R}$  resulta ser una medida de la *concentración* de un conjunto de ángulos. Hay que notar que cualquier conjunto de datos de la forma  $\theta_1, \dots, \theta_n, \theta_1 + \pi, \dots, \theta_n + \pi$  arrojará un valor de  $\bar{R} = 0$ , por lo que si  $\bar{R} \approx 0$  esto no implica que las direcciones estén dispersas alrededor de todo el círculo de manera uniforme.

En estadística direccional, para propósitos descriptivos e inferenciales,  $\bar{R}$  es la medida más relevante de concentración que cualquier otra medida de dispersión propuesta para datos circulares, ver por ejemplo Mardia y Jupp (2000) y Jammalamadaka y SenGupta (2001). Sin embargo, con propósitos de comparación con el análisis para datos sobre la recta a veces es útil considerar medidas de dispersión para datos circulares. A continuación se presentan dos de estas medidas. Para una revisión completa sobre medidas de dispersión asociadas a datos circulares el lector puede revisar, por ejemplo, Fisher (1993) y Mardia y Jupp (2000).

**Definición 1.2.3** Sea  $\theta_1, \dots, \theta_n$  un conjunto de ángulos. La *varianza circular* se define como

$$V = 1 - \bar{R}.$$

Por otro lado, una medida útil de la distancia entre dos ángulos  $\theta$  y  $\xi$  es

$$1 - \cos(\theta - \xi).$$

Así, una forma de medir la dispersión de un conjunto de ángulos  $\{\theta_1, \dots, \theta_n\}$  alrededor de un ángulo  $\alpha$  es a través de la cantidad

$$D(\alpha) = \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \alpha)\}. \quad (1.6)$$

**Proposición 1.2.1** *Si  $\theta_1, \dots, \theta_n$  son un conjunto de ángulos, la medida de dispersión  $D(\alpha)$ , definida en (1.6) alcanza su valor mínimo en  $\bar{\theta}$  y este valor mínimo es la varianza circular muestral,  $V$ .*

DEMOSTRACIÓN:

$$\begin{aligned} D(\alpha) &= \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \alpha)\} \\ &= 1 - \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \alpha) \end{aligned} \quad (1.7)$$

$\implies$

$$n D(\alpha) = n - n\bar{R} + (n\bar{R} - \sum_{i=1}^n \cos(\theta_i - \alpha)) \quad (1.8)$$

Por otro lado,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \bar{\theta})\} &= \frac{1}{n} \sum_{i=1}^n [\cos \theta_i \cos \bar{\theta} + \text{sen} \theta_i \text{sen} \bar{\theta}] \\ &= \bar{C} \cos \bar{\theta} + \bar{S} \text{sen} \bar{\theta} \\ &= \bar{R} \cos^2 \bar{\theta} + \bar{R} \text{sen}^2 \bar{\theta} \\ &= \bar{R}. \end{aligned} \quad (1.9)$$

De las ecuaciones (1.6) y (1.9) se tiene que

$$D(\bar{\theta}) = \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \bar{\theta})\} = 1 - \bar{R} = V.$$



Así, a partir de las ecuaciones (1.8) y (1.9) resulta

$$\begin{aligned}
n D(\alpha) &= (n - n\bar{R}) + n\bar{R} - \sum_{i=1}^n \cos(\theta_i - \alpha) \\
&= n(1 - \bar{R}) + \sum_{i=1}^n \{1 - \cos(\theta_i - \bar{\theta})\} - \sum_{i=1}^n \cos(\theta_i - \alpha) \\
&= nV + \cos \bar{\theta} \sum_{i=1}^n \cos \theta_i + \operatorname{sen} \bar{\theta} \sum_{i=1}^n \operatorname{sen} \theta_i - \cos \alpha \sum_{i=1}^n \cos \theta_i - \operatorname{sen} \alpha \sum_{i=1}^n \operatorname{sen} \theta_i \\
&= nV + n\bar{R} \cos^2 \bar{\theta} + n\bar{R} \operatorname{sen}^2 \bar{\theta} - n\bar{R} \cos \alpha \cos \bar{\theta} - n\bar{R} \operatorname{sen} \alpha \operatorname{sen} \bar{\theta} \\
&= nV + n\bar{R}[1 - \cos(\bar{\theta} - \alpha)] \\
&= nV + n2\bar{R} \left\{ \operatorname{sen} \left( \frac{\bar{\theta} - \alpha}{2} \right) \right\}^2.
\end{aligned}$$

De lo anterior, se obtiene

$$D(\alpha) = V + 2\bar{R} \left\{ \operatorname{sen} \left( \frac{\bar{\theta} - \alpha}{2} \right) \right\}^2,$$

por lo que  $D(\alpha)$  alcanza su valor mínimo en  $\bar{\theta}$ , y este es  $V$ .

□

La proposición anterior es análoga al resultado familiar que se tiene para observaciones  $x_1, \dots, x_n$  en la recta real: la cantidad

$$\frac{1}{n} \sum_{i=1}^n (x_i - a)^2$$

se minimiza cuando  $a = \bar{x}$  y el valor mínimo es la varianza muestral (con divisor  $n$ ).

## Momentos muestrales

En la sección 1.2.1 se puede apreciar que los momentos muestrales

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i \qquad \bar{S} = \frac{1}{n} \sum_{i=1}^n \operatorname{sen} \theta_i,$$

juegan un papel importante en la definición de la dirección media y la varianza muestral. En estadística direccional, estos momentos combinados aparecen en el primer momento trigonométrico alrededor de la dirección cero.

**Definición 1.2.4** *El  $p$ -ésimo momento trigonométrico, de un conjunto de ángulos  $\theta_1, \dots, \theta_n$ , alrededor de la dirección cero se define como*

$$m'_p = a_p + ib_p \quad \forall p = 1, 2, \dots$$

donde

$$a_p = \frac{1}{n} \sum_{i=1}^n \cos p\theta_i \quad b_p = \frac{1}{n} \sum_{i=1}^n \operatorname{sen} p\theta_i.$$

Entonces,

$$m'_p = \bar{R}_p e^{i\bar{\theta}_p} \tag{1.10}$$

donde  $\bar{R}_p$  y  $\bar{\theta}_p$  denotan la dirección media muestral y la longitud de la resultante promedio de los ángulos  $p\theta_1, \dots, p\theta_n$ .

De la definición anterior se puede notar que

$$m'_1 = \bar{C} + i\bar{S},$$

y

$$m'_1 = \bar{R} e^{i\bar{\theta}}.$$

De manera análoga se pueden definir los momentos trigonométricos alrededor de la dirección media.

**Definición 1.2.5** *El  $p$ -ésimo momento trigonométrico alrededor de la dirección media se define como*

$$m_p = \bar{a}_p + i\bar{b}_p \quad \forall p = 1, 2, \dots$$

donde

$$\bar{a}_p = \frac{1}{n} \sum_{i=1}^n \cos p(\theta_i - \bar{\theta}) \quad \bar{b}_p = \frac{1}{n} \sum_{i=1}^n \operatorname{sen} p(\theta_i - \bar{\theta}).$$

De manera particular, se puede mostrar que

$$m_1 = \bar{R}.$$

Las versiones poblacionales de los momentos trigonométricos anteriores juegan un papel relevante cuando se trabaja en teoría de distribuciones. Estas versiones poblacionales se definen en la siguiente sección.

## 1.2.2 La función característica

**Definición 1.2.6** La **función característica** de un ángulo aleatorio  $\Theta$ , con respecto a la función de distribución  $F$ , se define como la doble sucesión infinita de números complejos  $\{\psi_p : p = 0, \pm 1, \dots\}$  dada por

$$\psi_p = E(e^{ip\theta}) = \int_0^{2\pi} e^{ip\theta} dF(\theta), \quad p = 0, \pm 1, \dots$$

donde  $F(\theta)$  es la función de distribución, definida en el círculo unitario  $\mathbb{S}^2$ , del ángulo aleatorio  $\Theta$ .

Hay que notar que  $\psi_p$  se puede escribir como

$$\psi_p = \alpha_p + i\beta_p,$$

donde

$$\alpha_p = E(\cos p\theta) = \int_0^{2\pi} \cos p\theta dF(\theta)$$

y

$$\beta_p = E(\operatorname{sen} p\theta) = \int_0^{2\pi} \operatorname{sen} p\theta dF(\theta).$$

Así,  $\psi_p$ ,  $\alpha_p$  y  $\beta_p$  son las versiones poblacionales de los momentos trigonométricos muestrales  $m'_p$ ,  $a_p$  y  $b_p$ , definidos en la sección 1.2.1.

## Momentos poblacionales

**Definición 1.2.7** *El  $p$ -ésimo momento poblacional trigonométrico alrededor de la dirección cero de un ángulo aleatorio  $\Theta$  se define como*

$$\alpha_p = E(\cos p\theta), \quad \beta_p = E(\operatorname{sen} p\theta), \quad p = 0, \pm 1, \pm 2, \dots$$

Se debe notar que la sucesión  $\{(\alpha_p, \beta_p) : p = 0, \pm 1, \dots\}$  de momentos trigonométricos es equivalente a la función característica de  $\Theta$ ,  $\psi_p$ .

Así, para  $p \geq 0$

$$\psi_p = \rho_p e^{i \mu_p},$$

con  $\rho_p \geq 0$ , resulta ser la versión poblacional del momento muestral  $m'_p$  (ver ecuación 1.10)

Como el caso  $p = 1$  es el utilizado más frecuentemente, se acostumbra la siguiente notación

$$\alpha_1 = \alpha, \quad \beta_1 = \beta, \quad \rho_1 = \rho, \quad \mu_1 = \mu.$$

**Definición 1.2.8** *El  $p$ -ésimo momento poblacional trigonométrico alrededor de la dirección media se define como*

$$\bar{\alpha}_p + i \bar{\beta}_p$$

donde

$$\bar{\alpha}_p = E(\cos p(\theta - \mu)), \quad \bar{\beta}_p = E(\operatorname{sen} p(\theta - \mu)).$$

En el caso  $p = 1$ , se tiene que

$$\psi_1 = \rho e^{i\mu}$$

donde  $\mu$  se denomina la *dirección media* y  $\rho$  es la *longitud de la resultante media*. Hay que notar que  $\mu$  y  $\rho$  son las versiones poblacionales de  $\bar{\theta}$  y  $\bar{R}$ , respectivamente.

### 1.3 Modelos para datos direccionales

Por su construcción, los modelos de probabilidad para datos direccionales se pueden clasificar en tres grandes categorías. Los *modelos wrapped*, los *modelos tipo von Mises-Fisher* y los *modelos generados por proyecciones*.

Los modelos *wrapped* para datos circulares se obtienen al “envolver” alguna distribución definida originalmente sobre la recta real alrededor de la circunferencia del círculo unitario. Es decir, si  $Y$  es una variable aleatoria definida sobre la recta real, la correspondiente variable  $Y_w$  de la distribución *wrapped* está dada por

$$Y_w = Y(\text{mod } 2\pi).$$

Si el círculo unitario se identifica con el conjunto de números complejos con módulo uno, entonces el mapeo *wrapped*  $Y \mapsto Y_w$  se puede escribir como

$$Y \mapsto e^{2\pi i Y}.$$

Algunos de los modelos probabilísticos más relevantes dentro de esta familia son la distribución Normal *wrapped*, la distribución Cauchy *wrapped* y la distribución Poisson *wrapped*.

Los modelos de probabilidad tipo von Mises-Fisher, son modelos cuya construcción considera de manera natural la topología del correspondiente espacio muestral. Modelos relevantes dentro de esta familia son el modelo lattice, el modelo uniforme y el modelo von Mises-Fisher. Por su gran relevancia, en la modelación de datos direccionales, en la siguiente sección se describen algunas características del modelo von Mises-Fisher.

Los modelos generados por proyecciones se obtienen proyectando radialmente distribuciones definidas originalmente en  $\mathbb{R}^q$ . Es decir, dado un vector aleatorio  $\mathbf{Y}$  en  $\mathbb{R}^q$

tal que  $Pr(\mathbf{Y} = \mathbf{0}) = \mathbf{0}$ , entonces la correspondiente distribución proyectada sobre la esfera unitaria  $\mathbb{S}^q$  es la asociada al vector unitario  $\|\mathbf{Y}\|^{-1}\mathbf{Y}$ . Un caso particularmente importante es cuando  $\mathbf{Y}$  sigue una distribución normal  $q$ -variada con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\boldsymbol{\Lambda}$ ,  $N_q(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ . En este caso, se dice que el vector aleatorio  $\|\mathbf{Y}\|^{-1}\mathbf{Y}$  tiene una distribución *Normal proyectada  $q$ -variada*, denotada por  $PN_q(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ . Otros nombres para esta distribución son distribución Gaussiana angular u *offset*. En este trabajo se denotará por  $PN(\boldsymbol{\mu}, \boldsymbol{\Lambda})$  a la distribución generada al proyectar una densidad Normal bivariada y, se denominará simplemente distribución *Normal proyectada*. El modelo Normal proyectado de ninguna manera requiere que las direcciones observadas sean en realidad proyecciones de una respuesta multivariada. El modelo simplemente especifica una forma particular de la distribución de las direcciones.

Aunque anteriormente se ha considerado al círculo como un subconjunto de  $\mathbb{R}^2$  para propósitos de inferencia estadística (Jupp y Mardia, 1989), la distribución Normal bivariada proyectada ha sido poco trabajada y en consecuencia poco explotada, en parte, debido a su aparente complejidad (ver Jammalamadaka y SenGupta, 2001, pág. 43). Una notable excepción es el trabajo de Presnell *et al.* (1998). Ellos analizan una versión de este modelo para el caso de datos circulares desde un punto de vista clásico y ofrecen estimadores de máxima verosimilitud de los parámetros usando métodos iterativos como el de Newton-Raphson y el algoritmo EM. Por otro lado, Nuñez-Antonio y Gutiérrez-Peña (2005a) proponen un análisis Bayesiano completo de una versión del modelo Normal proyectado  $q$ -variado para datos circulares y esféricos.

Los modelos de dependencia propuestos en este trabajo tienen como fundamento a la distribución Normal proyectada. Por lo anterior, en el Capítulo 2 esta distribución es revisada ampliamente.

### 1.3.1 La distribución de von Mises-Fisher

Se dice que un vector aleatorio unitario  $q$ -dimensional  $\mathbf{Y}$  tiene una distribución von Mises-Fisher, con vector de dirección media  $\boldsymbol{\mu}$  y parámetro de concentración  $\kappa$ , si su función de densidad de probabilidad está dada por

$$f(\mathbf{y}|\boldsymbol{\mu}, \kappa) = \left(\frac{\kappa}{2}\right)^{q/2-1} \frac{1}{\Gamma(q/2)I_{q/2-1}(\kappa)} \exp\{\kappa\boldsymbol{\mu}'\mathbf{y}\} I_{\mathbb{S}^q}(\mathbf{y}) I_{\mathbb{S}^q}(\boldsymbol{\mu}) I_{(0,\infty)}(\kappa),$$

donde  $\mathbb{S}^q = \{\mathbf{y} \in \mathbb{R}^q : \mathbf{y}'\mathbf{y} = 1\}$ ,  $I_b(\cdot)$  denota la función de Bessel modificada de primer tipo de orden  $b$  y  $1_B(\cdot)$  es la función indicadora sobre el conjunto  $B$ . La distribución von Mises-Fisher para datos circulares,  $q = 2$ , es conocida como distribución von Mises. Esta es la distribución más importante en el análisis de datos circulares. Además, es el análogo natural sobre el círculo de la distribución normal sobre la línea real (ver, por ejemplo, Mardia y Jupp, 2000). De manera particular, se dice que un ángulo aleatorio  $\Theta$  sigue una distribución von Mises con dirección media  $\mu$  y parámetro de concentración  $\kappa$ , si su función de densidad está dada por

$$M(\theta|\mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta-\mu)}.$$

En el caso de datos esféricos,  $q = 3$ , la distribución von Mises-Fisher es conocida como distribución Fisher.

El intento más reciente de un análisis Bayesiano completo para la distribución von Mises-Fisher se debe a Nuñez-Antonio y Gutiérrez-Peña (2005b). Ellos consideran todos

los parámetros del modelo desconocidos y explotan el hecho de que esta distribución pertenece a una familia exponencial regular para generar muestras de la distribución final conjunta de  $(\boldsymbol{\mu}, \kappa)$  vía un esquema de muestreo-remuestreo por importancia.

Anterior al trabajo de Nuñez-Antonio y Gutiérrez-Peña (2005b), Damien y Walker (1999) desarrollan un muestreo de Gibbs basado en el uso de variables auxiliares. Sin embargo, la implementación de su metodología no es muy eficiente en algunos casos. Específicamente, la autocorrelación de las muestras de la distribución final es demasiado alta para valores grandes del parámetro de concentración ( $\kappa \geq 7$ , digamos), lo cual resulta en una convergencia lenta del algoritmo propuesto. Bagchi y Kadane (1991) desarrollaron aproximaciones de Laplace para las distribuciones finales asociadas a la von Mises, pero sólo producen estimadores Bayesianos (medias finales) para el coseno de la dirección bajo el supuesto de que el parámetro de concentración es conocido. Guttorp y Lockhart (1988) consideran el caso de ambos parámetros desconocidos, pero se ven obligados a usar la moda final como un estimador para el parámetro de concentración. El análisis Bayesiano de Bagchi (1987) y Bagchi y Guttman (1988) también restringe su atención a la distribución von Mises. Mardia y El-Atoum (1976) analizaron el modelo von Mises-Fisher, pero asumen el parámetro de concentración conocido y sólo ofrecen estimadores puntuales para el parámetro de la dirección media.

### 1.3.2 Relación entre algunas distribuciones

En esta parte se presentan relaciones entre algunos modelos probabilísticos para datos circulares.

En Mardia y Jupp (2000) se muestra que una distribución von Mises se puede aproximar por una distribución normal *wrapped* cuando  $\kappa \rightarrow \infty$ . Esta aproximación se



basa en la comparación de los primeros momentos trigonométricos. Por su parte Kent (1978) obtiene que

$$f_{vM}(\theta|\mu, \kappa) - f_{nw}(\theta|\mu, A(\kappa)) = O(k^{-1/2}) \quad \kappa \rightarrow \infty. \quad (1.11)$$

donde  $f_{vM}(\theta|\mu, \kappa)$  y  $f_{nw}(\theta|\mu, A(\kappa))$  denotan las densidades de una distribución von Mises y la normal *wrapped* que la aproxima, respectivamente. Por su parte, Stephens (1963) verifica numericamente que la aproximación (1.11) es satisfactoria para valores moderados de  $\kappa$ . En Mardia y Jupp (2000) se señala que la distribución von Mises también se puede aproximar por una distribución Cauchy *wrapped*, con la misma dirección media y la misma longitud de la resultante media.

Adicionalmente, en Fisher (1993) se señala que la distribución normal *wrapped* es una buena alternativa, por su similaridad, a la distribución von Mises para modelar datos unimodales simétricos.

Por su parte, Presnell *et al.* (1998) muestra gráficamente que la densidad de una distribución von Mises y la densidad de una distribución Normal proyectada con matriz de covarianzas la identidad, son muy cercanas. Para comparaciones adicionales en el caso circular y esférico, ver Watson (1983). Finalmente, en Mardia (1972), se exhibe la relación entre una distribución normal bivariada particular bajo proyecciones y una distribución Cauchy *wrapped* (*CW*). Específicamente, se muestra que

$$\Theta \sim PN(\mathbf{0}, \mathbf{\Lambda}) \Rightarrow 2\Theta \sim CW(\mu, \rho),$$

para ciertos valores de  $\mu$  y  $\rho$ .

Por lo expuesto anteriormente, en la práctica es común trabajar con cualquiera de las distribuciones von Mises, normal *wrapped*, Cauchy *wrapped* y Normal proyectada,

dependiendo del objetivo del estudio. Por ejemplo, para propósitos inferenciales, la distribución von Mises es la más utilizada en la literatura.

## 1.4 Modelos lineales para una respuesta circular

La teoría de modelos de dependencia cuando la variable de respuesta es angular no ha sido muy desarrollada a pesar de que sus aplicaciones no son poco comunes en diversas áreas, particularmente en biología, meteorología y ciencias del medio ambiente. Algunos ejemplos incluyen la dependencia de la dirección de movimiento de ciertas especies de animales sobre la distancia desplazada o recorrida, la dependencia de algún tipo de defecto o imperfección plana sobre el desplazamiento, la dependencia de direcciones de viento sobre la velocidad del mismo, etc.

### 1.4.1 Modelos de regresión

En el contexto de regresión el objetivo principal es tratar de modelar la variación de la dirección media de una variable de respuesta circular  $\mathbf{U}$  en términos de una o más variables explicativas o covariables.

Así, si  $(\mathbf{x}_1, \mathbf{u}_1), \dots, (\mathbf{x}_n, \mathbf{u}_n)$  son observaciones independientes donde  $\mathbf{x}_i$  es un vector de covariables observables y  $\mathbf{u}_i$  es la correspondiente respuesta direccional, con dirección media  $\boldsymbol{\eta}_i$  y longitud resultante media  $\rho_i$ , el problema consiste en modelar la dirección media  $\boldsymbol{\eta}_i$  en términos de las covariables  $\mathbf{x}_i$ . En el caso de datos circulares,  $\theta_i$  y  $w_i$  son las representaciones angulares de  $\mathbf{u}_i = (\cos \theta_i, \sin \theta_i)'$  y  $\boldsymbol{\eta}_i = (\cos w_i, \sin w_i)'$ , respectivamente, con  $E(\cos \theta_i | \mathbf{x}_i) = \rho_i \cos w_i$  y  $E(\sin \theta_i | \mathbf{x}_i) = \rho_i \sin w_i$ .

En la literatura se han propuesto varios modelos de regresión para abordar el proble-

ma planteado anteriormente. Entre las propuestas existentes se encuentran los modelos de Gould (1969), Johnson y Wehrly (1978), Fisher y Lee (1992) y de manera más reciente los de Presnell *et al.* (1998), Downs y Mardia (2002) y George y Ghosh (2006). Excepto Presnell *et al.* (1998), todos los modelos anteriores consideran que la variable respuesta  $\Theta_i$  sigue una distribución von Mises, donde la media angular  $w_i$  o el parámetro de concentración  $\kappa_i$  o ambos dependen de las covariables  $\mathbf{x}_i$ . A pesar de las diversas propuestas, los modelos planteados anteriormente presentan dificultades técnicas o limitaciones prácticas que los vuelven poco atractivos en la aplicación general para la modelación de este tipo de relaciones de dependencia.

En este trabajo se propone un análisis Bayesiano de un modelo de regresión, para el caso de una variable respuesta circular, que pretende subsanar algunas de las deficiencias de los modelos existentes en la literatura. El modelo de regresión propuesto se basa en una distribución Normal proyectada. Para realizar inferencias sobre los parámetros del modelo se introducen e implementan procedimientos basados en técnicas de simulación que tienen como base a los Métodos de Monte Carlos vía Cadenas de Markov.

### 1.4.2 Modelos para datos longitudinales

En la actualidad existe un vacío en la literatura que limita el establecimiento de un marco general para el análisis longitudinal de datos direccionales. Particularmente, los datos circulares longitudinales únicamente han sido estudiados en forma clásica usando técnicas semiparamétricas tales como la del esquema de ecuaciones estimadoras generalizadas (ver, Artes y Jørgensen, 2000 y Artes *et al.*, 2000). Dichos procedimientos semiparamétricos sufren de limitaciones que los hacen poco flexibles para realizar inferencias. Estas limitaciones incluyen la dificultad para el ajuste y comparación de

modelos, verosimilitudes multimodales, problemas en el cálculo de estimadores debido a los problemas de convergencia de los métodos iterativos empleados, etc. Lo anterior, debido quizá a que los métodos propuestos anteriormente consideran, esencialmente, las respuestas circulares como escalares subestimando su naturaleza periódica. Adicionalmente, como el enfoque de ecuaciones estimadoras generalizadas fundamentalmente se basa en supuestos sobre segundos momentos, una propuesta de tipo paramétrico, como la de este trabajo de investigación, podría ser más adecuada para el tratamiento de datos faltantes, los cuales son frecuentes en el contexto de datos longitudinales.

En este trabajo de investigación se considera una propuesta paramétrica para el análisis longitudinal de datos circulares que se espera contribuya a la modelación y entendimiento de este tipo de relaciones de dependencia.

## Capítulo 2

# La Distribución Normal Proyectada

Como se mencionó en el Capítulo 1, una forma simple de generar distribuciones de probabilidad sobre la esfera unitaria  $q$ -dimensional es proyectando radialmente distribuciones de probabilidad originalmente definidas sobre el espacio  $q$ -dimensional. Una distribución relevante, dentro de esta familia de modelos, es la distribución Normal proyectada  $q$ -variada. El modelo Normal proyectado  $q$ -variado se obtiene al proyectar radialmente una distribución de probabilidad normal multivariada.

### 2.1 La distribución Normal proyectada $q$ -variada

**Definición 2.1.1** *Se dice que un vector aleatorio unitario  $q$ -dimensional  $\mathbf{U} = \mathbf{Y}/\|\mathbf{Y}\|^{-1}$  tiene una distribución Normal proyectada  $q$ -variada  $PN_q(\boldsymbol{\mu}, \boldsymbol{\Lambda})$  si  $\mathbf{Y}$  tiene una distribución normal  $q$ -variada con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\boldsymbol{\Lambda} = \text{Var}(\mathbf{Y})^{-1}$ .*

Se debe notar que como  $\mathbf{U}$  es un vector unitario  $q$ -dimensional, éste también puede ser definido por  $q - 1$  ángulos. Así, para el caso de datos circulares ( $q = 2$ ), Mardia y Jupp (2000) presentan la siguiente definición.

**Definición 2.1.2** La función de densidad de probabilidad de una distribución Normal proyectada, para un ángulo aleatorio  $\Theta$ , está dada por

$$NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \frac{\varphi(\theta|\mathbf{0}, \boldsymbol{\Lambda}) + |\boldsymbol{\Lambda}|^{-1/2} D(\theta) \Phi(D(\theta)) \phi(|\boldsymbol{\Lambda}|^{-1/2} (\mathbf{u}' \boldsymbol{\Lambda} \mathbf{u})^{-1/2} \boldsymbol{\mu} * \mathbf{u})}{\mathbf{u}' \boldsymbol{\Lambda} \mathbf{u}} 1_{(0, 2\pi)}$$

donde  $\varphi(\cdot|\mathbf{0}, \boldsymbol{\Lambda})$ , denota la función de densidad de una  $N_2(\cdot|\mathbf{0}, \boldsymbol{\Lambda})$ ,  $\Phi(\cdot)$  y  $\phi(\cdot)$  denotan las funciones de distribución y de densidad de una Normal estándar, respectivamente,  $\mathbf{u} = (\cos \theta, \sin \theta)'$ ,  $\boldsymbol{\mu} * \mathbf{u} = \boldsymbol{\mu}' \mathbf{u}$  y

$$D(\theta) = \frac{\boldsymbol{\mu}' \boldsymbol{\Lambda}^{-1} \mathbf{u}}{(\mathbf{u}' \boldsymbol{\Lambda}^{-1} \mathbf{u})^{1/2}}.$$

Algunos casos particulares de interés son:

- Si  $\boldsymbol{\mu}' = \mathbf{0}$  y  $\boldsymbol{\Lambda} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1}$ , se sigue que

$$NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \frac{\sqrt{1 - \rho^2}}{2\pi(1 - \rho \sin 2\theta)}. \quad (2.1)$$

- Si  $\boldsymbol{\mu}' = (\mu, 0)$  y  $\boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$ , entonces

$$NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = (2\pi)^{-1/2} \phi(\mu) + \mu \cos \theta \phi(\mu \sin \theta) \Phi(\mu \cos \theta). \quad (2.2)$$

- Si  $\boldsymbol{\mu}' = \mathbf{0}$  y  $\boldsymbol{\Lambda} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}^{-1}$ , se obtiene

$$NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \frac{\sqrt{1 - h^2}}{2\pi(1 - h \cos 2\theta)}, \quad (2.3)$$

donde  $h = (\sigma_1^2 - \sigma_2^2)/(\sigma_1^2 + \sigma_2^2)$  con  $|h| \leq 1$ .

La distribución (2.2) aparece en el estudio de la potencia en cierto contraste de hipótesis (Klotz, 1964). En meteorología, (2.3) es empleada en el estudio de direcciones

de viento dado que ésta es la densidad de la dirección de viento, bajo el supuesto de que los componentes del vector  $\mathbf{Y}$  se distribuyen como dos variables aleatorias (vs.as.) normales independientes. Ver Mardia (1972).

La distribución Normal proyectada,  $NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda})$ , es muy versátil, ya que puede modelar comportamientos simétricos, asimétricos, unimodales y/o multimodales. Las Figuras 2.1, 2.2, 2.3 y 2.4, generadas en *Mathematica* v 5.2, muestran las curvas de nivel y las correspondientes distribuciones Normales proyectas (en un rango adecuado de longitud  $2\pi$ ) que se obtienen a partir de distribuciones Normales bivariadas con parámetros:

$$\boldsymbol{\mu}' = (2, 0), \quad \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

$$\boldsymbol{\mu}' = (2, 2), \quad \boldsymbol{\Lambda} = \begin{pmatrix} 2.25 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

$$\boldsymbol{\mu}' = (0, 1), \quad \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 9 \end{pmatrix}^{-1}$$

y

$$\boldsymbol{\mu}' = (0.5, 0.5), \quad \boldsymbol{\Lambda} = \begin{pmatrix} 4 & -0.36 \\ -0.36 & 4 \end{pmatrix}^{-1},$$

respectivamente.

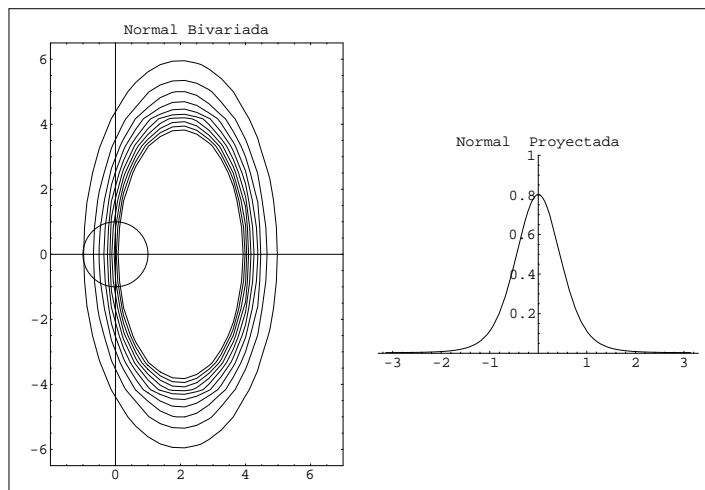


Figura 2.1: *Distribución Normal proyectada simétrica.*

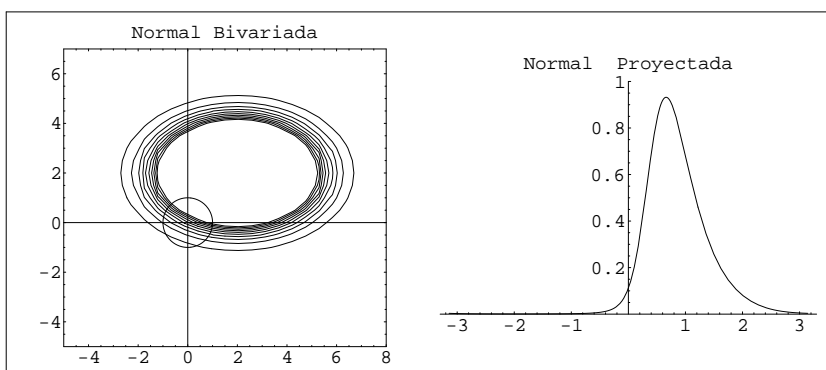


Figura 2.2: *Distribución Normal proyectada asimétrica.*



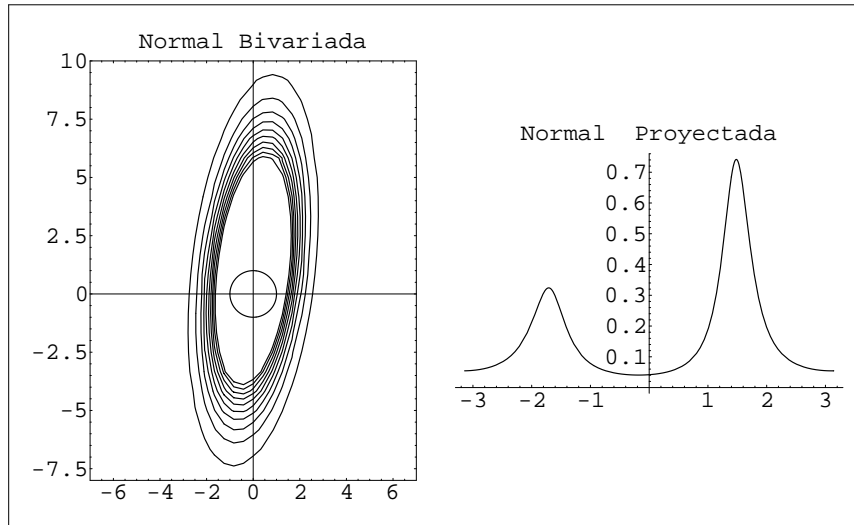


Figura 2.3: *Distribución Normal proyectada bimodal asimétrica.*

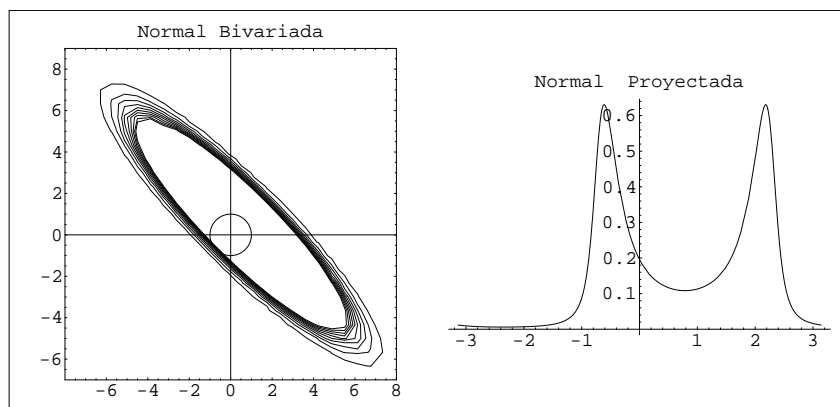


Figura 2.4: *Distribución Normal proyectada bimodal simétrica.*

## 2.2 Derivación de la distribución Normal proyectada $PN(\boldsymbol{\mu}, \boldsymbol{\Lambda})$

A continuación se presenta una forma de derivar distribuciones  $PN(\boldsymbol{\mu}, \boldsymbol{\Lambda})$  y en particular se muestra el caso de las densidades  $PN(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ .

