# *AJUSTE DE POLINOMIOS FRACCIONARIOS A CURVAS DE CRECIMIENTO*

Libia Cristina Acosta C.

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística
Bogotá, D.C.
febrero de 2010

# *AJUSTE DE POLINOMIOS FRACCIONARIOS A CURVAS DE CRECIMIENTO*

Libia Cristina Acosta C.

Trabajo de tesis para optar al título de
Magister en Ciencias Estadística

Director
Luis Guillermo Díaz M.
Profesor Asociado Departamento de Estadística

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística
Bogotá, D.C.
febrero de 2010

**Título en español**

Ajuste de polinomios fraccionarios a curvas de crecimiento.

**Title in English**

fit the fractional polynomials to curves growth.

**Resumen:** En esta tesis se propone una metodología que permite modelar curvas de crecimiento a través de polinomios fraccionarios, se utiliza la transformación propuesta por Potthoff y Roy (1964) que permite medir el efecto de la curva de crecimiento a través del tiempo. Al utilizar polinomios fraccionarios dentro del modelo ajustado en presencia de valores extremos se soluciona el problema que se presenta al ajustar curvas de crecimiento con polinomios clásicos de bajo orden, pues ofrecen una familia limitada de formas y polinomios de alto orden, pueden ajustar pobremente los valores extremos de las covariables. Además al realizar ajustes con polinomios fraccionarios de grado uno o dos se encuentran una variedad de formas de curvas. Para el ajuste de polinomios fraccionarios a curvas de crecimiento se hace la estimación de los parámetros, se estudian algunas propiedades de dichas estimaciones y se comparan los ajustes obtenidos a través del AIC. En general en presencia de ruido en algunos perfiles, el modelamiento a través de polinomios fraccionarios es más apropiado por lo que se obtiene un mejor $AIC$. Además a través de la extrapolación en modelos truncados, el pronóstico hecho con polinomios fraccionarios resulta con un mejor $MAPE$ cuando se presenta ruido en algunos perfiles.

**Abstract:** This thesis proposes a methodology that allows fit of growth curves by polynomials fractional, using the transformation proposed by Potthoff and Roy (1964) to measure the effect of the growth curve over time. Using fractional polynomials in the fitted model in the presence of extreme values, avoid the problem that occurs when adjusting growth curves classical with low-order polynomials, as they offer a family Limited forms and higher-order polynomials, they can adjust poorly extreme values of covariates. In addition to fit the fractional polynomials of degree one or two offer a variety of shapes of curves. To fit fractional polynomials to curves of growth becomes the parameter estimation, the study some properties of and compared those estimates obtained through the AIC. In general in the presence of extreme data, the fit of growth curve with polynomials fractional is more appropriate because compare the curves fits, through the AIC, the fit with fractional polynomials has a better AIC. Also through extrapolation models truncated, the forecast made with polynomial fractional results with a better $MAPE$ when introduced noise in some profiles.

**Palabras clave:** Datos longitudinales, curvas de crecimiento, modelos de regresión, estimación máxima verosimilitud, polinomios fraccionarios.

**Keywords:** Longitudinal data, growth curves, models regression, maximum likelihood estimation, polynomials Fractional.

## Nota de aceptación

Trabajo de tesis
Aprobado

---

Jurado
Liliana López K.

---

Jurado
Edilberto Cepeda C.

---

Director
Luis Guillermo Díaz M.

Bogota, D.C., febrero 2 de 2010

# Agradecimientos

Durante la elaboración de esta tesis conté con la ayuda de muchas personas, a todas ellas mis más sinceros agradecimientos. A mis padres, por siempre estar ahí en todo momento brindándome la ayuda y el apoyo incondicional. A cada uno de mis profesores quienes a través de sus conocimientos y experiencia académica aportaron en cada momento en el desarrollo coherente de los diferentes contenidos de esta tesis. De manera muy especial agradezco al profesor Luis Guillermo Díaz M., por sus acertadas observaciones, su apoyo continuo y sus esfuerzos en procura de mejorar cada uno de los elementos que contiene esta tesis de Maestría en Estadística.

Expreso mi gratitud y reconocimento a la Universidad Nacional de Colombia, por brindarme la oportunidad de mejorar mi nivel académico en pro de un mayor nivel profesional.

Por último y no menos importante doy gracias a Dios por haberme permitido llegar hasta este punto en mi vida, luego de muchas dificultades en el camino, de esta manera finalizo esta etapa de la vida y empiezo una nueva.

# Índice general

**Índice general** **I**

**Índice de figuras** **IV**

**Introducción** **VI**

**1. Curvas de crecimiento y polinomios fraccionarios** **1**

1.1. Introducción . . . . . 1

1.2. Curvas de crecimiento (CC) . . . . . 1

1.2.1. **Modelos de curvas de crecimiento** . . . . . 2

1.2.2. **Estimación de los parámetros** . . . . . 4

1.2.3. **Estructuras de covarianzas** . . . . . 8

**Estructuras de covarianzas de los modelos de CC.** . . . . . 8

1.2.4. **Selección basada en la quasi-verosimilitud predictiva** . . . . . 12

1.2.5. **Prueba de hipótesis** . . . . . 12

1.2.6. Verificación de los supuestos del modelo . . . . . 16

1.3. Polinomios fraccionarios . . . . . 17

1.3.1. **Transformación de Box-Tidwell** . . . . . 18

1.3.2. **Definición y notación de polinomios fraccionarios** . . . . . 19

1.3.3. **Relación entre la transformación Box-Tidwell, las funciones exponenciales y los polinomios fraccionarios** . . . . . 19

1.3.4. **Selección de exponentes** . . . . . 22

1.3.5. **Características de los PF1 y PF2** . . . . . 22

Máximo o mínimo de una función de $PF_2$ . . . . . 23

1.3.6. **Criterios para la selección de modelos** . . . . . 23

**2. Modelo propuesto** **25**

2.1. Modelo propuesto . . . . . 25

2.1.1. Estimación de los parámetros . . . . . 27

2.1.2. Propiedades de los estimadores . . . . . 27

2.1.3. Pruebas de hipótesis . . . . . 29

2.1.4. Seleccion del modelo . . . . . 30

**3. Aplicación** **31**

3.1. Datos de glucosa . . . . . 31

3.1.1. Análisis exploratorio de los datos de glucosa para pacientes obesos . . 31

3.1.2. Verificación de los supuestos del modelo . . . . . 33

3.1.3. Selección de la matriz de varianzas y covarianzas . . . . . 34

3.1.4. Estimación de los parámetros . . . . . 37

Estimación por máxima verosimilitud con polinomios clásicos . . . . . 37

Estimación máxima verosimilitud con polinomios fraccionarios . . . . 38

3.2. Datos de Ratones . . . . . 41

3.2.1. Análisis exploratorio de los datos de ratones . . . . . 41

3.2.2. Verificación de los supuestos del modelo . . . . . 43

3.2.3. Selección de la matriz de varianzas y covarianzas . . . . . 43

3.2.4. Estimación de los parámetros . . . . . 44

Estimación por máxima verosimilitud con polinomios . . . . . 44

Estimación máxima verosimilitud con polinomios fraccionarios . . . . 45

**4. Simulación** **48**

4.1. Introduccción . . . . . 48

4.1.1. Escenarios . . . . . 49

4.1.2. Modelos estimados . . . . . 50

4.2. Resultados . . . . . 50

4.2.1. Escenario I: . . . . . 51

4.2.2. Escenario II: . . . . . 57

4.2.3. Escenario III . . . . . 62

4.2.4. Escenario IV: . . . . . 67

4.2.5. Escenario V . . . . . 72

**5. Conclusiones y recomendaciones** **77**

**A. Datos de glucosa para pacientes obesos** **78**

A.1. Datos de ratones . . . . . 79

**B. Programación en R** **80**

B.1. Obtención de la matriz **X** para ajustes de curvas de crecimiento con polinomios fraccionarios. . . . . 80

B.2. Estimación de los parámetros . . . . . 81

B.2.1. Estimación por mínimos cuadrados generalizados . . . . . 81

B.2.2. Estimación por máxima verosimilitud . . . . . 82

Estructuras de covarianza . . . . . 82

B.3. Obtención de las estadísticas de prueba . . . . . 85

**C. Modelos estimados** **88**

**Bibliografía** **90**

# Índice de figuras

1.1. Curvas para algunos $PF_2$ . . . . . 21

3.1. Gráfico de perfiles para datos de glucosa en pacientes obesos. . . . . 32
3.2. Diagrama de caja para datos de glucosa en pacientes obesos en diferentes tiempos de medición . . . . . 33
3.3. Dispersogramas para datos de glucosa en pacientes obesos para los diferentes tiempos de medición . . . . . 34
3.4. Ajustes por máxima verosimilitud con polinomios convencionales y fraccionarios, para los datos de glucosa. . . . . 39
3.5. Gráfico de perfiles para el peso en libras de 13 ratones machos medidos en intervalos de 3 días de nacidos hasta los 21 días. . . . . 41
3.6. Diagrama de caja para datos de ratones en intervaloes 3 días hasta los 21 días. . . . . 42
3.7. Dispersogramas para los diferentes pesos medidos a los 13 ratones . . . . . 43
3.8. Ajustes por máxima verosimilitud con polinomios y fraccionarios, para los datos de ratones. . . . . 46

4.1. Diagrama de caja de los AIC promedio para los mejores modelos de grado uno. . . . . 51
4.2. Diagrama de caja de los AIC promedio para los mejores modelos de grado dos. . . . . 52
4.3. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno. . . . . 54
4.4. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados estimados de grado dos. . . . . 55
4.5. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y efecto campana. . . . . 57
4.6. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos estimados de grado dos y efecto campana. . . . . 58

4.7. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos estimados truncados de grado uno y dos y efecto campana. 60

4.8. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y efecto campana . . . . . 60

4.9. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y ruido en algunos perfiles . . . . . 62

4.10. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores estimados de grado dos y ruido en algunos perfiles . . . . . 63

4.11. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y ruido en algunos perfiles . . 65

4.12. Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y ruido en algunos perfiles . . 65

4.13. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y sin ruido . . . . . 67

4.14. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado dos y sin ruido . . . . . 68

4.15. Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado uno y sin ruido . . . . . 70

4.16. Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado dos y sin ruido . . . . . 70

4.17. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y ruido en algunos perfiles . . . . . 72

4.18. Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado dos y con ruido en algunos perfiles . . . . . 73

4.19. Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado uno y ruido en algunos perfiles . . . . . 75

4.20. Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado dos y ruido en algunos perfiles . . . . . 75

# Introducción

En el análisis de datos transversales la regresión es una de las técnicas más utilizadas en estadística, permite explorar la asociación entre variables independientes y una variable respuesta, para explicar la contribución de las variables explicativas y determinar su impacto sobre la variable dependiente. Los polinomios ajustados suelen ser de orden bajo, ofrecen una familia limitada de formas y polinomios de orden alto, pueden ajustar pobremente en los valores extremos de las covariables. Royston y Altman (1994 y 1997) concluyen, que en datos donde se presentan efectos en los extremos de las variables independientes, tales efectos son menos comunes y menos marcados con los polinomios fraccionarios. Otra desventaja es que los polinomios no tiene asíntotas y no pueden ajustar datos donde el comportamiento límite es esperado (McCullangh y Nelder). Una solución es propuesta por Box y Tidwell (1962), quienes desarrollan una aproximación para linealizar cada variable en modelos de regresión múltiple, ellos se concentran sobre la transformación de potencias de las variables independientes y muestran cómo obtener las potencias iterativamente.

En el análisis de datos longitudinales, en áreas de investigación como biología, medicina, agronomía, etc., se tienen variables que se miden a través del tiempo y con tratamientos tales como dosis, niveles de PH, etc., que se encuentran en forma cuantitativa, que de alguna manera alteran la respuesta a través del tiempo. Una(s) curva(s) de crecimiento, describe la evolución del tratamiento a través del tiempo.

Para describir el crecimiento humano, las curvas generalmente no son lineales. Por ejemplo Count (1942 y 1943) modela el crecimiento del cráneo en niños Americanos blancos y la estatura de niños Chinos de 3 meses a 7 años. Wingerd(1970) compara el modelo de Count con un modelo convencional cuadrático y con un modelo donde utiliza la transformación 0.5 en la variable independiente, concluye que modelos con más de dos términos son requeridos para éstos datos. Nelder (1966) usa un modelo para describir la relación entre el crecimiento de la planta y la concentración de fertilizante, sugiere un sistema de polinomios inversos, muestra alternativas para ajustar modelos con polinomios fraccionarios.

Las técnicas no paramétricas han sido propuestas para realizar suavizamiento. Estas son útiles en el modelamiento estadístico, cuando no se conoce la forma de la distribución de la cual provienen los datos, sin estar límitada a una cierta forma funcional como en el análisis paramétrico. Los spline cúbicos pueden ser vistos como un enlace entre los polinomios convencionales y los métodos modernos de suavizamiento no-paramétrico. Los splines fueron inicialmente desarrollados por Whittaker (1923) para interpolación, mucho más tarde Reinsh (1967) y Silverman (1985) los utilizaron como método para ajustar curvas a los datos. El método de suavizamiento no paramétrico via scatterplot, conduce a

formas funcionales apropiadas de los datos, (Hastie y Tibshirani, 1990). Más que imponer un rango limitado de formas de los datos, el suavizamiento es construido en cada dato por mínimos cuadrados ponderados, dentro de la vecindad del correspondiente valor de la covariable, el suavizamiento más conocido es el suavizamiento localmente ponderado via scatterplot (LOWESS). Estos procedimientos de tipo no parámetrico requieren un procesamiento computacional intensivo con sofware ampliamente disponible, pero conducen a ecuaciones complicadas para hacer predicciones.

El objetivo de este trabajo, es proponer una metodología para ajustar curvas de crecimiento, a datos longitudinales través de los polinomios fracccionarios cuando se presentan ruidos o efectos finales en los extremos de las variables independientes. Extendiendo el trabajo de Royston y Altman (1997) quienes ajustan polinomios fraccionarios a datos transversales.

Bajo el supuesto de normalidad se dispone del ajuste de curvas de crecimiento, se usan en situaciones donde se quiere estudiar la evolución de uno o más grupos de unidades a través del tiempo, para caracterizar esta tendencia, se utlizan en los ajustes polinomios clásicos (lineal, cuadrático o de orden superior), y es aquí donde los polinomios fraccionarios cobran gran interés, pues se puede extender el uso de polinomios fraccionarios a curvas de crecimiento en presencia de datos extremos. Combinando en la matriz $X$ de tiempos, un subconjunto de polinomios fraccionarios, obtenido del conjunto de polinomios fraccionarios propuesto por Royston y Altman (1997), para obtener ajustes más suaves, flexibles, faciles de comprender y parsimoniosos.

El trabajo se presenta en cuatro capítulos, en el capítulo uno se hace la revisión literaria, toma una serie de conceptos, modelos, métodos para la estimación de los parámetros y una serie de tópicos para el desarrollo de esta tesis, considerados necesarios para el desarrollo de los capítulos dos, tres y cuatro.

En el capítulo dos se presenta la metodología de ajuste de curvas de crecimiento con polinomios fraccionarios, para datos longitudinales distribuidos normalmente. Se detallan el proceso de estimación de los parámetros y algunas propiedades que cumplen sus estimaciones. Se introduce el modelo propuesto con la estructura de la matriz de tiempos $X$, esta metodología esta acompañada en el capítulo tres con una aplicación a un conjunto de datos reales. En el capitulo cuatro se ilustra el uso de la metodología propuesta en en capitulo dos a través de un estudio de simulación, donde se presentan ruidos en las medidas repetidas, para analizar la efectividad de los polinomios fraccionarios en presencia de valores extremos, se compara el desempeño de los modelos de curvas de crecimiento ajustados a través de polinomios y polinomios fraccionarios por medio del $AIC$. Por último en el capítulo cinco se presentan las conclusiones del trabajo y algunas recomendaciones para trabajos futuros.

# CAPÍTULO 1

# Curvas de crecimiento y polinomios fraccionarios

## 1.1. Introducción

En un estudio de datos longitudinales las unidades (individuos, animales, parcelas, etc.,) son medidas repetidas a través del tiempo respecto a uno o varios atributos. Los diseños de datos longitudinales pueden estar conformados por uno o más grupos de unidades medidas en una o más variables a lo largo de dos o más periodos de tiempo. Las mediciones en cada una de las unidades están correlacionadas entre sí. Las curvas de crecimiento son una herramienta para modelar este tipo de datos. Con ellas, se elabora un dispersograma, un gráfico de perfiles, con una serie de medidas tomadas en sucesivos intervalos de tiempo, en una o más muestras de sujetos, con el fin de examinar y explicar las posibles diferencias existentes entre las diferentes muestras de individuos y para describir la evolución de una variable en el tiempo.

Los polinomios fraccionarios propuestos por Royston y Altman (1997 y 2008) son restringuidas a un pequeño conjunto predefinido de valores enteros y no enteros. Los elementos enteros son exponentes que determinan polinomios y los elementos no enteros son exponentes fraccionarias que determinan los polinomios fraccionarios. Éstos polinomios fraccionarios se utilizan debido a que los polinomios de alto orden pueden ajustar pobremente los valores extremos y polinomios de bajo orden ofrecen una familia limitada de formas. Por tanto en modelos donde se presentan ruidos los polinomios fraccionarios pueden ajustar mejor los datos.

## 1.2. Curvas de crecimiento (CC)

Las curvas de crecimiento son utilizadas para datos registrados en diferentes tiempos, sobre individuos que reciben diferentes tratamientos o que están divididos en diversos grupos o poblaciones, en los que cada registro se forma por medidas sobre un número de variables. Las curvas de crecimiento muestran la tendencia o evolución de un individuo o un grupo(s) de individuos a lo largo de un periodo determinado de tiempo; el modelo ajustado generalmente está dado por un polinomio en función del tiempo. Los modelos de CC

son una herramienta fundamental para trabajar con datos longitudinales con correlación serial, como también para medidas repetidas. Estos modelos de curvas de crecimiento fueron inicalmente resumidos por Potthoff y Roy (1958) y posteriormente fueron utilizados por muchos autores incluyendo a Rao (1965, 1966 y 1967), Geisser (1980), Lee (1988), Keramidas y Lee (1995), Davis (2005), etc.

Por ejemplo, los datos de curvas de crecimientos de frutos se obtienen siguiendo las cambios de una o más variables (diamétro, peso, número de células, etc.) a través del tiempo. Una curva puede representar el crecimiento de un fruto a lo largo del tiempo; varía según la especie y las condiciones productivas de la zona de cultivo. Es por ello que resulta de interés ajustar una curva de crecimiento de un fruto adaptada a ciertas condiciones. Al ajustar la curva de crecimiento se tiene una idea de la tasa de crecimiento y del número de divisiones célulares que tuvieron los frutos a través del tiempo.

Si los periodos de tiempo se encuentran igualmente espaciados, el modelo de curva de crecimiento ajustado suele hacerse a través de polinomios ortogonales, los cuales se encuentran detallados en Díaz (2007),permite determinar el tipo de tendencia, la cual puede ser lineal, cuadrática o de grado superior en factores cuantitativos.

### 1.2.1. Modelos de curvas de crecimiento

El modelo de curvas de crecimiento involucra más de una variable dependiente o respuesta $\mathbf{Y}$. Considere la situación en la cual el vector respuesta de $t_p$ componentes es medido para cada una de las $n$ unidades o sujetos bajo condiciones particulares, donde la estructura de las medidas repetidas, de la matriz $\mathbf{Y}$ se presenta en la tabla 1.2.1

| | | Tiempos | | | |
|---|---|---|---|---|---|
| Grupos | Unidades | $t_1$ | $t_2$ | $\cdots$ | $t_p$ |
| 1 | 1 | $Y_{111}$ | $Y_{112}$ | $\cdots$ | $Y_{11t_p}$ |
| 1 | 2 | $Y_{121}$ | $Y_{122}$ | $\cdots$ | $Y_{12t_p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| 1 | $n_1$ | $Y_{1n_11}$ | $Y_{1n_12}$ | $\cdots$ | $Y_{1n_1t_p}$ |
| | medias | $\mu_{11}$ | $\mu_{12}$ | $\cdots$ | $\mu_{1t_p}$ |
| 2 | 1 | $Y_{2111}$ | $Y_{212}$ | $\cdots$ | $Y_{21t_p}$ |
| 2 | 2 | $Y_{221}$ | $Y_{222}$ | $\cdots$ | $Y_{22t_p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| 2 | $n_2$ | $Y_{2n_21}$ | $Y_{2n_22}$ | $\cdots$ | $Y_{2n_2t_p}$ |
| | medias | $\mu_{21}$ | $\mu_{22}$ | $\cdots$ | $\mu_{2t_p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| r | 1 | $Y_{r11}$ | $Y_{r12}$ | $\cdots$ | $Y_{r1t_p}$ |
| r | 2 | $Y_{r21}$ | $Y_{r22}$ | $\cdots$ | $Y_{r2t_p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| r | $n_r$ | $Y_{rn_r1}$ | $Y_{rn_r2}$ | $\cdots$ | $Y_{rn_rt_p}$ |
| | medias | $\mu_{r1}$ | $\mu_{r2}$ | $\cdots$ | $\mu_{rt_p}$ |

Suponga que $n$ individuos, son asignados en $r$ grupos con $n_j$ individuos en el j-ésimo grupo, de tal forma que, $\sum_{j=1}^{r} n_j = n$. Los individuos en cada uno de los grupos son medidos en los tiempos $t_1, t_2, ..., t_p$. Las $p$ variables sobre un mismo individuo no son independientes y se asumen con distribución normal multivariada. Las variables de diferentes individuos son asumidas para ser independientes.

Pan y Fang (2002) expresan el modelo de curva de crecimiento en forma matricial así:

$$\mathbf{Y}_{p\times n} = \mathbf{X}_{p\times m}\mathbf{B}_{m\times r}\mathbf{Z}_{r\times n} + \mathbf{E}_{p\times n} \tag{1.1}$$

donde $\mathbf{X}$ y $\mathbf{Z}$ son las matrices diseño de rango $m < p$ y $r < n$ respectivamente y $\mathbf{B}$ es la matriz de parámetros de tamaño $m \times r$. $\mathbf{X}$ es una matriz que relaciona el grado de la curva y los tiempos y $\mathbf{e}_{it} = (e_{i1}, \cdots, e_{it_p})^t$ es la parte aleatoria correspondiente a los errores del i-ésimo individuo en el t-ésimo tiempo.

Se denota,

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n)$$

A la matriz de respuestas de los individuos, de tamaño $p \times n$, donde $\mathbf{y}_i = (y_{1i}, y_{2i}, \cdots, y_{pi})^t$ es un vector p-variante de respuestas sobre el i-ésimo individuo en los $p$ puntos en el tiempo. La matriz diseño $\mathbf{X}$ de orden $p \times m$, se denota:

$$\mathbf{X} = \begin{pmatrix} 1 & t_1 & \cdots & t_1^{m-1} \\ 1 & t_2 & \cdots & t_2^{m-1} \\ 1 & t_3 & \cdots & t_3^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_p & \cdots & t_p^{m-1} \end{pmatrix}$$

La matriz diseño de orden $n \times r$ entre grupos, se denota:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{1_{n_1}} & 0 & \cdots & 0 \\ 0 & \mathbf{1_{n_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1_{n_r}} \end{pmatrix}$$

La matriz de parámetros de orden $(m-1) \times r$, se denota:

$$\mathbf{B} = \begin{pmatrix} \beta_{0,1} & \beta_{0,2} & \cdots & \beta_{0,r} \\ \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,r} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m-1,1} & \beta_{m-1,2} & \cdots & \beta_{m-1,r} \end{pmatrix} \tag{1.2}$$

$\mathbf{e}_{ri} = (e_{ri1}, e_{ri2}, \cdots, e_{rip})$ el vector de residuales del sujeto i-ésimo en el grupo r-ésimo. La matriz de residuales de orden $p \times n$, se denota:

$$\mathbf{E} = \begin{pmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & \ddots & \vdots \\ e_{p1} & \cdots & e_{pn} \end{pmatrix}$$

Bajo el supuesto que los $\mathbf{E} \sim \mathbf{N_{pn}}(\mathbf{0}, \mathbf{I} \otimes \mathbf{\Sigma})$ y $\mathbf{Y} \sim \mathbf{N_{pn}}(\mathbf{XBZ}, \mathbf{I} \otimes \mathbf{\Sigma})$, donde cada columna de la matriz $\mathbf{Y}$ de variables se distribuye independiente y es normal multivariada con matriz de covarianza $\mathbf{\Sigma}$ definida positiva,

$$E(\mathbf{Y}) = \mathbf{XBZ}$$

$$Var(\mathbf{Y}) = \mathbf{I}_n \otimes \mathbf{\Sigma}$$

La curva de crecimiento asociada con el j-ésimo grupo es descrita por el polinomio de grado $m-1$ definido por:

$$\mu_j = \beta_{0,j} + \beta_{1,j}t + \beta_{2,j}t^2 + ... + \beta_{m-1,j}t^{m-1} \tag{1.3}$$

donde $m-1$ es el entero preseleccionado que refleja la evolución del crecimiento promedio de los individuos.

### 1.2.2. Estimación de los parámetros

Una variedad de métodos de estimación han sido propuestos en la literatura. Entre estas, la estimación de mínimos cuadrados generalizados (EMCG) y la estimación máxima verosimilitudd (EMV). Sin embargo, los EMCG son idénticos a los EMV cuando se utilizan las estructuras de covarianzas simples de Rao (SCSR), entre estas, la estructura de covarianzas uniforme y la estructura de coeficientes de regresión aleatorios, estudiadas posteriormente y utilizadas en los trabajos de Geisser (1980), Lee (1988), Lee y Geisser (1995).

- **Mínimos cuadrados generalizados**

  Algunas inferencias estadísticas básicas del modelo de curvas de crecimiento se hacen a través de mínimos cuadrádos generalizados cuando los errores no tienen varianza constante y/o estan correlacionados ( Draper y Smith, 1998). En el caso que nos ocupa los datos estan correlacionados a través del tiempo y en algunos casos no se cumple el supuesto de homocedasticidad.

  En modelos de CC una estimación $\hat{\mathbf{B}}$ es la estimación de mínimos cuadrados generalizados de los coeficientes de regresión $\mathbf{B}$, que satisface la ecuación de estimación:

$$(\boldsymbol{X}^t\boldsymbol{X})\hat{\mathbf{B}}(\boldsymbol{Z}\boldsymbol{Z}^t) = \boldsymbol{X}^t\boldsymbol{Y}\boldsymbol{Z}^t \tag{1.4}$$

  esta ecuación de estimación tiene una única solución para los $\mathbf{B}$, si las matrices diseño $\mathbf{X}$ y $\mathbf{Z}$ en el modelo de CC son de rango completo (Von y Rosen, 1984 y Pan, 1988).

  La ecuación de estimación (1.4) es equivalente a:

$$\hat{\mathbf{B}} = (\boldsymbol{X^t X})^{-1}\boldsymbol{X}^t\boldsymbol{Y}\boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^{-1}$$

Es una función lineal de la variable respuesta $\mathbf{Y}$ en el sentido de una combinación lineal de $\mathbf{Y}$.

Por lo tanto el EMCG es una combinación lineal de la matriz de respuestas $\mathbf{Y}$. $\hat{\mathbf{B}}$, esta normalmente distribuido y su valor esperado es:

$$E(\hat{\mathbf{B}}) = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{X}\mathbf{B}\mathbf{Z}\boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^{-1} = \mathbf{B}$$

Así el EMCG, $\hat{\mathbf{B}}$ es un estimador insesgado de los coeficientes de $\mathbf{B}$ en el modelo de CC.

Además, su matriz de covarianza es:

$$Cov(\hat{\mathbf{B}}) = (\boldsymbol{Z}\boldsymbol{Z}^t)^{-1} \otimes (\boldsymbol{X}^t\boldsymbol{X})^{-1}(\boldsymbol{X}^t\boldsymbol{\Sigma}\boldsymbol{X})(\boldsymbol{X}^t\boldsymbol{X})^{-1} \tag{1.5}$$

Esta mide la variación entre los tiempos y $\hat{\boldsymbol{\Sigma}}$ puede ser reemplazado en la covarianza (1.5)

Para los coeficientes de regresión $\mathbf{B}$ en el modelo de CC (1.1) la distribución de la EMCG (1.4) de $\hat{\mathbf{B}}$, puede ser escrita como:

$$\hat{\mathbf{B}} \sim N_{m,r}(\mathbf{B}; (\boldsymbol{X}^t\boldsymbol{X})^{-1}(\boldsymbol{X}^t\boldsymbol{\Sigma}\boldsymbol{X})(\boldsymbol{X}^t\boldsymbol{X})^{-1})$$

Si las observaciones son independientes, la estructura de covarianzas es $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}_p$, donde $\sigma^2 > 0$, es desconocida. Los coeficientes de regresión pueden ser expresados así:

$$\hat{\mathbf{B}} \sim N_{m,r}(\mathbf{B};\ \boldsymbol{\sigma}^2(\boldsymbol{X}^t\boldsymbol{X})^{-1}) \tag{1.6}$$

Pan y Fang (2001) obtienen un estimador insesgado del componente de dispersión $\sigma^2$, dado por:

$$\hat{\sigma}^2 = \frac{1}{np - mr} tr[(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})^t]$$

donde $m$ y $r$ son los rangos de las matrices $\mathbf{X}$ y $\mathbf{Z}$, respectivamente y $\hat{\mathbf{B}}$ es el EMCG de los coeficientes de regresión $\mathbf{B}$.

En el caso de que las matrices $\mathbf{X}$ y $\mathbf{Z}$ no sean de rango completo, la solución de la ecuacion de estimación (1.4) es de la forma:

$$\hat{\mathbf{B}} = (\boldsymbol{X}^t\boldsymbol{X})^-\boldsymbol{X}^t\boldsymbol{Y}\boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^- + \mathbf{H} - (\boldsymbol{X}^t\boldsymbol{X})^-(\boldsymbol{X}^t\boldsymbol{X})\mathbf{H}(\boldsymbol{Z}\boldsymbol{Z}^t)(\boldsymbol{Z}\boldsymbol{Z}^t)^-$$

donde $\mathbf{H}$ es una matriz arbitratia $m \times r$ y $\mathbf{A}^-$ denota una inversa generalizada de la matriz $\mathbf{A}$.

- **Estimación por máxima verosimilitud**

  La estimación por máxima verosimilitud es una técnica muy utilizada en estadística inferencial, los estimadores obtenidos poseen propiedades como consistencia y distribución normal asintótica.

  La función de verosimilitud de los parámetros $\mathbf{B}$ y $\mathbf{\Sigma}$ en el modelo de CC es de la forma:

$$L(\mathbf{B},\mathbf{\Sigma}) = (2\pi)^{-pn/2}[det(\mathbf{\Sigma})]^{-n/2}\exp\{-\frac{1}{2}traza(\mathbf{\Sigma^{-1}}(\mathbf{Y}-\mathbf{XBZ})(\mathbf{Y}-\mathbf{XBZ})')\} \tag{1.7}$$

  Las funciones de verosimilitud de los coeficientes de regresión $\mathbf{B}$ y el componente de dispersión $\mathbf{\Sigma}$, se obtienen a través de las derivadas parciales de (1.7) respecto a $\mathbf{B}$ y $\mathbf{\Sigma}$ e igualando a cero, son:

$$\mathbf{X}^t\mathbf{\Sigma}^{-1}(\mathbf{Y\text{-}XBZ})\mathbf{Z}^t = \mathbf{0} \tag{1.8}$$

$$n\mathbf{\Sigma} - (\mathbf{Y\text{-}XBZ})(\mathbf{Y\text{-}XBZ})^t = \mathbf{0} \tag{1.9}$$

  Pan y Fang (2001) para el modelo de curva de crecimiento (1.1), con distribución normal de los $\boldsymbol{Y}$, demuestran que si $n > p + r$ y las matrices de diseño $\mathbf{X}$ y $\mathbf{Z}$ son de rango completo, entonces la solución para $(\hat{\mathbf{B}},\hat{\mathbf{\Sigma}})$ en las ecuaciones de verosimilitud (1.8 y 1.9), es única, y está dada por :

$$\hat{\boldsymbol{B}}_{\boldsymbol{(S)}} = (\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{Y}\boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^{-1} \tag{1.10}$$

  y

$$\hat{\mathbf{\Sigma}}_{\boldsymbol{(S)}} = \frac{1}{n}(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}}\mathbf{Z})(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}}\mathbf{Z})^t \tag{1.11}$$

  donde $\mathbf{S} = \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{Z^t})\mathbf{Y}^t$ y $\mathbf{P}_{Z^t} = \mathbf{Z}^t(\mathbf{Z}\mathbf{Z}^t)^{-1}\mathbf{Z}$.

