

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE GRADUADOS



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

**“MODELOS DE REGRESION BINARIA SKEW PROBIT
PARA EL CALCULO DE PROBABILIDAD DE DEFAULT
EN EL AMBITO DEL SISTEMA FINANCIERO”**

**TESIS PARA OPTAR EL GRADO DE MAGISTER EN
ESTADÍSTICA**

AUTOR

Luis Pantoja Marin

ASESOR

Dr. Jorge Luis Bazán Guzmán

JURADO:

Dra. Mery Elizabeth Doig Camino
Dr. Oscar Edgardo Millones Destefano

LIMA - PERU

2012

Dedicatoria

El presente trabajo de investigación está dedicado a aquellas personas que se especializan, como parte de su profesión, trabajo, al área de la investigación, en especial a aquellas personas y profesionales del área de la Estadística.



Agradecimientos

Agradecer primeramente a mi asesor por sus orientaciones y sus deseos de enseñanza a los alumnos de la escuela, en segundo orden agradecer a mi familia por su apoyo constante y espíritu de sacrificio hacia el desarrollo de proyectos por mas pequeños que estos sean.



Resumen

La presente investigación se fundamenta en el uso o aplicación de Modelos Skew Probit con enlace asimétrico desde un enfoque Bayesiano. Los modelos a usar incorporan la posibilidad de usar enlaces asimétricos para estimar la probabilidad de $y_i = 1$ en muestras no balanceadas (alta proporción de ceros y por ende pequeña proporción de unos). La formulación general de estos modelos es debida a Bazán, Bolfarine y Branco (2010).

Aunque en estos modelos inicialmente su computación es complicada se usaron Cadenas de Markov por Monte Carlo (MCMC) o muestreo Gibbs (para la aplicación de estos procedimientos ver Carlin y Polson, 1992) que hacen simple la formulación del modelo y por tanto simple su implementación usando el software WinBugs (los códigos de los diferentes modelos utilizados fueron obtenidos en el programa BRMUW propuesto por Bazán y Bayes, 2010).

De acuerdo al análisis y estudio de aplicación realizado sobre una muestra de clientes de préstamos pertenecientes a una entidad micro financiera, aquellos modelos Skew Probit BBB y Estándar presentan los mejores indicadores de eficiencia.

El análisis sobre datos reales señala que el modelo tradicional Probit presenta un 56.6% (371/664) de mala clasificación versus los modelos Estándar y BBB que en promedio muestran dicho indicador alrededor de 43% (290/664).

El análisis mediante curvas COR (Receiver Operating Characteristic) ratifica lo mencionado; el área debajo de las curvas superan el 0.74 de 1 para el modelo BBB, mientras que dicho dato es de 0.70 para el caso del modelo simétrico tradicional probit. Por tanto la sensibilidad y especificidad (eficiencia) es mayor para aquellos modelos Skew Probit (mejor modelo BBB).

Dentro de los modelos con Enlaces Asimétricos los modelos (SP) BBB y Estándar son los que presentan mejores indicadores de ajuste e información así como mejoran la sensibilidad y especificidad de un determinado modelo. Finalmente, se pretende la sistematización de la propuesta a nivel de la entidad micro financiera y su aplicación en la estimación de la probabilidad de default de créditos pero aplicado en todos los tipos de créditos.

Palabras-Clave: Regresión Binaria, Inferencia Bayesiana, Probabilidad de Default, Enlace Asimétrico, Modelos Skew Probit Generalizados, Variables Latentes, Deviance Information Criteria, Media de los Desvíos a Posteriori.

Abstract

This research is based on the use or application of Skew Probit Models with asymmetric link from a Bayesian approach. Using models incorporate the ability to use asymmetric links for estimating the probability of $y_i = 1$ unbalanced samples (high proportion of zeros and hence a small proportion). The general formulation of this model is due to Bazán, Bolfarine and Branco (2010).

Although these models in your computer is complicated initially used Markov Chain Monte Carlo (MCMC) and Gibbs sampling (for the application of these procedures see Carlin and Polson, 1992) that make easy the formulation of a simple model and therefore its implementation WinBugs using the software (the codes of the different models used were obtained in the program proposed by Bazán BRMUW and Bayes, 2010).

According to the analysis and implementation study on a sample of loan clients belonging to a micro financial institution, those Probit models BBB and Standard Skew presented the best indicators of efficiency.

The actual data analysis indicates that the traditional Probit model presents a 56.6% (371/664) versus misclassification BBB standard models and show that on average the indicator around 43% (290/664).

Analysis by ROC curves (Receiver Operating Characteristic) confirms the above, the area under the curve of 0.74 exceeds 1 for the BBB model, while the data is 0.70 for the case of the traditional symmetrical Probit model. For both sensitivity and specificity (efficiency) is greater for those Skew Probit models (best model BBB).

Among the models with asymmetric links models (SP) BBB and Standard are the top performers. Finally, it is intended to systematize the proposal at the level of micro financial institution and its application in estimating the probability of default of loans but applied in all types of credit.

Keywords: Binary Regression, Bayesian Inference, Probability of Default, Asymmetric Link, Generalized Probit Skew Models, Latent Variable, Deviance Information Criteria, Mean Deviation Afterwards.

Índice general

Lista de Abreviaturas	VIII
Lista de Símbolos	IX
Índice de figuras	X
Índice de cuadros	XI
1. Introducción	1
1.1. Consideraciones Preliminares	1
1.2. Objetivos	1
1.3. Organización del Trabajo	2
2. Distribución Normal y Skew Normal	3
2.1. Distribución Normal	3
2.1.1. Función de densidad de probabilidad	3
2.1.2. Propiedades	3
2.1.3. Función de Distribución Acumulada	3
2.1.4. Curtosis	5
2.1.5. Asimetría	5
2.2. Distribución Skew Normal	6
2.2.1. Función de Densidad de Probabilidad de la Skew Normal Estándar	6
2.2.2. Función de Distribución Acumulada de la Skew Normal Estándar	6
2.2.3. Estimación	7
2.2.4. Propiedades y Características de la Distribución Skew Normal de Tres Parámetros	8
3. Regresión Binaria	11
3.1. Regresión Binaria	11
3.2. Regresión Probit	12
3.2.1. Regresión - Probit Versión Aumentada	12
3.3. Revisión General de Clases de Enlaces Asimétricos	13
3.4. Modelos Skew Probit Generalizados	15
3.4.1. Modelo Skew Probit CDS.	15
3.4.2. Modelo Skew Probit BBB.	16
3.4.3. Modelo Skew Probit ESTANDAR.	17

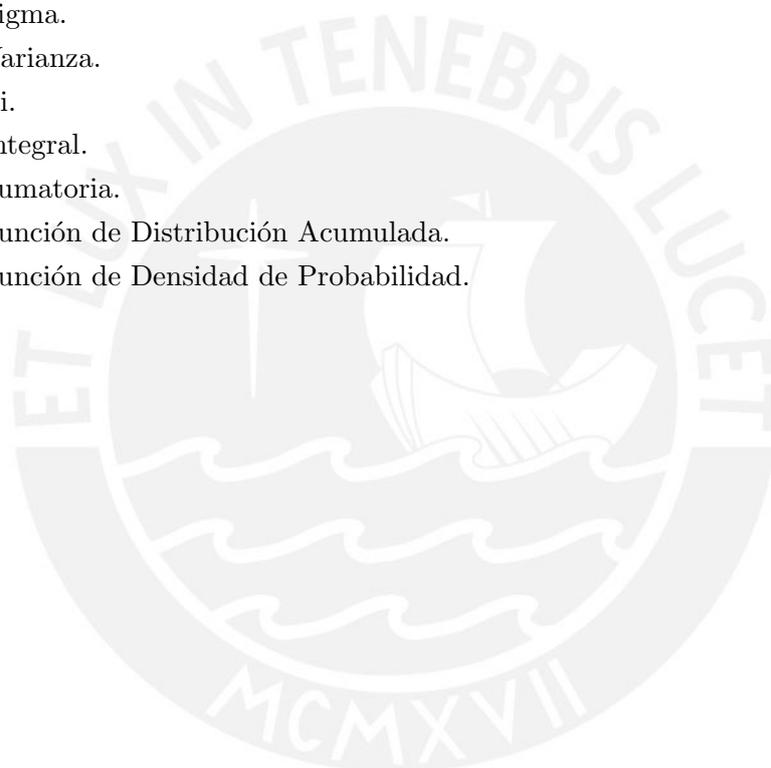
3.4.4. Regresión Skew Probit - Versión Aumentada	18
4. Inferencia Bayesiana para el Modelo Skew Probit	22
4.1. Verosimilitud del Modelo	22
4.2. Procedimiento Inferencial para el enlace Skew Probit	23
4.2.1. Verosimilitud Aumentada del Modelo Skew Probit	25
4.3. Especificación de Prioris	26
4.3.1. Elección de Prioris Alternativa	27
4.4. Selección de Modelos	28
4.5. Estructura Jerárquica de la Regresión Skew Probit	29
4.6. Implementacion Computacional	30
5. Aplicación: Cálculo de la Probabilidad de Default en el Ambito del Sistema Financiero	31
5.1. Modelo de Regresión Binaria	31
5.1.1. Descripción de los Datos	31
5.1.2. Formulación del Modelo General de Regresión Binaria	32
5.1.3. Resultados	33
5.1.4. Resumen de Análisis Gráfico Proporcionados por WinBUGS	34
5.2. Análisis Complementarios	37
5.2.1. Análisis según Curvas COR (Gráfico)	37
5.2.2. Análisis Según Tablas de Contingencia (numérico)	39
5.3. Estudio de Verificación	41
5.3.1. Diseño	41
5.3.2. Resultados	41
6. Conclusiones y Recomendaciones	44
6.1. Conclusiones	44
6.2. Recomendaciones	46
A. Gráficos Correspondientes a Numeral 5.1.4	47
A.1. Modelo Probit Simétrico - Muestra N=4745	47
A.2. Modelo Skew Probit Estándar Asimétrico - Muestra N=4745	49
A.3. Modelo Skew Probit CDS Asimétrico - Muestra N=4745	51
A.4. Modelo Skew Probit BBB Asimétrico - Muestra N=4745	53
B. Resumen de Programas Utilizados	55
Bibliografía	60

Lista de Abreviaturas

FDP	Función de densidad de probabilidad .
PBF	Pseudo factor de Bayes(<i>Pseudo bayes factor</i>).
BBB	Modelo Skew Probit BBB.
CDS	Modelo Skew Probit CDS.
Estandar	Modelo Skew Probit Estandar.
DIC	Deviance Information Criterio.
Dbar	Media de los Desvios a Posteriori.
BRMUW	Bayesian Regression Model Using WinBUGS.
CDF	Función de Distribución Acumulada.
SPG	Skew Probit Generalizado.
COR	Receiver Operating Characteristic, o Característica Operativa del Receptor.

Lista de Símbolos

μ	Media.
β	Beta.
λ	Lambda.
δ	Delta.
σ	Sigma.
σ^2	Varianza.
Φ	Fi.
\int	Integral.
Σ	Sumatoria.
FDA	Función de Distribución Acumulada.
FDP	Función de Densidad de Probabilidad.



Índice de figuras

2.1. Densidades de la distribución normal para diferentes valores de la media y la varianza	4
2.2. FDA's para distintos valores de la media y la varianza	4
2.3. Densidades de la distribución skew normal para diferentes valores de α	7
2.4. Probabilidades Acumuladas para diferentes valores de α	7
5.1. Curvas COR según Modelos Asimétricos o Modelo Simétrico	38
A.1. Gráficos de Historia de la Cadena y Densidades Modelo Probit Simétrico	47
A.2. Gráficos de Quantiles y Autocorrelacion Modelo Probit Simétrico	48
A.3. Gráficos de Historia de la Cadena y Densidades Modelo Estándar	49
A.4. Gráficos de Quantiles y Autocorrelacion Modelo Estándar	50
A.5. Gráficos de Historia de la Cadena y Densidades Modelo CDS	51
A.6. Gráficos de Quantiles y Autocorrelacion Modelo CDS	52
A.7. Gráficos de Historia de la Cadena y Densidades Modelo BBB	53
A.8. Gráficos de Quantiles y Autocorrelacion Modelo BBB	54

Índice de cuadros

5.1. Comparación de Modelos para Modelo Aplicativo de Otorgamiento de Préstamos (N = 4745)	33
5.2. Resumen de Aplicaciones Según Modelo Propuesto y Modelos Alternos	33
5.3. Estadísticas A Posteriori Modelo Skew Probit - BBB	36
5.4. Estadísticas A Posteriori Modelos Skew Probit - Estándar	36
5.5. Estadísticas A Posteriori Modelo Skew Probit - CDS	37
5.6. Estadísticas A Posteriori Modelo Probit	37
5.7. Tabla de Contingencia Modelo Asimétrico Skew Probit BBB	40
5.8. Tabla de Contingencia Modelo Asimétrico Skew Probit CDS	40
5.9. Tabla de Contingencia Modelo Asimétrico Skew Probit Estándar	40
5.10. Tabla de Contingencia Modelo Simétrico Probit	41
5.11. Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 4745)	42
5.12. Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 3157)	42
5.13. Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 1575)	42
5.14. Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 931)	43

Capítulo 1

Introducción

1.1. Consideraciones Preliminares

Los modelos binarios se aplican para evaluar el comportamiento probable de los consumidores, los cuales se enfrentan a una situación de elección binaria (Mc Fadden, 1974, 1981).

Los modelos de variable dependiente binaria (cero o uno) son especificados como una expresión lineal (modelo de probabilidad lineal) o como una expresión no lineal (modelos de probabilidad no lineal), estos últimos usan generalmente una función de enlace simétrica, siendo el enlace definido como una transformación del valor esperado de la variable respuesta.

De esta manera, un enlace simétrico es una función $F(\cdot)$ que se denota como una Función de Distribución Acumulada, siendo $F(\cdot)$ además una distribución simétrica; sin embargo en muchas aplicaciones las muestras no se encuentran balanceadas, esto es, una de las categorías de la variable dependiente es la menos preferida o la menos frecuente, aspecto que se presenta cuando se produce una condición de probabilidad de default (cuentas de crédito que presentan un comportamiento no balanceado), pocos clientes con alta probabilidad de default versus muchos clientes con baja probabilidad de default; evidencia empírica en el sector de las finanzas, microfinanzas y seguros señalan o confirman dicho comportamiento (estadísticas de la Superintendencia de Banca y Seguros). En esta situación el uso de un modelo con enlace simétrico no presenta buenos indicadores de ajuste e información (Bermudez et al, 2007), por el contrario el uso de un modelo con enlace sesgado o enlace asimétrico (Chen et al, 1999) ayudan a mejorar el ajuste y por ende la sensibilidad y especificidad de un modelo probit, lo cual proveerá buenos resultados para la proporción de sucesos en la matriz de confusión.

1.2. Objetivos

El objetivo general de la tesis es analizar, proponer y aplicar Modelos de Regresión Binaria Skew Probit para el Cálculo de Probabilidad de Default dentro de una institución Microfinanciera. De manera específica se realizaron las siguientes actividades:

- Revisar bibliografía literaria y otros sobre modelos Probit y Modelos Skew Probit.
- Evaluar y presentar propiedades de Skew Probit con Enlaces Asimétricos.
- Presentar e implementar la estimación Bayesiana correspondiente a los modelos propuestos.

- Analizar, proponer y aplicar el mejor modelo estudiado para el cálculo de probabilidades de default en una cartera de créditos y compararlo con el modelo sin asimetría o modelo tradicional.
- Aplicar datos reales correspondientes a una Micro Financiera (otorgamiento de préstamos).

1.3. Organización del Trabajo

En el Capítulo 2, llevamos a cabo una revisión sobre las distribuciones Normal y Skew Normal, se detallan los conceptos básicos de Función de Densidad y Función Acumulada, Valor Esperado, Varianza, Asimetría y Curtosis de dichas distribuciones; incluimos algunos gráficos.

En el Capítulo 3 revisamos conceptos sobre la Regresión Binaria, la Regresión Probit y su versión Aumentada, la Regresión Skew Probit Generalizada y sus casos particulares (versión Aumentada). En el capítulo 4 detallamos pormenores sobre la Inferencia Bayesiana, Verosimilitud y Funciones Prioris para los modelos propuestos, incluyendo la selección de modelos y su implementación computacional.

En el capítulo 5 se realiza una aplicación para el cálculo de la probabilidad de default en el ámbito del sistema financiero aplicando el modelo de regresión skew probit y considerando diversos análisis complementarios como la evaluación del mejor modelo en función a criterios como el DIC, Dbar, Curvas COR, Tablas de Contingencia, etc.

En el capítulo 6 se resumen las principales Conclusiones y Recomendaciones al estudio aplicado.

En los apéndices y anexo presentamos algunas pruebas de resultados en más detalle y también los programas utilizadas en las aplicaciones a conjuntos de datos reales.

Capítulo 2

Distribución Normal y Skew Normal

En el presente capítulo se realizará una breve descripción sobre aquellas funciones de distribución que de una u otra forma involucran conceptos, relaciones y criterios sobre y hacia la Regresión Skew Probit.

2.1. Distribución Normal

La distribución Normal o Gaussiana, en la Teoría de la Probabilidad, es una distribución de probabilidad continua que se utiliza a menudo como una primera aproximación para describir valores reales de variables aleatorias que tienden a agruparse en torno a un valor medio único (Media). La gráfica de la función de densidad de probabilidad tiene forma “acampanada”, y se conoce como la función de Gauss o curva Acampanada, contiene una media y varianza de $(0, 1)$ (Mendenhall y Wackerly, 2009).

2.1.1. Función de densidad de probabilidad

La función de densidad de probabilidad de una variable aleatoria X que sigue una distribución normal es dada por la siguiente expresión:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2.1)$$

Esta Ley de Probabilidad es bastante flexible para modelar este tipo de datos debido a que su función de densidad de probabilidad puede tomar diversas formas dependiendo del valor de los parámetros que caracterizan a esta distribución (ver figura 2.1).

2.1.2. Propiedades

La media y la varianza de una distribución normal son expresadas por

$$E(x) = \mu \quad Var(x) = \sigma^2 \quad (2.2)$$

2.1.3. Función de Distribución Acumulada

La función de distribución acumulada (FDA) de una distribución normal se halla expresada por

$$F(x|\mu, \sigma^2) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \quad (2.3)$$

La función de distribución acumulativa (FDA ó CDF, por siglas en inglés) describe las probabilidades de una variable aleatoria de caer en los intervalos de la forma $(-\infty, x]$. La FDA

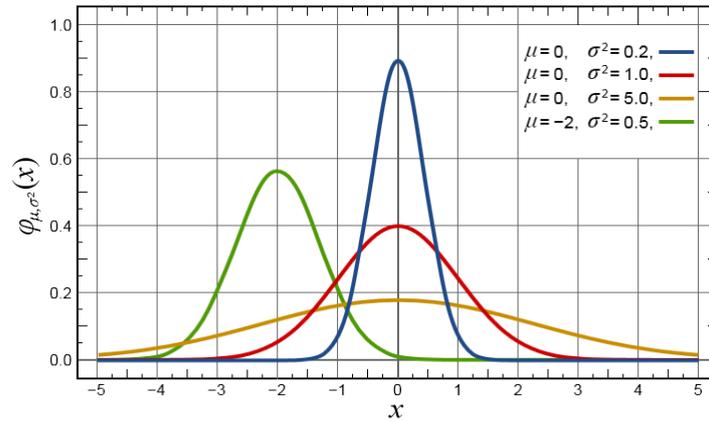


Figura 2.1: Densidades de la distribución normal para diferentes valores de la media y la varianza

de la distribución normal estándar se denota con la letra mayúscula griega Φ , y puede ser calculada como una parte de la integral de la función de densidad de probabilidad, (ver figura 2.2).