**Definición 2.2.1** Sea  $\mathbf{Y}$  un vector aleatorio bivariado con distribución de probabilidad normal con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\boldsymbol{\Lambda}$ . Entonces, la función de densidad de probabilidad de  $\mathbf{Y}$  está dada por

$$f(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \frac{|\boldsymbol{\Lambda}|^{1/2}}{2\pi} \exp\left\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Lambda}(\mathbf{y} - \boldsymbol{\mu})\right\}.$$

**Proposición 2.2.1** Sea  $\mathbf{Y}$  un vector aleatorio con distribución  $N_2(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ , si se define la transformación

$$\mathbf{y} = r(\cos\theta, \sin\theta)' = r\mathbf{v}'$$

donde  $\theta \in (0, 2\pi]$  y  $r \in \mathbb{R}^+$ . Entonces, la función de densidad conjunta de la transformación  $(r, \theta)$ , está dada por

$$f(r, \theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = r C_6 d^2 \exp\left(-\frac{1}{2}d^2 r^2 + d^2 b r\right),$$

donde  $d^2 = \mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}$ ,  $b = \frac{\mathbf{v}'\boldsymbol{\Lambda}\boldsymbol{\mu}}{\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}}$  y

$$C_6 = \frac{|\boldsymbol{\Lambda}|^{1/2} \exp\left(-\frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Lambda}\boldsymbol{\mu}\right)}{2\pi \mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}}.$$

DEMOSTRACIÓN: Empleando el teorema de cambio de variable, se tiene

$$\begin{aligned} f(r, \theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= (2\pi)^{-1} |\boldsymbol{\Lambda}|^{1/2} \exp\left\{-\frac{1}{2}(r\mathbf{v} - \boldsymbol{\mu})'\boldsymbol{\Lambda}(r\mathbf{v} - \boldsymbol{\mu})\right\} r \\ &= r(2\pi)^{-1} |\boldsymbol{\Lambda}|^{1/2} \exp\left\{-\frac{1}{2}[(\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v})r^2 - 2(\mathbf{v}'\boldsymbol{\Lambda}\boldsymbol{\mu})r + \boldsymbol{\mu}'\boldsymbol{\Lambda}\boldsymbol{\mu}]\right\} \\ &= r(2\pi)^{-1} |\boldsymbol{\Lambda}|^{1/2} \exp\left\{-\frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Lambda}\boldsymbol{\mu}\right\} \exp\left\{-\frac{1}{2}(\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v})\right\} \left[r^2 - 2\frac{\mathbf{v}'\boldsymbol{\Lambda}\boldsymbol{\mu}}{\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}} r\right] \\ &= r(2\pi)^{-1} |\boldsymbol{\Lambda}|^{1/2} \exp\left\{-\frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Lambda}\boldsymbol{\mu}\right\} \frac{1}{\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}} \times \\ &\quad (\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}) \exp\left\{-(1/2)(\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v})r^2 + (\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v})\left(\frac{\mathbf{v}'\boldsymbol{\Lambda}\boldsymbol{\mu}}{\mathbf{v}'\boldsymbol{\Lambda}\mathbf{v}}\right) r\right\} \\ &= r C_6 d^2 \exp\left(-\frac{1}{2}d^2 r^2 + d^2 b r\right). \end{aligned}$$

□

**Proposición 2.2.2** *Bajo las mismas condiciones de la Proposición (2.2.1), la función de densidad del ángulo aleatorio  $\Theta$ , es decir, la densidad de probabilidad de la correspondiente Normal proyectada, está dada por*

$$\begin{aligned} PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= \int_0^\infty f(r, \theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) dr \\ &= C_6 [1 + \frac{db}{\phi(db)} \Phi(db)] I_{(0,2\pi]}(\theta) \end{aligned}$$

donde  $d$ ,  $b$  y  $C_6$  son como en la Proposición 2.2.1,  $\Phi(\cdot)$  y  $\phi(\cdot)$  denotan las funciones de distribución y de densidad de una Normal estándar, respectivamente,

DEMOSTRACIÓN.

$$\begin{aligned} PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= \int_0^\infty f(r, \theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) dr \\ &= C_6 \int_0^\infty r d^2 \exp\{-\frac{1}{2}d^2 r^2\} \exp\{d^2 b r\} dr. \end{aligned}$$

Trabajando sólo con la integral y haciendo el cambio de variable

$$u = \exp\{d^2 b r\} \quad y \quad dv = r d^2 \exp\{-\frac{1}{2}d^2 r^2\},$$

se tiene que

$$v = \int r d^2 \exp\{-\frac{1}{2}d^2 r^2\} = -\exp\{-\frac{1}{2}d^2 r^2\}.$$

Por lo tanto,

$$\begin{aligned} PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= \int_0^\infty r C_6 d^2 \exp\{-\frac{1}{2}d^2 r^2\} \exp\{d^2 b r\} dr \\ &= [-\exp\{d^2 b r\} \exp\{-\frac{1}{2}d^2 r^2\}]_0^\infty \\ &\quad + \int_0^\infty \exp\{-\frac{1}{2}d^2 r^2\} \exp\{d^2 b r\} d^2 b dr \\ &= 1 + \int_0^\infty d^2 b \exp\{-(dr - db)^2/2 + d^2 b^2/2\} dr. \end{aligned}$$

Por otro lado, como

$$\begin{aligned} \int_0^\infty d^2 b \exp\{-(dr - db)^2/2 + d^2 b^2/2\} dr &= d^2 b \exp\{d^2 b^2/2\} \\ &\quad \times \int_0^\infty \exp\{-(dr - db)^2/2\} dr \\ &= d^2 b \exp\{d^2 b^2/2\} \sqrt{2\pi} \\ &\quad \times \int_{db}^\infty (2\pi)^{-1/2} \exp\{-s^2/2\} ds \\ &= db \Phi(db) [\phi(db)]^{-1}, \end{aligned}$$

entonces

$$PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = C_6 [1 + \frac{db}{\phi(db)} \Phi(db)] I_{(0,2\pi]}(\theta).$$

□

Una vez que se ha obtenido la función de densidad de probabilidad  $f(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}) = NP(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda})$  y se cuenta con la función de densidad conjunta  $f(r, \theta|\boldsymbol{\mu}, \boldsymbol{\Lambda})$ , se puede obtener en particular la densidad condicional de  $R$  dado  $\Theta$ . Esta densidad condicional es necesaria para llevar a cabo los procedimientos de inferencia propuestos en el resto del trabajo.

**Proposición 2.2.3** *Bajo las mismas condiciones de la Proposición (2.2.1), la función de densidad condicional de  $R$  dado  $\Theta$  está dada por*

$$f(r|\theta, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \frac{d^2 r \exp\left(-\frac{1}{2} d^2 [r^2 - 2 b r]\right)}{1 + \frac{db}{\phi(db)} \Phi(db)} 1_{(0, \infty)}(r).$$

Los parámetros del modelo Normal proyectado son estimables pero no identificables si no se consideran contrastes adicionales; lo anterior debido a que para toda  $a > 0$ , si se toma  $\boldsymbol{\mu}^* = a\boldsymbol{\mu}$  y  $\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}/a^2$  la distribución de las direcciones observadas es la misma.

**Proposición 2.2.4**  $\forall a \in \mathbb{R}, a > 0$ , se cumple que

$$PN(\theta|\boldsymbol{\mu}^*, \boldsymbol{\Lambda}^*) = PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}).$$

donde

$$\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}/a^2 \quad y \quad \boldsymbol{\mu}^* = a\boldsymbol{\mu}.$$

DEMOSTRACIÓN: De la Proposición (2.2.2) se tiene que

$$PN(\theta|\boldsymbol{\mu}^*, \boldsymbol{\Lambda}^*) = \frac{|\boldsymbol{\Lambda}^*|^{1/2} \exp\left(-\frac{1}{2} (\boldsymbol{\mu}^*)' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^*\right)}{2\pi \boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{v}} \left[1 + \frac{\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^*}{\sqrt{\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{v}}} \exp\left\{\frac{1}{2} \frac{(\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^*)^2}{\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{v}}\right\} \Phi\left(\frac{\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^*}{\sqrt{\boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{v}}}\right)\right]$$

Si  $\boldsymbol{\mu}^* = a\boldsymbol{\mu}$  y  $\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}/a^2$ , entonces

$$\begin{aligned} |\boldsymbol{\Lambda}^*|^{1/2} &= |\boldsymbol{\Lambda}|^{1/2}/a^2, \\ (\boldsymbol{\mu}^*)' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^* &= \boldsymbol{\mu}' \boldsymbol{\Lambda} \boldsymbol{\mu} \\ \boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{v} &= (\boldsymbol{v}' \boldsymbol{\Lambda} \boldsymbol{v})/a^2 \\ \boldsymbol{v}' \boldsymbol{\Lambda}^* \boldsymbol{\mu}^* &= (\boldsymbol{v}' \boldsymbol{\Lambda} \boldsymbol{\mu})/a \end{aligned}$$

de lo anterior, se sigue que

$$PN(\theta|\boldsymbol{\mu}^*, \boldsymbol{\Lambda}^*) = PN(\theta|\boldsymbol{\mu}, \boldsymbol{\Lambda}).$$

□

A partir de la demostración de la Proposición (2.2.4) se puede notar que una manera de atacar el problema de la falta de identificabilidad de los parámetros de una distribución Normal proyectada es considerando matrices de precisión tales que  $|\boldsymbol{\Lambda}| = 1$ . Una familia importante de distribuciones Normales proyectadas se obtiene cuando la distribución Normal bivariada que la genera tiene como matriz de precisión a la matriz identidad, es decir,  $\boldsymbol{\Lambda} = \boldsymbol{I}$ . La propuesta metodológica en este trabajo se basa en esta familia particular de distribuciones, por lo tanto, en la siguiente sección se presentan algunos resultados relevantes asociados a la distribución  $PN(\boldsymbol{\mu}, \boldsymbol{I})$ .

### 2.3 La distribución Normal proyectada $PN(\boldsymbol{\mu}, \boldsymbol{I})$

La distribución Normal proyectada  $PN(\boldsymbol{\mu}, \boldsymbol{I})$  es unimodal y rotacionalmente simétrica alrededor de su dirección media  $\boldsymbol{\eta}$ , la cual se puede mostrar que es igual a  $\boldsymbol{\mu}/\|\boldsymbol{\mu}\|$  (ver Apéndice A).

A continuación se presentan algunos resultados asociados a la distribución de probabilidad  $PN(\boldsymbol{\mu}, \boldsymbol{I})$ .

**Proposición 2.3.1** Sea  $\mathbf{Y}$  un vector aleatorio bivariado con distribución de probabilidad normal con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\mathbf{I}$ . Entonces,

1.

$$PN(\theta|\boldsymbol{\mu}, \mathbf{I}) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\|\boldsymbol{\mu}\|^2\right\} \left[1 + \frac{\mathbf{v}'\boldsymbol{\mu}}{\phi(\mathbf{v}'\boldsymbol{\mu})}\Phi(\mathbf{v}'\boldsymbol{\mu})\right] 1_{(0,2\pi]}(\theta) 1_{\mathbb{R}^2}(\boldsymbol{\mu})$$

2.

$$f(r|\theta, \boldsymbol{\mu}, \mathbf{I}) = \frac{r \exp\left\{-\frac{1}{2}[r^2 - 2(\mathbf{v}'\boldsymbol{\mu})r]\right\}}{1 + \frac{\mathbf{v}'\boldsymbol{\mu}}{\phi(\mathbf{v}'\boldsymbol{\mu})}\Phi(\mathbf{v}'\boldsymbol{\mu})} 1_{(0,\infty)}(r) 1_{\mathbb{R}^2}(\boldsymbol{\mu}).$$

3.  $f(r|\theta, \boldsymbol{\mu}, \mathbf{I})$  pertenece a una familia exponencial con parámetro canónico  $b = \mathbf{v}'\boldsymbol{\mu}$  y

$$E(R|\Theta) = b + \frac{\Phi(b)}{\phi(b) + b\Phi(b)}.$$

DEMOSTRACIÓN. Como  $\mathbf{Y} \sim N_2(\boldsymbol{\mu}, \mathbf{I})$ , entonces, de la Proposición (2.2.1), se tiene que

$$\begin{aligned} d^2 &= \mathbf{v}'\mathbf{I}\mathbf{v} = 1, \\ b &= \mathbf{v}'\boldsymbol{\mu} \quad y \\ C_6 &= (2\pi)^{-1} \exp\left\{-\frac{1}{2}\|\boldsymbol{\mu}\|^2\right\}. \end{aligned}$$

Los resultados 1 y 2 se siguen de las Proposiciones (2.2.2) y (2.2.3). Por otro lado,

$$\begin{aligned} f(r|\theta, \boldsymbol{\mu}, \mathbf{I}) &= r \exp\left\{-\frac{1}{2}[r^2 - 2(\mathbf{v}'\boldsymbol{\mu})r]\right\} \left[(2\pi)^{-1}|\boldsymbol{\Lambda}|^{1/2} \left[1 + \frac{\mathbf{v}'\boldsymbol{\mu}}{\phi(\mathbf{v}'\boldsymbol{\mu})}\Phi(\mathbf{v}'\boldsymbol{\mu})\right]\right]^{-1} \\ &= \exp\left\{-\frac{r^2}{2} + (\mathbf{v}'\boldsymbol{\mu})r + \ln r - \psi(\mathbf{v}'\boldsymbol{\mu})\right\} \\ &= \exp\left\{[r(\mathbf{v}'\boldsymbol{\mu}) - \psi(\mathbf{v}'\boldsymbol{\mu})] - \frac{r^2}{2} + \ln r\right\}. \end{aligned}$$

donde

$$\psi(b) = \ln\left\{1 + \frac{b\Phi(b)}{\phi(b)}\right\}.$$

Así,  $f(r|\theta, \boldsymbol{\mu}, \mathbf{I})$  pertenece a una familia exponencial con parámetro canónico  $b = \mathbf{v}'\boldsymbol{\mu}$  y

$$\begin{aligned} E(R|\Theta) &= \frac{\partial}{\partial b} \psi(b) \\ &= \frac{\partial}{\partial b} [\ln(\phi(b) + b\Phi(b)) - \ln \phi(b)] \\ &= b + \Phi(b) [\phi(b) + b\Phi(b)]^{-1}. \end{aligned}$$

□

Como se definió al inicio de este capítulo, un vector aleatorio unitario  $q$ -dimensional  $\mathbf{U} = \mathbf{Y} \|\mathbf{Y}\|^{-1}$ , tiene una distribución Normal proyectada  $q$ -variada  $PN_q(\boldsymbol{\mu}, \boldsymbol{\Lambda})$  si  $\mathbf{Y}$  tiene una distribución normal  $q$ -variada con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\boldsymbol{\Lambda}$ . En este trabajo, se considera primordialmente el caso de datos circulares ( $q = 2$ ), sin embargo, un análisis similar a las Proposiciones (2.2.1), (2.2.2) y (2.2.3) produce los siguientes resultados para el caso de datos esféricos ( $q = 3$ ).

**Proposición 2.3.2** *Sea  $\mathbf{U}$  un vector aleatorio sobre la esfera unitaria, entonces  $\mathbf{U}$  se puede definir como  $\mathbf{U} = (\cos \theta \sin \omega, \sin \theta \sin \omega, \cos \theta)$ . Además, la función de densidad de probabilidad conjunta de  $(\theta, \omega)$  y la función de densidad condicional de  $r | (\theta, \omega)$  están dadas por*

$$f(\theta, \omega | \boldsymbol{\mu}) = C(\theta, \boldsymbol{\mu}) \left[ \frac{b-1}{b} + \frac{b^2+1}{b} \left( 1 + \frac{b\Phi(b)}{\Psi(b)} \right) \right] I_{(0,\pi]}(\theta) I_{(0,2\pi]}(\omega) I_{\mathbb{R}^3}(\boldsymbol{\mu}),$$

y

$$f(r | \theta, \omega, \boldsymbol{\mu}) \propto r^{3-1} \exp\{-\frac{1}{2}r^2 + br\} I_{(0,\infty)}(r),$$

donde

$$C(\theta, \boldsymbol{\mu}) = (2\pi)^{-3/2} \sin \theta \exp\{-\frac{1}{2}\|\boldsymbol{\mu}\|^2\},$$

y  $b = \mathbf{u}'\boldsymbol{\mu}$ . En este caso, se puede especificar  $\boldsymbol{\eta} = (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \alpha)'$  por lo que  $(\alpha, \beta)$  también determina la dirección media del vector  $(\theta, \omega)$ .

# Capítulo 3

## Modelos de Regresión para Datos Circulares

### 3.1 Antecedentes

Los datos direccionales, particularmente los datos circulares pueden aparecer en modelos de regresión como la variable de respuesta. Específicamente, sea  $(\mathbf{x}_1, \theta_1), \dots, (\mathbf{x}_n, \theta_n)$  observaciones independientes, donde  $\mathbf{x}$  es un vector de covariables y  $\Theta$  es la correspondiente variable circular, con dirección media  $w$ . El problema central consiste en modelar la dirección media  $w$  en términos de las covariables  $\mathbf{x}$ .

La mayoría de los modelos propuestos en la literatura para una respuesta direccional presentan problemas como falta de identificabilidad y dificultades computacionales que los hacen difíciles de analizar en la práctica. Lo anterior resulta contrastante con los métodos simples de análisis típicamente disponibles para modelos con una respuesta lineal.



La teoría para modelos de regresión cuando la respuesta es una variable circular no se ha desarrollado completamente a pesar de que se pueden encontrar aplicaciones potenciales en varias áreas del conocimiento, particularmente en biología, geología y meteorología. En la literatura se han propuesto un número de modelos donde la variable respuesta es circular. Específicamente, Kato *et al.* (2008), Downs y Mardia (2002), y Rivest (1997) analizan el caso en el que se tienen covariables de naturaleza circular. Para una revisión sobre modelos de regresión circular-circular el lector se puede referir a Fisher (1993).

Recientemente, George and Ghosh (2006) discuten un modelo de regresión para una variable de respuesta circular y covariables lineales. Estos autores proponen un modelo de regresión semiparamétrico Bayesiano sobre una sola covariable y consideran el caso donde la variable respuesta tiene una distribución von Mises y ajustan el modelo utilizando técnicas de Monte Carlo vía cadenas de Markov. Anterior al trabajo de George and Ghosh (2006), Presnell *et al.* (1998) proponen un modelo basado en la distribución Normal bivariada bajo proyección y ofrecen estimadores de máxima verosimilitud para los parámetros involucrados usando métodos iterativos tales como Newton-Raphson y algoritmos EM. Sin embargo, como ellos mismos lo señalan, la aproximación asintótica a la varianza de los estimadores de los parámetros de regresión puede ser inadecuada o difícil de usar cuando la varianza verdadera es grande. Los modelos de regresión anteriores de Gould (1969), Johnson y Wehrly (1978), and Fisher y Lee (1992), consideran que la variable de respuesta sigue una distribución von Mises y que la media angular o el parámetro de concentración, o ambos, pueden depender de covariables.

Específicamente, el modelo propuesto por Gould (1969) considera  $w_i = w_0 + \mathbf{x}'_i \boldsymbol{\beta}$  y  $\kappa_i = \kappa \forall i = 1, \dots, n$  donde  $w_0$ ,  $\boldsymbol{\beta}$  y  $\kappa$  son parámetros desconocidos. El modelo anterior es inválido en muchas situaciones; Johnson y Wehrly (1978) y el mismo Gould (1969) señalan que se pueden presentar problemas de identificabilidad para ajustar este modelo, ya que para ciertos valores de las covariables la verosimilitud tiene una infinidad de máximos distintos.

Johnson y Wehrly (1978) consideran el caso de una sola covariable y proponen dos modelos, uno para el parámetro de localización y otro para el de concentración. Por su parte, Fisher y Lee (1992) proponen tres generalizaciones de los modelos de Johnson y Wehrly (1978); la más completa considera un modelo de la forma  $w_i = w + g(\mathbf{x}'_i \boldsymbol{\beta})$  y  $\kappa_i = h(\kappa)$ , para ciertas funciones  $g(\cdot)$  y  $h(\cdot)$ . Para ajustar dichos modelos, Fisher y Lee (1992) sugieren estimaciones de máxima verosimilitud empleando mínimos cuadrados ponderados iterados, sin embargo, su implementación presenta serias dificultades computacionales; la verosimilitud es multimodal y tiende a incrementarse cuando los componentes, del vector  $\boldsymbol{\beta}$ ,  $\beta_j \rightarrow \pm\infty$ . Lo anterior debido a que  $\beta_j = 0$  es indistinguible de  $\beta_j \approx \infty$  una vez que  $w$  es ajustada. Además, el máximo puede ocurrir en un pico demasiado estrecho. Presnell *et al.* (1998) señalan que la verosimilitud asociada a los modelos de Fisher y Lee (1992) presenta varios problemas de optimización numérica y que en su experiencia tales modelos pueden resultar inadecuados para uso general. Fisher y Lee (1992) sugieren una exploración gráfica de la verosimilitud como posible solución, lo cual puede resultar difícil o imposible cuando se tiene más de una covariable.

## 3.2 El modelo

El principal interés se enfoca en modelar la dependencia de la dirección media de una variable de respuesta circular  $\Theta$  sobre una o más variables explicativas, o covariables. Aunque los resultados presentados en este trabajo se pueden extender a otros casos, aquí sólo se discute el caso de covariables lineales las cuales serán denotadas por  $\mathbf{x} = (x_1, \dots, x_v)$ .

Sean  $(\mathbf{x}_1, \mathbf{u}_1), \dots, (\mathbf{x}_n, \mathbf{u}_n)$  observaciones independientes, donde  $\mathbf{x}_i$  es un vector de covariables y  $\mathbf{u}_i$  es la correspondiente variable circular, con dirección media  $\boldsymbol{\eta}_i$  y longitud resultante media  $\rho_i$ , la cual puede depender de  $\mathbf{x}_i$ . Para datos circulares,  $\theta_i$  y  $w_i$  son las representaciones angulares de  $\mathbf{u}_i = (\cos \theta_i, \sin \theta_i)'$  y  $\boldsymbol{\eta}_i = (\cos w_i, \sin w_i)'$ , respectivamente.

El modelo de regresión Normal proyectado considera la siguiente representación:

$$\mathbf{U}_i = \mathbf{Y}_i/R, \quad R = \|\mathbf{Y}_i\|, \quad i = 1, \dots, n,$$

donde  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  son vectores aleatorios independientes con distribución de probabilidad normal bivariada con vector de medias  $\boldsymbol{\mu}_i = E(\mathbf{Y}_i)$  y con matriz de precisión la matriz identidad  $\mathbf{I}$ . Si el vector de covariables  $\mathbf{x}_i$  es el mismo para ambas componentes de  $\boldsymbol{\mu}_i$  entonces se tiene la estructura estándar de un modelo de regresión normal multivariado, es decir  $\boldsymbol{\mu}_i = \mathbf{B}'\mathbf{x}_i$ , donde  $\mathbf{B} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2]$  es la matriz de coeficientes de regresión. Sin embargo, en la práctica cada uno de los dos componentes de  $\boldsymbol{\mu}_i$  puede depender de diferentes subconjuntos de covariables, en cuyo caso los vectores de coeficientes de regresión,  $\boldsymbol{\beta}^1$  y  $\boldsymbol{\beta}^2$ , pueden tener diferentes dimensiones.

Afortunadamente, en este trabajo se puede considerar este caso general ya que la matriz de precisión en el modelo es la matriz identidad, y así es posible especificar un modelo de regresión múltiple univariado de manera independiente sobre cada una de las componentes de  $\boldsymbol{\mu}_i$ . En esta forma, los componentes de  $\boldsymbol{\mu}_i$  están dados por

$$\mu_i^j = (\mathbf{x}_i^j)' \boldsymbol{\beta}^j, \quad \forall j = 1, 2$$

donde  $\mathbf{x}_i^j$  denota el subconjunto de covariables asociadas con la  $j$ -ésima componente de  $\boldsymbol{\mu}_i$ . Para facilitar la exposición, y con un ligero abuso de notación, se escribirá  $\mathbf{B} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2]$  y  $\boldsymbol{\mu}_i = \mathbf{B}' \mathbf{x}_i$  aún en el caso general, bajo el entendido que  $\mathbf{x}_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2\}$ , y que  $\boldsymbol{\beta}^1$  y  $\boldsymbol{\beta}^2$  pueden tener diferentes dimensiones.

Por otro lado,  $\mathbf{U}$  es un vector unitario 2-dimensional, por lo que también puede ser especificado utilizando un solo ángulo. Dado que  $\mathbf{Y} \sim N_2(\cdot | \boldsymbol{\mu}_x = \mathbf{B}' \mathbf{x}, \mathbf{I})$ , la densidad conjunta de  $(\Theta, R)$ ,  $f_{(\Theta, R)}(\theta, r)$ , se puede obtener haciendo  $R = \|\mathbf{Y}\|$  y posteriormente transformando a coordenadas polares. De lo anterior, omitiendo el subíndice en  $\boldsymbol{\mu}_x$ , se sigue que  $\Theta$  tiene una distribución Normal proyectada con función de densidad dada por

$$PN(\theta | \boldsymbol{\mu}, \mathbf{I}) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\|\boldsymbol{\mu}\|^2\right\} \left[1 + \frac{\mathbf{u}' \boldsymbol{\mu} \Phi(\mathbf{u}' \boldsymbol{\mu})}{\phi(\mathbf{u}' \boldsymbol{\mu})}\right] 1_{(0, 2\pi]}(\theta) 1_{\mathbb{R}^2}(\boldsymbol{\mu}),$$

donde  $\mathbf{u}' = (\cos \theta, \sin \theta)$  y  $1_A(\cdot)$  es la función indicadora sobre el conjunto  $A$ . Aquí,  $\phi(\cdot)$  y  $\Phi(\cdot)$  denotan la función de densidad de probabilidad y la función de distribución acumulada de la distribución normal estándar, respectivamente. Ver, Proposición (2.3.1) del Capítulo 2.

En un contexto Bayesiano, un modelo es completamente especificado por la distribución de probabilidad de las observaciones dados los parámetros, junto con una distribución inicial sobre tales parámetros. Así, el modelo de regresión Normal proyectado se completa con una especificación inicial  $f(\mathbf{B})$  sobre los coeficientes de regresión.

Este modelo considera las observaciones direccionales como proyecciones sobre el círculo unitario de vectores parcialmente no observados de una variable normal bivariada. El modelo Normal proyectado de ninguna forma requiere que las direcciones observadas sean realmente proyecciones de una respuesta multivariada, el modelo simplemente especifica una estructura conveniente. Este modelo es unimodal y rotacionalmente simétrico alrededor de la dirección media  $\boldsymbol{\eta}$ , la cual se puede mostrar que es igual a  $\boldsymbol{\mu}/\|\boldsymbol{\mu}\|$ . Aún más, cualquier distribución von Mises y ciertas distribuciones normales *wrapped* son comparables con una distribución Normal proyectada de esta forma (ver Sección 1.3.2)

En vista de las consideraciones anteriores, el modelo Normal proyectado  $PN(\cdot|\boldsymbol{\mu}, \mathbf{I})$  no pierde aplicabilidad práctica comparado con los modelos de regresión que consideran una distribución *wrapped* o una distribución von Mises para la variable respuesta, la cual es probablemente la distribución generalmente considerada para una variable respuesta circular.

Por la discusión presentada en esta sección, el problema principal se puede plantear de la siguiente manera: Dada una muestra  $\{\theta_1, \dots, \theta_n\}$  de  $PN(\cdot|\boldsymbol{\mu} = \mathbf{B}'\mathbf{x}, \mathbf{I})$ , y un conjunto de covariables  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , ¿cómo se pueden llevar a cabo inferencias sobre  $\mathbf{B}$  y, si es necesario, usar el modelo para otros fines tales como análisis de datos faltantes o predicción?

### 3.2.1 El problema de datos faltantes

Supóngase que una porción de los datos muestrales son faltantes. Específicamente, suponga que se tienen datos faltantes en la variable respuesta. En este caso, el muestreo de Gibbs ofrece una manera conveniente de analizar el problema. La idea se basa en el

método denominado *data augmentation* o método de variables auxiliares (ver Tanner y Wong, 1987), el cual se presenta a continuación.

Sea  $\mathbf{z} = (\mathbf{z}_{obs}, \tilde{\mathbf{z}})$  un conjunto de observaciones del modelo

$$\{p(\mathbf{z}|\boldsymbol{\alpha}), p(\boldsymbol{\alpha})\},$$

donde  $\mathbf{z}_{obs}$  denota los datos observados y  $\tilde{\mathbf{z}}$  representa los datos faltantes. Aquí,  $\boldsymbol{\alpha}$  representa un vector de parámetros desconocidos y  $p(\boldsymbol{\alpha})$  es una distribución inicial. Aunque en situaciones como ésta la forma de la densidad final  $p(\boldsymbol{\alpha}|\mathbf{z}_{obs})$  es usualmente bastante complicada,  $p(\boldsymbol{\alpha}|\mathbf{z})$  generalmente tiene una forma más simple (ver, por ejemplo, Tanner y Wong, 1987). Se debe notar que

$$p(\boldsymbol{\alpha}|\mathbf{z}_{obs}, \tilde{\mathbf{z}}) = p(\boldsymbol{\alpha}|\mathbf{z})$$

y

$$p(\tilde{\mathbf{z}}|\boldsymbol{\alpha}, \mathbf{z}_{obs}) = p(\tilde{\mathbf{z}}|\boldsymbol{\alpha}).$$

Así,  $\tilde{\mathbf{z}}$  se puede ver como un parámetro desconocido adicional y se puede incluir en un muestreo de Gibbs con densidades condicionales completas dadas por  $p(\boldsymbol{\alpha}|\mathbf{z})$  y  $p(\tilde{\mathbf{z}}|\boldsymbol{\alpha})$ . En esta forma, es posible generar observaciones de la distribución de interés,  $p(\boldsymbol{\alpha}|\mathbf{z}_{obs})$ .

Para el modelo de regresión Normal proyectado, se tomó  $\boldsymbol{\alpha} \equiv \mathbf{B}$  y  $\mathbf{z} = (\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$  un conjunto de observaciones del modelo

$$\{PN(\mathbf{z}|\mathbf{B}, \mathbf{I}), f(\mathbf{B})\},$$

donde  $\boldsymbol{\theta}$  y  $\tilde{\boldsymbol{\theta}}$  denotan datos observados y no observados, respectivamente. Como se describió previamente, si se considera  $\tilde{\boldsymbol{\theta}}$  como un parámetro adicional, entonces se puede incluir en un esquema de muestreo de Gibbs con las siguientes densidades condicionales

completas

$$f(\boldsymbol{\beta}^j | \theta_1, \dots, \theta_s, \tilde{\theta}_{s+1}, \dots, \tilde{\theta}_n) \quad \forall j = 1, 2,$$

$$f(\tilde{\boldsymbol{\theta}} | \mathbf{B}, \tilde{\mathbf{x}}, \boldsymbol{\theta}, \mathbf{r}) = f(\tilde{\boldsymbol{\theta}} | \mathbf{B}, \tilde{\mathbf{x}}, \mathbf{r}),$$

$$f(\tilde{\mathbf{r}} | \mathbf{z}, \mathbf{B}, \tilde{\mathbf{x}}, \mathbf{r}) = f(\tilde{\mathbf{r}} | \tilde{\boldsymbol{\theta}}, \mathbf{B}, \tilde{\mathbf{x}}),$$

y

$$f(\mathbf{r} | \mathbf{z}, \mathbf{B}, \tilde{\mathbf{x}}).$$

Así, se está en condiciones de generar observaciones de la distribución final  $f(\mathbf{B} | \boldsymbol{\theta})$ .

### 3.2.2 El problema de selección de modelos

El problema de selección de variables es uno de los más comunes en el contexto de regresión. En cierto modo, no existe una estrategia totalmente aceptada para la selección de modelos. Sin embargo, desde una perspectiva Bayesiana, el uso de criterios predictivos ha sido reconocido como una manera apropiada de atacar este problema. (ver, por ejemplo, Gelfand *et al.*, 1992, Gelfand y Dey, 1994, Gutiérrez-Peña, 1997, George, 2000, Barbieri y Berger, 2004, y las referencias incluidas ahí).

En este trabajo de investigación se considera el caso particular de seleccionar un modelo de entre un conjunto finito de modelos  $\mathbf{M} = \{M_1, \dots, M_m\}$ , usando un criterio predictivo Bayesiano. Se emplea el criterio denominado *high-structure* propuesto por Geisser y Eddy (1979), el cual está basado en una medida condicional predictiva. Más adelante se muestra cómo este criterio predictivo puede ser implementado en el contexto de regresión para datos circulares. Esencialmente, el criterio consiste en maximizar el producto de las densidades predictivas condicionales en lugar de maximizar la densidad

predictiva conjunta. Geisser y Eddy (1979) argumentan que, bajo ciertas condiciones, este procedimiento es asintóticamente equivalente al bien conocido criterio de Akaike.

Sean  $(\mathbf{x}_1, \mathbf{u}_1), \dots, (\mathbf{x}_n, \mathbf{u}_n)$  observaciones independientes del modelo  $M_k$ , donde  $\mathbf{x}_i$  es un vector de covariables y  $\mathbf{u}_i$  es la correspondiente respuesta direccional con distribución Normal proyectada  $PN(\cdot|\boldsymbol{\mu}, \mathbf{I})$ . Si  $\mathbf{u}_{(-j)} = (\mathbf{u}_1, \dots, \mathbf{u}_{j-1}, \mathbf{u}_{j+1}, \dots, \mathbf{u}_n)$  representa un conjunto de datos con la  $j$ -ésima observación  $\mathbf{u}_j$  omitida, y  $f(\mathbf{u}^*|\mathbf{u}, \mathbf{x}, M_k)$  es una densidad predictiva para una observación futura cuando el modelo  $M_k$  es verdadero, entonces el criterio predictivo condicional consiste en seleccionar aquel modelo que maximice la medida  $L_k$ , donde

$$L_k = \prod_{j=1}^n f_j(\mathbf{u}_j^*|\mathbf{u}_{(-j)}, \mathbf{x}_j, M_k) \quad \forall k = 1, \dots, m.$$

### 3.3 Inferencias vía métodos MCCM

Para realizar inferencias sobre el modelo propuesto, en este trabajo se considera la introducción de variables latentes apropiadas para definir una distribución conjunta aumentada con  $\mathbf{B}$ . Esta distribución conjunta se construye de tal forma que permita simular de todas las densidades condicionales requeridas para un muestreo de Gibbs.

Si  $(\Theta_1, R_1), \dots, (\Theta_n, R_n)$  pudieran ser observadas, entonces sería posible realizar inferencias sobre los coeficientes de  $\mathbf{B}$  de manera relativamente simple; sin embargo, el problema es que realmente sólo se observan las direcciones  $\{\theta_1, \dots, \theta_n\}$ . La estructura del modelo sugiere tratar los  $R_i = \|\mathbf{Y}_i\|$ ,  $i = 1, \dots, n$ , no observados como variables latentes (auxiliares). Así, el modelo para los datos completos (aumentados) resulta ser el usual modelo de regresión normal multivariado. Este fue el enfoque seguido por Presnell *et al.* (1998), quienes se enfocaron principalmente en estimación de máxima verosimilitud de



$\mathbf{B}$  vía algoritmos EM.

En el modelo de regresión normal multivariado usual, si  $\mathbf{D}_n = (\mathbf{y}_1, \dots, \mathbf{y}_n)$  es una muestra de observaciones independientes de  $N_2(\cdot | \boldsymbol{\mu} = \mathbf{B}'\mathbf{x}, \mathbf{I})$  y se considera la inicial conjugada  $N(\boldsymbol{\beta}^j | \boldsymbol{\beta}_0^j, \boldsymbol{\Lambda}_0^j)$ , entonces la distribución final de  $\boldsymbol{\beta}^j$  resulta ser

$$f(\boldsymbol{\beta}^j | \mathbf{D}_n) = N(\cdot | \boldsymbol{\mu}_F^j, \boldsymbol{\Lambda}_F^j), \quad \forall j = 1, 2,$$

con

$$\begin{aligned} \boldsymbol{\Lambda}_F^j &= \boldsymbol{\Lambda}_0^j + \mathbf{X}'\mathbf{X}, \\ \boldsymbol{\mu}_F^j &= (\boldsymbol{\Lambda}_F^j)^{-1}(\boldsymbol{\Lambda}_0^j\boldsymbol{\beta}_0^j + \mathbf{X}'\mathbf{y}^j), \end{aligned}$$

donde  $\mathbf{y}^j$  es el vector formado por las  $j$ -ésimas componentes de los vectores  $\mathbf{y}_1, \dots, \mathbf{y}_n$  y  $\mathbf{X}$  es la correspondiente matriz de diseño. Hay que recordar que  $\boldsymbol{\Lambda}_0^j$  y  $\boldsymbol{\Lambda}_F^j$  son matrices de precisión.

Si ahora se considera la variable latente  $R$  definida en  $(0, \infty)$ , dado que  $\mathbf{Y} \sim N_2(\cdot | \boldsymbol{\mu} = \mathbf{B}'\mathbf{x}, \mathbf{I})$  se puede definir su densidad conjunta con  $\Theta$  como

$$\begin{aligned} f(\theta, r | \boldsymbol{\mu} = \mathbf{B}'\mathbf{x}) &= N_2(r\mathbf{u} | \boldsymbol{\mu} = \mathbf{B}'\mathbf{x}, \mathbf{I}) |J| \\ &= (2\pi)^{-p/2} \exp\{-\frac{1}{2}\|\boldsymbol{\mu}\|^2\} \exp\{-\frac{1}{2}[r^2 - 2r(\mathbf{u}'\boldsymbol{\mu})]\} |J|, \end{aligned} \quad (3.1)$$

donde  $|J| = r$  es el Jacobiano de la transformación  $\mathbf{y} \mapsto (\theta, r)$  y  $\mathbf{u} = (\cos \theta, \sin \theta)'$ .

Para estar en condiciones de implementar un muestreo de Gibbs es necesario especificar todas las condicionales completas. Éstas se derivan en la siguiente sección.

### 3.3.1 Distribuciones condicionales

La densidad condicional de  $\boldsymbol{\beta}^j$ ,  $\forall j = 1, 2$ , está dada por

$$f(\boldsymbol{\beta}^j | \theta_1, \dots, \theta_n, \mathbf{r}) = N(\cdot | \boldsymbol{\mu}_F^j, \boldsymbol{\Lambda}_F^j), \quad (3.2)$$

donde  $\mathbf{r} = (r_1, \dots, r_n)$  es un vector  $n$ -dimensional.

Se debe notar que las  $R_i$  ( $i = 1, \dots, n$ ) son condicionalmente independientes dadas las  $\Theta_i$  ( $i = 1, \dots, n$ ). Así, se puede obtener la densidad condicional de  $R_i$  a partir de (3.1), ver Proposición (2.3.1) del Capítulo 2. Esta densidad condicional está dada, salvo una constante de proporcionalidad, por

$$f(r_i|\theta_i, \boldsymbol{\mu}_i = \mathbf{B}'\mathbf{x}_i) \propto r_i \exp\{-\frac{1}{2}r_i^2 + b_i r_i\} 1_{(0,\infty)}(r_i), \quad (3.3)$$

donde  $b_i = \mathbf{u}'_i \boldsymbol{\mu}_i$ . Se puede mostrar que  $f(r_i|\theta_i, \boldsymbol{\mu}_i)$  es logcóncava y pertenece a una familia exponencial con parámetro canónico  $b_i$ , ver Proposición (2.3.1) del Capítulo 2.

Claramente, es fácil muestrear de  $f(\boldsymbol{\beta}^j|\theta_1, \dots, \theta_n, \mathbf{r})$ . Por otro lado, se pueden generar realizaciones  $R_i$  de  $f(r_i|\theta_i, \boldsymbol{\beta}^1, \boldsymbol{\beta}^2)$  de tal forma que es factible simular un vector aleatorio  $\mathbf{R}$  de  $f(\mathbf{r}|\theta_1, \dots, \theta_n, \boldsymbol{\beta}^1, \boldsymbol{\beta}^2)$ . De hecho, este último paso se lleva a cabo vía un algoritmo de Metropolis-Hastings en su versión denominada de independencia.

Finalmente, se pueden usar las condicionales completas definidas a través de (3.2) y (3.3) en un muestreo de Gibbs para obtener una muestra de la distribución final conjunta  $f(\mathbf{B}, \mathbf{r}|\theta_1, \dots, \theta_n)$ . De lo anterior, se sigue que la muestra de  $\mathbf{B}$  obtenida de esta manera tiene la distribución (marginal) requerida  $f(\mathbf{B}|\theta_1, \dots, \theta_n)$ .