  Si las matrices diseño $\mathbf{X}$ y $\mathbf{Z}$ no son de rango completo la solución de las ecuaciones de verosimilitud, para los coeficientes de regresión $\mathbf{B}$, pueden no ser únicas, la forma general de la solución es:

$$\begin{aligned}\mathbf{B} = (\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{Y}\boldsymbol{P}_{Z^t}\hat{\boldsymbol{H}}^{-1}\boldsymbol{Z}^t(\boldsymbol{Z}\hat{\boldsymbol{H}}^{-1}\boldsymbol{Z}^t)^{-} + \boldsymbol{R} - (\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{X})^{-} \\ (\mathbf{X}^t\boldsymbol{S}^{-1}\mathbf{X})\mathbf{R}(\mathbf{Z}\hat{\mathbf{H}}^{-1}\mathbf{Z}^t)(\mathbf{Z}\hat{\mathbf{H}}^{-1}\mathbf{Z}^t)^{-}\end{aligned} \tag{1.12}$$

  donde $\mathbf{R}$ es una matriz arbitraria de tamaño $m \times r$ y

$$\hat{\mathbf{H}} = \mathbf{I}_n + \hat{\mathbf{V}}^t\mathbf{S}^{-1}\hat{\mathbf{V}}$$

$$\hat{\mathbf{V}} = \mathbf{Y}\mathbf{P}_{Z^t} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z}$$

  donde $\boldsymbol{P}_{\boldsymbol{Z}^t} = \boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^{-}\boldsymbol{Z}$

  La matriz $\mathbf{X}\hat{\mathbf{B}}\mathbf{Z}$ es la estimación del valor esperado de la matriz de medidas repetidas $\boldsymbol{Y}$ y no depende de las matrices $\mathbf{R}$ y $\hat{\mathbf{H}}$, es de la forma:

$$\boldsymbol{X}\hat{\mathbf{B}}\boldsymbol{Z} = \boldsymbol{X}(\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}^t\boldsymbol{S}^{-1}\boldsymbol{Y}\boldsymbol{P}_{\boldsymbol{Z}^t} \tag{1.13}$$

no depende de la selección específica de la inversa generalizada, por lo tanto la esperanza es única aunque la solución de los $\hat{\mathbf{B}}$ dado en (1.12) puede no ser única. En este caso, la solución del componente de dispersión $\boldsymbol{\Sigma}$ es:

$$\hat{\Sigma} = \frac{1}{n}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})^t \tag{1.14}$$

y tambien es única.

Para modelos de CC (1.1) con distribución normal, si $n > p + r$ entonces la solución de $\mathbf{B}$ y $\boldsymbol{\Sigma}$ en las ecuaciones de verosimilitud (1.8) y (1.9) dados por (1.10) y (1.11), son los estimadores máxima verosimilitud de los coeficientes de regresión $\mathbf{B}$ y el componente de dispersión $\boldsymbol{\Sigma}$, respectivamente. Si al sustituir en (1.7), se cumple:

$$L(\mathbf{B}, \boldsymbol{\Sigma}) \leqslant L(\hat{\mathbf{B}}, \hat{\boldsymbol{\Sigma}}), \tag{1.15}$$

para algún $\mathbf{B}$ y $\boldsymbol{\Sigma}$ implica que la solución $\hat{\mathbf{B}}$ y $\hat{\boldsymbol{\Sigma}}$ dados por (1.10) y (1.11) son realmente los estimadores máxima verosimilitud de $\mathbf{B}$ y $\boldsymbol{\Sigma}$ respectivamente.

En la literatura la estimación de los coeficientes de regresión $\mathbf{B}$ y el componente de dispersión $\boldsymbol{\Sigma}$ son estudiados a través de una variedad de métodos. Por ejemplo, Potthoff y Roy (1964) usan dos pasos, para obtener la estimación de los parámetros así:

- Como $\boldsymbol{\Sigma}$ es desconocida, ellos sugieren usar una estimación $\mathbf{G} > \mathbf{0}$, para reemplazar la componente de dispersión $\boldsymbol{\Sigma}$, por ejemplo se puede seleccionar $\mathbf{G} = \boldsymbol{\Sigma}_p$, obtenida por información pasada, si esta es disponible es decir $\mathbf{G} = \boldsymbol{\Sigma}_\mathbf{0}$ o construirla usando información de los datos, es decir $\boldsymbol{G} = \boldsymbol{S} = \boldsymbol{Y}(\boldsymbol{I}_n - \boldsymbol{P}_{\mathbf{Z}^t})\boldsymbol{Y}^t$. En el
- Inserta la matriz $\mathbf{G}$ en la estimación:

$$\hat{\mathbf{B}} = \left(\mathbf{X}^t\mathbf{G}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{G}^{-1}\mathbf{Z}^t(\mathbf{Z}\mathbf{Z}^t)^{-1}$$

Rao (1965) desarrolla alternativas para evitar la selección arbitraria de $\mathbf{G}$ y considera un modelo condicional en modelos de curvas de crecimiento, usando $p-m$ covariables, Khatri (1966) deriva el estimador máxima verosimitud de $\mathbf{B}$.

Cuando $\mathbf{G} = \mathbf{I}_p$ los resultados estimados por máxima verosimilitud (1.10, 1.11) son los mismos obtenidos en (1.5) por mínimos cuadrados generalizados.

La distribución de los EMV es complicada, por esto se centra sobre su valor esperado y su matriz de varianzas y covarianzas.

Para el modelo de curva de crecimiento, (1.1) con distribución normal , si $n > p + r$, los EMV de $\mathbf{B}$, los coeficientes de regresión dados por (1.10 y 1.11) son estimadores insesgados de los coeficientes de $\mathbf{B}$. Además EMV de las varianzas y covarianzas de $\hat{\mathbf{B}}$ pueden ser expresados como:

$$Cov(\hat{\mathbf{B}}) = \frac{n - r - 1}{n - r - (p - m) - 1}(\mathbf{Z}\mathbf{Z}^t)^{-1} \otimes (\mathbf{X}^t\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$$

Sea $\hat{\mathbf{\Sigma}}$ El EMV del componente de dispersión $\mathbf{\Sigma}$ en el modelo de curva de crecimiento dado por (1.11) y denote,

$$c = \frac{r}{n}.\frac{n-r-2(p-m)-1}{n-r-(p-m)-1}$$

Si $n > max\{p+r, 2(p-m)+r+1\}$ entonces

$$E(\hat{\mathbf{\Sigma}}) = \mathbf{\Sigma} + c\boldsymbol{X}(\boldsymbol{X}^t\mathbf{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^t$$

Un estimador insesgado de $\mathbf{\Sigma}$,

$$\hat{\mathbf{\Sigma}_0} = \mathbf{\Sigma} + c\boldsymbol{X}(\boldsymbol{X}^t\mathbf{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^t$$

En general el estimador del componente de dispersión que es insesgado no es único, por ejemplo:

$$\frac{1}{n-1}\boldsymbol{S} = \frac{1}{n-1}\boldsymbol{Y}(\boldsymbol{I}_n - P_{\boldsymbol{Z}^t})\boldsymbol{Y}^t \tag{1.16}$$

es otro estimador insesgado de $\mathbf{\Sigma}$, Von Rosen (1991).

Para determinar que estimador es mejor que otra, se comparan los errores cuadráticos medio de los estimadores de las matrices de varianzas y covarianzas de los estimadores insesgados $\boldsymbol{S}/(n-r)$ y $\mathbf{\Sigma}$, dados por (1.16).

Para el caso es necesario especificar algunas estructuras de covarianza, $\mathbf{\Sigma}$, utilizadas para modelos de curvas de crecimiento, las cuales se presentan en la siguiente sección.

### 1.2.3. Estructuras de covarianzas

Para seleccionar el modelo de varianzas y covarianzas que mejor se ajusta a las medidas repetidas es necesario tener en cuenta las estructuras de covarianzas, las cuales fueron originalmente estudiadas por Rao (1965) quien considero un modelo mixto, y fueron más tarde estudiadas por Geisser(1980), Kariya (1985) y Davidian (2005).

Para definir la estructura que sigue la matriz de varianzas y covarianzas, se hace necesario tener en cuenta la variación entre y dentro de las unidades experimentales. Además obtener el tipo de relación entre las observaciones (la matriz de covarianzas muestrales y/o la matriz de correlación).

#### Estructuras de covarianzas de los modelos de CC.

A continuación se plantean algunas estructuras de covarianzas utilizadas para análisis de datos longitudinales.

- **Matriz de covarianzas no estructurada** : cuando no hay patrón sistemático aparente de varianza y correlación se dice que la matriz de covarianzas corresponde a un modelo no estructurado, donde se deben estimar las varianzas asociadas a cada uno de los $p$ tiempos y $p(p-1)/2$ covarianzas para cada par de tiempos, por lo que se encuentran $p(p+1)/2$ parámetros describiendo la variación. Su estructura presenta la siguiente forma:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{pmatrix}$$

  De esta manera $\mathbf{\Sigma}$ es una matriz de covarianzas arbitraria, definida positiva.

- **Estructura de coeficientes de regresión aleatoria**: en este modelo la magnitud de correlación es alta para observaciones consecutivas y observaciones más lejanas es más pequeña.

$$\boldsymbol{\Sigma = X\Gamma X^t + \sigma^2 I_p} \tag{1.17}$$

  los componentes de dispersión $\mathbf{\Gamma}$ y $\boldsymbol{\sigma}$ en (1.17) son respectivamente dados por:

$$\hat{\mathbf{\Gamma}} = \frac{1}{n}(\mathbf{X^t X})^{-1}\mathbf{X^t S X}(\mathbf{X^t X})^{-1} - \hat{\sigma^2}(\mathbf{X^t X})^{-1} \tag{1.18}$$

$$\hat{\sigma^2} = \frac{1}{(p-m)n} tr[\mathbf{Y^t}(\mathbf{I_p} - \mathbf{P_X})\mathbf{Y}] \tag{1.19}$$

  donde $\mathbf{P_X}$ es la proyeccion matricial formada por $\mathbf{X}$, es decir $\mathbf{P_X} = \mathbf{X}(\mathbf{X^t X})^{-1}\mathbf{X^t}$

  Su estructura presenta la siguiente forma:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_{12} & \cdots & \rho^{p-1}\sigma_{1p} \\ \rho\sigma_{21} & \sigma_2^2 & \cdots & \rho\sigma_{2p}^{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1}\sigma_{p1} & \rho^{p-2}\sigma_{p2} & \cdots & \sigma_p^2 \end{pmatrix}$$

  Es útil en análisis de modelos mixtos lineales o modelos lineales con efectos aleatorios, también como herramientas estadísticas en el análisis de datos longitudinales y datos correlacionados (Searle, Casella y McCulloch (1992)).

- **Estructura de covarianzas uniforme**:

  La correlación entre las observaciones en el tiempo $t_j$ y $t_{j^t}$ y las varianzas en los diferentes tiempos son iguales, en este modelo se presenta la misma correlación entre todos los tiempos teniendo la misma varianza para cada uno de los tiempos.

$$\mathbf{\Sigma} = \sigma^2(1-\rho)\mathbf{I}_p + \rho\mathbf{1}_p\mathbf{1}_p^t \tag{1.20}$$

  Su estructura presenta la siguiente forma:

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

donde $\sigma^2 > 0$ y $-1/(p-1) < \rho < 1$ son parámetros de dispersión desconocidos.

Los componentes de dispersion $\sigma^2$ y $\rho$ en (1.20) se obtiene al sustituir el estimador MV, en $\mathbf{V}$, asi:

$$\hat{\mathbf{V}} = \mathbf{Y}\mathbf{P}_{\mathbf{Z^t}} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z} = ((\mathbf{I_p} - \mathbf{P_X})\mathbf{Y}\mathbf{P}_{\mathbf{Z^t}}) = \mathbf{P_Q}\mathbf{Y}\mathbf{P}_{\mathbf{Z^t}}$$

asi que $\mathbf{S} + \hat{\mathbf{V}}\hat{\mathbf{V}}^{\mathbf{t}} \equiv \mathbf{S}^*$ y reemplazando la matriz $\mathbf{S} + \mathbf{V}\mathbf{V}^{\mathbf{t}}$ con $\mathbf{S}^*$ y resolviendo las ecuaciones de estimación $\partial l/\partial\sigma^2 = 0$ y $\partial l/\partial\rho = 0$,los estimadores maxima verosimilitud de $\hat{\sigma^2}$ y $\hat{\rho}$ de los componentes de dispersion $\sigma^2$ y $\rho$ pueden ser obtenidos simultaneamente como:

$$\sigma^2 = \frac{1}{pn} tr(\mathbf{S}^*)$$

Y

$$\hat{\rho} = \frac{1}{p-1} \frac{\mathbf{1_p^t}\mathbf{S}^*\mathbf{1_p} - \mathbf{tr}(\mathbf{S}^*)}{tr(\mathbf{S}^*)}$$

donde $\mathbf{S}^* = \mathbf{S} + \mathbf{P_Q}\mathbf{Y}\mathbf{P}_{\mathbf{Z^t}}\mathbf{Y^t}\mathbf{P_Q}$ y $\mathbf{Q}$ esta en el espacio de matrices ortogonales $\mathbf{Q}$ de $\mathbf{X}$ esta estructura de covarianzas es usada para datos de CC, cuando las observaciones son una mezcla de varias poblaciones, ( Lee (1988)). Esta estructura de covarianzas se incluye en las estructuras de covarianzas de Rao, si se satisface:

$$\mathbf{X}^t\mathbf{1}_p\mathbf{1}_p^t\mathbf{Q} = \mathbf{0}$$

- **Estructura de covarianzas autoregresiva de orden uno (AR1) : $\mathbf{\Sigma} = \sigma^2\rho^{|i-j|}_{1\leqslant i,j\leqslant p}$**
  La forma de la estructura de covarianzas autoregresiva de orden uno es una estructura apropiada para procesos de crecimiento. Sus elementos sobre la diagonal son homocedásticos (con varianza $\sigma^2$), adicionalmente pares de errores tienen idéntica covarianza a bandas paralelas a la diagonal. Estas covarianzas son el producto de la varianza residual, $\sigma^2$ y un parámetro de autocorrelación del error, $\rho$ cuya magnitud es casi siempre menor que, o igual a la unidad. Usa solamente dos componentes de varianza $\sigma^2$ y $\rho$. Su estructura presenta la siguiente forma:

$$\Sigma = \sigma^2 \begin{pmatrix} \sigma^2 & \sigma^2\rho & \cdots & \sigma^2\rho^{p-1} \\ \sigma^2\rho & \sigma^2 & \cdots & \sigma^2\rho^{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2\rho^{p-1} & \sigma^2\rho^{p-2} & \cdots & \sigma^2 \end{pmatrix}$$

  En el modelo $AR1$ la magnitud de la correlación es alta para observaciones consecutivas y en observaciones más alejadas tiende a cero

- : **Estructura de covarianzas simple**: Cuando se considera la independencia de las observaciones$\mathbf{\Sigma} = \sigma^2 G$ $(G > 0$ es una estructura de covarianzas esférica).

$$\mathbf{\Sigma} = \sigma^2 \mathbf{I} \tag{1.21}$$

Su estructura presenta la siguiente forma:

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Para el modelo de CC (1.1) los estimadores máxima verosimilitud son completamente idénticos a los estimadores de mínimos cuadrados generalizados (1.5), si la componente de dispersión $\mathbf{\Sigma}$ es una estructura de covarianzas simples de Rao.

Las estructuras de varianzas de Rao (ECSR), estan dadas por:

$$\mathbf{\Sigma} = \boldsymbol{X}\mathbf{\Gamma}\boldsymbol{X}^t + \boldsymbol{Q}\mathbf{\Theta}\boldsymbol{Q}^t \tag{1.22}$$

donde $\mathbf{\Gamma}$ y $\mathbf{\Theta}$ son matrices definidas positivas desconocidas y la matriz $\boldsymbol{Q}$ viene de un espacio de matrices ortogonalas $Q$ de la matriz diseño $\boldsymbol{X}$. Es decir

$$\boldsymbol{Q} = \{Q : Q_{px(p-m)},\ rango = p - m,\ \boldsymbol{X}^t\boldsymbol{Q} = \mathbf{0}\}$$

Las ECSR (1.22) son comunmente usadas en análisis de curvas de crecimiento debido a su caracter parsimonioso, que hace la estadística inferencial más eficiente porque reduce la estadística inferencial no lineal a inferencias lineales. Dentro de estos modelos se cuentan la estructura de covarianzas uniforme y estructura de coeficiente de regresión aleatoria.

**Estimación de los componentes de dispersión**

Para los modelos de CC con estructuras de covarianzas de Rao (1.22) los EMV de los coeficientes de regresión $\mathbf{B}$ son equivalentes a los EMCG. Si $n > p + r$, los EMV de los componentes de los coeficientes de regresión $\mathbf{B}$ y los componentes de dispersión $\mathbf{\Gamma}$ y $\mathbf{\Theta}$, son de la forma:

$$\hat{\mathbf{B}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Y}\mathbf{Z}^t(\mathbf{Z}\mathbf{Z}^t)^{-1} \tag{1.23}$$

$$\hat{\mathbf{\Gamma}} = 1/n(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{S}\ \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1} \tag{1.24}$$

$$\hat{\mathbf{\Theta}} = (\mathbf{Q}^t\mathbf{Q})^{-1}\mathbf{Q}^t\mathbf{Y}\mathbf{Y}^t\mathbf{Q}(\mathbf{Q}^t\mathbf{Q})^{-1} \tag{1.25}$$

donde $\hat{\mathbf{\Sigma}}$ es el EMV del componente de dispersión, $\mathbf{\Sigma}$ y $\mathbf{Q}$ es una matriz ortogonal del espacio $\mathbf{Q}$ de $\mathbf{X}$. Cuando la componente de dispersión $\hat{\mathbf{\Sigma}}$ es no estructurada, la distribución de EMV para los $\mathbf{B}$ y los $\mathbf{\Sigma}$ no tienen una forma analítica disponible y los momentos son

complicados. Para modelos de CC con ECSR, se puede derivar de los EMCG, la distribución de los EMV de $\hat{\mathbf{B}}$ es normal es decir:

$$\hat{\mathbf{B}} \sim N_{m,n}(\mathbf{B}; (\boldsymbol{X^t X})^{-1}(\boldsymbol{X^t \Sigma X})(\boldsymbol{X^t X})^{-1})$$

Un problema en un conjunto de datos de crecimiento es como seleccionar la estructura de covarianzas más apropiada para el modelo de curvas de crecimiento, el cual puede ser visto como una selección de modelos con respecto a la matriz de covarianzas $\boldsymbol{\Sigma}$, una forma es la Selección basada en la quasi-verosimilitud predictiva.

### 1.2.4. Selección basada en la quasi-verosimilitud predictiva

Es posible obtener diferentes estructuras de covarianzas, para determinar cuál es la estructura de covarianzas que más se ajusta a los datos, se puede utilizar la quasi-verosimilitud predictiva, propuesta por Lee Jack (1991)

Sea $\boldsymbol{Y} = (\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_n})$ son las observaciones sobre $n$ casos independientes (o sujetos), con covariables $X_1, X_2, \cdots, X_N$ . Se asume que los posibles modelos $M_1, M_2, ..., M_q$ han sido generados por datos de curvas de crecimiento. La cuasi-verosimilitud predictiva es el producto de las densidades de cuasi-verosimilitudes predictivas, definidas por:

$$L_\alpha = \prod_{j=0}^{n} f_j(Y_j/Y_{(j)}, \mathbf{B}_{\alpha(j)}, \boldsymbol{\Sigma}_{\alpha,(j)}, M_\alpha)$$

donde $Y_{(j)}$=$(Y_1, ..., Y_{j-1}, Y_{j+1}, ..., Y_n)_{px(n-1)}$, y los parámetros $\mathbf{B}_{\alpha(j)}$ y $\hat{\Sigma}_{\alpha(j)}$ son los EMV de $\mathbf{B}$ y $\boldsymbol{\Sigma}$, respectivamente, bajo el modelo $M_\alpha$ con la muestra $Y_{(j)}$. La función

$$f_j(Y_j/Y_{(j)}, \mathbf{B}_{\alpha(j)}, \boldsymbol{\Sigma}_{\alpha,(j)}, M_\alpha)$$

denota la densidad condicional de $Y_j$ dado $Y_{(j)}$ y $\mathbf{B}_{\alpha,(j)}$ bajo el modelo $M_\alpha$. El criterio de selección es a través del valor de $\alpha^*$ que maximice la función de $log(L_\alpha)$ y se obtiene la estructura $M_\alpha^*$ como la más apropiada de los modelos que estan siendo considerados (Lee, 1991).

Una vez que se han estimado los parámetros, el interés es plantear las hipótesis para el análisis de curvas de crecimiento, para encontrar el tipo de relación entre los tiempos y las medidas repetidas.

### 1.2.5. Prueba de hipótesis

#### Curvas de crecimiento en una muestra

El ajuste de la curva se hace a través de un polinomio en función del tiempo. Se puede hacer a través de polinomios ortogonales, si los periodos de tiempo estan igualmente espaciados.

Un polinomio ortogonal es un caso especial de contraste, empleado para verificar tendencias de orden lineal, cuadrático o superior en factores cuantitativos. Esta metodología

es presentada detalladamente por Díaz (2007). También se pueden dar los contrastes a través de la matriz $X$ presentada en (1.2)

La hipótesis a probar es:

$$H_0 : \mathbf{X}\mu = \mathbf{0} \tag{1.26}$$

Se verifica mediante la estadística $T^2$ de Hotelling, presentada en detalle en Díaz (2007).

$$T^2 = n\bar{\mathbf{Y}}\mathbf{X}(\mathbf{X}\mathbf{S}\mathbf{X}^{\mathbf{t}})^{-\mathbf{1}}\mathbf{X}(\bar{\mathbf{Y}}\mathbf{X})^t \tag{1.27}$$

La cual se distribuye $T^2_{(k,n-1)}$, donde $k$ es el número de columnas de $\mathbf{X}$, $n$ el número de elementos de la muestra, $\bar{\mathbf{Y}}$ el vector de medias y $\mathbf{S}$ la matriz de covarianzas muestral.

Si se rechaza $H_0$, se realizan las pruebas sobre cada fila de $\mathbf{X}$ (polinomio), del tipo $X_i\mu$ y se hacen mediante la estadística:

$$t_i = \frac{\bar{Y}X_i}{\sqrt{X_i\bar{Y}{X_i}^t/n}} \tag{1.28}$$

para $i = 1, 2, /\text{cdots}, k$ esta estadística se distribuye como una t-Studente con $n-1$ grado de libertad.

Ahora se considera el caso cuando los periodos de tiempo no estan igualmente espaciados, es decir periodos de longitud diferente. Suponga que se observa una respuesta de un individuo en los tiempos $t_1, t_2, ..., t_p$ y que la media de la respuesta $\mu$, en cualquier punto en el tiempo $p$ es un polinomio sobre el tiempo de grado $m < p$, es decir,

$$\mu = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_{m-1}t^{m-1} \tag{1.29}$$

se tiene para cada punto $t_i$ con respuesta media $\mu_i$. La hipótesis se plantea asi:

$$H_0 = \begin{pmatrix} \mu_1 = \beta_0 + \beta_1 t_1 + \beta_2 t_1^2 + \cdots + \beta_{m-1}t_1^{m-1} \\ \mu_2 = \beta_0 + \beta_1 t_2 + \beta_2 t_2^2 + \cdots + \beta_{m-1}t_2^{m-1} \\ \vdots \\ \mu_p = \beta_0 + \beta_1 t_p + \beta 2 t_p^2 + \cdots + \beta_{m-1}t_p^{m-1} \end{pmatrix} \tag{1.30}$$

que equivale a

$$H_0 : \mu = \mathbf{X}\mathbf{B} \tag{1.31}$$

con $X$ dado en (1.2) y $B$ definidos (1.10) pero solamente una columna porque corresponde a un solo grupo o población.

Estimados los parámetros por máxima verosimilitud o por mínimos cuadrados generalizados, se verifica la hipótesis a través de la estadística siguiente:

$$T^2 = n(\bar{Y} - \mathbf{X}\mathbf{B})\mathbf{S}^{-1}(\bar{Y} - \mathbf{X}\mathbf{B})^t \tag{1.32}$$

la cual tiene distribución $T^2_{(p-m-1,n-1)}$. Los percentiles superiores de la distribución $T^2$ de Hotelling se encuentran en Díaz (2007)

**Curvas de crecimiento en r-muestras**

Para varias muestras o grupos, Davis (2002) considera la hipótesis general para el caso multivariado.

$$H_0 : \mathbf{ABC} = \mathbf{D} \tag{1.33}$$

- $\mathbf{B}$ : matriz de coeficientes presentada en (6)
- $\mathbf{A}$ : matriz de coeficientes de tamaño $a \times p$ de rango $a < p$, permite probar la hipótesis, dentro de los tiempos, es decir la hipótesis en los elementos dentro de las columnas de $\mathbf{B}$.
- $\mathbf{C}$ : matriz de coeficientes de tamano $p \times c$, rango $c \leqslant t \leqslant n - p$, permiten probar la hipótesis entre tiempos, es decir la hipótesis en los elementos dentro de las filas de $\mathbf{B}$.
- $\mathbf{D}$ :es una matriz de constantes, de tamaño $a \times c$

es una forma general de escribir las hipótesis.

La curva de crecimiento asociada al r-ésimo grupo es descrita por un polinomio de grado $m - 1$

$$\mu_{\mathbf{r,j}} = \beta_{r,0} + \beta_{r,1} j + \beta_{r,2} j^2 + \cdots + \beta_{r,m-1} j^{m-1}$$

Para la respuesta media en el grupo $r$ y el t-ésimo tiempo, $\mu_{r,t}$ Se tienen tres tipos de hipótesis de interés:

- $H_{01}$:Las $k$ curvas de crecimiento son de grado $(m - 1)$ o menos.
- $H_{01}$:Las $k$ curvas de crecimiento son iguales en los diferentes grupos.
- $H_{01}$:Las $k$ curvas de crecimiento son iguales excepto por una constante positiva.

Para probar las hipótesis anteriores Potthoff y Roy (1964) proponenen transformar el modelo de análisis de CC (2) a la forma usual del modelo de análisis de varianza multivariado, usando análisis de perfiles. Dicha transformación consiste en premultiplicar ambos lados del modelo (2) por la matriz $(\mathbf{X^t G^{-1} X})^{-1} \mathbf{X^t G^{-1}}$ de tamaño $p \times m$,donde $\mathbf{G}$ es una matriz simétrica y definida positiva, de tamaño $p \times p$, $\mathbf{X G^{-1} X^t}$ tiene rango $m$. El modelo queda expresado de la siguiente manera:

$$\mathbf{K} = \mathbf{BZ} + \mathbf{E^*} \tag{1.34}$$

donde $\mathbf{K} = \mathbf{Y G^{-1} X^t (X G^{-1} X^t)^{-1}}$ y $\mathbf{E^*} = E(\mathbf{X^t G^{-1} X})^{-1} \mathbf{X^t G^{-1}}$, de esta manera se encuentra un modelo de análisis de varianza multivariado, para el que se pueden hacer el mismo tipo de inferencias.

Para probar la hipótesis $H_0 : \mathbf{ABC} = \mathbf{D}$, existe una amplia variedad de estadísticas: Lambda de Wilks, La traza de Lawley-Hotelling, la traza de Bartlett-Nanda-Pillai y la máxima valor propio de Roy. Estas estadísticas son halladas usando:

- Matrices de suma cuadrados y producto cruzados.

$$\mathbf{Q_h} = (\mathbf{A\hat{B}C - D})^{\mathbf{t}}[\mathbf{A}(\mathbf{X^tX})^{-\mathbf{1}}\mathbf{A^t}]^{-\mathbf{1}}\mathbf{A\hat{B}C - D}) \tag{1.35}$$

  $Q_h$ es analoga al numerador de una prueba $F$ univariada.

- Matrices de suma cuadrados y producto cruzados de los residuales.

$$\mathbf{Q_e} = \mathbf{C^t}[\mathbf{Y^tY} - \mathbf{\hat{B}^t}(\mathbf{X^tX})\mathbf{\hat{B}}]\mathbf{C} \tag{1.36}$$

$\mathbf{Q_e}$ es analoga a la suma de cuadrados del error.

La estadística de razón de verosimilitud es:

$$\Lambda = \frac{|Q_e|}{|Q_h + Q_e|} = \prod \frac{1}{1+\lambda_i}$$

donde $\lambda_i$ son la solución de la ecuacion característica $|Q_h - \lambda Q_e| = 0$, esta estadística es conocida como $\Lambda$ de Wilks. Las otras estadísticas son obtenidas a través de los valores de $Q_e$ y $Q_h$, y el proceso de obtención se encuentra explicado en Davis (2002) y Srilivastava (2002).

En resumen las cuatro estadísticas son las siguientes:

- Lambda de Wilks: $\Lambda = \prod_{i=1}^{p} \frac{1}{(1+l_i)}$ $l_i$ son las raices de $|\mathbf{Q_h} - l\mathbf{Q_e}|$
- Traza de Lawley-Hotelling: $\mathbf{U} = \sum l_i = traza(\mathbf{Q_hQ_e})^{-1}$
- Traza de Bartlett-Nanda-Pillai: $\mathbf{V} = \sum_{i=1}^{p} \frac{l_i}{(1+l_i)} = traza(\mathbf{Q_h}(\mathbf{Q_e} + \mathbf{Q_h})^{-1})$
- Maximo valor propio de Roy: $\mathbf{R} = \frac{l_1}{(1+l_1)}$

La distribución exacta de $\Lambda$ de Wilks fue obtenida por Rao (1973) para algunos casos especiales, se encuentra resumida en la siguiente tabla:

| Numero de variables | Numero de poblaciones | Transformacion | Distribucion F |
|---|---|---|---|
| s=1 | $r \geqslant 2$ | $(\frac{1-\Lambda}{\Lambda})(\frac{N-r}{r-1})$ | $F_{(r-1,N-r)}$ |
| $s = 2$ | $r \geqslant 2$ | $\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}})(\frac{N-r-1}{r-1})$ | $F_{(2(r-1),2(N-r-1))}$ |
| $s = \geqslant 1$ | $r = 2$ | $\frac{1-\Lambda}{\Lambda})(\frac{N-r}{r-1})$ | $F_{(s,N-s-1)}$ |
| $s = \geqslant 1$ | $r = 3$ | $(\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}})(\frac{N-s-2}{p})$ | $F_{(2s,2(N-s-2))}$ |

Los valores críticos de una $\Lambda$ de Wilks se encuentran en Kshirsagar (1995)

Para muestras de tamaño grande, la hipótesis nula, que establece lo adecuado del polinomio de grado $m$, se rechaza si

$$\left(-N - \frac{1}{2}(p - m - r)\right) ln\Lambda > \chi^2(\alpha, (p - m - 1)r) \tag{1.37}$$

Los modelos de curvas de crecimiento cumplen supuestos que deben ser verificados a través de estadísticas que se detallan en la siguiente sección.

### 1.2.6. Verificación de los supuestos del modelo

En el modelo de curvas de crecimiento se deben cumplir los siguientes supuestos:

- Homogeneidad de varianzas para los diferentes grupos tratamiento.

  El estadístico de Bartlett permite verificar la igualdad de covarianzas en los diferentes grupos tratamiento.

  La hipótesis que se desea probar es:

$$H_0 : \Sigma_1 = \Sigma_2 = \cdots = \Sigma_r = \Sigma \tag{1.38}$$

  El estadístico de Bartlett para el caso que nos ocupa es:

$$\lambda = \left(\frac{|S_1|}{|S|}\right)^{n_1-1/2} \left(\frac{|S_2|}{|S|}\right)^{n_2-1/2} \cdots \left(\frac{|S_h|}{|S|}\right)^{n_r-1/2} \tag{1.39}$$

  donde $S$ es la matriz de varianzas y covarianzas de cada uno de los grupos de tratamientos $i = 1, 2, \cdots, r$

  El estadístico de prueba esta dado por:

$$Q{=}\frac{2}{N-1} m log\lambda \sim \chi^2_{1/2p(p+1)(r-1)} \tag{1.40}$$

  con

$$\alpha^* = (\sum_{r=1}^{k} l_r^{-1} - 1)\frac{2p^2+3p-1}{12(p+1)(r-1)}, \tag{1.41}$$

  $m = N - 1 - 2\alpha^*$ y $l_i = \dfrac{(n_i - 1)}{n-1}$

  Se rechaza la hipótesis nula si $Q \succ \chi^2_{p(p+1)(r-1),\alpha}$

- Distribución normal multivariada.