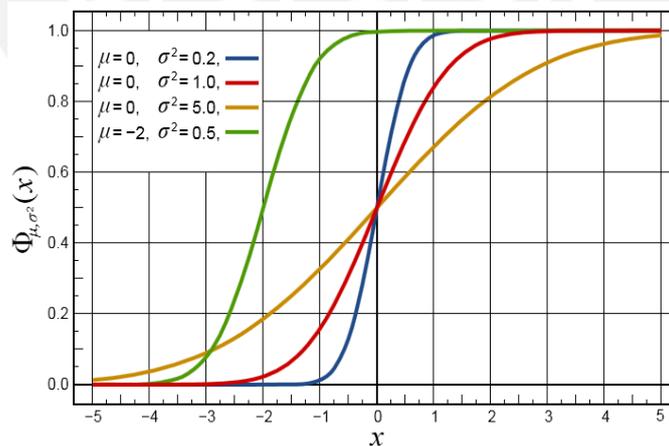


Figura 2.2: FDA's para distintos valores de la media y la varianza

2.1.4. Curtosis

La Curtosis es una medida de apuntamiento (Curtosis) de las distribuciones. Así las medidas de curtosis (también llamadas de apuntamiento o de concentración central) tratan de estudiar la mayor o menor concentración de frecuencias alrededor de la media y en la zona central de la distribución. El coeficiente de apuntamiento de uso más extendido es el basado en el cuarto momento con respecto a la media y se define como:

$$K = \frac{\mu_4^*}{\sigma^4} - 3 \quad (2.4)$$

Donde: μ_4^* es el cuarto momento centrado o con respecto a la media y σ es la desviación estándar.

Considerando entonces la distribución normal como referencia, una distribución puede ser:

- Mesocurtica, cuando la curva es similar a la curva normal; $K = 0$.
- Leptocurtica, cuando la curva es mas apuntada que la curva normal; $K > 0$.
- Platicurtica, cuando la curva es menos apuntada que la curva normal; $K < 0$.

Para el caso de la distribución normal se tiene que $k = 0$ y por tanto es considerada mesocurtica.

2.1.5. Asimetría

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin tener que hacer su representación gráfica. Como eje de simetría consideramos una recta paralela al eje de ordenadas que pasa por la media de la distribución. Si una distribución es simétrica, existe la misma área (número de valores) a la derecha como a la izquierda de la media, por tanto, el mismo número de desviaciones con signo positivo que con signo negativo. Decimos que hay asimetría positiva (o a la derecha) si la “cola” a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media a la derecha. Diremos que hay asimetría negativa (o a la izquierda) si la “cola” a la izquierda de la media es más larga que la de la derecha, es decir, si hay valores más separados de la media a la izquierda.

La medida de asimetría más utilizada parte del uso del tercer momento estándar. La razón de esto es que nos interesa mantener el signo de las desviaciones con respecto a la media, para obtener si son mayores las que ocurren a la derecha de la media que las de la izquierda. Sin embargo, no es buena idea tomar el momento estándar con respecto a la media de orden 1 (una simple suma de todas las desviaciones siempre debe ser cero); por ello, lo más sencillo es tomar las desviaciones al cubo.

El coeficiente de asimetría de Fisher, representado por AS , se define como:

$$AS = \frac{\mu_3^*}{\sigma^3} \quad (2.5)$$

donde μ_3^* es el tercer momento en torno a la media y σ es la desviación estándar. Por tanto:

- Si $AS = 0$, la distribución es simétrica

- Si $AS > 0$, la distribución es asimétrica positiva o a la derecha
- Si $AS < 0$, la distribución es asimétrica negativa o a la izquierda

Para el caso de la distribución normal se tiene que $AS = 0$ y por tanto es considerada simétrica.

2.2. Distribución Skew Normal

La distribución skew normal (asimétrica, sesgada) es una distribución de probabilidad continua; se generaliza la distribución normal para permitir una asimetría de cero; dados $\mu = 0$ y $\sigma^2 = 1$.

2.2.1. Función de Densidad de Probabilidad de la Skew Normal Estándar

Consideremos en primer lugar una v.a continua X con función de densidad de probabilidad de la siguiente forma (Azzalini, 2005):

$$f(x) = 2\phi(x)\Phi(\alpha x) \quad (2.6)$$

Donde α es un número fijo arbitrario y:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (2.7)$$

$$\Phi(\alpha x) = \int_{-\infty}^{\alpha x} \phi(t) dt \quad (2.8)$$

Tanto $\phi(x)$ y $\Phi(\alpha x)$, denotan la función de densidad normal estándar (Gauss) y su función de distribución acumulada (el último evaluado en el punto αx), respectivamente. El componente α se denomina parámetro de forma, ya que regula la forma de la función de densidad, como se ilustraría para casos donde $\alpha = 2$, $\alpha = 5$, $\alpha = -5$.

La densidad $f(x)$ muestra varias propiedades formales interesantes, es relativamente fácil comprobar que (ver figura 2.3):

- Cuando $\alpha = 0$, la asimetría se reduce y se obtiene la densidad normal estándar.
- A medida que α aumenta en valor absoluto, la asimetría de la distribución se incrementa.
- Cuando $\alpha \rightarrow \infty$, la densidad converge a la llamada función de densidad Half Normal (o normal plegado).
- Si el signo de α cambia, la densidad varía en sentido opuesto.

2.2.2. Función de Distribución Acumulada de la Skew Normal Estándar

La función de distribución acumulada está dada por: (ver figura 2.4)

$$\Phi(\alpha x) = \int_{-\infty}^{\alpha x} \phi(t) dt \quad (2.9)$$

En la sección 2.2.4 se presenta expresiones exactas para la función acumulada.

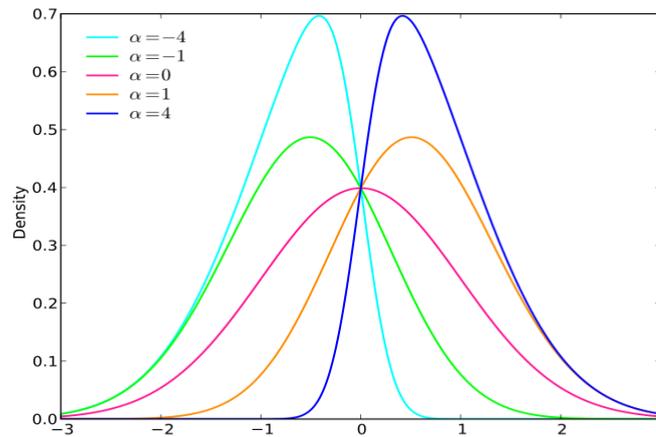


Figura 2.3: Densidades de la distribución skew normal para diferentes valores de α

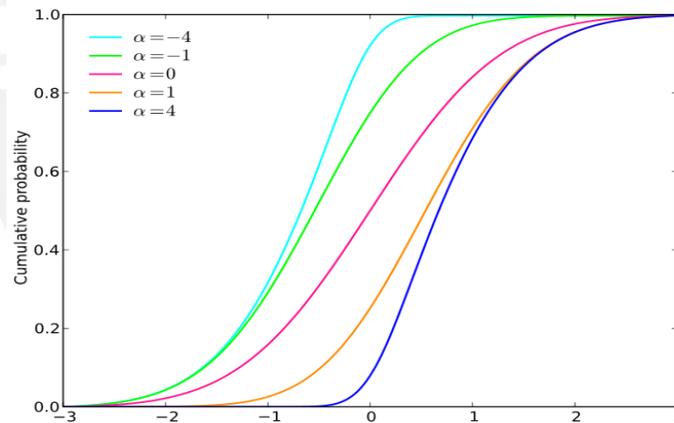


Figura 2.4: Probabilidades Acumuladas para diferentes valores de α

2.2.3. Estimación

Las Estimaciones para x y α se pueden calcular mediante Métodos de Estimación tales como el Método de Máxima Verosimilitud o via Métodos de Momentos (ver Azzalini, 1985).

2.2.4. Propiedades y Características de la Distribución Skew Normal de Tres Parámetros

En este caso consideramos lo dado por Azzalini A. 2010, si una variable aleatoria R sigue una distribución Skew Normal, con parámetro $\theta = (\mu, \sigma^2, \alpha)$, donde μ es un parámetro ($\mu \in \mathbb{R}$) de localización, $\sigma^2 > 0$ es un parámetro de escala y α , es un parámetro de asimetría ($\alpha \in \mathbb{R}$); si la función de densidad de probabilidad (pdf) de R es dado por:

$$f_{\theta} = \frac{2}{\sigma} \phi\left(\frac{r - \mu}{\sigma}\right) \Phi\left(\alpha \frac{r - \mu}{\sigma}\right) \quad (2.10)$$

Donde ϕ y Φ representan las funciones de densidad y distribución de la distribución normal estándar, entonces la notación usada es $R \sim SN(\theta)$.

Una parametrización alternativa considera reemplazar α por $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$, de los resultados $\delta \in [-1, 1]$. De acuerdo a Azzalini (1985), (2.10) posee algunas características y propiedades importantes:

1. La función generadora de momentos de R es:

$$M_R(t) = 2 \exp(\mu t + \frac{1}{2} \sigma^2 t^2) \Phi(\delta \sigma t)$$

2. $E[R] = \mu + \sqrt{\frac{2}{\pi}} \delta \sigma$

3. $Var[R] = (1 - \frac{2}{\pi} \delta^2) \sigma^2$

4. Si, $W \sim N(0, 1)$, y $V \sim HN(0, 1)$ son variables aleatorias independientes y $|\delta| < 1$, luego $R = \mu + \sigma[\delta V + (1 - \delta^2)^{\frac{1}{2}} W]$ posee una $SN(\mu, \sigma^2, \alpha)$, y $R|V$ posee una $N(\mu + \sigma \delta v, (1 - \delta^2) \sigma^2)$.

5. El índice de asimetría $\gamma = (\frac{2}{\pi})^{\frac{3}{2}} (2 - \frac{\pi}{2}) \text{sig}(\delta) \frac{\delta^3}{(1 - \frac{2}{\pi} \delta^2)^{\frac{3}{2}}}$, $\gamma \in [-0.995, 0.995]$.

6. También cuando $\alpha = 0$, la densidad de R en (2.10) se reduce a la densidad de $N(\mu, \sigma^2)$.

El caso especial con $\mu = 0$ y $\sigma^2 = 1$, es llamado por Azzalini como la distribución Skew - Normal Estándar, denotado por $S \sim SN_A(\alpha)$, siendo su función de densidad de probabilidad:

$$f_{\alpha}^A(s) = 2\phi(s)\Phi(\alpha s) \quad (2.11)$$

De otro lado cuando $\mu = 0$ y $\sigma^2 = 1 + \alpha^2$ en (2.10), sigue la distribución Skew-Normal Estándar dado por Chen et al.(1999) y Sahu et al.(2003). Se considera la siguiente notación $S \sim SN_s(\alpha)$, y su correspondiente función de densidad de probabilidad es:

$$f_{\alpha}^S(s) = \frac{2}{\sqrt{1 + \alpha^2}} \phi\left(\frac{s}{\sqrt{1 + \alpha^2}}\right) \Phi\left(\alpha \frac{s}{\sqrt{1 + \alpha^2}}\right) \quad (2.12)$$

Conexiones interesantes entre la distribución Skew - Normal Estándar de Azzalini y de Sahu

se pueden obtener realizando la transformación de $S^* = \sqrt{1 + \lambda^2}S$, donde $S \sim SN_A(\sigma)$. Entonces, $S^* \sim SN_s(\sigma)$. Asimismo por la distribución Skew - Normal de Azzalini $E[S] \in [-0.80, 0.80]$ y $\text{Var}[S] \in [0.36, 1]$. Por otro lado, por la distribución Skew - Normal Estándar de Sahu, $E[S] \in [-0.80|\alpha|, 0.80|\alpha|]$ y $\text{Var}[S] \in [1, 1 + 0.36\alpha^2]$. Por lo tanto, el valor esperado y la varianza toman valores más pequeños para la SN_A .

A continuación se muestra una representación especial para la función de distribución acumulada (FDA) de la distribución Skew - Normal, que será útil para establecer las relaciones entre los dos tipos de enlaces Skew Probit mostrados a continuación:

Según la proposición 3.1. (Bazan et al. 2010, Skew-Probit Links in Binary Regression); sea $R \sim SN_s(\theta)$. Entonces, la función de distribución acumulada (FDA) de R se puede escribir de dos formas diferentes:

$$F_\theta(r) = \int_0^\infty g(v)\Phi\left(\frac{r - \mu - \sigma\delta v}{\sigma\sqrt{1 - \delta^2}}\right)dv = \int_0^\infty g(v)\Phi\left(\frac{r - \mu}{\sigma}\sqrt{1 + \alpha^2} - \alpha v\right)dv \quad (2.13)$$

Ademas:

$$F_\theta(r) = 2\Phi_2(x; \mu, \Sigma), \quad (2.14)$$

Donde:

- $x = (s, 0)$
- $s = \frac{r - \mu}{\sigma}$
- $\Phi_2(\cdot)$, es la función de distribución acumulada de la distribución normal bivariada, con parámetros $\mu = (0, 0)'$ y $\Sigma = \text{Matriz}$ y coeficiente de correlación $-\delta$.

Donde:

Σ es:

$$\Sigma = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$$

La representación de la FDA Skew-Normal en (2.13) es similar al resultado dado en Chen et al. (1999). La representación de la función de distribución acumulada Skew - Normal dado en (2.14) indica que la distribución Skew - Normal evaluado en un punto r también se obtiene considerando una distribución normal bivariada con un vector de medias μ y una matriz de varianza - covarianza Σ evaluado en el punto $(r, 0)$. Este resultado es importante dado que algunos algoritmos computacionales eficientes están disponibles en computación relacionados con la distribución normal bivariada (integración). Otro algoritmo para evaluar la FDA (Función de Distribución Acumulada) de la distribución de Skew - Normal se basa en el uso de la función de Owen (Azzalini, 1985) y está disponible para los programas informáticos de R y Matlab. Además, la distribución de R se obtiene considerando $(U, W) N_2(\mu, \Sigma)$ μ

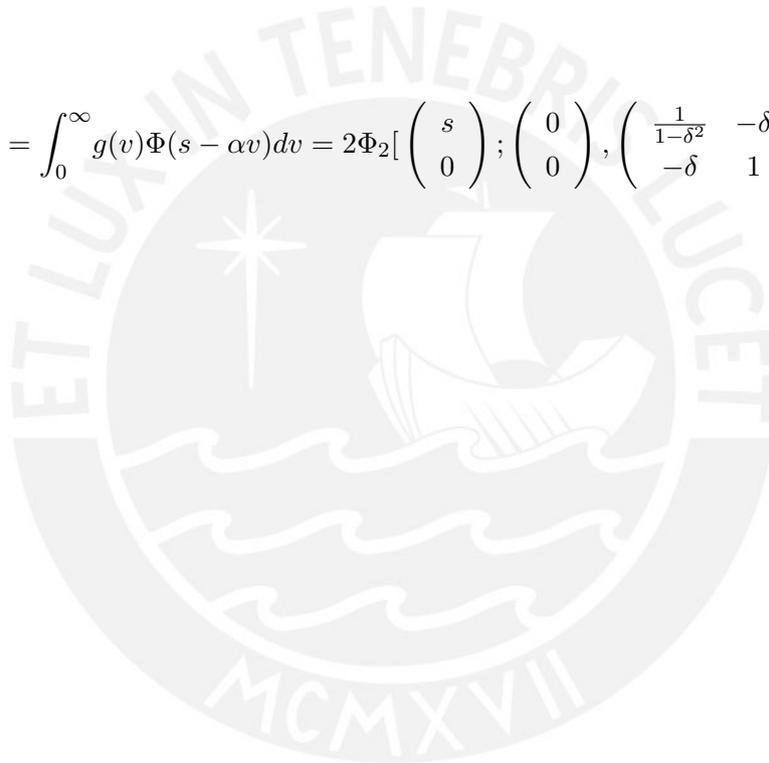
y Σ definidos anteriormente, siendo $R = U|W > 0$. Dos casos especiales de la proposición anterior se han establecido.

En correspondencia con la distribución Skew - Normal Estándar dado por Azzalini, se deduce que:

$$F_{\alpha}^A(s) = \int_0^{\infty} g(v)\Phi(s\sqrt{1+\alpha^2} - \alpha v)dv = 2\Phi_2\left[\begin{pmatrix} s \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (2.15)$$

Asimismo, en correspondencia con la distribución Skew - Normal Estándar dado por Sahu, se deduce que:

$$F_{\alpha}^S(s) = \int_0^{\infty} g(v)\Phi(s - \alpha v)dv = 2\Phi_2\left[\begin{pmatrix} s \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{1-\delta^2} & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (2.16)$$



Capítulo 3

Regresión Binaria

En estadística, la regresión binaria es una técnica en la que la respuesta (a menudo denominado como Y) es el resultado de una serie de ensayos de Bernoulli, o una serie de uno de dos resultados posibles disjuntos, tradicionalmente se denota el “éxito” con 1, y “fracaso” con 0, en la regresión binaria la probabilidad de éxito está relacionado con las variables explicativas; el concepto correspondiente en la regresión ordinaria es relacionar el valor de la media de la respuesta observada con las variables explicativas (Svetlozar T. y Rachev J, 2008).

Dentro de los modelos de regresión binaria se encuentran los siguientes:

1. Modelos con Variable Dependiente Categórica.
2. Modelos de Probabilidad Lineal.
3. Modelos Probit.
4. Modelos Logit.

3.1. Regresión Binaria

Definiendo una cartera de clientes de crédito como una colección de N individuos o cuentas de créditos donde una variable binaria y_i es observada para un cliente i , $i = 1, 2, \dots, N$, donde $y_i = 1$ si el cliente posee una probabilidad de default de 1 y $y_i = 0$ en otro caso. Es obvio que y_i sigue una distribución de Bernoulli, donde $Pr(y_i = 1) = p_i$ y $Pr(y_i = 0) = 1 - p_i$.

Así tenemos que: $E(y_i) = p_i$ y $Var(y_i) = p_i(1 - p_i)$

Sea $y = (y_1, y_2, \dots, y_n)'$ una muestra de $n \leq N$ observaciones, entonces la verosimilitud es dada por:

$$L(y) = \prod_{i=1}^n p_i^{y_i} \times (1 - p_i)^{1-y_i} \quad (3.1)$$

En un modelo de respuesta binaria o dicotómica, $p_i = F(x_i'\beta)$, donde $x = (x_{i1}, x_{i2}, \dots, x_{ik})'$ es el vector de variables explicativas o independientes; donde X_1 hasta X_8 son variables de comportamiento, financieras y socio demográficas; y $\beta = (\beta_0, \beta_1, \dots, \beta_k)'$ es el vector de coeficientes de regresión. La función de verosimilitud es escrita como:

$$L(y) = \prod_{i=1}^n [F(x_i'\beta)]^{y_i} \times [1 - F(x_i'\beta)]^{1-y_i} \quad (3.2)$$

Siendo:

$$x'_i\beta = \beta_0 + \beta_1X_1 + \beta_2X_2\dots\dots\dots\beta_kX_k \quad (3.3)$$

Para este caso se puede asumir diferentes casos de la función de enlace, $F^{-1}(\cdot)$. Por ejemplo puede ser la inversa de la acumulada de la función de distribución normal estándar (modelo Probit), o la inversa de la acumulada una función de distribución logística (modelo Logit). Una principal ventaja que se encuentra en el modelo logit sobre el modelo probit es que la interpretación de sus coeficientes es más fácil. Sin embargo el modelo probit presenta otras ventajas desde el punto de vista de estimación.

3.2. Regresión Probit

Como se ha indicado, cuando $F(\cdot) = \Phi(\cdot)$, es decir la acumulada de la distribución normal, presentada en el capítulo, estamos frente al modelo probit.