### 3.3.2 Datos faltantes

Como antes,  $\mathbf{z} = (\theta_1, \dots, \theta_s, \tilde{\theta}_{s+1}, \dots, \tilde{\theta}_n)$  denota un conjunto de observaciones del modelo  $\{PN(z|\mathbf{B}, \mathbf{I}), f(\mathbf{B})\}$ , donde  $\theta$  y  $\tilde{\theta}$  denotan datos observados y no observados, respectivamente. En el modelo para datos-aumentados las densidades condicionales

completas de los vectores  $\boldsymbol{\beta}^j$ , para  $j = 1, 2$ , están dadas por

$$f(\boldsymbol{\beta}^j | (r, \theta)_1, \dots, (r, \theta)_s, (\tilde{r}, \tilde{\theta})_{s+1}, \dots, (\tilde{r}, \tilde{\theta})_n) = N(\cdot | \boldsymbol{\mu}_F^j, \boldsymbol{\Lambda}_F^j).$$

En este caso, es necesario determinar las densidades condicionales completas para  $\tilde{\Theta}$  y todas las variables latentes  $R_i$  y  $\tilde{R}_i$ . Sin embargo, se puede tomar ventaja del esquema del muestreador de Gibbs para generar observaciones conjuntas de  $(\tilde{R}, \tilde{\Theta})$ . Lo anterior, reconociendo que después de una transformación apropiada, la densidad condicional completa de  $(\tilde{R}, \tilde{\Theta})$ ,  $f((\tilde{r}, \tilde{\theta}) | \mathbf{B}, \tilde{\mathbf{x}})$ , es el modelo normal bivariado, y por lo tanto

$$f((\tilde{r}, \tilde{\theta}) | \mathbf{B}, \tilde{\mathbf{x}}) = N_2(\tilde{r}\tilde{\mathbf{u}} | \boldsymbol{\mu} = \mathbf{B}'\tilde{\mathbf{x}}, \mathbf{I}).$$

Finalmente, las densidades condicionales completas de las  $R$ 's asociadas a los ángulos realmente observados están dadas por (3.3).

### 3.3.3 Selección de modelos

Por la discusión presentada en las secciones anteriores, dado un conjunto de observaciones  $\theta_1, \dots, \theta_n$ , se puede obtener una muestra  $\{\mathbf{B}^{(1)}, \dots, \mathbf{B}^{(N)}\}$  de la distribución final de  $\mathbf{B} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2]$  bajo cada uno de los modelos  $M_k$ . Así, se está en condiciones de calcular

$$f(\theta_i^* | \mathbf{B}^{(l)}, \mathbf{x}_i) = PN(\theta_i^* | \boldsymbol{\mu}_i^{(l)} = (\mathbf{B}^{(l)})^t \mathbf{x}_i)$$

y proponer una densidad predictiva mediante el estimador de Monte Carlo.

$$f(\theta_i^* | \theta_{(-i)}, \mathbf{x}_i, M_k) = \frac{1}{N} \sum_{l=1}^N PN(\theta_i^* | \boldsymbol{\mu}^{(l)}) \quad (3.4)$$

donde, como antes,  $\theta_{(-i)} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n)$  representa el conjunto de datos con la  $i$ -ésima observación omitida. Por lo tanto, la medida asociada al criterio de Geisser

and Eddy (1979) para cada modelos  $M_k$  se obtiene como

$$L_k = \prod_{i=1}^n f(\theta_i^* | \theta_{(-i)}, \mathbf{x}_i, M_k).$$

### 3.4 Ejemplos

En los siguientes ejemplos se utilizó el lenguaje y ambiente R (R Development Core Team, 2010) para simular los respectivos conjuntos de datos, así como para implementar todos los análisis propuestos.

**Ejemplo 3.1.** En este ejemplo se considera una muestra simulada de tamaño 100 de una distribución Normal proyectada con vector de medias  $\boldsymbol{\mu} = (\mu_1, \mu_2)'$  definido por

$$\begin{aligned} \mu_1 &= -3.0 + 1.0 x \\ \mu_2 &= 1.5 - 2 x + 3 x^2, \end{aligned}$$

donde los valores de la covariable  $x$  se tomaron como una secuencia regular en el intervalo  $(0, 2)$ . Se tomó  $\boldsymbol{\beta}_0^1 = \mathbf{0}$ ,  $\boldsymbol{\Lambda}_0^1 = \text{Diag}(0.0001, 0.0001)$ ,  $\boldsymbol{\beta}_0^2 = \mathbf{0}$ ,  $\boldsymbol{\Lambda}_0^2 = \text{Diag}(0.0001, 0.0001, 0.0001)$  y  $\mathbf{r} = (1, \dots, 1)$ . Estas especificaciones de  $\boldsymbol{\Lambda}_0^1$  y  $\boldsymbol{\Lambda}_0^2$  producen iniciales con varianzas todas iguales a  $1/0.0001 = 10,000$ . De esta manera, todas las densidades iniciales reflejan información vaga relativa a la que proporcionan los datos.

Las distribuciones marginales resultantes para cada componente de los vectores  $\boldsymbol{\beta}^1$  y  $\boldsymbol{\beta}^2$  se presentan en la Figura 3.1. Se puede notar que el verdadero valor de cada uno de los parámetros se encuentra contenido en la región de máxima densidad de la correspondiente densidad final.

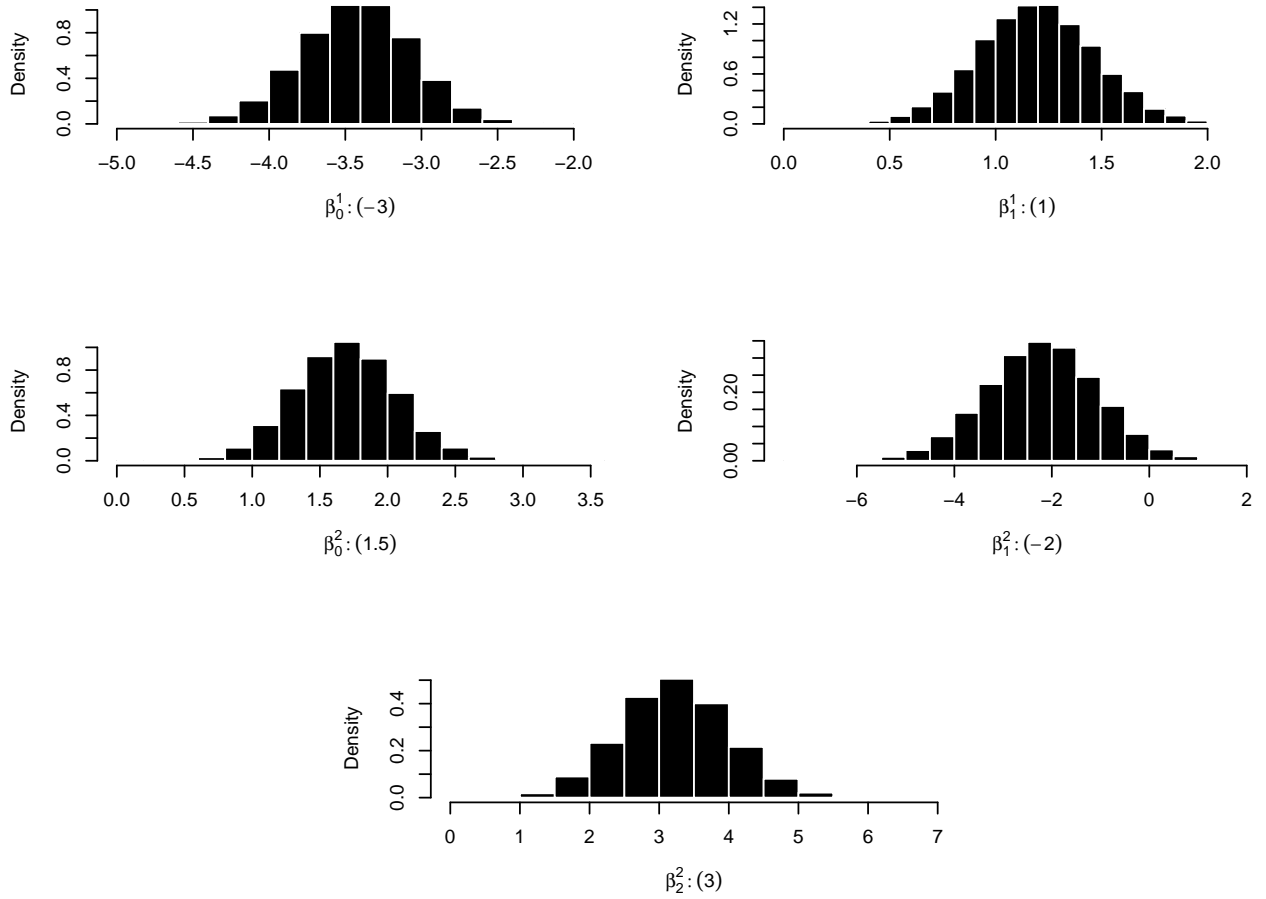


Figura 3.1: *Distribución final de  $\mathbf{B}$  (por componente) para el ejemplo 3.1. El verdadero valor de cada uno de los parámetros de regresión aparece entre paréntesis.*

**Ejemplo 3.2.** En este ejemplo, se simuló una muestra de tamaño 100 de una distribución von Mises( $\omega, \kappa$ ) con  $\omega = 0.1 + 5.0x$  y  $\kappa = 8$ , via la función `rvm` del paquete `CircStats` (ver, Jammalamadaka y SenGupta, 2001, o R, 2010). Aquí, los valores de la covariable  $x$  se tomaron como una secuencia regular sobre el intervalo  $(-1, 1)$ . Se tomó  $\beta_0^2 = \mathbf{0}$ ,  $\Lambda_0^2 = \text{Diag}(0.0001, 0.0001)$ ,  $\beta_0^1 = 0$ ,  $\Lambda_0^1 = 0.0001$  y  $\mathbf{r} = (1, \dots, 1)$ . En la Figura 3.2 se

muestran los datos y la dirección media circular ajustada con el modelo Normal proyectado propuesto. El método gráfico empleado despliega los 100 puntos de datos  $(x, \theta)$  en coordenadas Cartesianas, adicionalmente se grafican los puntos  $(x, \theta + 360)$  de manera que la relación (si es que existe) se muestre en la parte central, sin requerir que el ojo humano relacione los patrones de la parte superior con los patrones de la parte inferior de la gráfica. Esta manera de desplegar datos circulares fue sugerida por Fisher and Lee (1992).

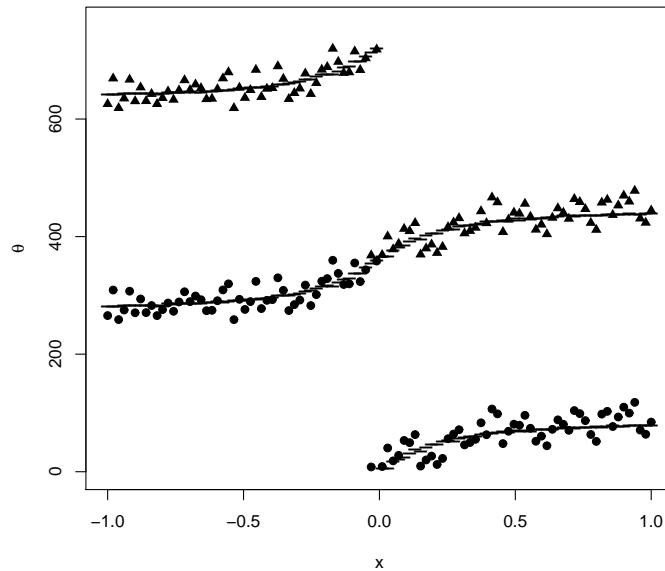


Figura 3.2: Conjunto de 100 observaciones (●) de un modelo de regresión von Mises. Los valores  $\theta + 360$  están representados por (▲). La línea representa la dirección media ajustada con base a un modelo Normal proyectado.

En este ejemplo se puede ver que el conjunto de datos simulados con una distribución von Mises se puede modelar razonablemente por una distribución Normal proyectada.

**Ejemplo 3.3.** Para este ejemplo se utilizó el modelo de regresión propuesto para analizar un conjunto de datos reales de tamaño 31, referentes a las distancias y direcciones tomadas por pequeños caracoles azules (*periwinkles*), después de que ellos fueron desplazados de la altura a la cual normalmente viven (ver Fisher, 1993, Apéndice B20). Se consideró la especificación propuesta por Presnell *et al.* (1998), la cual, empleando su notación, se define como

$$\begin{aligned}\mu_1 &= 1 + x \\ \mu_2 &= 1 + x.\end{aligned}$$

Aquí, el valor de los parámetros de las funciones de densidades  $N(\beta^j | \beta_0^j, \Lambda_0^j)$  para  $j = 1, 2$ , asociadas a la especificación inicial, se tomaron como  $\beta_0^j = \mathbf{0}$ ,  $\Lambda_0^j = \text{Diag}(0.0001, 0.0001)$  para  $j = 1, 2$  y  $\mathbf{r} = (1, \dots, 1)$ . Las distribuciones finales marginales para cada componente de los vectores  $\beta^1$  and  $\beta^2$  se presenta en la Figura 3.3.

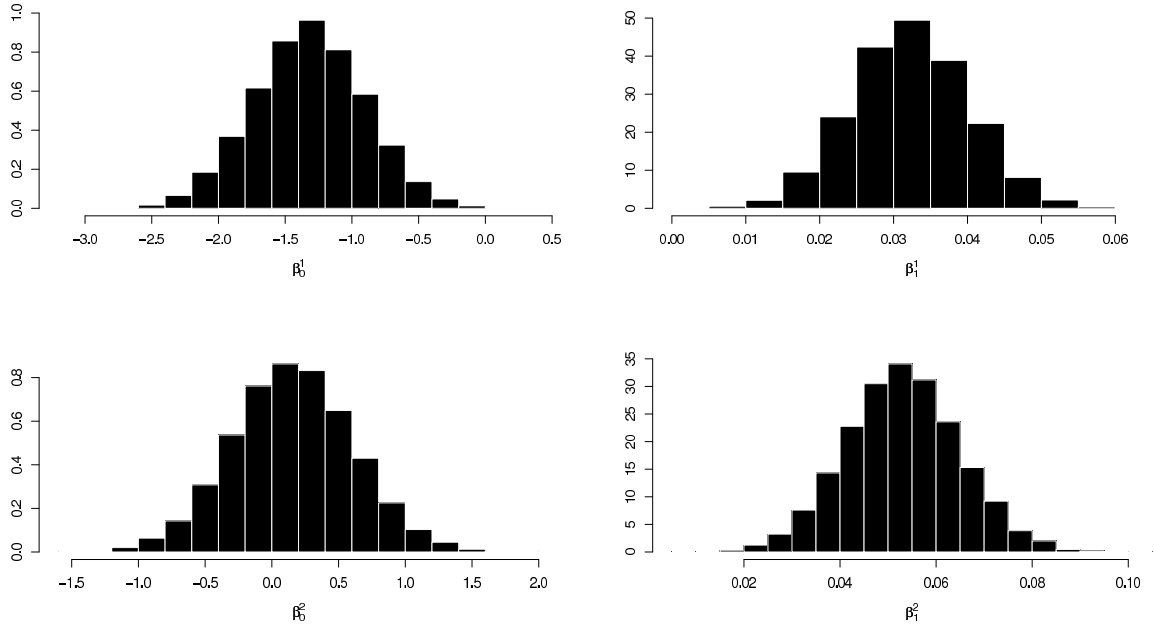


Figura 3.3: Densidades finales de  $\beta^1$  y  $\beta^2$  (por componente) para los datos de caracoles.

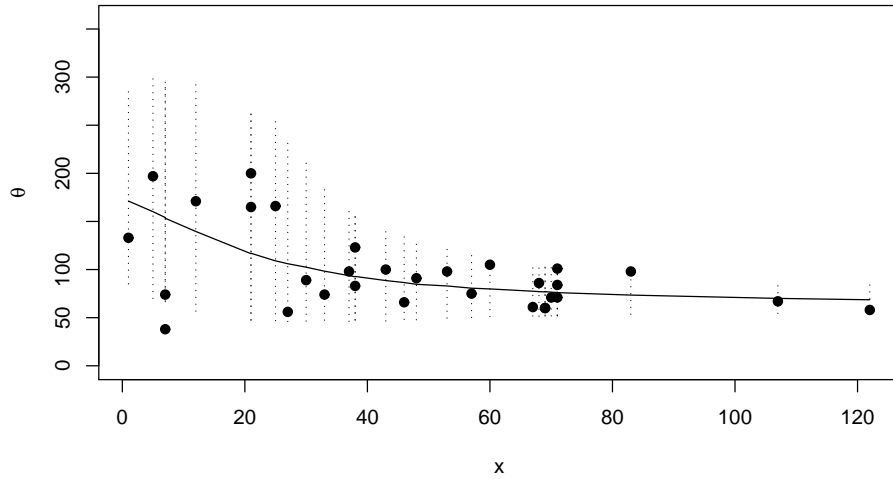


Figura 3.4: *Intervalos predictivos finales al 95% para los datos de caracoles. La línea continua representa la dirección media ajustada.*

Los resultados obtenidos concuerdan con los análisis previos de estos datos, ver Fisher y Lee (1992) y Presnell *et al.* (1998). Específicamente, los estimadores puntuales de los coeficientes de regresión así como sus correspondientes varianzas resultan ser del mismo orden que los obtenidos por Presnell *et al.* (1998). Por lo anterior, de manera particular, los intervalos de credibilidad obtenidos al 95% ofrecen las mismas conclusiones en términos de la significancia de los parámetros de regresión, comparados con los intervalos clásicos obtenidos por Presnell *et al.* (1998). Adicionalmente, en la Figura 3.4 se muestran los intervalos predictivos finales al 95% para cada valor de  $x$ , así como la dirección media ajustada con el correspondiente modelo Normal proyectado.

Con el fin de ilustrar el análisis del problema de datos faltantes, supóngase que las respuestas  $\theta_{13} = 197$ ,  $\theta_{23} = 75$  y  $\theta_1 = 67$  con  $x_{13} = 5$ ,  $x_{23} = 57$  y  $x_1 = 107$ , respectivamente, son datos faltantes. Utilizando los resultados de las secciones previas



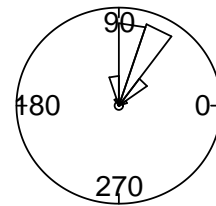
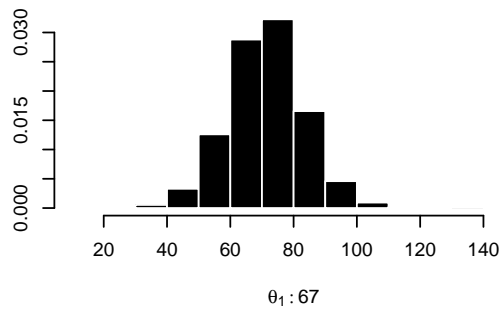
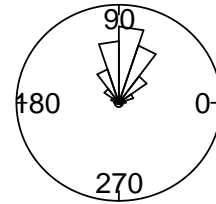
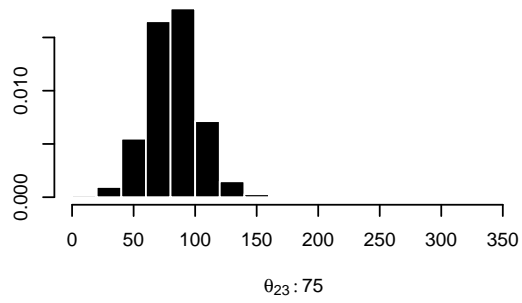
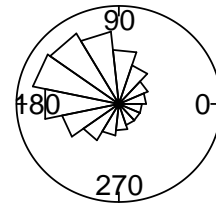
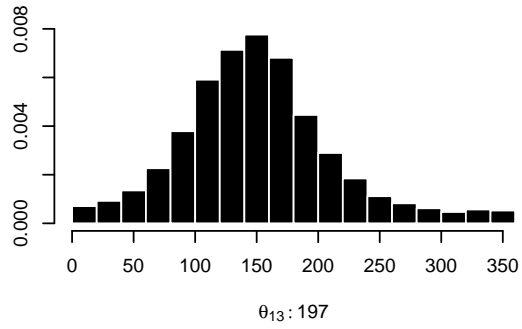


Figura 3.5: *Distribuciones finales de los valores faltantes  $\theta_{13}$  (en  $x = 5$ ),  $\theta_{23}$  (en  $x = 57$ ) y  $\theta_1$  (en  $x = 107$ ), para los datos de caracoles.*

se obtuvieron las distribuciones finales para  $\theta_{13}$ ,  $\theta_{23}$  y  $\theta_1$ . Estas distribuciones finales se muestran en la Figura 3.5. Se puede notar que la metodología propuesta produce inferencias apropiadas para los datos faltantes.

Model	$L_k$
$M_0 :$ $\mu_1 = 1$ $\mu_2 = 1$	$3.597 \times 10^{-15}$
$M_1 :$ $\mu_1 = 1$ $\mu_2 = 1 + x$	$1.138 \times 10^{-9}$
$M_2 :$ $\mu_1 = 1 + x$ $\mu_2 = 1 + x$	$1.219 \times 10^{-9}$
$M_3 :$ $\mu_1 = 1 + x$ $\mu_2 = 1 + x + x^2$	$2.854 \times 10^{-9}$
$M_4 :$ $\mu_1 = 1 + x + x^2$ $\mu_2 = 1 + x + x^2$	$3.52 \times 10^{-9}$
$M_5 :$ $\mu_1 = 1 + x + x^2 + x^3$ $\mu_2 = 1 + x + x^2 + x^3$	$6.785 \times 10^{-13}$

Tabla 3.1: *Medidas predictivas para selección de covariables para los datos de caracoles.*

Finalmente, con el objetivo de ilustrar el criterio propuesto por Geisser y Eddy (1979) para el caso de selección de covariables, se consideraron seis modelos. Bajo cada modelo propuesto  $M_k$ , se generó una muestra de la distribución final de  $\mathbf{B} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2]$ . Posteriormente, para cada dato  $\theta_i$ , se calculó la respectiva distribución predictiva de Monte Carlo (3.4) y se obtuvo la medida  $L_k$ . En la Tabla 3.1 se muestran los seis modelos así como el valor de la medida predictiva asociada para cada uno de ellos.

El correspondiente análisis sugiere que el modelo  $M_4$  ofrece mejores predicciones que cualquiera de los otros modelos bajo consideración.

**Ejemplo 3.4.** La dirección del viento puede tener un efecto significativo sobre los niveles de contaminantes del aire tales como el ozono (ver, por ejemplo, Jammalamadaka y Lund, 2006). Por esta razón, es importante contar con formas apropiadas de modelar las direcciones del viento, quizá como una etapa preliminar en el análisis de datos ambientales.

La Secretaría del Medio Ambiente de la Ciudad de México, a través de su sistema de monitoreo atmosférico (SMA), registra datos sobre el clima y ciertos contaminantes. Específicamente, el subsistema REDMET registra variables meteorológicas. REDMET consiste de 16 estaciones de monitoreo. La estación denominada EAC está localizada dentro del campus de la FES-Acatlán de la Universidad Nacional Autónoma de México (UNAM). Este campus se encuentra localizado en una zona de viviendas y servicios.

En este ejemplo, se consideró una muestra aleatoria estratificada de tamaño 480 tomada de la primera mitad del año 2007. Los datos fueron obtenidos del sitio web del SMA (REDMET, 2009). Se tomaron datos-hora sólo de la estación de monitoreo EAC. La base de datos se conformó con las siguientes variables: fecha, tiempo sobre un reloj de 24 horas (HRS), porcentaje de humedad relativa (RH), temperatura medida en grados Celsius (TMP), dirección azimuth del viento (WD) y velocidad del viento medida en metros por segundo (WSP). La variable respuesta fue WD.

Para modelar estos datos, se consideró una distribución Normal proyectada con  $\boldsymbol{\mu} = (\mu_1, \mu_2)'$  definida por

$$\begin{aligned}\mu_1 &= \beta_0^1 + \beta_1^1 RH + \beta_2^1 TMP + \beta_3^1 WSP + \beta_4^1 \cos(HRS) + \beta_5^1 \sin(HRS) \\ \mu_2 &= \beta_0^2 + \beta_1^2 RH + \beta_2^2 TMP + \beta_3^2 WSP + \beta_4^2 \cos(HRS) + \beta_5^2 \sin(HRS).\end{aligned}$$

Se uso una distribución inicial vaga con  $\boldsymbol{\beta}_0^j = \mathbf{0}$ ,  $\boldsymbol{\Lambda}_0^j = \text{Diag}(1, 1, 1, 1, 1, 1) * 0.0001$ , para  $j = 1, 2$  y  $\mathbf{r} = (1, \dots, 1)$ . Los intervalos de credibilidad finales al 95% para los vectores  $\boldsymbol{\beta}^1$  y  $\boldsymbol{\beta}^2$  se muestran en la Tabla 3.2. Los resultados sugieren que los efectos de  $\sin(HRS)$  (para  $\mu_1$ ) y  $RH$  (para  $\mu_2$ ) no son significativos.

	$\boldsymbol{\beta}^1$	$\boldsymbol{\beta}^2$
Intercepto	(-5.873462, -3.732854)	(-3.579.039, -1.139991)
$RH$	(0.0190294, 0.0373956)	<b>(-0.00663, 0.01327)</b>
$TMP$	(0.1244622, 0.2047609)	(0.0686012, 0.1560350)
$WSP$	(0.1523422, 0.4533813)	(-0.494002, -0.197336)
$\sin(HRS)$	<b>(-0.29125, 0.22229)</b>	(-0.866019, -0.347341)
$\cos(HRS)$	(-0.571925, -0.202993)	(-1.127180, -0.734448)

Tabla 3.2: *Intervalos finales de credibilidad al 95% para cada uno de los componentes de  $\boldsymbol{\beta}^1$  y  $\boldsymbol{\beta}^2$ .*

Adicionalmente, se calcularon los logaritmos de la medida predictiva asociada,  $L_k$ , para el modelo original ( $M_1$ ) así como para el modelo ( $M_2$ ) obtenido al omitir las covariables  $\sin(HRS)$  (para  $\mu_1$ ) y  $RH$  (para  $\mu_2$ ). Los resultados mostrados en la Tabla 3.3 también indican que el modelo más simple,  $M_2$ , es preferible, dado que  $L_1 < L_2$ .

Modelo		$\ln(L_k)$
$M_1 :$	$\mu_1 = \beta_0^1 + \beta_1^1 RH + \beta_2^1 TMP + \beta_3^1 WSP +$ $\beta_4^1 \cos(HRS) + \beta_5^1 \sin(HRS).$	-624.2146
	$\mu_2 = \beta_0^2 + \beta_1^2 RH + \beta_2^2 TMP + \beta_3^2 WSP +$ $\beta_4^2 \cos(HRS) + \beta_5^2 \sin(HRS).$	
$M_2 :$	$\mu_1 = \beta_0^1 + \beta_1^1 RH + \beta_2^1 TMP + \beta_3^1 WSP +$ $\beta_4^1 \cos(HRS).$	-621.6642
	$\mu_2 = \beta_0^2 + \beta_1^2 TMP + \beta_2^2 WSP + \beta_3^2 \cos(HRS) +$ $\beta_4^2 \sin(HRS).$	

Tabla 3.3: *Medidas predictivas para la selección de covariables, datos REDMET.*

### 3.5 Discusión

En esta parte del trabajo de investigación se presenta un análisis Bayesiano completo de un modelo de regresión para datos circulares basado en la distribución Normal proyectada. Aunque la versión de la Normal proyectada analizada no es la más general, es bastante flexible y no pierde aplicabilidad práctica comparada con los modelos de regresión que suponen una distribución von Mises para la variable respuesta. Además, esta metodología se puede llevar a cabo de manera relativamente simple vía el muestreo de Gibbs. Se debe resaltar que el problema de datos faltantes y de selección de modelos se pueden analizar de una manera natural y simple bajo el análisis Bayesiano propuesto. Adicionalmente, a diferencia de la mayoría de los modelos de regresión para datos circulares disponibles actualmente, la propuesta que se introduce en este trabajo no es

compleja de implementar computacionalmente y puede producir directamente inferencias sobre una variedad de cantidades de interés, incluyendo aquellas requeridas para predicción. Finalmente, aunque sólo se consideró el caso de una respuesta circular, la metodología propuesta se puede extender al caso de una respuesta direccional en dimensiones mayores, lo anterior, de manera más natural comparada con los procedimientos alternativos existentes.

# Capítulo 4

## Un Modelo Longitudinal para Datos Circulares

### 4.1 Introducción a los estudios longitudinales

Los estudios longitudinales presentan varias ventajas en la modelación de datos tipo panel. Por ejemplo, permiten al investigador analizar por separado los cambios en el tiempo dentro de los individuos de los cambios entre sujetos en la misma cohorte. Además, los estimadores de las tendencias individuales pueden ser usados para entender de mejor manera la heterogeneidad en la población y los factores de cambio a nivel de los individuos.

A pesar de las múltiples ventajas, el análisis de datos longitudinales presenta sus propios retos. En un estudio longitudinal, en principio, las observaciones no son independientes. Por lo anterior, se deben considerar métodos de análisis que tomen en cuenta la estructura de dependencia. Los métodos analíticos para este tipo de datos no están bien desarrollados del todo. Particularmente, cuando se trabaja con modelos que permiten distribuciones de probabilidad más generales para las medidas repetidas.

Una dificultad adicional que aparece en este tipo de estudios es la presencia de datos faltantes, lo cual hace más difícil su análisis. Otra característica peculiar de los datos longitudinales es que no sólo la variable respuesta cambia sobre el tiempo, sino que el valor de las covariables independientes o predictores también pueden cambiar en el tiempo. El tratamiento de las covariables que varían en el tiempo permite realizar inferencias estadísticas sobre la relación dinámica entre la variable respuesta y los predictores. Sin embargo, el precio resultante es la complejidad adicional en la modelación estadística.

#### 4.1.1 Diversos enfoques

En la literatura se han propuesto varios enfoques para el análisis de datos longitudinales en el caso de una respuesta escalar. En contraste, las propuestas para modelar relaciones de dependencia cuando las medidas repetidas son datos circulares son limitadas.

A continuación se presenta una exposición general de los diversos enfoques para el análisis de datos longitudinales cuando la respuesta es escalar. Lo anterior, con la finalidad de tener un marco de referencia para el análisis de datos circulares longitudinales. Para una revisión más detallada el lector se puede referir, por ejemplo, a Diggle *et al.* (1994) y de manera más reciente a Hedeker y Gibbons (2006).

La primera aproximación involucra la reducción de las medidas repetidas en una variable resumen. De hecho, después de esta reducción este enfoque es estrictamente no longitudinal ya que se obtiene una única medición por sujeto. Quizá el ejemplo más remoto del análisis de datos longitudinales, bajo este enfoque, fue presentado por Student (1908) en su ilustración de la prueba  $t$ .

Un segundo enfoque, tal vez el más simple y más restrictivo, es el modelo de análisis de varianza (ANOVA) para medidas repetidas, Winer (1971). Este modelo asume varian-



zas y covarianzas constantes en el tiempo. Tal supuesto puede tener poca validez para datos longitudinales, ya que a menudo las varianzas se incrementan con el tiempo debido a que algunos sujetos responden y otros no. Además, las covarianzas para ocasiones próximas pueden resultar más grandes que las covarianzas para ocasiones distantes.

Por su parte, los modelos de análisis de varianza multivariados (MANOVA) también han sido propuestos para el análisis de datos longitudinales (ver Bock, 1975). En el caso multivariado, las observaciones repetidas son generalmente transformadas en coeficientes polinomiales ortogonales y estos coeficientes son usados como una respuesta multivariada en un MANOVA. La principal desventaja de este enfoque es que no permite datos faltantes o diferentes ocasiones de medición para distintos sujetos.

Los modelos de patrones de covarianza (Jennrich y Schluchter, 1986) también se pueden utilizar para analizar datos longitudinales. En este enfoque, la matriz de varianzas-covarianzas de las observaciones de respuestas repetidas se modela directamente. Aquí, no existe un intento en distinguir entre la variación dentro de los sujetos y la variación entre sujetos, como en el caso de los modelos de regresión de efectos mixtos.

Los modelos de ecuaciones estimadoras generalizadas (GEE) son a menudo usados como una alternativa general y computacionalmente conveniente a los modelos de regresión de efectos mixtos. La desventaja de estos modelos es que el tratamiento de datos faltantes es limitado. Lo anterior restringe, en cierto sentido, a los modelos GEE en el caso de datos longitudinales incompletos.

Finalmente, los modelos de regresión de efectos mixtos actualmente son ampliamente utilizados en el análisis de datos longitudinales. Estos modelos pueden ser aplicados tanto a respuestas continuas distribuidas normalmente como a respuestas categóricas

y a otras con distribución diferente a la normal. Los modelos de regresión de efectos mixtos son bastante robustos a datos faltantes y a mediciones espaciadas de manera no regular. Además, pueden tratar de manera adecuada tanto covariables cuyo valor cambia en el tiempo como covariables invariantes. Por las características anteriores, los modelos de regresión de efectos mixtos se encuentran dentro de los métodos más generales para el análisis de datos longitudinales. Estos modelos son algunas veces denominados métodos de “verosimilitud-completa” debido a que hacen uso de toda la información disponible de cada sujeto en el estudio. Sin embargo, la desventaja de estos métodos es que son computacionalmente más demandantes que los denominados métodos de “pseudo-verosimilitud” tales como los modelos GEE.

La metodología introducida en este trabajo de investigación para el análisis longitudinal de datos circulares está basada en una distribución normal bivariada bajo proyecciones, la distribución Normal proyectada. En esta propuesta, cada componente del modelo normal proyectado se especifica por medio de un modelo lineal de efectos mixtos. Por este motivo y con la finalidad de comprender un poco más este tipo de modelos, en la siguiente sección se discuten algunos modelos pertenecientes a esta familia.

#### **4.1.2 Modelos de regresión de efectos-mixtos**

En la literatura se han ofrecido y desarrollado diversas variantes de los modelos de regresión de efectos mixtos; tales propuestas se conocen bajo diferentes nombres. Por ejemplo, modelos de dos estados (Laird y Ware, 1982), modelos de componentes de varianza (Dempster *et al.*, 1981), modelos multinivel (Goldstein, 1995), modelos lineales jerárquicos (Raudenbush y Bryk, 2002), modelos de coeficientes aleatorios (de Leeuw y Kreft, 1986), modelos mixtos (Wolfinger, 1993), etc. Se debe mencionar que, desde una

perspectiva Bayesiana, las variantes anteriores se pueden especificar y modelar teóricamente como familias de modelos jerárquicos. Una característica básica de todas las propuestas anteriores es la inclusión de efectos aleatorios individuales dentro de los modelos de regresión. Lo anterior, con el fin de considerar la influencia específica de los sujetos en las observaciones repetidas. De esta manera, estos efectos aleatorios individuales describen la tendencia particular de cada sujeto a través del tiempo y explican la estructura de dependencia de los datos longitudinales.

En este trabajo se utilizarán de manera indistinta los términos de modelos de efectos mixtos o modelos mixtos para referirnos a la clase de modelos que consideran efectos aleatorios individuales en su especificación.

### **Un modelo de intercepto aleatorio**

Para facilitar la introducción de los modelos mixtos considérese el modelo de regresión simple para la medición  $y$  del individuo  $i$ ,  $i = 1, \dots, N$ , en la ocasión o tiempo  $j$ ,  $j = 1, \dots, n_i$ , dado por

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \varepsilon_{ij}. \quad (4.1)$$

En este modelo de regresión los errores  $\varepsilon_{ij}$  se consideran distribuidos normal e *independientemente* con media cero y varianza  $\sigma^2$ . Este modelo (4.1) es inadecuado para el estudio de datos longitudinales porque las respuestas  $y$  son observadas repetidamente sobre el mismo individuo, por lo que es más razonable considerar que los errores dentro de un individuo estén correlacionados en cierto grado. Además, el modelo previo determina que el cambio a través del tiempo es el mismo para todos los individuos en vista de que los parámetros del modelo,  $\beta_0$  y  $\beta_1$ , no varían de individuo a individuo. Por las razones

anteriores, es útil incluir efectos específicos-individuales en el modelo que contemplen la dependencia de los datos y permitan diferenciar tendencias en el tiempo para individuos diferentes. Esto es precisamente lo que hacen los modelos de efectos mixtos.

Una extensión simple del modelo de regresión (4.1) que permite tener influencias particulares sobre las medidas repetidas de cada individuo está dada por

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + b_{0i} + \varepsilon_{ij},$$

donde  $b_{0i}$  representa este comportamiento específico de las observaciones repetidas para el individuo  $i$ . Hay que notar que si no se tienen influencias individuales sobre las medidas repetidas, entonces todos los términos  $b_{0i}$ ,  $i = 1, \dots, N$  serán iguales a cero.

Para entender de mejor manera cómo este modelo caracteriza la influencia individual sobre las observaciones, es útil emplear una representación jerárquica del modelo.

Nivel-1 o intra-sujetos,

$$y_{ij} = v_{0i} + v_{1i} t_{ij} + \varepsilon_{ij}, \tag{4.2}$$

Nivel-2 o entre-sujetos,

$$v_{0i} = \beta_0 + b_{0i},$$

$$v_{1i} = \beta_1.$$

Aquí, el Nivel-1 indica que la respuesta del individuo  $i$  al tiempo  $j$  es influenciada por su nivel inicial  $v_{0i}$  y la tendencia en el tiempo, o pendiente,  $v_{1i}$ . El Nivel-2 indica que el nivel inicial del individuo  $i$  es determinado por el valor poblacional  $\beta_0$  más una única contribución  $b_{0i}$  para ese individuo. Así, cada individuo tiene su propio nivel inicial. Este modelo también indica que cada pendiente individual es la misma, todas iguales a la pendiente poblacional  $\beta_1$ .

Otra manera de pensar lo anterior es que cada línea de tendencia individual es paralela a la tendencia poblacional determinada por  $\beta_0$  y  $\beta_1$  y, la diferencia entre cada tendencia individual y la tendencia poblacional está dada por  $b_{0i}$ . Debido a lo anterior, este modelo es algunas veces llamado *modelo de intercepto aleatorio*.