  Mardia (1970) define los coeficientes de simetría y curtosis multivariados, para un vector $X$ de tamaño $p \times 1$ con media $\mu$ y matriz de covarianzas $\Sigma$, mediante las siguientes expresiones:

$$\beta_{1,p} = E(\{(X-\mu)^t \Sigma^{-1} (Y-\mu)\}^3) \tag{1.42}$$

$$\beta_{2,p} = E(\{(X-\mu)^t \Sigma^{-1} (X-\mu)\}^3) \tag{1.43}$$

  Donde $X$ y $Y$ son independientes e identicamente distribuidos. Estas pruebas son invariantes para transformaciones lineales. si $X \sim N_p(\mu, \mathbf{\Sigma})$, entonces, los coeficientes de simetría y curtosis son, respectivamente, $\beta_{1,p} = 0$ y $\beta_{2,p} = p(p+2)$.

  La generalización de las medidas se observa porque $\sqrt{\beta_{1,1}} = \sqrt{\beta_1}$ y $\beta_1 = \beta_{2,1} = \beta_2$.

  Se puede contrastar la hipótesis sobre estos valores empleando los siguientes estimadores muestrales.

$$b_{1,p} = \frac{1}{n^2}\sum_{h=1}^{n}\sum_{i=1}^{n} g_{hi}^3 \tag{1.44}$$

y

$$b_{2,p} = \frac{1}{n}\sum_{i=1}^{n} g_{ii}^2 \tag{1.45}$$

$$g_{h,i} = (x_h - \bar{x})\{\frac{A}{n}\}^{-1}(x_i - \bar{x}) \tag{1.46}$$

$$A = \sum_i (x_i - \bar{x})(x_i - \bar{x})^t \tag{1.47}$$

Mardia en 1970 demuestra que bajo la hipótesis de distribución normal multivariada, se tiene la distribución asintótica de:

$$B_1 = \frac{n}{6} b_{1,p} \sim \chi_f^2 \tag{1.48}$$

donde $f = \frac{1}{6}p(p+1)(p+2)$ Se rechaza la hipótesis de simetría en torno a la media $(H_0 : \sqrt{\beta}_{1,1} = 0)$ si $B_1 \geqslant \chi^2_{\alpha,f}$.

Para verificar que el coeficiente de curtosis no es significativamente diferente de $p(p+2)$ se emplea la estadística

$$B_2 = \frac{b_{2,p} - p(p+2)}{(8p(p+2)/n)^{1/2}} \sim N(0,1) \tag{1.49}$$

Ahora se presentan algunas características de los polinomios fraccionarios, que son necesarias en el desarrollo de la metodología, ajuste de polinomios fraccionarios a curvas de crecimiento que se propone en el capitulo 2.

## 1.3. Polinomios fraccionarios

Los polinomios fraccionarios de grado uno son funciones monótonas y los de grado dos pueden presentar una variedad de formas de curvas, con al menos un máximo o un mínimo. Polinomios fraccionarios de grado tres o de más alto grado son raramente sensibles a valores extremos y la inestabilidad es mucho más alta. los polinomios fraccionarios de grado uno y dos presentan una variedad de formas y ajustan bien los valores extremos de las covariables. Los polinomios fraccionarios para datos transversales han sido estudiados ampliamente por Royston y Altman (1994 y 1997), Sauerbrei, Benner y Royston (2006), Royston y Sauerbrei (2008). Para datos longitudinales, Verbeke y Molengerghs (2006) incluyen polinomios fraccionarios para definir modelos más flexibles a los polinomios convencionales y describen los perfiles de un sujeto específico. Lesafre, Asefa y Verveke (1999) plantean un ejemplo donde los perfiles de un sujeto específico han sido modelados usando polinomios fraccionarios.

La transformación de Box-Tidwell se usa para transformar las variables predictoras, solamente utiliza los exponentes de las variables predictoras. Al utilizar un subconjunto de

exponentes fraccionarios y aplicar la transformación de Box-Tidwell se mejora la medida de bondad del ajuste al comparar varios modelos contendores.

### 1.3.1. Transformación de Box-Tidwell

Box y Tidwell (1962) desarrollan la transformación de Box-Tidwell que se emplea en la transformación de potencias, así:

$$X^{p_j} = \begin{cases} X^{p_j}, & \text{si } p_j \neq 0 \\ \ln X, & \text{si } p_j = 0 \end{cases}$$

Esta transformación incluye muchas de las formas mas comunes, la transformación reciproca y la transformación raíz cuadrada. En datos transversales se hacen transformaciones de las covariables $X$, para mejorar el ajuste de los datos (Tukey, 1957). Al definir $x^0$ como $\log x$, se obtienen transformaciones de funciones de las potencias de la forma $x^p$, para diferentes valores de $p$.

Box-Tidwell (1962), propusieron un método basado en la serie de Taylor, para transformar variables predictoras usando solamente potencias de ellas, ellos consideraron el modelo:

$$y = \beta_0 + \beta_1 w_1 + \cdots + \beta_k w_k + e$$

donde $w_j = x_j^{p_j}$ si $p_j \neq 0$ y $w_j = ln(x_j)$ si $p_j = 0$ El desarrollo de la serie de Taylor se hace con respecto a $\mathbf{p} = (p_1, p_2, \cdots, p_k)$ y alrededor de $\mathbf{p}_0 = (p_{1,0}, \cdots, p_{k,0}) = (1, \cdots, 1)$ luego se tiene

$$y \cong \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \gamma_1 z_1 + \gamma_2 z_2 + \cdots + \gamma_k z_k$$

donde $\gamma_j = (p_j - 1)\beta_j$ y $z_j = x_j \ln(x_j)$ para $j = 1, 2, \cdots, k$, se deben estimar los valores de $p_j$, Los valores de $p_j$ se estiman a través del siguiente procedimiento:

1. Hacer la regresión múltiple considerando las variables predictoras originales $x_j$ y denotar los coeficientes estimados por $b_j$.

2. Hacer una regresión lineal múltiple de $y$ versus las variables predictoras originales más las variables $z = x_j \ln x_j$ y denotar los coeficientes de $z_j$ por $\hat{\gamma}_j$

3. Calcular $\hat{\alpha}_j = \dfrac{\hat{\gamma}_j}{b_j} + 1$

   El procedimiento se debe repetir varias veces usando en cada etapa las nuevas variables transformadas y la siguiente relación:

   $$\hat{p_j}^{(m+1)} = \{\frac{\hat{\gamma_j}^{(m+1)}}{b_j^{(m)}} + 1\}\hat{p_j}^{(m)}$$

   El proceso termina cuando $|\hat{p}_j^{(m+1)} - p_j^{(m)}|$ tiende a cero.

Sin embargo, frecuentemente un solo paso es suficiente.

Mosteller y Tukey (1977) hicieron extensivo el uso de potencias transformadas en análisis de datos, escribiendo entonces una re-expresión de la escala de $x$.

### 1.3.2. Definición y notación de polinomios fraccionarios

$PF_m$, denota un polinomio fraccionario, el subindice $m$ que lo acompaña representa el grado, por ejemplo $PF_1$ denota el polinomio fraccionario de grado 1. Hay que distinguir entre una transformación de $PF$ y un modelo de $PF$. Una transformación $PF_1$ de un argumento positivo $x > 0$ con potencia $p$ definido como $x^p$, donde $p$ pertenece al conjunto de elementos de $S = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, este conjunto de exponentes fue propuesto por Royston y Altman (1994), con el fin de reducir a un conjunto finito de exponentes, las posibles transformaciones que se pueden realizar a las variables dependientes, argumentando las diferentes formas que poseen las curvas cuando se utilizan los diferentes elementos de $S$. El conjunto $S$ puede ser cambiado. Por conveniencia, $x^0$ (es decir con exponente $p = 0$) es igual al logaritmo natural de $x$ más que uno.

Royston y Altman (1994) definen inicialmente un modelo de $PF_1$ de la forma:

$$\varphi_1(x; p) = \beta_0 + \beta_1 x^p = \beta_0 + \varphi_1(x; p) \tag{1.50}$$

donde 1.50 es un polinomios fraccionarios de grado uno, $(PF_1)$ y $p$ es un subconjunto de elementos de $S$.

Posteriormente Royston y Altman (2008) extienden la definición a modelos de polinomios fraccionarios de grado $m$, donde $m$ es un entero mayor o igual que uno. Por ejemplo, un polinomio fraccionario de grado dos, toma un subconjunto de elementos de $S$, $\mathbf{p} = (p_1, p_2)$, donde $p_1, p_2 \epsilon S$, por tanto un modelo de $PF_2$, se define:

$$\varphi_2(x; \mathbf{p}) = \beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2} = \beta_0 + x^{\mathbf{p}}\beta = \beta_0 + \varphi_2(x; p)$$

son generalmente funciones cuadráticas, donde $\mathbf{p}$ es un subconjunto de elementos de $S$.

### 1.3.3. Relación entre la transformación Box-Tidwell, las funciones exponenciales y los polinomios fraccionarios

La transformación de Box-Tidwell y los modelos $PF_s$ tienen algunos elementos en común (por ejemplo:$p = -0.5, 0, 5, -1, 1$), pero en general ellos pueden diferir. Funciones multiexponenciales tales como $\beta_0 + \beta_1 \exp(p_1 x) + \beta_2 \exp p_2 x + \cdots + \beta_m \exp p_m x$ suministran otro enlace con las funciones de PF. Si $z = \exp(x)$, entonces una función multiexponencial con $m$ términos es vista para reemplazar la función $PF_m$ $\beta_0 + \beta_1 z^{p_1} + \beta_2 z^{p_2} + \cdots + \beta_m z^{p_m}$. La diferencia entre el criterio de Box-Tidwell, las funciones multiexponenciales y los $PF_s$ es que con los $PF_s$ los exponentes son restringidos a valores pertenecientes al conjunto $S$.

Una transformación de $PF_2$ de $x$ con los exponentes $\mathbf{p} = (p_1, p_2)$ o para $p_1 = p_2$ (llamados exponentes repetidos) $(p_1, p_1)$ es el vector $x^p$ con

$$x^{(\mathbf{p})} = x^{(p_1,p_2)} = \begin{cases} (x^{p_1}, x^{p_2}), & \text{si } p_1 \neq p_2, \\ (x^{p_1}, x^{p_1} \ln x & \text{si } p_1 = p_2 \end{cases} \tag{1.51}$$

Una modelo de $PF_2$ con vector de parámetros $\beta = (\beta_1, \beta_2)^T$ y exponentes $\mathbf{p}$ es

$$\varphi_2^*(x; \mathbf{p}) = \beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2} = \beta_0 + \varphi_2(x; p)$$

La constante $\beta_0$ es opcional y depende del contexto. Por ejemplo, $\beta_0$ es usualmente includio en modelos normales.

En el caso de exponentes iguales $p_j = p_i$ para al menos un par de índices distintos $(i, j) = (1, 2)$, con $1 \leq i, j \leq m$, $m = 2$ y $p = (p_1, p_2)$

La expresión del modelo puede ser dada por:

$$\beta_0 + \beta_1 x^{(p_j)} + \sum_{j=2}^{m} \beta_j x^{(p_j)} (\ln x)^{(j-1)}$$

La definición general de una función de $PF_m$ con exponentes $\mathbf{p} = (p_1 \leqslant \cdots \leqslant p_m)$. Si $h_0(X) = 1$ y $p_0 = 0$ entonces

$$\varphi_m(x; \beta, p) = \sum_{j=0}^{m} \beta_j h_j(x) \tag{1.52}$$

donde para $j = 1, \ldots, m$,

$$h_j(x) = \begin{cases} x^{p_j}, & \text{si } p_j \neq p_{j-1}, \\ h_{j-1}(x) \ln x & \text{si } p_j = p_{j-1} \end{cases} \tag{1.53}$$

$h_j(x)$ es un vector de funciones $\mathbf{h(x)} = (h_0, h_1, \ldots, h_m)$. Las expresiones (1.52)y (1.53) son una definición de los polinomios fraccionarios de grado $m$.

La figura 1.1 muestra cuatro formas de curvas de funciones de polinomios fraccionarios de grado dos, seleccionadas para dar una idea de la variedad de formas disponibles con una pocas combinaciones de los exponentes $p_1$ y $p_2$. Las curvas ilustran la región al rededor de un máximo o un mínimo y/o un punto silla y representan solamente una porción de valores en el eje $x$. En la gráfica de la figura 1.1, con potencias $\varphi_2 = (x; -0.5, 0)$ se representa una curva cuadrática asimétrica, y la gráfica con potencias, $\varphi_2 = (-1, 3)$ representa una curva cúbica asimétrica. Los dos ejemplos con potencias $(-1, -1)$ tienen asíntota en $x = \infty$, la primera se asemeja a una hipérbola rectangular y la segunda va creciendo rapidamente hacia un máximo.

La habilidad para generar una variedad de curvas, algunas de las cuales tiene asíntotas o tienen un fuerte aumento o caída de una parte y la otra parte casi plana, es una característica particularmente útil de los polinomios fraccionarios.

El uso de $PF$ de alto orden puede ser sensible en el modelamiento univariado, cuando la cantidad de ruido es cero. $PF_3$ o modelos de mas alto orden son raramente sensibles, siendo inestables en muchos casos.

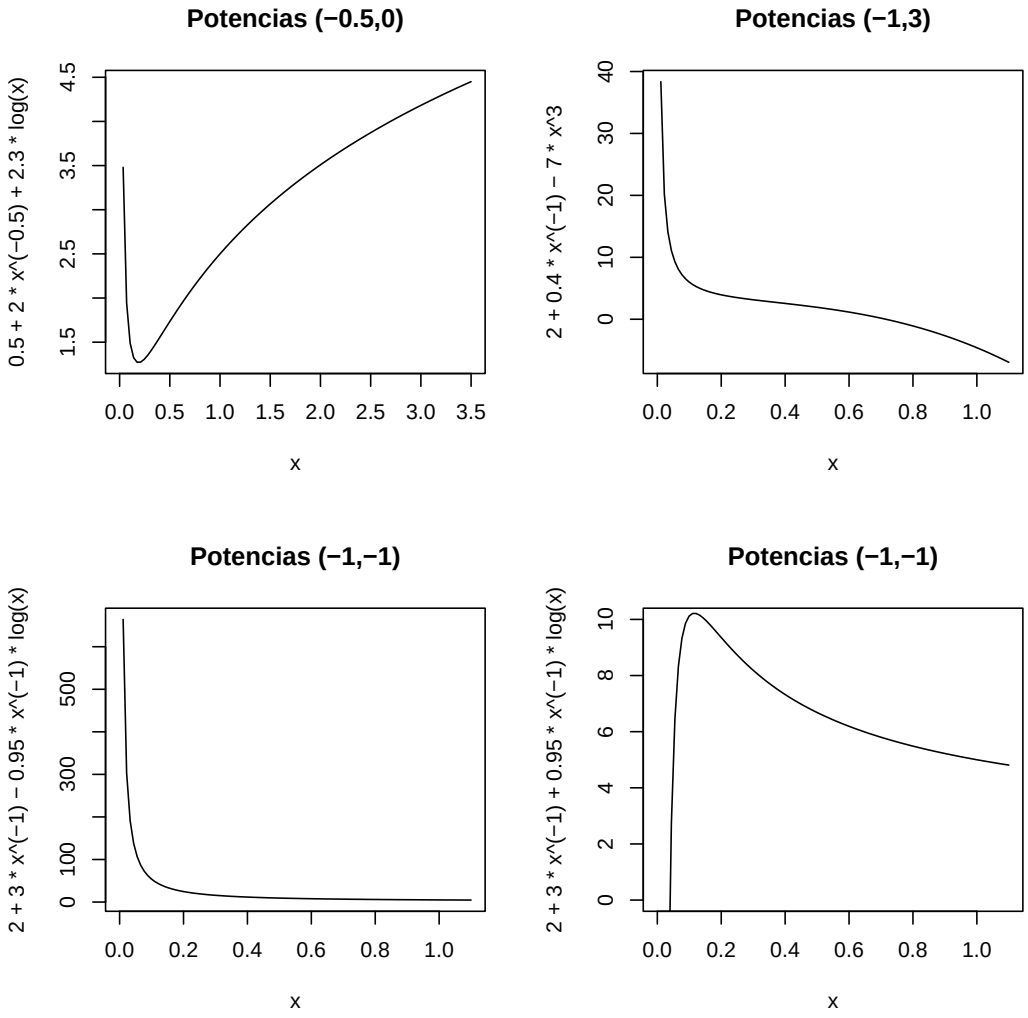


Figura 1.1: Curvas para algunos $PF_2$

Si los valores de $x$ no son positivos, una transformación preliminar de $x$ para asegurar positividad es necesaria. Una alternativa es seleccionar un origen diferente de cero, sea $\zeta < x$ al reescribir la ecuación (5) se obtiene:

$$\varphi_m(x; p_j) = \sum_{j=0}^{m} \beta_j h_j(x - \zeta),$$

Un caso común es cuando $x$ es una variable aleatoria positiva (como las cantidades físicas), pero debido a la imprecisión de las mediciones de los valores de $x$ puede ser cero. Una simple selección de $\zeta$ es menos el intervalo abierto de los valores muestrales de $x$, que garantiza positividad.

### 1.3.4. Selección de exponentes

Para modelar un conjunto de datos de tamaño $n$ usando polinomios fraccionarios es necesario determinar los mejores valores de $m$ y $p$. los valores candidatos de $p$ son todas las posibles m-tuplas seleccionadas con reemplazamiento de un conjunto fijo $S$. Por la experiencia se sugiere que:

$$S = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\} \tag{1.54}$$

el conjunto $S$ incluye los exponentes más comunmente usados. La estimación es mucho más fácil usando un pequeño subconjunto de $S$, por ejemplo para funciones de $PF_2$, solamente 36 modelos, 28 modelos con exponentes diferentes ($p_1 \neq p_2$) y 8 con exponentes iguales ($p_1 = p_2$); necesitan ser ajustados para encontrar las mejores potencias de $S$, en el apéndice B se presentan los diferentes modelos con $PF_1$ y $PF_2$. Los resultados computacionales tienen la ventaja de dar mayor estabilidad de la función y transportabilidad que otros conjuntos. Además modelos con valores extremos de potencias negativas o positivas son muy susceptibles a puntos extremos leves o valores cercanos a cero en la covariable $x$.

En ocasiones es necesario complementar los exponentes de $S$, por ejemplo incluir $1/3$ si la covariable tiene la dimensión del volumén, también si se desea restringir a potencias negativas si la función tiene una asintóta.

### 1.3.5. Características de los PF1 y PF2

Las funciones de $PF_1$ son siempre monótonas (es decir siempre tienen una pendiente positiva o negativa) y cuando sus exponentes son menores que cero ($p < 0$) tienen una asintóta a medida que $x$ tiende a infinito. Las funciones de PF2 pueden ser monótonas o unimodales (es decir tienen un máximo o un mínimo para algún valor positivo de $x$) y tienen una asintóta a medida que $x$ tiende a infinito cuando $p_1$ y $p_2$ son negativos.

Una función $PF_2$, $\beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2}$ con exponentes diferentes, $p_1 \neq p_2$, es monótona, cuando $signo(\beta_1\beta_2) * signo(p_2) = signo(p_1)$ y en otro caso es unimodal. El signo de cero es tomado como positivo. Por ejemplo

$-4x^{-2} + 3\ln x$ es monótona, si $(p_1, p_2) = (-2, 0)$ y

$$\begin{aligned} signo(\beta_1\beta_2) * signo(p_2) &= signo(-4 * 3) * signo(0) \\ &= (-) * (+) \\ &= (-) \\ signo(p_1) &= sig(-) = (-) \\ signo(\beta_1\beta_2) * signo(p_2) &= signo(p_1) \end{aligned}$$

Un modelo de $PF$ unimodal puede tener su máximo o su mínimo fuera del rango de la observacion de los datos, entonces esta función es monótona dentro del rango de observaciones pero no fuera de éste.

**Máximo o mínimo de una función de $PF_2$**

Un modelo de $PF_2$ con exponentes repetidos ($p_1 = p_2$) es siempre unimodal. Cuando una función de PF2 es unimodal, el valor de $x$ en el cual ocurre el máximo o el mínimo, depende de los exponentes $p_1$ y $p_2$ y de la relación $r = -\beta_1/\beta_2$. Por ejemplo la función $-4x^{-2} - 3\ln x$ tiene $r = -4/3$ y su máximo en $x = \{-4/3)(-2)^{-1/(-2)}\} = \sqrt{8/3} = 1.633$

Royston y Sauerbei (2008) definen la existencia de un máximo o un mínimo en $PF_2$, para valores particulares de $p_1$ y $p_2$ en la siguiente tabla:

| Condición de $p_2$ | $p_1 = 0$ | $p_1 \neq 0$ |
|---|---|---|
| $p_2 = 0$ $(p_2 \neq p_1)$ | - | $(rp_1)^{-1/p_1}$ |
| $p_2 \neq 0$ $(p_2 \neq p_1)$ | $(r/p_2)^{1/p_2}$ | $(rp_1/p_2)^{1/(p_2-p_1)}$ |
| $p_2 = p_1$ | $exp(r/2)$ | $exp(r - 1/p_1)$ |

Al ajustar los modelos de curvas de crecimiento de grado uno y dos, del conjunto de exponentes $S$, (1.54) se obtienen 43 modelos, de los cuales se debe seleccionar el modelo que se ajuste mejor a los datos, el criterio empleado para seleccionar el mejor modelo se explica en la siguiente sección.

### 1.3.6. Criterios para la selección de modelos

El método de máxima verosimilitud supone que la forma del modelo es conocida y sólo falta estimar los parámetros. Supongase que se desea estimar un vector de parámetros $\boldsymbol{\Theta} = (\theta_1, \theta_2, \cdots, \theta_p)^t$ y en lugar de un único modelo se tiene una secuencia de modelos $M_1, \cdots, M_i, \cdots, M_p$. Al estimar los parámetros bajo cada modelo por máxima verosimilitud, el modelo con mayor verosimilitud es $M_p$ es decir el modelo con mayor cantidad de parámetros libres. Esta limitación del método de máxima verosimilitud fue detectada por Fisher en 1936, indicando su limitación para comparar modelos distintos. Una solución para seleccionar entre varios modelos contendores es mediante el criterio de información de Akaike (AIC).

**Criterio de información de Akaike (AIC)** Akaike en 1985 propuso un enfoque para solucionar el problema de seleccionar un modelo, al minimizar, la siguiente expresión:

$$AIC = -2L(M_i) + 2p_i = D(M_i) + 2p_i \tag{1.55}$$

donde $p_i$ es el número de parámetros del modelo $M_i$ El criterio consiste en minimizar la suma de la desviación del modelo, que disminuirá si se introducen más parámetros, más el doble del número de parámetros en el modelo, que tiende a corregir este efecto.

Al extrapolar a un tiempo $p + 1$, la calidad del pronóstico del ajuste de polinomios fraccionarios a curvas de crecimiento, se puede medir por medio del Error medio absoluto porcentual (MAPE).

**Error medio absoluto porcentual (MAPE):**

El MAPE es el porcentaje promedio absoluto de error, mide la exactitud de los valores estimados. La exactitud se expresa como un porcentaje con $y_{i,p+1}$ igual al valor observado

en el tiempo $p$, $\hat{y}_{p+1}$ es el valor estimado en el tiempo $p+1$ y $n$ el número de observaciones.

$$MAPE_{(p+1)} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{i,p+1} - \hat{y}_{p+1}}{y_{i,p+1}}\right| 100\,\% \tag{1.56}$$

donde $y_{i,p+1} \neq 0$, para cada $i = 1, 2, \cdots, n$.

# CAPÍTULO 2

# Modelo propuesto

**Introducción**

En este capítulo se describe la metodología de estimación de los parámetros en un modelo de curva de crecimiento utilizando polinomios fraccionarios, bajo el supuesto de distribución normal. En la sección 2.1 se presenta el modelo propuesto, junto con la estructura y características de la matriz $\boldsymbol{X}$. En la sección 2.2 se estiman los parámetros por máxima verosimilitud. Luego en la sección 2.3 se dan las características de los estimadores obtenidos en 2.2. Posteriormente en la sección 2.4 se hacen las pruebas de hipótesis para los parámetros estimados, por último en la sección 2.5 se comparan los modelos obtenidos por medio del criterio de información de Akaike (AIC).

## 2.1. Modelo propuesto

El modelo que se propone para el ajuste de CC por medio de PF sigue la metodología presentada en el capítulo I para datos longitudinales. El grado del polinomio se obtiene como un subcojunto de elementos de $S$, (1.54)

El modelo de curvas de crecimiento se define:

$$\boldsymbol{Y} = \boldsymbol{XBZ} + \boldsymbol{E}$$

donde las matrices para cada uno de los componentes de la ecuación anterior se definieron en el capitulo I, excepto la matriz $\boldsymbol{X}$ que relaciona el grado de la curva y los tiempos. Los exponentes de los tiempos en la matriz $\mathbf{X}$ son un subconjunto de elementos de $S$, de particular interés polinomios fraccionarios de grado uno y dos, dadas las diferentes formas de curvas presentadas en la figura (1.1). Además las columnas de la matriz de errores, $\mathbf{E}$ son independientes, $\mathbf{E} \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma})$, por lo que $\mathbf{Y} \sim (\mathbf{XBZ}, \mathbf{\Sigma})$. Usualmente $p$ es el número de puntos en el tiempo observados sobre cada uno de los $n$ casos, $(m-1)$ es el grado del polinomio en el tiempo y $r$ es número de grupos tratamiento.

Para seleccionar un subconjunto de elementos de $S$ definido en la ecuación, (1.54 ) es conveniente observar el dispersograma y la matriz de correlación para detectar posibles

asíntotas, máximos y o mínimos y así seleccionar un subconjunto apropiado de $S$, evitando el uso de polinomios de orden alto.

Para definir la matriz $\boldsymbol{X}$ se tiene en cuenta la selección de potencias de $S$ que determina el grado del polinomio a ajustar y la matriz de tiempos $\boldsymbol{X}$.

En general para un modelo de polinomios fraccionarios de grado dos, $PF_2$, (1.53), si $\mathbf{p} = (p_1, p_2)$, $p_1 \neq p_2$, la matriz $\boldsymbol{X}$ en CC, es de la forma:

$$\boldsymbol{X} = \begin{pmatrix} 1 & t_1^{p_1} & t_1^{p_2} \\ 1 & t_2^{p_1} & t_2^{p_2} \\ 1 & t_3^{p_1} & t_3^{p_2} \\ \vdots & \vdots & \vdots \\ 1 & t_p^{p_1} & t_{p_2} \end{pmatrix}$$

El modelo de CC asociado al j-ésimo grupo, descrito por un $PF_2$ con elementos diferentes de $S$ es:

$$\mu_j = \beta_{0,j} + \beta_{1,j} t^{p1} + \beta_{2,j} t^{p_2}$$

De otra forma, si las potencias son iguales en un $PF_2$, $\mathbf{p} = (p_1, p_2)$, $p_1 = p_2$, la matriz $\boldsymbol{X}$ en CC es de la forma:

$$\boldsymbol{X} = \begin{pmatrix} 1 & t_1^{p_1} & t_1^{p_1} ln(t_1) \\ 1 & t_2^{p_1} & t_2^{p_1} ln(t_2) \\ 1 & t_3^{p_1} & t_3^{p_1} ln(t_3) \\ \vdots & \vdots & \vdots \\ 1 & t_p^{p_1} & t_p^{p_1} ln(t_p) \end{pmatrix}$$

El modelo de CC asociado al j-ésimo grupo, descrito por un $PF_2$ con potencias iguales es:

$$\mu_j = \beta_{0,j} + \beta_{1,j} t^{p1} + \beta_{2,j} t^{p_1} \ln(t)$$

Un $PF_3$, con $\mathbf{p} = (p_1, p_2, p_3)$, donde $p_1 \neq p_2$ y $p_2 = p_3$, la matriz $X$ en el modelo de CC con PF es de la forma:

$$\boldsymbol{X} = \begin{pmatrix} 1 & t_1^{p_1} & t_1^{p_2} & t_1^{p_2} ln(t_1) \\ 1 & t_2^{p_1} & t_2^{p_2} & t_2^{p_2} ln(t_2) \\ 1 & t_3^{p_1} & t_3^{p_2} & t_3^{p_2} ln(t_3) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_p^{p_1} & t_p^{p_2} & t_p^{p_2} ln(t_p) \end{pmatrix}$$

El modelo de CC asociado al j-ésimo grupo, descrito por un $PF_3$, donde $p_1 \neq p_2$ y $p_2 = p_3$ es:

$$\mu_j = \beta_{0,j} + \beta_{1,j} t^{p1} + \beta_{2,j} t^{p_2} + \beta_{3,j} t^{p_2} \ln(t)$$

La matriz $\boldsymbol{X}$ queda definida para cualquier subconjunto de $S$, siempre que los valores de $\boldsymbol{X}$ sean medidos en tiempos diferentes a cero, si esto se presenta se debe hacer un re-escalamiento de $\boldsymbol{X}$ que permita el uso de la exponente cero o el uso de potencias repetidas. Los polinomios fraccionarios son invariantes a cambios de escala de $\boldsymbol{X}$.

Por ejemplo, una matriz $\boldsymbol{X}$ de $t_p$ tiempos, igualmente espaciados para un subconjunto de $S$ de grado dos y $p = (0.5, 1)$, (determina la existencia de asintótas)

Por tanto la matriz $X$ es de la forma:

$$\boldsymbol{X} = \begin{pmatrix} 1 & t_1^{0.5} & t_1 \\ 1 & t_2^{0.5} & t_2 \\ 1 & t_3^{0.5} & t_3 \\ \vdots & \vdots & \vdots \\ 1 & t_p^{0.5} & t_p \end{pmatrix}$$

Cuando se tienen potencias iguales, $\mathbf{p} = (-0.5, -0.5)$ la matriz $X$ es de la forma:

$$\boldsymbol{X} = \begin{pmatrix} 1 & t_1^{-0.5} & t_1^{-0.5} ln(t_1) \\ 1 & t_2^{-0.5} & t_2^{-0.5} ln(t_2) \\ 1 & t_3^{-0.5} & t_3^{-0.5} ln(t_3) \\ \vdots & \vdots & \vdots \\ 1 & t_p^{-0.5} & t_p^{-0.5} ln(t_p) \end{pmatrix}$$

Con el subconjunto de $S$, se obtienen las posibles combinaciones para construir los posibles modelos de grado dos, $(PF_2)$. Definidas la matriz $\mathbf{X}$ de tiempos, la matriz diseño $\mathbf{Z}$ de unos y ceros, que define la cantidad de grupos y la matriz multivariada de respuestas $\mathbf{Y}$, se procede a estimar los parámetros.

### 2.1.1. Estimación de los parámetros

Cuando la distribución de los $\mathbf{Y}$ es normal multivariada, la estimación de los parámetros se puede hacer por máxima verosimitilud o por mínimos cuadrados generalizados. Si la estimación se hace por MV y las matrices diseño $\mathbf{X}$ y $\mathbf{Z}$ son de rango completo. Las funciones de verosimilitud (1.7) cuyas derivadas parciales con respecto a $\boldsymbol{B}$ y $\boldsymbol{\Sigma}$, se obtienen y se igualan a cero. Para obtener los estimadores de máxima verosimilitud, que son:

$$\hat{\boldsymbol{B}}_{(\boldsymbol{S})} = (\boldsymbol{X}^t \boldsymbol{S}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^t \boldsymbol{S}^{-1} \boldsymbol{Y} \boldsymbol{Z}^t (\boldsymbol{Z}\boldsymbol{Z}^t)^{-1} \tag{2.1}$$

y

$$\hat{\boldsymbol{\Sigma}}_{(\boldsymbol{S})} = \frac{1}{n} (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}\boldsymbol{Z})(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}\boldsymbol{Z})^t \tag{2.2}$$

donde $\boldsymbol{S} = \boldsymbol{Y}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{Z}^t})\boldsymbol{Y}^t$ y $\boldsymbol{P}_{Z^t} = \boldsymbol{Z}^t(\boldsymbol{Z}\boldsymbol{Z}^t)^{-1}\boldsymbol{Z}$.

### 2.1.2. Propiedades de los estimadores

Las propiedades de los estimadores de $\boldsymbol{B}$ son:

El valor esperado de los estimadores maxima verosimilitud de $\boldsymbol{B}$, para los coeficientes de regresión en el modelo de curva de crecimiento se hallan de la siguiente manera:

$$\begin{aligned} E(\hat{\boldsymbol{B}}) &= E_S[E_Y(\hat{\boldsymbol{B}}/\boldsymbol{S}] \\ &= E_S[(\boldsymbol{X^T S^{-1} X})^{-1}\boldsymbol{X^T S^{-1}} E_Y(\boldsymbol{Y/S})\boldsymbol{Z}^t(\boldsymbol{ZZ}^t)^{-1}] \\ &= E_S[(\boldsymbol{X^t S^{-1} X})^{-1}(\boldsymbol{X^t S^{-1} X})\boldsymbol{BZZ}^t(\boldsymbol{ZZ}^t)^{-1}] \\ &= E_S(\boldsymbol{B}) = \boldsymbol{B} \end{aligned}$$

Por tanto es un estimador insesgado.

