



CIMAT

Centro de Investigación en Matemáticas A. C.

**Algunas contribuciones al análisis  
de datos de degradación**

**Tesis**

que para obtener el grado de

**Doctor en Ciencias Con Orientación en  
Probabilidad y Estadística**

Presenta

**Abelardo Montesinos López**

Director de Tesis

**Dr. Enrique Raúl Villa Diharce**

Guanajuato, Gto., México, Septiembre de 2016.



Miembros del Jurado Dictaminador

**Presidente:** Dr. José Andrés Christen Gracia (CIMAT)

**Secretario:** Dr. Rogelio Ramos Quiroga (CIMAT)

**Vocal:** Dr. Luis Alberto Escobar Restrepo (LSU)

**Vocal:** Dr. Víctor Manuel Armando Aguirre Torres (ITAM)

**Vocal y director de tesis:** Dr. Enrique Raúl Villa Diharce (CIMAT)

Director de Tesis

---

Dr. Enrique Raúl Villa Diharce



# Índice general

<b>Índice general</b>	<b>I</b>
<b>Índice de cuadros</b>	<b>IV</b>
<b>Índice de figuras</b>	<b>VII</b>
<b>Agradecimientos</b>	<b>XI</b>
<b>1 Introducción</b>	<b>1</b>
<b>2 Modelo de degradación lineal de efectos aleatorios con errores correlacionados</b>	<b>7</b>
2.1 Introducción . . . . .	7
2.2 Modelo . . . . .	8
2.3 Media, densidad y función de distribución del tiempo de cruce . . . . .	9
2.4 Estimación por máxima verosimilitud . . . . .	12
2.5 Identificabilidad del modelo . . . . .	14
2.6 Aplicación del modelo a datos de degradación de lacas . . . . .	15
2.7 Conclusiones . . . . .	27
<b>Apéndices</b>	<b>28</b>
2.A Densidad y función de distribución del tiempo de cruce . . . . .	28
2.B Algoritmo EM para estimar los parámetros del modelo (2.1) . . . . .	36
2.B.1 Cálculos para la implementación . . . . .	38
2.B.2 Resultados de matrices . . . . .	40
2.C Matriz de varianzas y covarianzas del estimador de máxima verosimilitud del intercepto y pendiente del Browniano . . . . .	41
<b>3 Mezcla de procesos Gaussianos inversos</b>	<b>43</b>
3.1 Introducción . . . . .	43

3.2	Proceso Gaussiano inverso . . . . .	44
3.3	Mezcla de procesos Gaussianos inversos . . . . .	45
3.4	Estimación . . . . .	46
3.4.1	Máxima verosimilitud . . . . .	46
3.4.2	Ejemplo . . . . .	49
3.5	Conclusiones . . . . .	52
<b>Apéndices</b>		<b>54</b>
3.A	Proceso Gaussiano inverso Normal-Gamma . . . . .	54
3.B	La log-verosimilitud de la mezcla finita de procesos Gaussianos inversos es acotada . . . . .	54
3.C	Distribución empírica de la prueba de razón de verosimilitud para contrastar $H_0 : \lambda_1 = \lambda_2$ . . . . .	56
<b>4</b>	<b>Un método alternativo para la estimación de vida de anaquel</b>	<b>57</b>
4.1	Introducción . . . . .	57
4.2	Modelación de la función de probabilidad de aceptación . . . . .	59
4.3	Tiempos de rechazo . . . . .	60
4.4	Estimación . . . . .	61
4.5	Ejemplos . . . . .	64
4.5.1	Vida útil de tostadas . . . . .	64
4.5.2	Vida útil de pan . . . . .	67
4.6	Simulación . . . . .	70
4.7	Conclusiones . . . . .	72
<b>5</b>	<b>Comparación de algunos criterios de selección de modelos con datos de vida censurados</b>	<b>75</b>
5.1	Introducción . . . . .	75
5.2	Criterios de selección de modelos . . . . .	77
5.3	Estudio de simulación . . . . .	79
5.3.1	Implementación del mecanismo de censura aleatoria . . . . .	79
5.3.2	Resultados de simulación . . . . .	80
5.4	Conclusiones . . . . .	92
<b>Apéndices</b>		<b>95</b>
5.A	Mecanismos de generación de censura aleatoria por la derecha . . . . .	95
5.A.1	Mecanismo I . . . . .	95

5.A.2 Mecanismo II . . . . .	96
5.B Cuadros de resultados de simulación . . . . .	96
5.C Comportamiento asintótico del AIC . . . . .	114
<b>6 Conclusiones y comentarios generales</b>	<b>127</b>
<b>Referencias</b>	<b>129</b>

# Índice de cuadros

2.1	Ajuste de varios submodelos del modelo (2.1). Las estimaciones del primer modelo obtenidas con el algoritmo EM descrito en la Sección 2.4, mientras que la de los restantes modelos, numéricamente cuando las soluciones no tenían forma explícita. . . . .	21
3.1	Estimaciones de los parámetros de la mezcla de dos procesos Gaussianos inversos (MPGI) y del proceso Gaussiano inverso Normal-Gamma (PGING) (ver modelo en el Apéndice 3.A). . . . .	51
4.1	Datos de 6 consumidores de tostadas charras. . . . .	65
4.1	Promedio (Media), error estándar (EE) y sesgo relativo (SR) en la estimación de cuantiles $p = 0.05, 0.1, 0.5, 0.9$ y $0.95$ bajo la metodología de Hough et al. (2003) y la metodología propuesta en la Sección 4.2 (MDB). Para aproximar la verosimilitud en (4.5) se usaron 20 puntos de cuadratura. . . . .	71
5.B.1	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP. . . .	97
5.B.2	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP. . . . .	98
5.B.3	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP. . . . .	99
5.B.4	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY. . . .	100
5.B.5	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY. . . . .	101



5.B.6	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY. . . . .	102
5.B.7	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5. . . . .	103
5.B.8	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5. . . . .	104
5.B.9	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5. . . . .	105
5.B.10	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 1.5. . . . .	106
5.B.11	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 1.5. . . . .	107
5.B.12	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 1.5. . . . .	108
5.B.13	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 2.5. . . . .	109
5.B.14	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 2.5. . . . .	110
5.B.15	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 2.5. . . . .	111
5.B.16	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 3.5. . . . .	112
5.B.17	(Continuación). Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 3.5. . . . .	113

5.B.18 (Continuación). Proporción de veces que el AIC,  $AIC_c$ , BIC y  $BIC_M$  eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 3.5. . . . . 114

# Índice de figuras

2.1	Trayectorias de degradación del color de lacas a temperatura 167°C. . . . .	16
2.2	Coefficientes de regresión para cada trayectoria. . . . .	18
2.3	(a) Función de distribución empírica de los pseudotiempo (puntos) y función de distribución estimada de los submodelos en el Cuadro 2.1. La estimación que más se aleja del resto corresponde al modelo $Y(t) = \beta_0 + \eta_1 t + \epsilon_t$ . (b) Función de distribución empírica de los pseudotiempo (puntos) y función de distribución estimada junto con intervalos bootstrap (percentil corregido por sesgo)(Meeker y Escobar, 1998, página 333) de 95 % de confianza, de los tres modelos mejores de acuerdo al AIC (ver Cuadro 2.1). . . . .	20
2.4	Función de distribución del tiempo de cruce de los modelos $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$ y $Y(t) = \eta_1 t + \sigma_B B(t)$ . . . . .	23
2.5	Residuos de los modelos $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$ (izquierda) y $Y(t) = \eta_1 t + \sigma_B B(t)$ (derecha). . . . .	24
2.6	Gráfica de probabilidad normal de los residuos estandarizados de los modelos $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$ (izquierda) y $Y(t) = \eta_1 t + \sigma_B B(t)$ (derecha). . . . .	25
2.7	Residuos y gráfica de probabilidad de residuos estandarizados del modelo $Y(t) = \beta_0 + \beta_1 t + \epsilon_t$ . . . . .	26
2.8	Función de distribución del tiempo de cruce de tres modelos. . . . .	26
3.1	Trayectorias de degradación de datos de láser. . . . .	50
3.2	Distribución empírica de los pseudotiempo de falla (puntos negros), estimación de la función de distribución de los tiempos de cruce bajo la mezcla de dos procesos Gaussianos inversos (MPGI) y bajo el proceso Gaussiano inverso Normal-Gamma (PGING), junto con sus correspondientes intervalos bootstrap (percentil corregido por sesgo) de 95 % de confianza . . . . .	51
3.C.1	Distribución por simulación del estadístico de razón de verosimilitud bajo $H_0 : \lambda_1 = \lambda_2$ junto con la distribución $\chi_1^2$ . . . . .	56

4.1	Curva de aceptación de dos consumidores: Derecha con tiempo de rechazo asociado medio e izquierda con tiempo de rechazo pequeño. . . . .	60
4.1	Función de confiabilidad estimada con la metodología de Hough et al. (2003) (WEI-GH) bajo el modelo Weibull, el MDB (4.1) y el MDB (4.1) con consistencia constante ( $\sigma_i = \sigma$ ) (MDBR), todas con intervalo de confianza asintóticos (método delta). . . . .	66
4.2	Cuantiles 2.5 % y 97.5 % de la distribución de las proporciones muestrales de aceptación en cada unos de los tiempos de inspección con $n = 50$ consumidores (datos de tostadas) bajo el modelo MDB (a) y el modelo MDBR (b). . . . .	67
4.3	Función de confiabilidad estimada: a) metodología de Hough et al. (2003) (WEI-GH) bajo el modelo Weibull, MDB (4.1) y MDB (4.1) con consistencia constante ( $\sigma_i = \sigma$ ) (MDBR), todas con intervalos de confianza asintóticos (método delta). . . . .	68
4.4	Cuantiles 2.5 % y 97.5 % de la distribución de las proporciones muestrales de aceptación en cada unos de los tiempos de inspección con $n = 50$ consumidores bajo el MDB (a) y el MDBR (b). . . . .	69
4.5	“Estimación de efectos” individuales. . . . .	70
5.1	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución EXP. . . . .	82
5.2	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución RAY. . . . .	84
5.3	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5. . . . .	87
5.4	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 1.5. . . . .	89

5.5	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 2.5. . . . .	90
5.6	Proporción de veces que el AIC, $AIC_c$ , BIC y $BIC_M$ (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 3.5. . . . .	91
5.C.1	$h(\gamma)$ en la distribución EXP, $0 \leq \gamma \leq 0.5$ . . . . .	117
5.C.2	$h(\gamma)$ en la distribución RAY, $0 \leq \gamma \leq 0.5$ . . . . .	118
5.C.3	$h(\gamma)$ en la distribución EXP, $0.5 \leq \gamma < 1$ . . . . .	120
5.C.4	$h(\gamma)$ en la distribución RAY, $0.5 \leq \gamma < 1$ . . . . .	121



## Agradecimientos

A mi familia y amigos por su siempre apoyo incondicional.

A mi director de tesis, el Dr. Enrique Villa por su apoyo y sus valiosos consejos en la realización de este trabajo.

A mis sinodales, el Dr. Andrés Christen, el Dr. Luis Escobar, el Dr. Rogelio Ramos y el Dr. Víctor Aguirre, por la revisión y sus valiosas observaciones a este trabajo.

Al departamento de Probabilidad y Estadística del Centro de Investigaciones en Matemáticas (CIMAT), por el apoyo y la oportunidad que me brindó durante mi estancia.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico que recibí durante mis estudios de doctorado.





# Capítulo 1

## Introducción

El tema de confiabilidad se ha vuelto indispensable en las discusiones sobre la calidad de los productos que los fabricantes ofrecen a los consumidores, debido a que estos buscan tener en los productos que adquieren, altos niveles de desempeño a través del tiempo, esto es, exigen productos confiables. La confiabilidad, que se define como “la probabilidad de que un producto desempeñe, a un tiempo  $t$ , su función adecuadamente”, es un parámetro que debe estimarse en diferentes etapas en la vida de un producto.

El razonamiento actual que los fabricantes hacen sobre la confiabilidad de sus productos, considera que la confiabilidad debe imprimirse en los productos en el diseño y mejorarse aún más en la etapa de desarrollo. El proceso de introducir y mejorar la confiabilidad en un producto, desde su concepción hasta terminar su producción, se conoce como crecimiento de confiabilidad. A lo largo del tiempo en que se da el crecimiento de la confiabilidad, es necesario monitorear la confiabilidad del producto varias veces para verificar su incremento y el arribo al nivel de confiabilidad establecido como confiabilidad objetivo (target).

La medición de la confiabilidad de un producto no se requiere solamente durante la etapa de crecimiento de confiabilidad, sino además en otros momentos, en que se hagan cambios en el proceso de producción o en sus componentes, debido esto último a un cambio de proveedores, que con frecuencia se hace por consideraciones económicas.

Como hemos visto hay diferentes razones por las que es necesario hacer estudios de confiabilidad, en donde requerimos probar componentes para determinar sus tiempos de vida que luego analizamos para estimar la función de confiabilidad. Los estudios tradicionales de confiabilidad consisten en utilizar pruebas de vida acelerada para obtener, con cierta rapidez, tiempos de vida de los productos sometidos a prueba. En las pruebas de vida acelerada, las unidades trabajan en una condición acelerada hasta que fallan o se

censuran en tiempos preestablecidos por la duración de la prueba. Con estos tiempos de vida acelerada se estima la distribución de los tiempos de vida acelerada, que posteriormente se extrapola a las condiciones de operación del producto, para estimar finalmente los parámetros de confiabilidad requeridos, como percentiles o niveles de confiabilidad en tiempos dados.

Dos factores que se combinan para complicar enormemente la realización de las pruebas de vida acelerada, son **el desarrollo de la tecnología** que brinda la posibilidad de tener componentes de muy alta confiabilidad y la necesidad de **tomar decisiones en tiempos muy cortos**, para tener una ventaja competitiva importante en el mercado. En estos casos puede ser que las pruebas de vida acelerada no nos permitan tener resultados oportunamente y entonces nos vemos obligados a buscar procedimientos de prueba que nos permitan tener resultados en tiempos más cortos. Una alternativa importante a las pruebas de vida, son los estudios de degradación.

Para poder realizar un estudio de degradación que nos permita estimar la confiabilidad de un producto, es necesario contar con una característica de calidad de dicho producto, de tal manera que esta va creciendo o decreciendo a través del tiempo de uso y cuando cruza un valor umbral determinado, se declara la falla. En un estudio de degradación se establece la relación entre la trayectoria de degradación y los tiempos a la falla del producto.

Un modelo de degradación de uso frecuente es el modelo lineal, algunas veces porque el proceso de degradación es lineal, como en el caso de desgaste de llantas, desgaste de dispositivos láser, entre otros, y otras veces, porque las trayectorias de degradación originales no son lineales, pero pueden serlo después de transformar la variable de degradación y/o el tiempo. El software SPLIDA ampliamente difundido por Meeker y Escobar (1998), permite analizar datos de degradación, cuyas trayectorias se pueden linealizar aplicando algunas transformaciones simples en la variable de degradación y/o el tiempo.