Desde que los individuos en una muestra son típicamente pensados como representantes de una gran población, los efectos específicos-individuales  $b_{0i}$  son tratados como efectos aleatorios. Es decir,  $b_{0i}$  son considerados representativos de una distribución de efectos individuales en la población. En este caso, se asume que los  $b_{0i}$  siguen una distribución Normal y son independientes de los errores  $\varepsilon_{ij}$ , los cuales también se asumen con distribución Normal.

### **Un modelo de intercepto y pendiente aleatoria**

Un modelo más realista para el análisis de datos longitudinales es un modelo mixto que permite que tanto el intercepto como la tendencia en el tiempo varíen por individuo. Para lo anterior, se requiere que el Nivel-1 del modelo sea como en (4.2),

$$y_{ij} = v_{0i} + v_{1i}t_{ij} + \varepsilon_{ij},$$

pero el Nivel-2 sea definido como

$$\begin{aligned} v_{0i} &= \beta_0 + b_{0i}, \\ v_{1i} &= \beta_1 + b_{1i}. \end{aligned} \tag{4.3}$$

En este modelo,  $\beta_0$  es el intercepto poblacional general,  $\beta_1$  es la pendiente poblacional general,  $b_{0i}$  es la desviación del intercepto del individuo  $i$  y,  $b_{1i}$  es la desviación de la pendiente del sujeto  $i$ . Con dos efectos específicos-individuales, la distribución poblacional que se asume para las desviaciones de la pendiente e intercepto es una dis-

tribución normal bivariada  $N(\mathbf{0}, \mathbf{\Omega})$ , donde  $\mathbf{\Omega}$  es la matriz de precisión asociada a los efectos aleatorios.

Se debe notar que al combinar las ecuaciones definidas en (4.2) y (4.3) se obtienen las siguientes ecuaciones equivalentes para el modelo mixto,

$$\begin{aligned} y_{ij} &= v_{0i} + v_{1i}t_{ij} + \varepsilon_{ij} \\ y_{ij} &= (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + \varepsilon_{ij} \\ y_{ij} &= (\beta_0 + \beta_1 t_{ij}) + (b_{0i} + b_{1i} t_{ij}) + \varepsilon_{ij} \end{aligned}$$

En términos matriciales, la última ecuación se pueden ver como

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i,$$

la cual es la representación general de un modelo lineal de efectos mixtos.

## 4.2 Especificación de un modelo mixto

Laird y Ware (1982) proponen un modelo de efectos mixtos de dos-estados que incluye como casos especiales a los modelos para datos longitudinales y a los modelos para curvas de crecimiento. La especificación del modelo es la siguiente.

Para cada individuo, se tiene

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i \\ \mathbf{b}_i &\sim N_q(\mathbf{0}, \mathbf{\Omega}), \end{aligned}$$

donde  $\mathbf{Y}_i$  son vectores de  $n_i \times 1$  que contienen las observaciones del  $i$ -ésimo individuo, y los  $\boldsymbol{\varepsilon}_i$  son vectores de errores independientes de la misma dimension con distribución  $N_{n_i}(\mathbf{0}, \mathbf{R}_i)$ ,  $i = 1, \dots, N$ . En este modelo mixto,  $\mathbf{X}_i$  es una matriz de diseño ( $n_i \times p$ )

que contiene los valores de las covariables y  $\boldsymbol{\beta}$  es el correspondiente vector ( $p \times 1$ ) de efectos fijos. En contraste,  $\mathbf{Z}_i$  es una matriz de diseño ( $n_i \times q$ ), y  $\mathbf{b}_i$  es un vector ( $q \times 1$ ) de efectos aleatorios específicos de cada sujeto. En estos modelos, los  $\mathbf{b}_i$  modelan el comportamiento particular debido a los sujetos y le permiten al modelo capturar la dependencia entre las observaciones dentro del  $i$ -ésimo individuo.

Una simplificación del modelo se da cuando  $\mathbf{R}_i = \mathbf{I}$ . Laird y Ware (1982) denominan a este modelo *modelo de independencia condicional*, ya que implica que las  $n_i$  respuestas de cada individuo  $i$  son independientes, condicionalmente en  $\mathbf{b}_i$  y  $\boldsymbol{\beta}$ . Aplicaciones de este tipo de modelos se pueden encontrar, por ejemplo, en Laird y Ware (1982) y en Chib *et al.* (1999).

Desde una perspectiva Bayesiana la especificación jerárquica del modelo de Laird y Ware (1982) se puede completar definiendo una distribución inicial para los parámetros  $\boldsymbol{\beta}$  y  $\boldsymbol{\Omega}$  y  $\mathbf{R}_i$ . Para el modelo de independencia condicional se presentan a continuación algunos resultados.

Sea  $\mathbf{D}_n = \{\mathbf{Y}_1, \dots, \mathbf{Y}_N\}$  una m.a. del modelo  $N_{n_i}(\cdot | \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i, \mathbf{I})$ . Entonces, la función de verosimilitud está dada por

$$L(\boldsymbol{\beta}, \{\mathbf{b}_i\}, \boldsymbol{\Omega} | \mathbf{D}_n) \propto |\mathbf{I}|^{N/2} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)\right\}.$$

**Proposición 4.2.1** Sea  $\mathbf{D}_n = \{\mathbf{Y}_1, \dots, \mathbf{Y}_N\}$  una m.a. del modelo  $N_{n_i}(\cdot | \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i, \mathbf{I})$ . Si se considera la siguiente especificación inicial conjugada

$$f(\boldsymbol{\beta} | A) = N_p(\boldsymbol{\beta} | \mathbf{0}, A), \\ f(\mathbf{b}_i | \boldsymbol{\Omega}) = N_q(\mathbf{b}_i | \mathbf{0}, \boldsymbol{\Omega}), \quad \text{para cada } i = 1, \dots, N, \\ f(\boldsymbol{\Omega} | v, B) = Wi(\boldsymbol{\Omega} | v, B), \quad \text{con } E(\boldsymbol{\Omega}) = vB^{-1},$$

con  $A$ ,  $v$  y  $B$  parámetros específicos de cada una de las densidades iniciales. Entonces, la correspondiente distribución final conjunta (obtenida al multiplicar las densidades iniciales y la función de verosimilitud) está dada por

$$f(\boldsymbol{\beta}, \{\mathbf{b}_i\}, \boldsymbol{\Omega} | \mathbf{D}_n) \propto |\boldsymbol{\Omega}|^{(v-q-1+N)/2} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\ \left. - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i - \frac{1}{2} \text{tr}(B \boldsymbol{\Omega})\right\}.$$

**Proposición 4.2.2** Sea  $\mathbf{D}_n = \{\mathbf{Y}_1, \dots, \mathbf{Y}_N\}$  una m.a. del modelo  $N_n(\cdot | \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i, \mathbf{I})$ . Si se considera la especificación inicial dada en la Proposición 4.2.1, entonces las correspondientes densidades condicionales finales completas están dadas por

- i.  $f(\boldsymbol{\beta} | \{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta} | C^{-1} \sum_{i=1}^N \mathbf{X}_i' \mathbf{e}_i, C)$ .
- ii.  $f(\mathbf{b}_i | \{\boldsymbol{\beta}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_q(\mathbf{b}_i | D_i^{-1} \mathbf{Z}_i' \tilde{\mathbf{e}}_i, D_i) \quad \forall i = 1, \dots, N$ .
- iii.  $f(\boldsymbol{\Omega} | \{\mathbf{b}_i\}, \mathbf{D}_n) = Wi(v + N, B + \sum_{i=1}^N \mathbf{b}_i \mathbf{b}_i')$ ,

donde

$$\begin{aligned} C &= \sum_{i=1}^N \mathbf{X}_i' \mathbf{X}_i + A, \\ \mathbf{e}_i &= \mathbf{Y}_i - \mathbf{Z}_i \mathbf{b}_i, \\ D_i &= \mathbf{Z}_i' \mathbf{Z}_i + \boldsymbol{\Omega}, \\ \tilde{\mathbf{e}}_i &= \mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}_i \end{aligned}$$

DEMOSTRACIÓN: Ver Apéndice A.

## 4.3 Un modelo para datos longitudinales circulares

### 4.3.1 Antecedentes

El tratamiento de datos longitudinales para una respuesta escalar ha sido ampliamente analizado en la literatura. Sin embargo, como se mencionó en las secciones previas, el estudio de este tipo de dependencias es complejo. El caso de datos longitudinales cuando se tiene una respuesta direccional ha sido poco estudiado. Lo anterior se debe,



entre otras cosas, al problema que representa trabajar con distribuciones de probabilidad asociadas a este tipo de datos.

En la actualidad existe un vacío en la literatura que limita el establecimiento de un marco general para el análisis longitudinal de datos direccionales. Los datos circulares longitudinales han sido estudiados en forma clásica usando enfoques como el método de ecuaciones estimadoras generalizadas (GEE) propuesto por Liang y Zeger (1986) para datos lineales. Específicamente, Artes *et al.* (2000) derivan ecuaciones estimadoras para los parámetros de una familia de distribuciones circulares con dos parámetros. De manera particular, exhiben el caso para un modelo mixto. Utilizan métodos de pseudo-verosimilitud y obtienen estimadores asintóticos para los parámetros involucrados. Por su parte, Artes y Jørgensen (2000) extienden los métodos GEE a la clase de modelos de dispersion (Jørgensen, 1997a y 1997b) y aplican este enfoque para modelar datos longitudinales circulares. Artes y Jørgensen (2000) presentan un estudio de simulación para un modelo que considera sólo la dirección media y una única covariable. Señalan que en ciertas situaciones su propuesta puede tener problemas de convergencia. Concluyen que para obtener un buen desempeño su método se requiere de una alta correlación de las observaciones o de un tamaño de muestra grande.

Los procedimientos mencionados sufren de las limitaciones intrínsecas de los métodos GEE que los hacen poco flexibles para realizar inferencias. Estas limitaciones incluyen la dificultad para el ajuste y comparación de modelos, problemas en el cálculo de estimadores debido a los problemas de convergencia de los métodos iterativos empleados, etc. Lo anterior, debido en parte a que estos métodos carecen de una estructura paramétrica completa. Estos problemas hacen que estas propuestas resulten poco generales para su

aplicación en contextos como datos faltantes y selección de modelos, entre otros.

Por su parte, Song (2007) utiliza un enfoque de modelos lineales generalizados donde la componente aleatoria pertenece a la familia de modelos de dispersión (Jørgensen, 1997a y 1997b). Song (2007) sugiere el empleo de métodos de pseudo-verosimilitud penalizada y máxima verosimilitud restringida para evitar las dificultades analíticas que aparecen debido a la no linealidad de las funciones de puntaje asociadas. Sin embargo, la implementación práctica sufre de limitaciones que no permite realizar inferencias para todos los parámetros involucrados en los modelos. Para una ilustración, ver Ejemplo 4.2.

### 4.3.2 El modelo longitudinal propuesto

Para facilitar la exposición, se supone  $n_i = n \forall i = 1, \dots, N$ . Así, se tienen  $N$  sujetos o individuos en estudio con  $n$  observaciones angulares,  $\theta_j$ , para cada sujeto o individuo  $i$ .

Es decir,

$$\begin{pmatrix} \text{individuo} & & \text{observaciones} & & \\ & 1 & \theta_{11} & \dots & \theta_{1n} \\ & \vdots & \vdots & \vdots & \vdots \\ & i & \theta_{i1} & \dots & \theta_{in} \\ & \vdots & \vdots & \vdots & \vdots \\ & N & \theta_{N1} & \dots & \theta_{Nn} \end{pmatrix}$$

Con las mismas ideas del Capítulo 3, el primer paso para modelar este tipo de observaciones es proponer un modelo aumentado para estos datos vía la introducción de variables latentes  $R_{ij}$ , de manera que

$$\left( \begin{array}{cc} \text{individuo} & \text{observaciones} \\ 1 & \theta_{11} \quad \dots \quad \theta_{1n} \\ \vdots & \vdots \quad \vdots \quad \vdots \\ i & \theta_{i1} \quad \dots \quad \theta_{in} \\ \vdots & \vdots \quad \vdots \quad \vdots \\ N & \theta_{N1} \quad \dots \quad \theta_{Nn} \end{array} \right) \rightarrow \left( \begin{array}{cc} \text{individuo} & \text{observaciones} \\ 1 & \mathbf{Y}_{11} \quad \dots \quad \mathbf{Y}_{1n} \\ \vdots & \vdots \quad \vdots \quad \vdots \\ i & \mathbf{Y}_{i1} \quad \dots \quad \mathbf{Y}_{in} \\ \vdots & \vdots \quad \vdots \quad \vdots \\ N & \mathbf{Y}_{N1} \quad \dots \quad \mathbf{Y}_{Nn} \end{array} \right)$$

con  $\mathbf{Y}_{ij}$  vectores aleatorios con distribución normal bivariada.

Específicamente, la propuesta de modelo aumentado consiste en introducir variables latentes  $R_{ij}$  de tal manera que:

$$\mathbf{Y}_{ij} = \begin{pmatrix} Y_{ij}^I \\ Y_{ij}^{II} \end{pmatrix} = R_{ij} \times \begin{pmatrix} \cos \theta_{ij} \\ \text{sen } \theta_{ij} \end{pmatrix} \sim N_2(\boldsymbol{\mu}_{ij}, \mathbf{I}), \quad \begin{array}{l} i = 1, \dots, N \\ j = 1, \dots, n. \end{array}$$

donde  $R_{ij} = \|\mathbf{Y}_{ij}\|$ .

Adicionalmente, se propone la siguiente estructura para el vector de medias

$$\boldsymbol{\mu}_{ij} = \begin{pmatrix} \mu_{ij}^I \\ \mu_{ij}^{II} \end{pmatrix} = \mathbf{B}' \mathbf{x}_{ij} + \mathbf{Z}' \mathbf{b}_i,$$

donde

$$\mu_{ij}^k = (\mathbf{x}_{ij}^k)' \boldsymbol{\beta}^k + (\mathbf{z}_{ij}^k)' \mathbf{b}_i^k,$$

$$\begin{aligned} \forall k \in \{I, II\}, \\ i = 1, \dots, N, \text{ y} \\ j = 1, \dots, n. \end{aligned}$$

Aquí,  $\mathbf{B} = [\boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}]$  es la matriz de coeficientes del modelo y  $\mathbf{Z} = [\mathbf{b}^I, \mathbf{b}^{II}]$ . Por la discusión presentada en el Capítulo 3 cada uno de los componentes de  $\boldsymbol{\mu}_{ij}$  puede depender de diferentes subconjuntos de covariables, en cuyo caso los vectores  $\boldsymbol{\beta}^I$  y  $\boldsymbol{\beta}^{II}$  pueden tener diferentes dimensiones, al igual que los vectores  $\mathbf{b}^I$  y  $\mathbf{b}^{II}$ .

En lo que resta de este trabajo, a este modelo propuesto se le denominará modelo *longitudinal circular proyectado* (LCP). Se debe notar que la especificación del modelo LCP es equivalente a considerar un modelo lineal mixto en cada uno de los dos componentes del modelo. Es decir,

$$\begin{aligned} \mathbf{Y}_i^I &= \mathbf{X}_i^I \boldsymbol{\beta}^I + \mathbf{Z}_i^I \mathbf{b}_i^I + \boldsymbol{\varepsilon}_i^I. \\ \mathbf{Y}_i^{II} &= \mathbf{X}_i^{II} \boldsymbol{\beta}^{II} + \mathbf{Z}_i^{II} \mathbf{b}_i^{II} + \boldsymbol{\varepsilon}_i^{II}. \end{aligned}$$

donde los vectores  $\mathbf{Y}_i^I$  y  $\mathbf{Y}_i^{II}$  son vectores de dimensión  $n$ ,  $\forall i = 1, \dots, N$ .

Así, para los datos aumentados y considerando cada una de las dos componentes  $k \in \{I, II\}$ , la especificación jerárquica del modelo LCP es la siguiente.

- NIVEL-I. Para cada individuo  $i$ ,

$$\mathbf{Y}_i^k | \boldsymbol{\beta}^k, \{\mathbf{b}_i\}^k \sim N_n(\mathbf{X}_i^k \boldsymbol{\beta}^k + \mathbf{Z}_i^k \mathbf{b}_i^k, \mathbf{I}), \quad i = 1, \dots, N,$$

es decir, dado  $\boldsymbol{\beta}^k$  y  $\{\mathbf{b}_i\}^k$ ,

$$\mathbf{Y}_i^k = \mathbf{X}_i^k \boldsymbol{\beta}^k + \mathbf{Z}_i^k \mathbf{b}_i^k + \boldsymbol{\varepsilon}_i^k. \quad \forall i = 1, \dots, N,$$

donde  $\boldsymbol{\varepsilon}_i^k \sim N_n(\mathbf{0}, \mathbf{I})$ .

- NIVEL-II.  $\boldsymbol{\beta}^k$  y  $\mathbf{b}_i^k$  se consideran vectores independientes,  $\forall i = 1, \dots, N$ , con

$$\begin{aligned}\mathbf{b}_i^k | \boldsymbol{\Omega}^k &\sim N_q(\mathbf{0}, \boldsymbol{\Omega}^k) \quad \forall i = 1, \dots, N \\ \boldsymbol{\beta}^k &\sim N_{p_k}(\mathbf{0}, A^k).\end{aligned}$$

- NIVEL-III.

$$\boldsymbol{\Omega}^k \sim Wi(v^k, B^k), \quad v^k \geq q^k$$

donde  $q^k$  es la dimensión del vector  $\mathbf{b}_i^k$ . En esta parametrización  $E(\boldsymbol{\Omega}^k) = v^k (B^k)^{-1}$

### 4.3.3 Análisis del modelo

Para realizar inferencias sobre los parámetros del modelo LCP se considera la introducción de variables latentes apropiadas para poder definir una distribución conjunta aumentada con las matrices  $\mathbf{B}$  y  $\mathbf{b}$ . Esta distribución conjunta se construye de tal manera que se pueda simular de todas las densidades condicionales completas con el fin de poder implementar un muestreo de Gibbs. La construcción se presenta a continuación.

El Nivel-I del modelo LCP para cada observación  $ij$ -ésima se puede ver como

$$f(\mathbf{y}'_{ij} = r_{ij}(\cos \theta, \sin \theta) | \boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \mathbf{x}_{ij}^I, \mathbf{x}_{ij}^{II}, \mathbf{z}_{ij}^I, \mathbf{z}_{ij}^{II}) = N_2(\mathbf{y}_{ij} | \boldsymbol{\mu}_{ij}, \mathbf{I}).$$

Si se tuvieran los datos completos, es decir, si se pudieran observar realizaciones de las variables  $R_{ij}$ , se estaría en condiciones de realizar inferencias para los parámetros involucrados en el modelo. El problema es que sólo las direcciones  $\{\theta_{ij}\}$  son observables. Este problema se puede atacar introduciendo variables latentes  $R_{ij}$  mediante una transformación adecuada. Esto se presenta a continuación.

Si se introducen las variables latentes  $R_{ij}$ , a través de la transformación definida por

$$\mathbf{Y}_{ij} = R_{ij}(\cos \Theta_{ij}, \text{sen} \Theta_{ij})',$$

entonces, se puede obtener la densidad conjunta de  $(R_{ij}, \Theta_{ij})$ ,  $f(r_{ij}, \theta_{ij})$  (Ver Proposición (2.2.1)). Esta densidad está dada por

$$\begin{aligned} f((r_{ij}, \theta_{ij}) | \boldsymbol{\mu}_{ij}) &= (2\pi)^{-1} r_{ij} \\ &\times \exp\left\{-\frac{1}{2} [(r_{ij} \cos \theta_{ij} - \mu_{ij}^I)^2 + (r_{ij} \text{sen} \theta_{ij} - \mu_{ij}^{II})^2]\right\} \\ &= (2\pi)^{-1} r_{ij} \exp\left\{-\frac{1}{2} \|\boldsymbol{\mu}_{ij}\|^2\right\} \\ &\times \exp\left\{-\frac{1}{2} [r_{ij}^2 - 2 r_{ij} (\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij})]\right\} \end{aligned} \quad (4.4)$$

donde

$$\begin{aligned} \mathbf{v}'_{ij} &= (\cos \theta_{ij}, \text{sen} \theta_{ij}) \\ \|\boldsymbol{\mu}_{ij}\|^2 &= \sum_{k=1}^2 [(\mathbf{x}_{ij}^k)' \boldsymbol{\beta}^k + (\mathbf{z}_{ij}^k)' \mathbf{b}_i^k]^2. \end{aligned}$$

**Proposición 4.3.1** *Si la densidad conjunta de  $(R_{ij}, \Theta_{ij})$ ,  $f(r_{ij}, \theta_{ij})$ , está dada por (4.4),*

$$\boldsymbol{\mu}_{ij} = \begin{pmatrix} \mu_{ij}^I \\ \mu_{ij}^{II} \end{pmatrix} = \begin{pmatrix} (\mathbf{x}_{ij}^I)' \boldsymbol{\beta}^I + (\mathbf{z}_{ij}^I)' \mathbf{b}_i^I \\ (\mathbf{x}_{ij}^{II})' \boldsymbol{\beta}^{II} + (\mathbf{z}_{ij}^{II})' \mathbf{b}_i^{II} \end{pmatrix} \quad \begin{array}{l} i = 1, \dots, N, \\ j = 1, \dots, n, \end{array}$$

y  $\mathbf{v}'_{ij} = (\cos \theta_{ij}, \text{sen} \theta_{ij})$ , entonces, se tienen los siguientes resultados:

*i.*

$$f(\theta_{ij} | \boldsymbol{\mu}_{ij}, \mathbf{I}) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2} \|\boldsymbol{\mu}_{ij}\|^2\right\} \left[ 1 + \frac{\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}}{\phi(\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij})} \Phi(\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}) \right] I_{(0,2\pi]}(\theta_{ij}) I_{\mathbb{R}^2}(\boldsymbol{\mu}_{ij})$$

ii.

$$f(r_{ij}|\theta_{ij}, \boldsymbol{\mu}_{ij}, \mathbf{I}) = \frac{r_{ij} \exp\{-\frac{1}{2} [r_{ij}^2 - 2 (\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}) r_{ij}]\}}{1 + \frac{\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}}{\phi(\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij})} \Phi(\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij})} 1_{(0,\infty)}(r_{ij}) 1_{\mathbb{R}^2}(\boldsymbol{\mu}_{ij}).$$

iii.  $f(r_{ij}|\theta_{ij}, \boldsymbol{\mu}_{ij}, \mathbf{I})$  pertenece a una familia exponencial con parámetro canónico  $b_{ij} = \mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}$  y

$$E(R_{ij}|\Theta_{ij}, \boldsymbol{\mu}_{ij}, \mathbf{I}) = b_{ij} + \frac{\Phi(b_{ij})}{\phi(b_{ij}) + b_{ij} \Phi(b_{ij})}.$$

Donde, como antes,  $\phi$  y  $\Psi$  representan la función de densidad y de distribución de una v.a. normal estándar, respectivamente.

DEMOSTRACIÓN: Ver Proposición 2.3.1 en el Capítulo 2. □

## 4.4 Inferencias vía métodos MCCM

Una vez que se han completado (aumentado) los datos, vía la introducción de las variables latentes  $R_{ij}$ , para estar en condiciones de llevar a cabo inferencias para los parámetros del modelo LCP usando técnicas MCCM, se debe ser capaz de especificar todas las correspondientes densidades condicionales completas.

Las densidades condicionales finales completas, para todos los parámetros involucrados en el modelo LCP propuesto, se determinan en la siguiente proposición.

**Proposición 4.4.1** Sea  $\mathbf{D}_n = \{(r_{11}, \theta_{11}), \dots, (r_{Nn}, \theta_{Nn})\}$  un conjunto de observaciones del modelo

$$\mathbf{Y}_i^k | \boldsymbol{\beta}^k, \{\mathbf{b}_i\}^k \sim N_n(\mathbf{X}_i^k \boldsymbol{\beta}^k + \mathbf{Z}_i^k \mathbf{b}_i^k, \mathbf{I}), \quad \forall i = 1, \dots, N, \quad y \quad k \in \{I, II\}$$

Omitiendo el superíndice  $k$ , las densidades condicionales completas asociadas al modelo propuesto, para cada componente  $k \in \{I, II\}$ , son las siguientes.

- i.  $f(\boldsymbol{\beta}|\{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta}|C^{-1} \sum_{i=1}^N \mathbf{X}'_i \mathbf{e}_i, C)$ .
- ii.  $f(\mathbf{b}_i|\{\boldsymbol{\beta}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_q(\mathbf{b}_i|D_i^{-1} \mathbf{Z}'_i \tilde{\mathbf{e}}_i, D_i) \quad \forall i = 1, \dots, N$ .
- iii.  $f(\boldsymbol{\Omega}|\{\mathbf{b}_i\}, \mathbf{D}_n) = Wi(v + N, B + \sum_{i=1}^N \mathbf{b}_i \mathbf{b}'_i)$ .
- iv.  $f(r_{ij}|\{\theta_{ij}\}, \boldsymbol{\mu}_{ij}) \propto r_{ij} \exp\{-\frac{1}{2} [r_{ij}^2 - 2 (\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}) r_{ij}]\}$ .

donde  $C$ ,  $\mathbf{e}_i$ ,  $D_i$  y  $\tilde{\mathbf{e}}_i$  son, para cada componente  $k \in \{I, II\}$ , como en la Proposición 4.2.2

DEMOSTRACIÓN: Ver demostraciones de las Proposiciones 4.2.2 y 4.3.1. □

Se puede notar que no es difícil muestrear de  $f(\boldsymbol{\beta}|\{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n)$ ,  $f(\mathbf{b}_i|\{\boldsymbol{\beta}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n)$  y  $f(\boldsymbol{\Omega}|\{\mathbf{b}_i\}, \mathbf{D}_n)$ . Por otro lado, siguiendo las ideas del Capítulo 3, como las  $R_{ij}$  ( $\forall i = 1, \dots, N$  y  $\forall j = 1, \dots, n$ ) son condicionalmente independientes dado las  $\Theta_{ij}$ , se pueden generar realizaciones  $R_{ij}$  de  $f(r_{ij}|\theta_{ij}, \boldsymbol{\mu}_{ij})$  de tal forma que es factible simular una matriz aleatoria  $\mathbf{R}$  de  $f(\mathbf{R}|\{\theta_{ij}\}, \boldsymbol{\mu}_{ij})$ . De manera análoga a la propuesta del Capítulo 3, este último paso se lleva a cabo vía un algoritmo de Metropolis-Hastings.

Finalmente, se pueden usar las condicionales completas definidas en la Proposición (4.4.1) en un muestreo de Gibbs para obtener una muestra de la distribución final conjunta

$$f(\boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \boldsymbol{\Omega}^I, \boldsymbol{\Omega}^{II}, R_{ij} | \{\theta_{ij}\}), \quad (4.5)$$

En consecuencia, se está en condiciones de generar una muestra de las distribuciones finales de los parámetros de interés:  $\boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \boldsymbol{\Omega}^I$  y  $\boldsymbol{\Omega}^{II}$ . Un algoritmo que lleva a cabo esta metodología MCCM es:



#### ALGORITMO 1:

- Para cada componente  $k$ ,  $k \in \{I, II\}$ 
  1. Muestrear  $\beta^k$  de  $f(\beta^k | \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \Omega^I, \Omega^{II}, \mathbf{D}_n) = f(\beta^k | \{\mathbf{b}_i\}^k, \mathbf{D}_n)$ .
  2. Muestrear  $\mathbf{b}_i^k$  de  $f(\mathbf{b}_i^k | \beta^I, \beta^{II}, \Omega^I, \Omega^{II}, \mathbf{D}_n) = f(\mathbf{b}_i^k | \beta^k, \Omega^k, \mathbf{D}_n)$ ,  
para toda  $i = 1, \dots, N$ .
  3. Muestrear  $\Omega^k$  de  $f(\Omega^k | \beta^I, \beta^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \mathbf{D}_n) = f(\Omega^k | \{\mathbf{b}_i\}^k, \mathbf{D}_n)$ ,
- Muestrear  $R_{ij}$  de  $f(r_{ij} | \beta^I, \beta^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \Omega^I, \Omega^{II}, \{\theta_{ij}\})$   
 $\forall i = 1, \dots, N$  y  $\forall j = 1, \dots, n$ .
- Repetir los pasos anteriores utilizando los valores más recientes de las variables condicionadas.

## 4.5 Implementación

Por lo expuesto hasta el momento, parecería relativamente directo construir un algoritmo con el que se puedan obtener muestras de la distribución final conjunta (4.5) y realizar inferencias sobre los parámetros de interés. Sin embargo, se debe señalar que la implementación numérica del Algoritmo 1 requiere de varias consideraciones, entre las que se encuentran:

- i. La determinación del periodo de calentamiento del algoritmo.
- ii. La determinación del número de iteraciones (internas) del algoritmo de Metropolis-Hastings.
- iii. La evaluación de la convergencia de las cadenas simuladas.

Respecto al período de calentamiento se debe señalar que, en las iteraciones consideradas en este período, en lugar de simular de la densidad condicional de las variables  $R_{ij}$  se tomó la  $E(R_{ij}|\cdot) \forall i = 1, \dots, N, j = 1, \dots, n$ , definida en la Proposición (4.3.1). En lo referente al número de iteraciones internas del algoritmo de Metropolis-Hastings, después de varios análisis, se encontró que para algunos casos se pueden necesitar al menos 10 iteraciones en cada paso. Se analizó una estructura con las consideraciones anteriores, tanto para el período de calentamiento como para las iteraciones internas del algoritmo de Metropolis-Hastings, debido a que de otro modo el desempeño del algoritmo no era satisfactorio en términos de la convergencia. Aún con este esquema, la implementación del Algoritmo 1 es poco práctica, ya que se puede requerir de un número significativo de iteraciones para obtener resultados adecuados.

## 4.6 Simulación por Bloques

Algunos algoritmos que implementan métodos MCMC sufren de convergencia lenta cuando se trabajan modelos con un número considerable de parámetros y estos son altamente correlacionados *a posteriori*, o si la información en la verosimilitud es insuficiente para identificar completamente a los parámetros involucrados. Los modelos mixtos, que conforman cada componente del modelo LCP, no son la excepción. De manera particular, en el caso de los modelos lineales mixtos se han sugerido en la literatura diferentes

procedimientos que tratan de subsanar estos problemas. Gelfand *et al.* (1995) sugiere centrar los efectos aleatorios para reducir la correlación serial. Por otro lado, Vines *et al.* (1996) y Gilks y Roberts (1996) recomiendan reparametrizaciones muy generales para romper la correlación serial a través de la reducción del espacio de parámetros. Gelfand y Sahu (1999) imponen restricciones numéricas al final de cada iteración MCCM. Por su parte, Chib y Carlin (1999) proponen un procedimiento para actualizar en un único bloque los parámetros asociados al modelo lineal mixto y muestran que este procedimiento puede mejorar la convergencia de los procedimientos MCCM.

Para mejorar el desempeño del Algoritmo 1 en este trabajo se consideró una propuesta basada en la simulación por bloques. Esta consiste en simular en un grupo (bloque) los parámetros  $\boldsymbol{\beta}$  y  $\{\mathbf{b}_i\} \forall i = 1, \dots, N$  en cada uno de los dos componentes del modelo LCP. La siguiente proposición muestra las densidades condicionales completas para los parámetros  $\boldsymbol{\beta}^k$ ,  $k \in \{I, II\}$ , que se obtienen al considerar este bloqueo.

**Proposición 4.6.1** *Sea  $\mathbf{D}_n = \{(r_{11}, \theta_{11}), \dots, (r_{Nn}, \theta_{Nn})\}$  un conjunto de observaciones del modelo LCP*

$$\mathbf{Y}_i^k | \boldsymbol{\beta}^k, \{\mathbf{b}_i\}^k \sim N_n(\mathbf{X}_i^k \boldsymbol{\beta}^k + \mathbf{Z}_i^k \mathbf{b}_i^k, \mathbf{I}), \quad \forall i = 1, \dots, N, \quad y \quad k \in \{I, II\}$$

*Omitiendo el superíndice  $k$ , la densidad condicional completa para  $\boldsymbol{\beta}^k$ ,  $k \in \{I, II\}$ , está dada por*

$$f(\boldsymbol{\beta} | \boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta} | \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}_\beta),$$

*donde*

$$\begin{aligned} \boldsymbol{\mu}_\beta &= \boldsymbol{\Lambda}_\beta^{-1} \left( \sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{Y}_i \right), \\ \boldsymbol{\Lambda}_\beta &= \left( A + \sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right). \end{aligned}$$

*Las densidades condicionales para el resto de los parámetros del modelo son como en la Proposición (4.4.1). Es decir,*

$$f(\mathbf{b}_i | \{\boldsymbol{\beta}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_q(\mathbf{b}_i | D_i^{-1} \mathbf{Z}_i' \tilde{\mathbf{e}}_i, D_i) \quad \forall i = 1, \dots, N.$$

$$f(\boldsymbol{\Omega}|\{\mathbf{b}_i\}, \mathbf{D}_n) = Wi(v + N, B + \sum_{i=1}^N \mathbf{b}_i \mathbf{b}'_i).$$

$$f(r_{ij}|\{\theta_{ij}\}, \boldsymbol{\mu}_{ij}) \propto r_{ij} \exp\left\{-\frac{1}{2} [r_{ij}^2 - 2 (\mathbf{v}'_{ij} \boldsymbol{\mu}_{ij}) r_{ij}]\right\}.$$

donde  $\mathbf{v}'_{ij} = (\cos \theta_{ij}, \sin \theta_{ij})$ .

DEMOSTRACIÓN: Ver Apéndice A y Proposición (4.4.1). □

Un algoritmo que implementa esta metodología MCCM por bloqueo se puede especificar de la siguiente manera.

ALGORITMO 2 (GIBBS SAMPLER CON BLOQUEO):

- Para cada componente  $k$ ,  $k \in \{I, II\}$ ,
  1. Muestrear  $\boldsymbol{\beta}^k$  y  $\{\mathbf{b}_i\}$  de  $f(\boldsymbol{\beta}^k, \{\mathbf{b}_i\}^k | \boldsymbol{\Omega}^I, \boldsymbol{\Omega}^{II}, \mathbf{D}_n) = f(\boldsymbol{\beta}^k, \{\mathbf{b}_i\}^k | \boldsymbol{\Omega}^k, \mathbf{D}_n)$ ,  
muestreando
    - 1.1  $\boldsymbol{\beta}^k$  de  $f(\boldsymbol{\beta}^k | \boldsymbol{\Omega}^I, \boldsymbol{\Omega}^{II}, \mathbf{D}_n) = f(\boldsymbol{\beta}^k | \boldsymbol{\Omega}^k, \mathbf{D}_n)$ .
    - 1.2  $\mathbf{b}_i^k$  de  $f(\mathbf{b}_i^k | \boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \boldsymbol{\Omega}^I, \boldsymbol{\Omega}^{II}, \mathbf{D}_n) = f(\mathbf{b}_i^k | \boldsymbol{\beta}^k, \boldsymbol{\Omega}^k, \mathbf{D}_n)$ ,  
para toda  $i = 1, \dots, N$ .
  2. Muestrear  $\boldsymbol{\Omega}^k$  de  $f(\boldsymbol{\Omega}^k | \boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \mathbf{D}_n) = f(\boldsymbol{\Omega}^k | \{\mathbf{b}_i\}^k, \mathbf{D}_n)$ ,
- Muestrear  $R_{ij}$  de  $f(r_{ij} | \boldsymbol{\beta}^I, \boldsymbol{\beta}^{II}, \{\mathbf{b}_i\}^I, \{\mathbf{b}_i\}^{II}, \boldsymbol{\Omega}^I, \boldsymbol{\Omega}^{II}, \{\theta_{ij}\})$   
 $\forall i = 1, \dots, N$  y  $\forall j = 1, \dots, n$ .
- Repetir los pasos anteriores utilizando los valores más recientes de las variables condicionadas.

Aunque la diferencia del algoritmo anterior y el algoritmo 1 parece menor, en la implementación práctica se pueden observar beneficios al implementar este bloqueo en los métodos MCCM para el modelo LCP. Particularmente, de manera marginal este bloqueo mejora sustancialmente las inferencias para los modelos mixtos que conforman cada una de las dos componentes del modelo.

Hay que señalar que dependiendo de la especificación de las componentes del modelo LCP, los métodos propuestos en este trabajo, como cualquier otro procedimiento MCCM, pueden requerir de consideraciones distintas (en cuanto al número de iteraciones necesarias) para obtener resultados adecuados.

#### **4.6.1 Ejemplos**

Al igual que en el capítulo anterior, para los ejemplos de esta sección se utilizó el lenguaje y ambiente R (R Development Core Team, 2010) para simular los respectivos conjuntos de datos, así como, para implementar todos los análisis.

**Ejemplo 4.1.** En este ejemplo se considera una muestra longitudinal simulada que representa los datos de  $N = 60$  individuos observados en 5 ocasiones (tiempos). La muestra se obtuvo considerando la siguiente especificación del modelo LCP:

$$\begin{aligned} \mathbf{Y}_i^I | \boldsymbol{\beta}^I, & \sim N_5(\mathbf{X}_i^I \boldsymbol{\beta}^I, \mathbf{I}), \\ \mathbf{Y}_i^{II} | \boldsymbol{\beta}^{II}, \{\mathbf{b}_i\}^{II} & \sim N_5(\mathbf{X}_i^{II} \boldsymbol{\beta}^{II} + \mathbf{Z}_i^{II} \mathbf{b}_i^{II}, \mathbf{I}), \quad i = 1, \dots, 60, \end{aligned}$$

donde

$$\boldsymbol{\beta}^I = \begin{pmatrix} 100 \\ -4 \end{pmatrix}, \quad \boldsymbol{\beta}^{II} = \begin{pmatrix} 200 \\ -10 \end{pmatrix}.$$

$$\mathbf{b}_i^{II} | \boldsymbol{\Omega}^{II} \sim N_2(\mathbf{0}, \boldsymbol{\Omega}^{II}) \quad \forall i = 1, \dots, 60$$

con

$$(\boldsymbol{\Omega}^{II})^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}.$$

Para este ejemplo se tomó

$$X_i^I = X_i^{II} = Z_i^{II} = \begin{pmatrix} & \text{Tiempo} \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad \forall i = 1, \dots, 60.$$

La Figura 4.1 muestra el conjunto de datos que se obtiene bajo la especificación de este modelo LCP.