La matriz de varianzas y covarianza de el estimador máximo verosimil de $\hat{\boldsymbol{B}}$ se expresa como:

$$Cov(\hat{\boldsymbol{B}}) = \frac{n-r-1}{n-r-(p-m)-1}(\boldsymbol{ZZ}^t)^{-1} \otimes (\boldsymbol{X}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}$$

Para el componente de dispersión $\boldsymbol{\Sigma}$, la distribución de los estimadores MV $\hat{\boldsymbol{\Sigma}}$ es mucho mas complicada que $\hat{\boldsymbol{B}}$ (Gleser y Olkin, 1970), por esto solo se estudia el valor esperado y la matriz de varianzas y covarianzas de $\boldsymbol{\Sigma}$, obtenidas de von Rosen (1991). Sea $\hat{\boldsymbol{\Sigma}}$ el estimador MV del componente de dispersión $\boldsymbol{\Sigma}$ en el modelo de CC, si $n > max\{p+r, 2(p-m)+r+1\}$ entonces:

$$E(\hat{\boldsymbol{\Sigma}}) = \boldsymbol{\Sigma} - \frac{r}{c}\frac{n-r-2(p-m)-1}{n-r-(p-m)-1}\boldsymbol{X}(\boldsymbol{X^t\Sigma^{-1}X})^{-1}\boldsymbol{X^t} \tag{2.3}$$

$\hat{\boldsymbol{\Sigma}}$ es un estimador sesgado de $\boldsymbol{\Sigma}$ y su sesgo esta dado por:

$$\boldsymbol{\Sigma_k} = \frac{r}{c}\frac{n-r-2(p-m)-1}{n-r-(p-m)-1}\boldsymbol{X}(\boldsymbol{X^t\Sigma^{-1}X})^{-1}\boldsymbol{X}^t \tag{2.4}$$

que esta en función del parámetro desconocido $\boldsymbol{\Sigma}$. Al tomar $\hat{\boldsymbol{\Sigma}}$, se puede construir un estimador insesgado de $\boldsymbol{\Sigma}$. Si $n > max\{p+r, 2(p-m)+r+1\}$ entonces:

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{0}} = \hat{\boldsymbol{\Sigma}} + \frac{r}{n-r-(p-m)}\frac{n-r-2(p-m)-1}{n-r-(p-m)-1}\boldsymbol{X}(\boldsymbol{X}^t\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^t \tag{2.5}$$

$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{0}}$ es un estimador insesgado de $\hat{\boldsymbol{\Sigma}}$. El estimador insesgado del componente de dispersión, $\boldsymbol{\Sigma}$ no es único. Por ejemplo, si $\mathbf{S} = \boldsymbol{Y(I_n - P_Z^t)Y^t}$ sigue una distribución Wishard, $W_p(n-r, \Sigma)$, se tiene que $E(\boldsymbol{S}/(n-r)) = \hat{\boldsymbol{\Sigma}}$, es decir

$$\frac{1}{n-r}\boldsymbol{S} = \frac{1}{n-r}\boldsymbol{Y}(\boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{Z}^t})\boldsymbol{Y}^t \tag{2.6}$$

es otro estimador insesgado de $\boldsymbol{\Sigma}$. La estadística (2.6) es comunmente usada como un estimador de $\boldsymbol{\Sigma}$ en componentes principales, correlación canónica y análisis factorial. Sin embargo, si la estructura de valor medio como en el modelo de CC existente, entonces la estimación (2.6) es algo inusual porque $\boldsymbol{\Sigma}$ es la variación cuadrática alrededor del modelo $\boldsymbol{XBZ_i}$, donde $\boldsymbol{Z_i}$ es la i-ésima columna de $\boldsymbol{Z}$, mientras que en (2.6) describe solamente la variación alrededor de la media muestral. La media muestral varia alrededor del modelo estimado, pero este no es utilizado cuando se aplica (2.6). En la estimación por máxima

verosimilitud, se tiene en cuenta las estructuras de covarianza más apropiada para los datos. En el capítulo I, se citan las estructuras de covarianzas para modelos de CC. Entre dichas estructuras se encuentran las estructuras de covarianzas de Rao, que hacen que los estimadores de MCG coincidan con los obtenidos con MV. Dentro de las estructuras de covarianzas de Rao, utilizadas para modelos de curvas de crecimiento se encuentran la estructura de covarianzas uniforme y la estructura de coeficientes de regresión aleatoria.

### 2.1.3. Pruebas de hipótesis

**Hipótesis de interés** Se considera la hipótesis general para el caso multivariado.

$$H_0 : \mathbf{ABC} = \mathbf{D} \tag{2.7}$$

donde,

- $\mathbf{B^t}$: Matriz de coeficientes traspuesta, presentada en(1.23)
- $\mathbf{A}$: Matriz de coeficientes, permite probar la hipótesis dentro de los tiempos de orden (hipótesis sobre las elementos dentro de las columnas de $\mathbf{B^t}$) .
- $\mathbf{C}$: Matriz de coeficientes, permite probar la hipótesis entre tiempos (hipótesis sobre los elementos dentro de las filas de $\mathbf{B^t}$).
- $\mathbf{D}$: Es una matriz de constantes.

Para la respuesta media en el grupo $r$, en el t-ésimo tiempo, se consideran la hipótesis:

- $H_{01}$: Las $k$ curvas de crecimiento son de grado $(m-1)$

**Estadística de prueba**

Teniendo en cuenta que para ajustas curvas de crecimiento por medio de polinomios fraccionarios, se expresa grado uno cuando solamente se estima el modelo a través de una sola potencia, si se ajusta el modelo a través de dos potencias, se dice que el polinomio es de grado dos.

Para probar la hípótesis (2.7) existe una amplia variedad de estadísticas, halladas a través de (1.35) y (1.36), mostradas en detalle en el capitulo I. Las estadísticas utilizadas, han sido programas en el software R, se encuentra en el apéndice C y son:

- Lambda de Wilks:

$$\Lambda = \prod_{i=1}^{p} \frac{1}{(1+l_i)} \tag{2.8}$$

donde $l_i$ son las soluciones de la ecuación característica, $|\mathbf{Q_h} - l\mathbf{Q_e}|$

- Traza de Lawley-Hotelling:

$$\mathbf{U} = \sum l_i = traza(\mathbf{Q_h}\mathbf{Q_e})^{-1} \tag{2.9}$$

- Traza de Bartlett-Nanda-Pillai:

$$\mathbf{V} = \sum_{i=1}^{p} \frac{l_i}{(1+l_i)} = traza(\mathbf{Q_h}(\mathbf{Q_e}+\mathbf{Q_h})^{-1}) \tag{2.10}$$

- Maximo valor propio de Roy:

$$\mathbf{R} = \frac{l_1}{(1+l_1)} \tag{2.11}$$

Por medio de estas estadísticas se puede determinar si el modelo es significativo y además que parámetros son significativos en el modelo como se describió en el capítulo I.

### 2.1.4. Seleccion del modelo

Una forma de seleccionar entre varios modelos contendores es mediante el criterio de información de Akaike (AIC), este criterio consiste en minimizar la suma de la desviación del modelo, que disminuira si se introducen más parámetros, mas el doble del número de parámetros en el modelo, que tiende a corregir este efecto.

$$AIC = -2L(M_i) + 2p_i = D(M_i) + 2p_i \tag{2.12}$$

donde $p_i$ es el número de parámetros del modelo $M_i$

El criterio AIC permite obtener el modelo que proporcione mejores predicciones entre los modelos existentes.

Cuando se hacen pronósticos sobre modelos truncados, la calidad del ajuste se analiza a través del $MAPE$, definido en el capitulo I.

# CAPÍTULO 3

# Aplicación

## 3.1. Datos de glucosa

El departamento de Pediatría de la Escuela de Medicina en la Universidad de Colorado, en un estudio de la asociación de hiperglucemia e hiperinsulina , realizó pruebas sobre la tolerancia a la insulina. Estas pruebas se realizaron en 13 pacientes control y 20 pacientes obesos, en esta investigación se tomaron las mediciones de fosfato inorgánico en plasma (mg/dl), obtenidas a partir de muestras de sangre extraídas a 0, 0.5, 1, 1.5, 2, 3, 4 y 5 horas después de administrar la dosis estándar de glucosa. Estos datos revelaron en el estudio de Zerbe, que las curvas de pacientes obesos y pacientes control difieren significativamente solamente durante las primeras tres horas después de administrada la glucosa. La administración de glucosa determina una disminución en el fosfato inorgánico del plasma debido exlusivamente a la acción de la insulina.

Los datos de glucosa para pacientes obesos, que se encuentran en el apéndice B, fueron reportados incialmente por Chi y Reinsel en 1989 y Keramides y Lee en 1995. El objetivo para este conjunto de datos en esta tesis es ajustar modelos de curvas de crecimiento utilizando polinomios fraccionarios para luego comparar los modelos obtenidos con polinomios convencionales y polinomios fraccionarios.

### 3.1.1. Análisis exploratorio de los datos de glucosa para pacientes obesos

Se presentan algunos gráficos que permiten analizar las características de los datos.

**Gráfico de perfiles**

En la grafica 3.1, se observa en el tiempo cero niveles de fosfato un tanto atípicos, disminución del fosfato inorgánico después de la media hora de administrada la glucosa hasta la hora dos, luego en la hora tres se presenta un incremento en los niveles de fosfato inorgánico hasta la hora cinco.

**Diagrama de caja**

En el grafico (3.2), se observa la presencia de datos atípicos en el tiempo 0. También

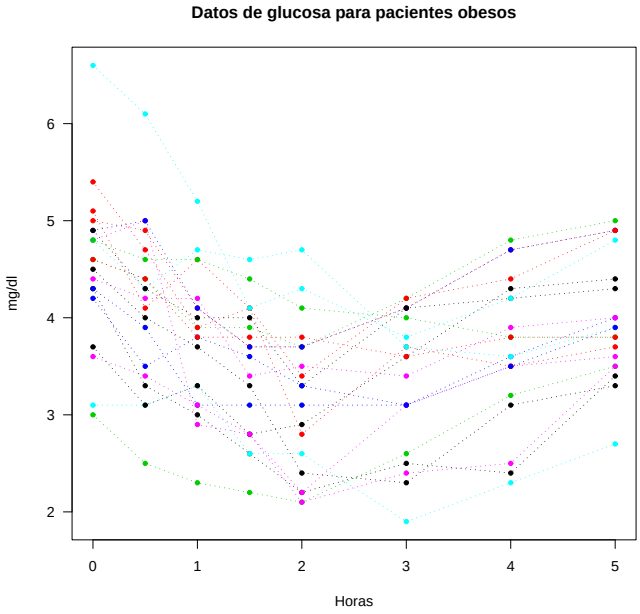


Figura 3.1: Gráfico de perfiles para datos de glucosa en pacientes obesos.

se observa una dispersión similar en cada uno de los tiempos, junto con una asimetría negativa en la mayoría de los tiempos, una disminución del fosfato inorgánico del plasma desde la media hora hasta la segunda hora, luego éste aumenta en la hora tres hasta la hora cinco.

**La matriz de varianzas y covarianzas**

$$\begin{pmatrix} 0.6453684 & 0.6155789 & 0.4401053 & 0.4027368 & 0.3670000 & 0.3771053 & 0.3200000 & 0.2942632 \\ 0.6155789 & 0.6951579 & 0.4182105 & 0.3782105 & 0.3849474 & 0.3721053 & 0.3468421 & 0.2851579 \\ 0.4401053 & 0.4182105 & 0.5069474 & 0.4311579 & 0.4704211 & 0.3405263 & 0.2873684 & 0.2487368 \\ 0.4027368 & 0.3782105 & 0.4311579 & 0.4564211 & 0.4398947 & 0.3710526 & 0.2873684 & 0.2234737 \\ 0.3670000 & 0.3849474 & 0.4704211 & 0.4398947 & 0.5899737 & 0.4030263 & 0.3647368 & 0.2879737 \\ 0.3771053 & 0.3721053 & 0.3405263 & 0.3710526 & 0.4030263 & 0.5103947 & 0.4831579 & 0.3909211 \\ 0.3200000 & 0.3468421 & 0.2873684 & 0.2873684 & 0.3647368 & 0.4831579 & 0.5505263 & 0.4378947 \\ 0.2942632 & 0.2851579 & 0.2487368 & 0.2234737 & 0.2879737 & 0.3909211 & 0.4378947 & 0.4055526 \end{pmatrix} \tag{3.1}$$

**Dispersogramas**

En el dispersograma (3.3) se observa que para algunos tiempos la correlación en tiempos adyacentes es mayor que en tiempos posteriores. En la matriz de correlación 3.2 se calcula la correlación exacta entre los diferentes tiempos.

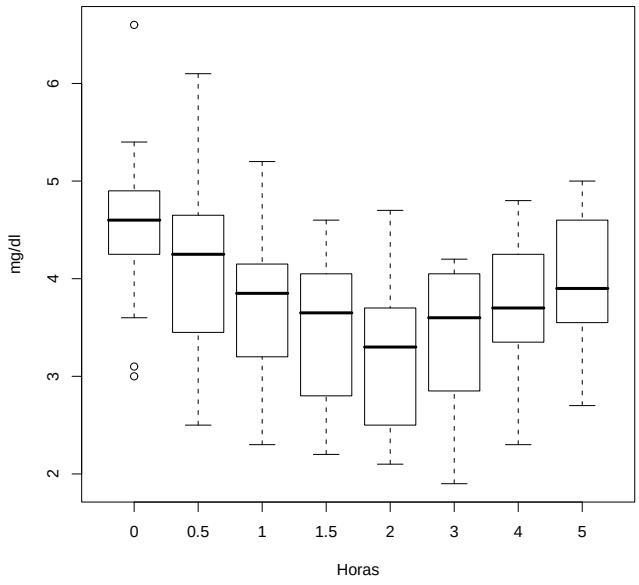


Figura 3.2: Diagrama de caja para datos de glucosa en pacientes obesos en diferentes tiempos de medición

$$\begin{pmatrix} 1.0000 & 0.9190 & 0.7694 & 0.7420 & 0.5947 & 0.6570 & 0.5368 & 0.5751 \\ 0.9190 & 1.0000 & 0.7044 & 0.6714 & 0.6010 & 0.6246 & 0.5606 & 0.5370 \\ 0.7694 & 0.7044 & 1.0000 & 0.8963 & 0.8601 & 0.6694 & 0.5439 & 0.5485 \\ 0.7420 & 0.6714 & 0.8963 & 1.0000 & 0.8477 & 0.7687 & 0.5732 & 0.5194 \\ 0.5947 & 0.6010 & 0.8601 & 0.8477 & 1.0000 & 0.7344 & 0.6399 & 0.5887 \\ 0.6570 & 0.6246 & 0.6694 & 0.7687 & 0.7344 & 1.0000 & 0.9114 & 0.8592 \\ 0.5368 & 0.5606 & 0.5439 & 0.5732 & 0.6399 & 0.9114 & 1.0000 & 0.9267 \\ 0.5751 & 0.5370 & 0.5485 & 0.5194 & 0.5887 & 0.8592 & 0.9267 & 1.0000 \end{pmatrix} \tag{3.2}$$

En la matriz de correlación 3.2, se observa una correlación alta en la primera banda por lo que la estructura de covarianzas $AR1$, resulta conveniente. Se puede identificar de una manera mas acertada el calculo del $AIC$ para cada una de las estructuras de covarianzas

Para poder realizar un estudio acertado de los datos se debe verificar los supuestos del modelo, para luego estimar los parámetros por máxima verosimilitud.

### 3.1.2. Verificación de los supuestos del modelo

- **Normalidad** La metodología planteada se desarrolla bajo el supuesto de normalidad multivariada, para verificarla se utiliza la estadística de Mardia que se encuentra detallada en el capitulo I.

  El estadístico de Mardia proporciona un p-valor de 0.2816682 por tanto no rechaza el supuesto de normalidad multivariada.

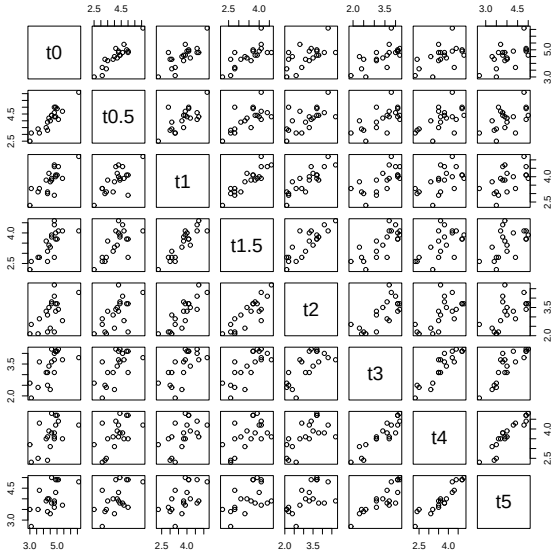


Figura 3.3: Dispersogramas para datos de glucosa en pacientes obesos para los diferentes tiempos de medición

- **Homocedasticidad** En la metodología planteada también es necesario la verificación del supuesto de homocedasticidad de varianzas, se hace a través del estadístico de Bartlett, que se mostro en el capítulo I. En esta aplicación no se calcula porque se tiene solamente una muestra.

Es posible hacer una traslación del tiempo una unidad, para que sea posible utilizar la transformación de Box-Tidwell, de tal manera que los tiempos quedan definidos así: 1, 1.5, 2, 2.5, 3, 4, 5 y 6.

Ahora se estima la estructura de covarianzas.

### 3.1.3. Selección de la matriz de varianzas y covarianzas

Las estructuras de covarianzas que se obtienen a continuación son las más utilizadas en modelos de curvas de crecimiento (no estructurada, uniforme, aleatoria, de Rao y autoregresiva de orden uno).

- **Estructura de covarianzas no estructurada, (NE)**

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 12.262 & 11.696 & 8.362 & 7.652 & 6.973 & 7.165 & 6.08 & 5.591 \\ 11.696 & 13.208 & 7.946 & 7.186 & 7.314 & 7.070 & 6.59 & 5.418 \\ 8.362 & 7.946 & 9.632 & 8.192 & 8.938 & 6.470 & 5.46 & 4.726 \\ 7.652 & 7.186 & 8.192 & 8.672 & 8.358 & 7.050 & 5.46 & 4.246 \\ 6.973 & 7.314 & 8.938 & 8.358 & 11.209 & 7.657 & 6.93 & 5.471 \\ 7.165 & 7.070 & 6.470 & 7.050 & 7.657 & 9.697 & 9.18 & 7.427 \\ 6.080 & 6.590 & 5.460 & 5.460 & 6.930 & 9.180 & 10.46 & 8.320 \\ 5.591 & 5.418 & 4.726 & 4.246 & 5.471 & 7.427 & 8.32 & 7.705 \end{pmatrix}$$

- **Estructura de covarianzas uniforme, (UNIFORME)**

  Se debe estimar $\rho$ y $\sigma^2$:

$$\hat{\rho} = 0.6707411 \tag{3.3}$$

$$\hat{\sigma^2} = 0.5275347 \tag{3.4}$$

  La estructura de covarianzas uniforme es:

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 0.527 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 \\ 0.353 & 0.527 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.3538 \\ 0.353 & 0.353 & 0.527 & 0.353 & 0.353 & 0.353 & 0.353 & 0.3538 \\ 0.353 & 0.353 & 0.353 & 0.527 & 0.353 & 0.353 & 0.353 & 0.3538 \\ 0.353 & 0.353 & 0.353 & 0.353 & 0.527 & 0.353 & 0.353 & 0.3538 \\ 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.527 & 0.353 & 0.3538 \\ 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.527 & 0.3538 \\ 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.5275 \end{pmatrix}$$

- **Estructura de covarianzas aleatoria, (ALEATORIA)**

  Se debe estimar $\Gamma$ y $\sigma^2$

$$\hat{\mathbf{\Gamma}} = \begin{pmatrix} 1.03619287 & -0.35340277 & 0.036931511 \\ -0.35340277 & 0.21981545 & -0.026671731 \\ 0.03693151 & -0.02667173 & 0.003529274 \end{pmatrix}$$

$$\hat{\sigma}^2 = 0.0814478 \tag{3.5}$$

  Luego la estructura de covarianzas aleatoria es:

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 0.65469 & 0.5103 & 0.4543 & 0.4052 & 0.3630 & 0.2993 & 0.2631 & 0.2545 \\ 0.5103 & 0.5560 & 0.4412 & 0.4103 & 0.3818 & 0.3322 & 0.2923 & 0.2621 \\ 0.4543 & 0.4412 & 0.5084 & 0.4116 & 0.3950 & 0.3585 & 0.3174 & 0.2717 \\ 0.4052 & 0.4103 & 0.4116 & 0.4904 & 0.4026 & 0.3782 & 0.3384 & 0.2832 \\ 0.3630 & 0.3818 & 0.3950 & 0.4026 & 0.4859 & 0.3912 & 0.3553 & 0.2968 \\ 0.2993 & 0.3322 & 0.3585 & 0.3782 & 0.3912 & 0.4787 & 0.3768 & 0.3298 \\ 0.2631 & 0.2923 & 0.3174 & 0.3384 & 0.3553 & 0.3768 & 0.4634 & 0.3707 \\ 0.2545 & 0.2621 & 0.2717 & 0.2832 & 0.2968 & 0.3298 & 0.3707 & 0.5009 \end{pmatrix}$$

- **Estructura de covarianzas de Rao, (ECR)** Se deben estimar $\mathbf{Q}\hat{\Theta}\mathbf{Q}^T$ y $\mathbf{X}\Gamma\mathbf{X}^T$.

$$\mathbf{Q}\hat{\Theta}\mathbf{Q}^T = \begin{pmatrix} 0.0206 & -0.016 & -0.009 & 0.0016 & -0.014 & 0.020 & 0.008 & -0.009 \\ -0.0169 & 0.084 & -0.049 & -0.0354 & -0.023 & 0.028 & 0.041 & -0.028 \\ -0.0094 & -0.049 & 0.069 & 0.0121 & 0.019 & -0.041 & -0.025 & 0.024 \\ 0.0016 & -0.035 & 0.012 & 0.0488 & -0.003 & -0.012 & -0.031 & 0.020 \\ -0.0146 & -0.023 & 0.019 & -0.0037 & 0.098 & -0.064 & -0.052 & 0.040 \\ 0.0202 & 0.028 & -0.041 & -0.0128 & -0.064 & 0.068 & 0.036 & -0.034 \\ 0.0084 & 0.041 & -0.025 & -0.0314 & -0.052 & 0.036 & 0.067 & -0.044 \\ -0.0099 & -0.028 & 0.024 & 0.0208 & 0.040 & -0.034 & -0.044 & 0.031 \end{pmatrix}$$

$$\mathbf{X}\Gamma\mathbf{X}^T = \begin{pmatrix} 0.6227 & 0.5411 & 0.4697 & 0.4085 & 0.3574 & 0.2857 & 0.2547 & 0.2644 \\ 0.5411 & 0.4973 & 0.4570 & 0.4203 & 0.3871 & 0.3313 & 0.2897 & 0.2623 \\ 0.4697 & 0.4570 & 0.4426 & 0.4265 & 0.4087 & 0.3679 & 0.3201 & 0.2655 \\ 0.4085 & 0.4203 & 0.4265 & 0.4271 & 0.4221 & 0.3953 & 0.3459 & 0.2741 \\ 0.3574 & 0.3871 & 0.4087 & 0.4221 & 0.4274 & 0.4136 & 0.3671 & 0.2881 \\ 0.2857 & 0.3313 & 0.3679 & 0.3953 & 0.4136 & 0.4229 & 0.3958 & 0.3323 \\ 0.2547 & 0.2897 & 0.3201 & 0.3459 & 0.3671 & 0.3958 & 0.4060 & 0.3979 \\ 0.2644 & 0.2623 & 0.2655 & 0.2741 & 0.2881 & 0.3323 & 0.3979 & 0.4851 \end{pmatrix}$$

La estructura de covarianzas de Rao es:

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 0.6434 & 0.5242 & 0.4603 & 0.4101 & 0.3428 & 0.3060 & 0.2632 & 0.2545 \\ 0.5242 & 0.5819 & 0.4072 & 0.3849 & 0.3634 & 0.3600 & 0.3310 & 0.2337 \\ 0.4603 & 0.4072 & 0.5118 & 0.4387 & 0.4286 & 0.3267 & 0.2950 & 0.2899 \\ 0.4101 & 0.3849 & 0.4387 & 0.4760 & 0.4184 & 0.3824 & 0.3144 & 0.2950 \\ 0.3428 & 0.3634 & 0.4286 & 0.4184 & 0.5255 & 0.3492 & 0.3149 & 0.3288 \\ 0.3060 & 0.3600 & 0.3267 & 0.3824 & 0.3492 & 0.4910 & 0.4320 & 0.2974 \\ 0.2632 & 0.3310 & 0.2950 & 0.3144 & 0.3149 & 0.4320 & 0.4733 & 0.3536 \\ 0.2545 & 0.2337 & 0.2899 & 0.2950 & 0.3288 & 0.2974 & 0.3536 & 0.5169 \end{pmatrix}$$

- **Estructura de covarianzas Autoregresiva de orden uno,(AR1)**

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 0.52753 & 0.35383 & 0.23733 & 0.1591 & 0.1067 & 0.07161 & 0.04803 & 0.03222 \\ 0.35383 & 0.52753 & 0.35383 & 0.2373 & 0.1591 & 0.10677 & 0.07161 & 0.04803 \\ 0.23733 & 0.35383 & 0.52753 & 0.3538 & 0.2373 & 0.15919 & 0.10677 & 0.07161 \\ 0.15919 & 0.23733 & 0.35383 & 0.5275 & 0.3538 & 0.23733 & 0.15919 & 0.10677 \\ 0.10677 & 0.15919 & 0.23733 & 0.3538 & 0.5275 & 0.35383 & 0.23733 & 0.15919 \\ 0.07161 & 0.10677 & 0.15919 & 0.2373 & 0.3538 & 0.52753 & 0.35383 & 0.23733 \\ 0.04803 & 0.07161 & 0.10677 & 0.1591 & 0.2373 & 0.35383 & 0.52753 & 0.35383 \\ 0.03222 & 0.04803 & 0.07161 & 0.1067 & 0.1591 & 0.23733 & 0.35383 & 0.52753 \end{pmatrix} \tag{3.6}$$

Para determinar cuál estructura de covarianzas de las dadas anteriormente se ajusta mejor a los datos, se compara los modelos obtenidos con cada estructura por medio del criterio de información de Akaike (AIC).

| **Estructura de covarianzas** | **AIC** |
|---|---|
| NE | 514.9952 |
| UNIFORME | 240.9987 |
| ALEATORIA | 386.3783 |
| AR1 | 224.9218 |

La estructura de covarianzas de AR1 presenta el menor AIC, por tanto es la estructura más apropiado para este conjunto de datos.

### 3.1.4. Estimación de los parámetros

Como se cumplen los supuestos de normalidad multivariada. Se procede a estimar los parámetros del modelo.

#### Estimación por máxima verosimilitud con polinomios clásicos

Se estiman todos los modelos de curvas de crecimiento de polinomios y polinomios fraccionarios definidos a través de subconjuntos de uno y dos elementos de $S$ y se seleccionan los que tienen menor $AIC$.

**Estimación máxima verosimilitud con polinomios** Para el modelo estimado con $p = (1, 2)$ y mayor $AIC$, $\hat{\mathbf{y}} = \hat{\beta}_{\mathbf{0}} + \hat{\beta}_{\mathbf{1}}\mathbf{t} + \hat{\beta}_{\mathbf{2}}\mathbf{t^2}$,se determina si sus parámetros son significativos. Se utiliza la estadística Lambda de Wilks. Las demás estadísticas citadas en el capítulo 2 se pueden obtener a través del software R. Los comandos se encuentra detallados en el apéndice C.

| Parámetro | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_0$ | 0.03542442 | 517.3531 | 1 | 19 | 2.997602e-15 |
| $\beta_1$ | 0.1621006 | 98.21118 | 1 | 19 | 6.085376e-09 |
| $\beta_2$ | 0.1477294 | 109.6135 | 1 | 19 | 2.500219e-09 |

El modelo estimado obtenido por máxima verosimilitud y con estructura de covarianzas de $AR1$, es:

$$\hat{\mathbf{y}} = 5.426663 - 1.10637\mathbf{t} + 0.1460229\mathbf{t}^2$$

Se calcula el error estándar para cada parámetro estimado.

| Parámetro | Estimación | Error estándar |
|---|---|---|
| $\beta_0$ | 5.426663 | 1.066975 |
| $\beta_1$ | -1.10637 | 0.4992723 |
| $\beta_2$ | 0.1460229 | 0.06237404 |

La matriz de covarianza de los $\hat{\mathbf{B}}$

$$\mathbf{cov}(\hat{\mathbf{B}}) = \begin{pmatrix} 0.063611700 & -0.019671871 & 0.0019004120 \\ -0.019671871 & 0.012536799 & -0.0015045413 \\ 0.001900412 & -0.001504541 & 0.0001946744 \end{pmatrix}$$

**Estimación máxima verosimilitud con polinomios fraccionarios**

Se ajustan polinomios fraccionarios de grado uno y dos. Se utilizan los comandos de R que se encuentran en el apéndice C; para obtener los diferentes subconjuntos de $S$ y las matrices de tiempos, $\mathbf{X}$ y los mejores modelos de curvas de crecimiento con polinomios fraccionarios.

La matriz $\mathbf{X}$ para $\mathbf{p} = (\mathbf{1}, \mathbf{1})$ es:

$$\mathbf{X} = \begin{pmatrix} 1 & 1.0 & 0.0000000 \\ 1 & 1.5 & 0.6081977 \\ 1 & 2.0 & 1.3862944 \\ 1 & 2.5 & 2.2907268 \\ 1 & 3.0 & 3.2958369 \\ 1 & 4.0 & 5.5451774 \\ 1 & 5.0 & 8.0471896 \\ 1 & 6.0 & 10.7505568 \end{pmatrix}$$

Se determina si los parámetros del modelo $\hat{\mathbf{y}} = \hat{\beta}_{\mathbf{0}} + \hat{\beta}_{\mathbf{1}}\mathbf{t} + \hat{\beta}_{\mathbf{2}}\mathbf{t}\ln(\mathbf{t})$ son significativos.

| | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_0$ | 0.04171585 | 436.4624 | 1 | 19 | 1.44329e-14 |
| $\beta_1$, $p = 1$ | 0.1479845 | 109.3919 | 1 | 19 | 2.841869e-09 |
| $\beta_2$, $p = 1$ | 0.1462061 | 110.9535 | 1 | 19 | 2.264054e-09 |

El intercepto y los coeficiente para el tiempo, al utiliza las potencias $p = (1, 1)$ son significativos.

El modelo estimado por máximo verosimil con potencias $(1, 1)$ es:

$$\hat{\mathbf{y}} = 6.6112475 - 2.0969653\mathbf{t} + 0.9312264\mathbf{t}ln(\mathbf{t})$$

En la tabla siguiente se dan los parámetros estimados junto con sus errores estándar:

| Parámetro | Estimación | Error estándar |
|---|---|---|
| $\beta_0$ | 6.6112475 | 1.415220925 |
| $\beta_1$, $p = 1$ | -2.0969653 | 0.896629 |
| $\beta_2$, $p = 1$ | 0.9312264 | 0.395366 |

La matriz de covarianzas para los parámetros estimados es:

$$\mathbf{cov}(\hat{\mathbf{B}}) = \begin{pmatrix} 0.10373736 & -0.05321224 & 0.021753082 \\ -0.05321224 & 0.04048692 & -0.017806166 \\ 0.02175308 & -0.01780617 & 0.007987907 \end{pmatrix}$$

Otro modelo de polinomios fraccionarios de mejor ajuste es el que se obtiene para $p = (0.5, 2)$ con matriz $\mathbf{X}$:

$$\mathbf{X} = \begin{pmatrix} 1 & 1.000000 & 1.00 \\ 1 & 1.224745 & 2.25 \\ 1 & 1.414214 & 4.00 \\ 1 & 1.581139 & 6.25 \\ 1 & 1.732051 & 9.00 \\ 1 & 2.000000 & 16.00 \\ 1 & 2.236068 & 25.00 \\ 1 & 2.449490 & 36.00 \end{pmatrix}$$

Se determina si los parámetros del modelo $\hat{\mathbf{y}} = \hat{\beta}_\mathbf{0} + \hat{\beta}_\mathbf{1}\mathbf{t}^{\mathbf{0.5}} + \hat{\beta}_\mathbf{2}\mathbf{t}^\mathbf{2}$ son significativos.

| Parámetro | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_0$ | 0.0480579 | 376.3565 | 1 | 19 | 5.540013e-14 |
| $\beta_1, p = 0.5$ | 0.1558084 | 102.9447 | 1 | 19 | 4.163889e-09 |
| $\beta_2, p = 2$ | 0.1497829 | 107.8503 | 1 | 19 | 2.853645e-09 |

El ajuste máximo verosimil del $PF_2$ con potencias $(0.5, 2)$ esta dado por:

$$\hat{\mathbf{y}} = 6.8892850 - 2.4786748\mathbf{t}^{0.5} + 0.0895537\mathbf{t}^2$$

En la grafica (3.4) se muestran los mejores modelos estimados con polinomios clásicos al utilizar $p = (1, 2)$ y con polinomios fraccionarios al utilizar $p = (1, 1)$ y $p = (0.5, 2)$.

