



**Universidad
del Tolima**

FACULTAD DE CIENCIAS

DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA

**ESTUDIO DEL BAJO PESO DE LOS NIÑOS
MEDIANTE GAMLSS EN IBAGUÉ AÑO 2014**

**Trabajo de grado para optar al título de Profesional en
Matemáticas con énfasis en Estadística**

Katherin Ramirez Rojas

Director:

Mg. Alfonso Sánchez Hernández

Departamento de Matemáticas y Estadística

Programa de Matemáticas con énfasis en Estadística

IBAGUÉ

2018

**ESTUDIO DEL BAJO PESO DE LOS NIÑOS
MEDIANTE GAMLSS EN IBAGUÉ AÑO 2014**

Autor:

Katherin Ramirez Rojas

Código:

070200172012

**Trabajo de grado para optar al título de Profesional en
Matemáticas con énfasis en Estadística**

Director:

Alfonso Sánchez Hernández

MSc. en Investigación Operativa y Estadística

**UNIVERSIDAD DEL TOLIMA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA
IBAGUÉ
2018**



UNIVERSIDAD DEL TOLIMA

FACULTAD DE CIENCIAS

PROGRAMA DE MATEMÁTICAS CON ÉNFASIS EN ESTADÍSTICA

ACTA DE SUSTENTACIÓN TRABAJO DE GRADO

TÍTULO: ESTUDIO DEL BAJO PESO DE LOS NIÑOS MEDIANTE GAMLSS EN IBAGUÉ, AÑO 2014

AUTORES: KATHERIN RAMIREZ ROJAS Céd. 070200172012

DIRECTOR: ALFONSO SÁNCHEZ HERNÁNDEZ

JURADOS: MIGUEL ARMANDO RODRÍGUEZ
JAIRO ALFONSO CLAVIJO

CALIFICACIÓN: 4.7 (CUATRO SIETE - MERITORIO)

APROBÓ

REPROBÓ

OBSERVACIONES:

FIRMAS


MIGUEL ARMANDO RODRÍGUEZ


JAIRO ALFONSO CLAVIJO

Jurado 1

Jurado 2


ALFONSO SÁNCHEZ HERNÁNDEZ
Director del Trabajo


YURI MARCELA GARCÍA S.
Directora del Programa

Ciudad y fecha: Ibagué, 22 de Octubre de 2018

Dedicatoria

Dedico este trabajo a Dios, por estar en cada paso de mi vida, por darme la oportunidad de llegar hasta aquí, y sé que seguirá guiándome en el camino que aún falta por recorrer.

Agradecimientos

Agradezco a Dios por permitirme aprender y formarme como profesional en esta, mi Universidad del Tolima, principalmente a la Facultad de Ciencias Básicas por contar con magníficos profesores, tal como el profesor Alfonso Sánchez Hernández, a él gracias por su dedicación y orientación en mi tesis, gracias a la Secretaria de Salud Municipal de Ibagué y al señor Carlos Bonilla por facilitar la base de datos con la que trabajé en este proyecto. Por otro lado Agradezco inmensamente a Haider Montes Másmela por su apoyo incondicional que me brindó día a día para seguir persistiendo y finalmente a Víctor Alexander Piamba por confiar en mí y brindarme la oportunidad de estudiar.

Índice General

RESUMEN	XI
INTRODUCCIÓN	XII
PLANTEAMIENTO DEL PROBLEMA	XIII
JUSTIFICACIÓN	XIV
OBJETIVOS	XV
1. REFERENTES TEÓRICOS	1
1.1. Modelos Lineales Generalizados (GLM)	1
1.1.1. Componentes de un Modelo Lineal Generalizado	1
1.1.2. Inferencia en un Modelo Lineal Generalizado (GLM)	2
1.1.3. Función Score y Matriz de Información	3
1.1.4. Estimación de β	4
1.1.5. Función Desvío	4
1.1.6. Estimación del parámetro de dispersión ϕ	5
1.1.7. Hipótesis	5
1.1.8. Residuos	6
1.1.9. Influencia Local	7
1.1.10. Criterios de Información	8

1.2.	Modelos Aditivos Generalizados (GAMs)	8
1.2.1.	Funciones suaves univariadas.	9
1.2.2.	Regresión Spline.	9
1.2.3.	Grado de suavizamiento mediante regresión spline penalizada.	10
1.2.4.	Validación cruzada para la obtención del parámetro de suavización λ	11
1.3.	Modelos Aditivos de Localización, Escala y Forma (GAMLSS)	11
1.3.1.	Ventajas de los GAMLSS	12
1.3.2.	Inferencia en GAMLSS	12
1.3.3.	Estimando β y γ con λ fijo	14
1.3.4.	El algoritmo RS	18
1.3.5.	La iteración outer (iteración GAMLSS)	18
1.3.6.	La iteración <i>inner</i> (iteración GLM o GLIM)	18
1.3.7.	El algoritmo modificado backfitting	19
1.4.	Muestreo Aleatorio Estratificado por Afijación Óptima	20
2.	Análisis de Resultados	24
2.1.	Selección de la muestra por Afijación Óptima	24
2.2.	Selección del GAMLSS	25
3.	Conclusiones	29
4.	Recomendaciones	30
	Apéndices	31

A. Apéndice de Tablas	31
A.1. Base de datos	31
A.2. Nombre de las variables	32
B. Apéndice de Gráficas	33
B.1. Gráfica de Ajuste Modelo 1	33
B.2. Gráfica de Ajuste Modelo 2	34
B.3. Gráfica de Ajuste Modelo 3	35
B.4. Gráfica de Variable semiparamétrica Talla	36

Índice de tablas

1.	Principales distribuciones de la familia exponencial	2
2.	Tamaños de muestra por estrato mediante afijación óptima	24
3.	Estimación del modelo paramétrico	25
4.	Estimación del modelo, usando Talla como variable semiparamétrica	26
5.	Estimación del mejor modelo, usando Talla como variable semiparamétrica	26
6.	Criterios de Información de Akaike y de Bayes para los tres modelos	27
7.	Probabilidad de éxito y fracaso para el Primer Ajuste.	27
8.	Probabilidad de éxito y fracaso para el Segundo Ajuste.	27
9.	Probabilidad de éxito y fracaso para el Tercer Ajuste.	28

Índice de figuras

1.	Como los algoritmos RS y CG alcanzan las estimaciones de máxima verosimilitud, modelo paramétrico. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 63).	14
2.	La iteración outer dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 65).	15
3.	La iteración <i>inner</i> (o iteración GLIM) dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 67).	16
4.	El algoritmo modificado backfitting dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 69).	17

RESUMEN

Los Modelos Lineales clásicos, dentro de los que se cuentan los modelos de regresión y análisis de varianza han sido muy útiles en el desarrollo de la ciencias, llámense puras o aplicadas, desde el descubrimiento del método de mínimos cuadrados, los que inicialmente fueron publicados por Legendre en *Principle of Means Squares* en 1805. La necesidad de utilizar técnicas más flexibles y expandir la posibilidad de emplear distribuciones diferentes a la normal, dió origen a los Modelo Lineales Generalizados, con el trabajo de Nelder y Wedderburn (1972, [4]) y se han convertido en una herramienta obligada para los usuarios e investigadores de la Estadística Teórica y Aplicada. Sus aportes al conocimiento y a la Ciencia han sido innumerables y por ende su aplicabilidad a problemas del entorno han sido exitosos y en este momento de tecnologías modernas, han permitido estudiar problemas de gran complejidad. Más adelante entre los años 1989 y 1990, Hastie y Tibshirani dieron paso a los Modelos Aditivos Generalizados, en los cuales algunas variables independientes del modelo se pueden utilizar como funciones suaves tipo Fourier ó Spline, Hastie y Tibshirani (1990, [17]). No obstante el gran avance en este campo y los métodos computacionales modernos, Rigby y Stasinopoulus (2017, [1], pp. 24) presentan a la comunidad científica, la versión más actualizada de modelos conocidos hasta hoy, los Modelos Aditivos Generalizados de Localización, Escala y Forma (GAMLSS), los cuales fusionan los modelos lineales generalizados y los modelos aditivos generalizados.

En el entorno social existen problemas como el de la nutrición y desde ese punto de vista se hace prescindible estudiar el bajo peso de los niños al nacer. En este trabajo se pretende estudiar esta temática utilizando Modelos Aditivos Generalizados de Localización, Escala y Forma (GAMLSS), usando la base de datos facilitada por la Secretaría de Salud de Ibagué, en el año 2014.

INTRODUCCIÓN

Los Modelos Lineales Generalizados (GLM), desde que fueron propuestos por Nelder y Wedderburn (1972, [4]), han tenido múltiples desarrollos y aplicaciones en todas las áreas de la ciencia, tanto pura como aplicada. Esto se debe en gran parte a la multiplicidad de distribuciones a las que puede pertenecer la variable respuesta y a la transformación de la media de la variable respuesta, llamada función de enlace ó función *link*, la cual permite relacionar la media de dicha variable con el predictor lineal.

Fué tan importante el descubrimiento de los Modelos Lineales Generalizados, que autores como Agresti (2015, [8]) consideran que un *Modelo Lineal clásico* es un Modelo Lineal Generalizado con función de enlace idéntica y forman un conjunto muy pequeño dentro de este gran universo de modelos.

Una modificación interesante de los Modelos Lineales Generalizados fué propuesta por Hastie y Tibshirani (1990, [17]). Estos autores propusieron incluir en el predictor lineal variables independientes en forma de funciones suaves, las cuales incluyen no linealidad y posibles correlaciones entre ellas. Paula (2013, [5]) denominó a estos modelos *semiparámetros*, debido a las variables independientes con este tipo de características. No obstante Hastie y Tibshirani los denominaron *Modelos Aditivos Generalizados* (GAMs).

En el campo del modelamiento estadístico se pretenden estudiar los parámetros de localización y escala, es decir la media y la varianza de la variable respuesta. Sin embargo Rigby y Stasinopoulos (2005, [3]) proponen los Modelos Aditivos Generalizados de Localización, Escala y Forma, en los cuales no sólo se estudia la media y la varianza, sino que se incluyen los parámetros de forma tales como la asimetría y la curtosis para su estudio, y en el que una gran variedad de distribuciones pueden ser estudiadas para la variable respuesta.

El presente trabajo se basa en la aplicación de Hachuel et. al (2006, [26]), quien utilizó Modelos Lineales Generalizados Mixtos, en el problema del bajo peso de los niños al nacer en Argentina, entre los años 2000 a 2003. Si bien existen muchos otros trabajos en el área, fué él el precursor en esta línea de trabajo.

Basados en la información suministrada por la Secretaría de Salud Municipal de la ciudad de Ibagué, en la cual se presentan los datos tomados de los certificados de nacido vivo, en Ibagué y municipios aledaños, actualizados por el DANE, se pretende ajustar un Modelo Aditivo Generalizado de Localización, Escala y Forma (GAMLSS), asociado al bajo peso, teniendo en cuenta el conjunto de covariables relacionadas con este problema. Este sería el tercer y último trabajo acerca de esta temática, el cual permitirá determinar los principales factores asociados al bajo peso de los neonatos.

PLANTEAMIENTO DEL PROBLEMA

El objetivo fundamental de este trabajo consiste en reevaluar el uso de Modelos Lineales Generalizados para el bajo peso al nacer, siguiendo los lineamientos de: Hachuel (2004, [26]), Jiménez y Lozano (1996, [27]), Suárez (2014, [30]), entre otros. El reto para la realización de este trabajo, permite formular las siguientes preguntas:

- Cuando los errores en un modelo estadístico no son normales, ¿Se pueden considerar otras metodologías alternativas que permiten estimar los parámetros del modelo, obteniendo resultados favorables?
- ¿Se puede pensar que los Modelos Aditivos Generalizados de Localización, Escala y Forma, pueden ser una alternativa razonable, para estudiar este tipo de problemas?
- ¿Se pueden incluir variables independientes que sean de naturaleza binaria, categóricas y de otra naturaleza en el modelo apropiado?
- ¿Es posible utilizar alguna técnica de diagnóstico, la cual permita identificar observaciones atípicas de los datos y la adecuada sensibilidad del modelo estimado?

JUSTIFICACION

El Bajo Peso al Nacer (BPN) de los niños es para los estudiosos en Ciencias Médicas y Sociales una problemática permanente de estudio. Lo anterior se suma a una cantidad de situaciones que van desde los embarazos no deseados, hasta malformaciones congénitas en el feto, cuyas causas pueden ser endógenas ó exógenas. Examinando un poco el ambiente social, se evidencia que muchos gobiernos intentan examinar leyes que permitan el aborto, el cual en la mayoría de los casos va en contra de principios morales, éticos o religiosos. Para la Organización Mundial de la Salud, es una situación que debe ser enfrentada por parte de los gobiernos y minimizada por organizaciones de salud y defensoras de los derechos humanos, teniendo en cuenta que la vida es un derecho inalienable.

En América Latina, más exactamente en Argentina Hachuel (2004) fué el pionero en tratar este problema utilizando Modelos Lineales Generalizados Mixtos. En ese estudio el mismo Hachuel propone utilizar como variable respuesta una relación binaria, en la cual se determina si el neonato tiene o no bajo peso y como covariables una serie de variables de tipo socio-económico, las cuales se registran en el certificado de nacido vivo. En Colombia este certificado es utilizado por el Departamento Administrativo Nacional de Estadística (DANE).

Este es el tercer trabajo acerca de la temática, y se pretende mediante el mismo establecer el uso del tipo de modelamiento más reciente en la literatura estadística y que corresponde al uso de Modelos Aditivos Generalizados de Localización, Escala y Forma, propuestos por Rigby y Stasinopoulos (2005, [3]). Se utiliza la base de datos suministrada por la Secretaría de Salud de Ibagué y que corresponde a los datos de los registros de nacido vivo en Ibagué y municipios aledaños durante el año 2014.