Para implementar la metodología propuesta en este capítulo se utilizó una distribución inicial vaga tomando  $A^I = \mathbf{0} = A^{II} = \mathbf{0}$ ,  $v^I = v^{II} = 2$  y  $B^I = B^{II} =$

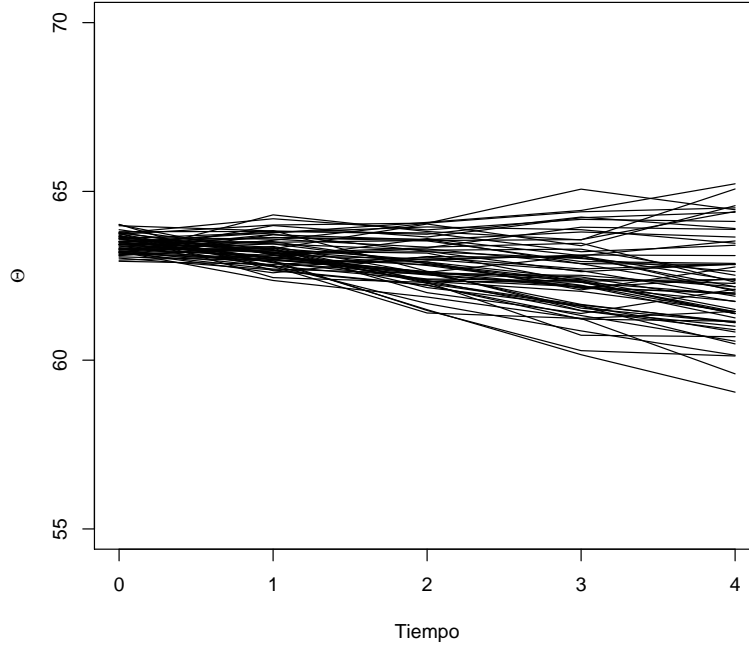


Figura 4.1: *Datos circulares longitudinales generados a partir del modelo LCP.*

Diag(0.001, 0.001). Los valores iniciales para el algoritmo de Gibbs fueron las matrices  $\mathbf{b} = \mathbf{0}$ ,  $\mathbf{\Omega}^{II} = \mathbf{I}$  y  $\mathbf{R} = \mathbf{1}$ . En la Figura 4.2 se presentan las distribuciones finales para los vectores  $\boldsymbol{\beta}^I$  y  $\boldsymbol{\beta}^{II}$ . Por su parte, en la Figura 4.3 se exhiben la distribuciones finales para los parámetros de la matriz de covarianzas asociadas a los efectos aleatorios. Es decir, para los elementos de la matriz

$$(\mathbf{\Omega}^{II})^{-1} = \begin{pmatrix} (\sigma_1^2)^{II} & \sigma_{12}^{II} \\ \sigma_{12}^{II} & (\sigma_2^2)^{II} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}.$$

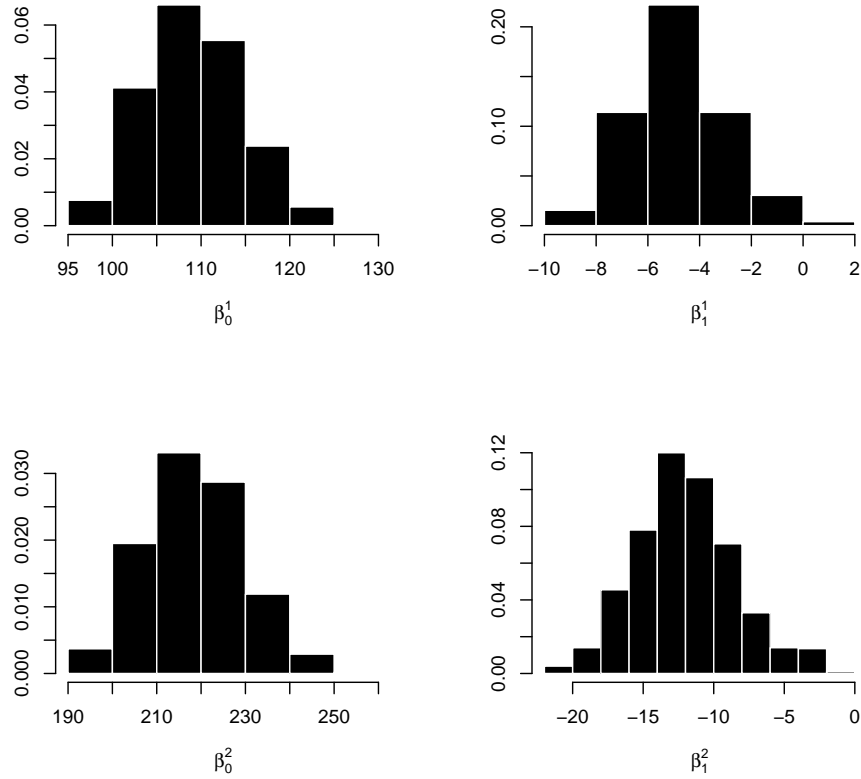


Figura 4.2: *Distribuciones finales de los vectores  $\beta^I$  y  $\beta^{II}$  para el Ejemplo 4.1.*

La Tabla 4.1 muestra los intervalos de credibilidad al 95 % para los parámetros

$\beta^I$	$\beta^{II}$	$(\Omega^{II})^{-1}$
(98.872, 119.515)	(198.826, 240.078)	( 0.000, 0.445)
( -8.404 , -0.920 )	( -18.399, -3.643 )	( 2.358, 6.318 )
		( -0.681, 1.519 )

Tabla 4.1: *Intervalos finales de credibilidad al 95% para cada uno de los componentes de los vectores  $\beta^I$  y  $\beta^{II}$ , así como para los elementos relevantes de la matriz  $(\Omega^{II})^{-1}$ .*



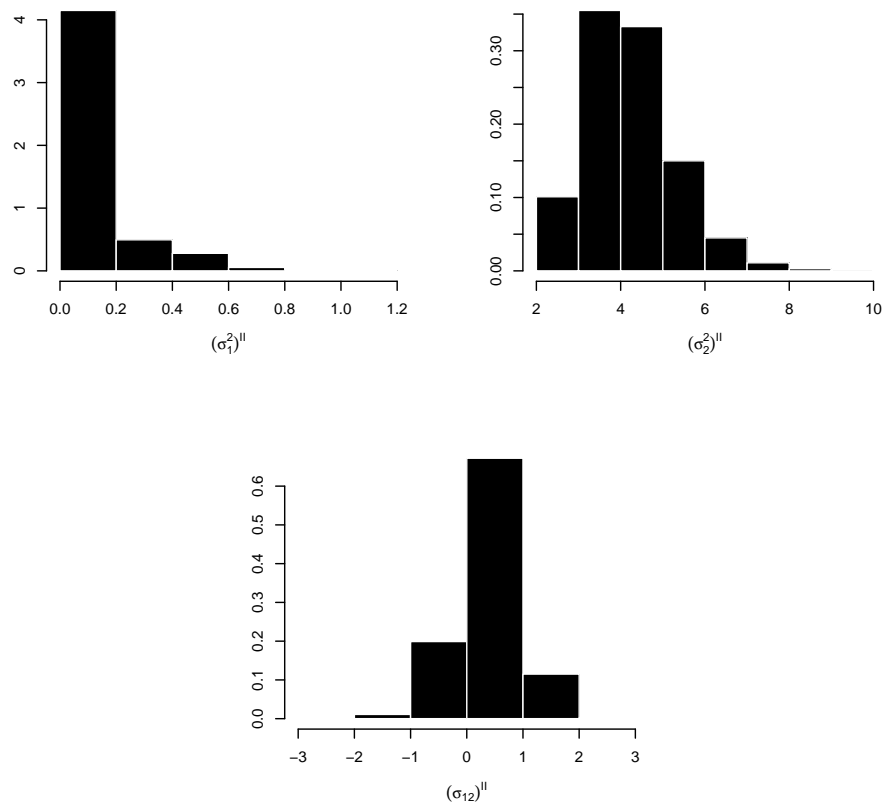


Figura 4.3: *Distribuciones finales de los parámetros de la matriz de covarianzas  $(\Omega^{II})^{-1}$  para el Ejemplo 4.1.*

del modelo LCP. Se puede notar que con la metodología empleada se pueden obtener inferencias razonables para el modelo longitudinal circular LCP. En particular, se puede observar que el verdadero valor de cada uno de los parámetros se encuentra contenido en el intervalo de máxima densidad de la correspondiente distribución final.

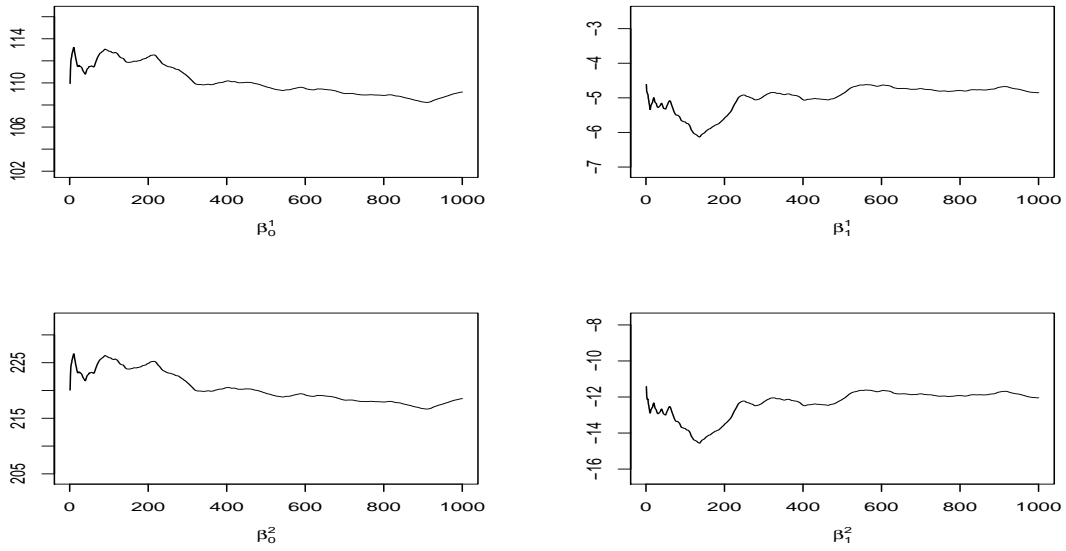


Figura 4.4: Promedio ergódicos de los vectores  $\beta^I$  y  $\beta^{II}$ .

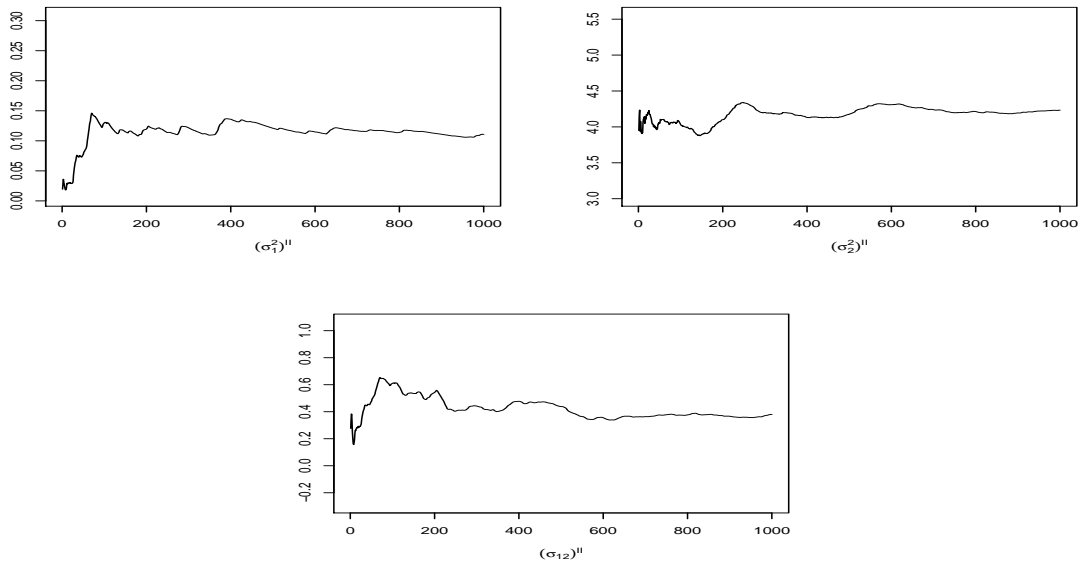


Figura 4.5: Promedios ergódicos para los parámetros de la matriz de covarianzas  $(\Omega^{II})^{-1}$ .

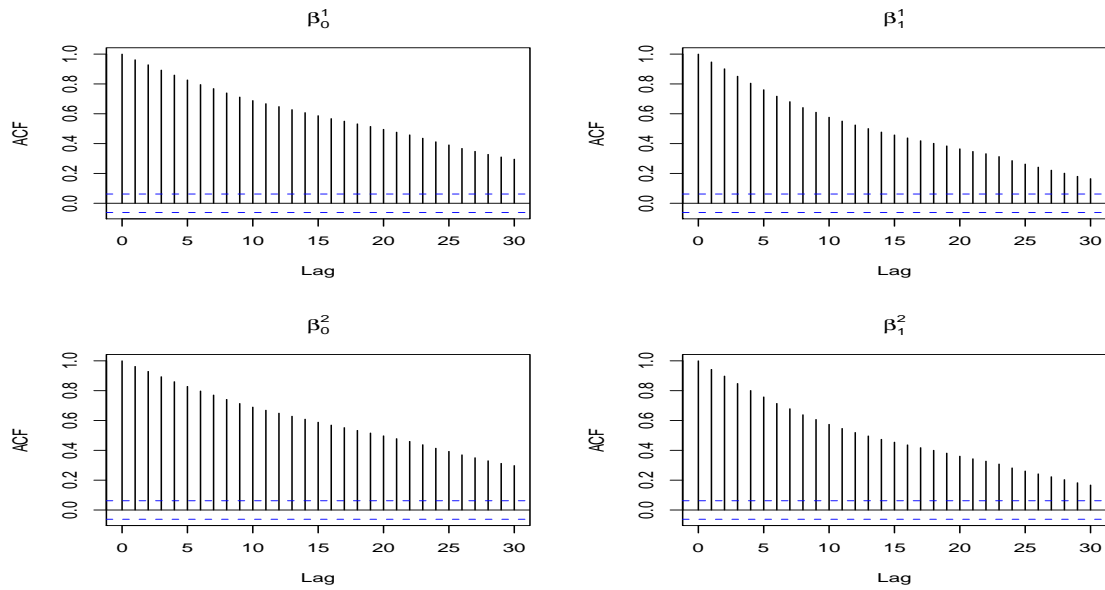


Figura 4.6: Autocorrelaciones para los elementos de los vectores  $\beta^I$  y  $\beta^{II}$ .

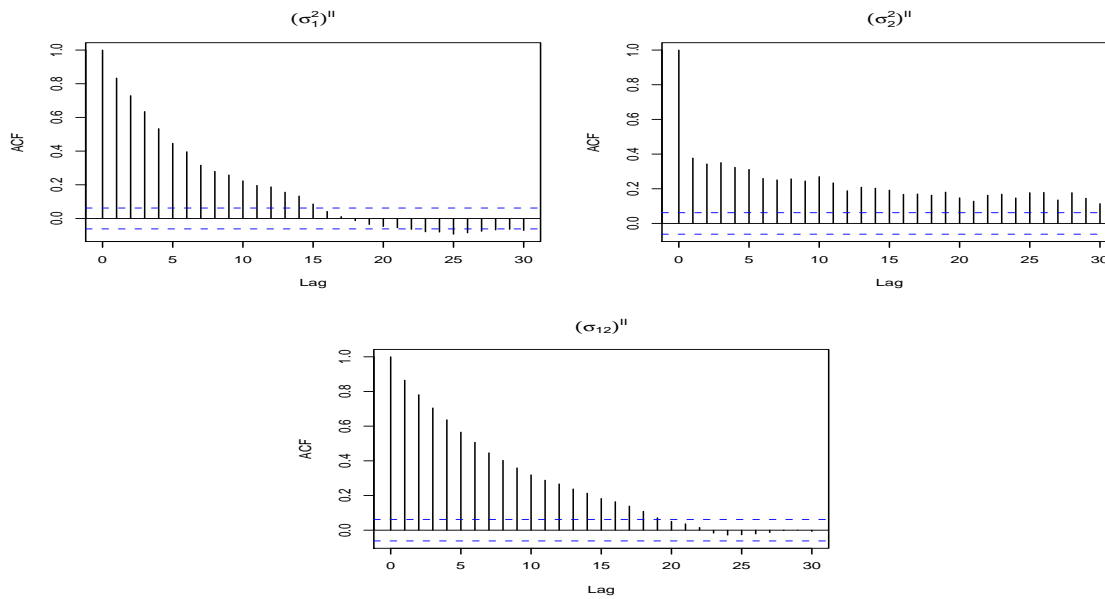


Figura 4.7: Autocorrelaciones para los parámetros de la matriz de varianzas y covarianzas  $(\Omega^{II})^{-1}$ .

En este ejemplo, un algoritmo que implementa el Gibbs sampler con bloque se corrió con un período de calentamiento de 100,000 iteraciones y posteriormente se obtuvo una muestra de tamaño 1000, donde cada unidad muestral se tomó cada 500 iteraciones de la cadena. Por su parte, el algoritmo de Metropolis-Hastings que simula cada una de las variables latentes  $R_{ij}$  ( $i = 1, \dots, 60$ ,  $j = 1, \dots, 5$ ) emplea 10 iteraciones en cada paso. Las Figuras 4.4 y 4.5 muestran los promedios ergódicos de las cadenas (promedios muestrales obtenidos con los valores acumulados en cada paso de la cadena) asociadas a los parámetros del modelo LCP ajustado. Por su parte las Figuras 4.6 y 4.7 exhiben las autocorrelaciones asociadas a cada una de las cadenas para los parámetros involucrados en el modelo. Hay que señalar que para este ejemplo fue necesario un número considerable de iteraciones del algoritmo para obtener inferencias razonables para los parámetros del modelo.

### **Estructuras longitudinales bajo el modelo LCP**

El modelo LCP resulta bastante flexible para modelar diversos comportamientos de dependencia longitudinal. Las Figuras 4.8, 4.9, 4.10 y 4.11 presentan algunas gráficas que muestran los diferentes comportamientos que se pueden tener bajo el modelo LCP.

Se puede observar que bajo este modelo se pueden tener los clásicos comportamientos de intercepto aleatorio, pendiente aleatoria e intercepto y pendiente aleatoria. Adicionalmente, bajo el modelo LCP se pueden tener dependencias más generales como la que se muestra en la Figura 4.11

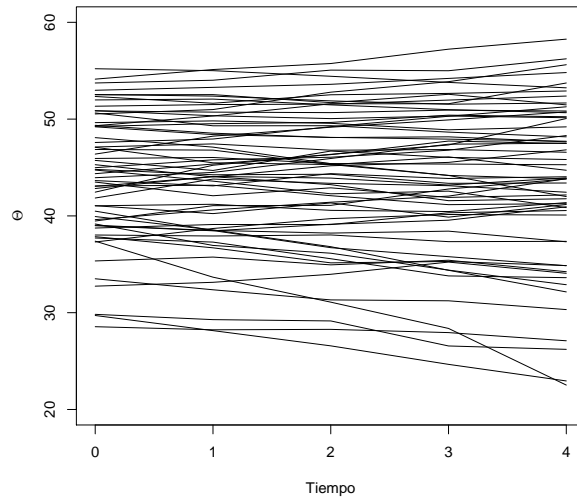


Figura 4.8: *Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.*

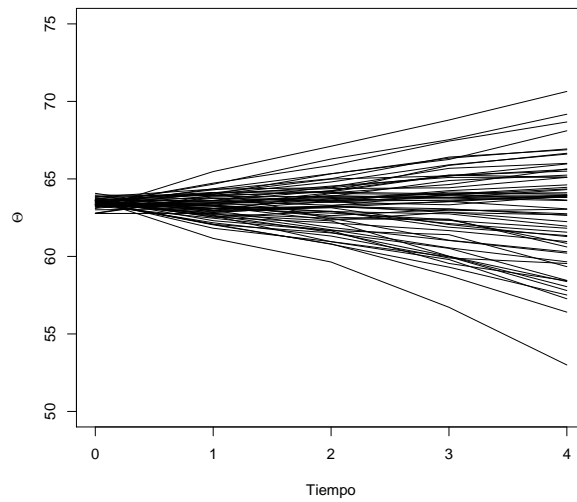


Figura 4.9: *Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.*

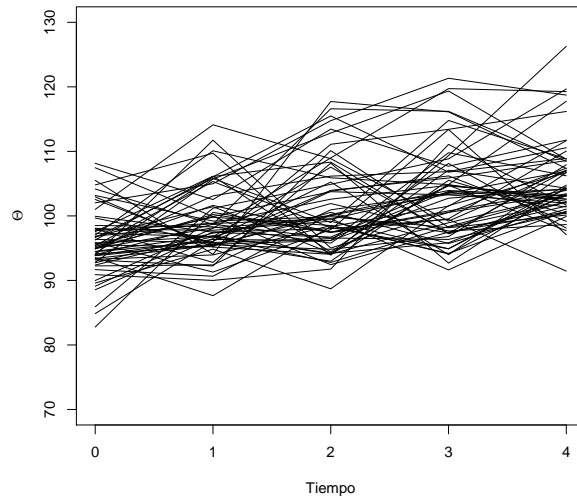


Figura 4.10: *Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.*

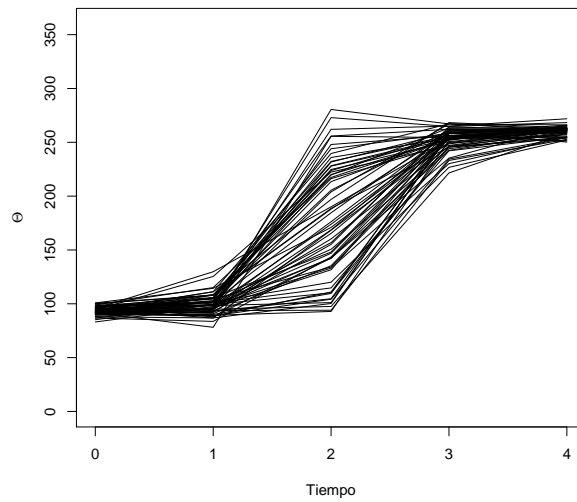


Figura 4.11: *Gráfica longitudinal para datos circulares que se puede obtener bajo el modelo LCP.*

**Ejemplo 4.2.** Para este ejemplo se consideró un conjunto de datos constituido por las direcciones de 65 *talitrus saltator* comúnmente llamadas pulgas de mar. Estos animales son pequeños y se relacionan con los cangrejos y langostas. Los *talitrus saltator* que viven en las costas del Mediterráneo están expuestos a dos principales fuerzas. Por un lado, deben alejarse de la orilla para evitar ser arrastrados por las olas, por el otro, deben mantenerse cerca del agua para evitar el riesgo de deshidratación bajo el sol. A lo largo del día, estos animalitos tienden a estar enterrados en la arena. Si son desplazados corren un riesgo mayor de deshidratación. Se cree que si los *talitrus saltator* son desplazados, se dirigirán hacia el mar tomando un curso perpendicular a la orilla de la playa, lo que se conoce como *dirección de escape teórica*.

Borgioli *et al.* (1999) y D'Elia *et al.* (2001) reportan un estudio longitudinal en cierta playa de Italia para entender el mecanismo asociado al escape de los *talitrus saltator* para evitar el riesgo de deshidratación. Los datos fueron tomados de Song (2007) y consisten en las direcciones de 65 *talitrus saltator* después de ser liberados secuencialmente en 5 ocasiones. Algunas covariables registradas en el estudio incluyen la velocidad del viento, la dirección azimuth del sol (Sun) y medidas oculares con las cuales se elabora un índice de simetría ocular (Eye). La dirección del viento fue transformada en cuatro categorías dependiendo del lugar donde provenía el viento (OS para offshore, LSE para longshore-east, LSW para lonshore-west y onshore). La Figura 4.12 muestra las 65 series con la correspondiente respuesta circular, la dirección de escape.

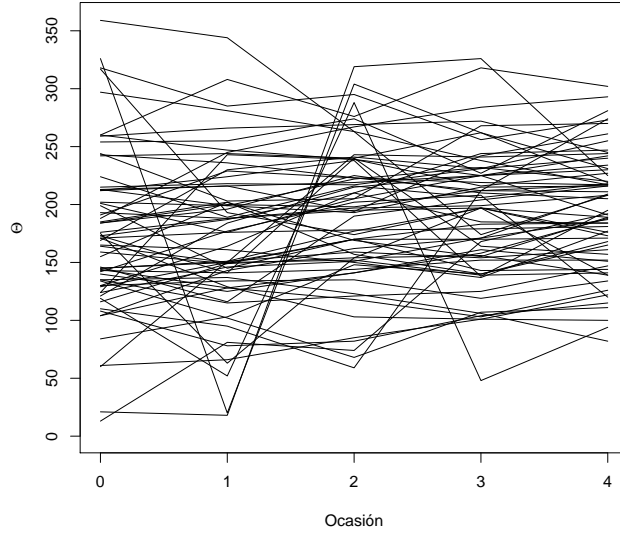


Figura 4.12: Gráfica longitudinal de las direcciones de escape de 65 *talitrus saltator* en 5 ocasiones consecutivas.

En D’Elia *et al.* (2001) y Song (2007) se analizan estos datos utilizando un enfoque de modelos lineales generalizados, donde la componente sistemática se define a través de un modelo lineal de efectos mixtos con intercepto aleatorio. En ambos casos, los autores consideran una distribución von Mises para el componente aleatorio. Sin embargo, de manera particular, Song (2007) no ofrece inferencias para todos los parámetros involucrados en el modelo.

Para analizar estos datos, en este trabajo se considera un modelo LCP con

$$\boldsymbol{\mu}_{ij}^I = \beta_0^I + \beta_1^I Sun + \beta_2^I Eye + \beta_3^I OS + \beta_4^I LSW + \beta_5^I LSE + \beta_6^I Tiempo \quad (4.6)$$

$$\boldsymbol{\mu}_{ij}^{II} = \beta_0^{II} + \beta_1^{II} Sun + \beta_2^{II} Eye + \beta_3^{II} OS + \beta_4^{II} LSW + \beta_5^{II} LSE + \beta_6^{II} Tiempo + b_{0i}$$

$$i = 1, \dots, 65.$$

En este ejemplo, para la especificación de la distribución inicial se tomó  $A^I = \mathbf{0}$ ,



$A^{II} = \mathbf{0}$ ,  $v^{II} = 2$  y  $B^{II} = 0.001$ . Los valores de  $\mathbf{b} = \mathbf{0}$ ,  $\mathbf{\Omega} = 1$  y  $\mathbf{R} = \mathbf{1}$  se usaron como valores iniciales para el Gibbs sampler. El Algoritmo 2 (Gibbs sampler con bloqueo) se corrió por 5000 iteraciones con un periodo de calentamiento de 1000 y en cada paso se consideraron sólo 2 iteraciones del algoritmo de Metropolis-Hastings.

Las Figuras 4.13 y 4.14 muestran las respectivas distribuciones finales para todos los parámetros de las componentes I y II del modelo LCP definido en (4.6). Por su parte, las Figuras 4.15, 4.16, 4.17 y 4.18 exhiben los promedios ergódicos y las autocorrelaciones asociados a cada uno de los parámetros de las componentes I y II del modelo LCP, respectivamente. Con base en el análisis correspondiente de las figuras 4.15 a 4.18, se puede concluir que con la especificación propuesta anteriormente para el Algoritmo 2 se está en condiciones de realizar inferencias para el modelo LCP (4.6).

En la Tabla 4.2 se muestran los intervalos de credibilidad al 95% para los parámetros de las dos componentes del modelo LCP, propuesto para los datos de *talitrus saltators*. Los resultados sugieren que el efecto de *Sun* y los efectos de *Eye*, *OS* y *LSW* no son relevantes para  $\boldsymbol{\mu}^I$  y  $\boldsymbol{\mu}^{II}$ , respectivamente. Como el parámetro  $\sigma^2$  resulta diferente de cero, la inclusión del efecto aleatorio asociado resulta necesario. Lo anterior indica una presencia de heterogeneidad entre los escapes de los *talitrus saltators*.

Se debe señalar que las inferencias obtenidas con este modelo LCP están en concordancia con análisis previos existentes en la literatura para estos datos, ver Song (2007) y D'Elia *et al.* (2001).

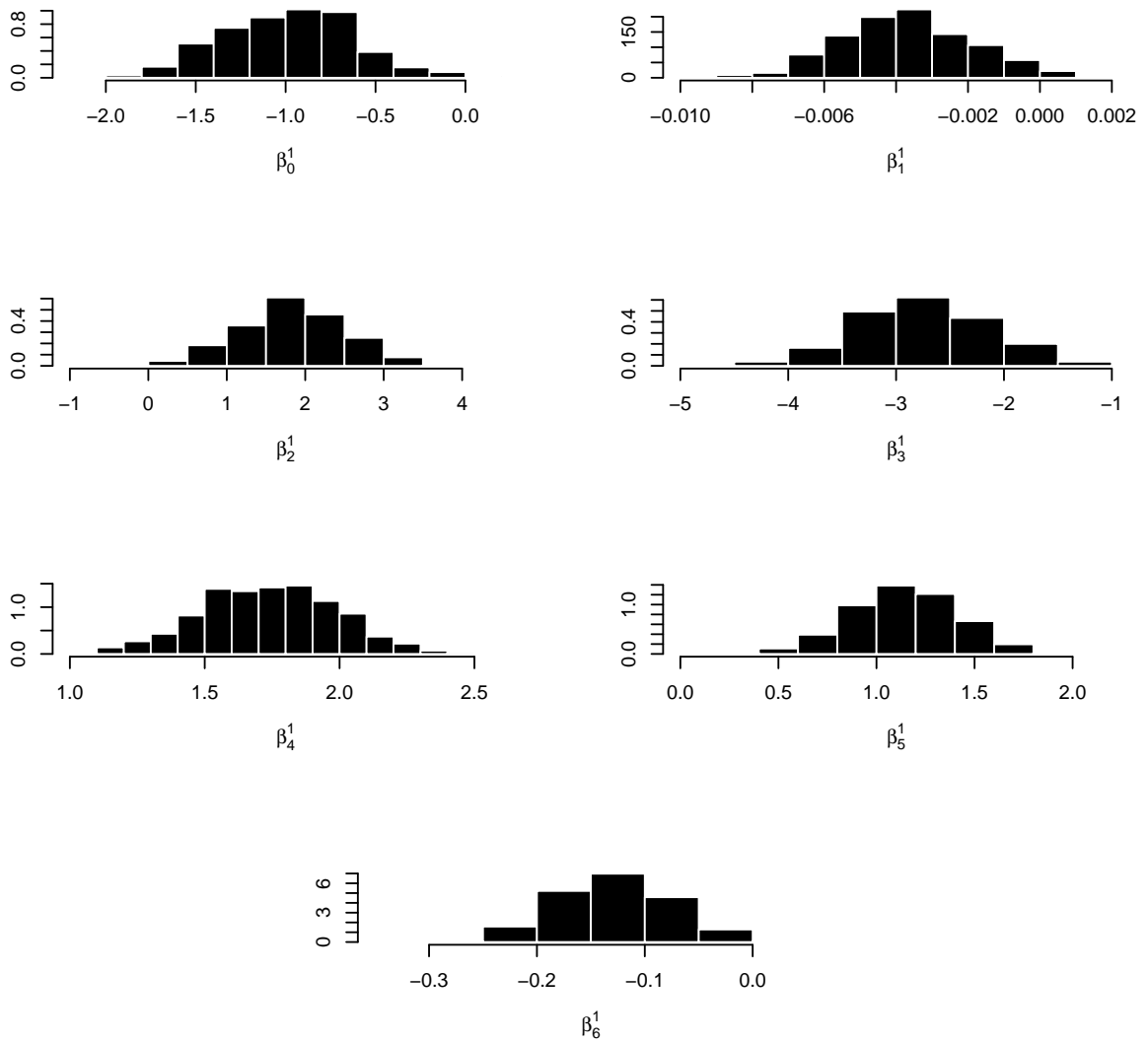


Figura 4.13: *Distribuciones finales de los parámetros en la componente I del modelo LCP para los datos de talitrus saltators.*

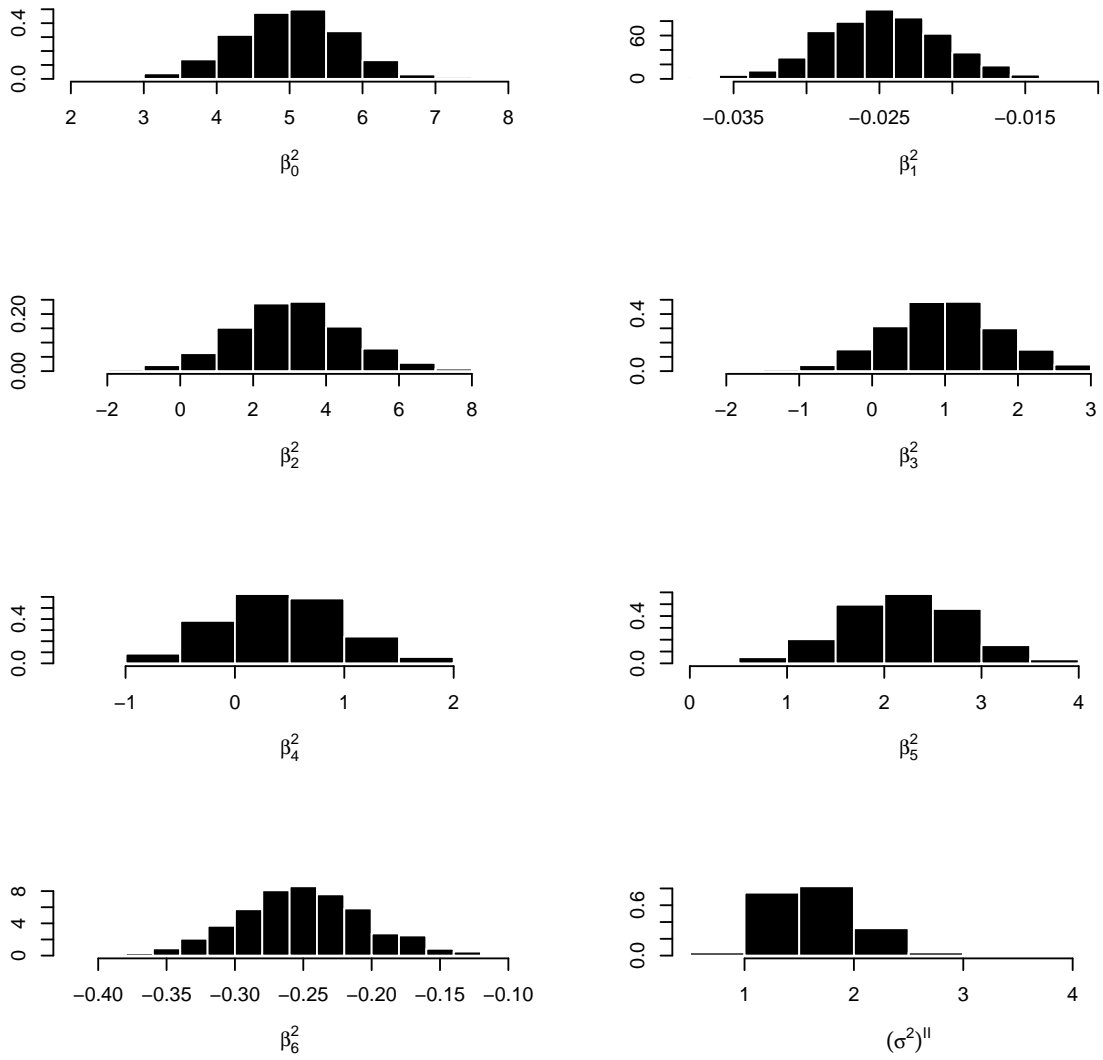


Figura 4.14: *Distribuciones finales de los parámetros en la componente II del modelo LCP para los datos de talitrus saltators.*