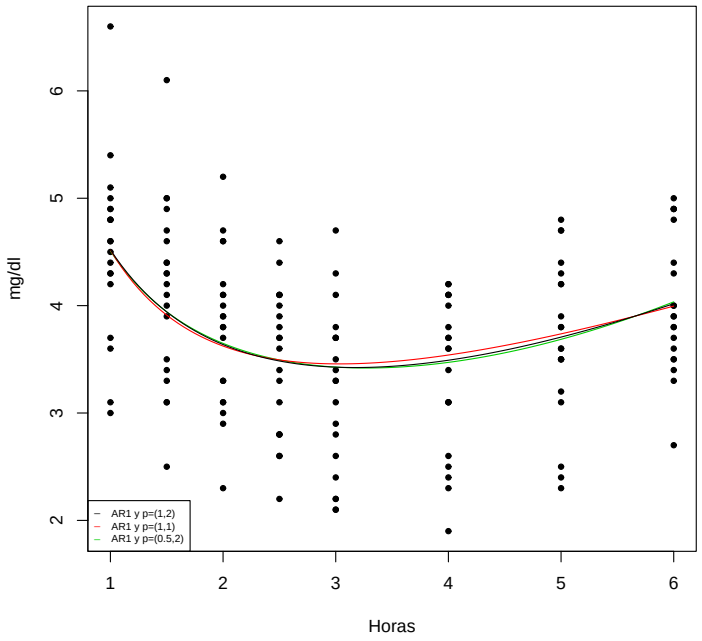


Figura 3.4: Ajustes por máxima verosimilitud con polinomios convencionales y fraccionarios, para los datos de glucosa.

Para determinar que tan buenos resultan los polinomios fraccionarios comparados con los polinomios convencionales, se calcula el $AIC$ de cada una de los modelos estimaciones.

| **Estimación** | **Potencias** | **AIC** |
|---|---|---|
| MV | (1, 2) | 224.9218 |
| MV | (1, 1) | 221.2249 |
| MV | (0.5, 2) | 222.2775 |

el mejor $AIC$ corresponde al ajuste por máxima verosimilitud con polinomios fraccionarios de grado dos y potencias $(1, 1)$.

En general las curvas ajustadas indican que los primeros tiempos después de aplicar la glucosa se observa una disminución en la cantidad promedio del fosfato en el plasma inorgánico y luego del tercer tiempo los niveles promedio de fosfato aumentan, lo que indica que la glucosa se debe aplicar en los tres primeros tiempos pues disminuye el fosfato y por tanto la enfermedad causada por el aumento de fosfato.

## 3.2. Datos de Ratones

Los datos de ratones fueron reportados por Williams y Izenman (1981) y analizados por Rao (1984, 1987) y later y Lee (1988, 1991). Son los pesos de 13 ratones machos medidos en intervalos de 3 días de nacidos hasta los 21 días, día del destete. Los tiempos se toman como $t = 1, 2, 3, 4, 5, 6, 7$

### 3.2.1. Análisis exploratorio de los datos de ratones

Se presentan algunos gráficos que permiten analizar las características de los datos.

**Gráfico de perfiles**

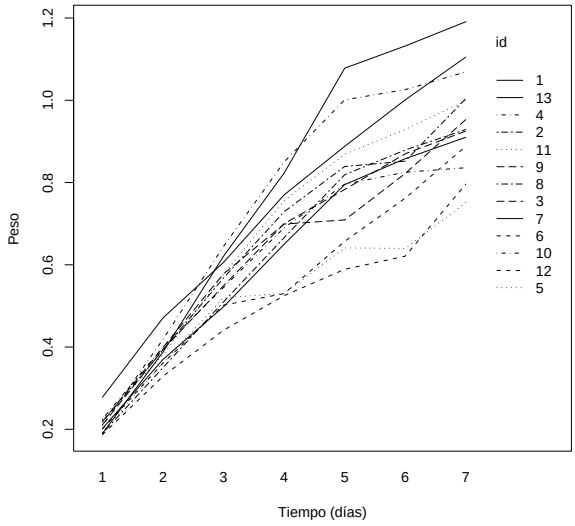


Figura 3.5: Gráfico de perfiles para el peso en libras de 13 ratones machos medidos en intervalos de 3 días de nacidos hasta los 21 días.

En la grafica (3.5), se observa el aumento de peso a medida que pasan los días hay una ganancia de peso promedio en los ratones. También se observa un incremento de la varianza con la edad para los 13 ratones.

**Diagrama de caja**

En el grafico (3.6), se observa un incremento de peso para los 13 ratones a medida que los días pasan, con una asimetría positiva en la mayoría de los tiempos.

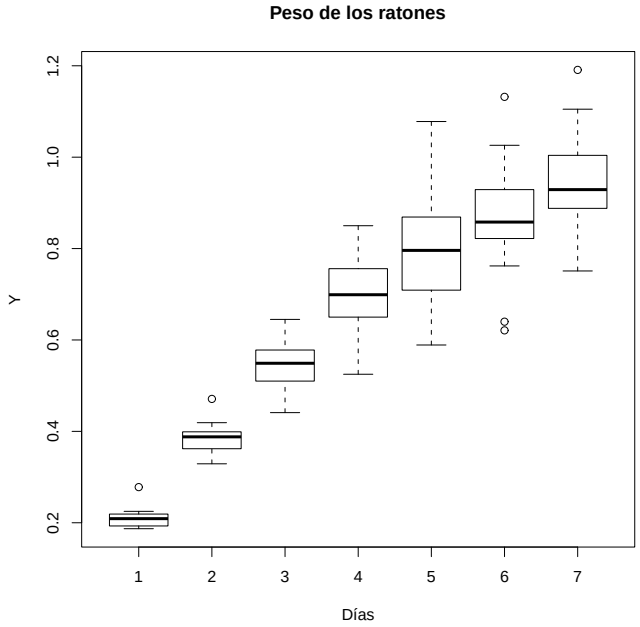


Figura 3.6: Diagrama de caja para datos de ratones en intervaloes 3 días hasta los 21 días.

**La matriz de varianzas y covarianzas**

$$\begin{pmatrix} 0.0005777 & 0.000735 & 0.0006369 & 0.001064 & 0.000752 & 0.00114 & 0.001102 \\ 0.0007351 & 0.001235 & 0.0015599 & 0.002602 & 0.002670 & 0.00311 & 0.002705 \\ 0.0006369 & 0.001559 & 0.0032038 & 0.005449 & 0.006491 & 0.00670 & 0.005726 \\ 0.0010641 & 0.002602 & 0.0054495 & 0.011345 & 0.013699 & 0.01400 & 0.011492 \\ 0.0007520 & 0.002670 & 0.0064915 & 0.013699 & 0.019218 & 0.01907 & 0.015285 \\ 0.0011444 & 0.003117 & 0.0067057 & 0.014008 & 0.019073 & 0.02036 & 0.016671 \\ 0.0011021 & 0.002705 & 0.0057264 & 0.011492 & 0.015285 & 0.01667 & 0.015367 \end{pmatrix} \tag{3.7}$$

**Dispersogramas**

En el dispersograma (3.7) se observa que para algunos tiempos la correlación en tiempos adyacentes es mayor que en tiempos posteriores. En la matriz de correlación 3.2 se calcula la correlación entre los diferentes tiempos.

$$\begin{pmatrix} 1.00000 & 0.87018 & 0.46814 & 0.41566 & 0.22569 & 0.3336 & 0.36989 \\ 0.87018 & 1.00000 & 0.78416 & 0.69522 & 0.54801 & 0.6214 & 0.62085 \\ 0.46814 & 0.78416 & 1.00000 & 0.90389 & 0.82728 & 0.8301 & 0.81609 \\ 0.41566 & 0.69522 & 0.90389 & 1.00000 & 0.92778 & 0.9215 & 0.87036 \\ 0.22569 & 0.54801 & 0.82728 & 0.92778 & 1.00000 & 0.9640 & 0.88943 \\ 0.33364 & 0.62147 & 0.83013 & 0.92157 & 0.96406 & 1.0000 & 0.94236 \\ 0.36989 & 0.62085 & 0.81609 & 0.87036 & 0.88943 & 0.9423 & 1.00000 \end{pmatrix} \tag{3.8}$$

En la matriz de correlación 3.8, se observa una correlación alta en la primera banda y luego empieza a disminuir, por lo que la estructura de covarianzas $AR1$, resulta conveniente. Más

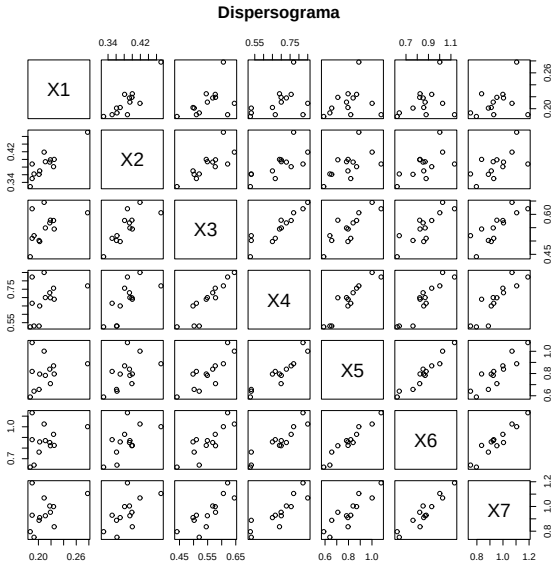


Figura 3.7: Dispersogramas para los diferentes pesos medidos a los 13 ratones

adelante con el criterio *AIC* se identifica la estructura de covarianzas más adecuada para las medidas repetidas.

Se verifican los supuestos del modelo de curva de crecimiento, para luego estimar los parámetros por máxima verosimilitud.

### 3.2.2. Verificación de los supuestos del modelo

- **Normalidad** La metodología planteada se desarrolla bajo el supuesto de normalidad multivariada, para verificarla se utiliza la estadística de Mardia que se encuentra detallada en el capitulo I.

  El estadístico de Mardia proporciona un p-valor de 0.22682 por tanto no rechaza el supuesto de normalidad multivariada.

- **Homocedasticidad** En la metodología planteada también es necesario la verificación del supuesto de homocedasticidad de varianzas, se hace a través del estadístico de Bartlett, que se mostro en el capítulo I. En esta aplicación no se calcula porque se tiene solamente una muestra.

Ahora se estima la estructura de covarianzas más apropiada para los datos.

### 3.2.3. Selección de la matriz de varianzas y covarianzas

Las estructuras de covarianzas que se utilizan en modelos de curvas de crecimiento son: no estructurada, uniforme, aleatoria y autoregresiva de orden uno. Para determinar cuál de estas estructuras de covarianzas se ajusta mejor a los datos, se comparan los modelos obtenidos con cada estructura por medio del criterio de información de Akaike (AIC).

| **Estructura de covarianzas** | **AIC** |
|---|---|
| AR1 | -238.6127 |
| UNIFORME | -209.8685 |
| ALEATORIA | -168.5712 |

**Estructura de covarianzas Autoregresiva de orden uno,(AR1)**

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 0.0094798 & 0.0058201 & 0.003573 & 0.002193 & 0.001346 & 0.0008269 & 0.0005076 \\ 0.0058201 & 0.0094798 & 0.005820 & 0.003573 & 0.002193 & 0.0013468 & 0.0008269 \\ 0.0035732 & 0.0058201 & 0.009479 & 0.005820 & 0.003573 & 0.0021938 & 0.0013468 \\ 0.0021938 & 0.0035732 & 0.005820 & 0.009479 & 0.005820 & 0.0035732 & 0.0021938 \\ 0.0013468 & 0.0021938 & 0.003573 & 0.005820 & 0.009479 & 0.0058201 & 0.0035732 \\ 0.0008269 & 0.0013468 & 0.002193 & 0.003573 & 0.005820 & 0.0094798 & 0.0058201 \\ 0.0005076 & 0.0008269 & 0.001346 & 0.002193 & 0.003573 & 0.0058201 & 0.0094798 \end{pmatrix}$$

La estructura de covarianzas $AR1$, (3.9) presenta el menor AIC, por tanto es la estructura más apropiado para los datos.

### 3.2.4. Estimación de los parámetros

Como se cumple el supuesto de normalidad multivariada para modelos de curvas de crecimiento. Se procede a estimar los parámetros del modelo.

#### Estimación por máxima verosimilitud con polinomios

Inicialmente se determina si los parámetros del modelo $\hat{\mathbf{y}} = \hat{\beta}_{\mathbf{1}}\mathbf{t} + \hat{\beta}_{\mathbf{2}}\mathbf{t^2}$ son significativos.

Se utiliza la estadística $\Lambda$ de Wilks. Las demás estadísticas citadas en el capítulo 2 se pueden obtener a través del software R. Los comandos se encuentra detallados en el apéndice C.

| Parámetro | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_1$ | 0.02338128 | 501.2311 | 1 | 12 | 3.723222e-11 |
| $\beta_2$ | 0.07437444 | 149.3458 | 1 | 12 | 3.946151e-08 |

El modelo estimado obtenido por máxima verosimilitud y con estructura de covarianzas de $AR1$, es:

$$\hat{\mathbf{y}} = 0.21710696 * \mathbf{t} - 0.01166174\mathbf{t}^2$$

Se calcula el error estándar para cada parámetro estimado.

| Parámetro | Estimación | Error estándar |
|---|---|---|
| $\beta_1$ | 0.21710696 | 0.03496443 |
| $\beta_2$ | -0.01166174 | 0.003440635 |

La matriz de covarianza de los $\hat{\mathbf{B}}$

$$\mathbf{cov}(\hat{\mathbf{B}}) = \begin{pmatrix} 4.379705e-05 & -4.235191e-06 \\ -4.235191e-06 & 7.731794e-07 \end{pmatrix}$$

**Estimación máxima verosimilitud con polinomios fraccionarios**

Se ajustan polinomios fraccionarios de grado uno y dos. Se utilizan los comandos de R que se encuentran en el apéndice C; para obtener los diferentes subconjuntos de $S$ de cardinal uno y dos, junto con las matrices de tiempos, $\mathbf{X}$. Se toman los modelos de curvas de crecimiento de polinomios fraccionarios con menor $AIC$.

Los modelos de curvas de crecimiento de polinomios fraccionarios con:

- $p = (-2, 0)$ y sin intercepto
- $p = (-0.5, -0.5)$ con intercepto

y estructura de covarianzas $AR1$, que poseen el mejor $AIC$ se describen a continuación.

Inicialmente para el modelo estimado con potencias $p = (-2, 0)$ se determina si los parámetros del modelo $\hat{\mathbf{y}} = \hat{\beta}_{\mathbf{1}}\mathbf{t}^{-\mathbf{2}} + \hat{\beta}_{\mathbf{2}}\ln(\mathbf{t})$ son significativos.

| | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_1$, $p = -2$ | 0.01358530 | 871.3075 | 1 | 12 | 1.426526e-12 |
| $\beta_2$, $p = 0$ | 0.01672501 | 705.4885 | 1 | 12 | 4.973355e-12 |

Los coeficientes para el tiempo $\beta_1$ y $\beta_2$, son significativos.

El modelo estimado por máximo verosimil con potencias $(-2, 0)$ es:

$$\hat{\mathbf{y}} = 0.2130882\mathbf{t}^{-\mathbf{2}} + 0.4853685 * \ln(t)$$

En la tabla siguiente se dan los parámetros estimados junto con sus errores estándar:

| Parámetro | Estimación | Error estándar |
|---|---|---|
| $\beta_1$, $p = -2$ | 0.2130882 | 0.02602827 |
| $\beta_2$, $p = 0$ | 0.4853685 | 0.06588678 |

La matriz de covarianzas para los parámetros estimados es:

$$\mathbf{cov}(\hat{\mathbf{B}}) = \begin{pmatrix} 4.304542e-05 & 7.407589e-05 \\ 7.407589e-05 & 1.745078e-04 \end{pmatrix} \tag{3.9}$$

La matriz 3.9 mide la exactitud de los coeficientes de regresión.

Otro modelo de curvas de crecimiento con polinomios fraccionarios y $p = (-0.5, -0.5)$, es:

$\hat{\mathbf{y}} = \hat{\beta}_{\mathbf{0}} + \hat{\beta}_{\mathbf{1}}\mathbf{t}^{\mathbf{-0.5}} + \hat{\beta}_{\mathbf{2}}\mathbf{t}^{\mathbf{-0.5}}$, se determina si los parámetros del modelo son significativos, a través de la estadística la $\Lambda$ de Wilks.

| Parámetro | Wilks | F | G.Num. | G.Den. | P-valor |
|---|---|---|---|---|---|
| $\beta_0$ | 0.03932397 | 293.1574 | 1 | 12 | 8.486046e-10 |
| $\beta_1, p = -0.5$ | 0.04494721 | 254.9798 | 1 | 12 | 1.896991e-09 |
| $\beta_2, p = -0.5$ | 0.06518398 | 172.0943 | 1 | 12 | 1.780958e-08 |

Los parámetros estimados son significativos.

El ajuste máximo verosimil del $PF_2$ con $p = (-0.5, -0.5)$ es:

$$\hat{\mathbf{y}} = 2.820301 - 2.608722\mathbf{t}^{-0.5} - 1.201086\mathbf{t}^{-0.5} * \ln(\mathbf{t})$$

$$\mathbf{cov}(\hat{\mathbf{B}}) = \begin{pmatrix} 0.02906493 & -0.02894509 & -0.016502479 \\ -0.02894509 & 0.02887452 & 0.016498701 \\ -0.01650248 & 0.01649870 & 0.009537544 \end{pmatrix} \tag{3.10}$$

La matriz 3.10 mide la exactitud de los coeficientes de regresión.

En la grafica (3.8) se muestran los mejores modelos estimados con polinomios, al utilizar $p = (1, 2)$ y con polinomios fraccionarios con elementos de $S$, $p = (-2, 0)$ y $p = (-0.5, -0.5)$.

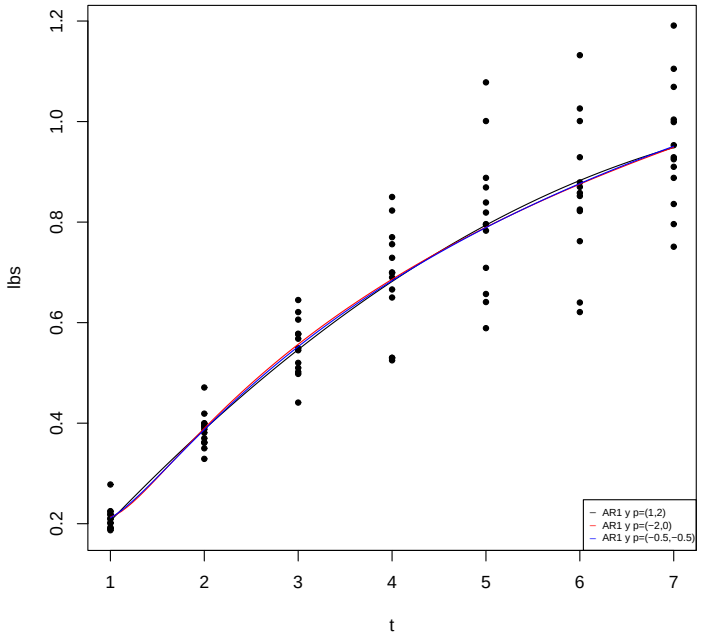


Figura 3.8: Ajustes por máxima verosimilitud con polinomios y fraccionarios, para los datos de ratones.

Para determinar que tan buenos resultan los polinomios fraccionarios comparados con los polinomios, se calcula el $AIC$ de cada una de los modelos estimaciones.

| **Estimación** | **Potencias** | **AIC** |
|---|---|---|
| MV | (1, 2) | -238.6127 |
| MV | (-2, 0) | -239.3172 |
| MV | (-0.5, -0.5) | -237.3434 |

el mejor $AIC$ corresponde al ajuste por máxima verosimilitud con polinomios fraccionarios de grado dos, sin intercepto y $p = (-2, 0)$.

En general las curvas de crecimiento ajustadas indican que hay una ganancia de peso promedio para los ratones, a medida que pasan los días los ratones aumentan su peso promedio.

# CAPÍTULO 4

# Simulación

## 4.1. Introduccción

En este capitulo se presenta un estudio de simulación con el fin de comparar el desempeño de los modelos de curvas de crecimiento ajustados a través de polinomios convencionales y polinomios fraccionarios.

Inicialmente, se generan simulaciones de datos de medidas repetidas, $\mathbf{Y}$ a partir de un diseño de la forma $\mathbf{Y} = \mathbf{XBZ} + \mathbf{E}$, donde:

- $\mathbf{X}$: Matriz que relaciona los tiempos y el grado del polinomio.
- $\mathbf{B}$: Matriz de parámetros.
- $\mathbf{Z}$: Matriz indicadora de grupos.
- $\mathbf{E}$: Matriz de errores aleatorios que se generan a partir de una distribución normal multivariada con vector de medias $\mathbf{0}_p$ y matriz de covarianzas $\mathbf{\Sigma_p}$. Además las filas de $\mathbf{E}$ son independiente

La forma en que se definen las matrices $\mathbf{X}$, $\mathbf{Z}$ y $\mathbf{B}$ determinan las características de un escenario.

Para cada uno de los seis escenarios se realizan 200 simulaciones, en cada una de ellas se lleva acabo la siguiente metodología:

- Se estiman varios modelos a través de la metodología de Potthoff y Roy, mediante el uso de diferentes matrices de diseño $\mathbf{X}$ y algunas estructuras de covarianza.
- Los modelos se estiman utilizando todos los tiempos y también truncando el último tiempo, con el objeto de visualizar la calidad del pronóstico del modelo.
- Con los modelos estimados se obtienen algunas medidas como el $AIC$ y $MAPE$ que determinan la bondad de ajuste y la calidad del pronóstico.

- Los resultados de las 200 simulaciones se evaluan mediante estadísticas descriptivas. Las cuales definen los mejores modelos.

### 4.1.1. Escenarios

A continuación se muestran las características de cada uno de los siete escenarios.

**Escenario I, II y III**

| Característica | Escenario I |
|---|---|
| Tamaño de muestra | 30 |
| Estructura de Covarianza | Uniforme |
| Tiempos | 7 |
| Descripción de Ruido | |
| Modelo General | $A$ |
| Característica | Escenario II |
| Tamaño de muestra | 30 |
| Estructura de Covarianza | Uniforme |
| Tiempos | 7 |
| Descripción de Ruido | Campana |
| Modelo General | $A$ |
| Característica | Escenario III |
| Tamaño de muestra | 30 |
| Estructura de Covarianza | Uniforme |
| Tiempos | 7 |
| Descripción de Ruido | Perfiles |
| Modelo General | $A$ |

**Escenario IV y V**

| Característica | Escenario IV |
|---|---|
| Tamaño de muestra | 30 |
| Estructura de Covarianza | AR1 |
| Tiempos | 7 |
| Descripción de Ruido | |
| Modelo General | $A$ |
| Característica | Escenario V |
| Tamaño de muestra | 30 |
| Estructura de Covarianza | AR1 |
| Tiempos | 7 |
| Descripción de Ruido | Perfiles |
| Modelo General | $A$ |

donde 4.1 define $A$, el modelo para los escenarios del uno al seis.

$$0.02201427 + 0.20843961t - 0.01077382t^2 \tag{4.1}$$

Dentro de los ruidos citados en los escenarios se encuentran:

- Sin ruido.
- Ruido sobre 2 perfiles en 3 tiempos.
- Ruido efecto campana, contamina cada uno de los perfiles en los tiempos de forma que se incremente la varianza a medida que pasa el tiempo.

Las estructuras de covarianza $AR1$ y Uniforme se describen detalladamente en el capítulo I.

### 4.1.2. Modelos estimados

Para los modelos estimados es necesario determinar:

- La matriz $\mathbf{X}$:

  Se Obtenen los posibles subconjuntos de uno y dos potencias diferentes de $S$ junto con potencias iguales, de tal manera que para subconjuntos de dos potencias se obtienen 36 modelos (28 modelos con potencias diferentes y 8 con potencias iguales), de una potencia 8 modelos. Necesitan ser ajustados 44 modelos para encontrar las mejores potencias de $S$ y así determinar la matriz $\mathbf{X}$.

- Formas de estimar los parámetros

  - Máxima verosimilitud: cuando se estiman los parámetros por máxima verosimilitud se tienen en cuenta las estructuras de covarianza:

  - Autoregresivo de orden uno.
  - Uniforme
  - No estructurada
  - Aletoria

- Truncamiento: para determinar la utilidad de los polinomios fraccionarios en curvas de alto grado se hace un truncamiento de los datos en el último tiempo, de tal forma que al realizar una extrapolación en el último tiempo los polinomios fraccionarios ajustan mejor, elimina efectos finales producidos por los polinomios convencionales de alto orden.

## 4.2. Resultados

Se obtienen los mejores modelos estimados por medio del $AIC$ promedio, cuando se modelo sobre todos los tiempos. De forma similar, para los modelos truncados estimados se escogen los mejores modelos a través del $AIC$ promedio y el $MAPE$ promedio.

A continuación, se presentan los resultados obtenidos en cada uno de los escenarios.

### 4.2.1. Escenario I:

Inicialmente se presenta los mejores modelos sin truncamiento, con una y dos potencias y sin ruido, con estructura de covarianzas uniforme para los siete tiempos.

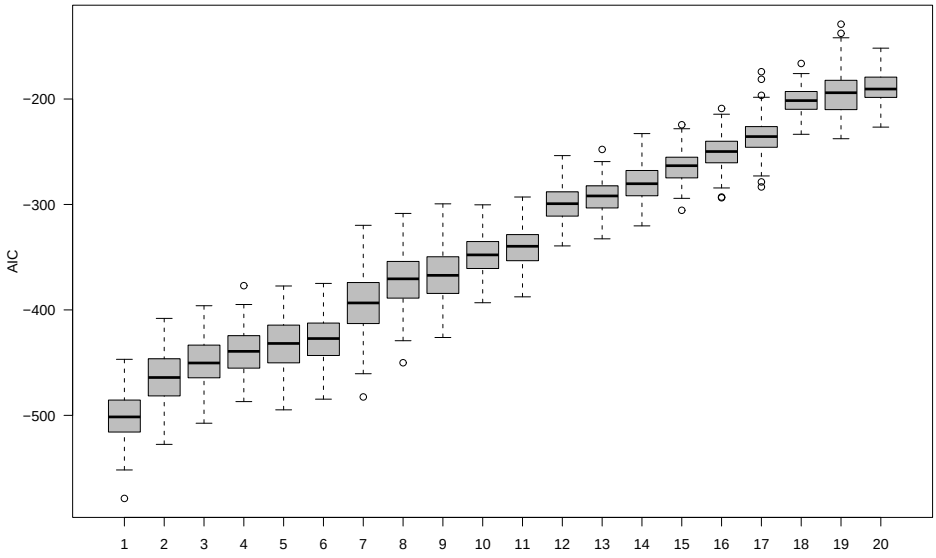


Figura 4.1: Diagrama de caja de los AIC promedio para los mejores modelos de grado uno.

En los diagramas 4.1 y 4.2 se dan los gráficos de caja para los mejores $AIC$ promedio, cuando en la estimación del modelo, se utiliza una y dos potencias respectivamente, en el eje $X$ se da la categoría del mejor modelo, en el eje $Y$ se muestra el $AIC$ promedio. Por lo tanto los primeros diagramas de caja muestran los mejores modelos estimados, expuestos detalladamente en la tabla 4.1

En la tabla 4.1 se dan los mejores modelos estimados de grado uno y dos, sin truncamiento. En los ajustes con polinomios de grado uno se observa que el mejor modelo estimado tiene potencia 0.5, estructura de covarianzas $Uniforme$ y su $AIC$ promedio es de $-502.08$. Para los ajustes de polinomios de grado dos, el mejor modelo estimado tiene potencias $(1, 2)$, estructura de covarianzas $Uniforme$ y su $AIC$ promedio de $-506.04$, el cual no difiere mucho del modelo estimado con potencias $(-1, 0)$ que tiene un $AIC$ promedio de $-505.37$.

A los modelos estimados se les calcula el promedio de los parámetros estimados y su desviación estándar, que indica que tan estable resultan las simulaciones. Estos resultados son presentados en la tabla 4.2

En la tabla 4.2 se presenta la estimación de los mejores modelos, en la primera parte de la tabla se muestran los modelos de grado uno ordenados por $AIC$ seguidos de los de grado dos. El mejor modelo estimado, con la potencia 0.5 muestra desviaciones estándar pequeñas para cada uno de sus parámetros. Del modelo estimado con potencias $p = (1, 2)$

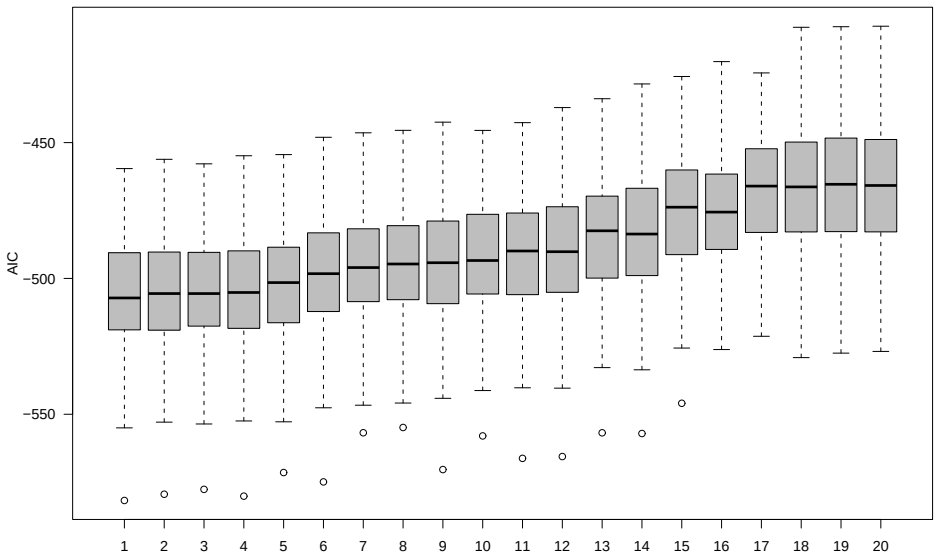


Figura 4.2: Diagrama de caja de los AIC promedio para los mejores modelos de grado dos.

también muestra desviaciones estándar pequeñas para cada uno de sus parámetros. lo que muestra una estabilidad en el proceso de simulación.

Ahora se exponen los mejores modelos de grado uno y dos, con truncamiento en el último tiempo y sin ruido.

En el diagrama 4.3 de la parte izquierda se muestran los $AIC$ promedio para los mejores modelos de grado uno y se hace truncamiento del último tiempo. Los diagramas de caja 4.3 en la parte derecha muestra los mejores modelos con truncamiento del tiempo siete. El $MAPE$ más pequeño hace pronósticos sobre el modelo extrapolando el tiempo siete. Para decidir que modelo truncado es el más apropiado se debe analizar conjuntamente el $AIC$ y el $MAPE$. Conviene analizar los primeros modelos que poseen el mejor $AIC$ y $MAPE$ conjuntamente.

En el diagrama 4.4 de la parte izquierda se muestran los $AIC$ promedio para los mejores modelos de grado dos con truncamiento del último tiempo. Los diagramas de caja 4.4 en la parte derecha muestran los mejores modelos al hacer truncamiento del tiempo siete. El $MAPE$ más pequeño hace pronósticos adecuados sobre el tiempo siete.