OBJETIVOS

OBJETIVO GENERAL

Estudiar el bajo peso de los niños al nacer en Ibagué y municipios aledaños durante el año 2014, mediante Modelos Aditivos Generalizados de Localización, Escala y Forma (GAMLSS).

OBJETIVOS ESPECIFICOS

- Dar a conocer algunos breves aspectos teóricos importantes, referentes a Modelos Aditivos Generalizados de Localización, Escala y Forma, útiles para su aplicación a problemas reales.
- Formular una metodología basada en muestreo, para el estudio de la base de datos y establecer un proceso metodológico para su respectivo análisis.
- Implementar la utilización de un Modelo GAMLSS para la estimación de los parámetros en un modelo que caracterice de manera eficiente el problema del bajo peso al nacer.
- Realizar un estudio de sensibilidad e influencia local para realizar la validación del modelo estimado.
- Interpretar de una manera estadística el comportamiento del bajo peso al nacer de los niños utilizando la base suministrada por la Secretaría de Salud Municipal de Ibagué para el año 2014.

1. REFERENTES TEÓRICOS

1.1. Modelos Lineales Generalizados (GLM)

Los Modelos Lineales Generalizados (GLM) fueron propuestos por Nelder y Wedderburn (1972, [4]), se crearon a partir del gran desarrollo computacional después de los años 70 y algunos modelos que exigían procesos iterativos para la estimación de sus parámetros, comenzaron a tener mayor utilidad. Por ejemplo modelos con estructura no lineal en los parámetros tuvieron un gran avance, Paula (2013, [5], pp. 1). Gran cantidad de textos especializados en este tópico han sido escritos, dentro de ellos se pueden citar: Nelder y McCullag (1989, [6]), Dobson (2000, [7]) y Agresti (2017, [8]) entre muchos otros.

1.1.1. Componentes de un Modelo Lineal Generalizado

Un modelo lineal generalizado está formado por tres partes esenciales:

- **Componente aleatoria:** representa la variable respuesta Y , que puede ser de valor real ó un vector $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)$. Cada variable ó componente del vector pertenece a la familia exponencial.
- **Componente sistemática:** corresponde al predictor lineal del modelo y se nota $\eta_i = \mathbf{X}_i^T \boldsymbol{\beta} = \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} = \sum_{j=1}^p X_{ij} \beta_j$, para la componente i -ésima del vector \mathbf{Y} , ó $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$ para el vector completo \mathbf{Y} .
- **Función de enlace o link:** relaciona la esperanza matemática de la variable dependiente con el predictor lineal, $g(\mu_i) = \eta_i$ para $i = 1, 2, \dots, n$. La función $g(\cdot)$ debe ser monótona, suave y diferenciable.

Una distribución pertenece a la familia exponencial, si su función de densidad de probabilidad (variable aleatoria continua) ó función de densidad discreta (variable aleatoria discreta) se puede escribir como:

$$f(y_i, \theta_i, \phi) = \exp[\phi\{y_i\theta_i - b(\theta_i)\} + c(y_i, \phi)]$$

donde b y c son funciones arbitrarias, ϕ es un parámetro de dispersión y θ_i es conocido como el parámetro canónico de la distribución.

$$E(Y_i) = \mu_i = b'(\theta_i) \quad \text{y} \quad \text{var}(Y_i) = \phi^{-1} V_i; \quad V_i = \frac{\partial \mu_i}{\partial \theta_i}$$

A V_i se le conoce como función varianza, también se nota $V(\mu_i)$.

Un caso particular ocurre cuando el parámetro canónico θ_i coincide con el predictor lineal η_i . Distribuciones en las que esto ocurre se denominan *distribuciones de enlace canónico*. En la tabla 1 se presentan las principales distribuciones pertenecientes a la familia exponencial.

Distribución	$b(\theta)$	θ	ϕ	$V(\mu)$
Normal	$\theta^2/2$	μ	σ^{-2}	1
Poisson	e^θ	$\log(\mu)$	1	μ
Binomial	$\log(1 + e^\theta)$	$\log\{\mu/(1 - \mu)\}$	n	$\mu(1 - \mu)$
Gamma	$-\log(-\theta)$	$-1/\mu$	$1/(\text{CV})^2$	μ^2
N. Inversa	$-\sqrt{-2\theta}$	$-1/2\mu^2$	ϕ	μ^3

Tabla 1: Principales distribuciones de la familia exponencial

Los enlaces canónicos para los modelos normal, binomial, Poisson, gamma e inversa gaussiana son respectivamente:

$$\mu = \eta, \quad \log(\mu) = \eta, \quad \log\left\{\frac{\mu}{1 - \mu}\right\} = \eta, \quad \mu^{-1} = \eta, \quad \text{y} \quad \mu^{-2} = \eta$$

1.1.2. Inferencia en un Modelo Lineal Generalizado (GLM)

La metodología estadística para la estimación de los parámetros en un GLM se basa en *máxima verosimilitud*. aquí se sigue estrictamente la metodología y notación utilizada por Paula (2004, [5], pp. 5).

El logaritmo de verosimilitud con respuestas independientes se expresa:

$$L(\boldsymbol{\beta}; \mathbf{y}) = \sum_{i=1}^n \phi\{y_i \theta_i - b(\theta_i)\} + \sum_{i=1}^n c(y_i, \phi)$$

Cuando el parámetro canónico θ_i coincide con el predictor lineal, es decir cuando $\theta_i = \eta_i = \sum_{j=1}^p x_{ij}\beta_j$, la ecuación anterior se convierte en:

$$L(\boldsymbol{\beta}; \mathbf{y}) = \sum_{i=1}^n \phi\left\{y_i \sum_{j=1}^p x_{ij}\beta_j - b\left(\sum_{j=1}^p x_{ij}\beta_j\right)\right\} + \sum_{i=1}^n c(y_i, \phi)$$

Al definir $S_j = \phi \sum_{i=1}^n Y_i x_{ij}$, la ecuación anterior se puede escribir:

$$L(\boldsymbol{\beta}; \mathbf{y}) = \sum_{j=1}^p \mathbf{s}_j \beta_j - \phi \sum_{i=1}^n \mathbf{b}\left(\sum_{j=1}^p \mathbf{x}_{ij}\beta_j\right) + \sum_{i=1}^n c(\mathbf{y}_i, \phi)$$

Y por el teorema de factorización, la estadística $S = (S_1, S_2, \dots, S_n)^T$ es una estadística suficiente y minimal para el vector de parámetros $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$. Los enlaces que corresponden a tales estadísticas son llamados *enlaces canónicos* y cumplen un papel importante en la teoría de los GLM, pues garantizan concavidad y resultados asintóticos importantes.

1.1.3. Función Score y Matriz de Información

Para obtener la función score del parámetro $\boldsymbol{\beta}$ se calcula la primera derivada de la función logaritmo de verosimilitud, Paula (2004, [5], pp. 15).

$$\begin{aligned} \partial L(\boldsymbol{\beta}; \mathbf{y}) / \partial \beta_j &= \sum_{i=1}^n \phi \left\{ y_i \frac{d\theta_i}{d\mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_j} - \frac{db(\theta_i)}{d\theta_i} \frac{d\theta_i}{d\mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_j} \right\} \\ &= \sum_{i=1}^n \phi \left\{ y_i V_i (d\mu_i / d\eta_i) x_{ij} - \mu_i V_i (d\mu_i / d\eta_i) x_{ij} \right\} \\ &= \sum_{i=1}^n \phi \left\{ \sqrt{\frac{\omega_i}{V_i}} (y_i - \mu_i) x_{ij} \right\} \end{aligned}$$

en donde $\omega_i = (d\mu_i / d\eta_i)^2 / V_i$ y la función score escrita en forma vectorial queda:

$$U(\boldsymbol{\beta}) = \frac{\partial L(\boldsymbol{\beta}; \mathbf{y})}{\partial \boldsymbol{\beta}} = \phi \mathbf{X}^T \mathbf{W}^{1/2} \mathbf{V}^{-1/2} (\mathbf{y} - \boldsymbol{\mu})$$

en donde \mathbf{X} es una matriz de orden $n \times p$ de rango completo, $\mathbf{W} = \text{diag}\{\omega_1, \omega_2, \dots, \omega_n\}$ es una matriz de pesos, $\mathbf{V} = \text{diag}\{V_1, V_2, \dots, V_n\}$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$ y $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$.

La matriz de información de Fisher se obtiene mediante el cálculo de las segundas derivadas parciales de la función logaritmo de verosimilitud ó de las primeras derivadas parciales de la función score:

$$\begin{aligned} \partial^2 L(\boldsymbol{\beta}; \mathbf{y}) / \partial \beta_j \partial \beta_l &= \phi \sum_{i=1}^n (y_i - \mu_i) \frac{d^2 \theta_i}{d\mu_i^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 x_{ij} x_{il} + \phi \sum_{i=1}^n (y_i - \mu_i) \frac{d\theta_i}{d\mu_i} \frac{d^2 \mu_i}{d\eta_i^2} x_{ij} x_{il} \\ &\quad - \phi \sum_{i=1}^n \frac{d\theta_i}{d\mu_i} \left(\frac{d\mu_i}{d\eta_i} \right) x_{ij} x_{il} \end{aligned}$$

de donde su valor esperado es:

$$E(\partial^2 L(\boldsymbol{\beta}; \mathbf{y}) / \partial \beta_j \partial \beta_l) = -\phi \sum_{i=1}^n \frac{d\theta_i}{d\mu_i} \left(\frac{d\mu_i}{d\eta_i} \right) x_{ij} x_{il} = -\phi \sum_{i=1}^n \frac{(d\mu_i / d\eta_i)^2}{V_i} x_{ij} x_{il} = -\phi \sum_{i=1}^n \omega_i x_{ij} x_{il}$$

y la matriz de información de Fisher en forma matricial se puede escribir:

$$\mathbf{K}(\boldsymbol{\beta}) = E \left\{ -\frac{\partial^2 L(\boldsymbol{\beta}; \mathbf{Y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \right\} = \phi \mathbf{X}^T \mathbf{W} \mathbf{X}$$

1.1.4. Estimación de β .

Se utiliza el proceso iterativo de Newton-Raphson. y la estimación de máxima verosimilitud de β se define expandiendo la función *score* $U(\beta)$ alrededor de un valor inicial $\beta^{(0)}$, es decir:

$$U(\beta) \approx U(\beta^{(0)}) + U'(\beta^{(0)})(\beta - \beta^{(0)})$$

en donde $U'(\beta)$ representa la primera derivada de la función *score*. Al despejar β y repetir iterativamente el proceso se llega a:

$$\beta^{(m+1)} = \beta^{(m)} + \{-U'(\beta^{(m)})\}^{-1}U(\beta^{(m)})$$

con $m = 0, 1, \dots$ y como la matriz $U'(\beta^{(m)})$ puede llegar a no ser definida positiva, ésta se cambia por la inversa de la matriz de información de Fisher, llegando al proceso iterativo:

$$\beta^{(m+1)} = \beta^{(m)} + \mathbf{K}^{-1}(\beta^{(m)})U(\beta^{(m)})$$

con $m = 0, 1, \dots$. Al realizar un poco de álgebra se llega a un proceso de mínimos cuadrados reponderados:

$$\beta^{(m+1)} = (\mathbf{X}^T \mathbf{W}^{(m)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(m)} \mathbf{z}^{(m)}$$

con $m = 0, 1, \dots$ y $\mathbf{z} = \eta + \mathbf{W}^{-1/2} \mathbf{V}^{-1/2} (\mathbf{y} - \boldsymbol{\mu})$ es una variable dependiente modificada.

Sen y Singer (1993, [10], Cap. 7) afirman que bajo condiciones de regularidad, el estimador de máxima verosimilitud de $\hat{\beta}$ es un estimador consistente y eficiente de β . esto significa:

$$\sqrt{n}(\hat{\beta} - \beta) \rightarrow_d N_p(\mathbf{0}, \boldsymbol{\Sigma}^{-1}(\beta)) \quad \text{cuando } n \rightarrow \infty$$

Además

$$\boldsymbol{\Sigma}(\beta) = \lim_{n \rightarrow \infty} \frac{\mathbf{K}(\beta)}{n}$$

siendo $\boldsymbol{\Sigma}(\beta)$ una matriz definida positiva.

1.1.5. Función Desvío

La función *desvío* sin pérdida de generalidad representa la discrepancia entre la verosimilitud del modelo saturado con n parámetros y el modelo estimado con p parámetros. Si se representa el logaritmo de verosimilitud para el modelo estimado por

$$L(\boldsymbol{\mu}; \mathbf{y}) = \sum_{i=1}^n L(\mu_i; y_i)$$

se sabe que $\mu_i = g^{-1}(\eta_i)$ y $\eta_i = \mathbf{X}_i^T \boldsymbol{\beta}$. Para el modelo saturado ($p = n$) la función logaritmo de verosimilitud es:

$$L(\mathbf{y}; \mathbf{y}) = \sum_{i=1}^n L(y_i; y_i)$$

La calidad del ajuste de un GLM se evalúa a través de la función *desvío*:

$$D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) = \phi D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = 2\{L(\mathbf{y}; \mathbf{y}) - L(\hat{\boldsymbol{\mu}}; \mathbf{y})\}$$

Interesantes propiedades de la función desvío pueden ser consultadas en Jorgensen (1987, [11]).