Los enfoques de modelación utilizados en los procesos de degradación son los conocidos como modelos de procesos aleatorios y modelos de efectos mixtos. En este trabajo, utilizamos el segundo enfoque, que consiste en modelar la variación entre unidades tomando como efectos aleatorios a los parámetros de las trayectorias de degradación. Para los errores de medición del nivel de degradación, consideramos dos alternativas, una consiste en suponer que son independientes e idénticamente distribuidos con distribución normal y en la otra opción tomamos los errores de medición como un proceso Browniano. Al considerar los errores de medición como un movimiento Browniano, estamos considerando que los errores tienen una estructura de dependencia. Hacemos un estudio comparativo de los modelos que resultan de las diferentes combinaciones de modelos de error y efectos aleatorios considerados.

Los procesos estocásticos, de Wiener, Gamma y Gaussiano inverso se han utilizado en el análisis de degradación. Además de aplicaciones, en la literatura se han reportado algunos estudios de dichos modelos. Estudios recientes han reportado una mayor flexibilidad del tercer modelo, respecto a los dos primeros. Sin embargo, en algunas situaciones reales, la consideración de procesos Gaussianos inversos estándar, no es suficiente para capturar la variabilidad entre unidades, porque los parámetros del proceso pueden variar entre unidades. Esta heterogeneidad puede deberse a una variabilidad existente en la materia prima, en el desempeño de los operadores del proceso, en las condiciones ambientales, además de otras causas. Esta variabilidad se ha considerado en los modelos, que los parámetros provienen de una distribución determinada, esto es, considerando efectos aleatorios en el proceso Gaussiano inverso. Cuando la inclusión de efectos aleatorios no es suficiente para capturar la heterogeneidad presente en los datos Yuan y Ji (2015) proponen la formulación general dada por Meeker y Escobar (1998), incluyendo efectos aleatorios que provienen de una mezcla continua de distribuciones. En este trabajo proponemos utilizar una mezcla finita en lugar de una continua, para modelar la heterogeneidad de un conjunto de datos (tiempos de vida de dispositivos láser) ampliamente conocidos en la literatura. Específicamente consideramos que las trayectorias de degradación provienen de uno de dos procesos Gaussiano inversos. Obtenemos los estimadores máximo verosímiles de los parámetros del modelo propuesto, utilizando un algoritmo EM.

El estudio de tiempos de vida de anaquel es de gran utilidad en la industria de alimentos, para administrar el movimiento de los productos, desde su fabricación hasta su consumo. Para el análisis de datos de vida de anaquel se han utilizado los enfoques de punto de corte y análisis de supervivencia. De acuerdo al análisis de supervivencia, el tiempo que transcurre desde que un producto alimenticio se fabrica hasta que este se rechaza por el consumidor (que denominamos tiempo de rechazo), se considera como una variable aleatoria, que se modela con frecuencia mediante una distribución Weibull o Lognormal. En los estudios de vida de anaquel, a diferencia de muchos estudios de confiabilidad, no observamos tiempos de rechazo exactos, sino datos censurados ya que los participantes en el estudio reciben solo un número finito de muestras del producto alimenticio que tienen diferentes tiempos de almacenamiento y en cada caso aceptan o rechazan el producto. De estos datos, se deducen los tiempos de rechazo censurados.

El procedimiento usual de análisis de datos de vida de anaquel, dado por Hough et al. (2003) consiste en estimar por máxima verosimilitud los parámetros de la función de supervivencia. En este procedimiento, se asume que los datos (en su gran mayoría censurados por intervalo), son dados por los jueces o consumidores participantes en el estudio son y no tienen error. Este supuesto difícilmente se cumple, ya que en este tipo de estudios, donde

las personas funcionan como instrumentos de medición, existe una consistencia limitada en las respuestas de los participantes, y los consumidores inconsistentes dan una respuesta con mayor incertidumbre que los participantes consistentes. En este trabajo, proponemos un método alternativo para modelar y analizar los datos de un estudio sensorial de vida de anaquel que nos permite expresar y cuantificar el tiempo de vida de anaquel y la consistencia de cada participante en el estudio. Inicialmente para cada consumidor estimamos su función de probabilidad de aceptación, de tal forma que queda parametrizada por dos parámetros, uno que corresponda al tiempo de rechazo y el otro a una medida de consistencia. Debido a la gran variabilidad que existe entre las respuestas de los consumidores, no tiene sentido construir una función de aceptación (o rechazo) para todos, sino una para cada uno. El modelo propuesto puede considerarse como un modelo de degradación binario, en donde la característica que observamos a través del tiempo es la aceptación o rechazo del producto, por parte de cada consumidor. El modelo que ajustamos a estas observaciones es una función de probabilidad, monótona decreciente, que inicialmente toma el valor uno y eventualmente va a cero. El valor umbral de la probabilidad de rechazo que tomamos para definir el tiempo de falla o rechazo para cada participante es  $p_f$ . Consideramos que para tiempos pequeños el consumidor aceptará el producto con alta probabilidad y para tiempos grandes el producto tendrá una probabilidad de aceptación muy baja. Suponemos que en un tiempo intermedio, se da el rechazo, cuando la probabilidad de aceptación es  $p_f=0.5$ . En este modelo propuesto, estamos considerando un estudio de vida de anaquel, similar a un proceso de degradación en donde la respuesta es binaria (aceptación-rechazo) y la trayectoria de degradación es la función latente  $p_i(t)$  que expresa la probabilidad de aceptación esperada. En este proceso se declara la falla cuando  $p_i(t)$  alcanza el valor 0.5. Para ilustrar el modelo propuesto se presentan dos ejemplos reales de estudios de vida de anaquel, provenientes de la industria de alimentos.

El análisis de datos de confiabilidad involucra como una etapa inicial muy importante la identificación del modelo para los datos de confiabilidad. Este es un paso difícil, sobre todo en el caso de datos censurados, cuando tenemos solo información parcial de los tiempos de vida. Otra razón por la que se complica la selección del modelo de confiabilidad, es que se tienen diferentes distribuciones de tiempos a la falla, similares que compiten por su similitud. Entre los criterios de selección de uso más frecuente, tenemos el Criterio de Selección de Akaike (AIC) y el Criterio de información Bayesiana (BIC), los cuales difieren solamente en los términos de penalización, aunque sus orígenes fueron diferentes. A través del tiempo, el uso de ambos criterios de selección de modelos ha crecido significativamente.

El interés de conocer que criterios de selección resultan mejores en diferentes situaciones, ha generado varios estudios teóricos y empíricos de estos índices. Aunque han existido

muchos estudios comparativos de los índices AIC y BIC, poca atención se le ha brindado al caso de muestras censuradas, a pesar de la amplia aplicación que podemos ver en la literatura. Es importante notar que estos índices no surgieron en situaciones de información parcial, ocasionada por tener datos censurados. Actualmente podemos encontrar la aplicación de estos índices en software como JMP (2013) que tiene implementados estos criterios, además de una modificación del AIC, no obstante, hasta hoy no se tiene un respaldo teórico ni empírico para estas situaciones. Motivado por esto, se presenta un estudio del efecto que tiene la fracción de censura en el desempeño de estos criterios.

El trabajo está organizado de la siguiente manera:

En el Capítulo 2, se presentan diferentes versiones de un modelo de degradación lineal, considerando modelos lineales de efectos aleatorios y diferentes formas del error de medición. Se estudia y compara el efecto que tienen los diferentes supuestos del modelo lineal, sobre el análisis de los datos, para un conjunto de datos de degradación de lacas.

En el Capítulo 3, se propone una mezcla finita de procesos Gaussianos inversos para modelar la heterogeneidad de trayectorias de degradación, como alternativa a la propuesta de Peng (2015) donde usa una mezcla continua de procesos Gaussianos inversos. La estimación de los parámetros es abordada por máxima verosimilitud empleando el algoritmo EM. El desempeño de éste modelo es ilustrado usando un conjunto de datos y comparado con el modelo de Peng (2015).

En el Capítulo 4, se propone un modelo de efectos mixtos, como modelo de degradación binaria, para datos de vida de anaquel. Este modelo es una alternativa al modelo de supervivencia propuesto por Hough et al. (2003) y tiene la ventaja de incluir en su estructura la inconsistencia de las respuestas de los consumidores participantes en el estudio de vida de anaquel. El modelo propuesto, además de dar como resultado la función de rechazo (o aceptación) del producto por parte de los consumidores, puede brindar información sobre la calidad de la información proporcionada por los participantes.

En el Capítulo 5 se presenta una comparación de 4 criterios de selección de modelos frecuentemente usados. La comparación es hecha por simulación considerando modelos de datos de vida con censura aleatoria, donde se estudian los efectos del número esperado de fallas y la fracción de censura en la muestra.

En el último Capítulo se dan conclusiones y comentarios generales.



# Capítulo 2

## Modelo de degradación lineal de efectos aleatorios con errores correlacionados

### 2.1. Introducción

Gracias a los avances tecnológicos cada vez se tienen productos más confiables, lo que hace difícil la evaluación de sus confiabilidades usando pruebas de vida tradicionales, debido a que se observan pocas fallas durante un experimento típico. Esto aún con métodos que consideran censuras y/o pruebas en las que el tiempo a la falla de los productos se aceleran con altos niveles de esfuerzo.

Por ello, debido a la necesidad de tomar decisiones en tiempos cortos para éste tipo de productos, una alternativa es el uso de datos de degradación cuando existe alguna característica de calidad de éste cuya degradación a través del tiempo puede relacionarse a su confiabilidad. Aun con pocas fallas y error de medición presente, el análisis de degradación produce estimaciones más precisas debido a que toma en cuenta la cantidad de degradación de las observaciones censuradas modelando directamente ésta con el tiempo (Lu y Meeker, 1993; Lu et al., 1996).

Los datos de degradación usualmente son analizados con modelos de trayectorias de degradación general (general degradation path model) que por medio de efectos aleatorios toman en consideración la variabilidad de las unidades. Referencias generales de ésta aproximación son por ejemplo Lu y Meeker (1993), Lu et al. (1996), Lu et al. (1997), Meeker y Escobar (1998), Robinson y Crowder (2000), Bae y Kvam (2004) y Hamada et al. (2008). Uno de éstos modelos frecuentemente usados en la estimación de tiempos de falla es el modelo lineal de efectos aleatorios con estructura de errores independientes (Gertsbakh y Kordonsky, 1969; Lu y Meeker, 1993; Lu et al., 1997; Rao y Zaino, 2004; Yáñez y Granada,

2006; Weaver et al., 2013; Rodríguez-Narciso, 2013; Peng, 2015). Sin embargo, dado que las mediciones de degradación usualmente son hechas a través del tiempo, tal supuesto en algunas situaciones puede no ser apropiado. Por ello, Peng y Tseng (2009) proponen una estructura dependiente de los errores por medio del movimiento Browniano, pero considerando que todos los individuos comienzan del mismo punto conocido. Dado que en algunos conjuntos de datos, los valores de las mediciones de las trayectorias inicialmente son diferentes de cero y estos difieren de unidad a unidad, en ésta tesis hemos considerado una extensión del modelo de Peng y Tseng (2009) donde se agrega un intercepto aleatorio. Esto permite que las trayectorias de degradación comiencen en diferentes valores de cero, además de que éstas difieran de unidad a unidad, como puede ser el caso en algunos conjuntos de datos.

El presente estudio está organizado como sigue. En la Sección 2.2 se presenta una extensión del modelo propuesto por Peng y Tseng (2009). En la Sección 2.3 se obtienen expresiones para la esperanza del tiempo de falla, la densidad y la distribución del tiempo de cruce. En la Sección 2.4 se describe la estimación de los parámetros por medio de máxima verosimilitud, y como alternativa se derivan las expresiones correspondientes del algoritmo EM, y en la Sección 2.5 se discute la identificabilidad de éste. En la Sección 2.6 se ilustra una aplicación del modelo extendido a datos de degradación de lacas. Finalmente se dan algunas conclusiones generales de éste estudio.

## 2.2. Modelo

Sea  $Y(t)$  la degradación acumulada de la característica de interés en el tiempo  $t$ , y supongamos que ésta es medida con cierto error de medición. Además suponiendo que las trayectorias de degradación son lineales y que éstas comienzan en niveles diferentes de cero variando de unidad a unidad, proponemos el siguiente modelo

$$Y(t) = \beta_0 + \beta_1 t + \sigma_B B(t) + \epsilon_t, \quad t \geq 0, \quad (2.1)$$

donde  $\boldsymbol{\beta} = (\beta_0, \beta_1)' \sim N_2(\boldsymbol{\eta}, \boldsymbol{\Psi})$ ,  $\boldsymbol{\eta} = (\eta_0, \eta_1)'$ ,  $\boldsymbol{\Psi} = \begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{bmatrix}$ ,  $B(t)$  denota el movimiento Browniano estándar y  $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$ , y asumimos que las partes aleatorias  $\boldsymbol{\beta}$ ,  $B(t)$  y  $\epsilon_t$  son estadísticamente independientes entre sí.

Cuando  $\sigma_{00} = 0$ , el modelo anterior se reduce a

$$Y(t) = \eta_0 + \beta_1 t + \sigma_B B(t) + \epsilon_t, \quad t \geq 0, \quad (2.2)$$



el cual corresponde al modelo de Peng y Tseng (2009) con intercepto igual a  $\eta_0$ , y además si  $\eta_0 = 0$ , el modelo en (2.1) es exactamente el modelo de Peng y Tseng (2009). Bajo supuestos similares se obtienen otros submodelos, como se describirán posteriormente.

### 2.3. Media, densidad y función de distribución del tiempo de cruce

En general, en el análisis de datos de degradación es difícil obtener la densidad o distribución del tiempo de vida de forma explícita, así que en muchos casos la evaluación de ésta es hecha por medio de simulación Monte Carlo (Lu y Meeker, 1993; Meeker y Escobar, 1998), sin embargo, en algunos casos sencillos esto es posible o al menos se pueden obtener aproximaciones de éstas.