El Modelo Probit tiene la siguiente expresión:

$$E[Y] = P(Y = 1) = \int_{-\infty}^{x'_i\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \quad (3.4)$$

La estimación de los parámetros se realiza mediante Estimación de Máxima Verosimilitud, en este tipo de modelos no resulta posible interpretar directamente las estimaciones de los parámetros β , ya que son modelos no lineales. En la práctica nos fijamos en el signo de los estimadores. Si el estimador es positivo, significará que incrementos en la variable asociada causan incrementos en $P(Y = 1)$ (aunque desconocemos la magnitud de los mismos). Por el contrario, si el estimador muestra un signo negativo, ello supondrá que incrementos en la variable asociada causarían disminuciones en $P(Y = 1)$.

3.2.1. Regresión - Probit Versión Aumentada

Las variables latentes, o rasgos latentes, se usan para modelar las características no observadas de los sujetos que determinan las respectivas probabilidades de éxito. Para los modelos con más de dos categorías de respuesta ordinales, las variables latentes se extienden naturalmente para describir las probabilidades de las observaciones en cada categoría.

Suponiendo que la probabilidad de éxito para una variable binaria y_i se modela como:

$$p_i = F(x'_i\beta), \quad (3.5)$$

Donde x_i es el vector de variables explicativas, β es un parámetro de regresión binaria, y F es una distribución conocida. Para introducir una estructura variable latente dentro de este modelo, para cada observación i se introduce una variable continua Z_i , donde:

$$Z_i = x'_i\beta + \epsilon_i, \quad (3.6)$$

Donde $\epsilon_1, \dots, \epsilon_n$ son independientes e idénticamente distribuidas según F . Asumimos que:

$$y_i = \begin{cases} 0, & \text{si, } Z_i \leq 0, \\ 1, & \text{si, } Z_i > 0 \end{cases} \quad (3.7)$$

Así, $Z_i > 0$ si y sólo si la observación i es un éxito. De ello se deduce que:

$$p_i = Pr(y_i = 1) = PR(Z_i > 0) = 1 - F(-x_i'\beta) \quad (3.8)$$

Para distribuciones F que son simétricos alrededor de 0, esto es equivalente a $p_i = F(x_i'\beta)$. Por lo tanto, la estructura de la variable latente se obtiene un modelo equivalente a una logística o modelo probit, y con la redefinición de menor importancia y evidente de los parámetros de regresión, también puede ser utilizado para motivar enlaces asimétricos como el log-log complementario.

Albert y Chib (1993) discute la utilidad de los modelos de variables latentes para data binaria y datos policotómicos tanto desde el punto de vista computacional y como herramientas de diagnóstico para la evaluación del modelo.

Para demostrar la utilidad del enfoque variable latente, vamos a limitar nuestra atención por el momento al modelo probit de modo que $F = \Phi$, y asumir una priori uniforme para β . En este caso, un esquema de muestreo de Gibbs se puede utilizar para simular de la distribución conjunta del vector de rasgos latentes Z y β por simulación sucesiva de la distribución posterior de Z dado β , y β dado Z . La distribución condicional de β dado Z es:

$$\beta \sim N(\beta_{LS}, (X'X)^{-1}), \quad (3.9)$$

Donde β_{LS} es el estimador mínimos cuadrados de β basado en Z , por ejemplo, $(X'X)^{-1}X'Z$ y X es la matriz de covarianza con filas x_1', \dots, x_n' .

Del mismo modo, se observa que la distribución condicional de cada componente de Z , Z_i , dado y_i y β , tiene una distribución normal truncada con media $x_i'\beta$. Truncamiento a la izquierda de 0 si $y_i = 1$, y a la derecha de 0 si $y_i = 0$. Equivalentemente, la densidad condicional de Z_i dado β es proporcional a:

$$Z_i | \beta, y_i \propto \begin{cases} \phi(z; x_i'\beta, 1)I(z \leq 0), & \text{si, } y_i = 0, \\ \phi(z; x_i'\beta, 1)I(z > 0), & \text{si, } y_i = 1 \end{cases} \quad (3.10)$$

Donde $\phi(\cdot; a, b)$ denota una densidad normal con media a y varianza b , y $I(\cdot)$ es la función indicadora. Estas observaciones motivan un algoritmo de muestreo de Gibbs para la simulación de m iteraciones de la densidad posterior resultante de un modelo probit con una priori uniforme.

3.3. Revisión General de Clases de Enlaces Asimétricos

Considerando un vector $y = (y_1, y_2, \dots, y_n)'_{n \times 1}$ de “ n ” variables aleatorias independientes binarias correspondientes a “ n ” observaciones de la variable respuesta, $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})'$ un vector de dimensión $k \times 1$ correspondiente a “ k ” variables explicativas para la observación

i , además $x_{i1} = (1, \dots, 1)'$. Sea X la matriz diseño $n \times k$ con filas x'_i , $\beta = (\beta_1, \dots, \beta_k)'$ un vector $k \times 1$ de coeficientes de regresión asociados a las variables explicativas. Considere que $y_i = 1$ con probabilidad p_i y $y_i = 0$ con probabilidad de $(1 - p_i)$.

Como hemos visto en los modelos de datos binarios es común asumir que:

$$p_i = F(\eta_i) = F(x'_i \beta), i = 1, \dots, n \quad (3.11)$$

En el cual $F(\cdot)$ denota la función de distribución acumulada (FDA). La función inversa F^{-1} se denomina comúnmente función de enlace y $\eta_i = x'_i \beta$, es el i -ésimo predictor lineal. En el caso que F sea una FDA de una distribución simétrica, entonces el enlace resultante es simétrico y p_i tiene una forma simétrica alrededor de $p_i = 0,5$. En el caso que $F(\cdot)$ sea una FDA de una distribución normal estándar, nosotros tenemos el enlace probit, y en el caso que $F(\cdot)$ sea la FDA de una distribución logística nosotros tenemos el enlace logit.

En el presente trabajo proponemos que los enlaces asimétricos pueden obtenerse considerando las siguientes modificaciones del modelo definido en (4.1); como referencia ver Bazán et. al 2008:

- Tomando F como la FDA de una distribución asimétrica (1er Caso - BBB, Estándar).
- Considerando una modificación del predictor lineal (2do Caso).
- Considerando F en una clase general de distribuciones de probabilidad, por ejemplo, en una clase de mezclas de distribuciones simétricas con asimétricas (3er Caso - CDS).

Otro ejemplo muy popular del primer caso es el enlace log-log complementario o cloglog, donde la FDA usada en el enlace corresponde a la distribución de Gumbel. En este caso, la FDA está completamente especificada, no depende de ningún parámetro adicional desconocido y no presenta, como caso particular, un enlace simétrico. Otros ejemplos del primer caso se obtienen, por ejemplo, considerando las siguientes FDA:

$$F(\eta_i) = 1 - (1 + \exp^{\eta_i})^{-\lambda}, \quad (3.12)$$

$$F(\eta_i) = (1 + \exp^{-\eta_i})^{-\lambda}, \lambda > 0. \quad (3.13)$$

Los enlaces que se obtienen de estas FDA fueron propuestos por Prentice (1976) y fueron popularizados en la literatura económica por Nagler (1994). Esos enlaces son logit asimetrizados y son llamados comúnmente como scobit y power logit, respectivamente, e incluyen al enlace logit como caso especial cuando el parámetro incluido es $\lambda = 1$.

También cae en este caso el enlace Probit Asimetrizado (Skew Probit) propuesto por Bazán, Branco y Bolfarine (2006), denominado aquí BBB Probit Asimetrizado y otro denominado estándar Probit Asimetrizado (Bazán, Bolfarine y Branco 2006).

El segundo caso se obtiene cuando se mantiene a F como una distribución simétrica pero se considera una modificación del predictor lineal η_i por $m(\eta_i, l)$, donde $m(\cdot)$ es una función

no lineal y continua, y λ es el parámetro que controla la asimetría. Por ejemplo, en el modelo de Stukel (Stukel 1988), F corresponde a la distribución logística y la especificación de $m(\eta_i, l)$ es diferente para $\eta_i \geq 0$ y para $\eta_i < 0$. Es decir, $m(\eta_i, l)$ es una función partida para diferentes rangos de η_i y λ .

Otro enlace más simple en el segundo caso se obtiene cuando el predictor lineal se reemplaza por una expresión polinomial, generalmente cuadrática o cúbica. Un caso conocido, denominado logit cuadrático, se obtiene cuando F corresponde a la distribución logística y el predictor lineal simple $\beta_0 + \beta_1 X$ es reemplazado por $\beta_0 + \beta_1 X + \beta_2 X^2$.

Un ejemplo del tercer caso ocurre cuando F está en la clase de mezcla de escala de distribuciones elípticas dada por:

$$F = \{F(\cdot) = \int_{[0, \infty]} H(\cdot | \nu) dG(\nu)\}. \quad (3.14)$$

donde G es la FDA en $[0, \infty >$ y H es una distribución elíptica propuesta por Basu y Mukhopadhyay (2000). También cae en este caso el enlace Probit Asimetrizado propuesto en Chen et al. (1999), denominado comúnmente CDS Probit Asimetrizado, el cual considera una clase de mezcla de distribuciones normales, donde la medida de mezcla es la distribución normal positiva que tiene función de densidad de probabilidad (FDP) dada por $g(x) = 2\phi(x)$, $x > 0$, donde $\phi(\cdot)$ es la FDP de la distribución normal estándar. Otro caso similar ocurre cuando se mezcla la normal positiva con la distribución logística que se conoce como logística asimetrizada o Skew Logit.

3.4. Modelos Skew Probit Generalizados

De acuerdo a Bazan et. al (2006), Bazan, J., Branco, M., Bolfarine, H. (2010) y otros respectivamente, podemos concluir que existen algunos enlaces especiales dentro de la clase Skew Probit Generalizados; a continuación se especifican los 03 existentes:

3.4.1. Modelo Skew Probit CDS.

Siendo:

$$p_i = F_\theta(\eta_i) = F_\theta(x_i' \beta) \quad (3.15)$$

$$\eta_i = x_i' \beta \quad (3.16)$$

$$\theta = (\mu; \sigma^2; \lambda) \quad (3.17)$$

F = Skew Normal

Modelo Skew Probit CDS:

$$\theta = (0; 1 + \lambda^2; -\lambda) \quad (3.18)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu - \sigma\delta v}{\sigma\sqrt{1 - \delta^2}}\right)dv \quad (3.19)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu}{\sigma}\sqrt{1 + \lambda^2} - \lambda v\right)dv \quad (3.20)$$

Por la proposición 2, (Bazán et al. 2006, pág, 4 - 7), se reescribe:

$$F_{\theta}(\eta_i) = 2\Phi_2\left[\begin{pmatrix} \eta_i \\ 0 \end{pmatrix}; \begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (3.21)$$

Donde:

- Φ_2 , es la Función de Distribución Normal Bivariada.
- $\delta = \frac{-\lambda}{\sqrt{1+\lambda^2}}$.
- $F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta)$
- $\mu = 0$.
- $\sigma^2 = 1 + \lambda^2$

Por tanto el modelo CDS se define como:

$$F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta) = 2\Phi_2\left[\begin{pmatrix} x'_i\beta \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 + \lambda^2 & \frac{\lambda}{\sqrt{1+\lambda^2}} \\ \frac{\lambda}{\sqrt{1+\lambda^2}} & 1 \end{pmatrix}\right] \quad (3.22)$$

3.4.2. Modelo Skew Probit BBB.

Siendo:

$$p_i = F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta); \quad (3.23)$$

$$\eta_i = x'_i\beta \quad (3.24)$$

$$\theta = (\mu; \sigma^2; \lambda) \quad (3.25)$$

Modelo Skew Probit BBB:

$$\theta = (0; 1; \lambda) \quad (3.26)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu - \sigma\delta v}{\sigma\sqrt{1 - \delta^2}}\right)dv \quad (3.27)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu}{\sigma}\sqrt{1 + \lambda^2} - \lambda v\right)dv \quad (3.28)$$

Por la proposición 2, (Bazán et al. 2006, pág, 4 - 7), reescribimos:

$$F_{\theta}(\eta_i) = 2\Phi_2\left[\begin{pmatrix} \eta_i \\ 0 \end{pmatrix}; \begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (3.29)$$

Donde:

- Φ_2 , es la Función de Distribución Normal Bivariada.
- $\delta = \frac{-\lambda}{\sqrt{1+\lambda^2}}$.
- $F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta)$
- $\mu = 0$.
- $\sigma^2 = 1$

Por tanto el Modelo BBB se define como:

$$F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta) = 2\Phi_2\left[\begin{pmatrix} x'_i\beta \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -\frac{\lambda}{\sqrt{1+\lambda^2}} \\ -\frac{\lambda}{\sqrt{1+\lambda^2}} & 1 \end{pmatrix}\right] \quad (3.30)$$

3.4.3. Modelo Skew Probit ESTANDAR.

Siendo:

$$p_i = F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta); \eta_i = x'_i\beta \quad (3.31)$$

$$\theta = (\mu; \sigma^2; \lambda) \quad (3.32)$$

Modelo Skew Probit ESTANDAR:

$$\theta = (0; 1 + \lambda^2; \lambda) \quad (3.33)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu - \sigma\delta v}{\sigma\sqrt{1 - \delta^2}}\right)dv \quad (3.34)$$

$$F_{\theta}(\eta_i) = \int_0^{\infty} g(v)\Phi\left(\frac{\eta_i - \mu - \sigma\delta v}{\sigma\sqrt{1 - \delta^2}}\right)dv \quad (3.35)$$

Por la proposición 2, (Bazán et al. 2006, pág, 4 - 7), se reescribe:

$$F_{\theta}(\eta_i) = 2\Phi_2\left[\begin{pmatrix} \eta_i \\ 0 \end{pmatrix}; \begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (3.36)$$

Donde:

- Φ_2 , es la Función de Distribución Normal Bivariada.
- $\delta = \frac{-\lambda}{\sqrt{1+\lambda^2}}$.

- $F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta)$
- $\mu = 0.$
- $\sigma^2 = \frac{1}{1-\delta^2}$

Por tanto el Modelo Estándar se define como:

$$F_{\theta}(\eta_i) = F_{\theta}(x'_i\beta) = 2\Phi_2\left[\begin{pmatrix} x'_i\beta \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 + \lambda^2 & -\frac{\lambda}{\sqrt{1+\lambda^2}} \\ -\frac{\lambda}{\sqrt{1+\lambda^2}} & 1 \end{pmatrix}\right] \quad (3.37)$$

$$\theta \approx \left(-\frac{\sqrt{2}\delta}{\sqrt{\pi - 2\delta^2}}, \frac{\pi}{\pi - 2\delta^2}, \lambda\right) \quad (3.38)$$

Donde:

θ = Probabilidad.

3.4.4. Regresión Skew Probit - Versión Aumentada

De la revisión literaria encontramos según Chen et al. (1999) que se propone un modelo Skew Probit con el uso de una variable latente auxiliar, el cual lo denota por CDS_{sp} (CDS Skew Probit), denotándose lo siguiente:

$$y_i = \begin{cases} 1, \text{ si } S_i > 0, \\ 0, \text{ si } S_i \leq 0 \end{cases} \quad (3.39)$$

Donde:

$$S_i = \eta_i + \lambda z_i + \epsilon_i; \quad (3.40)$$

Siendo $\eta_i = x'_i\beta$

Ademas:

- z_i
- ϵ_i

Son independientes.

Considerándose que:

$z_i \sim$ Half Normal (0, 1) y $\epsilon_i \sim$ Normal (0, 1).

Asi, la propuesta dada por Chen 1999, se expresa como:

$$p_i = \int_0^{\infty} \Phi(\eta_i + \lambda z_i) g(z_i) dv_i \quad (3.41)$$

- $\Phi(\cdot)$, es la FDA de una Normal Estándar
- $g(\cdot)$, es la FDP de una Half Normal
- λ , representa el parámetro de asimetría; $\lambda \in \mathfrak{R}$.

La expresión anterior es un caso especial del modelo SAHU et al. (2003), la cual es llamada representación estocástica.

Por otro lado, Bazán et al. (2005), considera a F (distribución conocida) como una FDA de una distribución skew normal, dada por Azzalini (1985). Dentro de la estructura de variable latente, se presenta a:

$$p_i = F_\lambda(\eta_i) = 2\Phi_2\left[\begin{pmatrix} \eta_i \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & -\delta \\ -\delta & 1 \end{pmatrix}\right] \quad (3.42)$$

Donde:

- $F_\lambda(\cdot)$, es la FDA de una distribución Skew Normal.
- λ es el parámetro de asimetría.
- $\Phi_2(\cdot)$, es la FDA de una Normal Estándar Bivariada
- $-\delta$, es el coeficiente de correlación.

Siendo: $\delta = -\frac{\lambda}{\sqrt{1+\lambda^2}}$, lo cual se encuentra en el rango de -1 a 1 .

Un aspecto importante que se rescata de las lecturas es respecto a los enlaces CDS_{sp} y los BBB_{sp} , se refiere a λ el cual esta asociado con $F(\cdot)$ y además es independiente del η_i , es por ello que se escribe $F_\lambda(\cdot)$.

Así también la principal diferencia que se encuentra entre los modelos CDS_{sp} y BBB_{sp} , en este ultimo no es necesario considerar una representación de una estructura lineal latente como si lo es necesario en un CDS_{sp} , por ello es suficiente especificar a p_i como una CDF de una distribución asimétrica y la expresión CDS_{sp} puede ser obtenida como una BBB_{sp} .

Bazán et al. (2006), propone un modelo Skew Probit Generalizado (GSP), una clase GSP se obtiene considerando lo siguiente:

$$p_i = F_\theta(\eta_i) = F_\theta(x'_i\beta); \quad (3.43)$$

Para $i=1,2,\dots,n$

Donde:

- $F_\theta(\cdot)$, es una FDA de la skew normal (Normal Asimétrica o Asimetrizada).
- $\theta = (\mu; \sigma^2; \lambda)$, es el vector de parámetros.

A manera de corolario y de acuerdo al valor obtenido de λ (Coeficiente de Asimetría), algunos enlaces especiales de GSP son dados como sigue:

- Si $\theta = (0, 1, 0)$ entonces se obtiene un probit clásico.

- Si $\theta = (0, 1 + \lambda^2, -\lambda)$, entonces se tiene una CDS_{sp} .
- Si $\theta = (0, 1, \lambda)$, entonces se tiene una BBB_{sp} .

Si se considera diferentes valores o expresiones para μ y σ^2 entonces se puede formular distintos modelos skew probit, lo cual se relaciona según sea el enlace.

Un modelo General puede ser especificado cuando los parámetros μ y σ^2 son estimados a partir de los datos observados, a este modelo se le conoce como Modelo General con enlace SP completo (Skew Probit Completo).

Otro interesante enlace skew probit es obtenido considerando la distribución skew normal con las siguientes especificaciones:

$$\mu = -\frac{\sqrt{2}\delta}{\sqrt{\pi - 2\delta^2}} \quad (3.44)$$

$$\sigma^2 = \frac{\pi}{\pi - 2\delta^2} \quad (3.45)$$

A estas representaciones se le conoce como skew probit estándar o estándar SP.

La función de verosimilitud para la clase GSP (General Skew Probit) está dado por:

$$\ell(\beta; \theta | y; x) = \prod_{i=1}^n [F_{\theta}(\eta_i)]^{y_i} \times [1 - F_{\theta}(\eta_i)]^{1-y_i} \quad (3.46)$$

- η_i , es el predictor lineal.
- X , es la matriz de diseño.
- $F_{\theta}(\eta_i) = F_{\theta}(x'_i \beta)$, $i=1, 2, \dots, n$.

Otra versión de la función de verosimilitud se obtiene cuando se considera una data aumentada, así:

$v_i \sim HN(0, 1)$ para $i = 1, 2, \dots, n$, siendo ésta la variable latente auxiliar.