1.1.6. Estimación del parámetro de dispersión ϕ

Se puede demostrar que los parámetros $\boldsymbol{\beta}$ y ϕ son ortogonales, esto se evidencia en $E[\partial^2 L(\boldsymbol{\beta}; \phi; \mathbf{y}) / \partial \boldsymbol{\beta} \partial \phi] = 0$. Una consecuencia inmediata es la independencia asintótica entre sus estimaciones $\hat{\phi}$ y $\hat{\boldsymbol{\beta}}$. Al derivar el logaritmo de verosimilitud con respecto al parámetro ϕ e igualando a cero, se llega a la solución:

$$\sum_{i=1}^n c'(y_i, \hat{\phi}) = \frac{1}{2} D(\mathbf{y}; \hat{\boldsymbol{\mu}}) - \sum_{i=1}^n \{y_i \hat{\theta}_i^0 - b(\hat{\theta}_i^0)\}$$

donde $D(\mathbf{y}; \hat{\boldsymbol{\mu}})$ representa la función desvío sobre el modelo bajo investigación. La estimación de máxima verosimilitud para ϕ en los modelos normal e inversa gaussiana están dados por: $\hat{\phi} = n/D(\mathbf{y}; \hat{\boldsymbol{\mu}})$. Para la distribución gamma es

$$2n\{\log(\hat{\phi}) - \psi(\hat{\phi})\} = D(\mathbf{y}; \hat{\boldsymbol{\mu}})$$

en donde $\psi(\phi) = \Gamma'(\phi)/\Gamma(\phi)$ es una función digamma.

Un estimador preferido de ϕ está basado en la estadística de Pearson

$$\hat{\phi}^{-1} = \sum_{i=1}^n \{(y_i - \hat{\mu}_i)/\hat{\mu}_i\}^2 / (n - p)$$

La condición para este estimador es que $\boldsymbol{\beta}$ debe haber sido estimado consistentemente.

1.1.7. Hipótesis

En este documento sólo se hace referencia a hipótesis simples. Supóngase que se desea probar la hipótesis:

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}^0 \quad \text{versus} \quad H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}^0$$

en donde $\boldsymbol{\beta}^0$ es un vector p -dimensional conocido y ϕ también se asume conocido. Las siguientes pruebas son utilizadas en GLMs:

- **Prueba de Razón de Verosimilitud:** se define

$$\epsilon_{RV} = 2\{L(\hat{\boldsymbol{\beta}}; \mathbf{y}) - L(\boldsymbol{\beta}^0; \mathbf{y})\}$$

Esta estadística se define como la diferencia entre dos funciones desvío:

$$\epsilon_{RV} = \phi\{D(\mathbf{y}; \hat{\boldsymbol{\mu}}^0) - D(\mathbf{y}; \hat{\boldsymbol{\mu}})\}$$

- **Prueba de Wald:** utiliza la estadística:

$$\epsilon_W = [\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^0]^T \hat{\text{Var}}(\hat{\boldsymbol{\beta}}) [\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^0]$$

donde $\hat{\text{Var}}(\hat{\boldsymbol{\beta}})$ representa la matriz de varianzas-covarianzas asintótica de $\hat{\boldsymbol{\beta}}$ estimada en $\hat{\boldsymbol{\beta}}$. Para GLMs $\hat{\text{Var}}(\hat{\boldsymbol{\beta}}) = \mathbf{K}^{-1}(\hat{\boldsymbol{\beta}})$, así que la estadística se puede reescribir:

$$\epsilon_W = [\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^0]^T (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}) [\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^0]$$

cuando el número de parámetros es igual a 1 la prueba de Wald es equivalente a una prueba t de Student.

- **Prueba Score:** conocida también como la prueba de Rao y se define cuando $U(\hat{\boldsymbol{\beta}}) = 0$. Para GLMs se define:

$$\epsilon_{SR} = \phi^{-1} \mathbf{U}(\beta^0)^T (\mathbf{X}^T \hat{\mathbf{W}}_0 \mathbf{X})^{-1} \mathbf{U}(\beta^0)$$

en donde \hat{W}_0 es estimada sobre H_0 .

- **Prueba F:** esta prueba se define con base en la función desvío, esto es

$$F = \frac{\{D(\mathbf{y}; \hat{\boldsymbol{\mu}}^0) - D(\mathbf{y}; \hat{\boldsymbol{\mu}})\}/p}{D(\mathbf{y}; \hat{\boldsymbol{\mu}})/(n-p)} \sim F_{p,(n-p)} \quad \text{cuando } \phi \rightarrow \infty$$

Finalmente y sobre la hipótesis nula se tiene que ϵ_{RV} , ϵ_W y $\epsilon_{SR} \sim \chi_p^2$ y una región de confianza basada en la prueba de Wald para $\boldsymbol{\beta}$ y con $(1 - \alpha)$ de confianza es:

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}) (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq \phi^{-1} \chi_p^2 (1 - \alpha)$$

1.1.8. Residuos

Siguiendo los principios de Cook (1986, [13]), se conocen tres tipos de residuos, estos son:

- **Residuos simples:** se definen como la diferencia entre el valor observado de la variable respuesta y el valor estimado por el modelo. Tomando como referencia el modelo lineal se definen:

$$\mathbf{r} = \mathbf{Y} - \hat{\mathbf{Y}}; \quad E(\mathbf{r}) = \mathbf{0}; \quad \text{y} \quad \text{Var}(\mathbf{r}) = (\mathbf{I} - \mathbf{H})\sigma^2$$

en donde \mathbf{I} es la matriz idéntica y $\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ es llamada matriz hat, tiene las propiedades de ser simétrica e idempotente, es decir $\mathbf{H}^T = \mathbf{H}$ y $\mathbf{H}^T \mathbf{H} = \mathbf{H} \mathbf{H}^T = \mathbf{H}$.

- **Residuos estandarizados:** son los mismos residuos simples estandarizados en media y varianza, para la i -ésima observación se notan y definen:

$$t_i = \frac{r_i}{s(1 - h_{ii})} \quad \text{con} \quad s^2 = \sum_{i=1}^n r_i^2 / (n - p) \quad \text{para} \quad i = 1, 2, \dots, n$$

- **Residuos estudentizados:** se definen como la diferencia entre el valor observado de la variable y el valor estimado, cuando la i -ésima observación ha sido eliminada, o sea $Y_i - \hat{Y}_{(i)}$, más exactamente se definen:

$$t_i^* = \frac{r_i}{s_{(i)}(1 - h_{ii})^{1/2}}$$

$s_{(i)}^2$ es la varianza estimada sin la i -ésima observación.

En Rao (1973, [14], p.185) se establecen las siguientes relaciones:

$$s_{(i)}^2 = s^2 \left(\frac{n - p - t_i^2}{n - p - 1} \right) \quad \text{y} \quad t_i^* \left(\frac{n - p - 1}{n - p - t_i^2} \right)$$

1.1.9. Influencia Local

Al multiplicar el vector de observaciones de un modelo por un vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_n)^T$, donde δ_j es un tipo de perturbación, definida tal que $0 \leq \delta_j \leq 1$. Cuando $\delta_j = 1 \forall j$ no hay perturbación en el modelo. Si $\delta_j = 0$ significa que la j -ésima observación fué excluida. El estimador de mínimos cuadrados para el modelo lineal, cuando el mismo modelo ha sido perturbado se escribe:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}} = (\mathbf{X}^T \boldsymbol{\Delta} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Delta} \mathbf{y}$$

en donde $\boldsymbol{\Delta} = \text{diag}(\delta_1, \delta_2, \dots, \delta_n)$ es la matriz diagonal de perturbaciones. La medida de influencia más conocida se basa en la región de confianza para el parámetro $\boldsymbol{\beta}$,

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T (\mathbf{X}^T \mathbf{X}) (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq p s^2 F_{p, (n-p)}(\alpha)$$

Se mencionan a continuación tres medidas de influencia, las cuales permiten medir la sensibilidad de un modelo estimado:

- **Distancia de Cook:** Excluyendo la i -ésima observación del modelo se define:

$$D_{(i)} = \frac{(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)})^T (\mathbf{X}^T \mathbf{X}) (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)})}{p s^2} = t_i^2 \frac{h_{ii}}{(1 - h_{ii})} \frac{1}{p}$$

- **DFFITs:** propuestos por Belsley, Kuh y Welsch (1980, [15]), se definen por:

$$DFFITs = \frac{|r_i|}{s_{(i)}(1 - h_{ii})^{1/2}} \left\{ \frac{h_{ii}}{(1 - h_{ii})} \right\}^{1/2} = |t_i^*| \left\{ \frac{h_{ii}}{(1 - h_{ii})} \right\}^{1/2}$$

- **Alternativa a los DFFITS:** Atkinson (1985, [16]) propone una alternativa a los DFFITS que se definen para la i -ésima observación como:

$$C_i = \left\{ \frac{(n-p)}{p} \frac{h_{ii}}{(1-h_{ii})} \right\}$$

1.1.10. Criterios de Información

La validez y calidad de un GLM se mide a través de la función *desvío* y los criterios de información.

- **Criterio de información de Akaike:** mide la calidad relativa del ajuste de un modelo estadístico a un conjunto de datos:

$$\mathbf{AIC} = 2k - 2\log(\mathbf{L})$$

en donde k es el número de parámetros del modelo y \mathbf{L} es el máximo valor de la función de verosimilitud. Se prefiere el modelo con menor **AIC**.

- **Criterio de información de Bayes:** criterio para la selección de un modelo, entre un conjunto finito de modelos,

$$\mathbf{BIC} = -2\ln(\mathbf{L}) + k\ln(n)$$

al tener dos modelos, se prefiere el que menor **BIC** tenga.

- **Criterio de información de Hannan - Quinn:** es un criterio alternativo al **AIC**, se define:

$$\mathbf{HQC} = n \log \left(\frac{\mathbf{RSS}}{n} \right) + 2k \log \log(n)$$

en donde **RSS** es la reducción en sumas de cuadrados del error del modelo estimado.

1.2. Modelos Aditivos Generalizados (GAMs)

Siguiendo a Rigby et.al (2005, [1]), las técnicas de suavización se hicieron muy populares a finales de los años 80s. Hastie y Tibshirani (1990, [17]) introdujeron los GLM en esa franja. Wood (2000, [18]) contribuyó enormemente a su teoría y desarrollo.

Un modelo aditivo generalizado GAM es definido por:

$$\mathbf{Y} \stackrel{\text{ind}}{\sim} \xi(\boldsymbol{\mu}, \boldsymbol{\phi})$$

$$\boldsymbol{\eta} = g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{s}_1(\mathbf{x}_1) + \mathbf{s}_2(\mathbf{x}_2) + \cdots + \mathbf{s}_J(\mathbf{x}_J)$$

donde \mathbf{s}_j es una función suave no paramétrica, aplicada a la covariable \mathbf{x}_j . La idea es llevar los datos a determinar la relación entre el predictor $\eta = g(\boldsymbol{\mu})$ y las variables explicativas, antes que forzarlos a una relación lineal ó polinomial. Los términos de suavizamiento introducen no linealidad en el modelo. La estimación de parámetros en este tipo de modelos es una versión *penalizada* de la estimación para GLMs.

1.2.1. Funciones suaves univariadas.

La mejor representación de una función suave univariada es a través del modelo:

$$y_i = f(x_i) + \epsilon_i \tag{1}$$

en donde y_i es una variable respuesta indexada para la i -ésima observación, x_i es una covariable. f es una función suave y los ϵ_i son variables aleatorias independientes e idénticamente distribuidas $N(0, \sigma^2)$. Además se supone que la variable independiente x_i está en el intervalo $[0, 1]$.

1.2.2. Regresión Spline.

Para estimar la función f , se requiere que esta función sea representada de tal manera que el modelo (1) siga siendo un modelo lineal en los parámetros. Esto puede ser posible escogiendo una *base*, definiendo a su vez un espacio de funciones de la cual f sea un elemento. Al escoger una de tales bases, cantidades de elementos de la misma pueden ser escogidas. Por ejemplo si $b_i(x)$ es la base i -ésima de algún espacio, f se puede representar por:

$$f(x) = \sum_{i=1}^q b_i(x) \beta_i \tag{2}$$

A manera de ejemplo si f es un polinomio de grado 4 y $b_1(x) = 1, b_2(x) = x, b_3(x) = x^2, \dots, b_5(x) = x^4$. El modelo (1) queda:

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \dots + \beta_5 x_i^4 + \epsilon_i$$

Es de aclarar que las funciones polinómicas son muy útiles en situaciones en las que el interés se enfoca en las propiedades de f en la vecindad de un punto especificado. pero cuando las preguntas se relacionan al dominio completo (específicamente el intervalo $[0, 1]$, las bases polinómicas pueden tener algunos problemas. Las bases *spline* se comportan muy bien en tales circunstancias, a causa de que tienen buenas aproximaciones teóricas. En este orden de ideas un *spline cúbico* es una curva hecha de secciones de polinomios cúbicos, juntados o conectados de manera que sean continuos y que se puedan evaluar la primera y segunda derivada de los mismos. Los puntos en los cuales las secciones se

juntan se denominan *knots* del spline. Típicamente estos *knots* deberían ser equiespaciados, a través del rango de los valores observados de x , ó cuantilas de la distribución de la variable x como si fuera una sola variable. Sean los nodos (knots) de localización notados por $\{x_i^* : i = 1, 2, \dots, q-2\}$.

Existen muchas formas de escribir una base para splines cúbicos, sinembargo la aproximación más general se puede encontrar en los libros de Wahba (1990, [24]) y Gu (2002, [25]) (citados por Wood). Para estas bases $b_1(x) = 1$, $b_2(x) = x$ y $b_{i+2}(x) = R(x, x_i^*)$ para $i = 1, \dots, q-2$.

$$R(x, z) = [(z - \frac{1}{2})^2 - \frac{1}{12}] [(x - \frac{1}{2})^2 - \frac{1}{12}]/4 - [(|x - z| - \frac{1}{2})^4 - \frac{1}{2}((|x - z| - \frac{1}{2})^2 + \frac{7}{240})]/24 \quad (3)$$

Al usar splines cúbicos para f , el modelo (1) en forma matricial se escribe $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde la i -ésima fila de la matriz \mathbf{X} se puede escribir:

$$\mathbf{X}_i = [1, x_i, R(x_i, x_1^*), R(x_i, x_2^*), \dots, R(x_i, x_{q-2}^*)]$$

En consecuencia el modelo puede ser estimado por mínimos cuadrados.