Asumiendo que una unidad falla cuando la degradación de ésta alcanza un nivel crítico  $D_f$ , definiendo a  $T$  como el tiempo en que una trayectoria alcanza dicho umbral, el valor esperado de  $T$  está dado por (ver Apéndice 2.A)

$$E(T) = -\frac{\sigma_{01}}{\sigma_{11}} + \sqrt{\frac{2}{\sigma_{11}}} \left( D_f - \eta_0 + \frac{\sigma_{01}}{\sigma_{11}} \eta_1 \right) D \left[ \frac{\eta_1}{\sqrt{2\sigma_{11}}} \right],$$

donde  $D(x) = \exp(-x^2) \int_0^x \exp(u^2) du$  es la integral de Dawson, y cuando la media del efecto aleatorio de la pendiente es mucho más grande que la desviación estándar de éste,  $\eta_1 \gg \sqrt{\sigma_{11}}$ , entonces ésta puede aproximarse como

$$E(T) \approx \frac{D_f - \eta_0}{\eta_1}.$$

La función de densidad del tiempo a la falla asociada bajo el modelo en (2.1) está dada por (ver Apéndice 2.A)

$$\begin{aligned}
f_T(t; \boldsymbol{\theta}) = & \frac{\sqrt{t(\sigma_\eta^2 t + \sigma_B^2)} \sigma_{00}}{2\pi t c_t} \exp \left\{ -\frac{(D_f - \eta_1 t - \eta_0)^2}{2c_t} - \frac{c_t \left(\frac{d_t}{c_t} - D_f\right)^2}{2t(\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}} \right\} \\
& + \frac{1}{\sqrt{2\pi t} \sqrt{c_t}} \left(\frac{d_t}{c_t} - D_f\right) \exp \left\{ -\frac{(D_f - \eta_1 t - \eta_0)^2}{2c_t} \right\} \\
& \times \left\{ 1 - 2\Phi \left[ -\frac{\left(\frac{d_t}{c_t} - D_f\right)}{\sqrt{\frac{t(\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}}{c_t}}} \right] \right\}, \tag{2.3}
\end{aligned}$$

donde  $\Phi(\cdot)$  es la distribución normal estándar y

$$\begin{aligned}
c_t &= \sigma_{00} + t^2 \sigma_{11} + t(2\sigma_{01} + \sigma_B^2), \\
d_t &= a_t b_t \sigma_{00} + \eta_0 t(\sigma_\eta^2 t + \sigma_B^2), \\
a_t &= D_f - \left(\eta_1 - \eta_0 \frac{\sigma_{01}}{\sigma_{00}}\right) t, \\
b_t &= 1 + \frac{\sigma_{01}}{\sigma_{00}} t, \\
\sigma_\eta^2 &= \sigma_{11} - \frac{\sigma_{01}^2}{\sigma_{00}}.
\end{aligned}$$

Si la varianza del intercepto y la correlación de éste con la pendiente tienden a cero ( $\sigma_{00}, \sigma_{01}, \sigma_{01}/\sigma_{00} \rightarrow 0$ ), entonces

$$\begin{aligned}
c_t &\rightarrow t(\sigma_{11} t + \sigma_B^2), \\
a_t &\rightarrow D_f - \eta_1 t, \\
b_t &\rightarrow 1, \\
d_t &\rightarrow \eta_0 t(\sigma_{11} t + \sigma_B^2), \\
\frac{d_t}{c_t} &\rightarrow \eta_0,
\end{aligned}$$

y así la densidad se reduce a la densidad de Peng y Tseng (2009) con intercepto igual a  $\eta_0$

$$f_T(t; \boldsymbol{\theta}) = \sqrt{\frac{D_f - \eta_0}{2\pi t^3 (\sigma_{11} t + \sigma_B^2)}} \exp \left\{ -\frac{(D_f - \eta_0 - \eta_1 t)^2}{2t(\sigma_{11} t + \sigma_B^2)} \right\}.$$

Para el caso en que  $\text{Var}(\beta_1|\beta_0) < \sigma_B^2(\sigma_B^2 + 4\sigma_{01})/(4\sigma_{00})$ , la función de distribución puede calcularse como (ver Apéndice 2.A)

$$\begin{aligned}
F_T(t; \boldsymbol{\theta}) = & \Phi \left[ \frac{\eta_1 t - (D_f - \eta_0)}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2 + \sigma_{00} + 2\sigma_{01} t + \frac{\sigma_{01}^2}{\sigma_{00}} t^2}} \right] \\
& + \sqrt{\frac{\sigma^2}{\sigma_{00}}} \exp(a_3 - a_2^2/a_1) \\
& \times \Phi \left( \frac{\frac{2\sigma_\eta^2 D_f + \sigma_B^2 \left[ \left( \eta_1 - \frac{\sigma_{01}}{\sigma_{00}} \eta_0 \right) t + D_f \right] - \frac{a_2}{a_1} \left[ 2\sigma_\eta^2 t + \sigma_B^2 \left( 1 - \frac{\sigma_{01}}{\sigma_{00}} t \right) \right]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \frac{\sigma^2 \left[ 2\sigma_\eta^2 t + \sigma_B^2 \left( 1 - \frac{\sigma_{01}}{\sigma_{00}} t \right) \right]^2}{\sigma_B^4 (\sigma_B^2 t + \sigma_\eta^2 t^2)}}}} \right), \tag{2.4}
\end{aligned}$$

donde

$$\begin{aligned}
a_1 &= \frac{2\sigma_\eta^2}{\sigma_B^4} - \frac{1}{2\sigma_{00}} - \frac{2\sigma_{01}}{\sigma_{00}\sigma_B^2}, \\
a_2 &= \frac{2\sigma_\eta^2 D_f}{\sigma_B^4} - \frac{\eta_0}{2\sigma_{00}} - \frac{\frac{\sigma_{01}}{\sigma_{00}} (D_f + \eta_0) - \eta_1}{\sigma_B^2}, \\
a_3 &= \frac{2\sigma_\eta^2 D_f^2}{\sigma_B^4} - \frac{\eta_0^2}{2\sigma_{00}} + \frac{2 \left( \eta_1 - \frac{\sigma_{01}}{\sigma_{00}} \eta_0 \right) D_f}{\sigma_B^2}, \\
\sigma_\eta^2 &= \sigma_{11} - \frac{\sigma_{01}^2}{\sigma_{00}}, \\
\sigma^2 &= 1/(-2a_1).
\end{aligned}$$

De forma similar a la densidad, si  $\sigma_{00}, \sigma_{01}, \sigma_{01}/\sigma_{00} \rightarrow 0$ , ésta se reduce a la función de distribución de los tiempos de cruce obtenida en Peng y Tseng (2009) pero con intercepto

$\eta_0$

$$F_T(t; \boldsymbol{\theta}) = \Phi \left( \frac{\eta_1 t - (D_f - \eta_0)}{\sqrt{\sigma_B^2 t + \sigma_{11} t^2}} \right) + \exp \left( \frac{2\eta_1 (D_f - \eta_0)}{\sigma_B^2} + \frac{2\sigma_{11} (D_f - \eta_0)^2}{\sigma_B^4} \right) \times \Phi \left( -\frac{2\sigma_{11} (D_f - \eta_0) t + \sigma_B^2 (\eta_1 t + D_f - \eta_0)}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_{11} t^2}} \right), \quad t \geq 0.$$

## 2.4. Estimación por máxima verosimilitud

Supongamos que  $n$  unidades son puestas a prueba y cada una de éstas son medidas periódicamente en los tiempos  $t_{1i}, \dots, t_{m_i i}, i = 1, \dots, n$ . Bajo el modelo anterior de forma vectorial la trayectoria para la  $i$ -ésima unidad está dada por

$$\mathbf{Y}_i(\mathbf{t}_i) = \mathbf{X}_i \boldsymbol{\beta}_i + \sigma_B \mathbf{B}_i(\mathbf{t}_i) + \boldsymbol{\epsilon}_i, \quad 1 \leq i \leq n \quad (2.5)$$

donde  $\mathbf{X}_i = [\mathbf{1}_i, \mathbf{t}_i]$ ,  $\mathbf{t}_i = (t_{1i}, \dots, t_{m_i i})'$ ,  $\mathbf{1}_i$  es el vector de unos de dimensión  $m_i \times 1$ ,  $\mathbf{Y}_i(\mathbf{t}_i) = (Y_i(t_{1i}), \dots, Y_i(t_{m_i i}))'$ ,  $\mathbf{B}_i(\mathbf{t}_i) = (B_i(t_{1i}), \dots, B_i(t_{m_i i}))'$  y  $\boldsymbol{\epsilon}_i = (\epsilon_{1i}, \dots, \epsilon_{m_i i})'$ .

Bajo los supuestos del modelo en (2.1), se puede mostrar que el vector de mediciones de degradación en la unidad  $i$ , tiene una distribución normal multivariada con  $\mathbf{X}_i \boldsymbol{\eta}$  y matriz de varianzas-covarianzas  $\boldsymbol{\Sigma}_i$ , es decir,

$$\mathbf{Y}_i(\mathbf{t}_i) \sim N_{m_i}(\mathbf{X}_i \boldsymbol{\eta}, \boldsymbol{\Sigma}_i),$$

donde  $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) = \mathbf{X}_i \boldsymbol{\Psi} \mathbf{X}_i' + \sigma_B^2 \mathbf{Q}_i + \sigma_\epsilon^2 \mathbf{I}_{m_i}$ ,  $\boldsymbol{\theta} = (\sigma_{00}, \sigma_{01}, \sigma_{11}, \sigma_B, \sigma_\epsilon)'$ ,  $\mathbf{Q}_i = [\text{mín}(t_{ki}, t_{li})]_{kl}$ ,  $k, l = 1, \dots, m_i$ , y  $\mathbf{I}_{m_i}$  es la matriz identidad de dimensión  $m_i \times m_i$ . Así, las estimaciones para los parámetros de este modelo se pueden obtener maximizando directamente la verosimilitud asociada, aunque reparametrizando el modelo el problema de estimación se puede hacer un poco más simple. Específicamente, definiendo  $\boldsymbol{\lambda} = (\boldsymbol{\eta}', \sigma_{00}, \sigma_{01}, \sigma_{11}, \sigma_B, \sigma_\epsilon)'$  al vector de parámetros a estimar y considerando la parametrización  $\boldsymbol{\lambda}^* = (\boldsymbol{\eta}', \sigma_{00}^*, \sigma_{01}^*, \sigma_{11}^*, \sigma_B^*, \sigma_\epsilon^*)'$ , donde  $\sigma_{00}^* = \sigma_{00}/\sigma_\epsilon^2$ ,  $\sigma_{01}^* = \sigma_{01}/\sigma_\epsilon^2$ ,  $\sigma_{11}^* = \sigma_{11}/\sigma_\epsilon^2$  y  $\sigma_B^* = \sigma_B/\sigma_\epsilon$ , la log-verosimilitud para este modelo es

$$\ell(\boldsymbol{\lambda}^*; \mathbf{y}) = -\frac{nm}{2} \log(2\pi) - \frac{mn}{2} \log(\sigma_\epsilon^2) - \frac{1}{2} \sum_{i=1}^n \log |\boldsymbol{\Sigma}_i^*| - \frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta})' (\boldsymbol{\Sigma}_i^*)^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta})$$

donde  $\mathbf{y}_i = \mathbf{y}_i(\mathbf{t}_i)$ ,  $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_n)'$ ,  $\Sigma_i^* = \mathbf{X}'_i \Psi^* \mathbf{X}'_i + \sigma_B^{*2} \mathbf{Q}_i + \mathbf{I}_{m_i}$ ,  $\Psi^* = \frac{1}{\sigma_\epsilon^2} \Psi$ , y  $m = \sum_{i=1}^n m_i$ .

Para valores de  $\Psi^*$  y  $\sigma_B^*$  dados, los estimadores de máxima verosimilitud de los restantes parámetros tienen las siguientes expresiones

$$\hat{\boldsymbol{\eta}}(\Psi^*, \sigma_B^*) = \left[ \sum_{i=1}^n \mathbf{X}'_i (\Sigma_i^*)^{-1} \mathbf{X}_i \right]^{-1} \left[ \sum_{i=1}^n \mathbf{X}'_i (\Sigma_i^*)^{-1} \mathbf{y}_i \right] \quad (2.6)$$

$$\hat{\sigma}_\epsilon(\Psi^*, \sigma_B^*) = \sqrt{\frac{1}{nm} \sum_{i=1}^n [\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\eta}}(\Psi^*, \sigma_B^*)]' (\Sigma_i^*)^{-1} [\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\eta}}(\Psi^*, \sigma_B^*)]} \quad (2.7)$$

y de aquí la verosimilitud perfil para  $\Psi^*$  y  $\sigma_B^*$  resulta ser

$$\ell_p(\Psi^*, \sigma_B^*; \mathbf{y}) = -\frac{nm}{2} \log(2\pi) - \frac{mn}{2} \log [\hat{\sigma}_\epsilon^2(\Psi^*, \sigma_B^*)] - \frac{1}{2} \sum_{i=1}^n \log |\Sigma_i^*| - \frac{nm}{2}.$$

Así, los estimadores de máxima verosimilitud pueden obtenerse primero maximizando la log-verosimilitud perfil anterior resultando estimaciones de máxima verosimilitud de  $\Psi^*$  y  $\sigma_B^*$ ,  $\hat{\Psi}^*$  y  $\hat{\sigma}_B^*$ , lo cual corresponde a un problema de optimización de dimensión 4, en vez de 7 sobre la verosimilitud original. Estimaciones de  $\boldsymbol{\eta}$  y  $\sigma_\epsilon$  se obtienen al sustituir  $\hat{\Psi}^*$  y  $\hat{\sigma}_B^*$  en (2.6) y (2.7), respectivamente.

Para el modelo en (2.1) los estimadores de máxima verosimilitud no se pueden obtener de forma analítica, y aunque el problema se puede reducir de dimensionalidad trabajando con la verosimilitud perfil como se describió anteriormente, esto también resulta difícil de implementar usando métodos numéricos que hacen uso de derivadas de primer y/o segundo orden. Para evitar ésta dificultad se propone usar el algoritmo EM (McLachlan y Krishnan, 2008; Pawitan, 2001), en el cual resultan expresiones explícitas. El algoritmo EM es un método iterativo de optimización para encontrar estimaciones de máxima verosimilitud de los parámetros, y cada iteración consiste de dos pasos: en el primero se calcula el valor esperado de la log-verosimilitud de los parámetros basada en los datos completos (datos observados y datos perdidos), con respecto a la distribución condicional de los datos perdidos dado los datos observados y valores actuales de los parámetros (expectation), mientras que en el segundo paso se maximiza ésta log-verosimilitud esperada con respecto a los parámetros (maximization).

Dados los valores actuales de los parámetros,  $\boldsymbol{\lambda}^{(t)} = (\boldsymbol{\eta}^{(t)}, \sigma_{00}^{(t)}, \sigma_{01}^{(t)}, \sigma_{11}^{(t)}, \sigma_B^{(t)}, \sigma_\epsilon^{(t)})'$ , el algoritmo EM consiste en iterar las siguientes expresiones hasta convergencia (ver Apéndice 2.B)

$$\boldsymbol{\eta}_{(t+1)} = \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i \left[ \mathbf{y}_i - \mathbf{Z}_i \widehat{\mathbf{b}}_i \right], \quad (2.8)$$

$$\sigma_{\epsilon(t+1)} = \sqrt{\frac{1}{m} \left[ \sum_{i=1}^n \widehat{\mathbf{e}}'_{i(t)} \widehat{\mathbf{e}}_{i(t)} + \sum_{i=1}^n \text{tr} \left( \mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)} \right) \right]}, \quad (2.9)$$

$$\sigma_{B(t+1)} = \sqrt{\frac{1}{m} \left\{ \sum_{i=1}^n \widehat{\mathbf{b}}'_{2i} \widehat{\mathbf{b}}_{2i} + \sigma_{B(t)}^2 \left[ m - \sigma_{B(t)}^2 \sum_{i=1}^n \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \right) \right] \right\}}, \quad (2.10)$$

$$\boldsymbol{\Psi}_{(t+1)} = \frac{1}{n} \sum_{i=1}^n \left( \mathbf{X}'_i \mathbf{B}_i \mathbf{X}_i + \boldsymbol{\Psi}_{(t)}^{-1} \right)^{-1} + \frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{b}}_{1i} \widehat{\mathbf{b}}'_{1i}, \quad (2.11)$$

donde  $\widehat{\mathbf{e}}_{i(t)} = \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t+1)} - \mathbf{Z}_i \widehat{\mathbf{b}}_i$ ,  $\widehat{\mathbf{b}}_i = \sigma_{\epsilon(t)}^{-2} \mathbf{F}_{i(t)} \mathbf{Z}'_i (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)})$ ,

$\mathbf{F}_{i(t)} = \boldsymbol{\Psi}_{i(t)}^* - \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^*$ ,  $\mathbf{V}_{i(t)} = \left( \mathbf{X}_i \boldsymbol{\Psi}_{i(t)} \mathbf{X}'_i + \sigma_{B(t)}^2 \mathbf{Q}_i + \sigma_{\epsilon(t)}^2 \mathbf{I}_i \right)$  y  $\mathbf{Z}_i = \begin{bmatrix} \mathbf{X}_i & \mathbf{Q}_i^{1/2} \end{bmatrix}$ . Así, la estimación de los parámetros consiste en iterar hasta convergencia las ecuaciones en (2.8-2.11).