Además:

$$\eta_i = \frac{\eta_i - \mu}{\sigma}$$

Obteniéndose:

$$\ell(\beta; \theta | y; X; v) = \prod_{i=1}^n [\Phi(\eta_i \sqrt{1 + \lambda^2} - \lambda v_i)]^{y_i} \times [1 - \Phi(\eta_i \sqrt{1 + \lambda^2} - \lambda v_i)]^{1-y_i} \times g(v_i) \quad (3.47)$$

Siguiendo la función de verosimilitud marginal para la clase GSP, viene expresado como:

$$\ell(\beta, \theta | y, X) = \prod_{i=1}^n \int_0^{\infty} [\Phi(\eta_i \sqrt{1 + \lambda^2} - \lambda v_i)]^{y_i} \times [1 - \Phi(\eta_i \sqrt{1 + \lambda^2} - \lambda v_i)]^{1-y_i} \times g(v_i) dv_i \quad (3.48)$$

Por último para obtener las funciones de verosimilitud para los enlaces BBB y CDS solo es necesario considerar los valores de θ dados anteriormente para $F_{\theta}(\eta_i)$, para $i = 1, 2, \dots, n$.

Dos versiones aumentadas del modelo serán presentadas en el siguiente capítulo junto con la especificación de Priors de modo de obtener la versión jerárquica del modelo, el cual es útil desde la perspectiva Bayesiana.



Capítulo 4

Inferencia Bayesiana para el Modelo Skew Probit

En este capítulo se abordan temas sobre Regresión Binaria basada en Inferencia Bayesiana para el modelo Skew Probit; en general seguiremos el trabajo de Bazan et al (2010).

4.1. Verosimilitud del Modelo

Las clases de enlaces skew-probit (sp) se obtiene considerando F en la especificación del modelo de regresión binaria y que se escribe como:

$$p_i = F_\theta(\eta_i) = F_\theta(x_i'\beta), \quad (4.1)$$

Para: $i = 1, \dots, n$.

Donde $F_\theta(\cdot)$ es la FDA de la distribución skew normal dado anteriormente con un vector de parámetros $\theta = (\mu, \sigma^2, \lambda)$.

Algunos enlaces especiales de la clase SP siguen de (4.1) considerando diferente parametrización. Cuando $\theta = (0, 1, 0)$ sigue el enlace probit, cuando $\theta = (0, 1 + \lambda^2, -\lambda)$ sigue el enlace skew probit CDS, y cuando $\theta = (0, 1, \lambda)$ sigue el enlace skew probit BBB. Nótese también que una variedad de diferentes enlaces skew-probit se pueden formular teniendo en cuenta otros valores para μ y σ^2 dado en (4.1). Un interesante enlace skew-probit se puede obtener considerando la distribución skew normal con $\mu = -\frac{\sqrt{2}\delta}{\sqrt{\pi-2\delta^2}}$ y $\sigma^2 = \frac{\pi}{\pi-2\delta^2}$ con resultados $E[R] = 0$ y $V(R) = 1$, este enlace será nombrado skew probit Estándar.

El modelo de regresión binaria asimétrica propuesto aquí se caracteriza por: (i) la probabilidad de éxito se obtiene considerando una función de distribución acumulada (FDA) evaluados en el predictor lineal, (ii) un parámetro de asimetría asociado con estas funciones (FDA) se introduce de forma independiente en el predictor lineal, y (iii) una estructura lineal latente no es necesario para la formulación de enlaces. Sin embargo, como se verá más adelante, dicha estructura se considera para fines de cálculo. La función de verosimilitud para la clase de los modelos Skew Probit está dada por:

$$L(\beta; \theta | y; x) = \prod_{i=1}^n [F_\theta(\eta_i)]^{y_i} \times [1 - F_\theta(\eta_i)]^{1-y_i} \quad (4.2)$$

Donde:

- η_i , es el predictor lineal.
- X , es la matriz de diseño.

- $F_{\theta}(\eta_i)$, es la función dada en (4.1).

Una función de verosimilitud aumentada para la clase Skew Probit, similar a dado por Chen et al. (1999), puede ser obtenido mediante $v_i \sim \text{HN}(0,1)$, $i=1, \dots, n$, con variables latentes, además con $\eta_i^* = \frac{\eta_i - \mu}{\sigma}$, así:

$$L(\beta; \theta | y, X, v) = \prod_{i=1}^n [\Phi(\eta_i^* \sqrt{1 + \lambda^2} - \alpha v_i)]^{y_i} \times [1 - \Phi(\eta_i^* \sqrt{1 + \alpha^2} - \alpha v_i)]^{1-y_i} g(v_i) \quad (4.3)$$

Siendo la función de verosimilitud aumentada para la clase de Skew Probit,

$$L(\beta; \theta | y, X) = \prod_{i=1}^n \int_0^{\infty} [\Phi(\eta_i^* \sqrt{1 + \alpha^2} - \alpha v_i)]^{y_i} \times [1 - \Phi(\eta_i^* \sqrt{1 + \alpha^2} - \alpha v_i)]^{1-y_i} g(v_i) dv_i \quad (4.4)$$

Para obtener las versiones de la función de verosimilitud para el enlace Skew Probit BBB y el enlace Skew Probit CDS, sólo es necesario tener en cuenta los valores correspondientes de θ . Nótese que la función de probabilidad aumentada marginal fue utilizado por Chen et al. (1999) para obtener estimaciones bajo enfoques bayesianos y clásicos.

4.2. Procedimiento Inferencial para el enlace Skew Probit

Se observa que en la clase de enlaces Skew Probit los parámetros θ y β tienen significados diferentes. Mientras θ es un vector de parámetros estructurales asociados con la elección de la función de enlace, el parámetro β es un vector de parámetros inherentes a los datos observados y no dependiente de la elección del modelo. Por tanto dos escenarios se pueden considerar. En el primer escenario θ y β son estimados, en el segundo escenario sólo β se permite que varíe y θ se fija en su valor “verdadero” θ_0 . Igual que Taylor J. y Siqueira A. (1996), nos referiremos a estos dos escenarios como Incondicional y Condicional respectivamente.

1. Escenario Condicional - 1ra Proposición.

Realizar Inferencia bajo el escenario condicional es más fácil tanto por máxima verosimilitud así como por enfoques Bayesianos. Sin embargo, se deberá imponer condiciones en la matriz de diseño para la existencia de estimadores ML (Máxima Verosimilitud o Likelihood Maximun) y la distribución posterior de β bajo prioris uniformes impropias. La siguiente proposición es una consecuencia directa de los resultados de Chen y Shao (2000).

Proposición 1: Al considerar el enfoque condicional para la clase SP, es decir, θ es conocida o fijo y β es el único parámetro de interés, y dejar que $t_i = 1$ si $y_i = 0$, y $t_i = -1$ Si $y = 1$, X el $n \times k$ la matriz de diseño conocida con “n” filas y se define X^* como la matriz con filas $T_i x_i'$. Luego bajo las condiciones:

- (C1) X es de rango completo;
- (C2) existe un vector positivo $a = (a_1, \dots, a_n)'$ tal que $X^{*'}a = 0$, se sigue que:
 - Para una priori uniforme impropia dado por β , $\pi(\beta) \propto 1$, la distribución posterior de β es adecuada, esto es, $\int_{R^k} L(\beta|y, X, \theta)d\beta < \infty$;
 - el estimador de máxima verosimilitud (MLE) de β existe.

La prueba de esta proposición es presentada en el anexo de Bazán et al. (2010)

2. Escenario Incondicional - 2da Proposición.

En el enfoque incondicional, el cálculo de los estimadores Máxima Verosimilitud utilizando las versiones de las funciones de verosimilitud dado en la sección anterior no es sencillo y es necesario desarrollar nuevas condiciones para la existencia de los estimadores Máxima Verosimilitud. Además, se estudia la propiedad de la distribución posterior bajo Prioris impropias uniformes para β y α pero teniendo en cuenta Prioris propias para μ y σ^2 . Estos escenarios consideran independencia entre Prioris, de manera que:

$$\pi(\beta, \theta) = \pi(\beta) \cdot \pi(\theta) \tag{4.5}$$

Donde: $\pi(\theta) = \pi(\alpha)\pi(\mu)\pi(\sigma^2)$,

La siguiente proposición es una consecuencia directa de los resultados en Chen et al. (1999).

Proposición 2. Al considerar el enfoque incondicional para la clase Skew Probit, es decir, θ y β son los parámetros de interés, y dejar que $S_0 = i : y_i = 0$, $S_1 = j : y_j = 1$, $m_0 = \text{card}(S_0)$, $m_1 = \text{card}(S_1)$ y X_0 y X_1 son las matrices con x'_i filas, $i \in S_0$ y x'_j , $j \in S_1$. Luego bajo las condiciones:

- C1) X_0 y X_1 , tienen rango completo;
- C2) existe un vector positivo $a_0 = (a_{01}, \dots, a_{0m_0})'$ que pertenece R^{m_0} y $a_1 = (a_{11}, \dots, a_{1m_1})'$ que pertenece R^{m_1} tal que $X_0^{*'}a_0 = 0$ y $X_1^{*'}a_1 = 0$, implica que:
 - Para una priori $\pi(\beta, \lambda)$ aproximado 1, y $\pi(\mu)$ y $\pi(\sigma^2)$ siendo prioris propias, la distribución posterior conjunta de (β, θ) , donde $\theta = (\mu, \sigma^2, \lambda)$, es propia;

$$\int_0^\infty \int_{-\infty}^\infty L(\mu, \sigma^2|y, X)\pi(\mu)\pi(\sigma^2)d\mu d\sigma^2 < \infty$$

Donde: $L(\mu, \sigma^2|y, X) = \int_{R^k} L(\beta, \theta|y, X)d\beta d\lambda$;

- El estimador de máxima verosimilitud (MLE) de β existe.

Las pruebas de los resultados de la presente sección se encuentran en el apéndice de Bazan, J., Branco, M., Bolfarine, H. (2010), páginas 693-695.

4.2.1. Verosimilitud Aumentada del Modelo Skew Probit

En esta sección, se presentan dos nuevas funciones de Verosimilitud de data - completa (o verosimilitudes aumentadas) para la clase skew probit. Comenzamos con una importante alternativa:

1. Verosimilitud Aumentada - 3ra Proposición.

Proposición 3. La clase de enlaces skew probit esta equivalentemente definido, teniendo en cuenta que:

$$y_i = I(s_i > 0) = \begin{cases} 1, s_i > 0, \\ 0, s_i \leq 0 \end{cases} \quad (4.6)$$

$$s_i \sim SN(\eta_i - \mu; \sigma^2; -\alpha)$$

$$i = 1; \dots; n$$

Donde $I(\cdot)$ es la Función Indicador Usual.

La prueba de esta proposición es presentada en el anexo de Bazán et al. (2010).

Las variables latentes s_i son introducidos para evitar trabajar con el tipo de probabilidades de Bernoulli y la representación muestra una estructura lineal latente que produce un modelo equivalente a la clase skew probit. Además, observe que el parámetro de asimetría para las variables latentes auxiliares es el negativo del valor del parámetro de asimetría. Por lo tanto, la función de verosimilitud para data completa para la clase skew probit está dada por:

$$L(\beta, \theta | s, y) = \prod_{i=1}^n f_{\theta}(s_i) p(y_i | s_i) \quad (4.7)$$

Donde, $p(y_i | s_i) = I(s_i > 0)I(y_i = 1) + I(s_i \leq 0)I(y_i = 0)$, $i = 1, \dots, n$. Cuando $\mu = \eta_i$, $\sigma^2 = 1$ y $\alpha = 0$, sigue el resultado correspondiente dado por Albert y Chib (1993) para el enlace probit.

Un enfoque adicional para el caso de data aumentada considera:

$$s_i = \eta_i + e_i, \quad (4.8)$$

Donde : $e_i \sim SN(-\mu, \sigma^2, -\alpha)$

Por la propiedad d, descrita para la distribución skew normal, se tiene:

$$e_i = -\mu + \sigma(-\delta v_i - \sqrt{1 - \delta^2} w_i) \quad (4.9)$$

Para, $i = 1; \dots; n$.

Considerándose además que $v_i \sim HN(0; 1)$, y $w_i \sim N(0; 1)$. La distribución condicional $e_i | v_i$ es una distribución normal con media $E[e_i | v_i] = -\sigma \delta v_i - \mu$ y varianza $Var[e_i | v_i] = (1 - \delta^2) \sigma^2$; por tanto $s_i^* = s_i | v_i \sim N(\eta_i - \mu - \sigma \delta v_i; (1 - \delta^2) \sigma^2)$.

De acuerdo a estos resultados la simulación lineal de S_i se puede realizar en dos pasos: i) simular v_i para $HN(0, 1)$ y luego simular s_i^* para $N(\eta_i - \mu - \sigma \delta v_i, (1 - \delta^2) \sigma^2)$

De acuerdo a ello se obtiene una función de verosimilitud de data completa skew probit:

$$L(\beta, \theta | s^*; v; y) = \prod_{i=1}^n \phi(s_i^*) g(v_i) p(y_i | s_i) \quad (4.10)$$

La densidad $p(y_i | s_i)$ puede ser sustituido por el correspondiente MCMC versión estimada $p(y_i | s_i^*)$, porque $I(S_i, y_i)$ implica que $I(S_i^*, y_i)$.

Nótese que los errores e_i son independientes y pueden ser considerados como residuos latentes (Albert y Chib, 1995). Los valores $e_i = s_i - \eta_i$, $i = 1; \dots; n$ pueden ser estimados y se utiliza para la verificación de los modelos. Para entender como los y_i cambian la distribución de residuos, tenga en cuenta las distribuciones condicionales posteriores de e_i con β , θ , s_i , y v_i , esto es $e_i^* = e_i | \beta; \theta; y_i; s_i; v_i$.

4.3. Especificación de Prioris

En el contexto de un análisis bayesiano, es necesario especificar las distribuciones aprioris de β, θ . Consideramos prioris independientes como se indica en (4.3.1) y también para cada componente de β tenemos en cuenta la independencia.

Dos prioris típicas son consideradas para el modelo probit, la priori normal $\beta_j \sim N(\mu_{\beta j}; \sigma_{\beta j})$ o priori uniforme $\pi(\beta) = 1$. La última priori es impropia, pero como se muestra en la Segunda Proposición los resultados de la distribución posterior es adecuada. Ahora bien, es importante considerar prioris propias para μ y σ , tales como $\mu \sim N(\mu_0, \tau_0^2)$ y $\sigma^2 \sim Inv - x^2(w, k)$, la escala de la distribución inversa-chi-cuadrado con $w > 0$, s grados de libertad y parámetro de escala $k > 0$. Podemos también trabajar con la definición $1/\sigma^2$, en cuyo caso la distribución a priori Gamma(w, k) puede ser especificada.

Por último, teniendo en cuenta la “parametrización delta” una priori no informativa sigue $\delta \sim U(-1; 1)$. No es difícil ver que, es equivalente a considerar $\alpha \sim T(0; 0,5; 2)$, donde $T(a, b, v)$ denota la distribución t de Student con localización a , escala b , y v grados de libertad (ver Bazán et al., 2006). Otra priori no informativa utilizada en la “parametrización alfa” es la Priori de Jeffreys.

Al considerar la primera función de verosimilitud de data completa dado en (4.7) y la estructura latente en (4.6), las condicionales completas de la clase skew probit puede ser obtenidas. Sin embargo, algunas de las condicionales completas no pueden ser directamente

muestreadas, requieren de algoritmos complejos como el de Metropolis-Hastings. Además, las distribuciones skew normales no están disponibles directamente en WinBugs.

Una segunda manera de aplicar el algoritmo MCMC es utilizar la segunda función de verosimilitud de data completa (4.10) que es fácil de implementar en WinBugs ya que todas las distribuciones están disponibles en dicho software. Seguidamente se desarrolla un procedimiento de cálculo para la clase Skew Probit con sede en esta segunda función de verosimilitud aumentada. La especificación de la estructura jerárquica para la parametrización delta puede obtenerse considerando:

$$\blacksquare S_i^* | v_i, x_i, y_i, \beta, \mu, \sigma^2, \delta \sim N(x_i' \beta - \mu - \sigma \delta v_i, (1 - \delta^2) \sigma^2) I(S_i^*, y_i)$$

Asumiendo que:

$$V_i \sim HN(0, 1), \beta \sim \pi(\beta), \mu \sim \pi(\mu), \sigma^2 \sim \pi(\sigma^2), \delta \sim \pi(\delta)$$

Para la especificación de la estructura jerárquica de la parametrización de alfa es necesario especificar las distribuciones aprioris para α y luego usar la transformación $\delta = \frac{\alpha}{1 + \alpha^2}$.

La estructura jerárquica anterior puede ser implementada en el software WinBugs. Considerando los enlaces Skew Probit CDS y BBB, las líneas cuarta y quinta en estructura jerárquica se eliminan. Además, cuando $\delta = 0$, la estructura jerárquica de la verosimilitud aumentada sigue el correspondiente al modelo probit normal.

4.3.1. Elección de Prioris Alternativa

Siendo el objetivo elegir una adecuada distribución conjunta apriori de las variables aleatorias β y λ , una posibilidad es el uso de las probabilidades condicionales asumiendo la dependencia entre ellos; Chen et al. (1999) proponen dos soluciones alternativas: el uso de distribuciones apriori no informativas o distribuciones apriori informativas para $\pi(\beta, \lambda)$. En el primer caso se requiere que la matriz $X_{n \times k}$, tenga un rango completo (véase el teorema 1, Chen et al. (1999)). En el presente planteamiento esta condición se cumple pues en nuestra base de datos se muestra que el numero de columnas es mucho menor que el numero de filas; por lo tanto, el desarrollo de nuestra investigación es guiada a través de una distribución apriori informativa para (β, λ) .

Dado que los coeficientes del modelo Probit no tienen una interpretación natural, hemos aceptado una manera simple los siguientes supuestos:

- (1) β y λ son independientes.
- (2) los odds ratios representados por $\text{odd} = \frac{p_i}{1-p_i} = \exp\{x_i' \beta\}$, se pueden obtener fácilmente.

Los odds ratios reflejan la probabilidad de que, el caso suceda dividido por la probabilidad de que el caso no suceda. Obviamente, si los odds son mayores que uno, entonces el evento es más probable que ocurra a que no ocurra. Estos valores odds ratios pueden ser utilizados para obtener la distribución normal apriori de la siguiente manera:

- Primero, tenemos en cuenta apriori el conocimiento sobre la media y la varianza de la odds ratio para cada covariable. En este sentido, podemos considerar el comportamiento de estas covariables observadas en otros estudios, o tener en cuenta el conocimiento o afirmación aproximada acerca de ellos. Suponiendo que la distribución a priori del vector de coeficientes β es normal, la distribución apriori de los odds ratios es log-normal.
- Segundo, podemos obtener la media y la varianza aprioris utilizando la relación mostrada como ecuación (siguiente):

$$\beta_k \sim N(\beta_k^0; \sigma_k^{-1}) \iff odd_k = \exp\{\beta_k\} \quad (4.11)$$

$$\sim LN(odd_k^0; \Sigma_{odd_k}^{-1}) \quad (4.12)$$

Donde LN denota la distribución log-normal, y los dos primeros momentos son:

$$E[odd_k] = odd_k^0 = \exp(\beta_k^0 + \frac{1}{2}\sigma_k^{-1}) \quad (4.13)$$

$$Var[odd_k] = \Sigma_{odd_k}^{-1} = \exp(2.\beta_k^0 + \sigma_k^{-1}).(\exp(\sigma_k^{-1}) - 1) \quad (4.14)$$

Los conocimientos aproximados o estudios previos nos pueden ayudar a obtener alguna información sobre la media y la varianza de los odds ratios. Entonces, podemos obtener información apriori acerca de los coeficientes β para la solución de las ecuaciones aprioris. Utilizando los odds proporcionados $\beta \sim N(\beta^0; \sigma_1 I)$, donde I representa la matriz identidad y $\beta^0 = (\beta_1^0; \dots; \beta_k^0)$ es el vector de coeficientes que incluye a β_0 , como valor constante. Ahora, via un modelo skewed asimétrico, λ debe ser estimado tomando valores no iguales a 0. Con este objetivo, tenemos una distribución normal simétrica con una desviación estándar grande, es decir, $\lambda \sim N(0, \sigma^2)$, con σ^2 grande. En este punto, es preciso mencionar que una varianza grande será menos informativa y tendrá menos impacto en la estimación.