1.2.3. Grado de suavizamiento mediante regresión spline penalizada.

Para controlar el suavizamiento alterando la dimensión de las bases, consiste en mantener fija la dimensión de la base tanto como sea razonable, pero controlando la suavidad del modelo añadiendo un ondulamiento penalizado a la función objetivo de *mínimos cuadrados*. Se pretende minimizar

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \int_0^1 [f''(x)]^2 dx$$

donde la integral de la segunda derivada al cuadrado penaliza el modelo. El parámetro λ es el parámetro de suavizamiento, cuando $\lambda \rightarrow \infty$ se lleva a una línea recta que estima f , cuando $\lambda = 0$ resulta una regresión spline no penalizada.

Como f es lineal en los parámetros, β_i , la función penalty puede ser escrita como una forma cuadrática:

$$\int_0^1 [f''(x)]^2 = \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta}$$

donde \mathbf{S} es una matriz de constantes conocida, un elemento de esta matriz se puede escribir $S_{i+2, j+2} = R(x_i^* x_j^*)$ para $i, j = 1, \dots, q-2$ mientras que las dos primeras filas y columnas de \mathbf{S} son cero. De esta forma la regresión spline penalizada consiste en minimizar:

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta} \quad (4)$$

Para la estimación del parámetro λ se hace necesario estimar primero el parámetro $\boldsymbol{\beta}$. Así después de un ejercicio algebraico sencillo pero engorroso se llega al estimador de mínimos cuadrados penalizados de $\boldsymbol{\beta}$.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T \mathbf{y} \quad (5)$$

Para este tipo de modelos la matriz de influencia, ó matriz hat, \mathbf{A} , para el modelo es:

$$\mathbf{A} = \mathbf{X} (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T$$

Se debe recordar que $\hat{\boldsymbol{\mu}} = \mathbf{A} \mathbf{y}$ y que la fórmula (4) también se expresa:

$$\left\| \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{X} \\ \sqrt{\lambda} \mathbf{B} \end{bmatrix} \boldsymbol{\beta} \right\|^2 = \|\mathbf{y} - \mathbf{X} \boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta}$$

donde \mathbf{B} es cualquier raíz cuadrada de la matriz \mathbf{S} tal que $\mathbf{B}^T \mathbf{B} = \mathbf{S}$. Esta matriz \mathbf{B} puede ser obtenida por descomposición espectral ó usando la descomposición de Cholesky.

1.2.4. Validación cruzada para la obtención del parámetro de suavización λ .

Si λ es muy grande los datos serán sobre suavizados, si por el contrario λ es pequeño los datos serán poco o nada suavizados. En los dos casos la estimación \hat{f} no será cerrada a la verdadera función f .

Un criterio para escoger λ puede ser minimizar:

$$M = \frac{1}{n} \sum_{i=1}^n (\hat{f}_i - f_i)^2 \quad \text{con} \quad \hat{f}_i \equiv \hat{f}(x_i) \quad \text{y} \quad f_i \equiv f(x_i)$$

Sea $\hat{f}^{[-i]}$ el modelo ajustado para todos los datos excepto y_i , se define el puntaje de *validación cruzada* ordinario por:

$$\nu_o = \frac{1}{n} \sum_{i=1}^n (\hat{f}_i^{[-i]} - y_i)^2$$

En Wood(2000, [18], pp. 129) se demuestra que:

$$\mathbf{E}(\nu_o) = \frac{1}{n} \mathbf{E} \left(\sum_{i=1}^n (\hat{f}_i - f_i)^2 \right) + \sigma^2$$

Se puede demostrar que para grandes muestras $\mathbf{E}(\nu_o) \approx \mathbf{E}(M) + \sigma^2$. Resulta un poco engorroso realizar una regresión spline cada vez que se elimina un dato, esto lleva a utilizar *validación cruzada generalizada* representada por:

$$\nu_g = \frac{n \sum_{i=1}^n (y_i - \hat{f}_i)^2}{[\text{tr}(\mathbf{I} - \mathbf{A})]^2}$$

La *validación cruzada generalizada* cual tiene ventajas computacionales sobre la *validación cruzada ordinaria* en términos de invarianza, ver Wahba (1990, [24], pp. 53).

1.3. Modelos Aditivos de Localización, Escala y Forma (GAMLSS)

Los Modelos Aditivos Generalizados de Localización, Escala y Forma fueron propuestos inicialmente por Rigby y Stasinopoulos (2005). Se definen:

$$\begin{aligned}
\mathbf{Y} &\overset{\text{ind}}{\sim} \mathbf{D}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \\
\boldsymbol{\eta}_1 &= g(\boldsymbol{\mu}) = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{s}_{11}(\mathbf{x}_{11}) + \mathbf{s}_{12}(\mathbf{x}_{12}) + \cdots, + \mathbf{s}_{1J_1}(\mathbf{x}_{1J_1}) \\
\boldsymbol{\eta}_2 &= g(\boldsymbol{\sigma}) = \mathbf{X}_2 \boldsymbol{\beta}_2 + \mathbf{s}_{21}(\mathbf{x}_{21}) + \mathbf{s}_{22}(\mathbf{x}_{22}) + \cdots, + \mathbf{s}_{2J_2}(\mathbf{x}_{2J_2}) \\
\boldsymbol{\eta}_3 &= g(\boldsymbol{\nu}) = \mathbf{X}_3 \boldsymbol{\beta}_3 + \mathbf{s}_{31}(\mathbf{x}_{31}) + \mathbf{s}_{32}(\mathbf{x}_{32}) + \cdots, + \mathbf{s}_{3J_3}(\mathbf{x}_{3J_3}) \\
\boldsymbol{\eta}_4 &= g(\boldsymbol{\tau}) = \mathbf{X}_4 \boldsymbol{\beta}_4 + \mathbf{s}_{41}(\mathbf{x}_{41}) + \mathbf{s}_{42}(\mathbf{x}_{42}) + \cdots, + \mathbf{s}_{4J_4}(\mathbf{x}_{4J_4})
\end{aligned}$$

Los parámetros $\boldsymbol{\nu}$ y $\boldsymbol{\tau}$ hacen referencia a la asimetría y curtosis de la distribución de los datos representados por la variable respuesta \mathbf{Y} . La estimación de parámetros en este tipo de modelos es similar a la de los GAMs.

1.3.1. Ventajas de los GAMLSS

- Crear una distribución *nueva* es relativamente *fácil*.
- Cualquier distribución puede ser *truncada* a derecha ó a izquierda.
- Una versión *censurada* de cualquier distribución puede ser creada.
- Cualquier distribución puede ser mezclada para crear mezclas finitas.
- Distribuciones continuas *discretizadas* pueden ser creadas para modelar variables respuesta.
- Cualquier distribución continua en el intervalo $(-\infty, \infty)$ puede ser transformada a una distribución en el intervalo $(0, \infty)$ ó el intervalo $(0, 1)$.

1.3.2. Inferencia en GAMLSS

Se debe tener en cuenta que para un modelo GAMLSS, donde $D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$ es la distribución de la variable respuesta \mathbf{Y} , \mathbf{X}_k son las matrices de diseño, incorporando los términos aditivos lineales en el modelo, $\boldsymbol{\beta}_k$ son los parámetros lineales y $s_{kj}(\mathbf{x}_{kj})$ representan las funciones suaves para las variables explicativas \mathbf{x}_{kj} , con $k = 1, 2, 3, 4$ y $j = 1, 2, \dots, J_k$. Nótese que las variables \mathbf{X} 's pueden ser las mismas ó diferentes de las \mathbf{x}_{kj} en los suavizadores. Los vectores $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_4$ son llamados los predictores de $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}$ respectivamente.

En lugar de las funciones suaves usadas dentro de los GAMLSS, se pueden escribir $s(\mathbf{x}) = \mathbf{Z}\boldsymbol{\gamma}$, donde \mathbf{Z} es la matriz base que depende de la variable explicativa \mathbf{x} y $\boldsymbol{\gamma}$ es un vector de parámetros a ser estimado utilizando la penalización cuadrática $\boldsymbol{\lambda}\boldsymbol{\gamma}^T \mathbf{G}\boldsymbol{\gamma}$, para una matriz conocida $\mathbf{G} = \mathbf{D}^T \mathbf{D}$ y donde el hiperparámetro $\boldsymbol{\lambda}$ regula la cantidad de suavizamiento necesitada para el ajuste. A veces se prefiere referirse a las funciones como funciones *suaves penalizadas* ó *suavizadores penalizados*. En Rigby,

Stasinopoulos y otros (2017, [1]), capítulo 9 se muestran distintas formulaciones de las matrices Z y D que llevan a diferentes formas de suavizamiento con variedad de propiedades estadísticas.

$$\begin{aligned}
\mathbf{Y}|\boldsymbol{\gamma} &\stackrel{\text{ind}}{\sim} \mathbf{D}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \\
\boldsymbol{\eta}_1 = g(\boldsymbol{\mu}) &= \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{Z}_{11} \boldsymbol{\gamma}_{11} + \mathbf{Z}_{12} \boldsymbol{\gamma}_{12} + \cdots, + \mathbf{Z}_{1J_1} \boldsymbol{\gamma}_{1J_1} \\
\boldsymbol{\eta}_2 = g(\boldsymbol{\sigma}) &= \mathbf{X}_2 \boldsymbol{\beta}_2 + \mathbf{Z}_{21} \boldsymbol{\gamma}_{21} + \mathbf{Z}_{22} \boldsymbol{\gamma}_{22} + \cdots, + \mathbf{Z}_{2J_2} \boldsymbol{\gamma}_{2J_2} \\
\boldsymbol{\eta}_3 = g(\boldsymbol{\nu}) &= \mathbf{X}_3 \boldsymbol{\beta}_3 + \mathbf{Z}_{31} \boldsymbol{\gamma}_{31} + \mathbf{Z}_{32} \boldsymbol{\gamma}_{32} + \cdots, + \mathbf{Z}_{3J_3} \boldsymbol{\gamma}_{3J_3} \\
\boldsymbol{\eta} = g(\boldsymbol{\tau}) &= \mathbf{X}_4 \boldsymbol{\beta}_4 + \mathbf{Z}_{41} \boldsymbol{\gamma}_{41} + \mathbf{Z}_{42} \boldsymbol{\gamma}_{42} + \cdots, + \mathbf{Z}_{4J_4} \boldsymbol{\gamma}_{4J_4}
\end{aligned} \tag{6}$$

Los *betas*, son los parámetros de efectos fijos:

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \boldsymbol{\beta}_3^T, \boldsymbol{\beta}_4^T)^T$$

Y los *gammas* son los parámetros de efectos aleatorios:

$$\boldsymbol{\gamma} = (\boldsymbol{\gamma}_{11}^T, \dots, \boldsymbol{\gamma}_{1J_1}^T; \cdots; \boldsymbol{\gamma}_{41}^T, \dots, \boldsymbol{\gamma}_{4J_4}^T)^T$$

Se asume además que los $\boldsymbol{\gamma}'_{kj}$ s son independientes dos a dos con distribución a priori:

$$\boldsymbol{\gamma}_{kj} \sim \mathbf{N}(\mathbf{0}, [\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})]^{-1}) \tag{7}$$

donde $[\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})]^{-1}$ es la inversa generalizada de una matriz simétrica $q_{kj} \times q_{kj}$, la cual depende del vector de hiperparámetros $\boldsymbol{\lambda}_{kj}$. Si $\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})$ es singular, se entiende que $\boldsymbol{\lambda}_{kj}$ tiene una densidad proporcional a $\exp(-\frac{1}{2} \boldsymbol{\gamma}^T \mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj}) \boldsymbol{\gamma})$. Un importante caso en la ecuación (7) es cuando $\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj}) = \boldsymbol{\lambda}_{kj} \mathbf{G}_{kj}$ y \mathbf{G}_{kj} es una matriz conocida para todo k, j . Si no existen efectos aleatorios en el modelo (6), éste se reduce a:

$$\begin{aligned}
\mathbf{Y} &\stackrel{\text{ind}}{\sim} \mathbf{D}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \\
\boldsymbol{\eta}_1 = g(\boldsymbol{\mu}) &= \mathbf{X}_1 \boldsymbol{\beta}_1 \\
\boldsymbol{\eta}_2 = g(\boldsymbol{\sigma}) &= \mathbf{X}_2 \boldsymbol{\beta}_2 \\
\boldsymbol{\eta}_3 = g(\boldsymbol{\nu}) &= \mathbf{X}_3 \boldsymbol{\beta}_3 \\
\boldsymbol{\eta}_4 = g(\boldsymbol{\tau}) &= \mathbf{X}_4 \boldsymbol{\beta}_4
\end{aligned} \tag{8}$$

El modelo (8) hace referencia a un modelo GAMLSS *paramétrico*, mientras que el modelo (6) hace referencia a un modelo GAMLSS de *efectos aleatorios*.

Ajustar el modelo GAMLSS paramétrico (8) requiere sólo el ajuste de los *betas*, mientras que el ajuste del modelo GAMLSS de efectos aleatorios (6) requiere el conocimiento de los parámetros $\boldsymbol{\beta}'_s$, $\boldsymbol{\gamma}'_s$ y $\boldsymbol{\lambda}'_s$ respectivamente.