## 2.5. Identificabilidad del modelo

La extensión dada en (2.1) está basada en agregar nuevos parámetros, por lo que en esta sección se procede a determinar la identificabilidad de éste modelo ya que ésta es uno de los pilares sobre los que se basa el enfoque clásico de la estimación estadística (Samaniego, 2010). Para más detalles sobre identificabilidad de modelos pueden consultarse Freedman (2009), Samaniego (2010) y Casella y Berger (2002).

Como en Cheng y Peng (2012), donde usan la Proposición 10 de Demidenko (2004) que muestra que la distribución normal está caracterizada únicamente por sus dos primeros momentos, para mostrar que el modelo en (2.5) es identificable basta con ver que  $\mathbf{X}_i \boldsymbol{\eta}_1 = \mathbf{X}_i \boldsymbol{\eta}_2$  y  $\boldsymbol{\Sigma}_i(\boldsymbol{\theta}_1) = \boldsymbol{\Sigma}_i(\boldsymbol{\theta}_2)$  implican que  $\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2$  y  $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ .

Como  $\mathbf{X}_i$  es de rango completo por columnas entonces solamente es necesario verificar la segunda implicación. Para mostrar la segunda implicación observe que si  $\boldsymbol{\theta}_k = \left( \sigma_{00}^{(k)}, \sigma_{01}^{(k)}, \sigma_{11}^{(k)}, \sigma_B^{(k)}, \sigma_{\epsilon}^{(k)} \right)'$ ,  $k = 1, 2$ , y  $\boldsymbol{\Sigma}(\boldsymbol{\theta}_1) = \boldsymbol{\Sigma}(\boldsymbol{\theta}_2)$ , entonces observado los elementos  $(i, j)$  y  $(i, j')$  de  $\boldsymbol{\Sigma}(\boldsymbol{\theta}_2) - \boldsymbol{\Sigma}(\boldsymbol{\theta}_1) = \mathbf{0}$ , donde  $i < j$ ,  $j' < j$ , se tiene que

$$\begin{aligned} \left( \sigma_{00}^{(2)} - \sigma_{00}^{(1)} \right) + (t_i + t_j) \left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_i t_j \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right) + t_i \left( \sigma_B^{(2)} - \sigma_B^{(1)} \right) &= 0 \\ \left( \sigma_{00}^{(2)} - \sigma_{00}^{(1)} \right) + (t_i + t_{j'}) \left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_i t_{j'} \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right) + t_i \left( \sigma_B^{(2)} - \sigma_B^{(1)} \right) &= 0 \end{aligned} \quad (2.12)$$

de donde se sigue que

$$(t_j - t_{j'}) \left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_i (t_j - t_{j'}) \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right) = 0$$

y como  $t_j \neq t_{j'}$ , entonces lo anterior equivale a

$$\left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_i \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right) = 0.$$

De forma similar, pero ahora usando los elementos  $(i', j)$  y  $(i', j')$  con  $i' < i$ , y  $j$  y  $j'$  los mismos que antes, se llega a la misma ecuación anterior para  $i'$ , por lo que

$$\left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_i \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right) = \left( \sigma_{01}^{(2)} - \sigma_{01}^{(1)} \right) + t_{i'} \left( \sigma_{11}^{(2)} - \sigma_{11}^{(1)} \right)$$

y de aquí dado que  $t_i \neq t_{i'}$ , se concluye que  $\sigma_{11}^{(2)} = \sigma_{11}^{(1)}$  y  $\sigma_{01}^{(2)} = \sigma_{01}^{(1)}$ . Luego, al sustituir esto en las ecuaciones en (2.12), con  $i'$  en vez de  $i$  en la segunda de estas, y después de igualar las dos ecuaciones resultantes, se obtiene que

$$\left( \sigma_{00}^{(2)} - \sigma_{00}^{(1)} \right) + t_i \left( \sigma_B^{(2)} - \sigma_B^{(1)} \right) = \left( \sigma_{00}^{(2)} - \sigma_{00}^{(1)} \right) + t_{i'} \left( \sigma_B^{(2)} - \sigma_B^{(1)} \right)$$

de donde nuevamente usando el hecho de que  $t_i \neq t_{i'}$  se concluye que  $\sigma_B^{(2)} = \sigma_B^{(1)}$  y  $\sigma_{00}^{(2)} = \sigma_{00}^{(1)}$ .

Por último, al sustituir en  $\Sigma(\boldsymbol{\theta}_2) - \Sigma(\boldsymbol{\theta}_1) = \mathbf{0}$  lo que hasta aquí hemos obtenido,  $\sigma_{00}^{(2)} = \sigma_{00}^{(1)}$ ,  $\sigma_{01}^{(2)} = \sigma_{01}^{(1)}$ ,  $\sigma_{11}^{(2)} = \sigma_{11}^{(1)}$  y  $\sigma_B^{(2)} = \sigma_B^{(1)}$ , notemos que  $\left( \sigma_\epsilon^{(2)} - \sigma_\epsilon^{(1)} \right) \mathbf{I}_m = \mathbf{0}$ , pero como  $\mathbf{I}_m$  es la matriz identidad entonces  $\sigma_\epsilon^{(2)} = \sigma_\epsilon^{(1)}$ . Así,  $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$  y por lo tanto el modelo dado en (2.5) es identificable.

## 2.6. Aplicación del modelo a datos de degradación de lacas

En ésta parte ilustramos la estimación del modelo propuesto en secciones anteriores, con datos de degradación de un cierto tipo de laca que se aplica en lámina de acero que se utiliza para fabricar estufas, estudiado por Peña de la Mora (2003). Para evaluar la confiabilidad de las lacas se diseñó una prueba de vida acelerada, donde la característica medida fue el color de éstas y el factor de aceleración fue la temperatura. Las temperaturas a las que se sometieron las lacas fueron 167, 175 y 180 grados Celsius ( $^{\circ}\text{C}$ ). Se definió la falla de

las lacas cuando el color de éstas alcanzan un color amarillo más intenso predeterminado de antemano, con valor 3.5 en su escala de medición. El color medido contempla tres características físicas de interés: luminosidad, tinte y saturación Peña de la Mora (2003).

Aunque el propósito del autor fue estimar la distribución del tiempo de vida de las lacas en condiciones normales (150 °C), aquí solamente consideramos los datos correspondientes a las unidades sometidas a 167 °C, los cuales consisten en las mediciones del color hechas en 10 tiempos a cada una de las 15 probetas a las que se ha aplicado laca y que conforman la muestra. Las trayectorias de degradación resultantes se muestran en la Figura 2.1, donde se puede apreciar el comportamiento creciente y lineal de éstas, además de que las magnitudes del color al inicio de la prueba difieren de unidad a unidad. De esta manera, el modelo en (2.1) parece adecuado para describir estos datos, ya que además de tomar en cuenta la variabilidad inicial de la degradación y la variabilidad de la tasa de degradación, también trata de explicar la posible dependencia entre las mediciones que no es explicada por los efectos aleatorios  $\beta_0$  y  $\beta_1$ .

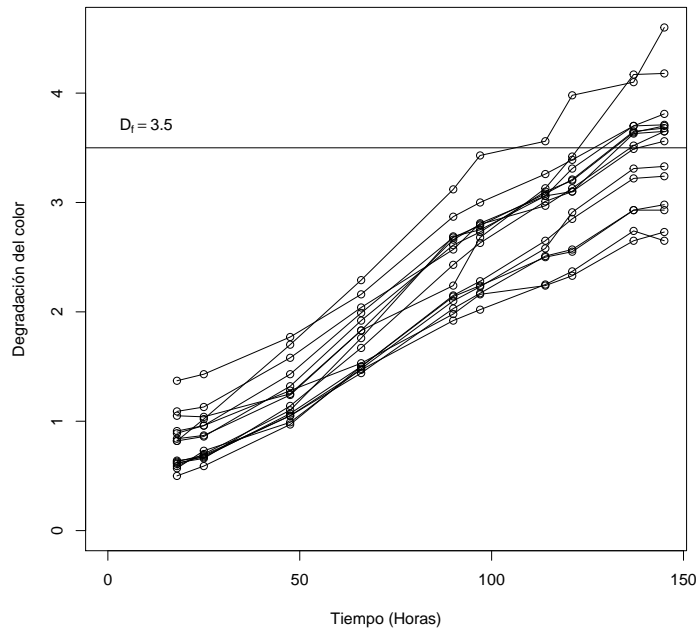


Figura 2.1: Trayectorias de degradación del color de lacas a temperatura 167°C.

Antes de implementar los métodos de estimación por máxima verosimilitud descritos en la Sección 2.4 para el modelo propuesto, primero presentamos los cálculos que se obtienen de estimar una regresión lineal simple a cada trayectoria. En la Figura 2.2 se muestran



los coeficientes de ésta regresión, en la cual se confirma la aleatoriedad de la tasa de degradación ( $\beta_1$ ) y de la degradación inicial ( $\beta_0$ ) entre la unidades. La media muestral y la varianza muestral de éstas estimaciones están dadas por

$$\begin{aligned}\bar{\tilde{\beta}} &= \frac{1}{n} \sum_{i=1}^n \tilde{\beta}_i = \begin{pmatrix} 0.3230 \\ 0.0224 \end{pmatrix}, \\ \tilde{\mathbf{S}}_{\beta} &= \frac{1}{n} \sum_{i=1}^n (\tilde{\beta}_i - \bar{\tilde{\beta}})' (\tilde{\beta}_i - \bar{\tilde{\beta}}) = \begin{bmatrix} 0.0716 & -4.8327 \times 10^{-4} \\ -4.8327 \times 10^{-4} & 1.5374 \times 10^{-5} \end{bmatrix}\end{aligned}$$

reflejando la correlación negativa entre el intercepto y la pendiente del modelo que se observa en la Figura 2.2. Estas estimaciones pueden usarse como valores de inicio en la búsqueda de los estimadores de máxima verosimilitud para los parámetros correspondientes del modelo en (2.1). Un valor de inicio razonable para  $\sigma_{\epsilon}$  se obtiene ponderando las sumas de cuadrados del error que se obtienen en cada regresión de cada trayectoria

$$\tilde{\sigma}_{\epsilon} = \sqrt{\frac{1}{m-2n} \sum_{i=1}^n SCE_i} = 0.1112$$

donde  $SCE_i = \sum_{j=1}^{m_i} (y_{ij} - \hat{y}_{ij})^2$ , y  $\hat{y}_{ij} = \hat{\beta}_{0i} + \hat{\beta}_{1i}t_j$  es el valor predicho de la degradación en  $t_j$  para cada trayectoria.

Tomando  $\tilde{\sigma}_B = \tilde{\sigma}_{\epsilon}$  como valor de inicio para el parámetro  $\sigma_B$  y para los restantes los valores de inicio antes descritos, al usar la verosimilitud perfil con la parametrización descrita en la Sección 2.4, las estimaciones que resultan son las siguientes:

$$\hat{\boldsymbol{\eta}} = \begin{pmatrix} 0.3856 \\ 0.0216 \end{pmatrix}, \hat{\boldsymbol{\Psi}} = \begin{bmatrix} 0.0406 & -2.3818 \times 10^{-4} \\ -2.3818 \times 10^{-4} & 7.1953 \times 10^{-6} \end{bmatrix}, \hat{\sigma}_B = 0.0295, \hat{\sigma}_{\epsilon} = 0.0535.$$

Notemos que los valores de  $\sigma_B$  y  $\sigma_{\epsilon}$  son del mismo orden, pero en la Sección 2.5 se estableció la identificabilidad del modelo, por lo que estos son estimados consistentemente sin la posibilidad de que éstos sean intercambiables en el sentido de que la verosimilitud también alcance el mismo máximo si los valores de estos son intercambiados. Con los mismos valores de inicio, estimaciones muy similares se obtienen con el algoritmo EM que se sugiere en la Sección 2.4:

$$\hat{\boldsymbol{\eta}} = \begin{pmatrix} 0.3856 \\ 0.0216 \end{pmatrix}, \hat{\boldsymbol{\Psi}} = \begin{bmatrix} 0.0406 & -2.3818 \times 10^{-4} \\ -2.3818 \times 10^{-4} & 7.1955 \times 10^{-6} \end{bmatrix}, \hat{\sigma}_B = 0.0295, \hat{\sigma}_{\epsilon} = 0.0535.$$

La función de distribución empírica de los pseudotiempo junto con la función de distribución (2.4) estimada del modelo en (2.1), se muestran en la Figura 2.3, donde se puede

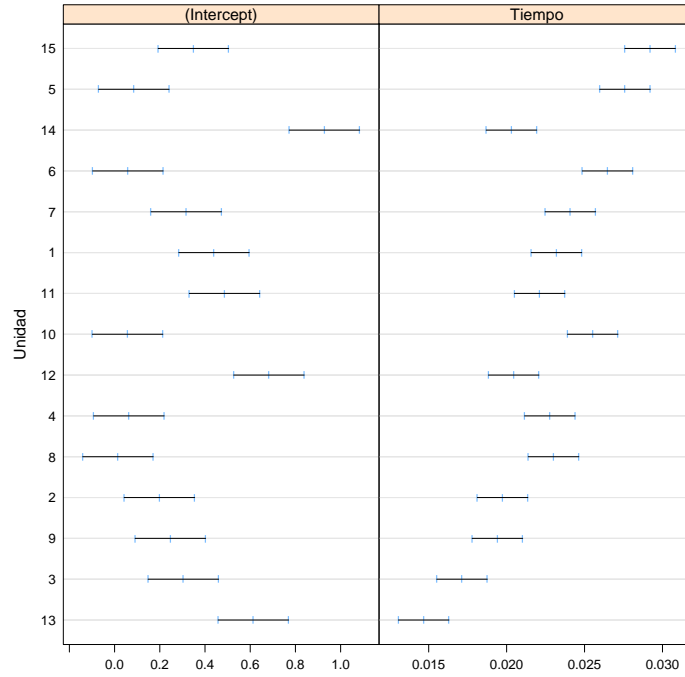


Figura 2.2: Coeficientes de regresión para cada trayectoria.

observar que la función de distribución estimada del modelo propuesto es muy similar a la distribución empírica de los pseudotiempo. Sin embargo, en el modelo propuesto estimado notemos que la desviación de la tasa de degradación es mucho menor que la media de ésta, lo que sugiere estimar submodelos de éste y de aquí elegir, en caso de que los datos lo indiquen, uno quizás más simple pero que siga describiendo de forma adecuada el comportamiento de las trayectorias de degradación observadas. Para tal motivo, en el Cuadro 2.1 se muestra el ajuste de varios submodelos del modelo en (2.1), donde además de mostrar las estimaciones se muestra el AIC para cada uno de ellos.