4.4. Selección de Modelos

A continuación se presenta el análisis correspondiente de criterios como fundamento para la selección, comprensión y evaluación de modelos, bajo modelos skew probit; entre los principales se tiene; Dbar (Media de los Desvíos a Posteriori), DIC (Criterios de Desviación de Información), Tablas de Contingencia y Análisis de Curvas COR (Característica Operativa del Receptor), tal como se propone en Bazan et al. (2010), Spiegelhalter J. (2002) y otros.

Los criterios señalados son validos particularmente en la selección de modelos bayesianos donde las distribuciones posteriores de los modelos han sido obtenidos vía simulación MCMC; el criterio se considera válido cuando la distribución posterior es aproximadamente normal. El criterio supone medición del modelo sobre datos reales; un mayor DIC o Dbar es el peor ajuste; es decir valores aproximados al verdadero valor de las medias.

El DIC así como el Dbar se estiman vía estimación MCMC, su forma de cálculo implica que:

Siendo $DIC = p_D + D^*$; se define la desviación $D(\theta) = -2\log(p(y|\theta)) + C$, donde “y” son los datos, θ son los parámetros desconocidos del modelo y $p(y|\theta)$ es la función de verosimilitud, C es una constante que se anula en los cálculos de comparación de modelos, por lo que necesita ser conocido.

La expectativa $D^* = E^\theta[D(\theta)]$ es una medida de ajuste del modelo respecto a datos reales; a mayor D^* peor es el ajuste. El número efectivo de parámetros del modelo se calcula como $p_D = D^* - D(\theta^*)$, donde θ^* es la expectativa de θ ; un dato más grande supone menos sensibilidad del modelo a sus datos reales (ajuste); según Spiegelhalter et al. (2002) también $DIC = Dbar + p_D$.

Los cálculos fueron resueltos mediante el uso de software WinBUGS y OpenBUGS así como el BRMUW, en la generación de la sintaxis y las posteriores.

El análisis según Curvas de Características Operativas del Modelo (COR) (Balakrishnan, N., 1991) es una representación gráfica de la sensibilidad frente a “1 - especificidad” para un sistema clasificador binario según se varía el umbral de discriminación, proporciona herramientas para seleccionar los modelos posiblemente óptimos y descartar modelos sub óptimos independientemente de (y antes de especificar) el coste de la distribución de las dos clases sobre las que se decide. Las tablas de contingencia permiten analizar y registrar la relación de dos o mas variables especialmente Cuantitativas o Cualitativas (Binarias especialmente versus pronósticos); aplica como complemento al análisis de curvas (COR) respecto a datos reales versus pronosticados; siendo el análisis via curvas COR una representación gráfica de la sensibilidad frente a la especificidad (restado 1) para un sistema clasificador binario según se varía el umbral de discriminación.

4.5. Estructura Jerárquica de la Regresión Skew Probit

De acuerdo al análisis presentado en el capítulo 3 y secciones precedentes del presente capítulo, el modelo puede ser escrito de la siguiente forma:

$$L(\beta, \theta | s^*; v; y) = \prod_{i=1}^n \phi(s_i^*) g(v_i) p(y_i | s_i) \quad (4.15)$$

La especificación de la estructura jerárquica para la parametrización delta puede obtenerse considerando ($\delta = \frac{\alpha}{1+\alpha^2}$):

- $S_i^* | v_i x_i, y_i, \beta, \mu, \sigma^2, \delta \sim N(x_i' \beta - \mu - \sigma \delta v_i, (1 - \delta^2) \sigma^2) I(S_i^*, y_i)$,

Asumiendo que:

- $V_i \sim HN(0, 1)$,

- $\beta \sim \pi(\beta)$,
- $\mu \sim \pi(\mu)$,
- $\sigma^2 \sim \pi(\sigma^2)$,
- $\delta \sim \pi(\delta)$,

4.6. Implementacion Computacional

De acuerdo a las propuestas y detalles de autores como Bazán J, Bolfarine H, Branco y Márcia D.(2010), así como Greene, W (2007); la implementación computacional se puede establecer de la siguiente manera:

1. La selección de la verosimilitud, se realiza en función al tipo de regresión planteada en el modelo inicial (Bernoulli) de regresión skew probit generalizado bajo sus 03 enfoques, Modelo CDS, Modelo BBB y Modelo Estándar, bajo una muestra de datos asimétricos; resultados que se compararan con aquel obtenido por aplicación del Modelo Probit tradicional, bajo una muestra de datos simétrica.
2. La definición de la priori, que detallamos en el capítulo 4; dado el tipo de información y muestra obtenida sobre la Micro Financiera se ha definido establecer una Priori informativa tanto para β así como para λ .
3. La sintaxis correspondiente para la evaluación de los modelos se describe en el apéndice B del presente trabajo de investigación. El soporte informático fue mediante el uso del software WinBUGS, BRMUW (Bazán. J. (2010). Manual de uso de BRMUW Software. Departamento de Ciencias. PUCP. Disponible en <http://argos.pucp.edu.pe/jlbazan/download/ManualdeusoBRMUW.pdf>) y el software OpenBUGS.
4. El uso de criterios de evaluación de modelos Dbar y DIC fueron contrastados con aquel análisis según Tablas de Contingencia o Curvas COR; grado de especificidad y sensibilidad de modelos.
5. Los análisis se complementaron con aquellos de Densidades, Auto correlación, Histogramas, Quantiles, Percentiles, Medias, Varianzas, etc; que las estadísticas a posteriori de cada modelo muestren como resultados.
6. Cabe indicar que para el caso de los modelos BBB, CDS y Estándar no se requiere explicitar las correspondientes prioris, tanto para μ así como para σ , dado que las mismas son específicas en cada modelo.
7. Finalmente, es preciso resaltar que los detalles específicos (formulas y sintaxis) para cada modelo, BBB, CDS y Estándar (Verosimilitud, Prioris, Posterior) se muestran en el Apéndice B del presente trabajo de investigación (Resumen de Programas Utilizados y Sintaxis de Modelos).

Capítulo 5

Aplicación: Cálculo de la Probabilidad de Default en el Ambito del Sistema Financiero

Siendo comúnmente los modelos de comportamiento de elección discreta aleatoria empleados en la literatura econométrica, así como la regresión binaria aplicada por ejemplo, para evaluar riesgos crediticios de conceder o no una línea de crédito, la posibilidad de otorgar un préstamo, o caso contrario de no hacerlo, etc. Entre los modelos de elección discreta aleatoria, la regresión binaria usando un enlace logístico o usando un enlace probit entre otros, proponen enlaces simétricos además de ser los usados actualmente; en el presente estudio la propuesta es la aplicación de modelos de datos binarios pero con enlaces asimétricos, condición que determinara finalmente la propuesta de un modelo ideal.

Para la implementación de los modelos se hará uso del programa BRMUW, el mismo que provee los diversos modelos revisados en los capítulos anteriores, entre otros. El método sera ilustrado considerando información para una muestra de clientes de créditos, Créditos tipo Consumo, de una cartera de clientes de créditos de una entidad microfinanciera a un periodo determinado (Diciembre 2010). Se hará uso de la sintaxis proveída para el programa WinBUGS.

5.1. Modelo de Regresión Binaria

5.1.1. Descripción de los Datos

De acuerdo a los datos proveídos en la pagina web de la Superintendencia de Banca y Seguros y AFP (SBS) sobre la entidad microfinanciera (objeto de estudio), para los años 2009 al 2011, el tipo de crédito Consumo muestra un nivel de morosidad o probabilidad de default de 3.46 % (Consumo - 2009), 2.87 % (Consumo - 2010) y 2.54 % (Consumo - 2011); dichos ratios son elevados aun cuando la entidad viene evaluando y estimando su probabilidad de default crediticio vía modelos de regresion binarios pero con enlaces simetricos, modelo Logistico (<http://www.sbs.gob.pe/app/stats/EstadisticaBoletinEstadistico.asp/>).

Para el estudio aplicado se ha analizado y validado datos de una muestra de clientes de créditos correspondientes a una entidad Micro Financiera, pertenecientes a cartera total de prestatarios de 31,371 clientes; se considero una muestra final de 4745 clientes correspondiente al 15 % de dicha cartera total; la validacion de variables se ha realizado mediante el uso de softwares como el SPSS y Eview's. Para efectos del presente estudio se restringe a la validacion de 28 variables (Socio Demograficas, De Comportamiento, etc), de las cuales solo 08 señalaron significancia para el estudio.

Se considera en el trabajo si es favorable o no el otorgamiento de préstamos para la entidad, dado un conjunto de factores asociados que incluyen características demográficas, aspectos sociales, financieras, entre otros aspectos. La proporción observada de préstamos otorgados y en situación normal o vigente es de 95 %.

Entre las variables explicativas consideramos, Saldo de Crédito, Numero de Cuotas Aprobadas, Periodo de Gracia, Tasa Efectiva Mensual, Numero de Entidades con Deuda, Categoría de Riesgo en Centrales, Ubigeo - Departamento y Riesgo De Sobreendeudamiento (uso o no de tarjetas o líneas de crédito).

5.1.2. Formulación del Modelo General de Regresión Binaria

Para la implementación de la metodología propuesta en el estudio y a nivel general proponemos el modelo de regresión binaria, dado por:

$$p_i = P(\text{default} = 1) = F_{\theta}[m(\eta_i, \lambda)], \quad (5.1)$$

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i}, \quad i = 1, 2, \dots, 4745, \quad (5.2)$$

Donde:

Variables de Comportamiento:

- X6 : Categoría de Riesgo.
- X8 : Riesgo de Sobreendeudamiento.

Variables Financieras:

- X1 : Saldo de Crédito.
- X2 : Numero de Cuotas.
- X3 : Días de Gracia.
- X4 : Tasa Efectiva Mensual.
- X5 : Numero de Entidades.

Variables Socio Demográficas:

- X7 : Ubicación Geográfica.

Los modelos a ser comparados serían establecidos de acuerdo al aspecto teórico revisados en los capítulos anteriores; Modelo Simétrico (Probit) versus aquellos Modelos con Enlaces Asimétricos (Skew Probit o Asimetrizados); BBB Probito Asimetrizado, Estándar Probito Asimetrizado y CDS Probito Asimetrizado. Las sintaxis de los modelos así como las bases de datos para su uso posterior, fueron obtenidos mediante el programa BRMUW (Bayesian Regression Modelo Using WINBUGS), e implementados en el programa WinBUGS.

5.1.3. Resultados

Para implementar la metodología propuesta para la data descrita anteriormente proponemos el siguiente modelo general de regresión binaria:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots \beta_8 X_{8i}, \tag{5.3}$$

Donde: $i = 1, 2, 3, \dots, 4745$,

Para la comparación de modelos en la Inferencia Bayesiana se emplea el criterio Deviance Information Criteria (DIC) y la media de los desvíos a posteriori (Dbar) propuestos por Spiegelhalter et al. (2002), los cuales indican que el mejor modelo es aquel que presenta el menor DIC o Dbar.

Los análisis serían complementados por aquellos según gráficos de Curvas COR y numérico según Tablas de Contingencia o Tablas de Confusión.

ENLACES	MODELOS	Bur in	Dbar	DIC
Simétricos	Probit - Clas	4,000	2,778.00	2,785.00
Asimétricos	CDS Probit	4,000	1,308.06	2,075.21
	Estándar Prob	4,000	2,628.83	2,643.58
	BBB Probit	4,000	2,743.13	2,750.83

Cuadro 5.1: Comparación de Modelos para Modelo Aplicativo de Otorgamiento de Préstamos (N = 4745)

Los resultados obtenidos indican que todos los modelos asimétricos implementados presentan mejor desempeño según los criterios de comparación de modelos a diferencia del Modelo Simétrico Probit tradicional. El modelo con mejor desempeño es el CDS Probit Asimetrizado.

Coefficientes	Variables	ESTANDAR	CDS	BBB	PROBIT
Beta[0]	Constante	-2.486000	-5.34100	-2.58900	-1.82100
Beta[1]	Saldo de Capital	-0.029950	-0.05295	-0.03151	-0.02035
Beta[2]	N° Cuotas	0.0023560	0.005614	0.002888	-0.00569
Beta[3]	Periodo de Gracia	0.0034210	0.005317	0.003532	0.003343
Beta[4]	Tasa Efectiva Mes	0.3052000	0.500600	0.298400	0.197000
Beta[5]	Numero Entidades	-0.111100	-0.19970	-0.11300	-0.13320
Beta[6]	Categoría de Riesgo	0.5335000	0.830900	0.529400	0.542000
Beta[7]	Departamento	0.0115700	0.017970	0.011860	0.000104
Beta[8]	Riesgo Sobredeuda RDS	-0.227600	-0.38600	-0.23400	-0.22140
Lambda	Coefficiente Asimetria	0.2436000	1.778000	-0.17750	N.A.

Cuadro 5.2: Resumen de Aplicaciones Según Modelo Propuesto y Modelos Alternos

De los resultados mostrados en el cuadro anterior se puede inferir que incrementos en la Numero de Cuotas, Periodo de Gracia, Tasa efectiva, Categoría de Riesgo y Colocaciones en Zona Lima; son variables que incrementan la Probabilidad de Default Crediticio.

Los resultados también señalan que las variables, Saldo de Capital, Numero de Entidades y Uso de Lineas, son variables que reducen la probabilidad de default crediticio.

De acuerdo al análisis de datos realizado sobre la cartera de clientes de Créditos de la Microfinanciera a un periodo determinado, se pudo apreciar que el modelo Probit (simétrico) presenta un 56.6% de mala clasificación (clientes malos pronosticados por el modelo como buenos clientes), mientras que el modelo Skew Probit BBB genera confusión en solo 43.5% de dicho tipo de clientes.

5.1.4. Resumen de Análisis Gráfico Proporcionados por WinBUGS

El análisis se realiza en función a las tendencias mostradas según los gráficos obtenidos mediante el programa WinBUGS y aplicados a la data objeto de estudio (Ver Apéndice A):

- Modelo BBB - Historia de la Cadena:

Los gráficos muestran una adecuada convergencia al valor estimado para cada Beta(s), el número de iteraciones (4000) para cada estimación es adecuado y suficiente para aseguramiento de la convergencia.

- Modelo BBB - Densidades:

Los gráficos sobre densidades muestran una correspondencia hacia una distribución normal para todas las estimaciones (Betas), el número de iteraciones realizadas se suponen adecuadas y se muestra para cada gráfico la convergencia al valor medio estimado (Betas) a medida que se incrementa el número de iteraciones.

- Modelo BBB - Quantiles:

Los gráficos sobre quantiles muestran para el caso de las Betas una estabilidad importante, pues se observa que las bandas (intervalos de credibilidad) marchan paralelas a la línea de la mediana.

- Modelo BBB - Autocorrelacion

Los Correlogramas muestran que la función de auto correlación decae en forma exponencial o converge hacia cero, lo cual nos indica que las estimaciones son estables (estacionarios).

Los gráficos muestran de acuerdo al análisis de auto correlación que la convergencia hacia el valor medio estimado no muestran un retraso significativo (0 - 20) para la Betas.

- Modelo Estándar - Historia de la Cadena:

Se muestra una adecuada convergencia al valor estimado para cada Beta (s), el número de iteraciones (4000) propuestos para cada estimación es adecuado y suficiente para aseguramiento de la convergencia.

- Modelo Estándar - Densidades:

Los gráficos muestran una correspondencia hacia una distribución normal para todas las estimaciones (Betas), el número de iteraciones realizadas se suponen adecuadas y se muestra en cada gráfico la convergencia al valor medio estimado (Betas) a medida que se incrementa el número de iteraciones.

- Modelo Estándar - Quantiles:

Los gráficos muestran estabilidad importante para las Betas, pues se observa que las bandas (intervalos de credibilidad) marchan paralelas a la línea de la mediana en todos los casos.

- Modelo Estándar - Autocorrelacion:

Los Correlogramas muestran que la función de auto correlación decae en forma exponencial o converge hacia cero para todos los casos, lo cual indica que las estimaciones son estables (estacionarios).

Los gráficos muestran de acuerdo al análisis de auto correlación que la convergencia hacia el valor medio estimado no muestran retrasos significativos (0 - 20) para la Betas.

- Modelo CDS - Historia de la Cadena:

Se muestra una adecuada convergencia al valor estimado para cada Beta (s), el número de iteraciones (4000) propuestos para cada estimación es adecuado y suficiente para aseguramiento de la convergencia. Para el caso de beta 7 (Categoría de Riesgo) se muestra una convergencia adecuada a partir de las 3500 iteraciones.

- Modelo CDS - Densidades:

Los gráficos muestran una correspondencia hacia una distribución normal para todas las estimaciones (Betas), el número de iteraciones realizadas se suponen adecuadas y se muestra en cada gráfico la convergencia al valor medio estimado (Betas) a medida que se incrementa el número de iteraciones. Para el caso de beta 7 no se muestra correspondencia o normalidad.

- Modelo CDS - Quantiles:

Los gráficos muestran una estabilidad importante para las Betas, se observa que las bandas (intervalos de credibilidad) marchan paralelas a la línea de la mediana en todos los casos; salvo para el caso de beta 7.

- Modelo CDS - Autocorrelacion:

Los Correlogramas muestran que la función de auto correlación decae en forma exponencial o converge hacia cero para todos los casos, lo cual indica que las estimaciones son estables (estacionarios).

Los gráficos muestran de acuerdo al análisis de auto correlación que la convergencia hacia el valor medio estimado no muestran retrasos significativos (0 - 20) para la totalidad de valores promedios estimados (Betas), excepto para el valor de Beta 2 y Beta 6 (Saldo de Capital y Numero de Entidades).

- Modelo Probit - Historia de la Cadena:

Los gráficos muestran adecuadas convergencias al valor estimado para cada Beta (s) solo para algunas variables; para las variables Beta 3, 5, 7 y 8 (Numero de Cuotas, Tasa Efectiva, Categoría de Riesgo y Departamento), el número de iteraciones (4000) propuestos para cada estimación no es suficiente para aseguramiento de la convergencia.

■ Modelo Probit - Densidades:

Los gráficos no muestran correspondencias hacia normalidad para las estimaciones (Betas) y el número de iteraciones realizadas se suponen inadecuadas, y se muestra en cada gráfico la no convergencia al valor medio estimado (Betas) a medida que se incrementa el número de iteraciones.

■ Modelo Probit - Quantiles:

Los gráficos muestran una estabilidad relativa para todas las Betas, se observa que las bandas (intervalos de credibilidad) marchan paralelas a la línea de la mediana en todos los casos, pero la relación es cambiante.

■ Modelo Probit - Autocorrelacion:

Los Correlogramas muestran que la función de auto correlación decae en forma exponencial o converge hacia cero solo para las variables Beta 4, 6, 7 y 9; en los demás casos las estimaciones muestran que no son estables (estacionarios).

Los gráficos muestran de acuerdo al análisis de auto correlación que la convergencia hacia el valor medio estimado muestran retrasos muy significativos (mayor a 100), para las variables Numero de Cuotas, Tasa Efectiva Mensual y Ubicación Geográfica.