Dentro de la metodología de modelos GAMLSS, el ajuste del modelo (8) se realiza a través de *máxima verosimilitud*, en tanto que el ajuste del modelo (6) se realiza por medio de *máxima verosimilitud penalizada*, el cual es equivalente a moda posterior ó *máximo a posteriori* (MAP), con respecto a $\boldsymbol{\beta}$ y $\boldsymbol{\gamma}$ para $\boldsymbol{\lambda}$ fijo. El logaritmo de verosimilitud para el modelo (8), bajo el supuesto de independencia de las observaciones es:

$$l = \sum_{i=1}^n \log f(y_i | \mu_i, \sigma_i, \nu_i, \tau_i) \tag{9}$$

Y la función logaritmo de verosimilitud para el modelo (6) es

$$l_p = l - \frac{1}{2} \sum_{k=1}^4 \sum_{j=1}^{J_k} \gamma_{kj}^T \mathbf{G}_{kj}(\lambda_{kj}) \gamma_{kj} \quad (10)$$

Los dos algoritmos básicos para el ajuste del modelo paramétrico con respecto a β y el modelo *no paramétrico*, con respecto a β, γ para λ fijo son el RS y el CG. Estos producen estimaciones de máxima verosimilitud en el modelo GAMLSS paramétrico y estimaciones MAP para β, γ con λ fijo.

1.3.3. Estimando β y γ con λ fijo

Rigby y Stasinopoulos (2005, [3]) proveyeron dos algoritmos para la estimación de β y γ para λ fijo.

- El algoritmo RS es una generalización del algoritmo usado años antes por Rigby y Stasinopoulos, para calcular media y dispersión en modelos aditivos (MADAM). La característica esencial es que este algoritmo no usa derivadas cruzadas en el logaritmo de verosimilitud.
- El algoritmo CG es una generalización del algoritmo de Cole y Green (1992). Este algoritmo requiere información de la primera y segunda derivadas cruzadas, de la función de verosimilitud, con respecto a la distribución de los parámetros μ, σ, ν y τ .

Rigby y Stasinopoulos demuestran que los dos algoritmos, para un λ fijo, llevan a estimaciones de máxima verosimilitud penalizadas de los *betas* y los *gammas*, $\hat{\beta}$ y $\hat{\gamma}$.

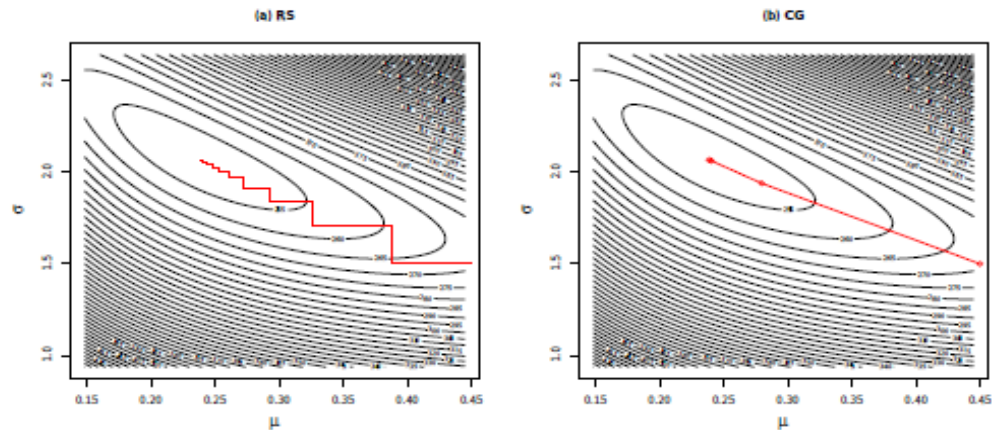


Figura 1: Como los algoritmos RS y CG alcanzan las estimaciones de máxima verosimilitud, modelo paramétrico. Fuente: Rigby y Stasinopoulos (2017, [1], pp. 63).

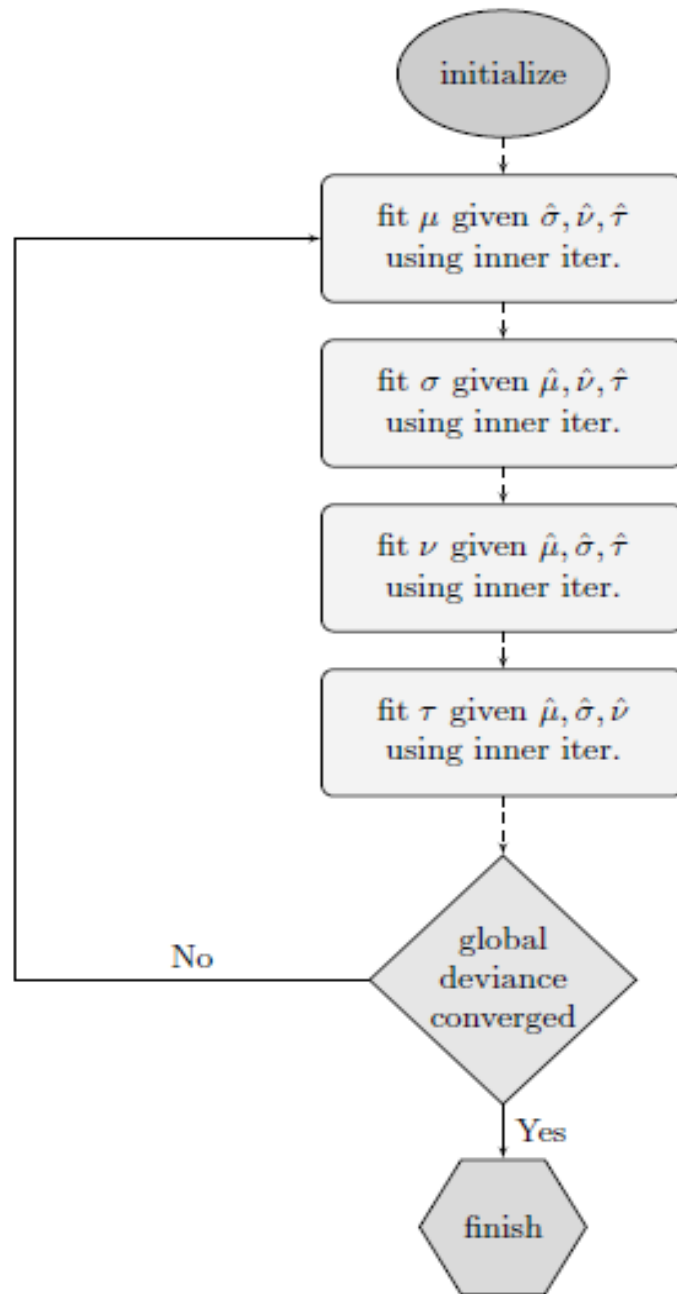


Figura 2: La iteración outer dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 65).

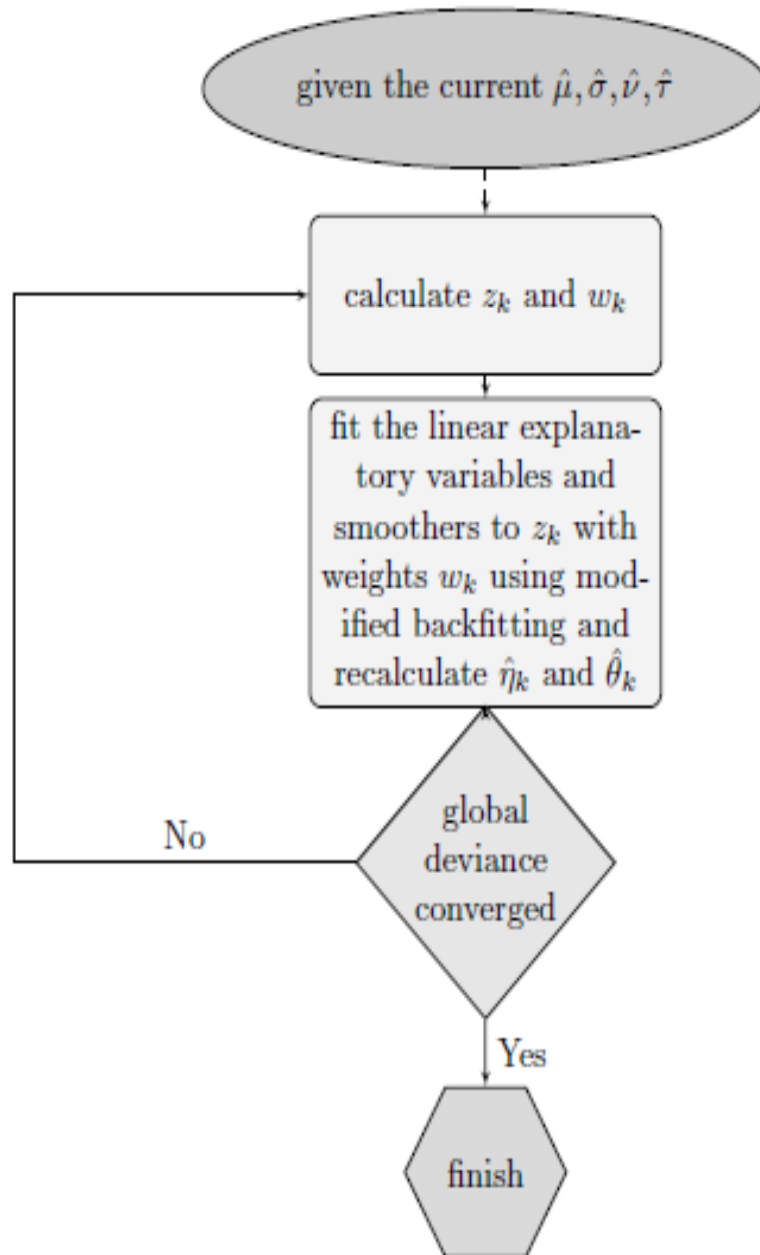


Figura 3: La iteración *inner* (o iteración GLIM) dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1]. pp. 67).

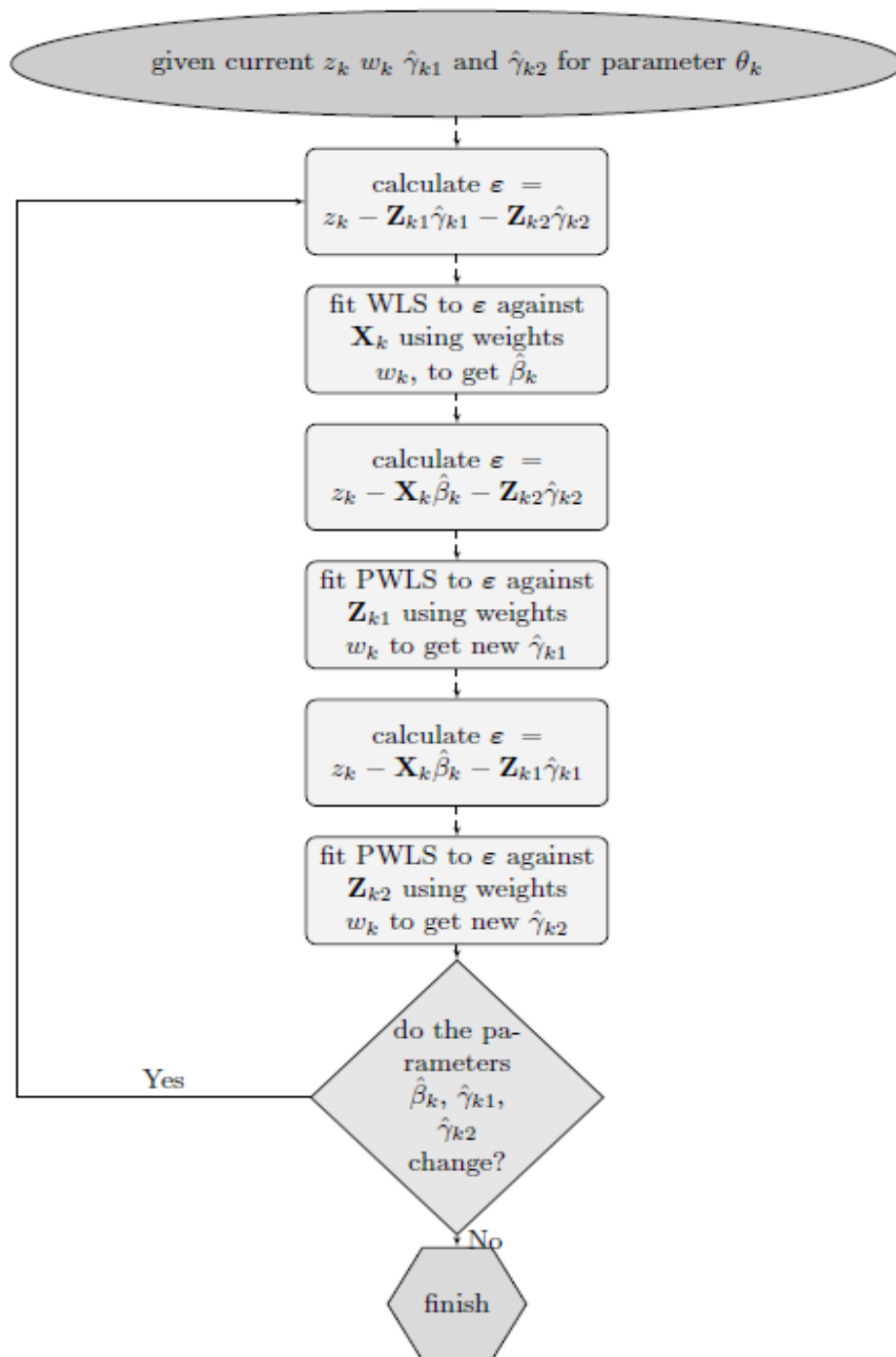


Figura 4: El algoritmo modificado backfitting dentro del algoritmo RS de GAMLSS. Fuente: Rigby y Stasinopoulos (2017, [1], pp. 69).