De acuerdo al AIC el “mejor” modelo es el que no contempla pendiente aleatoria,  $\beta_0 + \eta_1 t + \sigma_B B(t) + \epsilon_t$ , seguido por éste mismo modelo pero sin error de medición ( $\beta_0 + \eta_1 t + \sigma_B B(t)$ ), aunque observemos que las diferencias en los valores del AIC es bastante pequeña (menor que 1), e incluso también el modelo más completo (2.1) difieren poco de acuerdo a éste criterio (Cuadro 2.1). También observemos que la diferencia del AIC entre el modelo propuesto y el modelo lineal de efectos aleatorios (LME ( $\beta_0 + \beta_1 t + \epsilon_t$ )) es considerable, lo que parece indicar que el incorporar el componente del Browniano resulta significativo en el mejoramiento del modelo. Además, si solamente consideramos el intercepto o la pendiente aleatoria en el modelo LME, entonces la magnitud del AIC aumenta de forma drástica a

$-0.1536$  y  $-58.7813$ , respectivamente, por lo que desde esta perspectiva estos submodelos no son buena opción para estos datos, además de que el primero de estos queda descartado de forma más precisa si volteamos a ver las trayectorias de degradación (Figura 2.1), donde se observa que la variabilidad de las mediciones de éstas crece a medida que el tiempo lo hace. Aunque el modelo  $\beta_0 + \eta_1 t + \sigma_B B(t)$  de acuerdo al AIC es una buena alternativa, en caso de error de medición presente en el proceso de obtención de los datos, a menos que se tenga evidencia que la variabilidad de tal error es creciente, tal modelo no sería apropiado dado que la varianza de la parte del Browniano crece ( $\text{Var}(\sigma_B B(t)) = t\sigma_B^2$ ).

Las funciones de distribución de todos los submodelos que aparecen en el Cuadro 2.1 se muestran en la Figura 2.3. De aquí se puede observar que las diferencias comentadas en el AIC respecto a estos submodelos no se ven claramente reflejadas en éstas gráficas (Figura 2.3 (a)), a excepción del modelo  $\beta_0 + \eta_1 t + \epsilon_t$  que es el que más se separa de las restantes. Una comparación más clara de las funciones de distribución correspondientes a los tres mejores modelos de acuerdo al AIC, se muestra en ésta misma Figura (Figura 2.3 (b)), donde también se muestran intervalos bootstrap (percentil corregido por sesgo) (Meeker y Escobar, 1998, página 333) de 95 % de confianza. Podemos ver que para antes de aproximadamente  $t = 150$  horas, tanto puntual como por intervalo, las estimaciones de éstos tres difieren muy poco, y después de tal valor las diferencias tienden a ser ligeramente mayores, donde la incertidumbre es menor en el primero de éstos, y mayor en el último de éstos. Esto puede deberse a que el modelo más completo tiene más fuentes de incertidumbre que el resto, mientras que el segundo aunque tiene una fuente de incertidumbre menos que el primero, la estimación de  $\sigma_B$  resulta más grande al no contemplar el término de error.

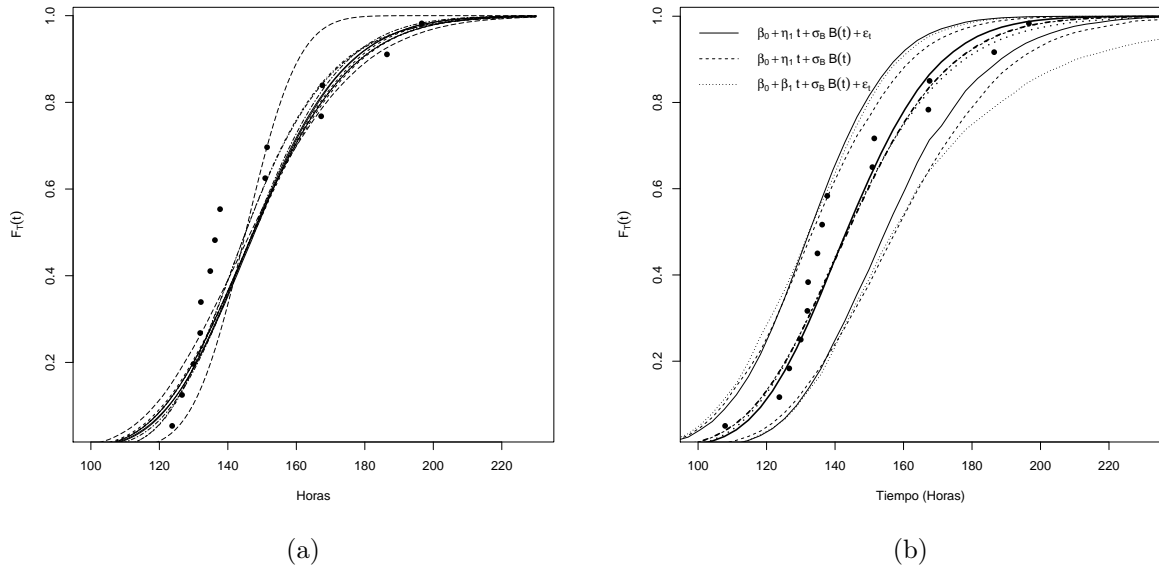


Figura 2.3: (a) Función de distribución empírica de los pseudotiempo (puntos) y función de distribución estimada de los submodelos en el Cuadro 2.1. La estimación que más se aleja del resto corresponde al modelo  $Y(t) = \beta_0 + \eta_1 t + \epsilon_t$ . (b) Función de distribución empírica de los pseudotiempo (puntos) y función de distribución estimada junto con intervalos bootstrap (percentil corregido por sesgo)(Meeker y Escobar, 1998, página 333) de 95 % de confianza, de los tres modelos mejores de acuerdo al AIC (ver Cuadro 2.1).

Modelo	$\eta_0$	$\eta_1$	$\sigma_{00}$	$\sigma_{01}$	$\sigma_{11}$	$\sigma_\epsilon$	$\sigma_B$	$loglik$	$p$	$AIC$
$Y(t) = \beta_0 + \beta_1 t + \sigma_B B(t) + \epsilon_t$	0.3856	0.0216	0.0406	$-2.3818 \times 10^{-4}$	$7.1953 \times 10^{-6}$	0.0535	0.0295	80.5360	7	-147.072
$Y(t) = \beta_0 + \eta_1 t + \sigma_B B(t) + \epsilon_t$	0.3969	0.0214	0.0300			0.0429	0.0340	79.0797	5	-148.1595
$Y(t) = \eta_0 + \beta_1 t + \sigma_B B(t) + \epsilon_t$	0.4026	0.0213			$2.7228 \times 10^{-6}$	0.0364	0.0371	76.5209	5	-143.0419
$Y(t) = \beta_0 + \beta_1 t + \epsilon_t$	0.3231	0.0224	0.0606	$-3.9340 \times 10^{-4}$	$1.3679 \times 10^{-5}$	0.1112		65.4872	6	-118.9745
$Y(t) = \beta_0 + \eta_1 t + \epsilon_t$	0.3230	0.0224	0.0914			0.2010		4.0768	4	-0.1536
$Y(t) = \eta_0 + \beta_1 t + \epsilon_t$	0.3230	0.0223			$1.1461 \times 10^{-5}$	0.1606		33.3906	4	-58.7813
$Y(t) = \beta_0 + \beta_1 t + \sigma_B B(t)$	0.4144	0.0212	0.0326	$-1.6904 \times 10^{-4}$	$3.3175 \times 10^{-6}$		0.0379	78.2920	6	-144.5841
$Y(t) = \beta_0 + \eta_1 t + \sigma_B B(t)$	0.4144	0.0212	0.0267				0.0386	77.9456	4	-147.8913
$Y(t) = \eta_0 + \beta_1 t + \sigma_B B(t)$	0.4144	0.0212			$1.3204 \times 10^{-6}$		0.0402	76.0466	4	-144.0934
$Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$	0.4144	0.0212					0.0404	76.0033	3	-146.0066
$Y(t) = \beta_1 t + \sigma_B B(t)$		0.0241			$2.7319 \times 10^{-14}$		0.0497	45.0997	3	-84.1994
$Y(t) = \eta_1 t + \sigma_B B(t)$		0.0241					0.0497	45.0997	2	-86.1994
$Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t) + \epsilon_t$	0.4054	0.0213				0.0322	0.0382	-76.3244	4	-144.6489

Cuadro 2.1: Ajuste de varios submodelos del modelo (2.1). Las estimaciones del primer modelo obtenidas con el algoritmo EM descrito en la Sección 2.4, mientras que la de los restantes modelos, numéricamente cuando las soluciones no tenían forma explícita.

Para explorar razones de porque dos modelos ajustados a estos datos difieren en gran medida en el AIC pero menos notorio en la estimación de la función de distribución de los tiempos de cruce, consideremos los siguientes modelos:

$$Y(t_{ij}) = \eta_0 + \eta_1 t_{ij} + \sigma_B B(t_{ij}) \quad (2.13)$$

$$Y(t_{ij}) = \eta_1 t_{ij} + \sigma_B B(t_{ij}). \quad (2.14)$$

Ya que de forma vectorial la distribución del vector de mediciones de degradación de la unidad  $i$ , bajo cada uno de los modelos anteriores queda expresada como

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\eta} + \sigma_B B(\mathbf{t}_i) \sim N_{m_i}(\mathbf{X}_i \boldsymbol{\eta}, \sigma_B^2 \mathbf{Q}_i),$$

$$\mathbf{Y}_i = \mathbf{t}_i \eta_1 + \sigma_B B(\mathbf{t}_i) \sim N_{m_i}(\mathbf{t}_i \eta_1, \sigma_B^2 \mathbf{Q}_i),$$

los estimadores de máxima verosimilitud de los parámetros para el primer modelo (2.13) están dados por

$$\hat{\boldsymbol{\eta}} = \left( \sum_{i=1}^n \mathbf{X}_i' \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \left( \sum_{i=1}^n \mathbf{X}_i' \mathbf{Q}_i^{-1} \mathbf{y}_i \right),$$

$$\hat{\sigma}_{B(1)}^2 = \frac{1}{m} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\eta}})' \mathbf{Q}_i^{-1} (\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\eta}}),$$

mientras que para el segundo (2.14)

$$\hat{\eta}_1 = \left( \sum_{i=1}^n \mathbf{t}_i' \mathbf{Q}_i^{-1} \mathbf{t}_i \right)^{-1} \left( \sum_{i=1}^n \mathbf{t}_i' \mathbf{Q}_i^{-1} \mathbf{y}_i \right)$$

$$\hat{\sigma}_{B(2)}^2 = \frac{1}{m} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{t}_i \hat{\eta}_1)' \mathbf{Q}_i^{-1} (\mathbf{y}_i - \mathbf{t}_i \hat{\eta}_1).$$

Estos estimadores se presentan en el Cuadro 2.1, donde podemos ver que el primero de éstos dos modelos también es altamente favorecido por el AIC, mientras que el segundo no (una diferencia de 60 en los valores de AIC). A pesar de ésto (y de hecho también por la prueba  $t$ ,  $t_c = 0.4144/0.0474 = 8.7426$ , ver cálculos en Apéndice 2.C), de la Figura 2.4 se puede observar que no hay una diferencia considerable en la estimación de la función de distribución, sin embargo, el intercepto estimado por el primer modelo es alrededor del 12% en magnitud del umbral de falla ( $D_f = 3.5$ ).

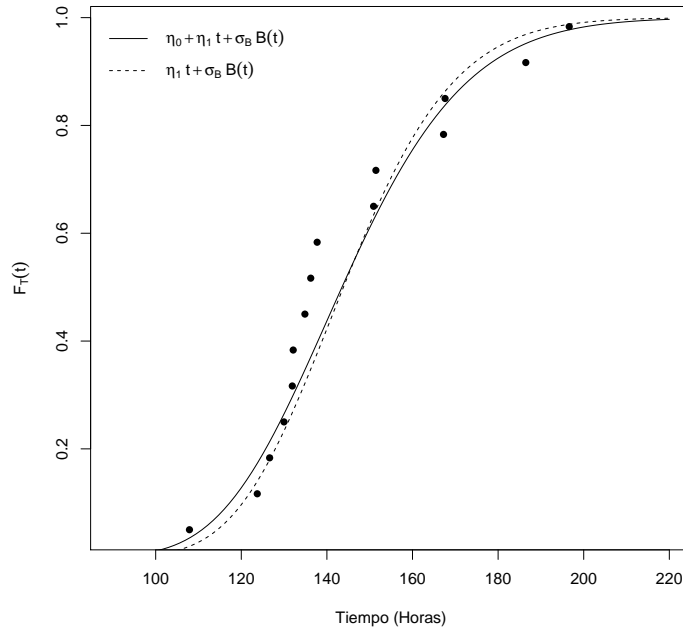


Figura 2.4: Función de distribución del tiempo de cruce de los modelos  $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$  y  $Y(t) = \eta_1 t + \sigma_B B(t)$ .

Los residuos de cada uno de los modelos se muestran en la Figura 2.5, donde se puede observar que los correspondientes al primer modelo tienen un mejor comportamiento dado que estos se encuentran distribuidos alrededor de cero en todas las horas inspeccionadas, mientras que los residuos del segundo modelo se comienzan a comportar de forma esperada aproximadamente después de 70 horas. De esta manera la gran diferencia en las log-verosimilitudes de ambos modelos (y en consecuencia del AIC) se debe a que el primero explica mejor el comportamiento de las trayectorias, aunque tal diferencia no se vea reflejada en la estimación de la función de distribución. Esto último al parecer porque los residuos se comportan mejor en la parte final de la inspección y por lo tanto la predicción implícita de los tiempos que hace el modelo no difieren mucho de las del primero.

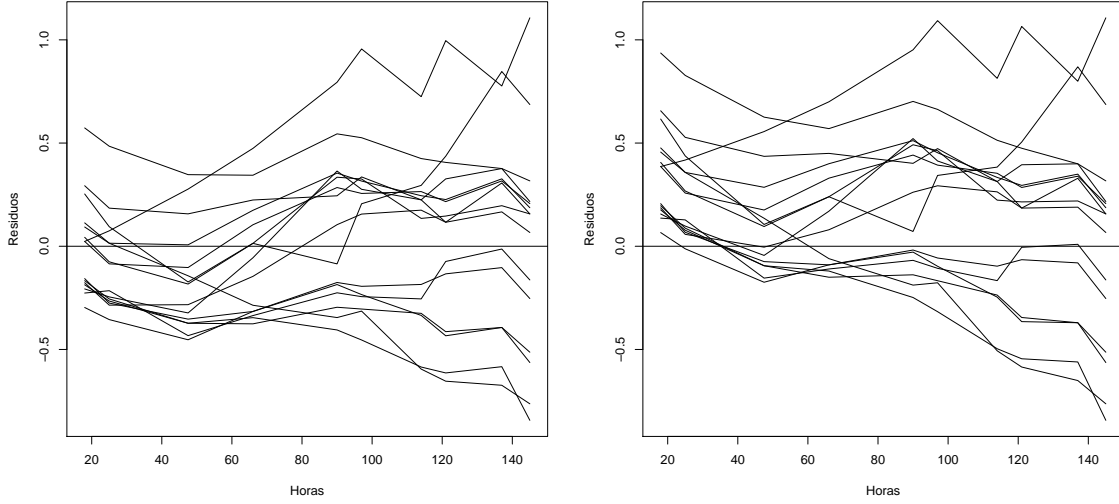


Figura 2.5: Residuos de los modelos  $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$  (izquierda) y  $Y(t) = \eta_1 t + \sigma_B B(t)$  (derecha).