Coefficientes	Variabes	Media	Varianza	p2.5	p97.5	Mediana
Beta[0]	Constante	-2.589000	0.056692	-3.041000	-2.113000	-2.592000
Beta[1]	Saldo de Capital	-0.031510	0.000448	-0.074540	0.009413	-0.031490
Beta[2]	N° Cuotas	0.002888	0.000048	-0.010760	0.016240	0.003023
Beta[3]	Periodo de Gracia	0.003532	0.000004	-0.000557	0.007680	0.003539
Beta[4]	Tasa Efectiva Mes	0.298400	0.002220	0.208900	0.391100	0.298100
Beta[5]	Numero Entidades	-0.111300	0.000888	-0.171000	-0.054550	-0.110500
Beta[6]	Categoría de Riesgo	0.529400	0.000491	0.485700	0.571200	0.530000
Beta[7]	Departamento	0.011860	0.000055	-0.002808	0.026150	0.011880
Beta[8]	Riesgo Sobredeuda RDS	-0.234000	0.006929	-0.396700	-0.066450	-0.235600
Lambda	Coefficiente Asimetria	-0.177500	0.031684	-0.511600	0.146300	-0.171300

Cuadro 5.3: Estadísticas A Posteriori Modelo Skew Probit - BBB

Coefficientes	Variabes	Media	Varianza	p2.5	p97.5	Mediana
Beta[0]	Constante	-2.486000	0.049151	-2.967000	-2.076000	-2.519000
Beta[1]	Saldo de Capital	-0.029950	0.000380	-0.069770	0.005372	-0.029670
Beta[2]	N° Cuotas	0.002356	0.000049	-0.010780	0.016520	0.002931
Beta[3]	Periodo de Gracia	0.003421	0.000005	-0.000690	0.007780	0.003466
Beta[4]	Tasa Efectiva Mes	0.305200	0.002147	0.215500	0.402500	0.308600
Beta[5]	Numero Entidades	-0.111000	0.000944	-0.170600	-0.051690	-0.111000
Beta[6]	Categoría de Riesgo	0.533500	0.000373	0.496400	0.571800	0.534900
Beta[7]	Departamento	0.011570	0.000055	-0.001904	0.027350	0.012140
Beta[8]	Riesgo Sobredeuda RDS	-0.227600	0.007914	-0.403500	-0.055810	-0.230300
Lambda	Coefficiente Asimetria	0.243600	0.045796	-0.099520	0.731000	0.378400

Cuadro 5.4: Estadísticas A Posteriori Modelos Skew Probit - Estándar

Para el caso de los modelos de mayor relevancia (SP - BBB y SP Estándar) las estadísticas tanto de Varianza así como de Percentiles muestran valores estables, es decir, significancia

Coefficientes	VARIABLES	Media	Varianza	p2.5	p97.5	Mediana
Beta[0]	Constante	-5.341000	0.352124	-4.735000	-2.572000	-3.533000
Beta[1]	Saldo de Capital	-0.052950	0.000611	-0.086250	0.010110	-0.036560
Beta[2]	N° Cuotas	0.005614	0.000077	-0.013390	0.021570	0.003471
Beta[3]	Periodo de Gracia	0.005317	0.000007	-0.001013	0.009033	0.003890
Beta[4]	Tasa Efectiva Mes	0.500600	0.004118	0.241400	0.498300	0.353800
Beta[5]	Numero Entidades	-0.199700	0.001554	-0.222700	-0.069820	-0.136700
Beta[6]	Categoría de Riesgo	0.830900	0.003319	0.524600	0.729200	0.610600
Beta[7]	Departamento	0.017970	0.000076	-0.001776	0.032450	0.014420
Beta[8]	Riesgo Sobredeuda RDS	-0.386000	0.010445	-0.468800	-0.069160	-0.262500
Lambda	Coefficiente Asimetria	1.778000	0.109760	0.267800	1.351000	0.862500

Cuadro 5.5: Estadísticas A Posteriori Modelo Skew Probit - CDS

Coefficientes	VARIABLES	Media	Varianza	p2.5	p97.5	Mediana
Beta[0]	Constante	-1.821000	0.242458	-2.329000	-0.581400	-2.046000
Beta[1]	Saldo de Capital	-0.020350	0.000573	-0.065680	0.030710	-0.021740
Beta[2]	N° Cuotas	-0.005696	0.000109	-0.030390	0.011220	-0.004044
Beta[3]	Periodo de Gracia	0.003343	0.000005	-0.000523	0.008146	0.003281
Beta[4]	Tasa Efectiva Mes	0.197000	0.005845	0.005996	0.287700	0.234000
Beta[5]	Numero Entidades	-0.133200	0.000898	-0.195000	-0.075520	-0.133600
Beta[6]	Categoría de Riesgo	0.542000	0.000992	0.501800	0.587800	0.542500
Beta[7]	Departamento	0.000104	0.000117	-0.024100	0.016120	0.000838
Beta[8]	Riesgo Sobredeuda RDS	-0.221400	0.007439	-0.385000	-0.051930	-0.218200
Lambda	Coefficiente Asimetria	NA	NA	NA	NA	NA

Cuadro 5.6: Estadísticas A Posteriori Modelo Probit

dentro del análisis de consistencia y relevancia de ambos modelos.

En el caso del modelo CDS las estadísticas de Varianza así como de Percentiles muestran también valores estables, es decir, significancia dentro del análisis de consistencia y relevancia del modelo; dicho análisis se complementa con aquel de Cuantiles.

En el modelo Probit las estadísticas a posteriori muestran estabilidad y consistencia respecto a sus valores de Media, Varianza y Percentiles pero ello no se condice con aquel de cuantiles descrito anteriormente.

A manera de resumen presentamos la correspondiente ecuación del Mejor Modelo (Skew Probit BBB): $\eta = -2,59 + 0,30Tasa - 0,11N^{\circ}Entidades + 0,53CategoríaRiesgo - 0,23RDS$.

5.2. Análisis Complementarios

5.2.1. Análisis según Curvas COR (Gráfico)

En el presente capítulo se realiza un análisis gráfico via Curvas COR y se realiza además un análisis numérico según lo estimado via Tablas de Contingencia o Tablas de Confusión, la muestra de datos es aquella utilizada para llevar a cabo el análisis de aplicación y consistente en una muestra de 4745 clientes extraídos de una base de datos de 31371 clientes, pertenecientes a una cartera de préstamos de clientes tipo Consumo (Créditos Personales Directos)

de una entidad micro financiera.

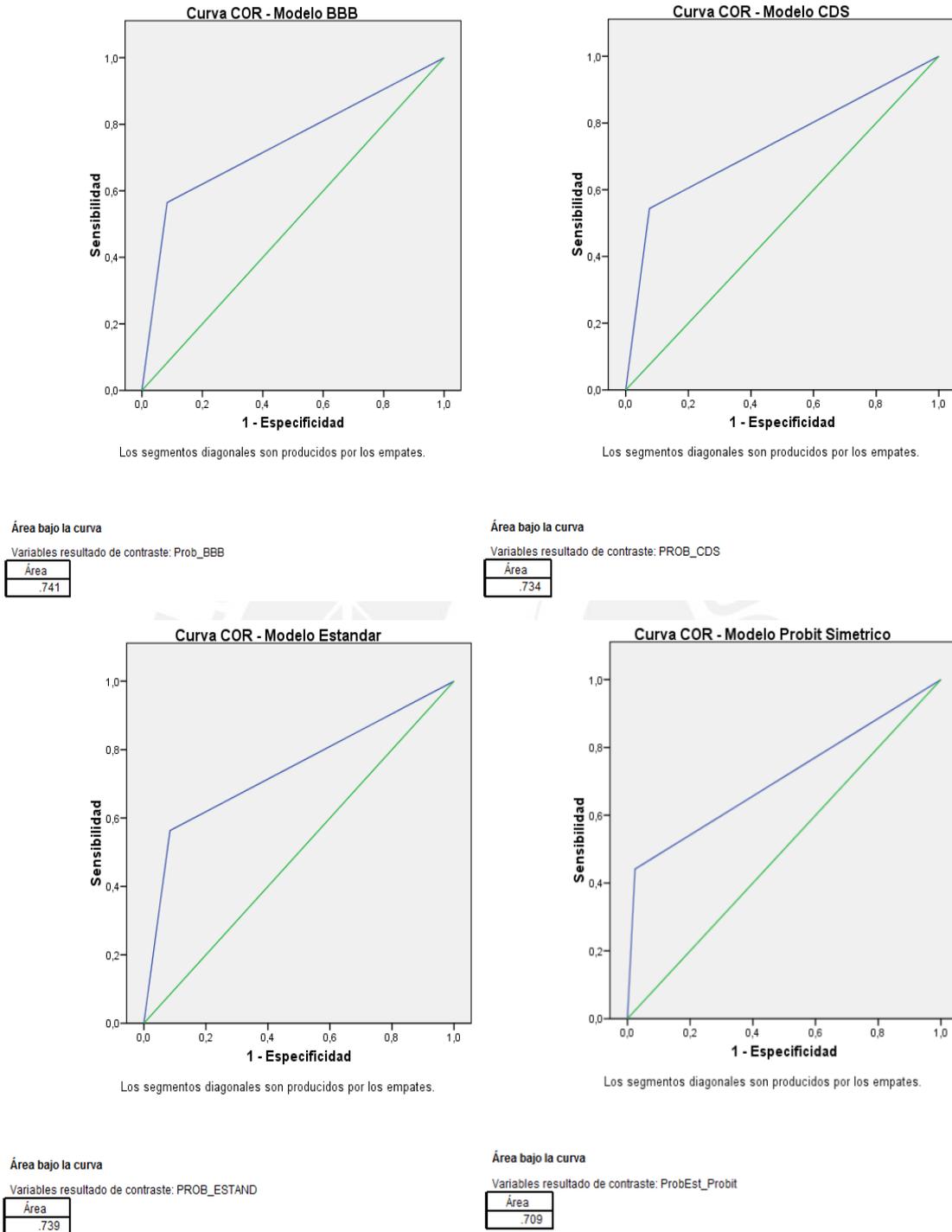


Figura 5.1: Curvas COR según Modelos Asimétricos o Modelo Simétrico

- Para el caso del Modelo Probit Simétrico y de acuerdo al análisis gráfico de Curvas COR o punto de corte se determina un punto de inflexión específico, el análisis muestra que el área debajo de la curva es aproximadamente a 0.709 respecto a 1.

- En el caso del Modelo Estándar Asimétrico y en función al análisis gráfico de Curvas COR o punto de corte se determina un punto de inflexión específico, el análisis muestra que el área debajo de la curva es aproximadamente a 0.739 respecto a 1.
- Para el Modelo CDS Asimétrico y en función al análisis gráfico de Curvas COR o punto de corte se determina un punto de inflexión específico, el análisis muestra que el área debajo de la curva es aproximadamente a 0.734 respecto a 1.
- En el caso del Modelo BBB Asimétrico y en función al análisis gráfico de Curvas COR o punto de corte se determina un punto de inflexión específico, el análisis muestra que el área debajo de la curva es aproximadamente a 0.741 respecto a 1.
- El análisis de sensibilidad (probabilidad de detección de casos de Default - 1 cuando realmente es Default - 1) versus el análisis de especificidad (probabilidad de detección de casos de No Default - 0 cuando realmente es No Default - 0), nos muestra que el área debajo de las curvas superan el 0.74 respecto a 1; por tanto la eficiencia de los modelos es significativa, principalmente para los modelos SPG tipo BBB, Estándar y CDS.
- En resumen el Modelo Skew Probit BBB obtuvo el mejor resultado, especialmente respecto al análisis de especificidad.

5.2.2. Análisis Según Tablas de Contingencia (numérico)

- Aquella tabla de contingencia para el Modelo Probit Simétrico nos muestra que el modelo discrimina correctamente para aquellos casos de No Default, pero existen 371 casos de un total de 664 considerados como Default que vienen siendo considerados como No Default.
- La tabla de contingencia para el Modelo Estándar Asimétrico nos muestra que el modelo discrimina correctamente para aquellos casos de No Default, pero existen 290 casos de un total de 664 considerados como Default siendo los mismos No Default.
- Aquella tabla de contingencia para el Modelo CDS Asimétrico nos muestra que el modelo discrimina correctamente para aquellos casos de No Default, pero existen 303 casos de un total de 664 considerados como Default siendo los mismos No Default.
- Aquella tabla de contingencia para el Modelo BBB Asimétrico nos muestra que el modelo discrimina convenientemente para aquellos casos de No Default, pero existen 289 casos de un total de 664 considerados como Default siendo los mismos No Default.
- Los modelos que muestran una mejor distribución de casos (Observados versus Pronosticados, diagonal), son los Modelos BBB y Estándar, las tablas correspondientes a dichos modelos muestran un detalle en la diagonal principal (0, 1) o (1, 0) cercanos a cero, lo cual demuestra que la eficiencia de ambos es significativa.
- Específicamente el modelo Skew Probit BBB muestra el mejor análisis discriminante, menor error de clasificación de datos pronosticados versus datos reales.

	Probabilidad BBB 0.00	Probabilidad BBB 1.00	Total
Y-Obs 0.00 Recuento	3,741	340	4,081
% del Total	78.80 %	7.20 %	86.00 %
Y-Obs 1.00 Recuento	289	375	664
% del Total	6.10 %	7.90 %	14.00 %
Total Recuento	4,030	715	4,745
% del total	84.90 %	15.10 %	100.00 %

Cuadro 5.7: Tabla de Contingencia Modelo Asimétrico Skew Probit BBB

	Probabilidad CDS 0.00	Probabilidad CDS 1.00	Total
Y-Obs 0.00 Recuento	3,772	309	4,081
% del Total	79.85 %	6.50 %	86.00 %
Y-Obs 1.00 Recuento	303	361	664
% del Total	6.40 %	7.60 %	14.00 %
Total Recuento	4,075	670	4,745
% del total	85.90 %	14.10 %	100.00 %

Cuadro 5.8: Tabla de Contingencia Modelo Asimétrico Skew Probit CDS

	Probabilidad Estándar 0.00	Probabilidad Estándar 1.00	Total
Y-Obs 0.00 Recuento	3,737	344	4,081
% del Total	78.80 %	7.20 %	86.00 %
Y-Obs 1.00 Recuento	290	374	664
% del Total	6.10 %	7.90 %	14.00 %
Total Recuento	4,027	718	4,745
% del total	84.90 %	15.10 %	100.00 %

Cuadro 5.9: Tabla de Contingencia Modelo Asimétrico Skew Probit Estándar

	Probabilidad Probit 0.00	Probabilidad Probit 1.00	Total
Y-Obs 0.00 Recuento	3,982	99	4,081
% del Total	83.90 %	2.10 %	86.00 %
Y-Obs 1.00 Recuento	371	293	664
% del Total	7.80 %	6.20 %	14.00 %
Total Recuento	4,353	392	4,745
% del total	91.70 %	8.30 %	100.00 %

Cuadro 5.10: Tabla de Contingencia Modelo Simétrico Probit

5.3. Estudio de Verificación

5.3.1. Diseño

En la presente sección se lleva a cabo un análisis de regresión binaria con enlace asimétrico vía simulación MCMC, pero bajo un escenario distinto de las condiciones realizadas, evaluadas y propuestas inicialmente.

Las condiciones propuestas para llevar a cabo las verificaciones de los resultados obtenidos en la sección 5.2 son:

- Se utilizara la muestra original de datos (Diciembre 2010), sin embargo el tamaño de muestra sera variable; 4547, 3157, 1575 y 931 respectivamente.
- El número de iteraciones para cada tamaño de muestra serían distintas; 4000, 2000, 1500 y 1500 respectivamente.
- Los modelos sometidos a evaluación bajo las nuevas condiciones son: 1. Probit, 2. BBB Skew Probit, 3. CDS Skew Probit y 4. Modelo Estándar Skew Probit.
- El análisis sobre los modelos se realizará en función a los criterios DIC (Criterion Deviance Information) y el Dbar (Media de los Desvíos a Posteriori); los resultados serán complementados con aquel gráfico obtenido mediante el software WinBUGS.
- Como se hizo anteriormente se utilizara el software WinBugs y el BRMUW para la generación de la sintaxis de los modelos.

5.3.2. Resultados

La muestra supone la existencia de 4547 clientes, dicha muestra ha sido extraída de una población de 31371 clientes (Créditos de Consumo y Personales), siendo aproximadamente el 10 % de Clientes en situación de default y el 90 % de clientes buenos pagadores. Los resultados de las verificaciones y estimación llevados a cabo se muestran en los cuadros 5.11 al 5.14.

Análisis de Modelo Propuesto (Ideal, Data No Homogénea con Enlace Asimétrico)

- Para todos los casos aquellos modelos con enlaces asimétricos presentan mejor desempeño en los criterios de comparación de modelos versus aquel modelo con enlace simétrico tradicional.
- De acuerdo al análisis de verificación realizado (Tamaño de Muestra y Número de Iteraciones menor a 4000), el modelo CDS presenta el mejor desempeño a medida que el tamaño de muestra se reduce. Sin embargo en el análisis según Curvas COR y Tablas de Contingencia se señala como mejor modelo a Estándar Probit.
- Respecto al análisis gráfico y de consistencia el modelo Estándar es aquel de mejor desempeño cuando la muestra es representativa y el número de iteraciones es igual o superior a 4000.

- Se debe indicar que para aquella muestra de 931 clientes y número de iteraciones de 1000, la totalidad de modelos muestran problemas de convergencia.
- El análisis muestra (gráficos 5.11 al 5.14), según el criterio DIC y Dbar, que el mejor modelo es el Skew Probit CDS, aun cuando las muestras son relativamente mas pequeñas.

ENLACES	MODELOS	Bur in	Dbar	DIC
Simétricos	Probit - Clas	4,000	2,778.00	2,785.00
Asimétricos	CDS Probito	4,000	1,308.06	2,075.21
	Estándar Prob	4,000	2,628.83	2,643.58
	BBB Probito	4,000	2,743.13	2,750.83

Cuadro 5.11: Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 4745)

ENLACES	MODELOS	Bur in	Dbar	DIC
Simétricos	Probit - Clas	2,000	1,932.00	1,938.00
Asimétricos	CDS Probito	2,000	1,268.97	1,715.70
	Estándar Prob	2,000	1,345.98	1,753.17
	BBB Probito	2,000	1,909.12	1,932.68

Cuadro 5.12: Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 3157)

ENLACES	MODELOS	Bur in	Dbar	DIC
Simétricos	Probit - Clas	1,500	916.0000	922.9000
Asimétricos	CDS Probito	1,500	872.0000	843.3000
	Estándar Prob	1,500	599.3000	812.4500
	BBB Probito	1,500	904.3700	911.8600

Cuadro 5.13: Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 1575)

ENLACES	MODELOS	Bur in	Dbar	DIC
Simétricos	Probit - Clas	1,500	531.7000	539.0000
Asimétricos	CDS Probito	1,500	474.9600	433.5400
	Estándar Prob	1,500	364.5500	507.2000
	BBB Probito	1,500	442.0200	490.8900

Cuadro 5.14: Comparación de Modelos Según Tamaño de Muestra y Número de Iteraciones (N = 931)

Capítulo 6

Conclusiones y Recomendaciones

6.1. Conclusiones

- De acuerdo al análisis desarrollado se infiere que los modelos con enlaces asimétricos muestran mejor desempeño que aquel comparativo modelo simétrico tradicional (análisis de criterios), para los datos de default.
- El análisis de verificación ratifica lo obtenido en el análisis de aplicación, los modelos con enlaces asimétricos muestran mejor desempeño que aquel modelo simétrico tradicional; especialmente cuando las muestras son grandes (iteraciones = 4000). Los valores de los criterios DIC, Dbar, Curvas COR y Tablas de Contingencia fundamentan lo señalado.
- Los modelos Skew Probit BBB y Estándar son los de mejor desempeño, se prefiere el Modelo BBB dado los resultados de la matriz de contingencia así como el análisis de sensibilidad y especificidad de ambos modelos. Cabe indicar que el modelo BBB bajo una muestra mayor no señala una toma de mayor tiempo al esperado para el cálculo de resultados computacionalmente, situación que si sucede para el caso de los modelos CDS y Estándar.
- Se concluye finalmente que los modelos con Enlaces Asimétricos presentan mejores indicadores de ajuste e información así como mejoran la sensibilidad y especificidad de un determinado modelo (regresión binaria); versus aquellos modelos simétricos tradicionales.
- El mejor modelo obtenido indica que las variables, Numero de Cuotas, Periodo de Gracia, Tasa Efectiva Mensual, Categoría de Riesgo y Departamento Geográfico, son significativas y positivas, mientras que las variables, Saldo de Capital, Numero de Entidades que Adeuda el Cliente y Riesgo de Sobreendeudamiento, son significativas y negativas. La ecuación del mejor modelo (Skew Probit BBB) se halla determinada por:
$$\eta = -2,59 + 0,30Tasa - 0,11N^{\circ}Entidades + 0,53CategoríaRiesgo - 0,23RDS.$$
- Se debe resaltar que el mejor modelo Skew Probit, es el único que no presenta retrasos significativos (computacionalmente) en el cálculo mediante el uso del WinBUGS u OpenBugs cuando las muestras son mayores a 5 mil clientes.