En la tabla 4.3 se dan cinco de los mejores modelos de grado uno y cinco de grado dos, el modelamiento se realizó con truncamiento del último tiempo. En los ajustes por polinomios de grado uno se observa que el mejor modelo tiene potencia 0.5, estructura de covarianzas $Uniforme$ y su $AIC$ promedio es de $-424.69$. Para los ajustes con polinomios de grado dos el mejor modelo estimado tiene potencias $(1, 2)$, estructura de covarianzas uniforme y su $AIC$ promedio es de $-425.47$, el cual no difiere significativamente del modelo estimado con potencias $(-0.5, 0)$ que tiene un $AIC$ promedio de $-425.28$. Al truncar el último tiempo el $MAPE$ que determina el mejor pronóstico para todo los datos incluyendo el

| structure | model | $mean.AIC$ | $sd.AIC$ |
|---|---|---|---|
| uniform | Int + t^0.5 | -502.08 | 22.52 |
| AR1 | Int + t^0.5 | -465.13 | 25.44 |
| uniform | Int + log(t) | -449.05 | 21.15 |
| uniform | Int + t^1 | -440.22 | 20.70 |
| AR1 | Int + log(t) | -432.08 | 22.93 |
| AR1 | Int + t^1 | -428.68 | 21.99 |
| random | Int + t^0.5 | -393.77 | 30.00 |
| uniform | Int + t^1 + t^2 | -506.04 | 21.95 |
| uniform | Int + t^-1 + log(t) | -505.37 | 22.17 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | -505.32 | 22.07 |
| uniform | Int + t^-0.5 + log(t) | -504.81 | 22.24 |
| uniform | Int + t^-1 + t^-0.5 | -503.16 | 21.93 |
| uniform | Int + log(t) + t^1 | -498.38 | 22.44 |
| uniform | Int + t^-1 + t^-1*log(t) | -495.84 | 21.75 |

Tabla 4.1: Tabla para los mejores modelos estimados de grado uno y dos

tiempo siete, (interpolación al tiempo siete). El modelo estimado con potencias $(-1, -0.5)$ presenta mejor $MAPE$ promedio, además su $AIC$ promedio esta cercano al mejor modelo estimado con potencias $(1, 2)$ para los seis tiempos iniciales. Por tanto el mejor modelo estimado para hacer pronósticos es el que posee potencias $(-1, -0.5)$

En la tabla 4.4 se muestran las estructuras de covarianza, los parámetros promedio ($mean_beta$) y la desviación estándar de cada uno de los parámetros ($sd_beta$), para los mejores modelos truncados de grado uno y dos. Para los modelos estimados truncados se debe escoger un modelo que tenga el mejor $AIC$ y el mejor $MAPE$. Resulta conveniete el modelo estimado con potencias $p = (-0.5, -0.5)$, que posee el menor $MAPE$ y no difiere mucho del mejor $AIC$.

En general en el escenario I, los polinomios fraccionarios no muestran mejores ajustes que los polinomios convencionales.

| structure | model | parameter | *mean.beta* | *sd.beta* |
|---|---|---|---|---|
| uniform | Int + t^0.5 | Int | -0.24 | 0.02 |
| uniform | Int + t^0.5 | t^0.5 | 0.46 | 0.01 |
| AR1 | Int + t^0.5 | Int | -0.23 | 0.02 |
| AR1 | Int + t^0.5 | t^0.5 | 0.45 | 0.01 |
| uniform | Int + log(t) | log(t) | 0.39 | 0.01 |
| uniform | Int + log(t) | Int | 0.17 | 0.02 |
| uniform | Int + t^1 + t^2 | t^1 | 0.21 | 0.01 |
| uniform | Int + t^1 + t^2 | t^2 | -0.01 | 0.00 |
| uniform | Int + t^1 + t^2 | Int | 0.02 | 0.02 |
| uniform | Int + t^-1 + log(t) | Int | -0.23 | 0.05 |
| uniform | Int + t^-1 + log(t) | t^-1 | 0.45 | 0.05 |
| uniform | Int + t^-1 + log(t) | log(t) | 0.58 | 0.02 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.63 | 0.09 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.22 | 0.07 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.85 | 0.09 |

Tabla 4.2: Tabla de los parámetros estimados para los mejores modelos.

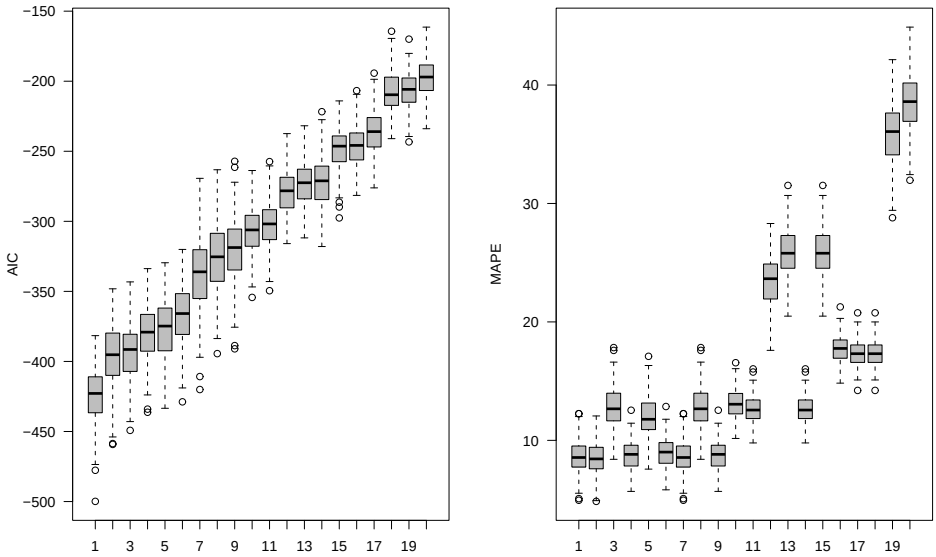


Figura 4.3: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno.

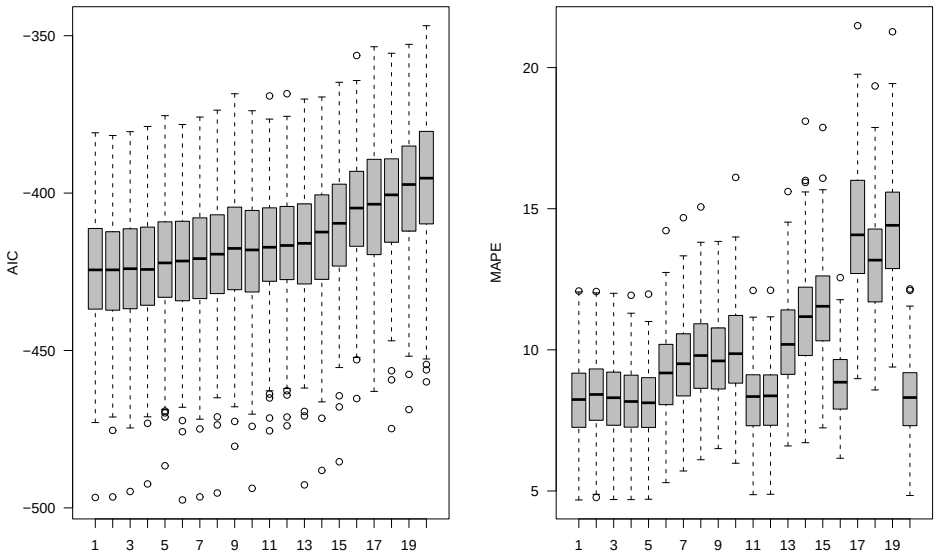


Figura 4.4: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados estimados de grado dos.

| structure | model | $mean.AIC$ | $sd.AIC1$ | $mean.MAPE$ | $sd.MAPE$ |
|---|---|---|---|---|---|
| uniform | Int + t^0.5 | -424.69 | 20.05 | 8.61 | 1.43 |
| AR1 | Int + t^0.5 | -396.30 | 22.76 | 8.49 | 1.40 |
| uniform | Int + t^1 | -393.75 | 18.66 | 12.78 | 1.67 |
| uniform | Int + log(t) | -379.73 | 19.51 | 8.79 | 1.22 |
| AR1 | Int + t^1 | -377.37 | 20.54 | 12.04 | 1.61 |
| AR1 | Int + log(t) | -366.67 | 20.84 | 8.99 | 1.24 |
| uniform | Int + t^1 + t^2 | -425.47 | 19.84 | 8.25 | 1.38 |
| uniform | Int + t^-0.5 + log(t) | -425.28 | 19.84 | 8.41 | 1.42 |
| uniform | Int + t^-1 + log(t) | -425.24 | 19.83 | 8.29 | 1.39 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | -424.81 | 19.82 | 8.19 | 1.35 |
| uniform | Int + t^-1 + t^-0.5 | -422.98 | 19.82 | 8.14 | 1.31 |
| uniform | Int + t^2 + t^2*log(t) | -419.03 | 19.78 | 9.76 | 1.58 |
| uniform | Int + t^-1 + t^-1*log(t) | -418.23 | 19.81 | 8.26 | 1.28 |

Tabla 4.3: Tabla para los mejores modelos truncado el último tiempo.

| structure | model | parameter | *mean.beta* | *sd.beta* |
|---|---|---|---|---|
| uniform | Int + t^0.5 | Int | -0.25 | 0.02 |
| uniform | Int + t^0.5 | t^0.5 | 0.47 | 0.01 |
| AR1 | Int + t^0.5 | Int | -0.24 | 0.02 |
| AR1 | Int + t^0.5 | t^0.5 | 0.46 | 0.01 |
| uniform | Int + t^1 | Int | 0.13 | 0.02 |
| uniform | Int + t^1 | t^1 | 0.13 | 0.00 |
| uniform | Int + t^1 + t^2 | t^2 | -0.01 | 0.00 |
| uniform | Int + t^1 + t^2 | t^1 | 0.21 | 0.01 |
| uniform | Int + t^1 + t^2 | Int | 0.02 | 0.02 |
| uniform | Int + t^-0.5 + log(t) | t^-0.5 | 1.27 | 0.16 |
| uniform | Int + t^-0.5 + log(t) | log(t) | 0.79 | 0.05 |
| uniform | Int + t^-0.5 + log(t) | Int | -1.05 | 0.16 |
| uniform | Int + t^-1 + log(t) | Int | -0.25 | 0.06 |
| uniform | Int + t^-1 + log(t) | t^-1 | 0.47 | 0.06 |
| uniform | Int + t^-1 + log(t) | log(t) | 0.59 | 0.03 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.64 | 0.11 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.86 | 0.11 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.23 | 0.08 |

Tabla 4.4: Tabla de los parámetros estimados para los mejores modelos truncado el último tiempo.

### 4.2.2. Escenario II:

Se empieza con los mejores modelos con efecto campana y sin truncamiento en el último tiempo.

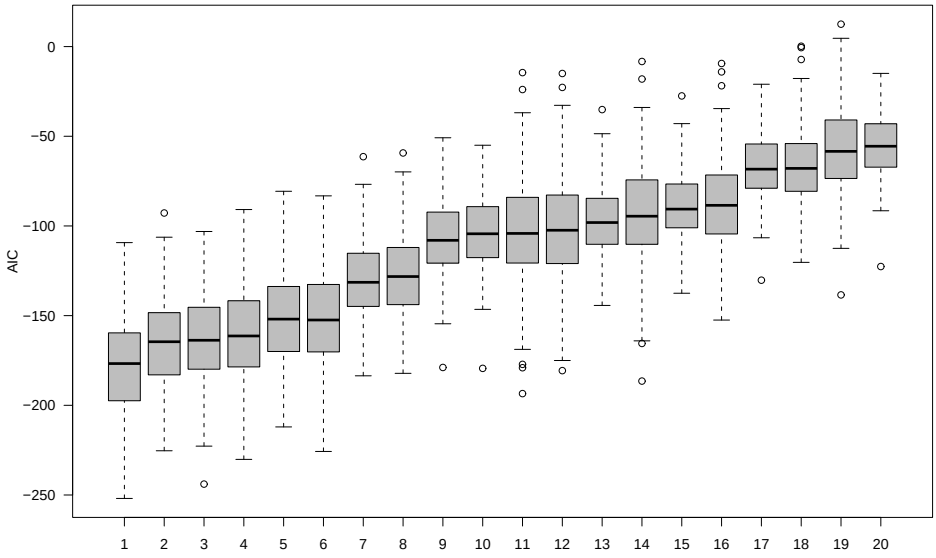


Figura 4.5: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y efecto campana.

En las gráficas 4.5 y4.6 se exponen los diagramas de caja para los mejores $AIC$ promedio para modelos estimados de grado uno y dos. En la tabla 4.5 se muestran los mejores modelos y se obtiene el $AIC$ promedio junto con su desviación estandar.

En la tabla 4.5, se evidencia que los mejores modelos estimados son los que tienen potencias $(0.5)$ y $(1, 2)$ con estructura de covarianzas uniforme, pues su $AIC$ promedio es el más pequeño. El modelo de polinomios fraccionariosde grado uno, con potencia $p = (0.5)$ resulta ser más parsimonioso.

Para determinar la estábilidad de los parámetros en cada simulación, en la tabla 4.6 se calcula el promedio de los parámetros estimados junto con su desviación estándar.

En la tabla 4.6 presenta los parámetros estimados para los mejores modelos de grado uno y dos, con los siete tiempos, donde las desviaciones estándar para los parámetros promedio estimados son pequeñas y no difieren ampliamente.

Ahora se presentan los reslutados para datos truncados en el último tiempo, con efecto campana.

En las gráficas 4.7 y 4.8 se exponen los mejores modelos truncados de grado uno y dos respectivamente, con su $AIC$ promedio y su $MAPE$ promedio.

En la tabla 4.7, se evidencia que el mejor modelo estimado es el que tiene $p = (0.5)$

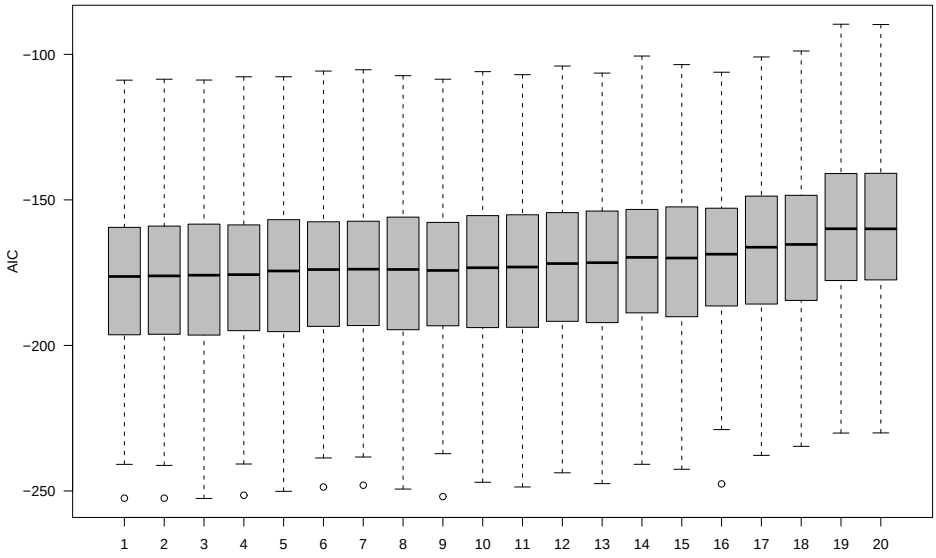


Figura 4.6: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos estimados de grado dos y efecto campana.

pues su $AIC$ es el más pequeño, seguido del polinomio convencional $p = (1, 2)$. El mejores modelos truncados con $p = (1, 2)$, $p = (-0.5, -0.5)$ y $p = (2, 2)$

En la tabla 4.8 presenta los parámetros estimados para los mejores modelos truncado el último tiempo junto con sus desviaciones estándar. Indica la estabilidad de las estimaciones de los parámetros de cada una de las simulaciones, pues su desviación estándar en general resulta ser pequeña.

En general en el escenario II, cuando se contamina los perfiles a través del efecto campana, el ajuste con polinomios fraccionarios se comporta un poco mejor que al utilizar polinomios convencionales.

| structure | model | $mean.AIC1$ | $sd.AIC1$ |
|---|---|---|---|
| uniform | Int + t^0.5 | -177.75 | 26.38 |
| uniform | Int + log(t) | -165.32 | 25.09 |
| uniform | Int + t^1 | -162.86 | 25.01 |
| AR1 | Int + t^0.5 | -160.84 | 26.33 |
| AR1 | Int + log(t) | -152.07 | 25.35 |
| uniform | Int + t^1 + t^2 | -177.10 | 26.45 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | -176.96 | 26.47 |
| uniform | Int + t^-0.5 + log(t) | -176.81 | 26.48 |
| uniform | Int + t^-1 + t^-0.5 | -176.55 | 26.41 |
| uniform | Int + log(t) + t^1 | -175.40 | 26.37 |

Tabla 4.5: Tabla para los mejores modelos sin truncamiento y efecto campana.

| structure | model | parameter | $mean.beta$ | $sd.beta$ |
|---|---|---|---|---|
| uniform | Int + t^0.5 | Int | -0.24 | 0.02 |
| uniform | Int + t^0.5 | t^0.5 | 0.46 | 0.01 |
| AR1 | Int + t^0.5 | Int | -0.23 | 0.02 |
| AR1 | Int + t^0.5 | t^0.5 | 0.45 | 0.01 |
| uniform | Int + log(t) | log(t) | 0.39 | 0.01 |
| uniform | Int + log(t) | Int | 0.17 | 0.02 |
| uniform | Int + t^1 + t^2 | t^1 | 0.21 | 0.01 |
| uniform | Int + t^1 + t^2 | t^2 | -0.01 | 0.00 |
| uniform | Int + t^1 + t^2 | Int | 0.02 | 0.02 |
| uniform | Int + t^-1 + log(t) | Int | -0.23 | 0.05 |
| uniform | Int + t^-1 + log(t) | t^-1 | 0.45 | 0.05 |
| uniform | Int + t^-1 + log(t) | log(t) | 0.58 | 0.02 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.63 | 0.09 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.22 | 0.07 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.85 | 0.09 |

Tabla 4.6: Tabla de los parámetros estimados para los mejores modelos.

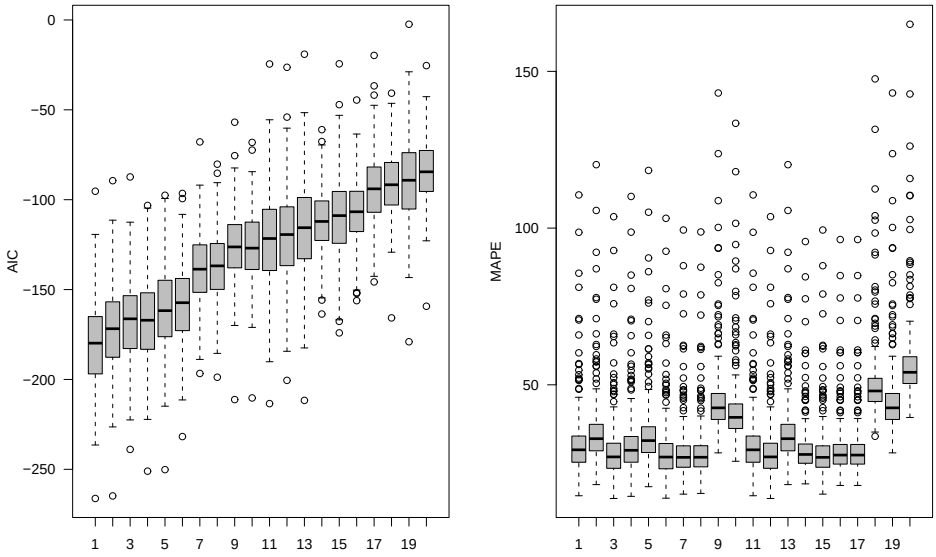


Figura 4.7: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos estimados truncados de grado uno y dos y efecto campana.

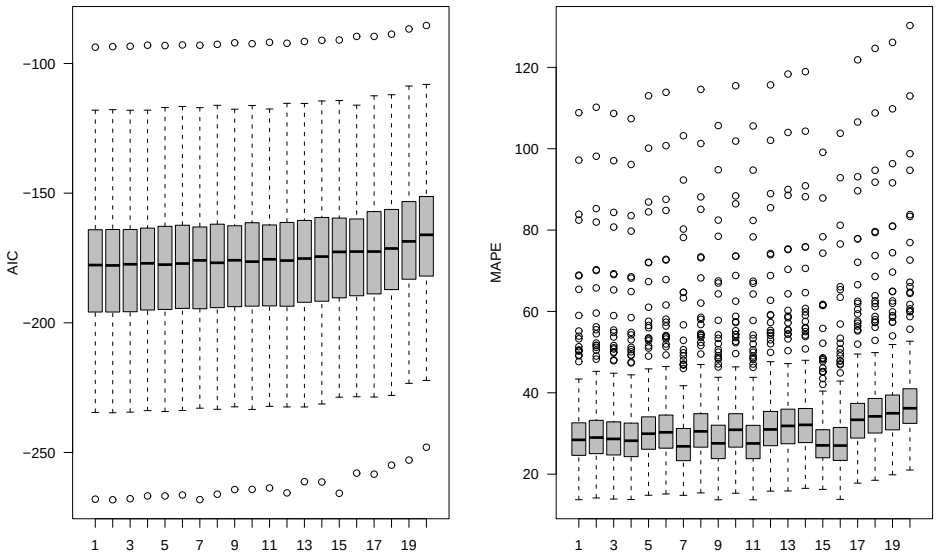


Figura 4.8: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y efecto campana .

| structure | model | *mean.AIC*1 | *sd.AIC*1 | *mean.MAPE* | *sd.MAPE* |
|---|---|---|---|---|---|
| uniform | Int + t^0.5 | -180.06 | 24.83 | 31.65 | 11.99 |
| uniform | Int + t^1 | -171.72 | 24.03 | 35.43 | 12.67 |
| uniform | Int + log(t) | -167.30 | 23.49 | 29.47 | 11.34 |
| AR1 | Int + t^0.5 | -166.85 | 23.93 | 31.47 | 11.96 |
| AR1 | Int + t^1 | -161.03 | 23.36 | 34.82 | 12.56 |
| uniform | Int + t^1 + t^2 | -178.85 | 24.93 | 30.71 | 11.89 |
| uniform | Int + t^-0.5 + log(t) | -178.78 | 24.93 | 31.31 | 11.97 |
| uniform | Int + t^-0.5 + t^-0.5*log(t) | -178.68 | 24.92 | 30.83 | 11.85 |
| uniform | Int + t^-1 + t^-0.5 | -178.24 | 24.87 | 30.46 | 11.74 |
| uniform | Int + log(t) + t^1 | -178.03 | 24.85 | 32.30 | 12.20 |
| uniform | Int + t^2 + t^2*log(t) | −177.32 | 24.77 | 29.27 | 11.27 |
| uniform | Int + t^-1 + t^-1*log(t) | -177.04 | 24.75 | 30.00 | 11.58 |

Tabla 4.7: Tabla para los mejores modelos truncado el último tiempo

| structure | model | parameter | *mean.beta* | *sd.beta* |
|---|---|---|---|---|
| uniform | Int + t^0.5 | Int | -0.25 | 0.02 |
| uniform | Int + t^0.5 | t^0.5 | 0.47 | 0.01 |
| AR1 | Int + t^0.5 | Int | -0.24 | 0.02 |
| AR1 | Int + t^0.5 | t^0.5 | 0.46 | 0.01 |
| uniform | Int + t^1 | Int | 0.13 | 0.02 |
| uniform | Int + t^1 | t^1 | 0.13 | 0.00 |
| uniform | Int + t^1 + t^2 | t^2 | -0.01 | 0.00 |
| uniform | Int + t^1 + t^2 | t^1 | 0.21 | 0.01 |
| uniform | Int + t^1 + t^2 | Int | 0.02 | 0.02 |
| uniform | Int + t^-0.5 + log(t) | t^-0.5 | 1.27 | 0.16 |
| uniform | Int + t^-0.5 + log(t) | log(t) | 0.79 | 0.05 |
| uniform | Int + t^-0.5 + log(t) | Int | -1.05 | 0.16 |
| uniform | Int + t^-1 + log(t) | Int | -0.25 | 0.06 |
| uniform | Int + t^-1 + log(t) | t^-1 | 0.47 | 0.06 |
| uniform | Int + t^-1 + log(t) | log(t) | 0.59 | 0.03 |

Tabla 4.8: Tabla de los parámetros estimados para los mejores modelos truncados.

### 4.2.3. Escenario III

El este escenario se generan datos con ruido en algunos perfiles. Inicialmente se presentan los mejores modelos sin truncamiento.

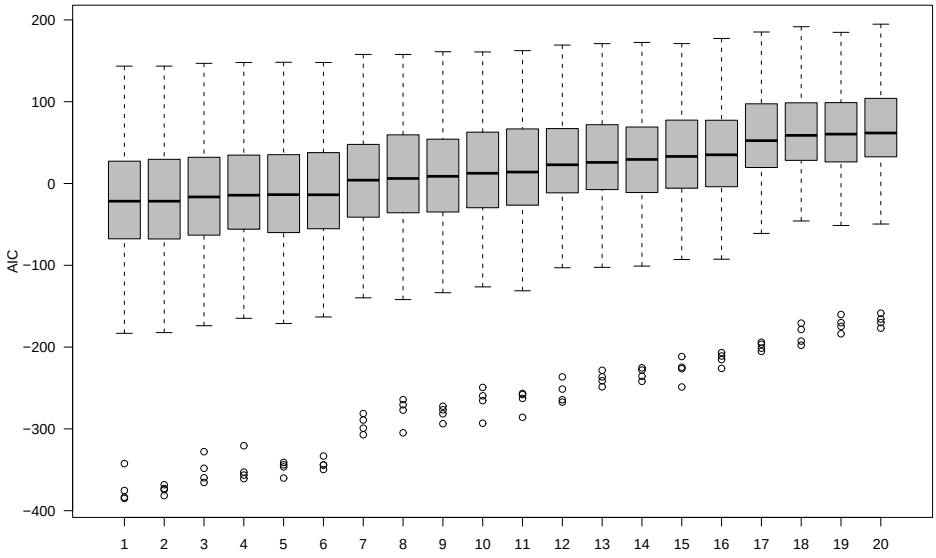


Figura 4.9: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y ruido en algunos perfiles

En las gráficas 4.9 y 4.10 se exponen los mejores modelos de grado uno y dos respectivamente, con su $AIC$ promedio.

La tabla 4.9, revela que los mejores modelos estimados tiene estructura de covarianzas $ALEATORIA$, y $p = (-2, -1)$, $p = (-2, -0.5)$ y $p = (3, 3)$.

En la tabla 4.10 se presentan los mejores modelos sin truncamiento, de grado uno y dos. Se observa la estabilidad del proceso de simulación, ya que los parámetros promedio presentan una desviación estándar pequeña.

Ahora se presentan los resultados cuando es truncado el último tiempo.

En las gráficas 4.11 y4.12 se exponen los mejores modelos truncados estimados de grado uno y dos, con su $AIC$ promedio y su $MAPE$ promedio.

La tabla 4.11, revela que los mejores modelos truncados son estimados con estructura de covarianzas $ALEATORIA$, y $p = (-1, 0.5)$, $p = (-2, 2)$ y $p = (-2, 0.5)$. El mejor de estos tres modelos con fines de pronóstico posee el menor $AIC$ promedio y a la vez el menor $MAPE$ promedio, es decir los modelos con potencias $(-2, 0.5)$ y $(-2, 0)$.

En las tablas 4.10 y 4.12 se muestran la estabilidad de la simulaciones ya que la desviación estándar es pequeña para cada uno de los parámetros en cada modelo estimado.

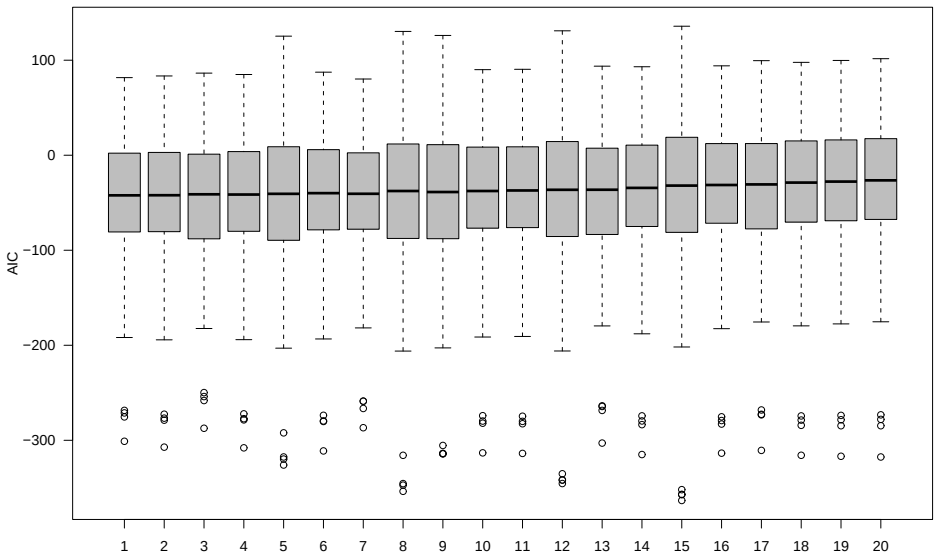


Figura 4.10: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores estimados de grado dos y ruido en algunos perfiles

En el escenario III, para datos no truncados tienen mejores ajustes que los polinomios convencionales y en algunos casos resultan más parsimoniosos. Para datos truncados en el último tiempo y ruido en algunos perfiles, los polinomios fraccionarios hacen pronósticos acertados de la muestra estudiada, a través de la interpolación de la curva en el último tiempo.

| structure | model | $mean.AIC1$ | $sd.AIC1$ |
|---|---|---|---|
| AR1 | Int + t^0.5 | -25.80 | 83.49 |
| uniform | Int + t^0.5 | -24.36 | 83.40 |
| AR1 | Int + log(t) | -19.43 | 81.01 |
| AR1 | Int + t^1 | -16.68 | 80.33 |
| uniform | Int + log(t) | -16.22 | 80.55 |
| random | Int + t^-2 + t^-1 | -44.64 | 64.17 |
| random | Int + t^-2 + t^-0.5 | -44.43 | 65.01 |
| random | Int + t^3 + t^3*log(t) | -44.08 | 64.58 |
| random | Int + t^-1 + t^-1*log(t) | -43.70 | 65.26 |
| AR1 | Int + t^3 + t^3*log(t) | -43.55 | 78.33 |

Tabla 4.9: Tabla para los mejores modelos con su $AIC$ promedio.

| structure | model | parameter | *mean.beta* | *sd.beta* |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.31 | 0.03 |
| AR1 | Int + t^0.5 | t^0.5 | 0.52 | 0.02 |
| uniform | Int + t^0.5 | Int | -0.33 | 0.03 |
| uniform | Int + t^0.5 | t^0.5 | 0.53 | 0.02 |
| AR1 | Int + log(t) | Int | 0.16 | 0.02 |
| AR1 | Int + log(t) | log(t) | 0.44 | 0.02 |
| AR1 | Int + t^1 | Int | 0.13 | 0.02 |
| AR1 | Int + t^1 | t^1 | 0.14 | 0.00 |
| random | Int + t^-2 + t^-1 | t^-2 | 1.94 | 0.13 |
| random | Int + t^-2 + t^-1 | t^-1 | -3.16 | 0.17 |
| random | Int + t^-2 + t^-1 | Int | 1.44 | 0.05 |
| random | Int + t^-2 + t^-0.5 | t^-0.5 | -2.68 | 0.15 |
| random | Int + t^-2 + t^-0.5 | Int | 2.04 | 0.08 |
| random | Int + t^-2 + t^-0.5 | t^-2 | 0.85 | 0.07 |
| random | Int + t^3 + t^3*log(t) | Int | 0.23 | 0.02 |
| random | Int + t^3 + t^3*log(t) | t^3 | 0.02 | 0.00 |
| random | Int + t^3 + t^3*log(t) | t^3*log(t) | -0.01 | 0.00 |

Tabla 4.10: Tabla de los parámetros estimados para los mejores modelos.