Para ver el comportamiento de la distribución de los residuos, consideremos las siguientes representaciones equivalentes de los modelos anteriores

$$\begin{aligned} \mathbf{Y}_i^* &= \mathbf{X}_i^* \boldsymbol{\eta} + \sigma_B \boldsymbol{\epsilon}_i \sim N_{m_i}(\mathbf{X}_i^* \boldsymbol{\eta}, \sigma_B^2 \mathbf{I}_i) \\ \mathbf{Y}_i^* &= \mathbf{t}_i^* \eta_1 + \sigma_B \boldsymbol{\epsilon}_i \sim N_{m_i}(\mathbf{t}_i^* \eta_1, \sigma_B^2 \mathbf{I}_i), \end{aligned}$$

donde  $\mathbf{Y}_i^* = \mathbf{Q}_i^{-1/2} \mathbf{Y}_i$ ,  $\mathbf{X}_i^* = \mathbf{Q}_i^{-1/2} \mathbf{X}_i$ ,  $\mathbf{t}_i^* = \mathbf{Q}_i^{-1} \mathbf{t}_i$  y  $\boldsymbol{\epsilon}_i = \mathbf{Q}_i^{-1/2} B(\mathbf{t}_i) \sim N(0, \sigma_B^2 \mathbf{I}_i)$ . De aquí otra forma de averiguar características de los modelos anteriores es por medio de la normalidad de los residuos

$$\begin{aligned} \hat{\boldsymbol{\epsilon}}_i^{(1)} &= \frac{1}{\hat{\sigma}_{B(1)}} (\mathbf{Y}_i^* - \mathbf{X}_i^* \hat{\boldsymbol{\eta}}) \\ \hat{\boldsymbol{\epsilon}}_i^{(2)} &= \frac{1}{\hat{\sigma}_{B(2)}} (\mathbf{Y}_i^* - \mathbf{t}_i^* \hat{\eta}_1), \end{aligned}$$

ya que bajo cada modelo estos residuos deberían comportarse aproximadamente como una muestra de una distribución normal estándar de tamaño  $m = \sum_{i=1}^n m_i$ . En la Figura 2.6 se muestran las gráficas de probabilidad normal de éstos residuos correspondientes a los modelos (2.13) y (2.14), donde se puede apreciar que nuevamente el primer modelo se desempeña mejor ya que los residuos correspondientes a éste caen más cerca de la línea recta de  $45^\circ$ , mientras que los residuos del segundo se alejan más en ambas colas. Por lo tanto, ésta otra perspectiva también resulta compatible con el AIC.



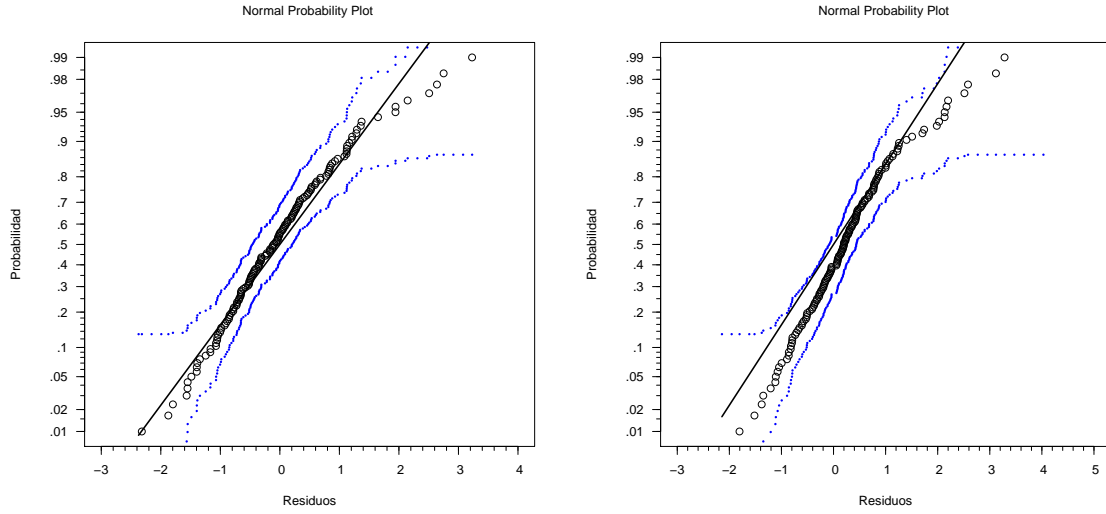


Figura 2.6: Gráfica de probabilidad normal de los residuos estandarizados de los modelos  $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$  (izquierda) y  $Y(t) = \eta_1 t + \sigma_B B(t)$  (derecha).

La función de distribución de los tiempos de falla correspondiente al modelo LME,  $\beta_0 + \beta_1 t + \epsilon_t$ , frecuentemente usado, como puede verse de la Figura 2.3 (a), y más precisamente en la Figura 2.8, tampoco difiere de forma sustancial con las de los modelos (2.13) y (2.14). En la Figura 2.7 se muestra la gráfica de los residuos y la gráfica de probabilidad normal del modelo LME ( $\beta_0 + \beta_1 t + \epsilon_t$ ). Bajo este modelo el supuesto de independencia de los errores no parece razonable dado que en algunos tiempos de inspección los residuos obtenidos están completamente por debajo de cero, mientras que otros por arriba, sin embargo, como se puede ver en la gráfica probabilidad normal, de forma conjunta parecen cumplir el supuesto de normalidad, y en particular con media cero. Esto quizás se debe a que aproximadamente se tiene un comportamiento alternante de los residuos a través de los tiempos de inspección, lo que hace una compensación para obtener aproximadamente media cero de los residuos combinados. Así, aunque no se nota una diferencia notoria en la distribución del tiempo de falla asociada a este modelo con la asociada al primer modelo antes considerado ( $\eta_0 + \eta_1 t + \sigma_B B(t)$ ) (ver Figura 2.6), tal vez éste comportamiento de los residuos es lo que hace la gran diferencia entre el AIC de éste modelo (LME) con el AIC del modelo  $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$ , donde los residuos de éste último tienen un comportamiento más cercano a lo esperado (Figuras 2.5 y 2.6).

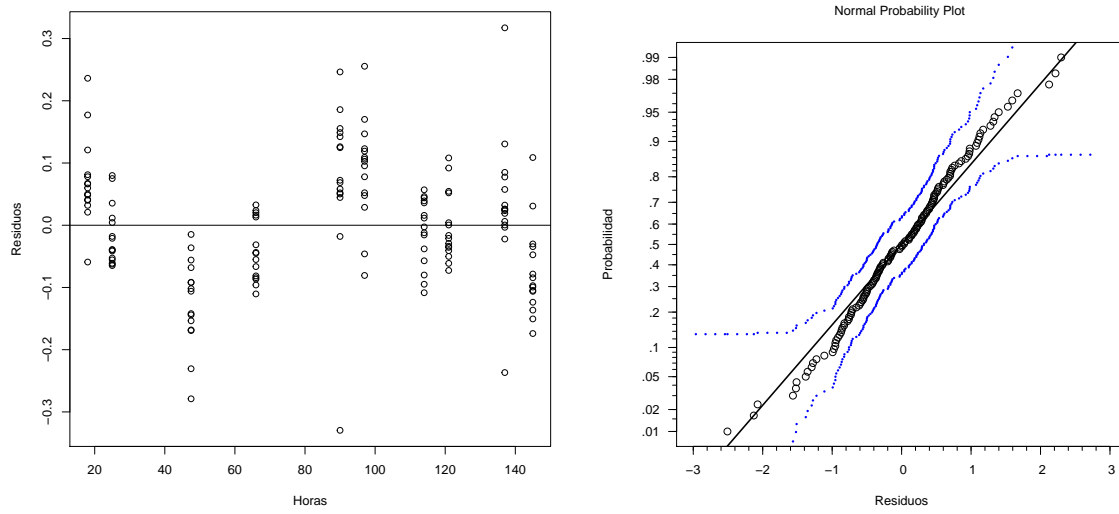


Figura 2.7: Residuos y gráfica de probabilidad de residuos estandarizados del modelo  $Y(t) = \beta_0 + \beta_1 t + \epsilon_t$ .

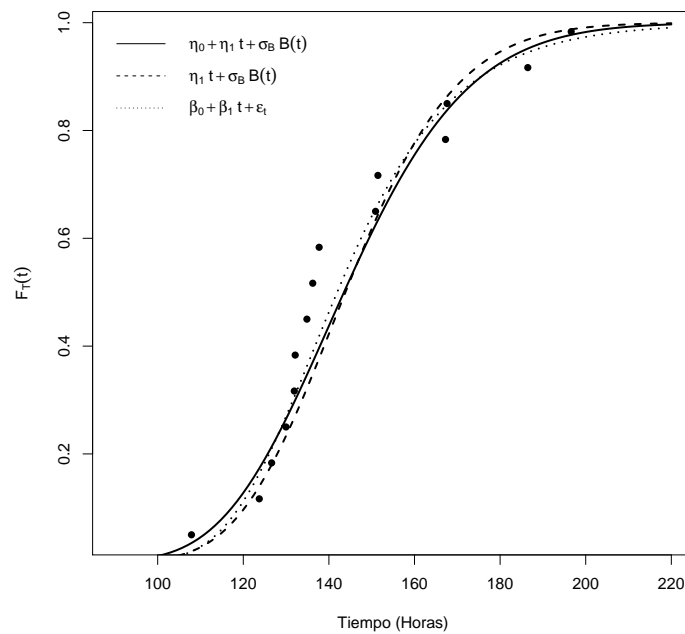


Figura 2.8: Función de distribución del tiempo de cruce de tres modelos.

## 2.7. Conclusiones

Se consideró una extensión del modelo de Peng y Tseng (2009) que permite que las trayectorias de degradación comiencen de un valor diferente de cero, además de que éstas difieran de unidad a unidad, como puede ser el caso en algunos conjuntos de datos. Se obtuvieron expresiones para la densidad y la función de distribución del tiempo de cruce que ésta extensión induce. Dado que la extensión implicó agregar nuevos parámetros, también se discutió la identificabilidad del modelo resultante. La estimación de los parámetros fue abordada por medio de máxima verosimilitud, y como alternativa se derivaron las expresiones correspondientes del algoritmo EM. Para la ilustración del modelo se usó un conjunto de datos reales, y además se ajustaron diferentes submodelos de tal extensión para elegir uno más parsimonioso. Con algunos submodelos se exploraron algunas razones de porque con el criterio AIC se obtuvieron grandes diferencias que no se vieron reflejadas en la estimación de la función de distribución. La mayoría de tales diferencias se debieron a que los modelos con menor AIC no representaban algunas características de los datos.

Así, aunque en general se debe tomar en cuenta la naturaleza de los datos en la elección de un modelo, por medio de simulación Monte Carlo podría averiguarse el efecto de la mala especificación (en tales submodelos) en la estimación de ciertas cantidades (cuantiles, tiempo medio, por ejemplo) variando diferentes factores que pudieran intervenir (tamaño de muestra, frecuencia y número de tiempos de inspección, etcétera) como fue abordado por Peng y Tseng (2009).

# Apéndices

## 2.A. Densidad y función de distribución del tiempo de cruce

Para obtener la función de densidad del tiempo a la falla de las unidades bajo éste modelo se asume que una unidad falla al tiempo  $t$  cuando la degradación de ésta alcanza un nivel crítico  $D_f$ . Notemos que como  $\boldsymbol{\beta} \sim N_2(\boldsymbol{\eta}, \boldsymbol{\Psi})$ , entonces por teoría de la distribución normal multivariada,  $\beta_1|\beta_0 \sim N[\eta_1 + (\sigma_{01}/\sigma_{00})(\beta_0 - \eta_0), \sigma_{11} - (\sigma_{01}^2/\sigma_{00})]$  y así condicionalmente sobre  $\beta_0$ ,  $T = \inf\{t : \beta_0 + \beta_1 t + \sigma_B B(t) > D_f\} = \inf\{t : L(t) > D_f - \beta_0\}$  tiene tanto densidad como distribución las obtenidas en Peng y Tseng (2009) pero con parámetros  $\eta = \eta_1 + (\sigma_{01}/\sigma_{00})(\beta_0 - \eta_0)$ ,  $\sigma_\eta^2 = \sigma_{11} - \sigma_{01}^2/\sigma_{00}$ ,  $\sigma_B^2 = \sigma_B^2$  y umbral  $w = D_f - \beta_0$ . Por teorema de probabilidad total la densidad marginal de  $T$  está dada por

$$\begin{aligned} f_T(t; \boldsymbol{\theta}) &= \int_{-\infty}^{\infty} f_T(t|\beta_0) f(\beta_0) d\beta_0 \\ &= \frac{1}{2\pi \sqrt{t^3(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}} \int_{-\infty}^{\infty} |\beta_0 - D_f| \exp\left\{-\frac{(a_t - b_t \beta_0)^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} - \frac{(\beta_0 - \eta_0)^2}{2\sigma_{00}}\right\} d\beta_0, \end{aligned}$$

donde  $h_1 = \frac{1}{2\pi \sqrt{t^3(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}}$ ,  $a_t = D_f - (\eta_1 - (\sigma_{01}/\sigma_{00})\eta_0)t$ ,  $b_t = 1 + (\sigma_{01}/\sigma_{00})t$ .