1.3.4. El algoritmo RS

Este algoritmo puede ser descrito siguiendo tres componentes anidadas

- La *iteración outer*, representada en la figura 2.
- La *iteración inner*, representada en la figura 3.
- El *algoritmo modificado backfitting*, representado en la figura 4.

La iteración *outer* llama en forma repetida la iteración *inner*, la cual a su vez llama repetidamente el *algoritmo backfitting*. La convergencia ocurre cuando los tres algoritmos han convergido.

1.3.5. La iteración outer (iteración GAMLSS)

La figura 2 representa diagramáticamente la iteración *outer*. Después de la inicialización del vector de parámetros $(\hat{\mu}, \hat{\sigma}, \hat{\nu}, \hat{\tau}) = (\mu_0, \sigma_0, \nu_0, \tau_0)$, la iteración *outer* procede como sigue:

1. Ajuste para μ (es decir maximice la función logaritmo de verosimilitud penalizada sobre μ) dadas las últimas estimaciones de $\hat{\sigma}$, $\hat{\nu}$ y $\hat{\tau}$, luego
2. Ajuste un modelo para σ , dadas las últimas estimaciones de $\hat{\mu}$, $\hat{\nu}$ y $\hat{\tau}$, luego
3. Ajuste un modelo para ν , dadas las últimas estimaciones de $\hat{\mu}$, $\hat{\sigma}$ y $\hat{\tau}$, y finalmente
4. Ajuste un modelo para τ , dadas las últimas estimaciones de $\hat{\mu}$, $\hat{\sigma}$ y $\hat{\nu}$.

Si luego de calcular el *desvío global*, éste converge, detenga el proceso, de lo contrario continúe.

1.3.6. La iteración inner (iteración GLM o GLIM)

Se prefiere utilizar la notación $\theta_1 = \mu$, $\theta_2 = \sigma$, $\theta_3 = \nu$ y $\theta_4 = \tau$. Para cada ajuste de una distribución paramétrica θ_k , la iteración *inner* es utilizada. Este es un algoritmo *local scoring*, similar al utilizado para el ajuste de un Modelo Lineal Generalizado (GLM), de ahí su nombre GLIM. GLIM fué un paquete perteneciente a la *Royal Statistical Society* disponible para ajustar GLMs. Esto indica que la primera versión de **gamlss** fué escrita a finales de los años 1990s en GLIM, la cual ya es una especie en vía de extinción.

La variable respuesta modificada para ajustar el parámetro $\boldsymbol{\theta}_k$ está dada por:

$$\mathbf{Z}_k = \boldsymbol{\eta}_k + \mathbf{W}_k^{-1} \circ \mathbf{U}_k \quad (11)$$

en donde:

- $\mathbf{Z}_k, \boldsymbol{\eta}_k, \mathbf{W}_k, \mathbf{U}_k$ son todos los vectores de longitud n . Por ejemplo el vector de pesos $\mathbf{W}_k = (w_{k1}, w_{k2}, \dots, w_{kn})^T$,
- $\mathbf{W}_k^{-1} \circ \mathbf{U}_k = (w_{k1}^{-1} u_{k1}, w_{k2}^{-1} u_{k2}, \dots, w_{kn}^{-1} u_{kn})$ es el producto Hadamar por elemento.
- $\boldsymbol{\eta}_k = g_k(\boldsymbol{\theta}_k)$ es el vector predictor del k -ésimo parámetro con $K = 1, 2, 3, 4$, correspondientes a los vectores $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}$ respectivamente.

$$\mathbf{U}_k = \frac{\partial l}{\partial \boldsymbol{\eta}_k} = \left(\frac{\partial l}{\partial \boldsymbol{\theta}_k} \right) \circ \left(\frac{d \boldsymbol{\theta}_k}{d \boldsymbol{\eta}_k} \right) \quad (12)$$

es la función *score* (vector de las primeras derivadas del logaritmo de verosimilitud con respecto al predictor). \mathbf{W}_k son los *pesos iterativos* para $k = 1, 2, 3, 4$, definidos como:

$$\mathbf{W}_k = -\mathbf{f}_k \circ \left(\frac{d \boldsymbol{\theta}_k}{d \boldsymbol{\eta}_k} \right) \circ \left(\frac{d \boldsymbol{\theta}_k}{d \boldsymbol{\eta}_k} \right) \quad (13)$$

donde las tres formas para definir \mathbf{f}_k dependen de la información disponible para la distribución específica.

$$\mathbf{f}_k = \begin{cases} \mathbb{E} \left(\frac{\partial^2 l}{\partial \boldsymbol{\theta}_k^2} \right) & \text{Si el valor esperado existe, llevando al algoritmo de Fisher Scoring.} \\ \frac{\partial^2 l}{\partial \boldsymbol{\theta}_k^2} & \text{Llevando al algoritmo estándar de Newton-Raphson Scoring.} \\ - \left(\frac{\partial l}{\partial \boldsymbol{\theta}_k} \right) \circ \left(\frac{\partial l}{\partial \boldsymbol{\theta}_k} \right) & \text{Llevando al algoritmo quasi-Newton Scoring.} \end{cases}$$

1.3.7. El algoritmo modificado backfitting

Este algoritmo trabaja como sigue: se desean ajustar las variables explicativas y los suavizadores para \mathbf{Z}_k , con pesos de trabajo \mathbf{W}_k . La figura 4 demuestra como el proceso trabaja dentro del algoritmo RS. \mathbf{X}_k representa la matriz de diseño para la parte lineal del modelo con coeficientes $\boldsymbol{\beta}_k$, y por simplicidad se asumen solamente dos suavizadores, con parámetros γ_{k1} y γ_{k2} , con matrices base \mathbf{Z}_{k1} y \mathbf{Z}_{k2} respectivamente.

Para pesos iterativos dados \mathbf{W}_k , se trabaja la variable respuesta \mathbf{Z}_k y valores previamente inicializados o estimados para los coeficientes de los suavizadores $\hat{\gamma}_{k1}$ y $\hat{\gamma}_{k2}$, se calculan los residuales parciales $\boldsymbol{\epsilon}$ para los parámetros $\boldsymbol{\beta}_k$, y se ajusta un estimador por mínimos cuadrados ponderados para obtener un nuevo valor de $\hat{\boldsymbol{\beta}}$. Luego se obtienen los residuales parciales con respecto al primer suavizador, y se usan mínimos cuadrados reponderados para obtener un nuevo estimador $\hat{\gamma}_{k1}$. Posteriormente se usan los residuales con respecto al segundo suavizador y se vuelven a utilizar mínimos cuadrados reponderados para obtener $\hat{\gamma}_{k2}$. Se repite este proceso hasta que $\hat{\boldsymbol{\beta}}, \hat{\gamma}_{k1}$ y $\hat{\gamma}_{k2}$ converjan.

1.4. Muestreo Aleatorio Estratificado por Afijación Óptima

Teniendo en cuenta que la base de información suministrada a través de la base de datos es demasiado grande, se hace necesario establecer una estrategia de muestreo para el análisis de la información. El *muestreo aleatorio estratificado por afijación óptima* es una metodología adecuada que permite abordar la problemática del bajo peso de los niños al nacer. Esta metodología se encuentra referenciada en Solanilla et.al (2010, [29]), Arthanari y Dodge (1981, [23]) y Jiménez y Lozano (2006, [27]). Sea $U = \{U_1, U_2, \dots, U_N\}$ una población particionada en subpoblaciones o estratos. Una característica de la población se infiere a partir de muestras de cada uno de los estratos (Esto es una ganancia de precisión).

Sea N_i el número de unidades del estrato i -ésimo y $\sum_{i=1}^L N_i = N$, donde L es el número de estratos, N es el número de unidades en la población, mientras que n_i el tamaño de la muestra del i -ésimo estrato. Se asume que las muestras se toman independientemente en cada estrato.

El problema de afijación óptima es determinar los n_i , el objetivo es minimizar la varianza (ganancia de precisión) de la estimación de la característica de la población bajo estudio, la restricción es el número de muestras tomadas (presupuesto disponible).

Considérese la estimación insesgada de la media poblacional, \bar{Y} , donde Y es la característica bajo estudio. Sea \bar{y}_i la estimación insesgada de la media \bar{Y}_i (en el estrato i -ésimo), esto es,

$$\bar{y}_i = \frac{1}{n_i} \sum_{h=1}^{L} y_{ih},$$

sea \bar{y}_{est} dado por

$$\bar{y}_{est} = \frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i,$$

el cual es un estimador insesgado de la media poblacional \bar{Y} puesto que $E(\bar{y}_{est}) = \bar{Y}$, veamos lo siguiente:

$$\begin{aligned} E(\bar{y}_{est}) &= E \left[\frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i \right] = \frac{1}{N} \sum_{i=1}^L E(\bar{y}_i) N_i \\ &= \frac{1}{N} \sum_{i=1}^L \bar{Y}_i N_i = \frac{1}{N} \sum_{i=1}^L \sum_{h=1}^{N_i} y_{ih} \\ &= \frac{1}{N} \sum_{i=1}^L Y_i = \bar{Y} \end{aligned}$$

La precisión de esta estimación está medida por la varianza de la estimación muestral. Así considere-

mos $V(\bar{y}_{est})$, o sea, el promedio de $(\bar{y}_{est} - \bar{Y})^2$ sobre todas las posibles muestras. Ahora

$$\begin{aligned} (\bar{y}_{est} - \bar{Y})^2 &= \left[\frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i - \frac{1}{N} \sum_{i=1}^L N_i \bar{Y}_i \right]^2 \\ &= \frac{1}{N^2} \left[\sum_{i=1}^L N_i (\bar{y}_i - \bar{Y}_i) \right]^2 \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 (\bar{y}_i - \bar{Y}_i)^2 + \frac{2}{N^2} \sum_{i>j} N_i N_j (\bar{y}_i - \bar{Y}_i) (\bar{y}_j - \bar{Y}_j) \end{aligned}$$

promediando y teniendo en cuenta que los productos cruzados se anulan tenemos:

$$\begin{aligned} V(\bar{y}_{est}) &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 E(\bar{y}_i - \bar{Y}_i)^2 \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i V(\bar{y}_i) \end{aligned}$$

donde

$$V(\bar{y}_i) = \frac{1}{N_i - 1} \sum_{h=1}^{N_i} (y_{ih} - \bar{Y}_i)^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right)$$

Problema 1.1

$$\text{mín } V(\bar{y}_{est}) = \sum_{i=1}^L W_i^2 S_i^2 x_i \quad (\text{Función Objetivo})$$

$$\text{s. a. } \sum_{i=1}^L n_i = n \quad (\text{Función Lineal}) \tag{R1}$$

$$n_i \in \mathbb{Z}^+, 1 \leq n_i \leq N_i, 1 \leq i \leq L \tag{R2}$$

donde $W_i = \frac{N_i}{N}$, $S_i^2 = \frac{1}{N_i - 1} \sum_{h=1}^{N_i} (y_{ih} - \bar{Y}_i)^2$ y $x_i = \frac{1}{n_i} - \frac{1}{N_i}$.

La función objetivo no es lineal, si olvidamos (R2) podemos usar *Multiplicadores de Lagrange*. Sea

$$\begin{aligned} V(\bar{y}_i) &= \sum_{i=1}^L W_i^2 S_i^2 x_i = \sum_{i=1}^L a_i x_i \\ &= \sum_{i=1}^L a_i \left(\frac{1}{n_i} - \frac{1}{N_i} \right) = \sum_{i=1}^L \frac{a_i}{n_i} - \sum_{i=1}^L \frac{a_i}{N_i} \\ &= \sum_{i=1}^L \frac{a_i}{n_i} + \text{constante} \end{aligned}$$

donde $a_i = W_i^2 S_i^2$. Reescribiendo el problema 1.1 tenemos:

$$\begin{aligned} \text{mín} \quad Z &= \sum_{i=1}^L \frac{a_i}{n_i} \quad (\text{Función Objetivo}) \\ \text{s. a.} \quad \sum_{i=1}^L n_i &= n \end{aligned} \tag{R1}$$

$$n_i \in \mathbb{Z}^+, 1 \leq n_i \leq N_i, 1 \leq i \leq L \tag{R2}$$

Minimizar la $V(\bar{y}_{est})$ es equivalente a minimizar:

$$f(n_1, n_2, \dots, n_L) = \sum_{i=1}^L \frac{a_i}{n_i}.$$

Hallando ∇f tenemos:

$$\nabla f = \left(-\frac{a_1}{n_1^2}, -\frac{a_2}{n_2^2}, \dots, -\frac{a_L}{n_L^2} \right)$$

con restricción $g(n_1, n_2, \dots, n_L) = (\sum_{i=1}^L n_i - n) = 0$, hallando ∇g tenemos:

$$\nabla g = (1, 1, \dots, 1) = \mathbf{1}.$$

Nótese que $\nabla g \neq 0$. Por condición de lagrange $\nabla f - \lambda \nabla g = \left(-\frac{a_1}{n_1^2}, -\frac{a_2}{n_2^2}, \dots, -\frac{a_L}{n_L^2} \right) - \lambda \mathbf{1} = 0$, bajo la restricción $g(n_1, n_2, \dots, n_L) = 0$, tenemos:

$$\left. \begin{aligned} -\frac{a_1}{n_1^2} - \lambda &= 0 \\ -\frac{a_2}{n_2^2} - \lambda &= 0 \\ &\vdots \\ -\frac{a_L}{n_L^2} - \lambda &= 0 \end{aligned} \right\}$$

$$n_1 + n_2 + \dots + n_L = n$$

Considerando λ negativo y solucionando las ecuaciones tenemos:

$$\left. \begin{aligned} n_1 &= \sqrt{\frac{a_1}{\lambda}} \\ &\vdots \\ n_L &= \sqrt{\frac{a_L}{\lambda}} \end{aligned} \right\}$$

$$n = \frac{\sqrt{a_1} + \sqrt{a_2} + \dots + \sqrt{a_L}}{\sqrt{\lambda}}.$$

Despejando la solución se tiene:

$$\sqrt{\lambda} = \frac{\sum_{i=1}^L \sqrt{a_i}}{n}.$$

Respuesta:

$$n_i = \frac{\sqrt{a_i}}{\sqrt{\lambda}} = \frac{n\sqrt{a_i}}{\sum_{k=1}^L \sqrt{a_k}}$$

El valor n_i es el valor de la muestra para cada estrato i , con $i = 1, 2, \dots, L$. El valor n es el valor de la afijación óptima, es decir el tamaño de muestra sobre toda la población, la cual representa todos los estratos en la muestra, debido a que $n = n_1 + n_2 + \dots + n_L$. El valor n puede obtenerse utilizando inicialmente *muestreo aleatorio simple*.