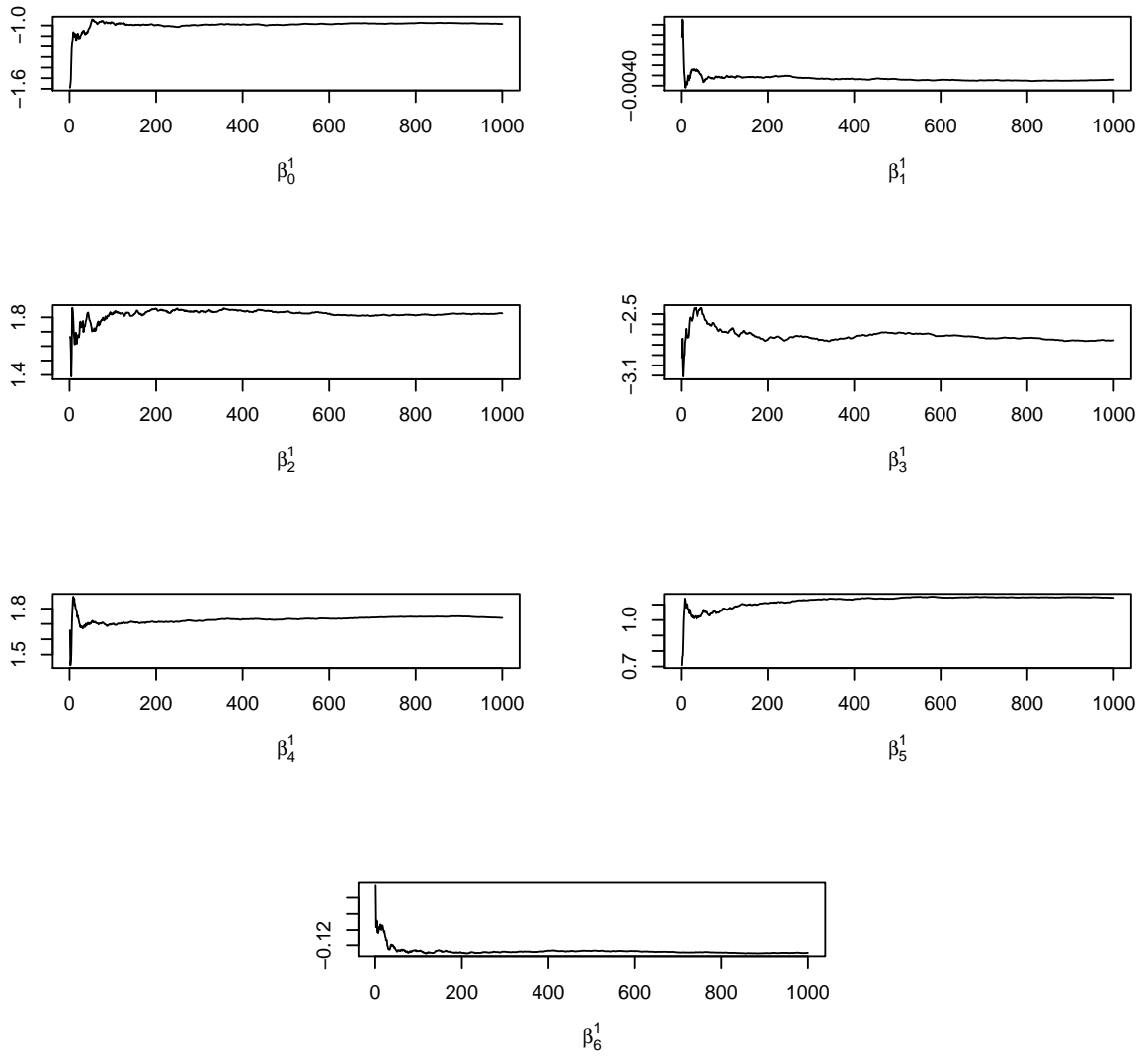


Figura 4.15: *Promedios ergódicos para los parámetros en la componente I del modelo LCP para los datos de talitrus saltators.*

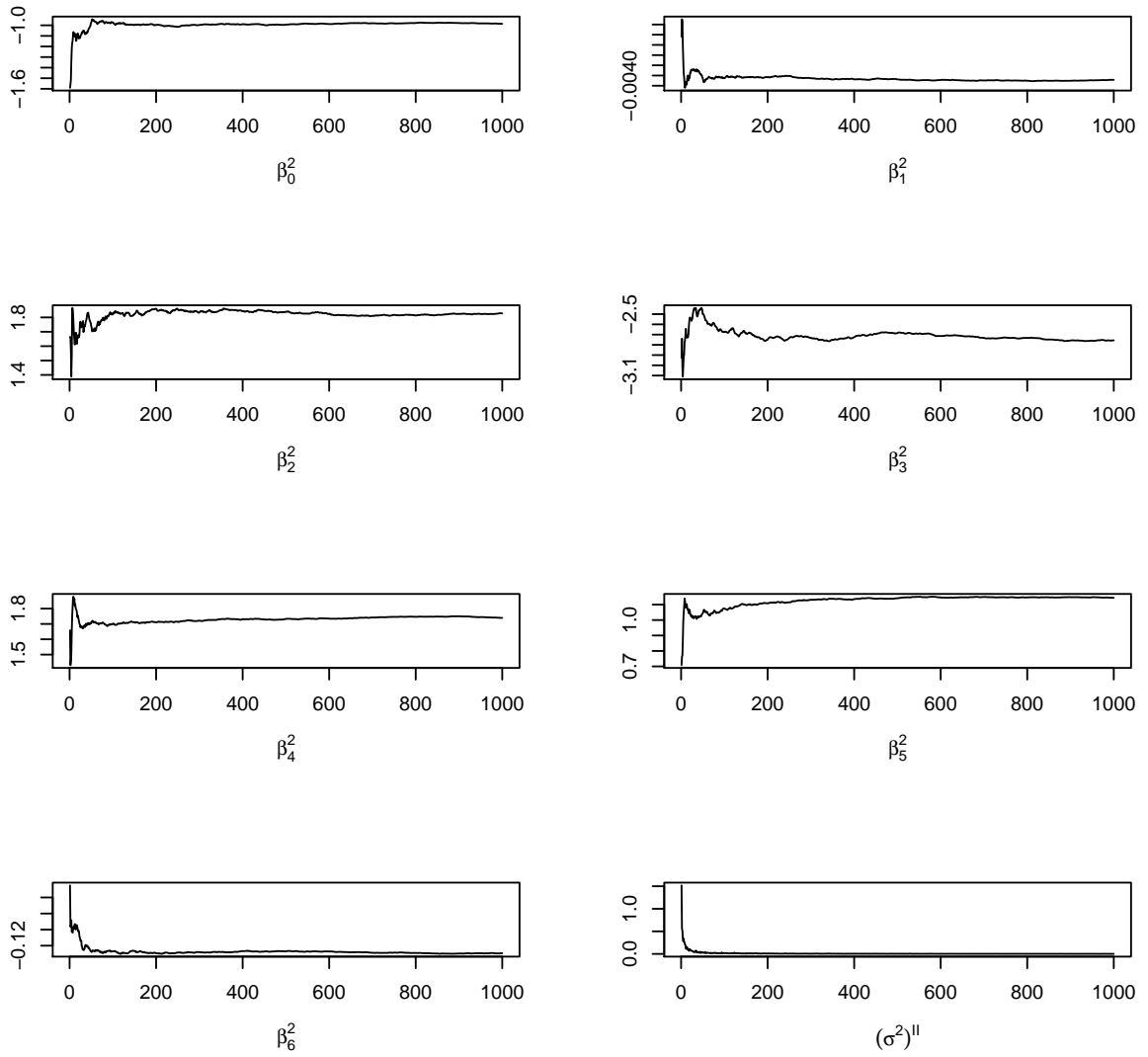


Figura 4.16: *Promedios ergódicos para los parámetros en la componente II del modelo LCP para los datos de talitrus saltators.*

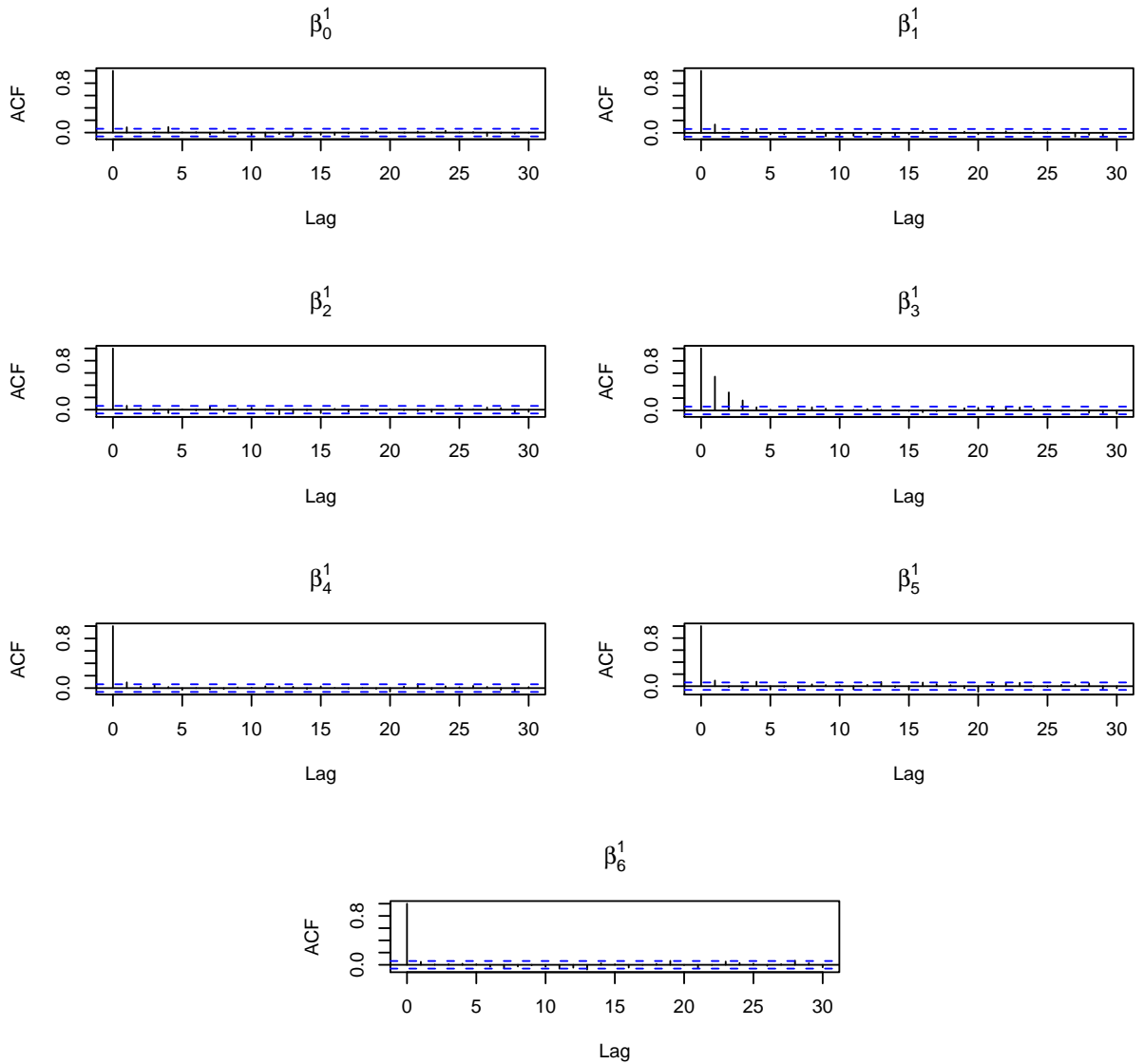


Figura 4.17: Autocorrelaciones de los parámetros en la componente I del modelo LCP para los datos de talitrus saltators.

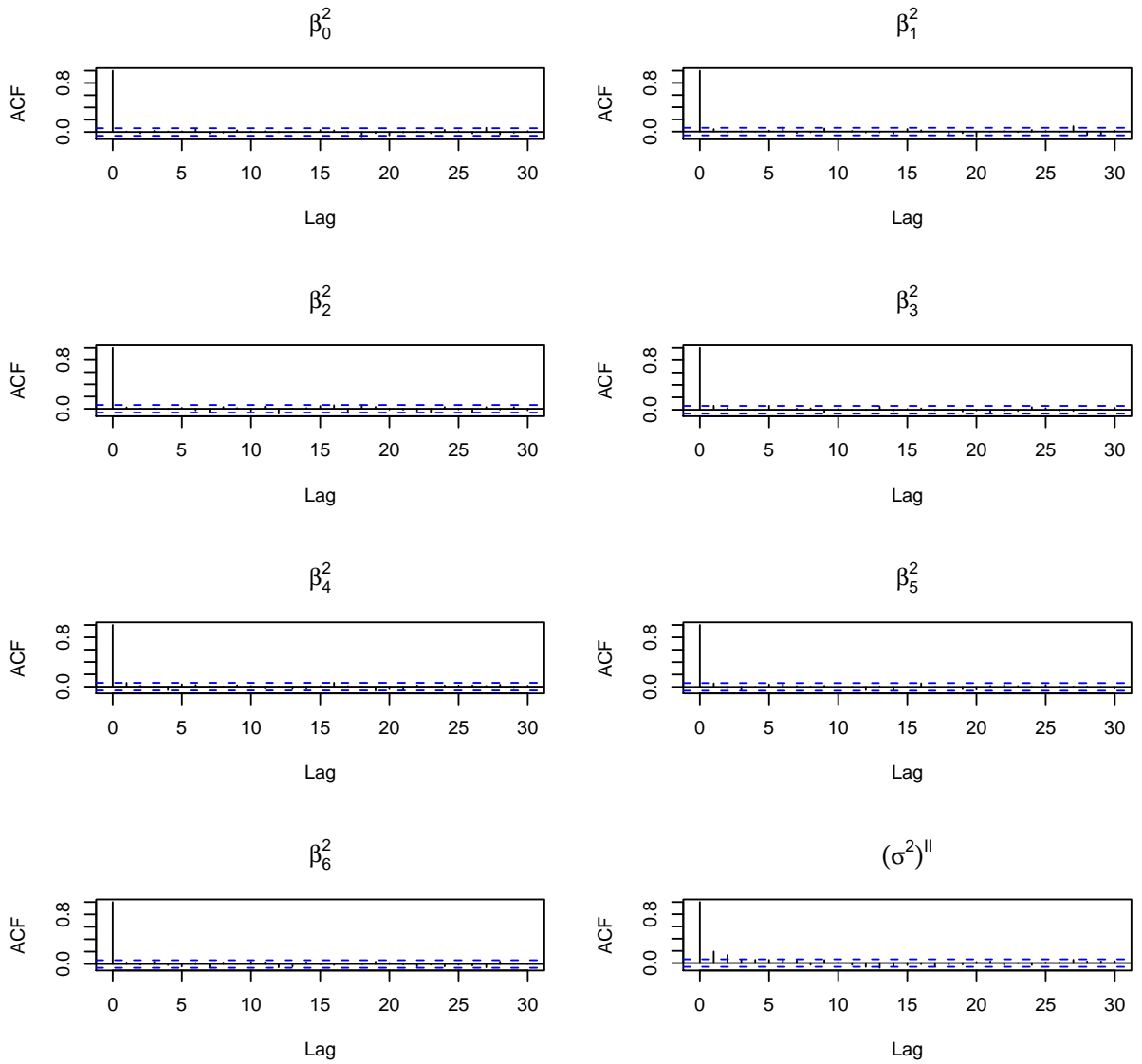


Figura 4.18: Autocorrelaciones de los parámetros en la componente II del modelo LCP para los datos de talitrus saltators.

	Componente I	Componente II
$\beta_0$	(-1.7041 , -0.2797)	(3.5300 , 6.4445)
$\beta_1(Sun)$	(-0.0069 , 0.0002)	(-0.0326 , -0.0167)
$\beta_2(Eye)$	(0.5108 , 3.1228 )	(-0.2894 , 6.1439)
$\beta_3(OS)$	(-4.0097 , -1.6492)	(-0.5842 , 2.4310)
$\beta_4(LSW)$	(1.2645 , 2.2534)	(-0.6257 , 1.5565)
$\beta_5(LSE)$	(0.6042 , 1.6781)	(0.9985 , 3.4888)
$\beta_6(Tiempo)$	(-0.2260 , -0.0277)	(-0.3420 , -0.1565)
$\sigma^2$		( 0.9825 , 2.4505 )

Tabla 4.2: *Intervalos finales de credibilidad al 95% para cada uno de los componentes del modelo LCP para los datos de talitrus saltators.*

## 4.7 Discusión

Los datos longitudinales de tipo circular han sido poco analizados en la literatura. Lo anterior se debe, por un lado, al reto de modelar la estructura de dependencia que se puede presentar al observar a un mismo conjunto de individuos en diferentes ocasiones. Por otro lado, al trabajar con datos de tipo direccional se deben emplear procedimientos de análisis que consideren la naturaleza periódica de este tipo de observaciones. En este trabajo se ha introducido un modelo, denominado LCP, para analizar datos longitudinales donde la variable de respuesta es de tipo circular. Dicha propuesta se basa en la distribución Normal proyectada, donde cada componente de la distribución normal bivariada que la genera se especifica a través de un modelo lineal mixto.



Aunque el modelo LCP asume una estructura de independencia condicional en cada uno de sus componentes, por lo discutido en este capítulo, resulta un buen candidato para modelar una amplia variedad de comportamientos longitudinales. El modelo LCP no considera explícitamente algún componente serial, por lo que inicialmente podría resultar poco atractivo para modelar series largas. Sin embargo, el modelo resulta útil para analizar series longitudinales cortas, donde uno de los principales objetivos es entender la asociación entre las covariables y la variable de respuesta observada repetidamente. No obstante, el modelo LCP ofrece bases para el estudio de modelos longitudinales para series largas de datos direccionales.

Los modelos existentes en la literatura para el análisis de datos longitudinales de tipo direccional asumen una distribución de probabilidad von Mises para la variable respuesta y atacan el problema desde un enfoque semiparamétrico como los métodos GEE o de modelos lineales generalizados. Sin embargo, estos enfoques presentan diversos problemas de maximización numérica. Por su parte, el modelo LCP es relativamente fácil de analizar bajo la metodología propuesta y no pierde aplicabilidad práctica comparado con los modelos que asumen una distribución von Mises para la variable circular. Adicionalmente, esta metodología ofrece la flexibilidad de realizar inferencias de manera directa para otros parámetros asociadas al modelo, por ejemplo, aquellos obtenidos bajo alguna reparametrización.

Se debe señalar que la metodología MCCM analizada para realizar inferencias sobre el modelo LCP puede requerir, en algunos casos, de un número relevante de iteraciones para su adecuada convergencia. Sin embargo, a pesar de esta aparente dificultad, tanto el modelo longitudinal circular proyectado (LCP) como la metodología MCCM sugerida resultan en una buena opción a tomar en cuenta en la modelación de datos longitudinales donde la variable de respuesta es de tipo circular. Adicionalmente, con la metodología propuesta se está en condiciones de atacar problemas como el de datos faltantes, comparación de modelos y pronóstico, entre otros.

# Capítulo 5

## Conclusiones y Perspectivas

El estudio de datos con estructuras de dependencia representa un reto ya sea en términos técnicos o desde el punto de vista aplicado. Aún más, cuando se trabaja con datos donde la variable de respuesta es de tipo direccional, proponer modelos que permitan describir adecuadamente relaciones de dependencia, requiere de un esfuerzo adicional.

En el contexto de regresión, en la literatura se han propuesto varios modelos que consideran una respuesta de tipo direccional. Sin embargo, la mayoría de ellos presentan limitaciones para realizar inferencias adecuadas para todos los parámetros involucrados en el modelo. Lo anterior, aunado a los diversos problemas numéricos que surgen en su implementación práctica. Por su parte, cuando se trabaja con conjuntos de datos direccionales de tipo longitudinal el problema se complica. Por un lado, hay que considerar la dependencia natural que se tiene cuando se trabaja con medidas repetidas para un mismo conjunto de individuos. Por otro lado, el tratamiento de la naturaleza direccional de los datos implica trabajar con modelos probabilísticos especiales; para los cuales los métodos inferenciales asociados aún se encuentran actualmente en etapa de

continuo desarrollo. Lo anterior da como resultado un vacío tanto en el planteamiento de modelos, que describan adecuadamente datos longitudinales de tipo circular, como en propuestas para realizar inferencias en esta clase de modelos.

Este trabajo de investigación pretende contribuir al análisis de modelos lineales para datos direccionales. Particularmente, se propone un análisis Bayesiano de un modelo de regresión donde la variable respuesta es circular. Este modelo de regresión se basa en la distribución de probabilidad que se obtiene al proyectar radialmente una distribución normal bivariada específica. Los procedimientos empleados para hacer inferencias sobre los parámetros de este modelo de regresión fundamentalmente descansan en métodos de Monte Carlo vía Cadenas de Markov (MCCM). Con la metodología propuesta los problemas de datos faltantes y de selección de covariables se pueden atacar de manera directa. Adicionalmente, esta metodología permite realizar inferencias para todos los parámetros del modelo y no presenta los problemas numéricos de la mayoría de los modelos de regresión propuestos en la literatura para una respuesta circular.

Para el tratamiento de datos longitudinales donde la variable de respuesta es de tipo circular, en este trabajo se introduce un modelo que también se basa en una versión de la distribución Normal proyectada. A este modelo se le denominó modelo longitudinal circular proyectado (LCP). En la literatura existen diversas propuestas para trabajar con datos longitudinales estándar. Sin embargo, debido a los beneficios que ofrecen los modelos mixtos para describir este tipo de observaciones repetidas, en este proyecto se optó por especificar cada uno de los componentes del modelo LCP a través de un modelo lineal mixto.

Para estar en condiciones de llevar a cabo inferencias vía métodos MCCM inicialmente se tuvo que introducir conjuntos convenientes de variables latentes, tanto para el modelo de regresión como para el modelo LCP. Lo anterior, con el objetivo de construir un modelo aumentado con la característica de que las distribuciones conjuntas para los parámetros involucrados fueran relativamente fáciles de simular. Específicamente, se trabajaron esquemas como el muestreo de Gibbs y el algoritmo de Metropolis-Hastings.

En el caso del modelo de regresión los algoritmos planteados son fáciles de implementar y su desempeño numérico es adecuado, particularmente en términos de la convergencia de las cadenas simuladas. En el caso del modelo longitudinal para datos circulares, hay que proceder con cautela. Debido a que el modelo LCP está compuesto por modelos lineales mixtos, éste hereda las bondades y ventajas relacionadas a la flexibilidad para describir datos longitudinales. Sin embargo, también hereda en cierta manera los problemas relacionados al ajuste de los propios modelos mixtos. El enfoque Bayesiano analizado en este trabajo permite atacar los problemas inferenciales para los modelos mixtos y ofrece una manera de realizar inferencias sobre el modelo LCP. Adicionalmente, dado que el enfoque propuesto permite obtener muestras de la distribución final para todos los parámetros del modelo LCP, ofrece condiciones para tratar con problemas como el de datos faltantes, selección de covariables y predicción, entre otros.

Los modelos lineales propuestos en este trabajo se fundamentan en la distribución Normal proyectada, la cual se obtiene al proyectar radialmente una distribución normal bivariada con matriz de covarianzas la matriz identidad. Inicialmente, podría parecer que al tener esta matriz particular de covarianzas el modelo resultaría muy restringido. Sin embargo, como se ha discutido en este trabajo, en el contexto de regresión el

modelo Normal proyectado no pierde aplicabilidad práctica, comparado con los modelos propuestos en la literatura que asumen una distribución von Mises o una distribución normal *wrapped* para la variable de respuesta.

En el caso del modelo longitudinal circular, la distribución Normal proyectada resulta atractiva en la construcción de modelos para describir series longitudinales cortas. Además, esta distribución concuerda con los modelos denominados de independencia condicional cuando se trabaja con datos longitudinales. Se debe reconocer que otras estructuras de la matriz de covarianzas resultarían en una familia más rica de distribuciones, la cual permitiría modelar en principio comportamientos de dependencia más complejos como los de las series longitudinales largas. Sin embargo, hay que señalar que las distribuciones Normales proyectadas generadas a partir de normales bivariadas con matrices de covarianzas distintas a la matrix identidad, pueden presentar problemas de falta de identificabilidad de sus parámetros.

A pesar de las limitaciones y/o desventajas que se puedan tener, los modelos y las metodologías desarrolladas en este trabajo de investigación representan una opción a considerar para la descripción de relaciones de dependencia cuando la variable de respuesta es circular. Finalmente, dado que los modelos discutidos aquí se basan en una distribución normal bivariada bajo proyección, éstos se pueden extender de manera más natural al estudio de datos direccionales en dimensiones mayores.

# Referencias Bibliográficas

1. Arnold, B.C. y SenGupta, A. (2006). Recent advances in the analyses of directional data in ecological and environmental sciences. *Environmental and Ecological Statistics*, **13**, 253-256.
2. Artes, R. y Jørgensen, B. (2000). Longitudinal data estimating equations for dispersion models. *Scand. J. Stat.*, **27**, 321-334.
3. Artes, R., Paula, G. A. y Ranvaud, R. (2000). Analysis of circular longitudinal data based on generalized estimating equations. *Aust. N. Z. J. Stat.*, **42**, 347-358.
4. Bagchi, P. (1987). Bayesian analysis of directional data. Ph.D Thesis, Universidad de Toronto.
5. Bagchi, P. y Guttman, I. (1988). Theoretical considerations of the von Mises-Fisher distribution. *J. Appl. Statist.*, **15**, 149-169.



6. Bagchi, P. y Kadane, J.B. (1991). Laplace approximations to posterior moments and marginal distributions on circle, spheres, and cylinders. *Canad. J. of Statist.*, **19**, 67-77.
7. Barbieri, M.M. y Berger, J.O. (2004). Optimal predictive model selection. *The Annals of Statistics*, **32**, 3, 870-897.
8. Bernardo, J.M. y Smith, A.F.M. (1994). *Bayesian Theory*. Chichester: Wiley.
9. Bock R.D. (1975). *Multivariate Statistical Methods in Behavioral Research*. New York: McGraw-Hill.
10. Borgioli, C., Martelli, M., Porri, F., D'Elia, A., Marcheti, G.M. y Scapini, F. (1999). Orientation in *talitrus saltator* (montagu): trends in intrapopulations variability related to enviromental and intrinsic factors. *Journal of Experimental Marine Biology and Ecology*, **238**, 29-47.
11. Chib, S. y Carlin, B.P. (1999). On MCMC sampling in hierarchical longitudinal models. *Statistics and Computing*, **9**, 17-26.
12. Damien, P. y Walker, S.G. (1999). A full Bayesian analysis of circular data using the von Mises distribution. *Canad. J. Statist.*, **2**, 291-298.
13. D'Elia, A., Borgioli, C. y Scapini, F. (2001). Orientation of sandhoppers under conditions in repeated trials: An analysis using longitudinal directional data. *Estuarine, Coastal and Shelf Science*, **56**, 839-847.
14. de Leeuw, J. y Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, **11**, 57-85.

15. Dempster A.P., Rubin, D.B. y Tsutakawa, R.K. (1981). Estimation in covariance component models. *Journal of the American Statistical Society*, **76**, 341-353.
16. Diggle P. J., Liang K. y Zeger S.L. (1994). *Analysis of Longitudinal Data*. New York: Oxford University Press, Inc.
17. Downs, T.D. y Mardia, K.V. (2002). Circular regression. *Biometrika*, **89**, 3, 683-697.
18. Fisher, N.I. (1993). *Statistical Analysis of Circular Data*. Cambridge: University Press.
19. Fisher, N.I. y Lee, A.J. (1992). Regression models for an angular response. *Biometrics*, **48**, 665-677.
20. Fisher, N.I., Lewis, T. y Embleton, B.J.J. (1987). *Statistical Analysis of Spherical Data*. Cambridge: University Press.
21. Geisser, S. y Eddy, W.F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association*, **74**, 365, 153-160.
22. Gelfand, A.E., Sahu, S.K. (1999). Identifiability, improper priors and Gibbs sampling for generalized linear models. *J. Amer. Statist. Assoc.* **94**, 445, 247-253.
23. Gelfand, A.E. y Dey, D.K. (1994). Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society, Series B*, **56**, 3, 501-514.
24. Gelfand, A.E., Dey D.K. y Chang H. (1992). Model determination using predictive distributions with implementation via sampling-based methods (with discussion).

- In Bernardo J.M., Berger J.O., Dawid A.P. and Smith A.F.M. eds. *Bayesian Statistics 4*. Oxford: Oxford University Press. 147-167.
25. Gelfand, A.E., Sahu, S.K. y Carlin, B.P. (1995). Efficient parametrizations for normal linear mixed models. *Biometrika*, **82**, 479-488.
  26. George, E.I. (2000) The variable selection problem. *Journal of the American Statistical Association*, **95**, 452, 1304-1308.
  27. George, B.J. y Ghosh, K. (2006). A semiparametric Bayesian model for circular-linear regression. *Communications in Statistics (Simulations and Computation)*, **35**, 4, 911-923.
  28. Gilks, W.R. y Roberts, G.O. (1996). Strategies for improving MCMC. In *Markov Chain Monte Carlo in Practice*, W.R. Gilks, S. Richardson y D.J. Spiegelhalter, D.J., eds, London: Chapman and Hall, 89-114.
  29. Goldstein, H. (1995). *Multilevel Statistical Models, 2nd edition*. New York: Halstead Press.
  30. Gould, A.L. (1969). A regression technique for angular variates. *Biometrics*, **25**, 683-700.
  31. Gutiérrez-Peña, E. (1997). A Bayesian predictive semiparametric approach to variable selection and model comparison in regression. *Bulletin of the International Statistical Institute*, Tome LVII. (Proceedings of the 51st Session of the ISI, Invited Papers, Book 1.) Istanbul, Turkey, 17-29.

32. Guttorp, P. y Lockhart, R.A. (1988). Finding de location of a signal: A Bayesian analysis. *J. Amer. Statist. Assoc.*, **83**, 322-329.
33. Hedeker D. y Gibbons R.D. (2006). *Longitudinal Data Analysis*. New Jersey: John Wiley & Sons, Inc.
34. Jammalamadaka, S.R. y Lund, U.J. (2006). The effect of wind direction on ozone levels: a case study. *Environmental and Ecological Statistics*, **13**, 3, 287-298.
35. Jammalamadaka, S.R. y SenGupta, A. (2001). *Topics in Circular Statistics*. Singapore: World Scientific.
36. Jennrich, R.I. y Schluchter, M.D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, **42**, 805-820.
37. Johnson, R.A. y Wehrly, T.A. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, **73**, 602-606.
38. Jørgensen, B. (1997a). Proper dispersion models (with discussion). *Braz. J. Probab. Statist.*, **11**, 89-140.
39. Jørgensen, B. (1997b). *The theory of dispersion models*. Chapman and Hall: London.
40. Jupp, P.E. y Mardia, K.V. (1989). A unified view of the theory of directional statistics, 1975 - 1989. *Int Stat Rev.*, **57**, 261-294.
41. Kent, J.T. (1978). Limiting behaviour of the von Mises-Fisher distribution. *Math. Proc. Cambridge Phil. Soc*, **84**, 531-536.

42. Kato, S., Kunio, S. y Grace, S.S. (2008). A circular-circular regression model. *Statistica Sinica*, **18**, 633-645.
43. Klotz, J. (1964). Small sample power of the bivariate sign test of Blumen and Hodges. *Ann. Math. Statist*, **35**, 1576-1582.
44. Kume, A. y Walker, S.G. (2006). Sampling from compositional and directional distributions. *Statistics and Computing*, **16**, 3, 261-265.
45. Kume, A. y Walker, S.G. (2009). On the Fisher-Bingham distribution. *Statistics and Computing*, **9**, 2, 167-172.
46. Laird, N.M. y Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 963-974.
47. Liang, K-Y. y Zeger, S.L. (1986). Longitudinal analysis using generalized linear models. *Biometrika*, **73**, 13-22.
48. Mardia, K.V. (1972). *Statistics of Directional Data*. Academic Press, London.
49. Mardia, K.V. (1975). Statistics of directional data. *J. Roy. Statist. Soc. Ser B*, **37**, 97-133.
50. Mardia, K.V. y El-Atoum, S.A.M. (1976). *Bayesian Inference for the von Mises-Fisher distribution*. *Biometrika*, **63**, 203-205.
51. Mardia, K.V. y Jupp, P.E. (2000). *Directional Statistics*. Chichester: Wiley.
52. Nuñez-Antonio, G. y Gutiérrez-Peña, E. (2005a). A Bayesian analysis of directional data using the projected normal distribution. *Journal of Applied Statistics*,

**32**, 10, 995-1001.

53. Nuñez-Antonio, G. y Gutiérrez-Peña, E. (2005b). A Bayesian Analysis of Directional Data Using the von Mises-Fisher Distribution. *Communications in Statistics-Simulation and Computation*, **34**, 989-999.
54. Presnell, B., Morrisson S.P. and Littell, R.C. (1998). Projected multivariate linear model for directional data. *Journal of the American Statistical Association*, **93**, 443, 1068-1077.
55. R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
56. Raudenbush, W. y Bryk A.S. (2002). *Hierarchical Linear Models, 2nd edition*. Thousand Oaks: Sage Publications (CA).
57. REDMET (2009). Atmospheric Monitoring System of Mexico City, URL <http://www.sma.df.gob.mx/sma/>. Site accessed March, 2009.
58. Rivest, L.P. (1997). A decentred predictor for circular-circular regression. *Biometrika*, **84**, 717-726.
59. Song, X-K. P. (2007). *Correlated Data Analysis: Modeling Analytics, and Applications*. Springer: New York.
60. Stephens, M.A. (1963). Random walk on a circle. *Biometrika*, **50**, 385-390.
61. Student (1908). The probable error of the mean. *Biometrika*, **6**, 1-25.

62. Tanner, M.A. y Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, **82**, 398, 528-540.
63. Vines, S.K., Gilks, W.R. y Wild, P. (1996). Fitting Bayesian multiple random effects models. *Statistics and Computing*, **6**, 337-346.
64. Watson, G.S. (1983). *Statistics on Sphere*, New York: Wiley.
65. Winer, B.J. (1971). *Statistical Principles in Experimental Design*, 2nd. edition. New York: McGraw-Hill.
66. Wolfinger R.D. (1993). Covariance structure selection in general mixed models. *Communications in Statistics, Simulation and Computation*, **22**, 1079-1106.

# Apéndice A

En este Apéndice se presentan las demostraciones de algunas proposiciones que aparecen a lo largo de este trabajo.

**Proposición 4.2.2** *Sea  $\mathbf{D}_n = \{\mathbf{Y}_1, \dots, \mathbf{Y}_N\}$  una m.a. del modelo  $N_n(\cdot | \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i, \mathbf{I})$ . Si se considera la siguiente especificación inicial dada en la Proposición 4.2.1,*

$$\begin{aligned} f(\boldsymbol{\beta} | A) &= N_p(\boldsymbol{\beta} | \mathbf{0}, A), \\ f(\mathbf{b}_i | \boldsymbol{\Omega}) &= N_q(\mathbf{b}_i | \mathbf{0}, \boldsymbol{\Omega}), \quad \text{para cada } i = 1, \dots, N, \\ f(\boldsymbol{\Omega} | v, B) &= Wi(\boldsymbol{\Omega} | v, B), \quad \text{con } E(\boldsymbol{\Omega}) = vB^{-1}. \end{aligned}$$

*Entonces las correspondientes densidades condicionales finales completas están dadas por*

- i.  $f(\boldsymbol{\beta} | \{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta} | C^{-1} \sum_{i=1}^N \mathbf{X}'_i \mathbf{e}_i, C)$
- ii.  $f(\mathbf{b}_i | \{\boldsymbol{\beta}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_q(\mathbf{b}_i | D_i^{-1} \mathbf{Z}'_i \tilde{\mathbf{e}}_i, D_i) \quad \forall i = 1, \dots, N$
- iii.  $f(\boldsymbol{\Omega} | \{\mathbf{b}_i\}, \mathbf{D}_n) = Wi(v + N, B + \sum_{i=1}^N \mathbf{b}_i \mathbf{b}'_i)$ .



donde

$$\begin{aligned}
C &= \sum_{i=1}^N \mathbf{X}'_i \mathbf{X}_i + A, \\
\mathbf{e}_i &= \mathbf{Y}_i - \mathbf{Z}_i \mathbf{b}_i, \\
D_i &= \mathbf{Z}'_i \mathbf{Z}_i + \Omega, \\
\tilde{\mathbf{e}}_i &= \mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}_i
\end{aligned}$$

DEMOSTRACIÓN:

- *Demostración del inciso i: Densidad condicional final completa para  $\boldsymbol{\beta}$ .*

De la Proposición (4.2.1) se tiene que la densidad final conjunta, de los parámetros del modelo, está dada por

$$\begin{aligned}
f(\boldsymbol{\beta}, \{\mathbf{b}_i\}, \Omega | \mathbf{D}_n) &\propto |\Omega|^{(v-q-1+N)/2} \\
&\times \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\
&\left. - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} - \frac{1}{2} \sum_{i=1}^N \mathbf{b}'_i \Omega \mathbf{b}_i - \frac{1}{2} \text{tr}(B\Omega)\right\}.
\end{aligned}$$

De lo anterior se sigue que la densidad condicional completa de  $\boldsymbol{\beta}$ , salvo una constante de proporcionalidad, está dada por

$$\begin{aligned}
f(\boldsymbol{\beta} | \{\mathbf{b}_i\}, \Omega, \mathbf{D}_n) &\propto \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\
&\left. - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} \right\} \\
&= \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{e}_i - \mathbf{X}_i \boldsymbol{\beta})' (\mathbf{e}_i - \mathbf{X}_i \boldsymbol{\beta}) - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta}\right\} \quad (5.1)
\end{aligned}$$

Trabajando con el exponente se obtiene

$$\begin{aligned}
& -\frac{1}{2} \sum_{i=1}^N (\mathbf{e}_i - \mathbf{X}_i \boldsymbol{\beta})' (\mathbf{e}_i - \mathbf{X}_i \boldsymbol{\beta}) - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} \\
& = \sum_{i=1}^N [\mathbf{e}_i' \mathbf{e}_i - \mathbf{e}_i' \mathbf{X}_i \boldsymbol{\beta} - (\mathbf{X}_i \boldsymbol{\beta})' \mathbf{e}_i + (\mathbf{X}_i \boldsymbol{\beta})' \mathbf{X}_i \boldsymbol{\beta}] + \boldsymbol{\beta}' A \boldsymbol{\beta} \\
& = \sum_{i=1}^N \mathbf{e}_i' \mathbf{e}_i - 2 \sum_{i=1}^N \mathbf{e}_i' \mathbf{X}_i \boldsymbol{\beta} + \sum_{i=1}^N \boldsymbol{\beta}' \mathbf{X}_i' \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\beta}' A \boldsymbol{\beta} \\
& \propto -2 \left( \sum_{i=1}^N \mathbf{e}_i' \mathbf{X}_i \right) \boldsymbol{\beta} + \boldsymbol{\beta}' \left( \sum_{i=1}^N \mathbf{X}_i' \mathbf{X}_i + A \right) \boldsymbol{\beta}
\end{aligned}$$

Completando la forma cuadrática,

$$-2 \left( \sum_{i=1}^N \mathbf{e}_i' \mathbf{X}_i \right) \boldsymbol{\beta} + \boldsymbol{\beta}' \left( \sum_{i=1}^N \mathbf{X}_i' \mathbf{X}_i + A \right) \boldsymbol{\beta} \propto (\boldsymbol{\beta} - C^{-1} \boldsymbol{\mu}_\beta)' C (\boldsymbol{\beta} - C^{-1} \boldsymbol{\mu}_\beta)$$

donde

$$\begin{aligned}
C & = \sum_{i=1}^N \mathbf{X}_i' \mathbf{X}_i + A, \\
\boldsymbol{\mu}'_\beta & = \sum_{i=1}^N \mathbf{e}_i' \mathbf{X}_i.
\end{aligned}$$

Así, resulta

$$\begin{aligned}
f(\boldsymbol{\beta} | \{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) & \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\
& \left. \propto \exp \left\{ (\boldsymbol{\beta} - C^{-1} \boldsymbol{\mu}_\beta)' C (\boldsymbol{\beta} - C^{-1} \boldsymbol{\mu}_\beta) \right\} \right.
\end{aligned}$$

Es decir,  $f(\boldsymbol{\beta} | \{\mathbf{b}_i\}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta} | C^{-1} \sum_{i=1}^N \mathbf{X}_i' \mathbf{e}_i, C)$ .