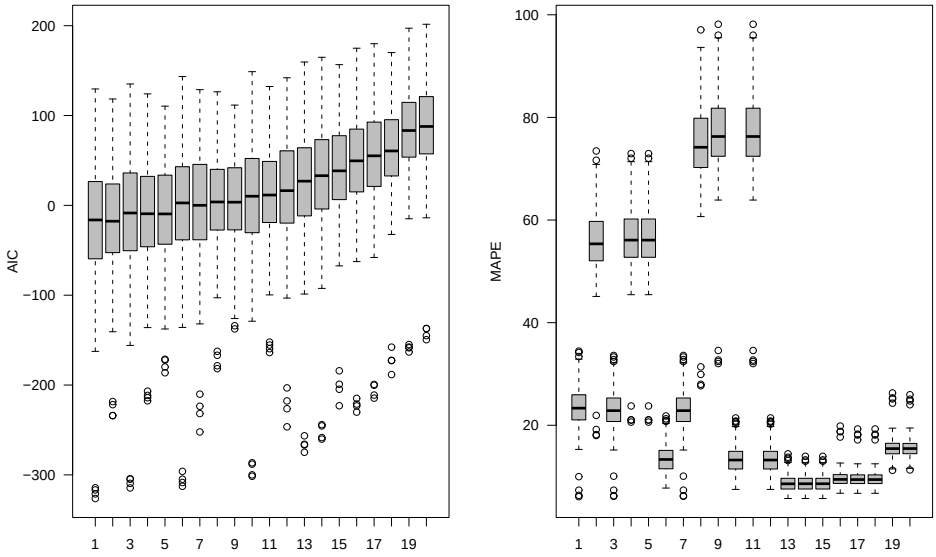


Figura 4.11: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y ruido en algunos perfiles

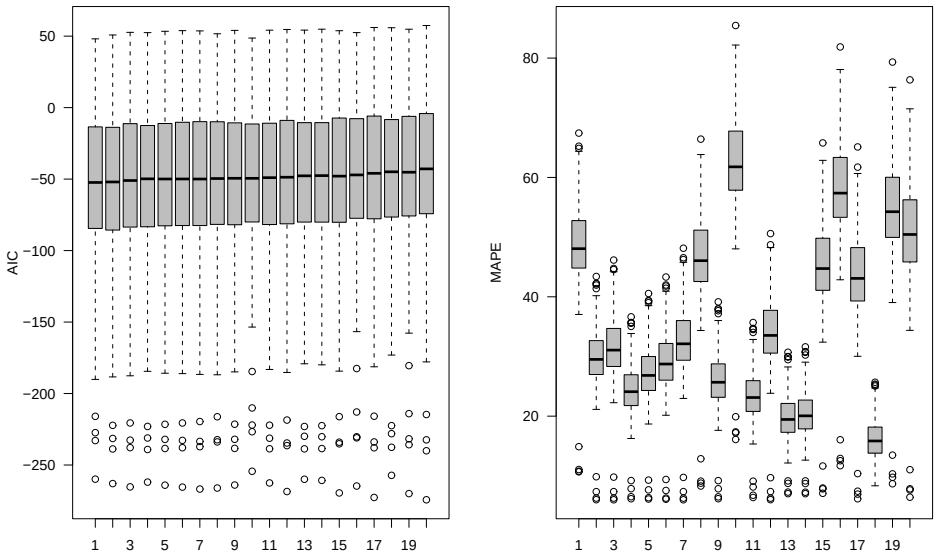


Figura 4.12: Diagrama de caja de los AIC promedio y los MAPE promedio para los mejores modelos truncados de grado uno y dos y ruido en algunos perfiles

| structure | model | $mean.AIC1$ | $sd.AIC1$ | $mean.MAPE$ | $sd.MAPE$ |
|---|---|---|---|---|---|
| AR1 | Int + t^0.5 | -20.32 | 73.37 | 23.40 | 4.34 |
| AR1 | Int + t^2 | -17.86 | 61.96 | 55.45 | 7.77 |
| uniform | Int + t^0.5 | -11.57 | 72.82 | 23.02 | 4.20 |
| uniform | Int + t^2 | -9.16 | 61.44 | 56.10 | 7.48 |
| random | Int + t^2 | -7.83 | 55.82 | 56.10 | 7.48 |
| random | Int + t^-2 + t^2 | -53.78 | 53.63 | 48.34 | 8.06 |
| random | Int + t^-2 + t^0.5 | -53.35 | 54.07 | 29.61 | 5.60 |
| random | Int + t^-1 + t^0.5 | -52.06 | 54.50 | 31.27 | 5.98 |
| random | Int + t^-2 + log(t) | -51.51 | 53.89 | 24.31 | 4.82 |
| random | Int + t^-1 + log(t) | -51.38 | 54.31 | 27.03 | 5.32 |

Tabla 4.11: Tabla para los mejores modelos y su $AIC$ promedio.

| structure | model | parameter | $mean.beta$ | $sd.beta$ |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.43 | 0.05 |
| AR1 | Int + t^0.5 | t^0.5 | 0.60 | 0.03 |
| AR1 | Int + t^2 | Int | 0.28 | 0.02 |
| AR1 | Int + t^2 | t^2 | 0.02 | 0.00 |
| uniform | Int + t^0.5 | Int | -0.44 | 0.05 |
| uniform | Int + t^0.5 | t^0.5 | 0.60 | 0.03 |
| random | Int + t^-2 + t^2 | Int | 0.38 | 0.03 |
| random | Int + t^-2 + t^2 | t^2 | 0.02 | 0.00 |
| random | Int + t^-2 + t^2 | t^-2 | -0.18 | 0.03 |
| random | Int + t^-2 + t^0.5 | Int | -0.74 | 0.12 |
| random | Int + t^-2 + t^0.5 | t^0.5 | 0.74 | 0.07 |
| random | Int + t^-2 + t^0.5 | t^-2 | 0.23 | 0.06 |
| random | Int + t^-1 + t^0.5 | t^0.5 | 0.83 | 0.09 |
| random | Int + t^-1 + t^0.5 | t^-1 | 0.41 | 0.10 |
| random | Int + t^-1 + t^0.5 | Int | -1.01 | 0.19 |

Tabla 4.12: Tabla de los parámetros estimados para los mejores modelos truncados.

### 4.2.4. Escenario IV:

Inicialmente se muestran los mejores modelos sin truncamiento.

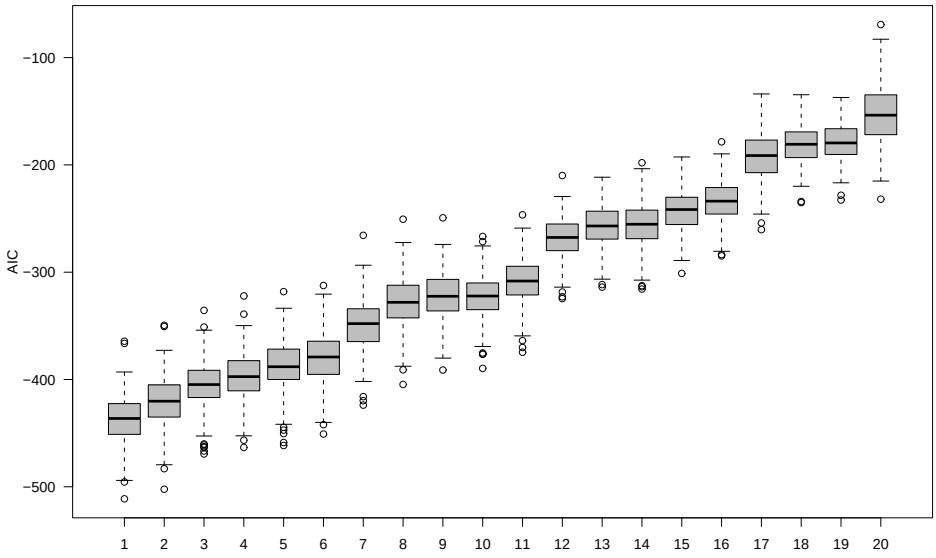


Figura 4.13: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y sin ruido

En las diagramas de caja 4.13 y 4.14 se muestran los mejores $AIC$ promedio obtenidos de los modelos estimados de grado uno y dos respectivamente. Al estimar los modelos de grado uno los $AIC$ promedio crecen, mostrando diferencias entre los diferentes modelos ajustados. Cuando se estiman los modelos de grado dos los $AIC$ muestran un comportamiento homogeneo para los primeros modelos, esto indica la gama de modelos de polinomios fraccionarios que ajustan bien.

En la tabla 4.13 se presentan los mejores modelos estimados con su estructura de covarianzas, su $mean.AIC$ promedio y la desviación estándar del $sd.AIC$. En esta tabla se evidencia que no existen muchas diferencias al ajustar polinomios convencionales y polinomios fraccinarios. Pues al estimar los modelos con $p = (1, 2)$ o $p = (-0.5, -0.5)$ sus $AIC$ promedio son casí iguales.

En la tabla 4.14 se muestran los mejores modelos estimados, la media y la desviación estándar de los parámetros estimados. Esta tabla permite verificar la estabilidad de las simulaciones realizadas, pues la desviación estándar para cada uno de los parámetros promedio en cada modelo estimado es pequeña.

Ahora se presentan los resultados para el escenario IV cuando los datos son truncados en el último tiempo.

En las gráficas 4.15 y 4.16 se muestran los diagramas de caja para los mejores modelos truncados de grado uno y dos respectivamente.

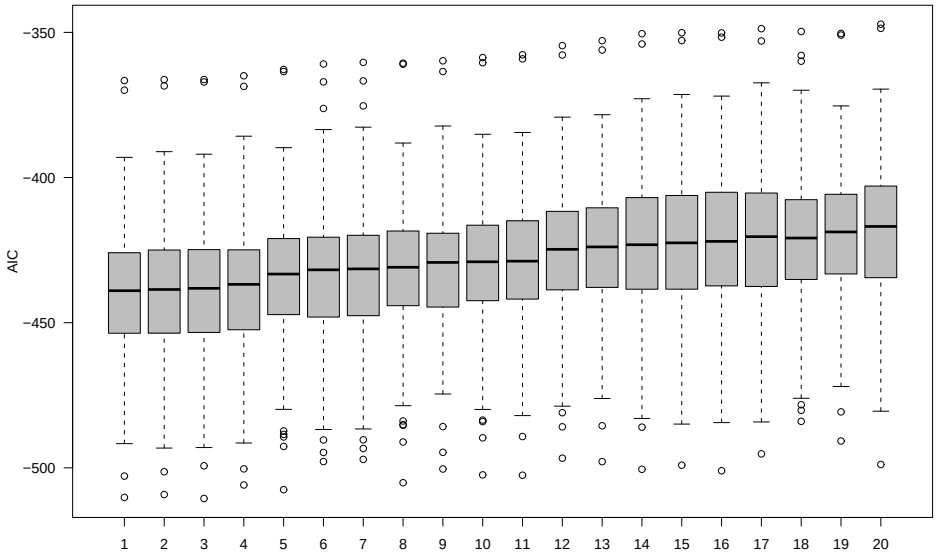


Figura 4.14: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado dos y sin ruido

Para la gráfica 4.15 de los $AIC$ para los modelos estimados de grado uno, se observa el crecimiento del $AIC$, luego los primeros modelos truncados son los mejores. Para hacer pronósticos a través de la extrapolación al último tiempo es necesario revisar el comportamiento del $MAPE$ y del $AIC$. Los modelos con el mejor $MAPE$ y $AIC$ conjuntamente son los que tienen $p = (0.5)$, $p = (0)$,$p = (0.5)$ y $p = (0)$ y estructura de covarianzas $AR1$ y $uniforme$, $AR1$, $uniforme$ y $aleatoria$ respectivamente.

En la gráfica 4.16 se observan los diagramas de caja del $AIC$ promedio para los mejores modelos, mostrando un comportamiento casí constante de los $AIC$, además el gráfico del $MAPE$ permite seleccionan los modelos con: $p = (1, 2)$, $p = (-0.5, 0)$,$p = (-0.5, -0.5)$,$p = (-1, -0.5)$, $p = (-1, -1)$ y $p = (-2, -0.5)$ pues el valor del $MAPE$ en estos modelos es el mejor.

En la tabla 4.15 se dan los mejores modelos truncados estimados. El mejor modelo se estima con la potencia $p = (0.5)$ y estructura de covarianzas $AR1$, pero su $AIC$ y su $MAPE$ no difieren mucho del modelo estimado con potencias $(1, 2)$.

Para los mejores modelos truncados, la tabla 4.16 muestra el promedio de los párametros estimados y su desviación estándar. La desviación estándar muestra la estabilidad de las simulaciones.

En general en el escenario IV los modelos ajustados por polinomios convencionales son muy similares a los obtenidos con polinomios fraccionarios.

| structure | model | $mean.AIC1$ | $sd.AIC1$ |
|---|---|---|---|
| AR1 | Int + t^0.5 | -437.70 | 22.74 |
| uniform | Int + t^0.5 | -420.75 | 24.58 |
| AR1 | Int + log(t) | -404.19 | 22.82 |
| AR1 | Int + t^1 | -397.75 | 22.06 |
| uniform | Int + log(t) | -387.93 | 24.32 |
| AR1 | Int + t^1 + t^2 | -440.58 | 22.89 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | -439.84 | 23.11 |
| AR1 | Int + t^-0.5 + log(t) | -439.36 | 23.04 |
| AR1 | Int + t^-1 + t^-0.5 | -438.60 | 23.16 |
| AR1 | Int + log(t) + t^1 | -434.88 | 22.91 |

Tabla 4.13: Tabla para los mejores modelos con su $AIC$ promedio.

| structure | model | parameter | $mean.beta$ | $sd.beta$ |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.24 | 0.03 |
| AR1 | Int + t^0.5 | t^0.5 | 0.46 | 0.02 |
| uniform | Int + t^0.5 | Int | -0.24 | 0.03 |
| uniform | Int + t^0.5 | t^0.5 | 0.46 | 0.02 |
| AR1 | Int + log(t) | Int | 0.18 | 0.02 |
| AR1 | Int + log(t) | log(t) | 0.38 | 0.01 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.21 | 0.11 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.84 | 0.14 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.62 | 0.15 |
| AR1 | Int + t^-0.5 + log(t) | Int | -0.96 | 0.19 |
| AR1 | Int + t^-0.5 + log(t) | log(t) | 0.76 | 0.07 |
| AR1 | Int + t^-0.5 + log(t) | t^-0.5 | 1.18 | 0.19 |
| AR1 | Int + t^-1 + t^-0.5 | Int | 2.09 | 0.08 |
| AR1 | Int + t^-1 + t^-0.5 | t^-1 | 1.83 | 0.16 |
| AR1 | Int + t^-1 + t^-0.5 | t^-0.5 | -3.70 | 0.24 |

Tabla 4.14: Tabla de los parámetros estimados para los mejores modelos estimados.

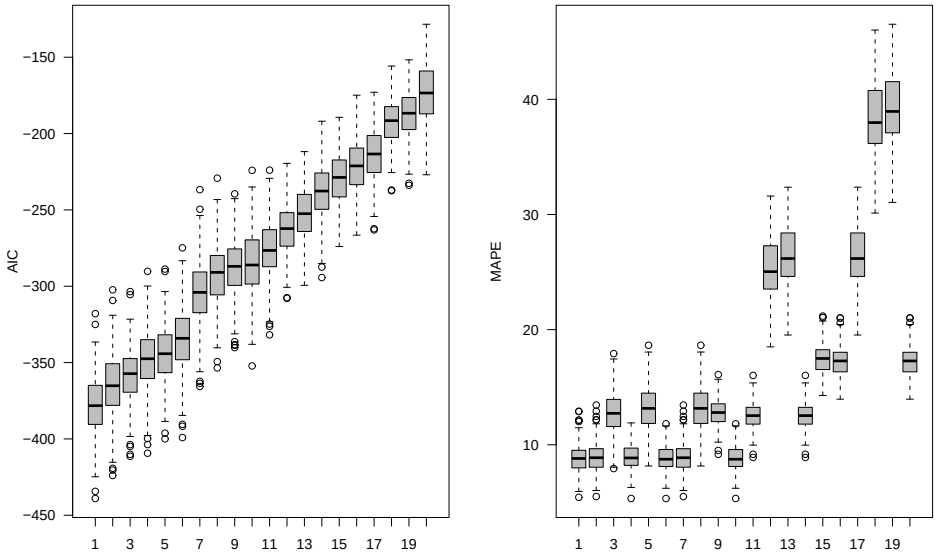


Figura 4.15: Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado uno y sin ruido

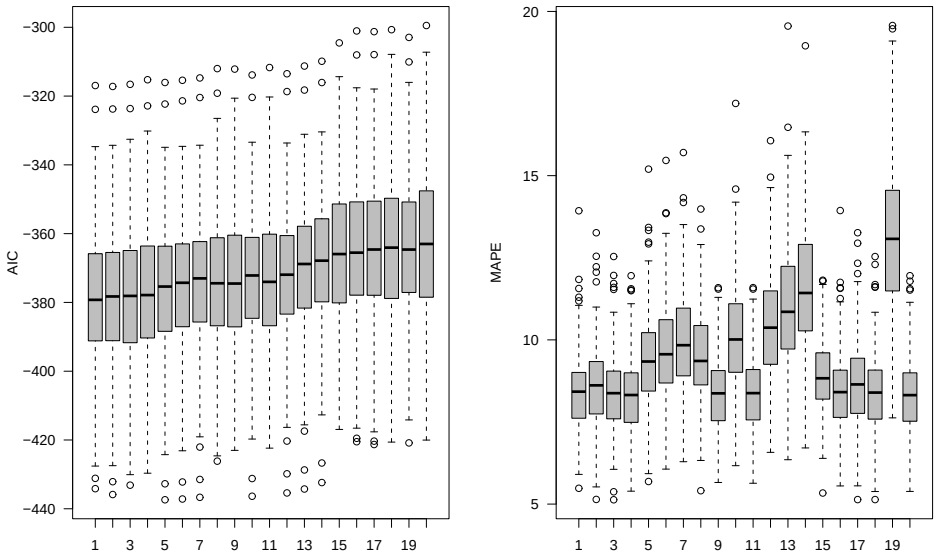


Figura 4.16: Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado dos y sin ruido

| structure | model | $mean.AIC1$ | $sd.AIC1$ | $mean.MAPE$ | $sd.MAPE$ |
|---|---|---|---|---|---|
| AR1 | Int + t^0.5 | -378.99 | 19.78 | 8.81 | 1.33 |
| uniform | Int + t^0.5 | -365.12 | 21.45 | 8.90 | 1.36 |
| AR1 | Int + t^1 | -358.74 | 18.92 | 12.82 | 1.84 |
| AR1 | Int + log(t) | -348.13 | 19.95 | 8.93 | 1.11 |
| uniform | Int + t^1 | -344.53 | 20.06 | 13.22 | 1.92 |
| AR1 | Int + t^1 + t^2 | -379.46 | 20.04 | 8.42 | 1.26 |
| AR1 | Int + t^-0.5 + log(t) | -379.24 | 20.09 | 8.61 | 1.31 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | -378.92 | 20.26 | 8.37 | 1.24 |
| AR1 | Int + t^-1 + t^-0.5 | -377.63 | 20.39 | 8.29 | 1.19 |
| AR1 | Int + t^-1 + t^-1*log(t) | -374.28 | 20.53 | 8.36 | 1.13 |
| AR1 | Int + t^-2 + t^-0.5 | -373.71 | 20.56 | 8.37 | 1.13 |

Tabla 4.15: Tabla para los mejores modelos truncados con su $AIC$ y su $MAPE$ respectivamente.

| structure | model | parameter | $mean.beta$ | $sd.beta$ |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.25 | 0.03 |
| AR1 | Int + t^0.5 | t^0.5 | 0.47 | 0.02 |
| uniform | Int + t^0.5 | Int | -0.26 | 0.03 |
| uniform | Int + t^0.5 | t^0.5 | 0.47 | 0.02 |
| AR1 | Int + t^1 | t^1 | 0.13 | 0.00 |
| AR1 | Int + t^1 | Int | 0.11 | 0.02 |
| hline AR1 | Int + t^1 + t^2 | t^2 | -0.01 | 0.00 |
| AR1 | Int + t^1 + t^2 | t^1 | 0.21 | 0.02 |
| AR1 | Int + t^1 + t^2 | Int | 0.02 | 0.03 |
| AR1 | Int + t^-0.5 + log(t) | Int | -1.06 | 0.21 |
| AR1 | Int + t^-0.5 + log(t) | log(t) | 0.80 | 0.07 |
| AR1 | Int + t^-0.5 + log(t) | t^-0.5 | 1.28 | 0.21 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.23 | 0.12 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.86 | 0.16 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.64 | 0.17 |

Tabla 4.16: Tabla de los parámetros estimados para los mejores modelos truncados.

### 4.2.5. Escenario V

En el escenario VI se dan los resultados en presencia de ruido en algunos perfiles.

Inicialmente se presentan los resultados sin truncamiento.

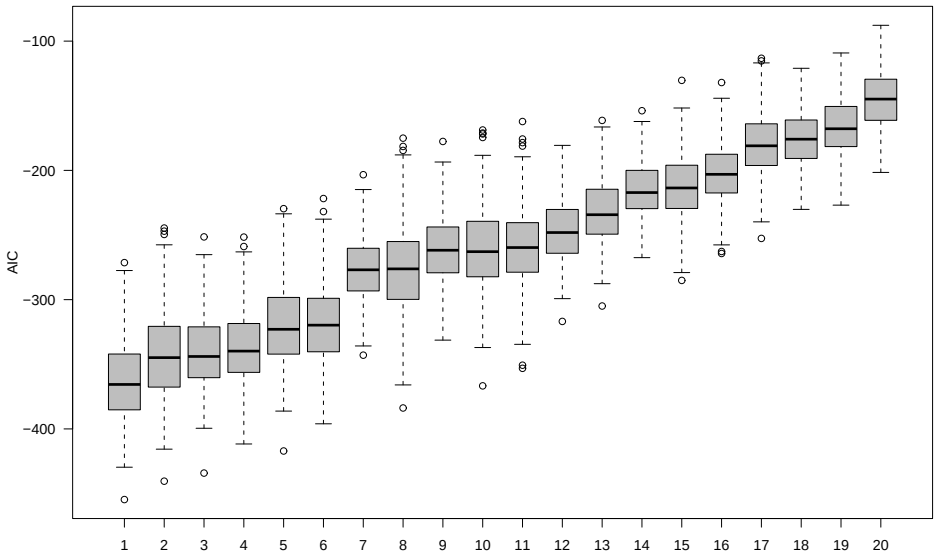


Figura 4.17: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado uno y ruido en algunos perfiles

En la gráfica 4.17 se exponen los mejores modelos con su $AIC$ promedio, de este conjunto de modelos se toman los que poseen menor $AIC$ promedio.

En el diagrama de caja 4.18, se presentan los mejores modelos estimados de grado dos, los primeros modelos muestran un comportamiento del $AIC$ promedio casí constante, por tanto existe una amplia gama de modelos estimados con un buen ajuste.

| structure | model | $mean.AIC1$ | $sd.AIC1$ |
|---|---|---|---|
| AR1 | Int + t^0.5 | -363.35 | 31.36 |
| uniform | Int + t^0.5 | -344.18 | 34.24 |
| AR1 | Int + log(t) | -337.32 | 29.02 |
| uniform | Int + t^1 | -321.27 | 32.19 |
| AR1 | Int + t^-0.5 + log(t) | -362.08 | 31.45 |
| AR1 | Int + log(t) + t^1 | -361.87 | 31.36 |
| AR1 | Int + t^-0.5 + t^1 | -361.11 | 31.26 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | -360.82 | 31.34 |
| AR1 | Int + log(t) + t^2 | -360.43 | 31.13 |

Tabla 4.17: Tabla para los mejores modelos estimados con su $AIC$ promedio.

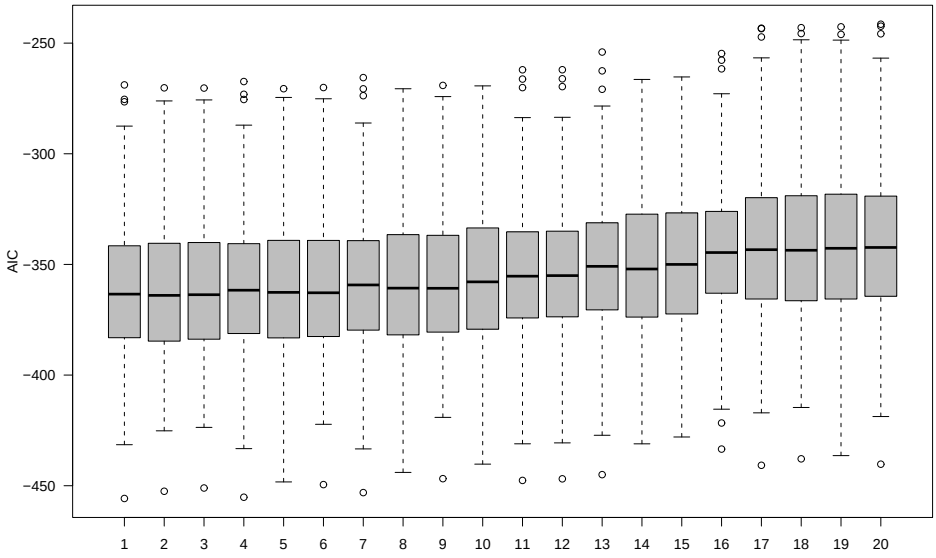


Figura 4.18: Diagrama de caja de los AIC promedio para los mejores modelos estimados de grado dos y con ruido en algunos perfiles

En la tabla 4.17 se observa que el mejor modelo tiene potencia (0.5) y estructura de covarianzas $AR1$, seguido del modelo con potencias $p = (-0.5, 0, )$ con estructura de covarianzas $AR1$.

A través de la tabla 4.18 se puede verificar la estabilidad de la simulación, pues los valores de la desviación estándar en la estimación de cada uno de los parámetros de los mejores modelos es pequeña.

En seguida se analiza el escenario VI, cuando hay truncamiento del último tiempo.

En la grafica 4.19 se exponen los mejores modelos truncados en el último tiempo de grado uno, con ruido en algunos perfiles , muestra que se pueden descartar los modelos que poseen alto $AIC$ promedio y alto $MAPE$. Los modelos con menor $MAPE$ y $AIC$ se deben seleccionar pues permiten hacer mejores pronósticos.

En la grafica 4.20 se muestran los mejores modelos truncados estimados de grado dos, se observa un valor casí constante de los $AIC$ promedio. Para hacer pronósticos se deben seleccionar los $MAPE$ y $AIC$ más pequeños que son los modelos estimados con $p = (-0.5, -0.5)$, $p = (-1, -0.5)$, $p = (-0.5, 1)$ y $p = (-1, 1)$ con estructura de covarianzas $AR1$ y $p = (-0.5, -0.5)$ con estructura de covarianzas $uniforme$.

En la tabla 4.19 se exponen los mejores modelos de grado uno y dos, se caracterizan por tener simultaneamente un valor pequeño de $AIC$ y de $MAPE$.

Para determinar si las simulaciones tienen estabilidad, en la tabla 4.20 se dan los parámetros estimados acompañados de su desviación estándar para los mejores modelos truncados estimados. Mostrando estabilidad pues la desviación estándar al estimar cada

| structure | model | parameter | *mean.beta* | *sd.beta* |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.22 | 0.03 |
| AR1 | Int + t^0.5 | t^0.5 | 0.44 | 0.01 |
| uniform | Int + t^0.5 | Int | -0.22 | 0.03 |
| uniform | Int + t^0.5 | t^0.5 | 0.44 | 0.02 |
| AR1 | Int + t^1 | Int | 0.15 | 0.02 |
| AR1 | Int + t^1 | t^1 | 0.12 | 0.00 |
| AR1 | Int + log(t) | Int | 0.19 | 0.02 |
| AR1 | Int + log(t) | log(t) | 0.37 | 0.01 |
| AR1 | Int + t^-0.5 + log(t) | Int | -0.96 | 0.21 |
| AR1 | Int + t^-0.5 + log(t) | log(t) | 0.75 | 0.07 |
| AR1 | Int + t^-0.5 + log(t) | t^-0.5 | 1.18 | 0.21 |
| AR1 | Int + log(t) + t^1 | t^1 | 0.06 | 0.01 |
| AR1 | Int + log(t) + t^1 | Int | 0.16 | 0.02 |
| AR1 | Int + log(t) + t^1 | log(t) | 0.18 | 0.03 |
| AR1 | Int + t^-0.5 + t^1 | t^-0.5 | -0.36 | 0.07 |
| AR1 | Int + t^-0.5 + t^1 | Int | 0.50 | 0.08 |
| AR1 | Int + t^-0.5 + t^1 | t^1 | 0.08 | 0.01 |

Tabla 4.18: Tabla de los parámetros estimados, (mean.beta) para los mejores modelos.

parámetro promedio es pequeña.

En el escenario V se puede concluir que las muestras generadas con ruido en algunos perfiles presentan los mejores ajustes cuando se utilizan polinomios fraccionarios.

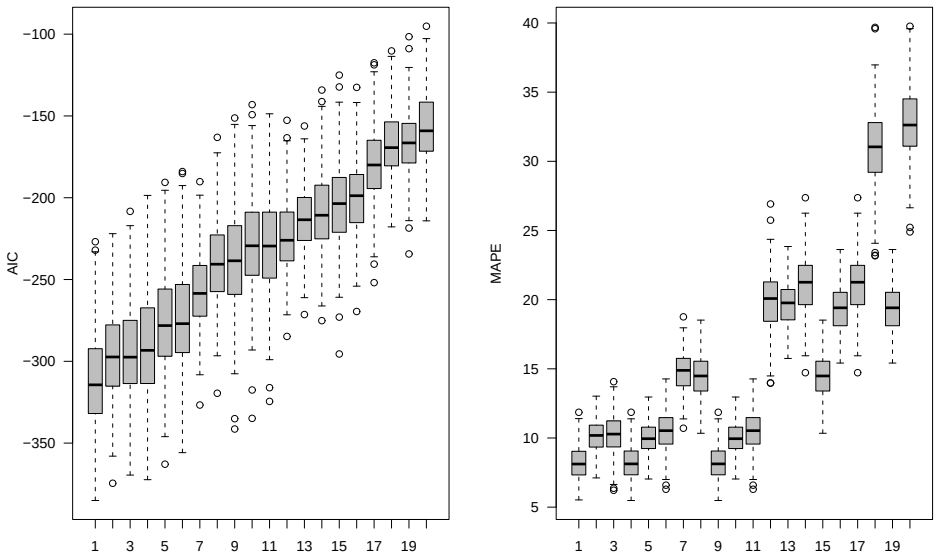


Figura 4.19: Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado uno y ruido en algunos perfiles

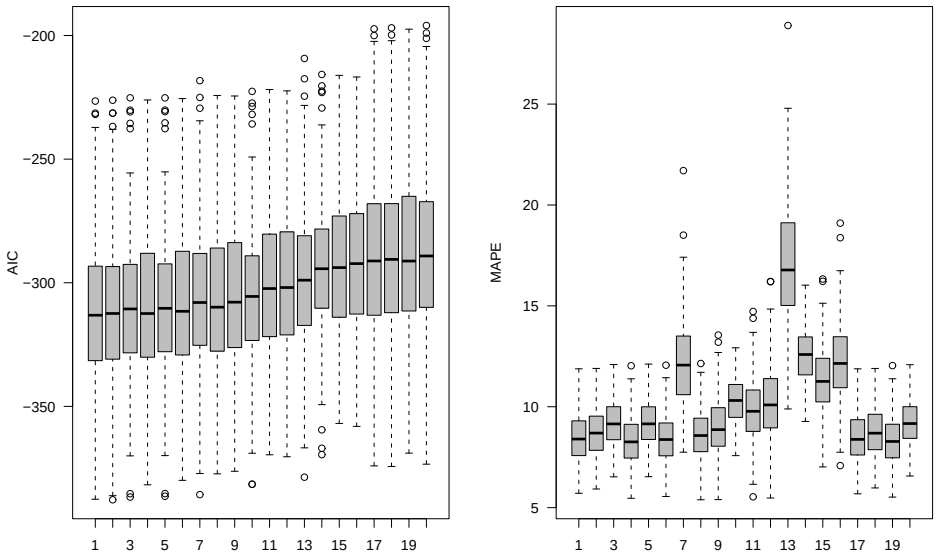


Figura 4.20: Diagrama de caja de los AIC promedio para los mejores modelos truncados estimados de grado dos y ruido en algunos perfiles

| structure | model | $mean.AIC1$ | $sd.AIC1$ | $mean.MAPE$ | $sd.MAPE$ |
|---|---|---|---|---|---|
| AR1 | Int + t^0.5 | -311.57 | 29.27 | 8.20 | 1.22 |
| AR1 | Int + log(t) | -296.38 | 27.28 | 10.16 | 1.21 |
| AR1 | Int + t^1 | -293.62 | 28.29 | 10.24 | 1.51 |
| uniform | Int + t^0.5 | -290.72 | 32.15 | 8.21 | 1.23 |
| uniform | Int + log(t) | -277.11 | 30.01 | 10.00 | 1.21 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | -311.18 | 29.23 | 8.45 | 1.19 |
| AR1 | Int + t^-1 + t^-0.5 | -310.71 | 29.13 | 8.70 | 1.17 |
| AR1 | Int + t^-1 + t^-1*log(t) | -309.16 | 28.89 | 9.21 | 1.16 |
| AR1 | Int + t^-0.5 + t^1 | -309.03 | 29.11 | 8.31 | 1.29 |
| AR1 | Int + t^-2 + t^-0.5 | -308.96 | 28.85 | 9.23 | 1.16 |

Tabla 4.19: Tabla para los mejores modelos truncados con el $AIC$ promedio y el $MAPE$ promedio.

| structure | model | parameter | $mean.beta$ | $sd.beta$ |
|---|---|---|---|---|
| AR1 | Int + t^0.5 | Int | -0.22 | 0.03 |
| AR1 | Int + t^0.5 | t^0.5 | 0.44 | 0.02 |
| AR1 | Int + log(t) | Int | 0.20 | 0.02 |
| AR1 | Int + log(t) | log(t) | 0.35 | 0.01 |
| AR1 | Int + t^1 | t^1 | 0.13 | 0.00 |
| AR1 | Int + t^1 | Int | 0.13 | 0.02 |
| uniform | Int + t^0.5 | Int | -0.22 | 0.03 |
| uniform | Int + t^0.5 | t^0.5 | 0.44 | 0.02 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5 | -2.39 | 0.17 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | t^-0.5*log(t) | -1.08 | 0.12 |
| AR1 | Int + t^-0.5 + t^-0.5*log(t) | Int | 2.62 | 0.17 |
| AR1 | Int + t^-1 + t^-0.5 | t^-0.5 | -3.30 | 0.27 |
| AR1 | Int + t^-1 + t^-0.5 | Int | 1.93 | 0.09 |
| AR1 | Int + t^-1 + t^-0.5 | t^-1 | 1.59 | 0.18 |
| AR1 | Int + t^-1 + t^-1*log(t) | t^-1*log(t) | -1.23 | 0.10 |
| AR1 | Int + t^-1 + t^-1*log(t) | t^-1 | -1.18 | 0.05 |
| AR1 | Int + t^-1 + t^-1*log(t) | Int | 1.40 | 0.05 |
| AR1 | Int + t^-0.5 + t^1 | t^-0.5 | -0.36 | 0.08 |
| AR1 | Int + t^-0.5 + t^1 | Int | 0.50 | 0.09 |
| AR1 | Int + t^-0.5 + t^1 | t^1 | 0.09 | 0.01 |
| AR1 | Int + t^-1 + t^1 | t^1 | 0.32 | 0.05 |
| AR1 | Int + t^-1 + t^1 | Int | 1.93 | 0.09 |
| AR1 | Int + t^-1 + t^1 | t^-1 | -0.19 | 0.04 |

Tabla 4.20: Tabla de los parámetros estimados para los mejores modelos truncados.