La parte más interesante de esta técnica es que el procedimiento matemático para la obtención de la muestra está basado en *programación matemática*.

Se puede presentar los siguientes problemas:

- $n \notin \mathbb{Z}^+$, entonces, se redondea.
- $n_i > N_i$, entonces, aquí se tendría que redistribuir, ya que se presenta un sobremuestreo.
- $n_i < 1$, entonces, tendríamos que tomar como mínimo una unidad en cada estrato.

2. Análisis de Resultados

2.1. Selección de la muestra por Afijación Óptima

Utilizando la base de datos suministrada por la Secretaria de Salud Municipal de los niños nacidos en Ibagué y municipios aledaños en el año 2014. Se procedió a depurar dicha base. Para su análisis, se tomó como variable de referencia el nivel educativo de la madre, variable compuesta por 6 categorías.

1. Primaria incompleta.
2. Primaria completa.
3. Secundaria incompleta.
4. Secundaria completa.
5. Universitaria.
6. No sabe, no responde (N.S.N.R)

La base de datos (ver tabla A1 del Apéndice de Tablas.) se filtró teniendo en cuenta la variable peso, con el fin de estratificar la información y utilizar un muestreo por Afijación Óptima, $n_i = \frac{n\sqrt{a_i}}{\sum_{k=1}^L \sqrt{a_k}}$ donde $a_i = W_i^2 * S_i^2$ y $W_i = \frac{N_i}{N}$ es el peso del respectivo estrato¹.

Estrato	N_i	\bar{x}	s^2	W_i	a_i	n_i	muestra
N1	13927	3203	271387,3	0,7733	162320,6255	696,8761	697
N2	298	3282	229216	0,01654	62,7691	13,7038	14
N3	3384	3207	274254,451	0,1879	9684,6388	107,2199	171
N4	20	2982	277603,6842	0,0001	0,3424	1,0121	2
N5	115	3109	260649,603	0,0083	18,0845	7,3556	8
N6	229	3160	260167,069	0,0127	42,1689	11,2322	12
Total	18008						904

Tabla 2: Tamaños de muestra por estrato mediante afijación óptima

¹Solanilla et al (2010, [29])

En la última columna de la anterior tabla, se muestran los tamaños de muestra obtenidos para cada estrato. Con base en estos tamaños de muestra, después de calcular un filtro ordenado para cada estrato, se realizó una selección del tamaño de muestra en cada estrato utilizando muestreo aleatorio simple.

2.2. Selección del GAMLSS

Con base en el tamaño y las muestras seleccionadas de la base, se procedió a realizar el ajuste del Modelo Aditivo Generalizado de Localización, Escala y Forma, que permita representar en forma significativa los datos, en donde la variable respuesta Y de carácter binario (1: Bajo peso, 0: Peso normal), se asocia mediante un enlace *logit* con las variables independientes:

- Tiempo de gestación (tgest).
- Tipo de parto (tparto).
- Talla de nacimiento (talla).
- Número de consultas (nconsul)
- Multiplicidad de parto (mparto).

El análisis de este modelo arrojó los siguientes resultados :

	Estimate	Std. Error	t value	Pr(> t)	Signif.
(Intercept)	57.34	6.67	8.60	0.00	***
tgest	-0.87	0.13	-6.62	0.00	***
tparto	-1.46	0.52	-2.79	0.01	**
talla	-0.57	0.10	-5.58	0.00	***
mparto	1.61	0.75	2.15	0.03	*

Tabla 3: Estimación del modelo paramétrico

Se evidencia según la Tabla 3 que las variables paramétricas Tiempo de gestación, Tipo de parto, Talla y Multiplicidad de parto son altamente significativas, tomando como variable repuesta Y de carácter binario y utilizando como distribución para la variable respuesta la binomial.

Teniendo en cuenta que los GAMLSS incluyen variables independientes continuas, las cuales entran al modelo como funciones suaves, se modifica la inclusión de la variable independiente Talla a entrar al modelo como variable semiparamétrica, obteniendo los siguientes resultados:

	Estimate	Std. Error	t value	Pr(> t)	Signif.
(Intercept)	49.01	6.80	7.20	0.00	***
tgest	-0.88	0.14	-6.48	0.00	***
tparto	-1.61	0.55	-2.93	0.00	**
pb(talla)	-0.40	0.10	-3.85	0.00	***
nconsul	0.13	0.10	1.35	0.18	
mparto	1.50	0.76	1.98	0.05	*

Tabla 4: Estimación del modelo, usando Talla como variable semiparamétrica

Se observa en la Tabla 4 que no hubo cambios sustanciales con respecto al modelo anterior, también la variable número de consultas no es significativa. Por tanto se propone un nuevo modelo en el que a la media se le agregan las variables más significativas y la variable talla sigue ingresando al modelo de manera no paramétrica. Esto conduce a la obtención del mejor modelo.

	Estimate	Std. Error	t value	Pr(> t)	Signif.
(Intercept)	47.62	6.70	7.11	0.00	***
tgest	-0.82	0.13	-6.36	0.00	***
tparto	-1.38	0.52	-2.68	0.01	**
pb(talla)	-0.40	0.10	-3.83	0.00	***
mparto	1.48	0.74	2.00	0.05	*

Tabla 5: Estimación del mejor modelo, usando Talla como variable semiparamétrica

En la Tabla 5 la significancia no cambia, sin embargo se observa una pequeña mejora en los parámetros, lo que asegura menores criterios de información de Akaike, de Bayes y Desvío global. También en las gráficas de ajuste en el apéndice de gráficas se puede apreciar la calidad del ajuste de cada uno

de los tres modelos GAMLSS que han sido expuestos.

Ajustes	AIC	BIC	Desvío Global
Primer Ajuste	180.432	204.4606	170.432
Segundo Ajuste	179.9534	212.0476	166.5967
Tercer Ajuste	179.8068	207.8288	168.1449

Tabla 6: Criterios de Información de Akaike y de Bayes para los tres modelos

Variables	$\hat{\pi}_i$	$1 - \hat{\pi}_i$	Odds Ratio
tgest	0.70	0.30	2.34
tparto	0.81	0.19	4.26
talla	0.63	0.37	1.70
mparto	0.17	0.83	0.2049

Tabla 7: Probabilidad de éxito y fracaso para el Primer Ajuste.

En la tabla 7 se observa el aporte en probabilidad estimada del primer ajuste GAMLSS estimado, para cada una de las variables significativas del modelo ajustado.

Variables	$\hat{\pi}_i$	$1 - \hat{\pi}_i$	Odds Ratio
tgest	0.70	0.30	2.33
tparto	0.83	0.17	4.88
p(talla)	0.60	0.40	1.50
nconsul	0.47	0.53	0.88
mparto	0.19	0.81	0.23

Tabla 8: Probabilidad de éxito y fracaso para el Segundo Ajuste.

Se nota en la Tabla 8 que al incluir la variable talla en forma *no paramétrica*, el valor del parámetro mejora, no obstante la variable número de consultas no es significativa.

VARIABLES	$\hat{\pi}_i$	$1 - \hat{\pi}_i$	Odds Ratio
tgest	0.70	0.30	2.34
tparto	0.80	0.20	4.00
p(talla)	0.60	0.40	1.50
mparto	0.19	0.81	0.23

Tabla 9: Probabilidad de éxito y fracaso para el Tercer Ajuste.

En la tabla 9 se presenta el mejor modelo, una vez las variables más significativas se le incluyen a la media del modelo y se excluye la variable número de consultas.

3. Conclusiones

Con base en los análisis propuestos en el presente trabajo, se pueden escribir entre otras las siguientes conclusiones:

- El mejor modelo obtenido para explicar el problema del bajo peso al nacer de los neonatos en Ibagué durante el año 2014 es $\text{Log}\left(\frac{\pi}{1-\pi}\right) = 47,62 - 0,82 * tgest - 1,38 * tparto - 0,40 * pb(talla) + 1,48 * mparto$.
- Por cada unidad (semana) de tiempo de gestación menos que tenga la madre aumenta la probabilidad de que el niño recién nacido tenga bajo peso, lo cual representa 1,34, casi 2 individuos con bajo peso al nacer.
- Cada vez que una madre no tenga parto normal, la población de individuos con bajo peso puede aumentar en 3.
- La talla en centímetros de un recién nacido, si bien es una variable altamente significativa en el modelo, para el conjunto de datos disminuye ostensiblemente entre las 40 y 50 centímetros. Se evidencia una talla con alta dispersión entre los 30 y 40 cms. En promedio con poca dispersión se encuentran las tallas entre 45 y 50 cms. (ver gráfica B.3 del apéndice de)
- Con la variable multiplicidad de parto ocurre todo lo contrario, es decir por cada embarazo más que tenga una mujer, se ganaría casi un individuo para tener peso normal.

4. Recomendaciones

- Se recomienda explorar nuevos modelos GAMLSS, utilizando variables de tipo socioeconómico que en lo posible fuesen aleatorias, con el objeto de utilizar modelos GAMLSS mixtos .
- Examinar para la variable respuesta del modelo, otro tipo de distribuciones teóricas o construídas, tales como las mixtas ó jerárquicas.
- Tratar de implementar el manejo de Datos Funcionales, o Modelos Generalizados Funcionales a este tipo de problemas.

Apéndices

A. Apéndice de Tablas

A.1. Base de datos

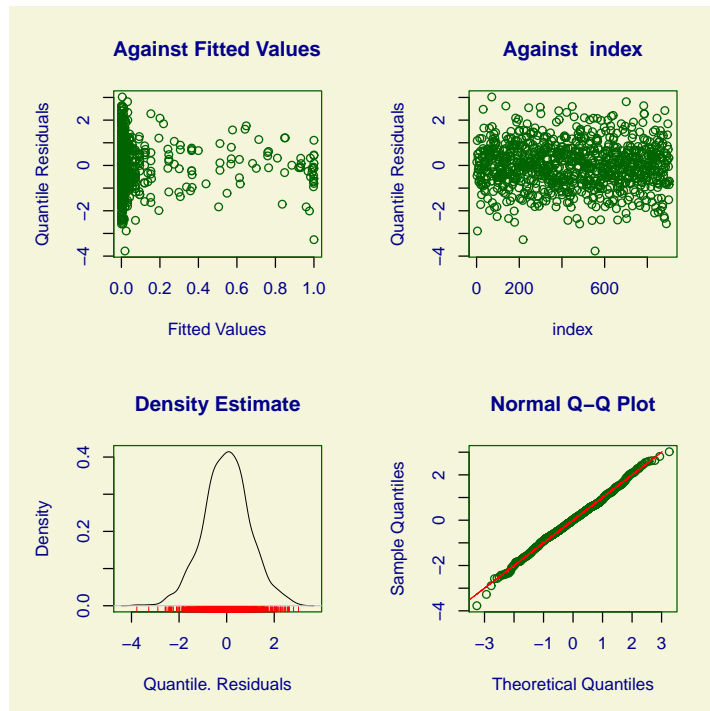
	sexo	peso	Y	talla	año	mes	atnparto	tgest	nconsul	tparto
1	1	3400	0	51	2014	10	1	39	9	1
2	2	3200	0	53	2014	7	1	39	4	1
3	2	3580	0	53	2014	3	1	40	5	1
4	1	3200	0	50	2014	5	1	38	6	1
5	2	3550	0	50	2014	7	1	38	6	2
6	2	2800	0	49	2014	5	1	39	8	1
7	2	2800	0	49	2014	9	1	39	2	1
8	2	3785	0	51	2014	11	1	39	7	2
9	2	3000	0	48	2014	12	1	38	7	1
10	2	4430	0	53	2014	7	1	40	9	2
11	1	3305	0	50	2014	1	1	40	7	1
12	2	3950	0	55	2014	9	1	40	6	1
13	2	3400	0	49	2014	1	1	39	6	1
14	1	2580	0	49	2014	8	1	39	7	1
15	2	2600	0	50	2014	1	1	39	5	1
16	2	3490	0	50	2014	3	1	39	6	2
17	1	3000	0	50	2014	10	1	38	8	1
18	1	4250	0	56	2014	7	1	41	6	1
19	2	2990	0	51	2014	12	1	39	6	1
20	1	4130	0	56	2014	11	1	39	4	1