Luego, como

$$\frac{(a_t - b_t \beta_0)^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{(\beta_0 - \eta_0)^2}{2\sigma_{00}} = \frac{(\beta_0 - b_t^*/a_t^*)^2}{2\sigma_t^2} + c^{**},$$

donde  $a_t^* = \frac{b_t^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{1}{2\sigma_{00}}$ ,  $b_t^* = \frac{a_t b_t}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{\eta_0}{2\sigma_{00}}$ ,  $c^* = \frac{a_t^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{\eta_0^2}{2\sigma_{00}}$ ,  
 $c^{**} = c^* - \frac{(b_t^*)^2}{a_t^*}$  y  $\sigma_t^2 = \frac{1}{2a_t^*}$ , entonces

$$f_T(t; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi t^3(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}} \exp(-c^{**}) \sqrt{\sigma_t^2} E(|X - D_f|),$$

donde  $X \sim N(b_t^*/a_t^*, \sigma_t^2)$ . Pero como  $X - D_f \sim N(b_t^*/a_t^* - D_f, \sigma_t^2)$  entonces  $|X - D_f|$  se distribuye de acuerdo a una variable aleatoria Folded Normal (Leone et al., 1961) con parámetros  $b_t^*/a_t^* - D_f$  y  $\sigma_t^2$ , es decir,  $|X - D_f| \sim FN(b_t^*/a_t^* - D_f, \sigma_t^2)$ , y de esto  $E|X - D_f| = \sigma_t \sqrt{\frac{2}{\pi}} \exp\left[-\frac{\left(\frac{b_t^*}{a_t^*} - D_f\right)^2}{2\sigma_t^2}\right] + \left(\frac{b_t^*}{a_t^*} - D_f\right) \left[1 - 2\Phi\left[-\frac{\frac{b_t^*}{a_t^*} - D_f}{\sigma_t}\right]\right]$  (Elandt, 1961). Por lo tanto la densidad de  $T$  está dada por

$$f_T(t; \boldsymbol{\theta}) = \frac{1}{\pi \sqrt{t^3(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}} \exp(-c^{**}) \sigma_t^2 \exp\left[-\frac{\left(\frac{b_t^*}{a_t^*} - D_f\right)^2}{2\sigma_t^2}\right] + \frac{1}{\sqrt{2\pi t^3(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}} \exp(-c^{**}) \sqrt{\sigma_t^2} \left(\frac{b_t^*}{a_t^*} - D_f\right) \left[1 - 2\Phi\left[-\frac{\frac{b_t^*}{a_t^*} - D_f}{\sigma_t}\right]\right],$$

pero como  $\sigma_t^2 = \frac{1}{2a_t^*} = \frac{t(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}{b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)}$ ,

$$f_T(t; \boldsymbol{\theta}) = \frac{\sqrt{t(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}}{\pi t [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]} \exp\left[-c^{**} - \frac{\left(\frac{b_t^*}{a_t^*} - D_f\right)^2}{2\sigma_*^2}\right] + \sqrt{\frac{1}{2\pi t^2 [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]}} \exp(-c^{**}) \times \left(\frac{b_t^*}{a_t^*} - D_f\right) \left[1 - 2\Phi\left[-\frac{\frac{b_t^*}{a_t^*} - D_f}{\sigma_*}\right]\right].$$

Además

$$\frac{b_t^*}{a_t^*} = \frac{a_t b_t \sigma_{00} + \eta_0 t (\sigma_\eta^2 t + \sigma_B^2)}{b_t^2 \sigma_{00} + t (\sigma_\eta^2 t + \sigma_B^2)},$$

y

$$\begin{aligned} c^{**} &= c^* - \frac{(b_t^*)^2}{a_t^*} = \frac{a_t^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{\eta_0^2}{2\sigma_{00}} - \frac{\left[\frac{a_t b_t}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{\eta_0}{2\sigma_{00}}\right]^2}{\frac{b_t^2}{2t(\sigma_\eta^2 t + \sigma_B^2)} + \frac{1}{2\sigma_{00}}} \\ &= \frac{a_t^2 t (\sigma_\eta^2 t + \sigma_B^2) \sigma_{00} + \eta_0^2 b_t^2 \sigma_{00} t (\sigma_\eta^2 t + \sigma_B^2) - 2a_t b_t \sigma_{00} \eta_0 t (\sigma_\eta^2 t + \sigma_B^2)}{2 [b_t^2 \sigma_{00} + t (\sigma_\eta^2 t + \sigma_B^2)] t (\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}} \\ &= \frac{a_t^2 + \eta_0^2 b_t^2 - 2a_t b_t \eta_0}{2 [b_t^2 \sigma_{00} + t (\sigma_\eta^2 t + \sigma_B^2)]} = \frac{(a_t - \eta_0 b_t)^2}{2 [b_t^2 \sigma_{00} + t (\sigma_\eta^2 t + \sigma_B^2)]}, \end{aligned}$$

por lo que

$$\begin{aligned}
f_T(t; \boldsymbol{\theta}) = & \frac{\sqrt{t(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}}{\pi t [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]} \exp \left\{ -\frac{(a_t - \eta_0 b_t)^2}{2 [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]} \right\} \times \\
& \exp \left\{ -\frac{[b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)] \left( \frac{a_t b_t \sigma_{00} + \eta_0 t (\sigma_\eta^2 t + \sigma_B^2)}{b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)} - D_f \right)^2}{2t (\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}} \right\} + \\
& \sqrt{\frac{1}{2\pi t^2 [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]}} \exp \left\{ -\frac{(a_t - \eta_0 b_t)^2}{2 [b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)]} \right\} \times \\
& \left( \frac{a_t b_t \sigma_{00} + \eta_0 t (\sigma_\eta^2 t + \sigma_B^2)}{b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)} - D_f \right) \times \\
& \left\{ 1 - 2\Phi \left[ -\frac{\left( \frac{a_t b_t \sigma_{00} + \eta_0 t (\sigma_\eta^2 t + \sigma_B^2)}{b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)} - D_f \right)}{\sqrt{\frac{t(\sigma_\eta^2 t + \sigma_B^2)\sigma_{00}}{b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2)}}} \right] \right\}.
\end{aligned}$$

Como  $a_t = D_f - (\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t$ ,  $b_t = 1 + (\sigma_{01}/\sigma_{00}) t$  y  $\sigma_\eta^2 = \sigma_{11} - \sigma_{01}^2/\sigma_{00}$ , entonces

$$a_t - \eta_0 b_t = D_f - (\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t - \eta_0 (1 + (\sigma_{01}/\sigma_{00}) t) = D_f - \eta_1 t - \eta_0,$$

y

$$\begin{aligned}
b_t^2 \sigma_{00} + t(\sigma_\eta^2 t + \sigma_B^2) &= (1 + (\sigma_{01}/\sigma_{00}) t)^2 \sigma_{00} + t^2 (\sigma_{11} - \sigma_{01}^2/\sigma_{00}) + t\sigma_B^2 \\
&= \sigma_{00} + 2\sigma_{01}t + \sigma_{01}^2/\sigma_{00}t^2 + t^2\sigma_{11} - \sigma_{01}^2/\sigma_{00}t^2 + t\sigma_B^2 \\
&= \sigma_{00} + 2\sigma_{01}t + t^2\sigma_{11} + t\sigma_B^2 \\
&= \sigma_{00} + t^2\sigma_{11} + t(2\sigma_{01} + \sigma_B^2),
\end{aligned}$$

y así

$$\begin{aligned}
f_T(t; \boldsymbol{\theta}) = & \frac{\sqrt{t(\sigma_\eta^2 t + \sigma_B^2)} \sigma_{00}}{\pi t c_t} \exp \left\{ -\frac{(D_f - \eta_1 t - \eta_0)^2}{2c_t} - \frac{c_t \left(\frac{d_t}{c_t} - D_f\right)^2}{2t(\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}} \right\} + \\
& \frac{1}{\sqrt{2\pi t} \sqrt{c_t}} \left(\frac{d_t}{c_t} - D_f\right) \exp \left\{ -\frac{(D_f - \eta_1 t - \eta_0)^2}{2c_t} \right\} \times \\
& \left\{ 1 - 2\Phi \left[ -\frac{\left(\frac{d_t}{c_t} - D_f\right)}{\sqrt{\frac{t(\sigma_\eta^2 t + \sigma_B^2) \sigma_{00}}{c_t}}} \right] \right\}, \tag{2.15}
\end{aligned}$$

donde

$$\begin{aligned}
c_t &= \sigma_{00} + t^2 \sigma_{11} + t(2\sigma_{01} + \sigma_B^2) \\
d_t &= a_t b_t \sigma_{00} + \eta_0 t (\sigma_\eta^2 t + \sigma_B^2) \\
a_t &= D_f - (\eta_1 - \eta_0 (\sigma_{01}/\sigma_{00})) t \\
b_t &= 1 + (\sigma_{01}/\sigma_{00}) t \\
\sigma_\eta^2 &= \sigma_{11} - \sigma_{01}^2/\sigma_{00}.
\end{aligned}$$

Si  $\sigma_{00}, \sigma_{01}, \sigma_{01}/\sigma_{00} \rightarrow 0$ , entonces  $c_t \rightarrow t(\sigma_{11}t + \sigma_B^2)$ ,  
 $a_t = D_f - (\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t \rightarrow D_f - \eta_1 t$  y  $b_t = (1 + (\sigma_{01}/\sigma_{00}) t) \rightarrow 1$ ,  
 $d_t \rightarrow (D_f - \eta_1 t) (0) + \eta_0 t (\sigma_{11}t + \sigma_B^2)$  y  $\frac{d_t}{c_t} \rightarrow \eta_0$ , y así la densidad se reduce a la densidad de Peng y Tseng (2009) con intercepto igual a  $\eta_0$

$$f_T(t) = \sqrt{\frac{D_f - \eta_0}{2\pi t^3 (\sigma_{11}t + \sigma_B^2)}} \exp \left\{ -\frac{(D_f - \eta_0 - \eta_1 t)^2}{2t(\sigma_{11}t + \sigma_B^2)} \right\}.$$

Nuevamente, por el teorema de probabilidad total la función de distribución puede

obtenerse como sigue:

$$\begin{aligned}
F_T(t; \boldsymbol{\theta}) &= \int_{-\infty}^{\infty} F_T(t|\beta_0) f(\beta_0) d\beta_0 \\
&= \underbrace{\int_{-\infty}^{\infty} \Phi \left( \frac{\eta t - w}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right) \frac{1}{\sqrt{\sigma_{00}}} \phi \left( \frac{\beta_0 - \eta_0}{\sqrt{\sigma_{00}}} \right) d\beta_0}_{I_1} + \\
&\quad \underbrace{\int_{-\infty}^{\infty} \exp \left( \frac{2\eta w}{\sigma_B^2} + \frac{2\sigma_\eta^2 w^2}{\sigma_B^4} \right) \Phi \left( -\frac{2\sigma_\eta^2 w t + \sigma_B^2 (\eta t + w)}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right) \frac{1}{\sqrt{\sigma_{00}}} \phi \left( \frac{\beta_0 - \eta_0}{\sqrt{\sigma_{00}}} \right) d\beta_0}_{I_2}.
\end{aligned} \tag{2.16}$$

Como  $\eta = \eta_1 + (\sigma_{01}/\sigma_{00})(\beta_0 - \eta_0)$  y  $w = D_f - \beta_0$ , haciendo el cambio de variable  $z = (\beta_0 - \eta_0)/\sqrt{\sigma_{00}}$ ,  $I_1$  puede re-expresarse como

$$\begin{aligned}
I_1 &= \int_{-\infty}^{\infty} \Phi \left( \frac{[\eta_1 + (\sigma_{01}/\sigma_{00})(\beta_0 - \eta_0)] t - (D_f - \beta_0)}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right) \frac{1}{\sqrt{\sigma_{00}}} \phi \left( \frac{\beta_0 - \eta_0}{\sqrt{\sigma_{00}}} \right) d\beta_0 \\
&= \int_{-\infty}^{\infty} \Phi \left( \frac{\eta_1 t - (D_f - \eta_0) + [\sqrt{\sigma_{00}} + (\sigma_{01}/\sqrt{\sigma_{00}}) t] z}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right) \phi(z) dz,
\end{aligned}$$

y usando el Lema 2.B.1 se tiene que

$$\begin{aligned}
I_1 &= \Phi \left[ \frac{\frac{\eta_1 t - (D_f - \eta_0)}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \left[ \frac{\sqrt{\sigma_{00}} + (\sigma_{01}/\sqrt{\sigma_{00}}) t}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right]^2}} \right] \\
&= \Phi \left[ \frac{\eta_1 t - (D_f - \eta_0)}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2 + \sigma_{00} + 2\sigma_{01} t + (\sigma_{01}^2/\sigma_{00}) t^2}} \right].
\end{aligned} \tag{2.17}$$

Para calcular  $I_2$  primero notemos que



$$\begin{aligned}
\frac{2\eta w}{\sigma_B^2} + \frac{2\sigma_\eta^2 w^2}{\sigma_B^4} - \frac{(\beta_0 - \eta_0)^2}{2\sigma_{00}} &= \frac{2(\eta_1 + (\sigma_{01}/\sigma_{00})(\beta_0 - \eta_0))(D_f - \beta_0)}{\sigma_B^2} + \frac{2\sigma_\eta^2(D_f - \beta_0)^2}{\sigma_B^4} - \\
&\quad \frac{(\beta_0 - \eta_0)^2}{2\sigma_{00}} \\
&= -\frac{(\beta_0 - a_2/a_1)^2}{2\sigma^2} + a_3 - a_2^2/a_1, \tag{2.18}
\end{aligned}$$

donde  $a_1 = \frac{2\sigma_\eta^2}{\sigma_B^4} - \frac{1}{2\sigma_{00}} - \frac{2\sigma_{01}}{\sigma_{00}\sigma_B^2}$ ,  $a_2 = \frac{2\sigma_\eta^2 D_f}{\sigma_B^4} - \frac{\eta_0}{2\sigma_{00}} - \frac{(\sigma_{01}/\sigma_{00})(D_f + \eta_0) - \eta_1}{\sigma_B^2}$ ,  $a_3 = \frac{2\sigma_\eta^2 D_f^2}{\sigma_B^4} - \frac{\eta_0^2}{2\sigma_{00}} + \frac{2(\eta_1 - (\sigma_{01}/\sigma_{00})\eta_0)D_f}{\sigma_B^2}$  y  $\sigma^2 = 1/(-2a_1)$ . Así que de (2.18) y Lema 2.B.1, cuando  $a_1 < 0$  se tiene que

$$\begin{aligned}
I_2 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_{00}}} \exp\left(-\frac{(\beta_0 - a_2/a_1)^2}{2\sigma^2} + a_3 - a_2^2/a_1\right) \times \\
&\quad \Phi\left(-\frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00})\eta_0)t + D_f] - \beta_0 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00})t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}\right) d\beta_0 \\
&= \frac{\sqrt{\sigma^2}}{\sqrt{\sigma_{00}}} \exp(a_3 - a_2^2/a_1) E_{\tilde{\beta}_0} \left[ \Phi\left(-\frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00})\eta_0)t + D_f]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} + \right. \right. \\
&\quad \left. \left. \frac{\tilde{\beta}_0 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00})t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}\right) \right],
\end{aligned}$$

donde  $\tilde{\beta}_0 \sim N(a_2/a_1, \sigma^2)$ . Luego usando nuevamente el Lema 2.B.1 se tiene que

$$\begin{aligned}
& E_{\tilde{\beta}_0} \left[ \Phi \left( \frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t + D_f] - \tilde{\beta}_0 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right) \right] \\
&= \Phi \left( \frac{\frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t + D_f] - a_2/a_1 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \left( \frac{\sigma [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}} \right)^2}} \right) \\
&= \Phi \left( \frac{\frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t + D_f] - a_2/a_1 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \frac{\sigma^2 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]^2}{\sigma_B^4 (\sigma_B^2 t + \sigma_\eta^2 t^2)}}} \right).
\end{aligned} \tag{2.19}$$