- La implementación Bayesiana de estos modelos es relativamente fácil usando WinBUGS, sin embargo se tiene inconvenientes cuando la muestra y numero de variables son grandes.



6.2. Recomendaciones

- Evaluar el mejor modelo usando las variables que resulten significativas.
- Realizar estudios de validación del mejor modelo incluyendo análisis de validación cruzada (usándose diferentes tamaños de muestras).
- Desarrollar el mejor modelo (ideal) dentro de la entidad objeto de estudio, para el cálculo de probabilidades como herramienta o aplicativo para el otorgamiento de prestamos. Lo anterior supondría la validación del modelo mediante la aprobación de prestamos en forma práctica.
- Sistematización de la propuesta del mejor modelo asimétrico (Skew Probit BBB o Estándar) en la entidad objeto de estudio, cuya finalidad sera de mejorar los procesos y toma de decisión en la gestión de riesgos crediticios principalmente.
- Complementar el análisis teórico y práctico de los demás modelos dentro de la gama de productos y servicios de la entidad; Cartera Hipotecaria, Cartera Mayorista, etc; verificándose y estimándose su aplicación y desarrollo para dichos productos crediticios.

Apéndice A

Gráficos Correspondientes a Numeral 5.1.4

A.1. Modelo Probit Simétrico - Muestra N=4745

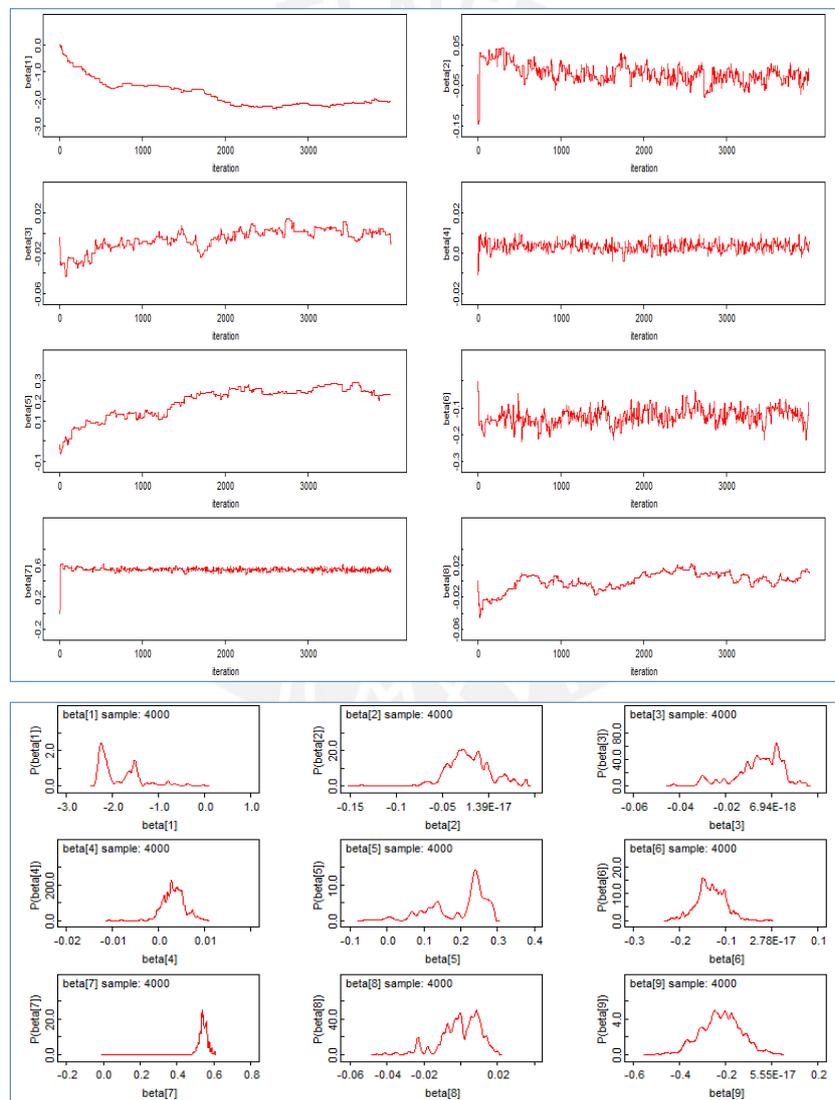


Figura A.1: Gráficos de Historia de la Cadena y Densidades Modelo Probit Simétrico