- *Demostración del inciso ii: Densidades condicionales finales completas para cada  $\mathbf{b}_k$ , con  $k = 1, \dots, N$ .*

Como la densidad final conjunta, de los parámetros del modelo es

$$f(\boldsymbol{\beta}, \{\mathbf{b}_i\}, \boldsymbol{\Omega} | \mathbf{D}_n) \propto |\boldsymbol{\Omega}|^{(v-q-1+N)/2} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\ \left. - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i - \frac{1}{2} \text{tr}(B \boldsymbol{\Omega})\right\}.$$

Entonces, la densidad condicional completa de  $\mathbf{b}_k$ , resulta ser

$$f(\mathbf{b}_k | \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{D}_n) \propto \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)\right\} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^N \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i\right\} \\ \propto \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i)' (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i) - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i\right\} \\ \propto \exp\left\{-\frac{1}{2} (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i)' (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i) - \frac{1}{2} \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i\right\} \quad (5.2)$$

Se puede notar que esta última expresión (5.2) es similar a la ecuación (5.1). Por lo tanto, siguiendo los desarrollos del inciso anterior, se tiene

$$f(\mathbf{b}_k | \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{D}_n) \propto \exp\left\{-\frac{1}{2} (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i)' (\tilde{\mathbf{e}}_i - \mathbf{Z}_i \mathbf{b}_i) - \frac{1}{2} \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i\right\}$$

Es decir,  $f(\mathbf{b}_k | \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{D}_n) = N_q(\mathbf{b}_k | D^{-1} \mathbf{Z}_i' \tilde{\mathbf{e}}_i, D_i) \forall k = 1, \dots, N$ , con

$$D_i = \mathbf{Z}_i' \mathbf{Z}_i + \boldsymbol{\Omega}, \\ \tilde{\mathbf{e}}_i = \mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}_i.$$

- *Demostración del inciso iii: Densidad condicional final completa para  $\boldsymbol{\Omega}$ .*

Como la densidad final conjunta, de los parámetros del modelo, está dada por

$$f(\boldsymbol{\beta}, \{\mathbf{b}_i\}, \boldsymbol{\Omega} | \mathbf{D}_n) \propto |\boldsymbol{\Omega}|^{(v-q-1+N)/2} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i)' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{Z}_i \mathbf{b}_i) \right. \\ \left. - \frac{1}{2} \boldsymbol{\beta}' A \boldsymbol{\beta} - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i' \boldsymbol{\Omega} \mathbf{b}_i - \frac{1}{2} \text{tr}(B \boldsymbol{\Omega})\right\}$$

Entonces, la densidad condicional completa de  $\Omega$ , resulta ser

$$\begin{aligned}
f(\Omega|\beta, \{\mathbf{b}_i\}, \mathbf{D}_n) &\propto |\Omega|^{(v-q-1+N)/2} \\
&\times \exp\{-\frac{1}{2} \sum_{i=1}^N \mathbf{b}'_i \Omega \mathbf{b}_i - \frac{1}{2} \text{tr}(B\Omega)\} \\
&= |\Omega|^{(v-q-1+N)/2} \\
&\times \exp\{-\frac{1}{2} \sum_{i=1}^N \text{tr}(\Omega \mathbf{b}_i \mathbf{b}'_i) - \frac{1}{2} \text{tr}(B\Omega)\} \\
&= |\Omega|^{(v-q-1+N)/2} \\
&\times \exp\{-\frac{1}{2} \text{tr}[\sum_{i=1}^N \Omega \mathbf{b}_i \mathbf{b}'_i + B \Omega ]\} \\
&= |\Omega|^{(v-q-1+N)/2} \\
&\times \exp\{-\frac{1}{2} \text{tr}[(\sum_{i=1}^N \mathbf{b}_i \mathbf{b}'_i + B) \Omega ]\}
\end{aligned}$$

Por lo tanto,  $f(\Omega|\beta, \{\mathbf{b}_i\}, \mathbf{D}_n) = Wi(v + N, B + \sum_{i=1}^N \mathbf{b}_i \mathbf{b}'_i)$ .

□

**Proposición 4.6.1** Sea  $\mathbf{D}_n = \{(r_{11}, \theta_{11}), \dots, (r_{Nn}, \theta_{Nn})\}$  un conjunto de observaciones del modelo LCP

$$\mathbf{Y}_i^k | \beta^k, \{\mathbf{b}_i\}^k \sim N_n(\mathbf{X}_i^k \beta^k + \mathbf{Z}_i^k \mathbf{b}_i^k, \mathbf{I}), \quad \forall i = 1, \dots, N, \quad y \quad k \in \{I, II\}$$

Omitiendo el superíndice  $k$ , la densidad condicional completa para  $\beta^k$ ,  $k \in \{I, II\}$ , está dada por

$$f(\beta|\Omega, \mathbf{D}_n) = N_p(\beta|\mu_\beta, \Lambda_\beta),$$

donde

$$\begin{aligned}
\mu_\beta &= \Lambda_\beta^{-1} ( \sum_{i=1}^N \mathbf{X}'_i \mathbf{V}_i^{-1} \mathbf{Y}_i ), \\
\Lambda_\beta &= ( A + \sum_{i=1}^N \mathbf{X}'_i \mathbf{V}_i^{-1} \mathbf{X}_i ).
\end{aligned}$$

DEMOSTRACIÓN:

Debido a la estructura Gaussiana condicional del modelo LCP se tiene que

$$\mathbf{Y}_i | \beta, \Omega \sim N_n(\mathbf{X}_i \beta, \mathbf{V}_i^{-1}),$$

donde  $\mathbf{V}_i = \mathbf{Z}_i \Omega^{-1} \mathbf{Z}'_i + \mathbf{I}$ . Por otro lado, si se considera una distribución inicial  $N_p(\beta|\mathbf{0}, A)$  para el parámetro  $\beta$  de cada componente, entonces

$$\begin{aligned}
f(\boldsymbol{\beta}|\boldsymbol{\Omega}|\mathbf{D}_n) &\propto \exp\left\{-\frac{1}{2}\sum_{i=1}^N(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta}) - \frac{1}{2}\boldsymbol{\beta}'A\boldsymbol{\beta}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\sum_{i=1}^N\mathbf{Y}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\boldsymbol{\beta} + \sum_{i=1}^N\boldsymbol{\beta}'\left(A + \sum_{i=1}^N\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\right)\boldsymbol{\beta}\right)\right\} \\
&\propto \exp\left\{(\boldsymbol{\beta} - \boldsymbol{\Lambda}_\beta\boldsymbol{\mu}_\beta)'\mathbf{C}(\boldsymbol{\beta} - \boldsymbol{\Lambda}_\beta\boldsymbol{\mu}_\beta)\right\}
\end{aligned}$$

Así, resulta

$$f(\boldsymbol{\beta}|\boldsymbol{\Omega}, \mathbf{D}_n) = N_p(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}_\beta)$$

con

$$\boldsymbol{\mu}_\beta = \boldsymbol{\Lambda}_\beta^{-1}\left(\sum_{i=1}^N\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{Y}_i\right)$$

y

$$\boldsymbol{\Lambda}_\beta = \left(A + \sum_{i=1}^N\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\right).$$

□

**Proposición.** Sea  $\mathbf{X}$  un vector aleatorio con distribución normal bivariada con vector de medias  $\boldsymbol{\mu}$  y matriz de precisión  $\boldsymbol{\Lambda}$ . Si  $\mathbf{U} = \mathbf{X}/\|\mathbf{X}\|$  denota la Normal proyectada inducida, con dirección media  $\boldsymbol{\eta} = E(\mathbf{U})/\|E(\mathbf{U})\|$ , entonces  $\boldsymbol{\eta}$  también puede ser determinada como  $\boldsymbol{\eta} = \boldsymbol{\mu}/\|\boldsymbol{\mu}\|$ .

DEMOSTRACIÓN:

Primero, hay que notar que

$$\begin{aligned}
\boldsymbol{\eta} &= E[\mathbf{U}]/\|E[\mathbf{U}]\| = E[\mathbf{X}/\|\mathbf{X}\|]/\|E[\mathbf{X}/\|\mathbf{X}\|]\| \\
&= E[(\cos \Theta, \sin \Theta)]/\|E[(\cos \Theta, \sin \Theta)]\|
\end{aligned}$$

donde  $\Theta$  es el ángulo aleatorio correspondiente a la dirección aleatoria  $\mathbf{U}$ . Así,  $\boldsymbol{\eta} = (\cos w, \sin w)$  para algún  $w \in (0, 2\pi]$ , dado que  $E[(\cos \Theta, \sin \Theta)]/\|E[(\cos \Theta, \sin \Theta)]\|$  es un vector unitario.

Por otro lado,

$$\begin{aligned}
\boldsymbol{\mu}/\|\boldsymbol{\mu}\| &= E[\mathbf{X}]/\|E[\mathbf{X}]\| = E_{r,\Theta}[r(\cos \Theta, \text{sen}\Theta)]/\| E_{r,\Theta}[r(\cos \Theta, \text{sen}\Theta)] \| \\
&= E_r E_{\Theta|r}[r(\cos \Theta, \text{sen}\Theta)]/\| E_r E_{\Theta|r}[r(\cos \Theta, \text{sen}\Theta)] \| \\
&= E_r[r E_{\Theta|r}[(\cos \Theta, \text{sen}\Theta)]]/\| E_r[r E_{\Theta|r}[(\cos \Theta, \text{sen}\Theta)]] \| \\
&= E_r[r r^*(r)(\cos w, \text{sen}w)]/\| E_r[r r^*(r)(\cos w, \text{sen}w)] \|
\end{aligned}$$

para algún  $r^*(r) > 0$ . Así,

$$\begin{aligned}
\boldsymbol{\mu}/\|\boldsymbol{\mu}\| &= (\cos w, \text{sen}w) E_r[r r^*(r)] / \| (\cos w, \text{sen}w) E_r[r r^*(r)] \| \\
&= (\cos w, \text{sen}w) \\
&= \boldsymbol{\eta}.
\end{aligned}$$

□

# Apéndice B

Este Apéndice contiene los programas necesarios para implementar la metodología propuesta en este trabajo de investigación. Además, estos programas permiten llevar a cabo y reproducir todos los resultados y ejemplos obtenidos. Los programas `RegressionFunctions.R`, `MissingData.R` y `CircularRegression.R` corren en R en su versión 2.9.1. Los programas requieren de los R-paquetes `MASS`, `CircStat` y `TeachingDemos`.

Al final del este apéndice se incluye el programa `Normal-proyectada.nb` que incluye el código en `Mathematica 5.2` con el que se pueden reproducir las gráficas exploratorias (2.1), (2.2), (2.3) y (2.4) del Capítulo 2.

## Contenido del Apéndice B

Programa	Página
<code>CircularRegression.R</code> .....	132
<code>RegressionFunctions.R</code> .....	136
<code>MissingData.R</code> .....	144
<code>Normal-proyectada.nb</code> .....	151
<code>LCP.R</code> .....	155

## CircularRegression.R

```
# Este programa contiene las funciones correspondientes y el código
# para reproducir los análisis presentados en el Capítulo 3 y
# publicados en el artículo 'A Bayesian Regression Model for Circular
# Data based on the projected Normal Distribution' de Nuñez-Antonio G.,
# Gutiérrez-Peña E., y Escarela G.
#
# Autor(es): G. Nuñez-Antonio y
#           E. Gutiérrez-Peña < eduardo@sigma.iimas.unam.mx >
# Responsable: G. Nuñez-Antonio < gab.nuneza@gmail.com >
# Fecha: Julio, 2009.
#-----
source(file="RegressionFunctions.R")
#-----

Ejemplo 3.1: Datos: Muestra de tamaño 100 de una distribución
normal proyectada

Example1<-read.table(file="Example1.dat")
E1<-CircReg(theta~1+x,theta~1+x+x2,data=Example1,tm=1000,t.lag=20)
B1<-E1$BI
B2<-E1$BII
if(0){
  # Figura 3.1
  nf<-layout(rbind(c(1,1,2,2),c(3,3,4,4),c(0,5,5,0)))
  hist(B1[,1],freq=F,col="black", border="white", main="",
  xlab=expression(beta[0]^1:(-3.0)))
  hist(B1[,2],freq=F,col="black", border="white", main="",
  xlab=expression(beta[1]^1:(1.0)))
  hist(B2[,1],freq=F,col="black", border="white", main="",
  xlab=expression(beta[0]^2:(1.5)))
  hist(B2[,2],freq=F,col="black", border="white", main="",
  xlab=expression(beta[1]^2:(-2.0)))
}
```



```

hist(B2[,3],freq=F,col="black", border="white", main="",
xlab=expression(beta[2]^2:(3.0)
}
#-----
# Ejemplo 3.2: Datos: Muestra de tamaño 100 de una distribución
# von Mises.

Example2<-read.table(file="Example2.dat")
E2<-CircReg(theta~1,theta~1+x,data=Example2,tm=1000,t.lag=20)
B1<-E2$BI
B2<-E2$BII
if(0){
  # Figura 3.2
  X.1.aux<-model.matrix(theta~1,Example2)
  p.1<-length(X.1.aux[1,])
  X.1<-matrix(X.1.aux,ncol=p.1)
  #
  X.2.aux<-model.matrix(theta~1+x,Example2)
  p.2<-length(X.2.aux[1,])
  X.2<-matrix(X.2.aux,ncol=p.2)
  #
  nn<-10000
  mu.star<-matrix(0,tm,2)
  xy.star1<-rep(0,nn)
  xy.star2<-rep(0,nn)
  theta.star<-matrix(0,n,nn)
  theta.star.q<-matrix(0,n,2)
  medias.star<-rep(0,n)
  #
  n<-100    #sample size
  #
  for(j in 1:n)
    {

```

```

    mu.star[,1]<-c(B1**X.1[j,])
    mu.star[,2]<-c(B2**X.2[j,])
    xy.star1<-rnorm(nn,mean(mu.star[,1]),1)
    xy.star2<-rnorm(nn,mean(mu.star[,2]),1)
    theta.star[j,]<-atan2(xy.star2,xy.star1)**(2*pi)
    medias.star[j]<-circ.stats(theta.star[j,],1)
  }
Mdat<-matrix(c(Example2$x,medias.star),n,2)
o<-order(Mdat[,1])
Mdat.o<-cbind(Mdat[o,1],Mdat[o,2])
#
par(mfrow=c(1,1))
plot(Example2$x,(Example2$theta)*180/pi,ylim=c(0,760),pch=19,xlab="x",
ylab=expression(theta),main="",cex=1)
points(Example2$x,360+(Example2$theta)*180/pi,pch=19,cex=1)
points(Mdat.o[,1],Mdat.o[,2],pch="-",cex=1.1)
points(Mdat.o[,1],Mdat.o[,2]+360,pch="-",cex=
}
#-----
# Ejemplo 3.3: Datos de caracoles (Periwinkles)
periwinkles<-read.table(file="Periwinkles.dat")
per<-CircReg(theta~1+x,data=periwinkles,tm=1000,t.lag=20,
              burn=5000,flag.lk=TRUE)
B1<-per$BI
B2<-per$BII
if(0){
  # Figura 3.3
  par(mfrow=c(2,2))
  hist(B1[,1],freq=F,col="black",border="white",ylab="",main="",
xlab=expression(beta[0]^1))
  hist(B1[,2],freq=F,col="black",border="white",ylab="",main="",
xlab=expression(beta[1]^1))
  hist(B2[,1],freq=F,col="black",border="white",ylab="",main="",

```

```

    xlab=expression(beta[0]^2))
    hist(B2[,2],freq=F,col="black",border="white",ylab="",main="",
    xlab=expression(beta[1]^2))
}
# Nota: Las Figuras 3.4 y 3.5 se obtienen con el programa MissingData.R
#-----
# Ejemplo 3.4: Datos REDMET.
red2m.aux<-read.csv(file="red2m.csv",header=T)
red2m.dat<-red2m.aux[,-1]
red2<-CircReg(theta~1+rh+tmp+wsp+hora.s+hora.c,data=red2m.dat,
              tm=1000,t.lag=20)
B1<-red2$BI
B2<-red2$BII
if(0){
  # Tabla 3.1
  library(TeachingDemos)
  emp.hpd(B1[,1],conf=0.95)
  emp.hpd(B1[,2],conf=0.95)
  emp.hpd(B1[,3],conf=0.95)
  emp.hpd(B1[,4],conf=0.95)
  emp.hpd(B1[,5],conf=0.95)
  emp.hpd(B1[,6],conf=0.95)
  emp.hpd(B2[,1],conf=0.95)
  emp.hpd(B2[,2],conf=0.95)
  emp.hpd(B2[,3],conf=0.95)
  emp.hpd(B2[,4],conf=0.95)
  emp.hpd(B2[,5],conf=0.95)
  emp.hpd(B2[,6],conf=0.95)  }
#-----

```

## RegressionFunctions.R

```
# Programa: RegressionFunctions.R
# En este programa se definen las funciones necesarias para la
# implementación de todos los programas.
#
# Autor(es): G. Nuñez-Antonio y
#            E. Gutiérrez-Peña < eduardo@sigma.iimas.unam.mx >
# Responsable: G. Nuñez-Antonio < gab.nuneza@gmail.com >
# Fecha: Julio, 2009.
#-----
CircReg<-function(mod1,mod2=mod1,data,tm,t.lag=1,burn=2000,flag.lk=FALSE)
{
# Esta función implementa un muestreo de Gibbs para obtener muestras
# de la distribución final de los parámetros de regresión en B.
#
# Argumentos:
#   mod1, mod2: Especificación de las componentes I y II, respectively.
#   data: Base de datos de tipo 'frame'.
#   tm: Tamaño de muestra total.
#   t.lag: La muestra es tomada cada t.lag-ésima iteración.
#   burn: Número de iteraciones tomadas como periodo de calentamiento
#         para el algoritmo de Gibbs.
#   flag.lk: indicador lógico; si se ajusta al valor TRUE, la medida
#           predictiva lk es calculada.
# Salida:
#   CircReg regresa una lista con las siguientes componentes:
#   BI: Una matriz que contiene una muestra de los coeficientes de la
#       componente I del modelo.
#   BII: Una matriz que contiene una muestra de los coeficientes de la
#        componente II del modelo.
#   Lk: El valor de la medida predictiva lk.
```

```

#
# Detalles:
# mod1 y mod2 son objetos de clase "formula". Ver, R (2010), por ejemplo.
# Si mode2 no se especifica, este asume la misma estructura que mod1.
# Si t.lag o burn o flag.lk no es especificado, asume el valor por
# omisión de 1, 2000 y FALSE, respectivamente.
# Si flag.lk no es TRUE, el valor final retornado de lk es lk=0.
#
# Autor(es): G. Nuñez-Antonio y
#             E. Gutiérrez-Peña < eduardo@sigma.iimas.unam.mx >
# Responsable: G. Nuñez-Antonio < gab.nuneza@gmail.com >
# Fecha: Julio, 2009.
#-----
library(MASS)
# Matrices de diseño
datose<-cbind(cos(data$theta),sin(data$theta))
n<-length(data$theta)
# Design Matrices
X.1.aux<-model.matrix(mod1,data)
p.1<-length(X.1.aux[1,])
X.1<-matrix(X.1.aux,ncol=p.1)
#
X.2.aux<-model.matrix(mod2,data)
p.2<-length(X.2.aux[1,])
X.2<-matrix(X.2.aux,ncol=p.2)
#
# Especificación inicial para la "matriz" B.
#
m1<-rep(0,p.1)
L1<-0.0001*diag(p.1)
m2<-rep(0,p.2)
L2<-0.0001*diag(p.2)
#

```

```

# Especificación final para la "matriz" B.
#
XtX.1<-crossprod(X.1)
XtX.2<-crossprod(X.2)
Lstar1<-L1+XtX.1
Lstar2<-L2+XtX.2
Sigma1<-chol2inv(chol(Lstar1))
Sigma2<-chol2inv(chol(Lstar2))
L1m1<-crossprod(L1,m1)
L2m2<-crossprod(L2,m2)
#
###----- MUESTREO DE GIBBS -----
#cat("Empieza el muestreador de Gibbs ....", fill = TRUE)
### Valores iniciales
r<-rep(1,n)
### Número de iteraciones
kk<-tm*t.lag
#print(paste(" Total iterations =", kk, " ..."))
### Matrices para obtener la muestra final.
B1<-matrix(0,tm,p.1)
B2<-matrix(0,tm,p.2)
#----- Periodo de calentamiento -----
for(k in 1:burn)
{
Y<-r*datose
# Sampling of vectors beta1 and beta2.
XtY1<-crossprod(X.1,Y[,1])
XtY2<-crossprod(X.2,Y[,2])
mstar1<-c( crossprod(Sigma1,((L1m1)+XtY1) ) )
mstar2<-c( crossprod(Sigma2,((L2m2)+XtY2) ) )
beta1<-mvrnorm(1,mstar1,Sigma1)
beta2<-mvrnorm(1,mstar2,Sigma2)
# Sampling of vector r

```

```

for(j in 1:n)
  {
    t.aux<-data$theta[j]
    mu.b1<-c(crossprod(beta1,X.1[j,]))
    mu.b2<-c(crossprod(beta2,X.2[j,]))
    b<-Dbd(t.aux,mu.b1,mu.b2)
    r[j]<-b+( pnorm(b)/( dnorm(b)+b*pnorm(b) ) )
  }
}
### ----- Iteraciones -----
for(k in 1:kk)
  {
    flag1<-(k/1000)-trunc(k/1000)
    if(flag1==0){print(k)}
    # Sampling of vectors beta1 and beta2.
    XtY1<-crossprod(X.1,Y[,1])
    XtY2<-crossprod(X.2,Y[,2])
    mstar1<-c( crossprod(Sigma1,((L1m1)+XtY1) ) )
    mstar2<-c( crossprod(Sigma2,((L2m2)+XtY2) ) )
    beta1<-mvrnorm(1,mstar1,Sigma1)
    beta2<-mvrnorm(1,mstar2,Sigma2)
    # Sampling of vector r
    for(j in 1:n)
      {
        t.aux<-data$theta[j]
        mu.b1<-c(crossprod(beta1,X.1[j,]))
        mu.b2<-c(crossprod(beta2,X.2[j,]))
        r[j]<-Metro(1lrt,t.aux,mu.b1,mu.b2,1,5)
      }
    Y<-r*datose
    #----- Valores de cada iteración -----
    flag<-(k/t.lag)-trunc(k/t.lag)
    if(flag==0)

```

```

{
ii<-k/t.lag

B1[ii,]<-beta1 B2[ii,]<-beta2 }
#----- Termina el algoritmo de Gibbs -----
}
#-----
# Cálculo de la medida predictiva Lk. Lk<-0.0
flag.LK<-ifelse(flag.lk=="TRUE",1,0) if(flag.LK){ mu<-matrix(0,tm,2)
predictiva<-rep(0,n) for(i in 1:n){
mu[,1]<-c(crossprod(t(B1),X.1[i,]))
mu[,2]<-c(crossprod(t(B2),X.2[i,])) norm2.mu.i<-norm2.row(mu)
vtmu.i<-c(crossprod(t(mu),datose[i,]))
predic.k<-(1/(2*pi))*exp(-0.5*norm2.mu.i)
          *( 1 + ((vtmu.i*pnorm(vtmu.i))/dnorm(vtmu.i)) )
predictiva[i]<-mean(predic.k)
}
Lk<-prod(predictiva)
}
#-----
# Salida:
B<-list(BI=B1,BII=B2,Lk=Lk)
drop(B)
}
#-----
#-----
#
Dbd<-function(t,mu1,mu2) { cos(t)*mu1+sin(t)*mu2 }
#
# Logaritmo natural del kernel de la densidad f(ln r/theta).
llrt<-function(y,t,mu1,mu2) {
          2*y-0.5*exp(y)*(exp(y)-2*Dbd(t,mu1,mu2) )
}

```



```

## Valores iniciales para el algoritmo de Metropolis.
#
media0<-function(t,mu1,mu2) {
  log( ( Dbd(t,mu1,mu2) + ( (Dbd(t,mu1,mu2)^2) + 8 ) ^0.5 )/2 )
}
#
var0<-function(m0) {
  ( 2 + exp(2*m0) )^(-1) }
#
# Norma Euclidiana al cuadrado de la matriz A, renglón por renglón.
norm2.row<-function(A){
  n<-length(A[,1])
  A2<-A^2
  norm2<-rep(0,n)
  norm2<-sapply((1:n),function(x){sum(A2[x,])})
  drop(norm
}
#-----
## Algoritmo de Metropolis.
Metro<-function(f,t,mu1,mu2,tamuestra,nodeite)
{
  N<-tamuestra
  ite<-nodeite
#
  m0<-media0(t,mu1,mu2)
  v0<-var0(m0)
  y0<-rnorm(N,m0,sqrt(v0))
#
  for (i in 1:ite)
  {
    y1<-rnorm(N,m0,sqrt(v0))
    lfy1 <- f(y1,t,mu1,mu2)
    ldny1 <- log(dnorm(y1,m0,sqrt(v0)))

```

```

w1<-(lfy1-ldny1)
lfy0 <- f(y0,t,mu1,mu2)
ldny0 <- log(dnorm(y0,m0,sqrt(v0)))
w0<-(lfy0 - ldny0)
lalpha<-(w1-w0)
u<-runif(N,0,1)
aux<-ifelse(log(u)<=lalpha,y1,y0)
y0<-aux
}
rdt<-exp(y0)
drop(rdt)
}
#-----
rNPxy.1<-function(M,V){
# Esta función simula una observación (r,theta)
# de una 2-dimensional distribución normal proyectada,
# con vector de medias M y matriz de covarianza V.
#
library(MASS)
xy<-mvrnorm(1,M,V)
rr<-sqrt(sum(xy^2))
ttheta<-(atan2(xy[2],xy[1]))%(2*pi)
xy.rt<-c(rr,ttheta)
drop(xy.rt)
}
#-----
circ.stats<-function(data,flag) {
# Esta función calcula la dirección media muestral (en grados)
# de una muestra de ángulos.
#
# data: Muestra de ángulos.
# flag: Si los datos están en grados, ajustar flag=0.
#       Si los datos están en radianes, ajustar flag=1.

```

```
#
if (flag==0) data<-data*(pi/180)
Cbarra<-mean(cos(data))
Sbarra<-mean(sin(data))
Rbarra<-sqrt(Cbarra^2 + Sbarra^2)
auxmean<-atan(Sbarra/Cbarra)
meanphi<-0*c(1:length(data))
if (Cbarra<0) meanphi<-auxmean+pi else {
if (auxmean<0) meanphi<-auxmean+(2*pi) else meanphi<-auxmean }
theta.barra<-meanphi*180/pi
drop(theta.barra)
}
```

---

```
#-----
```

## MissingData.R

```
# Programa MissingData.R
#
# Este programa simula una muestra de la distribución final de
# respuestas faltantes en el contexto de regresión. Ver, Capítulo 3,
# en particular ejemplo 3.3.
#
# Autor(es): G. Nuñez-Antonio y
#           E. Gutiérrez-Peña < eduardo@sigma.iimas.unam.mx >
# Responsable: G. Nuñez-Antonio < gab.nuneza@gmail.com >
# Fecha: Julio, 2009.
#-----
#rm(list=ls())
#.libPaths("/usr/local/R/contrib")
library(MASS)
library(CircStats)
source(file="RegressionFunctions.R")
#-----
### Datos
# Datos de caracoles (Periwinkles) sin los datos theta_13=197 (x_13=5),
# theta_23=75 (x_23=57) y theta_1=67 (x_1=107).

x<-c(#107,
46,33,67,122,69,43,30,12,25,37,69,#5,
83,68,38,21,1,71,60,71,71,#57,
53,38,70,7,48,7,21,27,
5,57,107)
thetaa<-c(#67,
66,74,61,58,60,100,89,171,166,98,60,#197,
98,86,123,165,133,101,105,71,84,#75,
98,83,71,74,91,38,200,56,
```

```

-1000,-1000,-1000)
theta.barra<- 91.162

#-----
Id2<-diag(2)
Id3<-diag(3)
Id4<-diag(4)

index.miss<-which(thetaa==-1000)
thetaa[index.miss]<-theta.barra
n.miss<-length(index.miss)

theta<-(pi/180)*thetaa
n<-length(theta)

datose<-matrix(0,n,2)
datose<-cbind(cos(theta),sin(theta))

# Especificación inicial para la "matriz" B.

m1<-c(0.0,0.0)
L1<-Id2*0.0001
m2<-c(0.0,0.0)
L2<-Id2*0.0001

# Matriz e diseño

X<-matrix(c(rep(1,n),x),n,2)
# X<-matrix(c(rep(1,n),x,x^2),n,3)

# Especificación final para la "matriz" B.

XtX<-t(X)%*%X

```

```

Lstar1<-L1+XtX
Lstar2<-L2+XtX
Sigma1<-solve(Lstar1)
Sigma2<-solve(Lstar2)
L1m1<-L1%%m1
L2m2<-L2%%m2

###----- MUESTREO DE GIBBS -----

### Valores iniciales

r<-rep(1,n)
Y<-r*datose

### Número de iteraciones

kk<-20000

### Matrices para recabar la muestra final (número de iteraciones/t.lag)

t.lag<-1
tm<-kk/t.lag
B1<-matrix(0,tm,2)
B2<-matrix(0,tm,2)
ttheta.miss<-rep(0,n.miss)
T.miss<-matrix(0,tm,n.miss)

#----- Periodo de calentamiento (10\%*iterations) -----

for(k in 1:(kk/10))
{
# Muestreo de los vectores beta1 y beta2.

```

```

XtY1<-t(X)%*%Y[,1]
XtY2<-t(X)%*%Y[,2]
mstar1<-c( Sigma1%*(L1m1 + XtY1) )
mstar2<-c( Sigma2%*(L2m2 + XtY2) )
beta1<-mvrnorm(1,mstar1,Sigma1)
beta2<-mvrnorm(1,mstar2,Sigma2)

# Muestreo de las respuestas faltantes, del vector r y del
# vector r.tilde.

for(j in 1:(n-n.miss) )
{
  t.aux<-theta[j]
  mu.b1<-c(beta1%*X[j,])
  mu.b2<-c(beta2%*X[j,])
  b<-Dbd(t.aux,mu.b1,mu.b2)
  r[j]<-b+( pnorm(b)/( dnorm(b)+b*pnorm(b) ) )
}
for(jj in (n-n.miss+1):n )
{
  mu.b1<-c(beta1%*X[jj,])
  mu.b2<-c(beta2%*X[jj,])
  rt.miss<-rNPxy.1(c(mu.b1,mu.b2),Id2)
  r[jj]<-rt.miss[1]
  datose[jj,]<-c(cos(rt.miss[2]),sin(rt.miss[2]))
}
Y<-r*datose
}
### ----- Iteraciones -----

for(k in 1:kk)
{
  flag1<-(k/1000)-trunc(k/1000)

```

```

if(flag1==0){print(k)}

# Muestreo de los vectores beta1 y beta2.

XtY1<-t(X)%*%Y[,1]
XtY2<-t(X)%*%Y[,2]
mstar1<-c( Sigma1%*(L1m1 + XtY1) )
mstar2<-c( Sigma2%*(L2m2 + XtY2) )
beta1<-mvrnorm(1,mstar1,Sigma1)
beta2<-mvrnorm(1,mstar2,Sigma2)

# Muestreo de las respuestas faltantes, del vector r y del vector r.tilde.

for(j in 1:(n-n.miss) )
{
  t.aux<-theta[j]
  mu.b1<-c(beta1%*X[j,])
  mu.b2<-c(beta2%*X[j,])
  r[j]<-Metro(1lrt,t.aux,mu.b1,mu.b2,1,10)
}
ii<-k
for(jj in (n-n.miss+1):n )
{
  mu.b1<-c(beta1%*X[jj,])
  mu.b2<-c(beta2%*X[jj,])
  rt.miss<-rNPxy.1(c(mu.b1,mu.b2),Id2)
  ttheta.miss[jj]<-rt.miss[2]
  r[jj]<-rt.miss[1]
  datose[jj,]<-c(cos(rt.miss[2]),sin(rt.miss[2]))
  T.miss[ii,jj-(n-n.miss)]<-ttheta.miss[jj]
  rt.miss<-c(0.0,0.0)
}
Y<-r*datose

```



```

#----- Valores de cada iteración -----
flag<-(k/t.lag)-trunc(k/t.lag)
if(flag==0)
{
ii<-k/t.lag
B1[ii,]<-beta1
B2[ii,]<-beta2
}
#----- Termina el algoritmo de Gibbs -----
}
###----- Salidas -----
if(0){
# Figura 3.4
nn<-10000
mu.star<-matrix(0,tm,2)
xy.star1<-rep(0,nn)
xy.star2<-rep(0,nn)
theta.star<-matrix(0,n,nn)
theta.star.q<-matrix(0,n,2)
medias.star<-rep(0,n)
for(j in 1:n)
{
mu.star[,1]<-c(B1%*%X[j,])
mu.star[,2]<-c(B2%*%X[j,])
xy.star1<-rnorm(nn,mean(mu.star[,1]),1)
xy.star2<-rnorm(nn,mean(mu.star[,2]),1)
theta.star[j,]<-atan2(xy.star2,xy.star1)%%(2*pi)
medias.star[j]<-circ.stats(theta.star[j,],1)
theta.star.q[j,]<-quantile(theta.star[j,],probs=c(0.05,0.975),
names=F)
}
plot(x,theta*180/pi,ylim=c(0,360),col="red",pch=19,
ylab=expression(theta),main="Predictive Distribution")

```

```

points(x,medias.star,col="blue",pch=19)
for(j in 1:n)
{
lines(rep(x[j],2),theta.star.q[j,]*180/pi,col="red",lty="dotted")
}
bringToTop()
}
#-----
if(0){
# Figura 3.5
par(mfrow=c(3,2))
hist(T.miss[,1]*180/pi,col="black",border="white",
xlab=expression(theta[13]: 197),main="",freq=F ,ylab="")
rose.diag(T.miss[,1],main="",bins=15,prop=2)
hist(T.miss[,2]*180/pi,col="black",border="white",
xlab=expression(theta[23]: 75),main="" ,freq=F, ylab="")
rose.diag(T.miss[,2],main="",bins=20,prop=1.3)
hist(T.miss[,3]*180/pi,col="black",border="white",
xlab=expression(theta[1]: 67),main="" ,freq=F, ylab="")
rose.diag(T.miss[,3],main="",bins=20,prop=1.3)
bringToTo
}
#-----
# Última: Actualización, Julio 2009, -gna.

```

## Normal Bivariada Proyectada

A continuación se presenta el Programa Normal-proyectada.nb, este código se muestra respetando la tipografía de Mathematica 5.2 de tal manera que se puede correr directamente, sin necesidad de editarlo.