# CAPÍTULO 5

# Conclusiones y recomendaciones

Se extendíó la metodología propuesta de Potthoff y Roy de curvas de crecimiento utilizando polinomios fraccionarios, modificando convenientemente la matriz $\mathbf{X}$. A través de la simulación se encontro que al ajustar curvas de crecimiento con polinomios fraccionarios se evitan los problemas que se presentan al ajustar polinomios de bajo orden, que ofrecen una familia limitada de formas y polinomios de alto orden pueden ajustar pobremente los valores extremos. Se mejoran el ajuste en presencia de valores extremos en las covariables en particular cuando existe ruido en algunos perfiles. La extrapolación de un modelo estimado en $p$ tiempos a $p+1$ tiempos, permite hacer pronósticos acertados en el tiempo $p+1$.

El problema analizado en esta tesis de maestría, ofrece elementos que pueden estudiarse en investigaciones futuras. Algunos puntos de interés son los siguientes:

- Estudiar el ajuste de polinomios fraccionarios cuando no se asume distribución normal para la variable respuesta, a través de las ecuaciones de estimación generalizada (EEG), propuestas por Liang y Zeger en 1986.
- Ajustar polinomios fraccionarios a curvas de crecimiento modelando la matriz de varianzas y covarianzas a través de la descomposición de Sholesky.
- En medidas repetidas de variables respuesta binarias sobre el tiempo y bajo diferentes condiciones, se puede caracterizar el patrón de la variable respuesta sobre el tiempo, a través del ajuste de polinomios fraccionarios utilizando el método mínimos cuadrados ponderados para respuestas binarias.

# APÉNDICE A

## Datos de glucosa para pacientes obesos

Los datos de glucosa para pacientes obesos muestran la adimisibilidad en una prueba de tolerancia para 20 pacientes obesos. Se Miden los niveles de fosfato inorgánico del plasma, obtenidos a partir de muestras de sangre extraídas a 0, 0.5, 1, 1.5, 2, 3,4 horas después de administrar la dosis estándar de glucosa.

| Horas (Obeso) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| No | 0 | 0.5 | 1 | 1.5 | 2 | 3 | 4 | 5 |
| 1 | 4.3 | 3.3 | 3.0 | 2.6 | 2.2 | 2.5 | 2.4 | 3.4 |
| 2 | 5.0 | 4.9 | 4.1 | 3.7 | 3.7 | 4.1 | 4.7 | 4.9 |
| 3 | 4.6 | 4.4 | 3.9 | 3.9 | 3.7 | 4.2 | 4.8 | 5.0 |
| 4 | 4.3 | 3.9 | 3.1 | 3.1 | 3.1 | 3.1 | 3.6 | 4.0 |
| 5 | 3.1 | 3.1 | 3.3 | 2.6 | 2.6 | 1.9 | 2.3 | 2.7 |
| 6 | 4.8 | 5.0 | 2.9 | 2.8 | 2.2 | 3.1 | 3.5 | 3.6 |
| 7 | 3.7 | 3.1 | 3.3 | 2.8 | 2.9 | 3.6 | 4.3 | 4.4 |
| 8 | 5.4 | 4.7 | 3.9 | 4.1 | 2.8 | 3.7 | 3.5 | 3.7 |
| 9 | 3.0 | 2.5 | 2.3 | 2.2 | 2.1 | 2.6 | 3.2 | 3.5 |
| 10 | 4.9 | 5.0 | 4.1 | 3.7 | 3.7 | 4.1 | 4.7 | 4.9 |
| 11 | 4.8 | 4.3 | 4.7 | 4.6 | 4.7 | 3.7 | 3.6 | 3.9 |
| 12 | 4.4 | 4.2 | 4.2 | 3.4 | 3.5 | 3.4 | 3.9 | 4.0 |
| 13 | 4.9 | 4.3 | 4.0 | 4.0 | 3.3 | 4.1 | 4.2 | 4.3 |
| 14 | 5.1 | 4.1 | 4.6 | 4.1 | 3.4 | 4.2 | 4.4 | 4.9 |
| 15 | 4.8 | 4.6 | 4.6 | 4.4 | 4.1 | 4.0 | 3.8 | 3.8 |
| 16 | 4.2 | 3.5 | 3.8 | 3.6 | 3.3 | 3.1 | 3.5 | 3.9 |
| 17 | 6.6 | 6.1 | 5.2 | 4.1 | 4.3 | 3.8 | 4.2 | 4.8 |
| 18 | 3.6 | 3.4 | 3.1 | 2.8 | 2.1 | 2.4 | 2.5 | 3.5 |
| 19 | 4.5 | 4.0 | 3.7 | 3.3 | 2.4 | 2.3 | 3.1 | 3.3 |
| 20 | 4.6 | 4.4 | 3.8 | 3.8 | 3.8 | 3.6 | 3.8 | 3.8 |

## A.1. Datos de ratones

Los datos de ratones muestran el peso en libras de 13 ratones medidos en intervalos de tres días (3, 6, 9, 12, 15, 18 y 21 días) desde el nacimiento hasta el destete.

| | Días | | | | | | |
|---|---|---|---|---|---|---|---|
| No | 3 | 6 | 9 | 12 | 15 | 18 | 21 |
| 1 | .190 | .388 | .621 | .823 | 1.078 | 1.132 | 1.191 |
| 2 | .218 | .393 | .568 | .729 | .839 | .852 | 1.004 |
| 3 | .211 | .394 | .549 | .700 | .783 | .870 | .925 |
| 4 | .209 | .419 | .645 | .850 | 1.001 | 1.026 | 1.069 |
| 5 | .193 | .362 | .520 | .530 | .641 | .640 | .751 |
| 6 | .201 | .361 | .502 | .530 | .657 | .762 | .888 |
| 7 | .202 | .370 | .498 | .650 | .795 | .858 | .910 |
| 8 | .190 | .350 | .510 | .666 | .819 | .879 | .929 |
| 9 | .219 | .399 | .578 | .699 | .709 | .822 | .953 |
| 10 | .225 | .400 | .545 | .690 | .796 | .825 | .836 |
| 11 | .224 | .381 | .577 | .756 | .869 | .929 | .999 |
| 12 | .187 | .329 | .441 | .525 | .589 | .621 | .796 |
| 13 | .278 | .471 | .606 | .770 | .888 | 1.001 | 1.105 |

# APÉNDICE B

# Programación en R

## B.1. Obtención de la matriz X para ajustes de curvas de crecimiento con polinomios fraccionarios.

En el programa que se encuentra acontinuación se da la rutina que permite obtener los diferentes subconjuntos de potencias de $S$, junto con la matriz $X$.

```
(X.model <- function (tm ,f, S=c(-2,-1,-0.5,0,0.5,1,2,3) ) {  ### f
número de elementos de S seleccionados.

    m    <- length(S)
    tn   <- length(tm)
    temp <- data.frame(t= tm)

    ## pot<-  t(outer( tm , S , FUN="^") )
    ## print(pot)

    A <-1:m

    if(f>1){
    for(i  in 1:(f-1) ){
       A  <- merge( A  , 1:m )
       k  <- dim(A)[2]
       names(A) <- paste("x",1:k,sep="")
    }}

    ### print(A)
    nomA <- rep( NA ,dim(A)[1])

    matrices <- NULL
```

```
    class(matrices) <- "list"


    for( i in 1:dim(A)[1] ){
      ## i <- 1

          w   <- S[ as.numeric(A[i,]) ]
          c01 <- data.frame(table(w))
          c01$text <- ""
          c01$text <- paste( " + I(t^", c01$w  ,")", sep = "" )

            if(max(c01$Freq) >= 2){
            for(j in 2:max(c01$Freq)){
               cond01  <-  c01$Freq >= j
               c01$text[cond01] <- paste( c01$text[cond01]," +
               I(t^",c01$w[cond01]," * log(t)^",
                j-1 ," )", sep = "" )
             }}

          wc <- paste( "~ + 1" , paste( c01$text , collapse = "") , sep = ""  )
          model00 <- model.matrix( as.formula( wc ) , temp )
           print(wc)
          ##ss <- paste("_",paste(as.numeric(A[i,]),collapse="_"),sep="" )

          nomA[i]       <- paste("C_",paste(as.numeric(A[i,]),collapse="_"),sep="" )
          matrices[[i]] <- model00
    }

    names(matrices) <- nomA
    return(matrices)

} )
```

## B.2. Estimación de los parámetros

### B.2.1. Estimación por mínimos cuadrados generalizados

```
(betas01 <- function(X_, Y_, Z_ ){

  X1   <- X_
  Y1   <- as.matrix(Y_)
  Z1   <- as.matrix(Z_)

  beta0  <- solve( t(X1)%*%X1 )%*%t(X1)%*%t(Y1)%*%Z1%*%solve(t(Z1)%*%Z1 )
  SIGMA  <- diag(dim(Y)[1])
```

```
  K      <- solve(t(X1)%*%solve(SIGMA)%*% X1 ) %*%t(X1)%*%solve(SIGMA)%*% t(Y1)
  SALIDA <- list( beta = beta0, SIGMA = SIGMA, K=K)
  SALIDA

} betas01(X,t(Y),Z)
```

### B.2.2. Estimación por máxima verosimilitud

```
betas.mv <- function(X_, Y_, Z_){

  X1   <- X_
  Y1   <- as.matrix(Y_)
  Z1   <- as.matrix(Z_)
  COV  <- as.matrix(COV_)
  SIGMA  <- diag(dim(Y)[1])
  K      <- solve(t(X1)%*%solve(SIGMA)%*% X1 ) %*%t(X1)%*%solve(SIGMA)%*% t(Y1)
  SALIDA <- list( beta = beta0, SIGMA = SIGMA, K=K)
  SALIDA


  beta.mv  <- solve( t(X1)%*%solve(COV)%*%X1 )%*%t(X1)%*%solve(COV)
  %*%(Y1)%*%t(Z1)%*%solve((Z1)%*%t(Z1) )
  K      <- solve(t(X1)%*%solve(COV)%*% X1 ) %*%t(X1)%*%solve(COV)%*% (Y1)
  SALIDA <- list( beta.mv = beta.mv,K=K)
  SALIDA

}
 betas.mv(X,(Y),Z, COV))
```

#### Estructuras de covarianza

```
(\textbf{Aleatorios}

randomSigma <- function(X_, Y_, Z_ ){

   X1   <- X_
   Y1   <- as.matrix(Y_)
   Z1   <- as.matrix(Z_)

   n <- dim(Y1)[1]
   p <- dim(Y1)[2]
   m <- dim(X1)[2]-1

   Px     <- X1 %*% solve(t(X1)%*%X1) %*% t(X1)
```

```
   Pz     <- Z1 %*% solve(t(Z1)%*%Z1) %*% t(Z1)
   SS     <- t(Y1)  %*%  (diag(1,n) - Pz)  %*%  Y1
   sigma2 <- ( 1 / ( (p-m) * n ) ) * sum(diag(SS))
   gamm <- 1/n * solve( t(X1)%*%X1 )%*% t(X1) %*% SS %*% X1 %*%
     solve( t(X1)%*%X1 ) - sigma2*solve( t(X1)%*%X1 )
   SIGMA  <- (X1 %*% gamm %*% t(X1) ) +  sigma2 * diag(1,p)
   beta0  <- solve( t(X1) %*% solve(SIGMA) %*% X1 )%*%t(X1)%*%
   solve(SIGMA)%*%t(Y1)%*%Z1%*%solve(t(Z1)%*%Z1 )
   K      <- solve(t(X1)%*%solve(SIGMA)%*% X1 )
    %*%t(X1)%*%solve(SIGMA)%*% t(Y1)


  SALIDA <- list( beta=beta0, SIGMA=SIGMA, K=K,gamm=gamm , sigma2=sigma2 )
  SALIDA


}

randomSigma(X, t(Y), Z )

\textbf{Uniforme}

uniformSigma <- function(X_, Y_, Z_ ){

   X1   <- X_
   Y1   <- as.matrix(Y_)
   Z1   <- as.matrix(Z_)

   n <- dim(Y1)[1]
   p <- dim(Y1)[2]
   m <- dim(X1)[2]-1

   uno  <- as.matrix(rep(1,p))

   Px     <- X1 %*% solve(t(X1)%*%X1) %*% t(X1)
   Pz     <- Z1 %*% solve(t(Z1)%*%Z1) %*% t(Z1)
   SS     <- t(Y1)  %*%  (diag(1,n) - Pz)  %*%  Y1

   V         <- (diag(1,p) - Px)%*% t(Y1) %*% Pz
   Sast      <- SS + V %*% t(V)
   sigma.est <- (1/(p*n))* sum(diag(Sast))
   rho       <- 1/(p-1) * (t(uno) %*% Sast %*%
    uno - sum(diag(Sast)) ) /  sum(diag(Sast))
   rho       <- as.numeric(rho)

   SIGMA     <- sigma.est* ( ( 1-rho )* diag(1,p) +  rho*uno%*%t(uno) )
   beta0  <- solve( t(X1) %*% solve(SIGMA) %*% X1 )%*%t(X1)%*%
```

```
   solve(SIGMA)%*%t(Y1)%*%Z1%*%solve(t(Z1)%*%Z1 )
   K      <- solve(t(X1)%*%solve(SIGMA)%*% X1 ) %*%t(X1)%*%solve(SIGMA)%*% t(Y1)


  SALIDA <- list( beta=beta0, SIGMA=SIGMA , K=K   )
  SALIDA


}

   uniformSigma(X, t(Y), Z )

\textbf{estructuras de covarianzas de Rao}


beta.Rao <- function(X_, Y_, Z_ ){

   X1   <-   as.matrix(X_)
   Y1   <- t(as.matrix(Y_))
   Z1   <- t(as.matrix(Z_))

   n <- dim(Y1)[1]
   p <- dim(Y1)[2]
   m <- dim(X1)[2]-1

   Px     <- X1 %*% solve(t(X1)%*%X1) %*% t(X1)
   Pz     <- Z1 %*% solve(t(Z1)%*%Z1) %*% t(Z1)
   SS     <- t(Y1)  %*%  (diag(1,n) - t(Pz))  %*%  Y1
     gama   <- 1/n * solve( t(X1)%*%X1 )%*% t(X1) %*% SS
     %*% X1 %*%  solve( t(X1)%*%X1 )
   xgm    <- X1%*%gama%*%t(X1)
   QO    <- diag(1,p)-X1%*% solve(t(X1)%*%X1)%*%t(X1)
   QOQ   <-1/n * QO %*%Y %*% t(Y) %*% QO
   SIG    <- xgm + QOQ
   betaRAO  <- solve( t(X1) %*% solve(SIG) %*% X1 )%*%t(X1)%*%
   solve(SIG)%*%t(Y1)%*%Z1%*%solve(t(Z1)%*%Z1 )
   K      <- solve(t(X1)%*%solve(SIG)%*% X1 ) %*%t(X1)%*%solve(SIG)%*% t(Y1)

   SALIDA <- list( betaRAO=betaRAO,SIG=SIG, xgm=xgm, QOQ=QOQ, gama=gama, K=K)
  SALIDA

}

beta.Rao(X, Y, Z )


\textbf{estructuras de covarianzas no estructurado}
```

```
beta.NE <- function(X_, Y_, Z_ ){

   X1   <-   as.matrix(X_)
   Y1   <- t(as.matrix(Y_))
   Z1   <- t(as.matrix(Z_))

   n <- dim(Y1)[1]
   p <- dim(Y1)[2]
   m <- dim(X1)[2]-1

   Px     <- X1 %*% solve(t(X1)%*%X1) %*% t(X1)
   Pz     <- Z1 %*% solve(t(Z1)%*%Z1) %*% t(Z1)
   SS     <- t(Y1)  %*%  (diag(1,n) - t(Pz))  %*%  Y1
   betaNE  <- solve( t(X1) %*% solve(SS) %*% X1 )%*%t(X1)%*%
   solve(SS)%*%t(Y1)%*%Z1%*%solve(t(Z1)%*%Z1 )
   K      <- solve(t(X1)%*%solve(SS)%*% X1 ) %*%t(X1)%*%solve(SS)%*% t(Y1)

   SALIDA <- list( betaNE=betaNE,SS=SS, K=K)
  SALIDA

}

beta.random(X, Y, Z ) beta.uniform(X, Y, Z ) beta.Rao(X, Y, Z )
beta.NE(X, Y, Z ))
```

## B.3. Obtención de las estadísticas de prueba

El siguiente programa en R, permite obtener las estadísticas:

```
(\begin{itemize}
\item{}Lambda de Wilks
\item{}La traza de Lawley de Hotelling
\item{}La traza deBatlett-Nanda-Pillai
\item{}Criterio del maximo valor propio de Roy

\end{itemize}

Estas estadísticas se obtienen en una tabla que se presenta en una
salida que contiene: el valor de cada estadística, con sus grados de
libertad del numerador, del  denominador y su P-valor.


 MANOVA<-function(formula,Y,Z,A,C,D){
```

```
Manova.HipotSimple<-manova(formula)
#Betas<-solve(t(Z)%*%Z)%*%t(Z)%*%Y
Qh<-t((A%*%Betas%*%C)-D)%*%solve( A%*%solve(t(Z)%*%Z)%*%t(A) )
%*%((A%*%Betas%*%C)-D)
Qe<-t(C)%*%( (t(Y)%*%Y)-(t(Betas)%*%(t(Z)%*%Z)%*%Betas)  )%*%C

Wilks<- (det(Qe))/(det(Qh+Qe)) p<-qr(Qh+Qe)$rank
q<-qr( A%*%t(Z)%*%Z%*%t(A) )$rank

v<-Manova.HipotSimple$df.residual #Grados de libertad error r<-v-(
(p-q+1)/2 ) u<-((p*q)-2)/4

t<-1 if( ( (p^2)+(q^2)-5 )>0  ) t<- sqrt(      ( ((p*q)^2 ) -4  ) /
(  (p^2)+(q^2)-5   )    ) m<-((abs(p-q))-1)/2 n<-(v-p-1)/2
constante<-( (r*t)-(2*u) )/(p*q) s <- min(p,q)

F.Wilks<-constante*(  ( 1-( Wilks^(1/t) ) )/( Wilks^(1/t) )   )
pvalor.Wilks<-1-pf(F.Wilks, p*q, (r*t)-(2*u))
SalidaWilks<-list(Wilks,F.Wilks,p*q,(r*t)-(2*u),pvalor.Wilks)

Pillai<-sum(diag( Qh%*%solve(Qh+Qe))) # Pillai.
s<-min(p,q) F.Pillai<-(  ( (2*n)+s+1  )/( (2*m)+s+1 ) )*(
Pillai/(s-Pillai)) pvalor.Pillai<-1-pf(F.Pillai,
s*((2*m)+s+1),s*((2*n)+s+1))
SalidaPillai<-list(Pillai,F.Pillai,s*((2*m)+s+1),s*
((2*n)+s+1),pvalor.Pillai)

Lawley<-sum(diag(Qh%*%solve(Qe))) # Lawley-Hotelling.

if(n>1){ b<-(p+(2*n))*(q+(2*n))/(2*((2*n)+1)*(n-1))
c<-(2+((p*q)+2)/(b-1))/(2*n)

F.Lawley<- (Lawley/c)*((4 + ((p*q)+2)/(b-1))/((p*q)))
pvalor.Lawley<-1-pf(F.Lawley, p*q, 4 + ((p*q)+2)/(b-1))
SalidaLawley<-list(Lawley,F.Lawley,p*q,4 + ((p*q)+2)/(b-1)
,pvalor.Lawley)

} if(n<=1){ F.Lawley<- ((2*((s*n)+1)*Lawley)/((s^2) *((2*m)+s+1)))
pvalor.Lawley<-1-pf(F.Lawley, s*((2*m)+s+1), 2*((s*n)+1) )
SalidaLawley<-list(Lawley,F.Lawley,s*((2*m)+s+1), 2*((s*n)+1)
,pvalor.Lawley) }

Roy<-max(eigen(Qh%*%solve(Qe))$values) %Roy
r<-max(p,q) F.Roy<-Roy*((v-r+q)/r) pvalor.Roy<-1-pf(F.Roy, r, v-r+q)
SalidaRoy<-list(Roy,F.Roy,r,v-r+q,pvalor.Roy)
names(SalidaRoy)<-c("Roy","F","Grados Num.","Grados Den.","Pvalor")
```

```
SALIDA<-rbind(SalidaWilks,SalidaPillai,SalidaLawley,SalidaRoy)
colnames(SALIDA)<-c("Estadísticos","F","GradosNum.","GradosDen.","P-valor")
rownames(SALIDA)<-c("Wilks","Pillai","Lawley","Roy" ) SALIDA })
```

# APÉNDICE C

## Modelos estimados

**Polinomios de grado uno y dos**

| | element | listname | model |
|---|---|---|---|
| 1 | 1 | C_1 | ~ + 1 + I(t^-2) |
| 2 | 2 | C_2 | ~ + 1 + I(t^-1) |
| 3 | 3 | C_3 | ~ + 1 + I(t^-0.5) |
| 4 | 4 | C_4 | ~ + 1 + I( log(t) ) |
| 5 | 5 | C_5 | ~ + 1 + I(t^0.5) |
| 6 | 6 | C_6 | ~ + 1 + I(t^1) |
| 7 | 7 | C_7 | ~ + 1 + I(t^2) |
| 8 | 8 | C_8 | ~ + 1 + I(t^3) |

| | element | listname | model |
|---|---|---|---|
| 9 | 9 | C_1_1 | ~ + 1 + I(t^-2) + I(t^-2 * log(t)^1 ) |
| 10 | 10 | C_1_2 | ~ + 1 + I(t^-2) + I(t^-1) |
| 11 | 11 | C_1_3 | ~ + 1 + I(t^-2) + I(t^-0.5) |
| 12 | 12 | C_1_4 | ~ + 1 + I(t^-2) + I( log(t) ) |
| 13 | 13 | C_1_5 | ~ + 1 + I(t^-2) + I(t^0.5) |
| 14 | 14 | C_1_6 | ~ + 1 + I(t^-2) + I(t^1) |
| 15 | 15 | C_1_7 | ~ + 1 + I(t^-2) + I(t^2) |
| 16 | 16 | C_1_8 | ~ + 1 + I(t^-2) + I(t^3) |
| 17 | 17 | C_2_2 | ~ + 1 + I(t^-1) + I(t^-1 * log(t)^1 ) |
| 18 | 18 | C_2_3 | ~ + 1 + I(t^-1) + I(t^-0.5) |
| 19 | 19 | C_2_4 | ~ + 1 + I(t^-1) + I( log(t) ) |
| 20 | 20 | C_2_5 | ~ + 1 + I(t^-1) + I(t^0.5) |
| 21 | 21 | C_2_6 | ~ + 1 + I(t^-1) + I(t^1) |
| 22 | 22 | C_2_7 | ~ + 1 + I(t^-1) + I(t^2) |
| 23 | 23 | C_2_8 | ~ + 1 + I(t^-1) + I(t^3) |
| 24 | 24 | C_3_3 | ~ + 1 + I(t^-0.5) + I(t^-0.5 * log(t)^1 ) |
| 25 | 25 | C_3_4 | ~ + 1 + I(t^-0.5) + I( log(t) ) |
| 26 | 26 | C_3_5 | ~ + 1 + I(t^-0.5) + I(t^0.5) |
| 27 | 27 | C_3_6 | ~ + 1 + I(t^-0.5) + I(t^1) |
| 28 | 28 | C_3_7 | ~ + 1 + I(t^-0.5) + I(t^2) |
| 29 | 29 | C_3_8 | ~ + 1 + I(t^-0.5) + I(t^3) |
| 30 | | | ~ + 1 + I( log(t) ) + I(t^0 * log(t)^1 ) |
| 31 | 30 | C_4_5 | ~ + 1 + I( log(t) ) + I(t^0.5) |
| 32 | 31 | C_4_6 | ~ + 1 + I( log(t) ) + I(t^1) |
| 33 | 32 | C_4_7 | ~ + 1 + I( log(t) ) + I(t^2) |
| 34 | 33 | C_4_8 | ~ + 1 + I( log(t) ) + I(t^3) |
| 35 | 34 | C_5_5 | ~ + 1 + I(t^0.5) + I(t^0.5 * log(t)^1 ) |
| 36 | 35 | C_5_6 | ~ + 1 + I(t^0.5) + I(t^1) |
| 38 | 37 | C_5_8 | ~ + 1 + I(t^0.5) + I(t^3) |
| 39 | 38 | C_6_6 | ~ + 1 + I(t^1) + I(t^1 * log(t)^1 ) |
| 40 | 39 | C_6_7 | ~ + 1 + I(t^1) + I(t^2) |
| 41 | 40 | C_6_8 | ~ + 1 + I(t^1) + I(t^3) |
| 42 | 41 | C_7_7 | ~ + 1 + I(t^2) + I(t^2 * log(t)^1 ) |
| 43 | 42 | C_7_8 | ~ + 1 + I(t^2) + I(t^3) |
| 44 | 43 | C_8_8 | ~ + 1 + I(t^3) + I(t^3 * log(t)^1 ) |

# Bibliografía


[1] Berkey, C. S., and Reed, R. B. A model for describing normal and abnormal growth in early chilhood. Hum. Biol., **59**, 973-987, 1987.

[2] Box, G. E. P., and Tidwell, P. W. Transformation of the independent variables. Techometrics **4**, 531-549, 1962.

[3] Count, E. W. A quantitative analysis of growth in certain human skull dimensions. Hum. Biol., **14**, 143-165, 1942.

[4] Count, E. W. growth patterns of human physique: an approach to kinetic anthrpometryo. Hum. Biol., **15**, 1-32, 1943.

[5] Davidian, M. Appplied longitudinal data analysis. Chapman and Hall, 2005.

[6] Davis, C. S. Statistics methods for analysis of repeated measures. Springer, 73-120, 2005.

[7] Díaz, M. L. Estadística multivariada: inferencia y métodos. Facultad de ciencias,Universidad Nacional de Colombia, 150-172, 2007.

[8] Diggle, P., Heagerty, P., Liang, K., and Zeger, S. Analysis of longitudinal data. Oxford Statistical Science Series, 33-113, 2003.

[9] Draper, N. R., and Smith, H. Applied regression analysis. Jonh Wiley and Sons, 1998.

[10] Gasser, T., Muller, H., Kohler, W., Molinari, L., and Prader, A. Nonparametric regression analysis of growth curves. The Annals of Statistics, 1984.

[11] Geisser, S. Growth curves analysis. Handbook of Statistics, **1**, 89-115, 1980.

[12] Geisser, S. Applications of growth curves prediction. Sankhya, **37**, 239-256, 1975.

[13] Gleser, L. J., and Olkin, I. Linear models in multivariate analysis. Chapell Hill. University North Caroline, 1970.

[14] Guerrero, S., y Melo, O. Optimización de curvas de crecimiento a través de análisis univariado. Revista Colombiana de estadística, 31(2), 193-209, 2008.

[15] Hastie, T., and Tibshirani R. Generalized additive models. Chapman and Hall, 1990.

[16] Kariya, T. Testing in the multivariate general general linear model. Kinokuniya printing company, Tokio, Japon, 1985.

[17] Keramidas, E. M., and Lee, L. C. Selection of a covariance structure for growth curves. Biometrical, **37**, 782-797, 1995.

[18] Kratri, C. A. A note on a MANOVA model applied to problems in growth curves. Ann. Inst. Stat. Math., **18**, 75-86, 1966.

[19] Kshirsagar, A. M., and Boyce, S. Growth curves. STATISTICS: textboooks and monographs. Marcel Dekker, 1-142, 1995.

[20] Lee, J. C. Prediction and estimation of growth curves with special covariance structures. Amer. Statist. Assoc., **83**, 432-440, 1988.

[21] Lee, J. C. Test and model selection for the general growth curve model. Biometrics, **47**, 147-159, 1991.

[22] Lessafre, E., Asefa, M., and Verbeke, G. Assessing the goodness of fit of the laird y ware model: an example: the Jimma infant survival differential longitudinal study. Statistics in Medicine, **18**, 834-854, 1999.

[23] Macchiavelli, R. Analisis de datos longitudinales. 2003.

[24] McCullang, P., and Nelder, J. A. Generalized linear models. Chapman adn Hall, 1983.

[25] Mosteller, F., and Tukey, J. W. Data analysis and regression. Adison Wesley , **18**, 1977.

[26] Pan, J. X., and Fang, K. T. Growth curve models and statistical diagnostics. Springer Series in Statistics, 1-140, 2002.

[27] Pan, W. Akaike's information criterion in generalized estimating equation. Biometrics, **57**, 120-125, 2001.

[28] Peña, D. Analisis de datos multivariantes. McGraw Hill, 304-372, 2002.

[29] Potthoff, R., and Roy, S. A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika, **51** 313-326, 1964.

[30] Rao, C. R. Linear statistical inference and its applications. Jonh Wiley and Sons, 1973.

[31] Rao, C. R. Some statistical methods for comparison of growth curve. Biometrika, **14**, 1-17, 1958.

[32] Rao, C. R. The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, **52**, 447-458, 1965.

[33] Reinsh, C. H. Smoothing by Spline Functions. Numerische Mathematik, **10**, 177-183, 1967.

[34] von Rosen, D. Moments for a multivariate linear model with application to the growth model. J. Multi. Analysis, **35**, 243-259, 1991.

[35] Royston, P., and Altman, D. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Royal Statistical Society, **43**, 429-467, 1994.

[36] Royston, P., and Altman, D. Aproximating statistical functions by using fractional polynomial regression. The Statistician, **46**, 411-422, 1997.

[37] Royston, P., and Sauerbrei, W. Multivariate model-building. A pragmatic approach to regression analysis based on polynomials for modelling continuous variables. John Wiley and Sons, 71-113, 2008.

[38] Sauerbrei, W., Hirmer, M., Benner, A., and Royston, P. Multivariable regression model building by using fractional polynomials: description of SAS, STATA and R programs. Computational Statistics and Data Analysis, 3464-3485, 2006.

[39] Searle, S. R. Linear models. John Wiley and Sons, 1997.

[40] Searly, S. R. Matrix algebra useful for statistics. John Wiley and Sons, 1982.

[41] Searly, S. R., Casella, G., and McCullonch, C. E. Variance components. John Wiley and Sons, 1992.

[42] Singer, J. D., and Willett, J. B. Applied longitudinal data analysis. Oxford, 2003.

[43] Silverman, B. W. Some aspects of the spline smoothing approach to non-parametric curve fitting. Royal Statistics. Soc., **47**, 1-21, 1985.

[44] Srilivastava, M. S. Methods of multivariate statistics. Wiley Series in Probability and Statistics, 365-391, 2002.

[45] Tukey, J. W. On the comparative anatomy of transformations. Annals of Mathematical Statistics, **28**, 602-632, 1957.

[46] Venables W. N. and Ripley B. D. . Modern Applied Statistics with S. Springer, New York, ISBN 0-387-95457-0, http://www.stats.ox.ac.uk/pub/MASS4,2002.

[47] Verveke, G., and Molenberghs, G. Linear mixed models for longitudinal data. Springer Series in Statistics, 3464-3485, 2006.

[48] Whittaker, J. On a new method of graduation. Proc. Edinburgh Math. Soc., **41**, 63-75, 1923.

[49] Wingerd, J. The relation of growtu from birth to two years to sex, parental size and other factors,using Rao's method of the transformed time scale. Hum Biol., **42**, 105-131, 1970.