A.2. Nombre de las variables

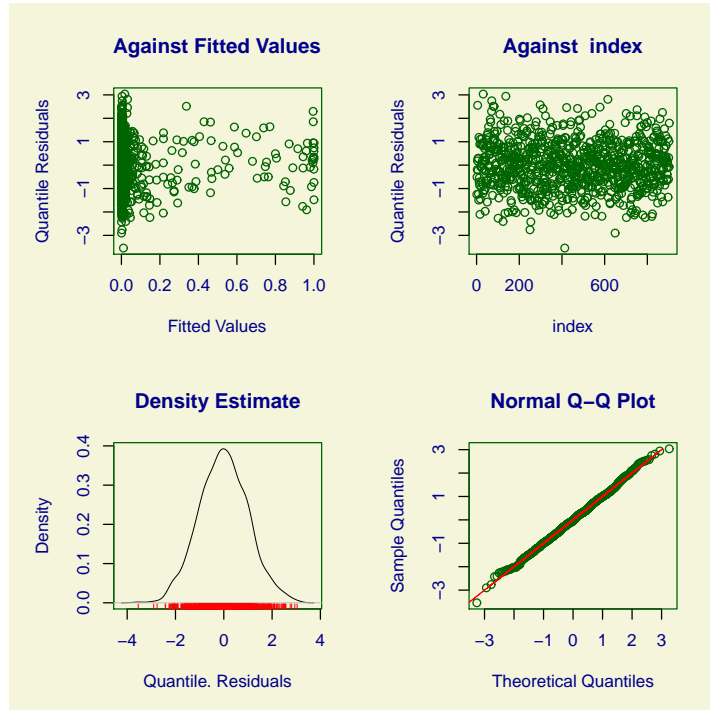
Variables	
Sexo	sexo
Peso	peso
Talla	talla
Año	Año
Mes	mes
Atención del parto	atnparto
Tiempo de Gestación	tgest
Número de consultas	nconsul
Tipo de Parto	tparto
Multiplicidad del Parto	mparto
Edad de la Madre	edadm
Estado Civil de la Madre	ecivilm
Nivel Educativo de la Madre	neducm
Lugar de residencia	lresi
Número de Hijos	nhijos
Número de Embarazos	nembar
Edad del Padre	edadp
Nivel Educativo del Padre	neducp

B. Apéndice de Gráficas

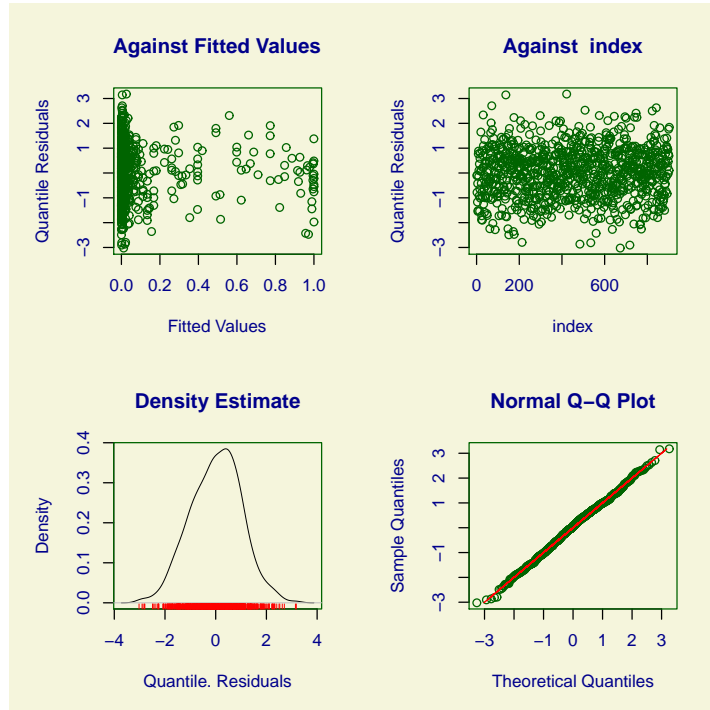
B.1. Gráfica de Ajuste Modelo 1



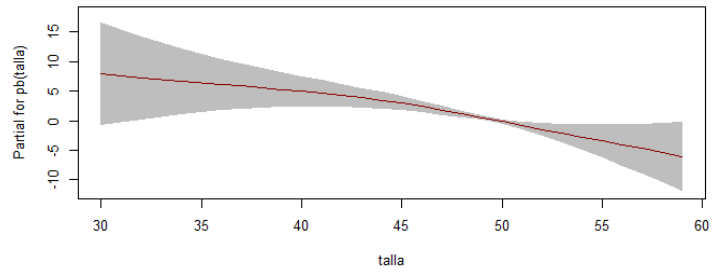
B.2. Gráfica de Ajuste Modelo 2



B.3. Gráfica de Ajuste Modelo 3




B.4. Gráfica de Variable semiparamétrica Talla



Bibliografia

- [1] STASINOPOULOS, M.K.; RIGBY, R.A.; HELLER, G.Z.; VOUDOURIS, V.; BASTIANI, F.D.: (2017) *Flexible Regression and Smoothing Using GAMLSS in R*. CRC Press, a Chapman & Hall Book, London.
- [2] EISENHART, C. (1961). *Boscovitch and the Combination of Observation*. In Whyte, LL(Ed). *Roger Joseph Boscovitch*, Fordham University Press, New York.
- [3] RIGBY, R.A. and STASINOPOULOS, D.M. «Generalized Additive Models for location, scale and shape», *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**(3), págs. 251–258, 2005.
- [4] NELDER, J.A, AND WEDDERBURN, R.W.M *Generalized Linear Models*. Journal of the Royal Statistical Society. Series A (General), Vol. 135, No. 3, pp. 370-384, 1972.
- [5] PAULA, G.A. *Modelos de Regressão com apoio computacional*. Versões 2004 é 2013. Universidad de São Paulo, S.P. 2013.
- [6] NELDER, J.A, AND MCCULLAGH, P. *Generalized Linear Models*. Chapman & Hall, London, 1989.
- [7] DOBSON, A. *An Introduction to Generalized Linear Models*. CRC Press Book, Third Edition, Sydney, 2008.
- [8] AGRESTI, A. *Foundations of Linear and Generalized Linear Models*. Jhon Wiley & Sons Inc., New York, 2015.
- [9] ARANDA-ORDAZ, F.J. (1981). On the two families of transformations to additivity for binary response data. *Biometrika*. **68**, 375-364.
- [10] SEN, P.K. AND SINGER, J.M. (1993). *Large Sample Methods in Statistics: An Introduction with Applications*. Chapman & Hall, London.
- [11] JORGENSEN, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society*. B **49**, 127-162.
- [12] COOK, R.D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society*. B **48**, 133-169.
- [13] COOK, R.D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society*. B **48**, 133-169.
- [14] RAO, C.R. (1973). *Linear Statistical Inference and Its Applications*. Second Edition. Wiley, New York.

- [15] BELSLEY, D.A.; KUH, E. AND WELSCH, R.E (1980). *Regression Diagnostics*. Second Edition. John Wiley, New York.
- [16] ATKINSON, A.C. (1985). *Plots, Transformations and Regressions*. Oxford Statistical Sciences Series. Oxford.
- [17] HASTIE, T. AND TIBSHIRANI, R. (1990). *Generalized Additive Models*. Chapman & Hall, London.
- [18] WOOD, S.N. (2000). *Generalized Additive Models: an introduction with R*. Chapman & Hall/CRC Texts in Statistical Science, London.
- [19] LEE, E.T. AND WANG, J.W. (2003). *Statistical Methods for Survival Data Analysis*. John Wiley & Sons Inc., New York.
- [20] KLEINBAUM, D. (1994). *Logistic Regression*. Springer-Verlag, New York.
- [21] LAMBERT, D. (1992). Zero-inflated Poisson Regression with an Applications to Defects Manufacturing. *Technometrics*. Vol. 34, 1, pp. 1-14.
- [22] LITTLE, R.J.A. AND RUBIN, D.B. (1987). *Statistical Analysis with Missing Data*. John Wiley & Sons Inc., New York.
- [23] ARTHANARI, T.S. AND DODGE, Y. (1981). *Mathematical Programming in Statistics*. John Wiley & Sons Inc., New York.
- [24] WAHBA, G. (1990). Spline Models for observational data. *Philadelphia: SIAM*.
- [25] GU, C. (2002). *Smoothing Spline ANOVA Models*. Springer, New York.
- [26] Hachuel, L. Boggio, G. y Wojdyla, D. (2004). Modelos Logit Mixtos: una aplicación en el área de la Salud. *Novenas Jornadas de Ciencias Económicas y Estadística*, Rosario, Argentina.
- [27] Jimenez, L. y Lozano, M. (2006). Un Modelo Lineal Generalizado para el bajo peso al nacer de los niños en Ibague 2000 - 2004. Tesis de pregrado en Matemáticas y Estadística. U.T.
- [28] Solanilla, L., Clavijo, J.A., Sánchez, A. and Zambrano, A.J. (2015). Cook's Local Influence in Generalized Linear Models via the Shape Operator. *Revista Comunicaciones en Estadística, US-TA*. Vol. 8, 1, pp. 99 - 110.
- [29] Solanilla, L., Clavijo, J.A., Sánchez, A. y Zambrano, A.J. (2010). Afijación Óptima de tamaños de muestra en muestreo aleatorio estratificado vía programación matemática. *Revista Comunicaciones en Estadística*. Vol 3, 1. pp. 7 - 23.
- [30] Suarez, H. (2014). Enfoque Bayesiano mediante Regresión Beta, para el bajo peso de los niños al nacer en Pitalito (Huila), 2010. Tesis de pregrado en Matemáticas con énfasis en Estadística, Universidad del Tolima.

	PROCEDIMIENTO DE FORMACIÓN DE USUARIOS AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONAL	Página 1 de 3
		Código: GB-P04-F03
		Versión: 03
		Fecha Aprobación: 15 de Febrero de 2017

Los suscritos:

<u>Piquel Armando Rodríguez Vargas</u>	con C.C.N°	<u>11332739</u>
<u>Alonso Sánchez Hernández</u>	con C.C.N°	<u>7218554</u>
<u>JAIRO A CIAVIJO</u>	con C.C.N°	<u>11331132</u>
_____	con C.C.N°	_____
_____	con C.C.N°	_____

Manifiesto (en) la voluntad de:

Autorizar

No Autorizar Motivo: _____


La consulta en físico y la virtualización de **mi OBRA**, con el fin de incluirlo en el repositorio institucional de la Universidad del Tolima. Esta autorización se hace sin ánimo de lucro, con fines académicos y no implica una cesión de derechos patrimoniales de autor.

Manifestamos que se trata de una OBRA original y como de la autoría de LA OBRA y en relación a la misma, declara que la UNIVERSIDAD DEL TOLIMA, se encuentra, en todo caso, libre de todo tipo de responsabilidad, sea civil, administrativa o penal (incluido el reclamo por plagio).

Por su parte la UNIVERSIDAD DEL TOLIMA se compromete a imponer las medidas necesarias que garanticen la conservación y custodia de la obra tanto en espacios físico como virtual, ajustándose para dicho fin a las normas fijadas en el Reglamento de Propiedad Intelectual de la Universidad, en la Ley 23 de 1982 y demás normas concordantes.

La publicación de:

Trabajo de grado	<input checked="" type="checkbox"/>	Artículo	<input type="checkbox"/>	Proyecto de Investigación	<input type="checkbox"/>
Libro	<input type="checkbox"/>	Parte de libro	<input type="checkbox"/>	Documento de conferencia	<input type="checkbox"/>
Patente	<input type="checkbox"/>	Informe técnico	<input type="checkbox"/>		<input type="checkbox"/>
Otro: (fotografía, mapa, radiografía, película, video, entre otros)					<input type="checkbox"/>

 Universidad del Tolima	PROCEDIMIENTO DE FORMACIÓN DE USUARIOS AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONAL	Página 2 de 3
		Código: GB-P04-F03
		Versión: 03
		Fecha Aprobación: 15 de Febrero de 2017

Producto de la actividad académica/científica/cultural en la Universidad del Tolima, para que con fines académicos e investigativos, muestre al mundo la producción intelectual de la Universidad del Tolima. Con todo, en mi condición de autor me reservo los derechos morales de la obra antes citada con arreglo al artículo 30 de la Ley 23 de 1982. En concordancia suscribo este documento en el momento mismo que hago entrega del trabajo final a la Biblioteca Rafael Parga Cortes de la Universidad del Tolima.

De conformidad con lo establecido en la Ley 23 de 1982 en los artículos 30 “**...Derechos Morales. El autor tendrá sobre su obra un derecho perpetuo, inalienable e irrenunciable**” y 37 “**...Es lícita la reproducción por cualquier medio, de una obra literaria o científica, ordenada u obtenida por el interesado en un solo ejemplar para su uso privado y sin fines de lucro**”. El artículo 11 de la Decisión Andina 351 de 1993, “**los derechos morales sobre el trabajo son propiedad de los autores**” y en su artículo 61 de la Constitución Política de Colombia.

- Identificación del documento:

Título completo: Estudio del bajo peso de los niños al nacer mediante GAMLSS año 2014

- Trabajo de grado presentado para optar al título de:

Profesional en Matemáticas con Énfasis en Estadística


- Proyecto de Investigación correspondiente al Programa (No diligenciar si es opción de grado “Trabajo de Grado”):

- Informe Técnico correspondiente al Programa (No diligenciar si es opción de grado “Trabajo de Grado”):

- Artículo publicado en revista:

- Capítulo publicado en libro:

- Conferencia a la que se presentó:

 Universidad del Tolima	PROCEDIMIENTO DE FORMACIÓN DE USUARIOS	Página 3 de 3
	AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONAL	Código: GB-P04-F03
		Versión: 03
		Fecha Aprobación: 15 de Febrero de 2017

Quienes a continuación autentican con su firma la autorización para la digitalización e inclusión en el repositorio digital de la Universidad del Tolima, el:

Día: 24 Mes: Octubre Año: 2018

Autores:

Firma

Nombre:	<u>Kathern Ramirez Rojas</u>	<u>Kathern Ramirez Rojas</u>	C.C. <u>1082775890</u>
Nombre:	_____	_____	C.C. _____
Nombre:	_____	_____	C.C. _____
Nombre:	_____	_____	C.C. _____

El autor y/o autores certifican que conocen las derivadas jurídicas que se generan en aplicación de los principios del derecho de autor.