---

### Densidad de $\Theta$ bajo una Normal Bivariada Proyectada

El Caso General :

$$C1[s1_, s2_, rho_] := \frac{1}{2\pi s1 s2 (1-\rho^2)^{0.5}}$$

$$C2[mu_, v_, s1_, s2_, rho_] := \left(\frac{\mu}{s1}\right)^2 + \left(\frac{v}{s2}\right)^2 - \frac{2\rho\mu v}{s1 s2}$$

$$C3[t_, s1_, s2_, rho_] := \left(\frac{\text{Cos}[t]}{s1}\right)^2 + \left(\frac{\text{Sin}[t]}{s2}\right)^2 - \frac{\rho \text{Sin}[2t]}{s1 s2}$$

$$C4[t_, mu_, v_, s1_, s2_, rho_] := \frac{s2^2 \text{Cos}[t] \mu + s1^2 \text{Sin}[t] v - \rho s1 s2 (\text{Cos}[t] v + \text{Sin}[t] \mu)}{(s1 s2)^2}$$

$$C6[t_, mu_, v_, s1_, s2_, rho_] := \frac{C1[s1, s2, \rho] e^{-\frac{C2[\mu, v, s1, s2, \rho]}{2(1-\rho^2)}} (1-\rho^2)}{C3[t, s1, s2, \rho]}$$

$$d2[t\_ , s1\_ , s2\_ , rho\_ ] := \frac{C3[t, s1, s2, \rho]}{1 - \rho^2}$$

$$b[t\_ , mu\_ , v\_ , s1\_ , s2\_ , rho\_ ] := \frac{C4[t, \mu, v, s1, s2, \rho]}{C3[t, s1, s2, \rho]}$$

$$Dbd[t\_ , mu\_ , v\_ , s1\_ , s2\_ , rho\_ ] := d2[t, s1, s2, \rho]^{0.5} b[t, \mu, v, s1, s2, \rho]$$

$$\text{DensNormal}[x] := \frac{e^{-\frac{x^2}{2}}}{(2\pi)^{0.5}}$$

$$\text{PhiD}[t\_ , mu\_ , v\_ , s1\_ , s2\_ , rho\_ ] := \int_{-\infty}^{\text{Dbd}[t, \mu, v, s1, s2, \rho]} \text{DensNormal}[x] dx$$

$$ftheta[t\_ , mu\_ , v\_ , s1\_ , s2\_ , rho\_ ] := C6[t, \mu, v, s1, s2, \rho] (1 + (2\pi)^{0.5} \text{Dbd}[t, \mu, v, s1, s2, \rho]$$

$$e^{0.5 \text{Dbd}[t, \mu, v, s1, s2, \rho]^2} \text{PhiD}[t, \mu, v, s1, s2, \rho])$$

$$\text{Normal2}[x\_ , y\_ , mu\_ , v\_ , s1\_ , s2\_ , rho\_ ] := C1[s1, s2, \rho] e^{-\frac{\left(\frac{x-\mu}{s1}\right)^2 + \left(\frac{y-v}{s2}\right)^2 - \frac{2\rho(x-\mu)(y-v)}{s1s2}}{2(1-\rho^2)}}$$

```
graf6 = Plot[ftheta[t, 0, 1, 1, 3, 0.3], {t, -π, π}, PlotLabel → "Normal Proyectada"];
```

```
graf3 = ContourPlot[Normal2[x, y, 0, 1, 1, 3, 0.3], {x, -5, 5}, {y, -10, 10},
```

```
ContourShading → False, PlotLabel → "Normal(0,1,1,3,0.3)", DisplayFunction → Identity]
```

```
Circ1[x_]:= (1 - x2)0.5
```

```
Circ2[x_]:= - (1 - x2)0.5
```

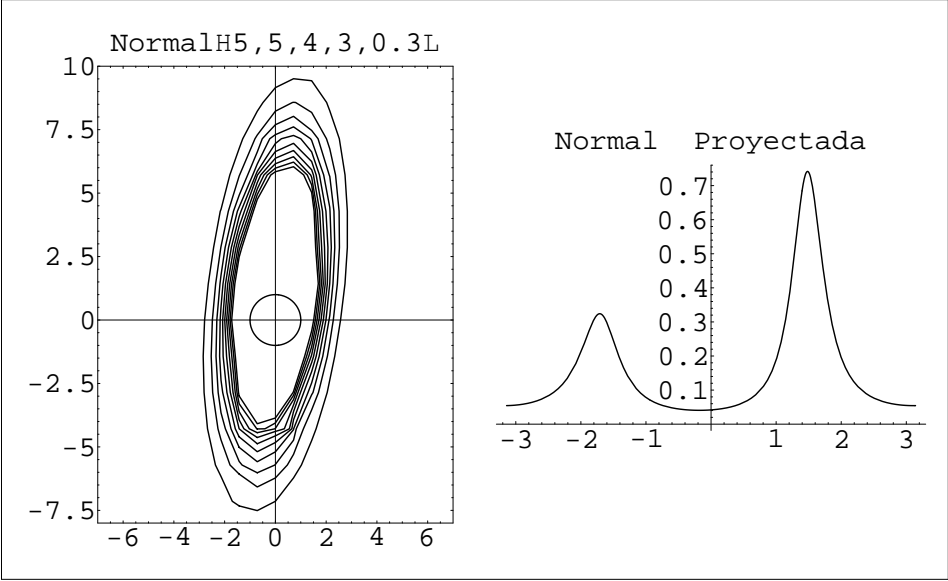
```
graf4 = Plot[Circ1[x], {x, -1, 1}, DisplayFunction → Identity]
```

```
graf5 = Plot[Circ2[x], {x, -1, 1}, DisplayFunction → Identity]
```

```
gmix = Show[graf3, graf4, graf5, DisplayFunction → $DisplayFunction, Frame → True,
```

```
Axes → True, PlotRange → {{-7, 7}, {-8, 10}}, AspectRatio → Automatic]
```

```
Show[GraphicsArray[{gmix, graf6}], Frame → True]
```



## LCP.R

```
# Programa: LCP.R
# Este programa es igual al MN11.R (basado en el MN1.R)
#-----
# Este programa implementa el muestreo de Gibbs para el modelo LCP.
# Los datos simulados fueron generados con el programa SimulateData2.R
# Notas:
# 1. El programa simula en bloque las beta's y los coeficientes bi's.
#    Ver, algoritmo 2 de Chib y Carlin (1999).
# 2. El programa es tal que cada componente
#    puede o no tener efectos aleatorios.
# 3. El programa considera Xi diferentes y Zi=Z para toda i.
#    Así, se puede implementar el ejercicio 2 del capítulo 4,
#    referente a los datos de sanhoppers.
#-----
rm(list=ls())
set.seed(150)
time.ini<-Sys.time()
#-----
library(MASS)
library(bayesSurv)
#-----
source("S:/Nightwish/Example2/Basicfunctions0.R")
#-----
# Data:
sh<-read.csv(file="S:/Docto09/Datasets/DElia/Sandhoppers/sh.csv",row.names=1)
Theta<-sh[,c(2:6)]*pi/180
N<-length(Theta[,1])
n<-length(Theta[1,])
DataC<-cos(Theta)
DataS<-sin(Theta)
#-----
```

```

# Matrices de diseño para cada componente.
p1<-7 # Dimension del vector de efectos fijos.
p2<-7
q2<-1 # Dimension del vector de efectos aleatorios.
XI<-array(0,c(N,n,p1))
XII<-array(0,c(N,n,p2))
for(i in 1:N){
XI[i, ,1]<-rep(1,n)
XI[i, ,2]<-rep(sh$sun[i],n)
XI[i, ,3]<-rep(sh$eye[i],n)
XI[i, ,4]<-rep(sh$w1[i],n)
XI[i, ,5]<-rep(sh$w2[i],n)
XI[i, ,6]<-rep(sh$w3[i],n)
XI[i, ,7]<-c(1:5)
XII[i, ,1]<-rep(1,n)
XII[i, ,2]<-rep(sh$sun[i],n)
XII[i, ,3]<-rep(sh$eye[i],n)
XII[i, ,4]<-rep(sh$w1[i],n)
XII[i, ,5]<-rep(sh$w2[i],n)
XII[i, ,6]<-rep(sh$w3[i],n)
XII[i, ,7]<-c(1:5) }
Z<-matrix(c(rep(1,n)))
#-----
# Parámetros de la especificación inicial.
A1<-matrix(0,p1,p1)
A2<-matrix(0,p2,p2)
v2<-q2
B2<-0.001
#-----
# Algunos objetos para la especificación de las distribuciones finales.
# En este caso, son iguales para cada componente.
XtX.I<-0.0
XtX.II<-0.0

```



```

for(i in 1:N){
XtX.I<-t(XI[i,,])%*%XI[i,,]+ XtX.I      # sum of { t(XI)*XI }
XtX.II<-t(XII[i,,])%*%XII[i,,]+ XtX.II  # sum of { t(XII)*XII }
}
ZtZ<-c(t(Z)%*%Z)
#-----
# Algunas funciones para el algoritmo de Gibbs.
betaI.f<-function(Y){
# Simulación del vector beta I
e<-matrix(0,N,n)
C<-XtX.I + A1
#e<-Y-t(Z%*%t(b))
e<-Y
invC<-chol2inv(chol(C))
sXte<-rowSums( sapply( (1:N), function(w){t(XI[w,,])%*%t(e[w,,])} ) )
betaF<-c(invC%*%sXte)
beta.aux<-mvrnorm(1,betaF,invC)
drop(beta.aux)
}
#-----
betaIIBlock.f<-function(Omega,Y){
# Simulación en bloque para el vector beta II
# Lo siguiente es posible ya que en este caso Vi=V ( Zi= Z para toda i ).
invOmega<-(1.0/Omega)
Vi<-(Z%*%invOmega%*%t(Z))+ diag(n)
invVi<-chol2inv(chol(Vi))
#
sXtVX<-diag(p2)*0.0
aux<-lapply( (1:N), function(w){t(XII[w,,])%*%invVi%*%XII[w,,]} )
for(i in 1:N){sXtVX<-sXtVX+aux[[i]]}
Var.beta<-chol2inv(chol(A2 + sXtVX ))
sXtVY<-rowSums( sapply( (1:N), function(w){t(XII[w,,])%*%invVi%*%t(Y[w,,])} ) )
#

```

```

betaF<-c( Var.beta%%sXtVY )
beta.aux<-mvrnorm(1,betaF,Var.beta)
drop(beta.aux)
}
#-----
if(0){
b.f<-function(Omega,beta,Y){
# Sampling of b (i.e. {bi} vectors, i=1,...,N)
# Lo siguiente es posible ya que en este caso Zi = Z para toda i.
# Nota: Este código es pensando matricialmente.
#
etilde<-matrix(0,N,n)
b.aux<-matrix(0,N,q2)
bF<-matrix(0,N,q2)
#
D<-(ZtZ) + Omega
invD<-solve(D)
#
etilde<- Y - t( sapply( (1:N), function(w){X[w,,]%%beta} ) )
#
bF<-t(invD%%t(Z)%%t(etilde))
b.aux<- as.matrix( sapply( (1:N), function(w){mvrnorm(1,bF[w,],invD)} ) )
# Nota: Si la dimensión de b, cada {bi}, es mayor a 1, entonces de debe usar:
# b.aux<-t ( as.matrix( sapply( (1:N), function(w){mvrnorm(1,bF[w,],invD)} ) ) )
drop(b.aux)
} }
b.f<-function(Omega,beta,Y){
# Muestreo de b (i.e. {bi} vectores, i=1,...,N)
# Lo siguiente es posible ya que en este caso Zi = Z para toda i.
#
etilde<-matrix(0,N,n)
b.aux<-c(1:N)*0.0
bF<-c(1:N)*0.0

```

```

D<-(ZtZ) + Omega
invD<-(1.0/D)
etilde<- Y - t( sapply( (1:N), function(w){XII[w,]%*%beta}  ) )
bF<-as.vector(t(invD*t(Z)%*%t(etilde)))
b.aux<- mvrnorm(1,bF,diag(N)*invD)
drop(b.aux)
}
#-----
Omega.f<-function(b){
# Muestreo de Omega
#
bb.sum<-c(t(b)%*%b)
vc<-1.0/(B2+bb.sum)
Omega.aux<-rWishart(1,v2+N,vc)
drop(Omega.aux)
}
#-----
# Empieza el muestreo de Gibbs.
#-----
# Valores iniciales para el muestreo de Gibbs.
# En este caso estos se toman iguales para cada una de las componentes.
b<-c(1:N)*0.0
Omega<-1.0
R<-matrix(1,N,n)
#-----
# Iterations: kk
kk<-5000
burn<-1000
#-----
# Matrices para recabar la muestra final de tamaño tm.
n.lag<-5
tm<-kk/n.lag
Beta.I<-matrix(0,tm,p1)

```

```

Beta.II<-matrix(0,tm,p2)
VCov<-c(1:tm)*0.0
#-----
# Periodo de Calentamiento.
Omega.II<-Omega
b.II<-b
#
for(k in 1:burn)
{
Y.I<-R*DataC
Y.II<-R*DataS
#Simulations for beta, bi's, Omega.
beta.I<-betaI.f(Y.I)
beta.II<-betaIIBlock.f(Omega.II,Y.II)
b.II<-b.f(Omega.II,beta.II,Y.II)
Omega.II<-Omega.f(b.II)
#Simulations for each (i,j) of R.
for(i in 1:N)
{
for(j in 1:n)
{
t.aux<-Theta[i,j]
mu.ij.I<-c(beta.I%%XI[i,j,])
mu.ij.II<-c(beta.II%%XII[i,j,] + b.II[i])
bb<-Dbd(t.aux,mu.ij.I,mu.ij.II)
R[i,j]<- bb + ( pnorm(bb)/ ( dnorm(bb)+(bb*pnorm(bb)) ) )
}
}
#-----
# El periodo de calentamiento es completado.
}
#-----
#Comienzan las iteraciones.

```

```

for(k in 1:kk)
{
if((k%%500)==0) {print(k)}
# if(k%%100==0) metro.ite<-10 else metro.ite<-3
# Simulaciones para beta, bi's y Omega.
beta.I<-betaI.f(Y.I)
beta.II<-betaIIBlock.f(Omega.II,Y.II)
b.II<-b.f(Omega.II,beta.II,Y.II)
Omega.II<-Omega.f(b.II)
# Simulaciones para cada elemento (i,j) de la matriz R.
for(i in 1:N)
{
for(j in 1:n)
{
t.aux<-Theta[i,j]
mu.ij.I<-c(beta.I%%XI[i,j,])
mu.ij.II<-c(beta.II%%XII[i,j,] + b.II[i])
R[i,j]<- Metro(1l1rt,t.aux,mu.ij.I,mu.ij.II,1,2)
}
}
Y.I<-R*cos(Theta)
Y.II<-R*sin(Theta)
#-----
# Valores de cada iteración.
#flag<-(k/n.lag)-trunc(k/n.lag)
if((k%%n.lag)==0)
{
ii<-k/n.lag
Beta.I[ii,]<-beta.I
Beta.II[ii,]<-beta.II
VCov[ii]<-1/Omega.II
}
#-----

```

```

}
# Terminan las iteraciones
#-----
#Out
#write("S:/Nightwish/Example2/Beta.I,file="Beta.I")
#write("S:/Nightwish/Example2/Beta.II,file="Beta.II")
#write("S:/Nightwish/Example2/VCov,file="VCov")
time.fin<-Sys.time()
time.tot<-time.fin-time.ini
time.tot
#-----
#Windows Outs
if(0){
par(mfrow=c(4,2))
hist(Beta.I[,1],freq=F,xlab=expression(beta[0]^1),main=" ")
hist(Beta.I[,2],freq=F,xlab=expression(beta[1]^1),main=" ")
hist(Beta.I[,3],freq=F,xlab=expression(beta[2]^1),main=" ")
hist(Beta.I[,4],freq=F,xlab=expression(beta[3]^1),main=" ")
hist(Beta.I[,5],freq=F,xlab=expression(beta[4]^1),main=" ")
hist(Beta.I[,6],freq=F,xlab=expression(beta[5]^1),main=" ")
hist(Beta.I[,7],freq=F,xlab=expression(beta[6]^1),main=" ")
par(mfrow=c(4,2))
hist(Beta.II[,1],freq=F,xlab=expression(beta[0]^2),main=" ")
hist(Beta.II[,2],freq=F,xlab=expression(beta[1]^2),main=" ")
hist(Beta.II[,3],freq=F,xlab=expression(beta[2]^2),main=" ")
hist(Beta.II[,4],freq=F,xlab=expression(beta[3]^2),main=" ")
hist(Beta.II[,5],freq=F,xlab=expression(beta[4]^2),main=" ")
hist(Beta.II[,6],freq=F,xlab=expression(beta[5]^2),main=" ")
hist(Beta.II[,7],freq=F,xlab=expression(beta[6]^2),main=" ")
hist(VCov,freq=F,xlab=expression((sigma^2)),main=" ")
med.erg1<-cbind(cumsum(Beta.I[,1]),cumsum(Beta.I[,2]),
cumsum(Beta.I[,3]),cumsum(Beta.I[,4]),cumsum(Beta.I[,5]),
cumsum(Beta.I[,6]),cumsum(Beta.I[,7]),cumsum(Beta.I[,1]),

```

```

cumsum(Beta.I[,2]), cumsum(Beta.I[,3]), cumsum(Beta.I[,4]),
cumsum(Beta.I[,5]), cumsum(Beta.I[,6]), cumsum(Beta.I[,7]), VCov )/(1:tm)
par(mfrow=c(4,2))
plot(med.erg1[,1], type="l", xlab="iteraciones", ylab=expression(beta[0]^1))
plot(med.erg1[,2], type="l", xlab="iteraciones", ylab=expression(beta[1]^1))
plot(med.erg1[,3], type="l", xlab="iteraciones", ylab=expression(beta[2]^1))
plot(med.erg1[,4], type="l", xlab="iteraciones", ylab=expression(beta[3]^1))
plot(med.erg1[,5], type="l", xlab="iteraciones", ylab=expression(beta[4]^1))
plot(med.erg1[,6], type="l", xlab="iteraciones", ylab=expression(beta[5]^1))
plot(med.erg1[,7], type="l", xlab="iteraciones", ylab=expression(beta[6]^1))
par(mfrow=c(4,2))
plot(med.erg1[,8], type="l", xlab="iteraciones", ylab=expression(beta[0]^2))
plot(med.erg1[,9], type="l", xlab="iteraciones", ylab=expression(beta[1]^2))
plot(med.erg1[,10], type="l", xlab="iteraciones", ylab=expression(beta[2]^2))
plot(med.erg1[,11], type="l", xlab="iteraciones", ylab=expression(beta[3]^2))
plot(med.erg1[,12], type="l", xlab="iteraciones", ylab=expression(beta[4]^2))
plot(med.erg1[,13], type="l", xlab="iteraciones", ylab=expression(beta[5]^2))
plot(med.erg1[,14], type="l", xlab="iteraciones", ylab=expression(beta[6]^2))
plot(med.erg1[,15], type="l", xlab="iteraciones", ylab=expression(sigma^2))
par(mfrow=c(4,2))
acf(Beta.I[,1])
acf(Beta.I[,2])
acf(Beta.I[,3])
acf(Beta.I[,4])
acf(Beta.I[,5])
acf(Beta.I[,6])
acf(Beta.I[,7])
par(mfrow=c(4,2))
acf(Beta.II[,1])
acf(Beta.II[,2])
acf(Beta.II[,3])
acf(Beta.II[,4])
acf(Beta.II[,5])

```

```

acf(Beta.II[,6])
acf(Beta.II[,7])
acf(VCov)
library(TeachingDemos)
emp.hpd(Beta.I[,1],conf=0.95)
emp.hpd(Beta.I[,2],conf=0.95)
emp.hpd(Beta.I[,3],conf=0.95)
emp.hpd(Beta.I[,4],conf=0.95)
emp.hpd(Beta.I[,5],conf=0.95)
emp.hpd(Beta.I[,6],conf=0.95)
emp.hpd(Beta.I[,7],conf=0.95)
emp.hpd(Beta.II[,1],conf=0.95)
emp.hpd(Beta.II[,2],conf=0.95)
emp.hpd(Beta.II[,3],conf=0.95)
emp.hpd(Beta.II[,4],conf=0.95)
emp.hpd(Beta.II[,5],conf=0.95)
emp.hpd(Beta.II[,6],conf=0.95)
emp.hpd(Beta.II[,7],conf=0.95)
}
# Última Actualización: Febrero 11, 2010. -gna

```



# Apéndice C

## Datos de Tortugas ( $\theta$ )

---

8	9	13	13	14	18	22	27	30	34	38	38	40	44	45
47	48	48	48	48	50	53	56	57	58	58	61	63	64	64
64	65	65	68	70	73	78	78	78	83	83	88	88	88	90
92	92	93	95	96	98	100	103	106	113	118	138	153	153	155
204	215	223	226	237	238	243	244	250	251	257	268	285	319	343
350														

---

## Datos: Ejemplo 3.1

	$\theta$	x	x <sup>2</sup>
1	2.57	0.00	0.0000
2	1.98	0.02	0.0004
3	2.78	0.04	0.0016
4	2.32	0.06	0.0036
5	2.86	0.08	0.0064
6	2.36	0.10	0.01
7	2.81	0.12	0.01
8	2.78	0.14	0.02
9	2.74	0.16	0.03
10	2.83	0.18	0.03
11	3.26	0.20	0.04
12	2.75	0.22	0.05
13	2.90	0.24	0.06
14	2.80	0.26	0.07
15	3.54	0.28	0.08
16	3.19	0.30	0.09
17	2.95	0.32	0.10
18	3.09	0.34	0.12
19	2.64	0.36	0.13
20	2.55	0.38	0.15
21	2.47	0.40	0.16
22	2.85	0.42	0.18
23	2.65	0.44	0.20
24	3.10	0.46	0.22
25	2.04	0.48	0.24

	$\theta$	x	x <sup>2</sup>
26	2.52	0.51	0.26
27	2.50	0.53	0.28
28	2.41	0.55	0.30
29	2.27	0.57	0.32
30	2.12	0.59	0.34
31	2.49	0.61	0.37
32	2.14	0.63	0.39
33	2.13	0.65	0.42
34	2.02	0.67	0.44
35	2.23	0.69	0.47
36	2.16	0.71	0.50
37	3.02	0.73	0.53
38	3.18	0.75	0.56
39	2.93	0.77	0.59
40	2.40	0.79	0.62
41	2.18	0.81	0.65
42	2.45	0.83	0.69
43	2.80	0.85	0.72
44	2.63	0.87	0.75
45	2.59	0.89	0.79
46	2.35	0.91	0.83
47	2.23	0.93	0.86
48	2.15	0.95	0.90
49	1.87	0.97	0.94
50	2.59	0.99	0.98

*continuación . . .*

	$\theta$	x	x2		$\theta$	x	x2
51	2.45	1.01	1.02	76	1.67	1.52	2.30
52	2.27	1.03	1.06	77	1.88	1.54	2.36
53	2.42	1.05	1.10	78	2.17	1.56	2.42
54	2.06	1.07	1.15	79	1.70	1.58	2.48
55	2.42	1.09	1.19	80	1.72	1.60	2.55
56	1.97	1.11	1.23	81	1.99	1.62	2.61
57	2.47	1.13	1.28	82	1.98	1.64	2.68
58	2.31	1.15	1.33	83	1.84	1.66	2.74
59	1.81	1.17	1.37	84	1.81	1.68	2.81
60	2.04	1.19	1.42	85	1.58	1.70	2.88
61	2.14	1.21	1.47	86	1.93	1.72	2.95
62	1.98	1.23	1.52	87	1.76	1.74	3.02
63	1.87	1.25	1.57	88	1.60	1.76	3.09
64	2.26	1.27	1.62	89	1.58	1.78	3.16
65	2.18	1.29	1.67	90	1.88	1.80	3.23
66	2.02	1.31	1.72	91	1.83	1.82	3.31
67	1.81	1.33	1.78	92	1.68	1.84	3.38
68	2.04	1.35	1.83	93	1.57	1.86	3.45
69	1.74	1.37	1.89	94	1.59	1.88	3.53
70	1.90	1.39	1.94	95	1.61	1.90	3.61
71	1.86	1.41	2.00	96	1.54	1.92	3.6864
72	1.52	1.43	2.06	97	1.67	1.94	3.7636
73	1.83	1.45	2.12	98	1.82	1.96	3.8416
74	1.94	1.47	2.17	99	1.89	1.98	3.9204
75	1.74	1.49	2.23	100	1.62	2.00	4.0000

## Datos: Ejemplo 3.2

	$\theta$	$x$
1	4.64	-1.00
2	5.39	-0.98
3	4.52	-0.96
4	4.80	-0.94
5	5.36	-0.92
6	4.72	-0.90
7	5.13	-0.88
8	4.72	-0.86
9	4.93	-0.84
10	4.64	-0.82
11	4.81	-0.80
12	5.01	-0.78
13	4.76	-0.76
14	5.04	-0.74
15	5.34	-0.72
16	5.05	-0.70
17	5.22	-0.68
18	5.11	-0.66
19	4.78	-0.64
20	4.79	-0.62
21	5.08	-0.60
22	5.40	-0.58
23	5.58	-0.56
24	4.52	-0.54
25	5.12	-0.52

	$\theta$	$x$
26	4.82	-0.49
27	5.05	-0.47
28	5.65	-0.45
29	4.84	-0.43
30	5.09	-0.41
31	5.11	-0.39
32	5.76	-0.37
33	5.38	-0.35
34	4.79	-0.33
35	4.96	-0.31
36	5.10	-0.29
37	5.54	-0.27
38	4.93	-0.25
39	5.26	-0.23
40	5.66	-0.21
41	5.73	-0.19
42	6.28	-0.17
43	5.89	-0.15
44	5.56	-0.13
45	5.58	-0.11
46	6.20	-0.09
47	5.65	-0.07
48	6.00	-0.05
49	0.14	-0.03
50	6.26	-0.01

*continuación . . .*

	$\theta$	x		$\theta$	x
51	0.15	0.01	76	1.38	0.52
52	0.70	0.03	77	1.68	0.54
53	0.32	0.05	78	1.29	0.56
54	0.49	0.07	79	0.91	0.58
55	0.93	0.09	80	1.06	0.60
56	0.86	0.11	81	0.77	0.62
57	1.10	0.13	82	1.26	0.64
58	0.17	0.15	83	1.54	0.66
59	0.35	0.17	84	1.41	0.68
60	0.46	0.19	85	1.23	0.70
61	0.21	0.21	86	1.82	0.72
62	0.39	0.23	87	1.72	0.74
63	0.98	0.25	88	1.52	0.76
64	1.12	0.27	89	1.11	0.78
65	1.25	0.29	90	0.90	0.80
66	0.80	0.31	91	1.71	0.82
67	0.87	0.33	92	1.79	0.84
68	0.96	0.35	93	1.34	0.86
69	1.45	0.37	94	1.63	0.88
70	1.10	0.39	95	1.92	0.90
71	1.86	0.41	96	1.74	0.92
72	1.72	0.43	97	2.06	0.94
73	0.84	0.45	98	1.24	0.96
74	1.20	0.47	99	1.11	0.98
75	1.41	0.49	100	1.47	1.00

### Datos: Ejemplo 3.3

	$\theta$	x	x2	x3
1	1.17	107	11449	1225043
2	1.15	46	2116	97336
3	1.29	33	1089	35937
4	1.06	67	4489	300763
5	1.01	122	14884	1815848
6	1.05	69	4761	328509
7	1.75	43	1849	79507
8	1.55	30	900	27000
9	2.98	12	144	1728
10	2.90	25	625	15625
11	1.71	37	1369	50653
12	1.05	69	4761	328509
13	3.44	5	25	125
14	1.71	83	6889	571787
15	1.50	68	4624	314432
16	2.15	38	1444	54872

	$\theta$	x	x2	x3
17	2.88	21	441	9261
18	2.32	1	1	1
19	1.76	71	5041	357911
20	1.83	60	3600	216000
21	1.24	71	5041	357911
22	1.47	71	5041	357911
23	1.31	57	3249	185193
24	1.71	53	2809	148877
25	1.45	38	1444	54872
26	1.24	70	4900	343000
27	1.29	7	49	343
28	1.59	48	2304	110592
29	0.66	7	49	343
30	3.49	21	441	9261
31	0.98	27	729	19683

## Descripción de datos: Ejemplo 3.4

	fecha	hora	rh	tmp	$\theta$	wsp	hora.rad	sen(hora)	cos(hora)
1	2007-01-01	3	77.900	10.450	5.411	2.370	0.785	0.707	0.707
2	2007-01-01	5	80.270	10.110	5.480	2.090	1.309	0.966	0.259
3	2007-01-02	2	85.480	8.070	4.294	0.850	0.524	0.500	0.866
4	2007-01-02	13	53.270	17.760	1.187	1.850	3.403	-0.259	-0.966
5	2007-01-02	23	68.120	10.960	4.171	0.770	6.021	-0.259	0.966
:	:	:	:	:	:	:	:	:	:
250	2007-04-10	9	44.320	15.280	0.559	0.260	2.356	0.707	-0.707
251	2007-04-11	5	43.550	8.480	3.927	0.990	1.309	0.966	0.259
252	2007-04-11	6	45.180	7.550	4.049	1.060	1.571	1.000	0.000
253	2007-04-11	13	19.330	23.040	0.192	1.370	3.403	-0.259	-0.966
:	:	:	:	:	:	:	:	:	:
477	2007-06-29	21	79.230	15.230	5.498	2.360	5.498	-0.707	0.707
478	2007-06-30	1	81.650	13.660	4.032	0.680	0.262	0.259	0.966
479	2007-06-30	14	49.000	23.320	0.436	0.470	3.665	-0.500	-0.866
480	2007-06-30	15	57.020	20.620	5.201	2.110	3.927	-0.707	-0.707

## Datos: Ejemplo 4.1

	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
1	63.12	63.22	63.04	61.95	62.67
2	63.54	63.39	62.30	62.45	61.57
3	63.13	62.57	62.49	61.97	61.19
4	63.59	63.04	63.12	62.55	61.83
5	63.32	63.52	63.47	64.16	63.67
6	63.53	63.09	62.98	63.27	63.30
7	63.86	63.22	63.31	62.73	62.33
8	63.16	63.91	64.03	64.23	64.00
9	63.24	62.64	62.42	62.35	61.95
10	63.52	62.93	63.35	62.47	62.16
11	63.29	63.41	63.08	62.13	61.46
12	63.53	63.15	63.37	62.82	62.86
13	63.42	63.10	62.37	61.71	61.51
14	63.79	62.85	62.68	62.24	61.96
15	63.71	63.16	62.24	61.04	60.51
16	63.46	63.67	63.24	63.39	62.92
17	63.35	63.51	63.09	62.23	62.55
18	63.53	63.47	63.63	63.01	63.97
19	63.47	62.58	62.40	62.40	62.22
20	63.58	64.20	63.77	64.00	64.47
21	63.56	63.06	62.61	61.66	61.56
22	63.15	62.99	62.72	62.57	63.17
23	63.48	63.31	63.47	63.11	62.59
24	63.42	63.52	63.87	63.75	64.39
25	63.38	63.37	62.55	63.12	62.52
26	63.31	63.04	62.38	61.91	60.91
27	63.55	63.23	63.66	63.93	63.85
28	63.38	63.07	62.93	61.97	61.66
29	63.44	62.70	62.81	61.77	61.65
30	63.75	63.28	63.15	62.49	62.00



continuación . . .

	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
31	63.40	63.24	63.04	62.78	62.47
32	63.47	63.11	62.78	62.54	62.27
33	63.94	63.00	62.60	62.75	62.68
34	63.30	63.01	61.87	61.12	59.72
35	63.53	63.11	62.58	62.88	62.15
36	63.87	63.73	64.24	64.38	64.80
37	63.54	62.39	61.25	60.25	58.48
38	63.41	63.15	62.60	61.86	61.24
39	63.86	63.07	62.42	62.31	61.75
40	63.68	63.12	62.37	61.42	59.68
41	63.44	62.56	61.21	60.27	58.85
42	63.64	63.37	63.08	63.35	63.04
43	63.78	62.93	63.10	62.54	61.96
44	63.33	63.23	62.96	62.49	62.80
45	63.45	63.05	63.36	63.12	62.65
46	63.55	62.95	62.65	61.63	61.23
47	63.44	62.45	62.56	61.98	60.50
48	63.60	63.27	62.89	63.06	63.16
49	63.74	63.27	63.04	62.82	62.63
50	63.54	63.04	62.18	62.30	61.43
51	63.15	63.32	62.77	62.66	62.35
52	63.10	63.15	63.13	62.17	62.30
53	63.25	63.07	63.36	62.67	62.95
54	63.47	62.62	63.13	63.06	62.28
55	63.85	63.61	63.69	62.36	62.85
56	63.57	63.03	63.40	62.84	62.66
57	63.19	63.34	63.07	62.72	62.58
58	63.39	63.10	62.90	62.85	62.59
59	63.60	63.81	63.62	63.70	63.14
60	63.11	63.38	62.67	62.17	61.40

## Datos: Ejemplo 4.2

	sun	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	temp.s	wd	eye	w1	w2	w3	temp	hum
1	125	215	218	217	231	240	17	225	0.00	0	0	0	19	90
2	126	173	141	240	164	157	18	225	0.35	0	0	0	19	91
3	128	135	137	123	106	82	18	225	0.07	0	0	0	19	96
4	130	60	151	177	197	176	18	225	-0.06	0	0	0	18	96
5	132	119	52	288	48	94	18	225	0.18	0	0	0	18	97
6	134	159	149	170	170	171	18	204	0.00	0	0	0	18	97
7	136	175	63	154	197	139	18	204	0.00	0	0	0	19	97
8	139	145	150	243	241	244	19	204	-0.14	0	0	0	19	96
9	141	146	116	190	204	216	19	188	-0.00	0	0	0	20	90
10	143	140	125	141	158	188	20	188	-0.07	0	0	0	20	90
11	144	110	95	59	212	274	20	188	-0.07	0	0	0	20	90
12	147	317	193	174	197	189	20	188	-0.05	0	0	0	20	90
13	240	297	281	263	238	217	23	154	0.10	0	1	0	22	78
14	241	142	177	202	208	120	22	154	-0.22	0	1	0	22	77
15	247	259	266	269	272	245	21	154	-0.02	0	1	0	22	78
16	249	260	308	276	318	302	21	154	0.07	0	1	0	21	78
17	250	260	247	239	225	234	21	154	0.00	0	1	0	21	78
18	251	318	285	295	256	273	21	154	-0.07	0	1	0	21	78
19	253	326	20	304	262	219	21	154	0.06	0	1	0	21	78
20	254	21	18	319	326	225	21	154	-0.06	0	1	0	20	79
21	257	124	243	239	216	227	20	154	0.00	0	1	0	20	80
22	258	193	188	195	197	173	20	154	0.06	0	1	0	20	80
23	259	212	216	219	225	230	20	154	0.00	0	1	0	20	80
24	124	171	129	103	101	100	23	210	0.14	0	0	0	21	75
25	127	185	202	169	155	189	23	210	0.15	0	0	0	21	73
26	127	108	78	82	104	115	22	210	0.29	0	0	0	21	77
27	128	155	202	157	140	165	22	210	-0.07	0	0	0	21	75
28	129	13	81	74	168	192	22	210	-0.00	0	0	0	21	75
29	131	130	164	210	185	184	21	210	0.00	0	0	0	21	76
30	132	104	143	162	171	209	22	210	-0.06	0	0	0	21	76
31	134	104	128	118	103	126	22	210	-0.06	0	0	0	21	76
32	135	176	188	197	199	208	22	210	-0.07	0	0	0	21	76
33	136	135	115	121	125	147	22	210	0.00	0	0	0	21	76

continuación . . .

	sun	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	temp.s	wd	eye	w1	w2	w3	temp	hum
34	138	199	145	170	183	184	22	210	0.05	0	0	0	21	76
35	139	116	149	154	156	139	22	210	0.00	0	0	0	21	76
36	249	184	196	225	210	205	23	237	0.24	0	0	0	23	69
37	250	244	203	205	268	270	23	237	0.00	0	0	0	23	69
38	253	184	200	198	212	217	22	237	-0.02	0	0	0	22	71
39	253	224	189	216	206	218	22	237	0.00	0	0	0	22	73
40	255	191	229	223	212	211	22	237	-0.09	0	0	0	22	73
41	256	213	199	222	232	220	22	237	0.06	0	0	0	22	72
42	257	201	176	205	225	191	22	237	-0.08	0	0	0	22	72
43	258	242	244	240	262	231	22	237	0.00	0	0	0	21	72
44	260	212	224	238	138	195	21	237	-0.13	0	0	0	21	72
45	264	254	255	274	227	281	20	237	-0.08	0	0	0	20	74
46	264	169	230	241	231	255	20	237	0.16	0	0	0	20	75
47	265	243	236	223	219	217	20	232	-0.07	0	0	0	20	75
48	124	129	150	170	143	144	24	61	0.05	1	0	0	25	57
49	131	144	128	141	161	151	24	61	0.06	1	0	0	25	53
50	136	124	146	151	171	187	24	90	0.06	1	0	0	25	51
51	137	171	176	178	179	181	25	90	-0.06	1	0	0	25	52
52	144	164	150	152	139	141	23	240	-0.07	0	0	0	22	53
53	145	130	134	135	119	134	22	240	-0.06	0	0	0	22	53
54	148	145	151	159	154	161	22	240	0.06	0	0	0	23	53
55	149	143	141	144	152	152	22	240	0.13	0	0	0	22	24
56	259	213	216	194	242	261	24	264	0.05	0	0	1	24	61
57	260	134	102	68	107	111	24	264	0.36	0	0	1	25	61
58	261	165	161	150	137	177	24	264	0.00	0	0	1	25	61
59	263	121	184	215	225	242	23	264	0.00	0	0	1	24	62
60	264	202	199	193	215	216	23	264	-0.18	0	0	1	24	62
61	266	84	103	148	140	168	23	264	0.09	0	0	1	24	63
62	266	61	66	85	101	122	23	264	0.13	0	0	1	24	63
63	267	188	245	267	284	293	22	264	0.00	0	0	1	24	63
64	268	359	344	263	174	208	22	264	-0.08	0	0	1	24	63
65	269	173	187	210	244	247	22	264	0.10	0	0	1	24	61



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

# ACTA DE DISERTACIÓN PÚBLICA

No. 00013

Matricula: 206381008

ANALISIS BAYESIANO DE  
MODELOS LINEALES PARA DATOS  
DIRECCIONALES CONSIDERANDO  
LA DISTRIBUCION NORMAL BAJO  
PROYECCION

En México, D.F., se presentaron a las 8:00 horas del día 1 del mes de junio del año 2010 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. ALBERTO CASTILLO MORALES  
DR. RUSSELL JAMES BOWATER  
DR. EDUARDO GUTIERREZ PEÑA  
DR. MANUEL MENDOZA RAMIREZ  
DRA. BLANCA ROSA PEREZ SALVADOR

Bajo la Presidencia del primero y con carácter de Secretaria la última, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

DOCTOR EN CIENCIAS (MATEMÁTICAS)

DE: GABRIEL NUÑEZ ANTONIO

y de acuerdo con el artículo 78 fracción IV del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

*Aprobar*

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.



*Gabriel Nuñez Antonio*

GABRIEL NUÑEZ ANTONIO  
ALUMNO

REVISO

*[Signature]*

LIC. JULIO CÉSAR DE LARA ISASSI  
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI

*[Signature]*  
DR. JOSÉ ANTONIO DE LOS REYES  
BEREDIA

PRESIDENTE

*[Signature]*  
DR. ALBERTO CASTILLO MORALES

VOCAL

*[Signature]*  
DR. RUSSELL JAMES BOWATER

VOCAL

*[Signature]*  
DR. EDUARDO GUTIERREZ PEÑA

VOCAL

*[Signature]*  
DR. MANUEL MENDOZA RAMIREZ

SECRETARIA

*[Signature]*  
DRA. BLANCA ROSA PEREZ SALVADOR