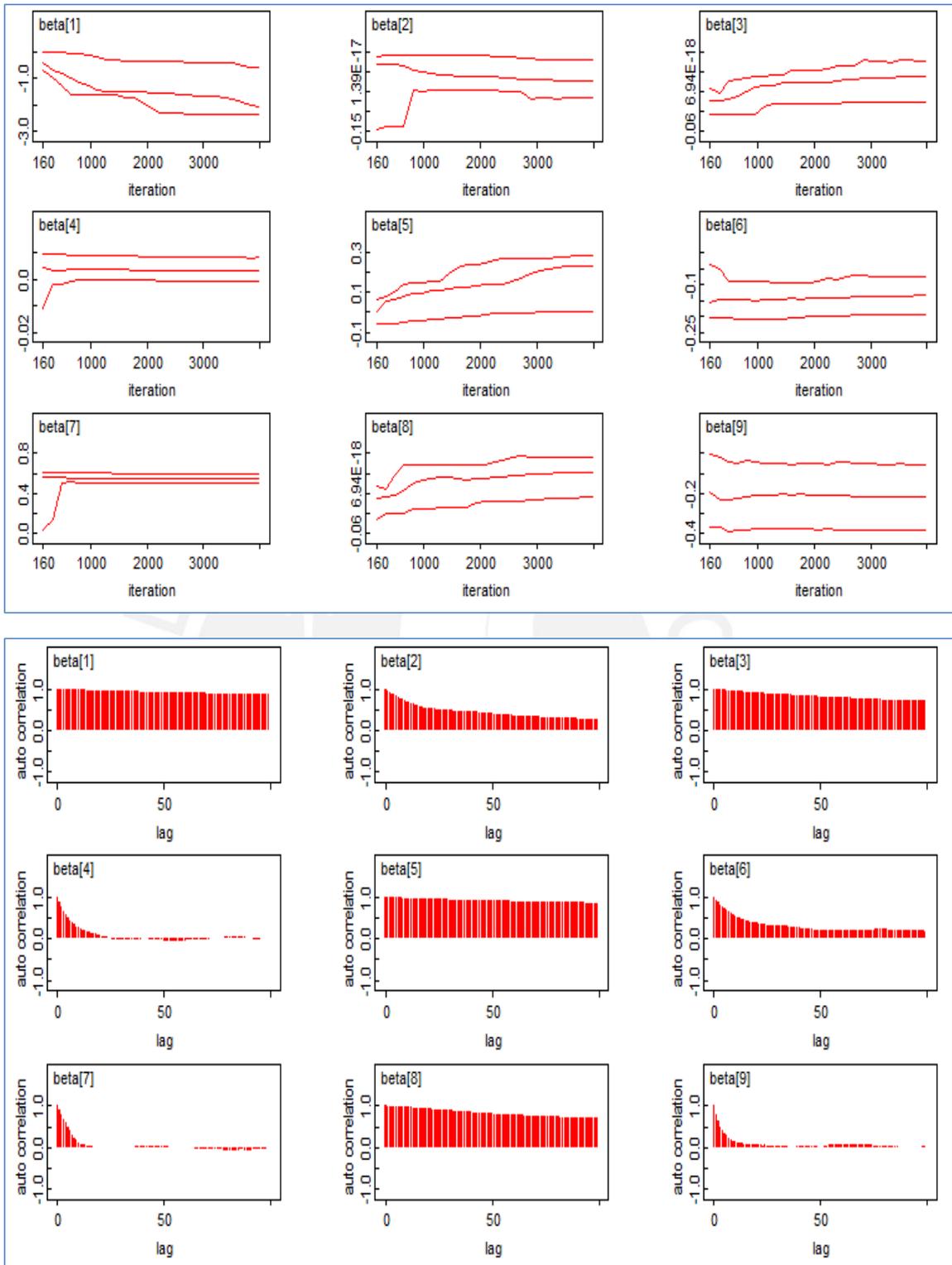


Figura A.2: Gráficos de Cuantiles y Autocorrelacion Modelo Probit Simétrico

A.2. Modelo Skew Probit Estándar Asimétrico - Muestra N=4745

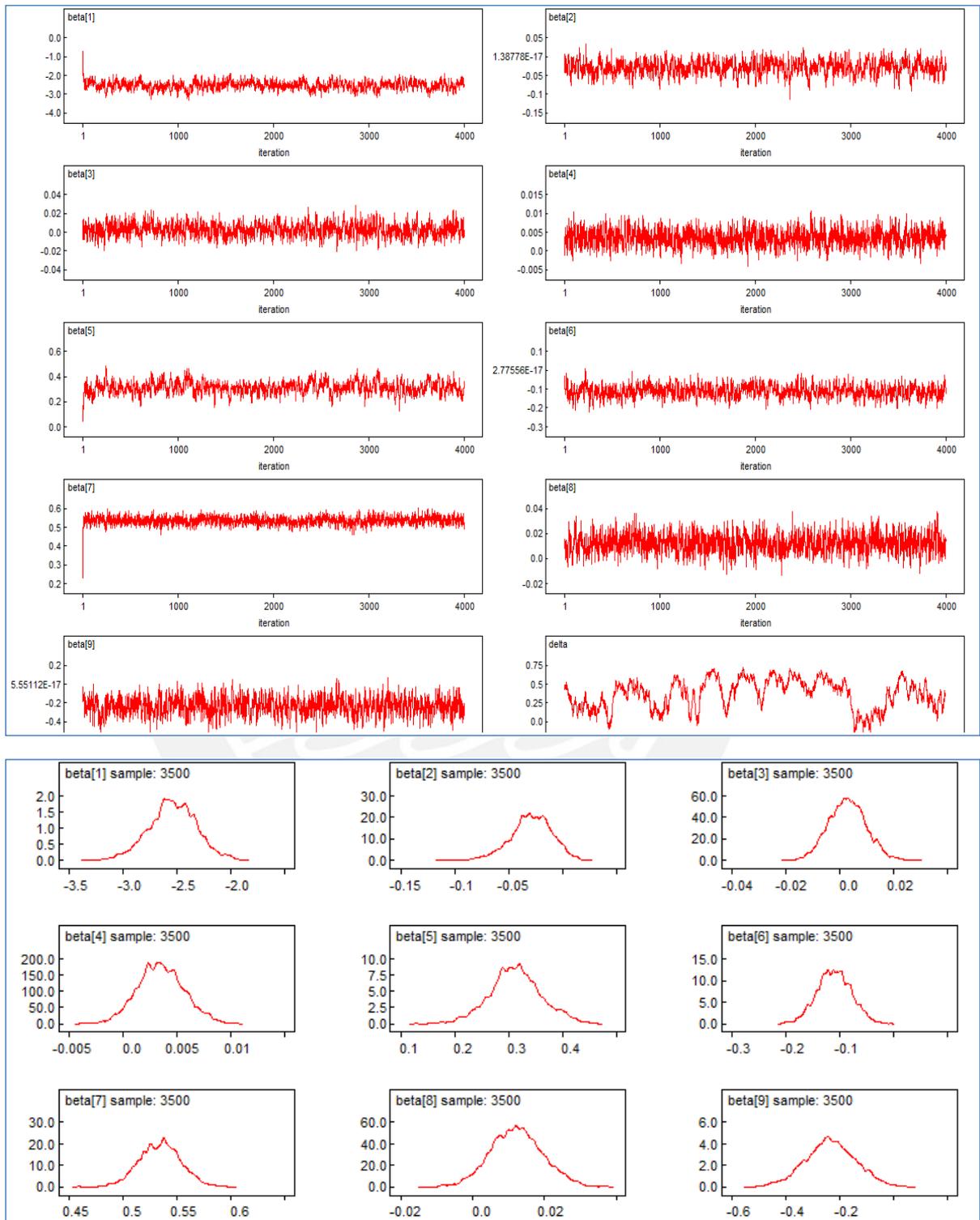


Figura A.3: Gráficos de Historia de la Cadena y Densidades Modelo Estándar

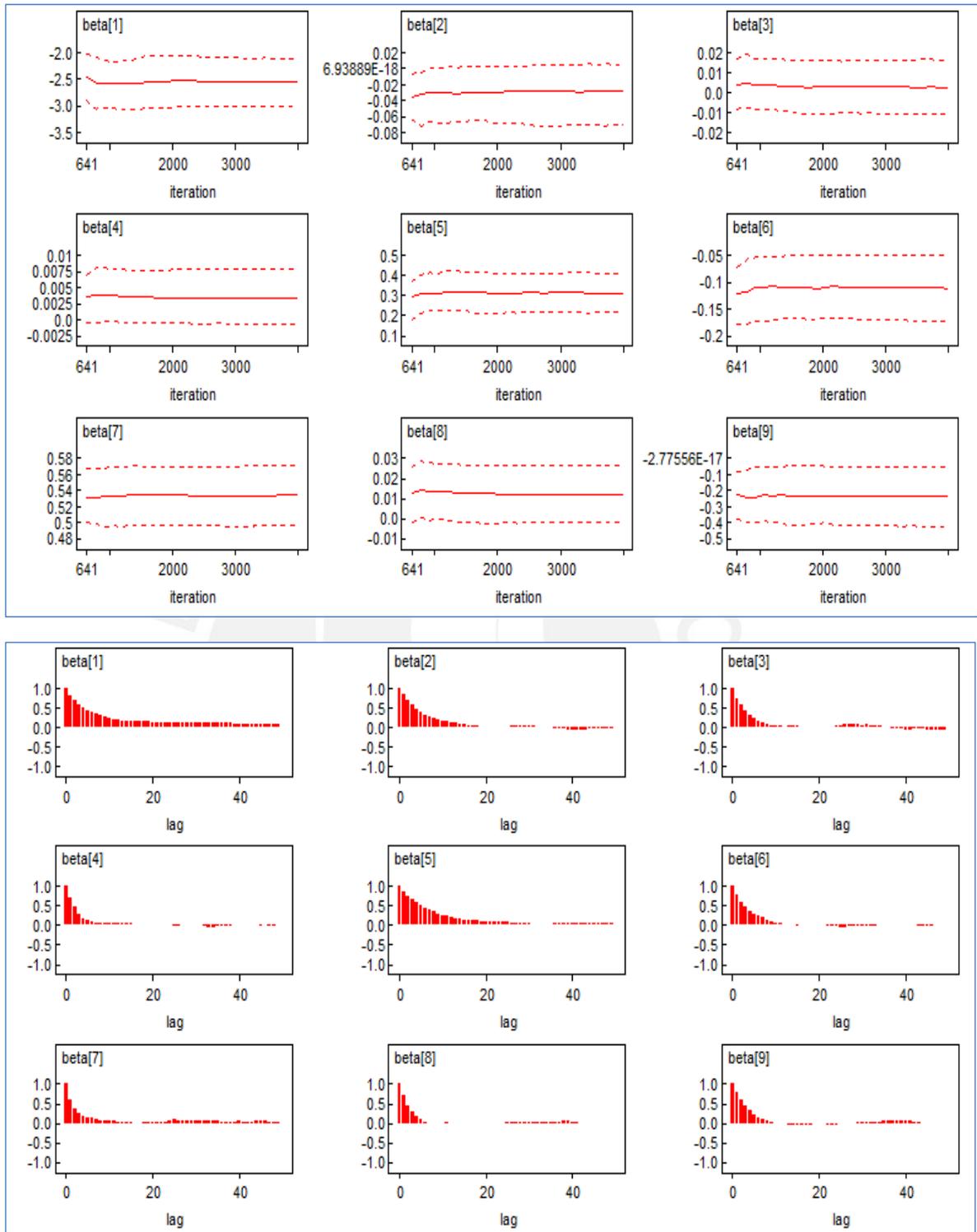


Figura A.4: Gráficos de Cuantiles y Autocorrelacion Modelo Estándar

A.3. Modelo Skew Probit CDS Asimétrico - Muestra N=4745

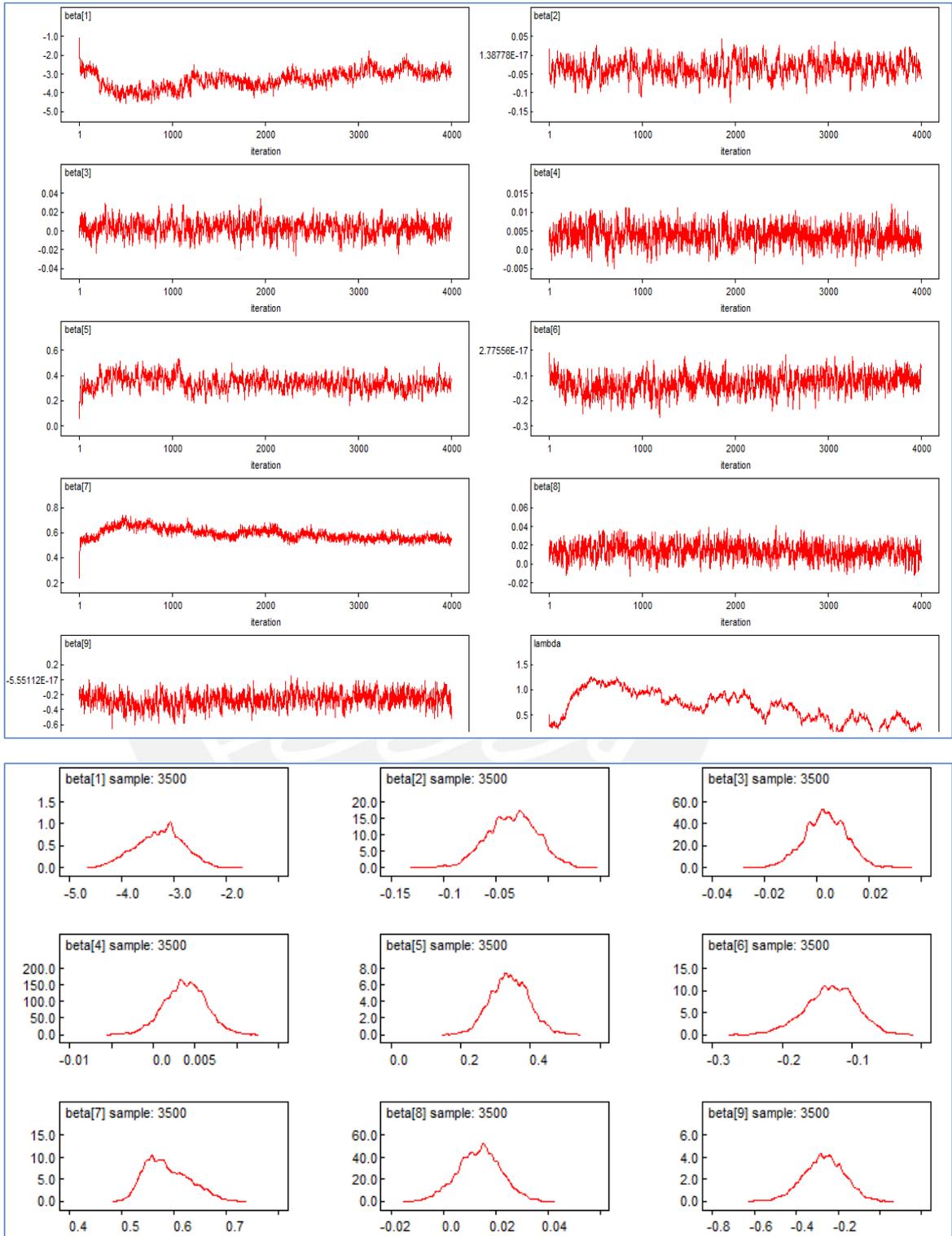


Figura A.5: Gráficos de Historia de la Cadena y Densidades Modelo CDS

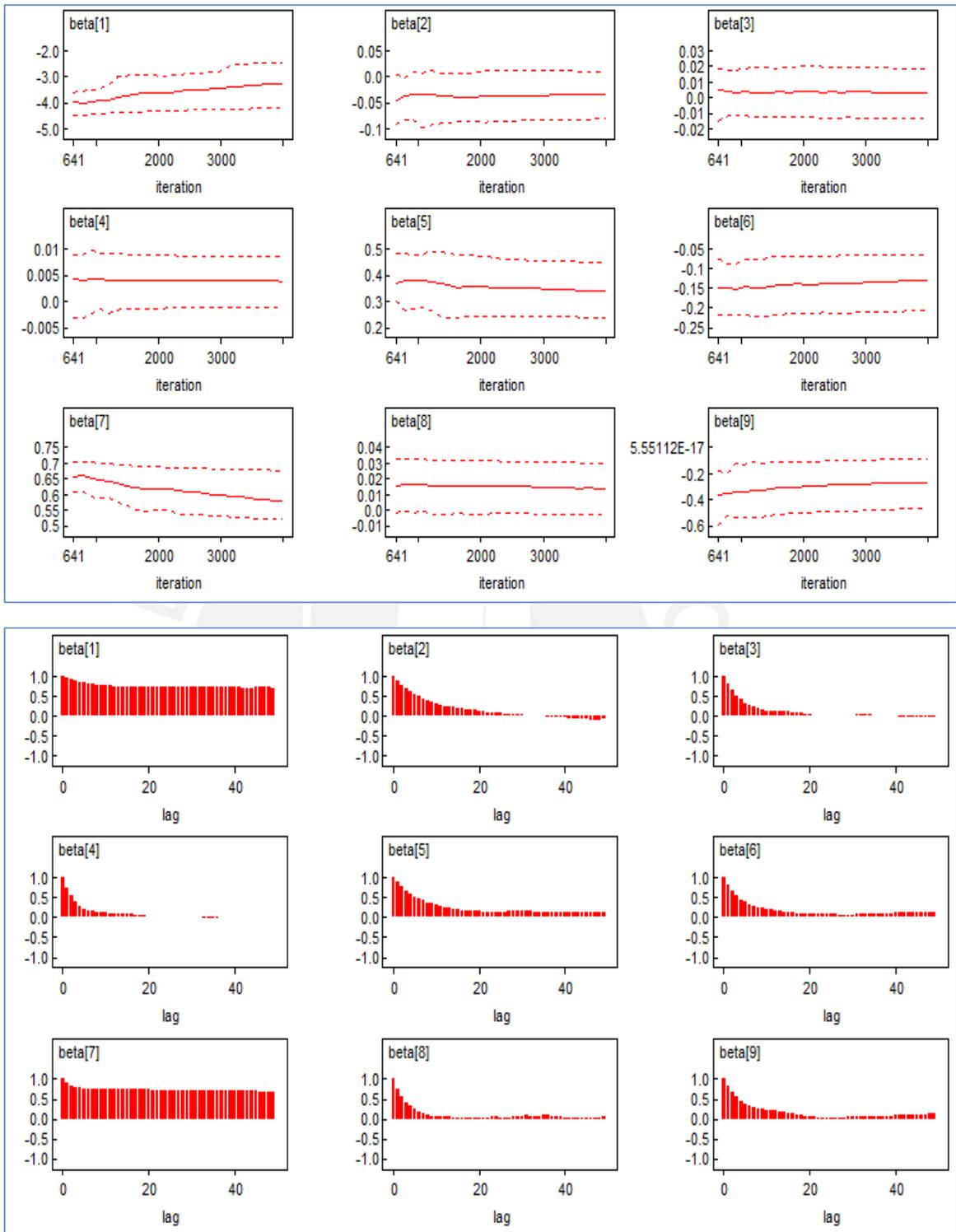


Figura A.6: Gráficos de Cuantiles y Autocorrelacion Modelo CDS

A.4. Modelo Skew Probit BBB Asimétrico - Muestra N=4745

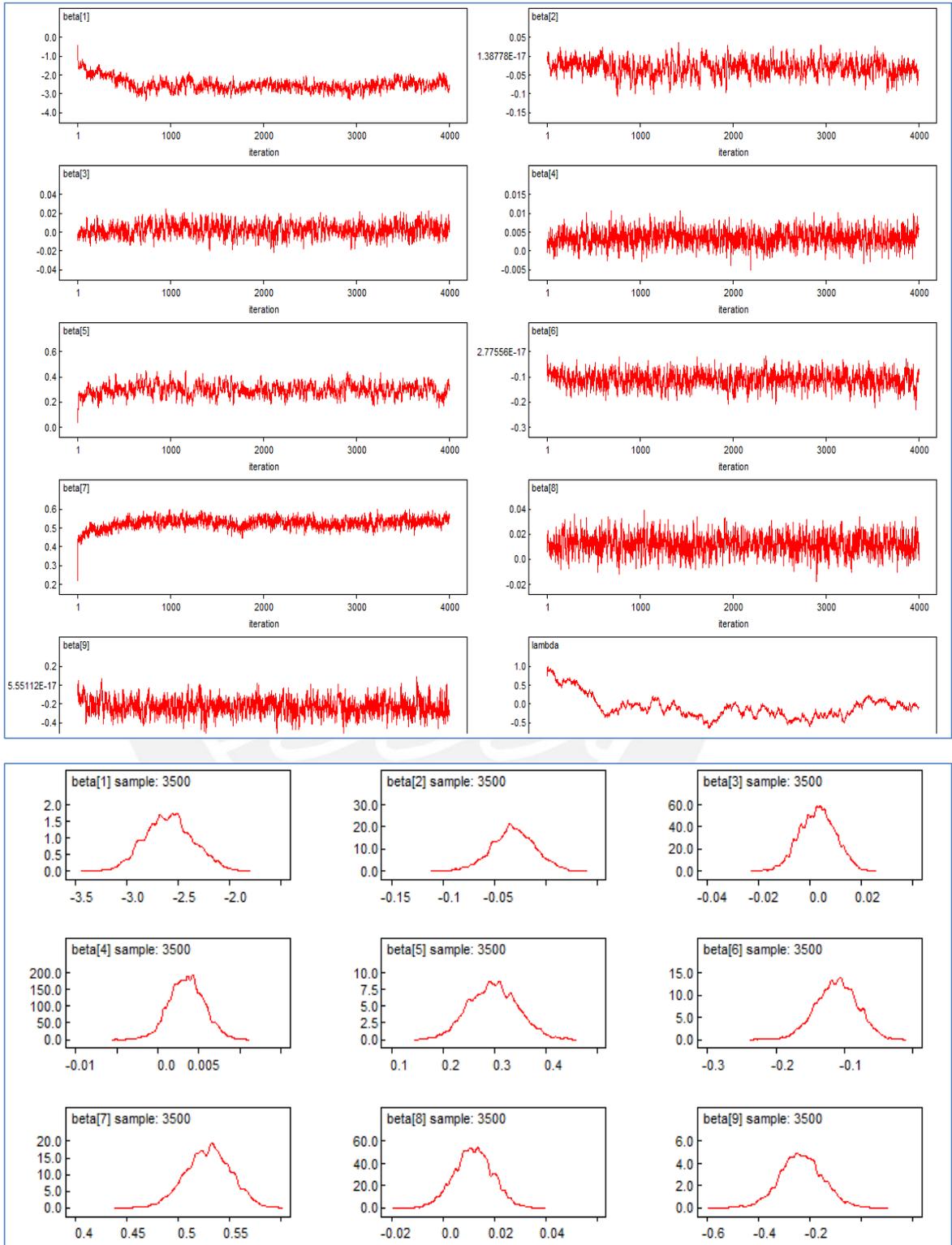


Figura A.7: Gráficos de Historia de la Cadena y Densidades Modelo BBB

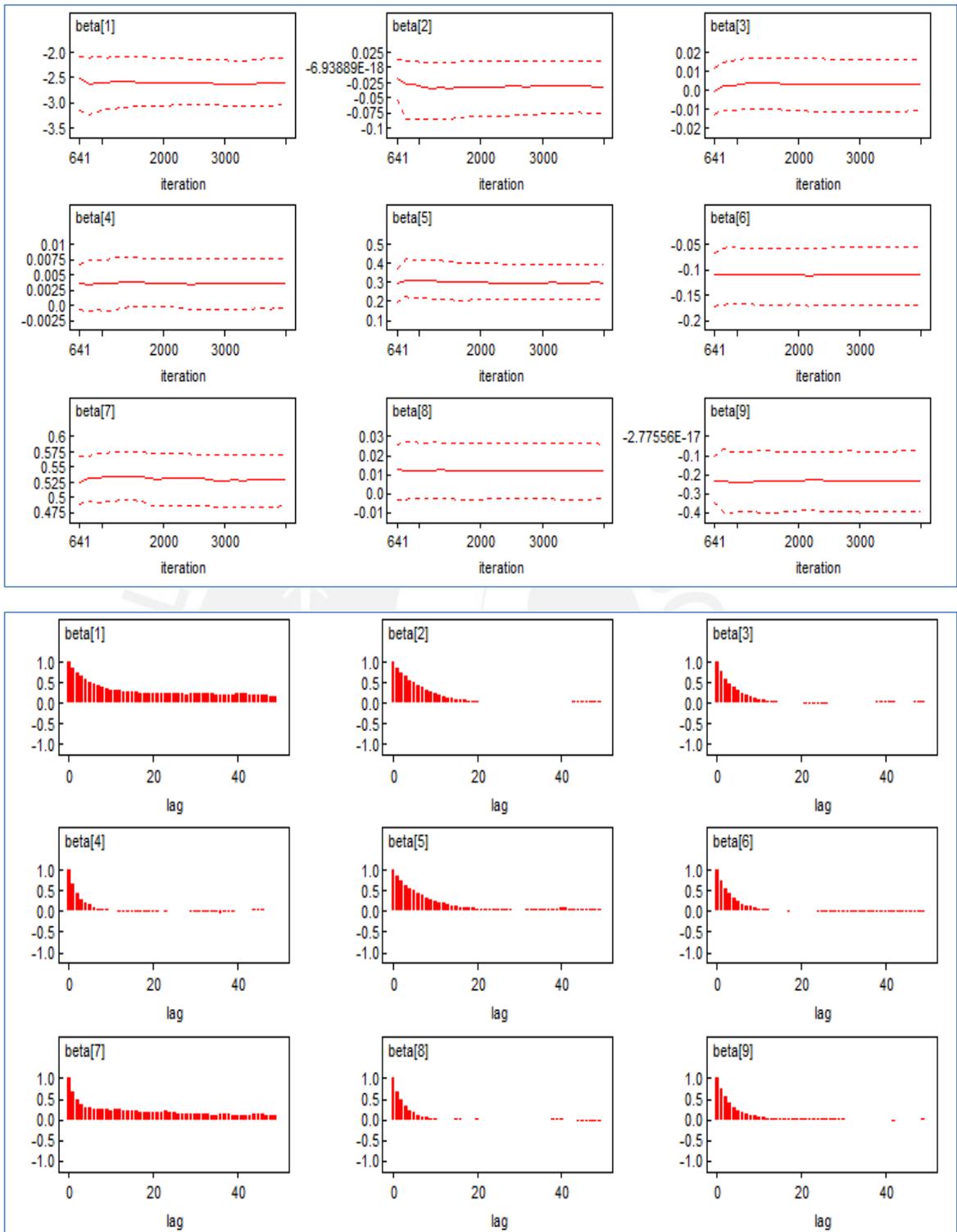


Figura A.8: Gráficos de Quantiles y Autocorrelacion Modelo BBB

Apéndice B

Resumen de Programas Utilizados

Sintaxis para Modelo BBB (Programa BRMUW)

```

model
{
  for(i in 1:n) {
    m[i] <- beta[1]+beta[2]*X1[i]+beta[3]*X2[i]+beta[4]*X3[i]+beta[5]*X4[i]
    +beta[6]*X5[i]+beta[7]*X6[i]+beta[8]*X7[i]+beta[9]*X8[i]
    muz[i]<- m[i]-delta*V[i]
    zs[i]~dnorm(muz[i],preczs)I(lo[Y[i]+1],up[Y[i]+1])
    V[i] ~ dnorm(0,1)I(0,)
  }

  for (j in 1:k) {beta[j] ~ dnorm(0.0,1.0E-3)}
  delta ~ dunif(-1,1)
  preczs<- 1/(1-pow(delta,2))
  lambda<-delta/sqrt(1-pow(delta,2))
  lo[1]<- -50; lo[2]<- 0;
  up[1]<- 0; up[2]<-50;
}

Inits
list(beta=c(0,0,0,0,0,0,0,0,0),delta=0.5)

Data
list(n=4745,k=9)

```

Sintaxis para Modelo CDS (Programa BRMUW)

```

model
{

```

```

for(i in 1:n) {
  m[i] <- beta[1]+beta[2]*X1[i]+beta[3]*X2[i]+beta[4]*X3[i]+beta[5]*X4[i]
  +beta[6]*X5[i]+beta[7]*X6[i]+beta[8]*X7[i]+beta[9]*X8[i]
  muz[i]<- m[i]+delta*sigma*V[i]
  zs[i]~dnorm(muz[i],1)I(lo[Y[i]+1],up[Y[i]+1])
  V[i] ~ dnorm(0,1)I(0,)
}

for (j in 1:k) {beta[j] ~ dnorm(0.0,1.0E-3)}
delta ~ dunif(-1,1)
lambda<-delta/sqrt(1-pow(delta,2))
sigma<-1/sqrt(1-pow(delta,2))
lo[1]<- -50; lo[2]<- 0;
up[1]<- 0; up[2]<-50;
}

Inits
list(beta=c(0,0,0,0,0,0,0,0,0),delta=0.5)

Data
list(n=4745,k=9)

Sintaxis para Modelo ESTANDAR (Programa BRMUW)

model
{
  for(i in 1:n) {
    m[i] <- beta[1]+beta[2]*X1[i]+beta[3]*X2[i]+beta[4]*X3[i]+beta[5]*X4[i]
    +beta[6]*X5[i]+beta[7]*X6[i]+beta[8]*X7[i]+beta[9]*X8[i]
    muz[i]<- m[i]-mu-sigma*delta*V[i]
    zs[i]~dnorm(muz[i],preczs)I(lo[Y[i]+1],up[Y[i]+1])
    V[i] ~ dnorm(0,1)I(0,)
  }

  for (j in 1:k) {beta[j] ~ dnorm(0.0,1.0E-3)}
  delta ~ dunif(-1,1)
  preczs<-1/((1-pow(delta,2))*sigma2)
  sigma2<-3.1415/(3.1415-2*pow(delta,2))
  sigma<-sqrt(sigma2)
  mu<--1.4142*delta/sqrt(3.1415-2*pow(delta,2))
  lo[1]<- -50; lo[2]<- 0;
}

```

```

up[1]<- 0; up[2]<-50;
}

```

```

Inits
list(beta=c(0,0,0,0,0,0,0,0,0),delta=0.5)

```

```

Data
list(n=4745,k=9)

```

Sintaxis para Modelo Probit (Programa BRMUW)

```

model
{
  for(i in 1:n) {
    Y[i] ~ dbern(p[i])
    p[i]<- phi(m[i])
    m[i] <- beta[1]+beta[2]*X1[i]+beta[3]*X2[i]+beta[4]*X3[i]+beta[5]*X4[i]
    +beta[6]*X5[i]+beta[7]*X6[i]+beta[8]*X7[i]+beta[9]*X8[i]
  }
  for (j in 1:k) {beta[j] ~ dnorm(0.0,1.0E-3)}
}

```

```

Inits
list(beta=c(0,0,0,0,0,0,0,0,0))

```

```

Data
list(n=4745,k=9)

```

Sintaxis Modelo SKEW PROBIT BBB (Calculo de Probabilidades - Programa R)

Funcion General a Usar:

```
psn(x, location = 0, scale = 1, shape = lambda)
```

Donde:

```
x=(XBeta=X1Beta1, X2Beta2.....)
```

```
location=media (Para el caso = 0)
scale=varianza (Para el caso = 1)
shape=lambda (Coeficiente de Asimetria)
```

```
xbeta <- c(0.912830, -2.489328, .....)
```

```
Probsp_BBB <-psn(xbeta, location=0, scale = 1, shape = -0.17750)
```

```
Probsp_BBB
```

Sintaxis Modelo SKEW PROBIT CDS (Calculo de Probabilidades - Programa R)

Funcion a Usar:

```
psn(x, location = 0, scale = 1, shape = lambda)
```

Donde:

```
x=(XBeta=X1Beta1, X2Beta2.....)
```

```
location=media (Para el caso = 0)
scale=varianza (Para el caso = 4.1612)
shape=lambda (Coeficiente de Asimetria)
```

```
xbeta <- c(0.912830, -2.489328, .....)
```

```
Probsp_CDS <-psn(xbeta, location=0, scale = 4.1612, shape = -1.7780)
```

```
Probsp_CDS
```

Sintaxis Modelo SKEW PROBIT ESTANDAR (Calculo de Probabilidades - Programa R)

Funcion a Usar:

```
psn(x, location = 0, scale = 1, shape = lambda)
```

Donde:

```
x=(XBeta=X1Beta1, X2Beta2.....)
```

```
location=media, Para el caso = [(2^(0.50))*delta]/((pi - 2*delta^2)^(0.50))]
scale=varianza, Para el caso = [pi/(pi - 2*(delta^2))]
shape=lambda (Coeficiente de Asimetria)
delta= [lambda / (1 + lambda^2)^(0.50)]
```

```
xbeta <- c(0.912830, -2.489328, .....)
```

```
Probsp_ESTANDAR <-psn(xbeta, location=-0.192302, scale = 1.036980, shape = 0.2436)
```

```
Probsp_ESTANDAR
```



Bibliografía

- Albert, J. y Chib, S. (1993). *Bayesian Analysis of binary and polytomous response data*, American Statistical Association.
- Albert, J. y Johnson, J. (1999). *Ordinal Data Modeling*, Springer.
- Azzalini, A. (2005). *Introdution to the skew normal distribution*, Department of Statistical Sciences - University of Padua.
- Azzalini, A. y Regoli, G. (2010). *Some properties of skew-symmetric distributions*, Padova - University.
- Bazan, J. y Bayes, C. (2010). *Inferencia Bayesiana en Modelos de Regresion Binaria Usando BRMUW*, Department of Sciences PUCP.
- Bazan, J. y Bolfarine, H. (2006). *A Generalized skew probit class link for binary regression*, Department of Sciences PUCP.
- Bermudez, L., Vasquez, F. y Gomez, E. (2007). *A Bayesian Dichotomous Model with Asymmetric Link for Fraud in Insurance*, Elsevier.
- Carlin, B. y Polson, N. (1992). *Monte Carlo Bayesian Methods for Discrete Regression Models and Categorical Time Series*, Oxford - University.
- Chen, M., Shao, M. y Ibrahim, J. (1999). *A New Skewed Link Model for Dichotomous Quantal Response Data*, Springer.
- Fadden, M. y Manski, C. (1981). *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge - University.
- Mendenhall, W. y Wackerly, D. (2009). *Estadística Matemática con Aplicaciones*, CENGAGE - Learning.
- Nagler, J. (1994). *Scobit: An Alternative Estimator to Logit and Probit*, American Journal of Political Science.
- Sahu, S. y D. Branco, M. (2003). *A new class of multivariate skew distributions with applications to Bayesian regression models*, The Canadian Journal of Statistics.
- Spiegelhalter, D. y Best, N. (2002). *Bayesian measures of model complexity and fit*, Journal of the Royal Statistical Society.
- Svetlozar, T. y Rachev, J. (2008). *Bayesian Methods in Finance*, Wiley.