Sustituyendo (2.17) y (2.19) en (2.16) se tiene que la función de distribución de los tiempo de cruce está dada por

$$\begin{aligned}
F_T(t; \boldsymbol{\theta}) = & \Phi \left[ \frac{\eta_1 t - (D_f - \eta_0)}{\sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2 + \sigma_{00} + 2\sigma_{01} t + (\sigma_{01}^2/\sigma_{00}) t^2}} \right] + \\
& \sqrt{\frac{\sigma^2}{\sigma_{00}}} \exp(a_3 - a_2^2/a_1) \times \\
& \Phi \left( \frac{\frac{2\sigma_\eta^2 t D_f + \sigma_B^2 [(\eta_1 - (\sigma_{01}/\sigma_{00}) \eta_0) t + D_f]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \frac{\sigma^2 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]^2}{\sigma_B^4 (\sigma_B^2 t + \sigma_\eta^2 t^2)}}} + \right. \\
& \left. \frac{\frac{a_2/a_1 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\eta^2 t^2}}}{\sqrt{1 + \frac{\sigma^2 [2\sigma_\eta^2 t + \sigma_B^2 (1 - (\sigma_{01}/\sigma_{00}) t)]^2}{\sigma_B^4 (\sigma_B^2 t + \sigma_\eta^2 t^2)}}} \right), \tag{2.20}
\end{aligned}$$

donde  $a_1 = \frac{2\sigma_\eta^2}{\sigma_B^4} - \frac{1}{2\sigma_{00}} - \frac{2\sigma_{01}}{\sigma_{00}\sigma_B^2}$ ,  $a_2 = \frac{2\sigma_\eta^2 D_f}{\sigma_B^4} - \frac{\eta_0}{2\sigma_{00}} - \frac{(\sigma_{01}/\sigma_{00})(D_f + \eta_0) - \eta_1}{\sigma_B^2}$ ,  $a_3 =$

$\frac{2\sigma_\eta^2 D_f^2}{\sigma_B^4} - \frac{\eta_0^2}{2\sigma_{00}} + \frac{2(\eta_1 - (\sigma_{01}/\sigma_{00})\eta_0) D_f}{\sigma_B^2}$ ,  $\sigma^2 = 1/(-2a_1)$ , y  $\sigma_\eta^2 = \sigma_{11} - \sigma_{01}^2/\sigma_{00}$ . El cálculo de  $I_2$  es válida solamente si  $a_1 < 0$ , lo cual equivale a que

$$\begin{aligned} a_1 &= \frac{2\sigma_\eta^2}{\sigma_B^4} - \frac{1}{2\sigma_{00}} - \frac{2\sigma_{01}}{\sigma_{00}\sigma_B^2} < 0 \\ \Leftrightarrow \frac{2(\sigma_{11} - \sigma_{01}^2/\sigma_{00})}{\sigma_B^4} - \frac{1}{2\sigma_{00}} - \frac{2\sigma_{01}}{\sigma_{00}\sigma_B^2} &< 0 \\ \Leftrightarrow 4(\sigma_{11} - \sigma_{01}^2/\sigma_{00})\sigma_{00} - \sigma_B^4 - 4\sigma_{01}\sigma_B^2 &< 0 \\ \Leftrightarrow 4(\sigma_{11}\sigma_{00} - \sigma_{01}^2) &< \sigma_B^2(\sigma_B^2 + 4\sigma_{01}) \\ \Leftrightarrow \text{Var}(\beta_1|\beta_0) &< \frac{\sigma_B^2(\sigma_B^2 + 4\sigma_{01})}{4\sigma_{00}}. \end{aligned}$$

Para obtener el tiempo medio de vida notemos que dado  $\beta_0$  y  $\beta_1$ ,  $T$  tiene distribución Gaussiana inversa,  $GI\left[\frac{D_f - \beta_0}{\beta_1}, \left(\frac{D_f - \beta_0}{\sigma_B}\right)^2\right]$ , y además dado  $\beta_1$ ,  $\beta_0$  tiene distribución  $N[\eta_0 + (\sigma_{01}/\sigma_{11})(\beta_1 - \eta_1), \sigma_{00} - \sigma_{01}^2/\sigma_{11}]$ , y entonces

$$\begin{aligned} E(T) &= E_{\beta_1} \left[ E_{\beta_0|\beta_1} \left[ E_{T|\beta_0, \beta_1} [T] \right] \right] = E_{\beta_1} \left[ E_{\beta_0|\beta_1} \left[ \frac{D_f - \beta_0}{\beta_1} \right] \right] \\ &= E_{\beta_1} \left[ \frac{D_f - \eta_0 - (\sigma_{01}/\sigma_{11})(\beta_1 - \eta_1)}{\beta_1} \right] = -\sigma_{01}/\sigma_{11} + E_{\beta_1} \left[ \frac{D_f - \eta_0 + (\sigma_{01}/\sigma_{11})\eta_1}{\beta_1} \right], \end{aligned}$$

y como  $\beta_1 \sim N(\eta_1, \sigma_{11})$ , siguiendo Peng y Tseng (2009) se tiene que

$$E(T) = -\sigma_{01}/\sigma_{11} + \sqrt{\frac{2}{\sigma_{11}}} (D_f - \eta_0 + (\sigma_{01}/\sigma_{11})\eta_1) D \left[ \frac{\eta_1}{\sqrt{2\sigma_{11}}} \right],$$

donde  $D(x) = \exp(-x^2) \int_0^x \exp(u^2) du$  es la integral de Dawson. Luego como  $D(x) \approx \frac{1}{2x}$  para  $x$  suficientemente grande (Peng y Tseng, 2009), si  $\eta_1 \gg \sqrt{\sigma_{11}}$ , entonces

$$\begin{aligned} E(T) &\approx -\sigma_{01}/\sigma_{11} + \frac{D_f - \eta_0 + (\sigma_{01}/\sigma_{11})\eta_1}{\eta_1} \\ &\approx \frac{D_f - \eta_0}{\eta_1}. \end{aligned}$$

Notemos que el valor esperado de  $T$  no depende de  $\sigma_B$ , mientras que el valor esperado aproximado de  $T$  no depende de  $\sigma_B$  ni de las varianzas y covarianzas de los efectos aleatorios, solamente de las medias de los efectos aleatorios, y como era de esperarse del umbral  $D_f$ .

## 2.B. Algoritmo EM para estimar los parámetros del modelo (2.1)

Como  $\mathbf{Y}_i(\mathbf{t}_i) \sim N_{m_i}(\mathbf{X}_i\boldsymbol{\eta}, \boldsymbol{\Sigma}_i)$ , donde  $\boldsymbol{\Sigma}_i = \mathbf{X}_i\boldsymbol{\Psi}\mathbf{X}_i' + \sigma_B^2\mathbf{Q}_i + \sigma_\epsilon^2\mathbf{I}_{m_i}$ ,  $\mathbf{X}_i = [\mathbf{1}_i, \mathbf{t}_i]$  y  $\boldsymbol{\theta} = (\sigma_{00}, \sigma_{01}, \sigma_{11}, \sigma_B, \sigma_\epsilon)'$ , entonces observemos que

$$\mathbf{Y}_i(\mathbf{t}_i) \stackrel{d}{=} \mathbf{X}_i\boldsymbol{\eta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i,$$

donde  $\mathbf{Z}_i = [\mathbf{X}_i \mathbf{Q}_i^{1/2}]$ ,  $\mathbf{b}_i = \begin{pmatrix} \mathbf{b}_{1i} \\ \mathbf{b}_{2i} \end{pmatrix} \sim N_{p+m_i} \left( \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \boldsymbol{\Psi}_i^* = \begin{bmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0}' & \sigma_B^2\mathbf{I}_{m_i} \end{bmatrix} \right)$ ,  $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}_{m_i}, \sigma_\epsilon^2\mathbf{I}_{m_i})$ ,  $\mathbf{b}_i \perp \boldsymbol{\epsilon}_i$  y  $\mathbf{Q}_i^{1/2}$  es la matriz raíz cuadrada de  $\mathbf{Q}_i$ . Luego siguiendo Pawitan (2001, Capítulo 12 página 358) (con  $\boldsymbol{\Sigma}_i = \sigma_\epsilon^2\mathbf{I}_i$  y  $\mathbf{D}_i = \boldsymbol{\Psi}_i^*$ ) se tiene que el paso E del algoritmo E-M para este modelo puede expresarse como

$$\begin{aligned} Q(\boldsymbol{\eta}, \boldsymbol{\theta} | \boldsymbol{\eta}^{(t)}, \boldsymbol{\theta}^{(t)}) &= - \left( \sum_{i=1}^n m_i + \frac{np}{2} \right) \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log |\boldsymbol{\Sigma}_i| - \frac{1}{2} \sum_{i=1}^n \log |\mathbf{D}_i| - \\ &\quad \frac{1}{2} \sum_{i=1}^n \hat{\boldsymbol{\epsilon}}_i' \boldsymbol{\Sigma}_i^{-1} \hat{\boldsymbol{\epsilon}}_i - \frac{1}{2} \sum_{i=1}^n \hat{\mathbf{b}}_i' \mathbf{D}_i^{-1} \hat{\mathbf{b}}_i - \frac{1}{2} \sum_{i=1}^n \text{tr} [(\mathbf{Z}_i' \boldsymbol{\Sigma}_i^{-1} \mathbf{Z}_i + \mathbf{D}_i^{-1}) \mathbf{F}_{i(t)}] \\ &= - \left( m + \frac{np}{2} \right) \log(2\pi) - N [\log(\sigma_\epsilon) + \log(\sigma_B)] - \frac{1}{2} n \log |\boldsymbol{\Psi}| - \\ &\quad \frac{\sigma_\epsilon^{-2}}{2} \sum_{i=1}^n \hat{\boldsymbol{\epsilon}}_i' \hat{\boldsymbol{\epsilon}}_i - \frac{1}{2} \sum_{i=1}^n \hat{\mathbf{b}}_{1i}' \boldsymbol{\Psi}^{-1} \hat{\mathbf{b}}_{1i} - \frac{\sigma_B^{-2}}{2} \sum_{i=1}^n \hat{\mathbf{b}}_{2i}' \hat{\mathbf{b}}_{2i} - \\ &\quad \frac{1}{2} \sum_{i=1}^n \text{tr} [(\sigma_\epsilon^{-2} \mathbf{Z}_i' \mathbf{Z}_i + \boldsymbol{\Psi}_i^{*-1}) \mathbf{F}_{i(t)}], \end{aligned} \quad (2.21)$$

donde  $m = \sum_{i=1}^n m_i$ ,  $\mathbf{F}_{i(t)} = \left( \sigma_{\epsilon(t)}^{-2} \mathbf{Z}_i' \mathbf{Z}_i + \boldsymbol{\Psi}_{i(t)}^{*-1} \right)^{-1}$ ,  $\hat{\boldsymbol{\epsilon}}_i = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\eta} - \mathbf{Z}_i\hat{\mathbf{b}}_i$ ,  $\hat{\mathbf{b}}_i = \sigma_{\epsilon(t)}^{-2} \mathbf{F}_{i(t)} \mathbf{Z}_i' (\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\eta}_{(t)})$ , y  $\boldsymbol{\theta} = (\boldsymbol{\theta}'_{\boldsymbol{\Psi}}, \sigma_B)'$ . Notemos que  $\hat{\boldsymbol{\epsilon}}_i$  depende de  $\boldsymbol{\eta}$ .

Para continuar con el paso M primero observemos que usando (2.29) se tiene que

$$\mathbf{F}_{i(t)} = \boldsymbol{\Psi}_{i(t)}^* - \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}_i' \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^*, \quad (2.22)$$

donde

$$\mathbf{V}_{i(t)} = (\mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}_i' + \sigma_{\epsilon(t)}^2 \mathbf{I}_i) = (\mathbf{X}_i \boldsymbol{\Psi}_{i(t)} \mathbf{X}_i' + \sigma_{B(t)}^2 \mathbf{Q}_i + \sigma_{\epsilon(t)}^2 \mathbf{I}_i), \quad (2.23)$$

y así

$$\begin{aligned} \text{tr}(\boldsymbol{\Psi}_i^{*-1} \mathbf{F}_{i(t)}) &= \text{tr}(\boldsymbol{\Psi}_i^{*-1} \boldsymbol{\Psi}_{i(t)}^*) - \text{tr} \left[ \boldsymbol{\Psi}_i^{*-1} \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}_i' \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \right] \\ &= \text{tr}(\boldsymbol{\Psi}_i^{-1} \boldsymbol{\Psi}_{i(t)}) + \sigma_B^{-2} \sigma_{B(t)}^2 m_i - \text{tr} \left[ \boldsymbol{\Psi}_i^{*-1} \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}_i' \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \right], \end{aligned}$$

pero como

$$\begin{aligned}
\text{tr} \left[ \Psi_i^{*-1} \Psi_{i(t)}^* \mathbf{Z}'_i \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \Psi_{i(t)}^* \right] &= \text{tr} \left\{ \Psi_i^{*-1} \begin{bmatrix} \Psi_{(t)} \mathbf{X}'_i \\ \sigma_{B(t)}^2 \mathbf{Q}_i^{1/2} \end{bmatrix} \mathbf{V}_{i(t)}^{-1} \begin{bmatrix} \mathbf{X}_i \Psi_{(t)} & \sigma_{B(t)}^2 \mathbf{Q}_i^{1/2} \end{bmatrix} \right\} \\
&= \text{tr} \left\{ \begin{bmatrix} \Psi^{-1} \Psi_{(t)} \mathbf{X}'_i \\ \sigma_B^{-2} \sigma_{B(t)}^2 \mathbf{Q}_i^{1/2} \end{bmatrix} \mathbf{V}_{i(t)}^{-1} \begin{bmatrix} \mathbf{X}_i \Psi_{(t)} & \sigma_{B(t)}^2 \mathbf{Q}_i^{1/2} \end{bmatrix} \right\} \\
&= \text{tr} \left\{ \mathbf{V}_{i(t)}^{-1} \left[ \mathbf{X}_i \Psi_{(t)} \Psi^{-1} \Psi_{(t)} \mathbf{X}'_i + \sigma_B^{-2} \sigma_{B(t)}^4 \mathbf{Q}_i \right] \right\} \\
&= \text{tr} \left\{ \Psi^{-1} \Psi_{(t)} \mathbf{X}'_i \mathbf{V}_{i(t)}^{-1} \mathbf{X}_i \Psi_{(t)} \right\} + \sigma_B^{-2} \sigma_{B(t)}^4 \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \right),
\end{aligned}$$

entonces

$$\begin{aligned}
\text{tr} \left( \Psi_i^{*-1} \mathbf{F}_{i(t)} \right) &= \sigma_B^{-2} \sigma_{B(t)}^2 \left[ m_i - \sigma_{B(t)}^2 \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \right) \right] + \\
&\quad \text{tr} \left\{ \Psi^{-1} \left[ \Psi_{(t)} - \Psi_{(t)} \mathbf{X}'_i \mathbf{V}_{i(t)}^{-1} \mathbf{X}_i \Psi_{(t)} \right] \right\},
\end{aligned}$$

y usando nuevamente la identidad (2.29) esto se reduce a

$$\text{tr} \left( \Psi_i^{*-1} \mathbf{F}_{i(t)} \right) = \sigma_B^{-2} \sigma_{B(t)}^2 \left[ m_i - \sigma_{B(t)}^2 \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \right) \right] + \text{tr} \left\{ \Psi^{-1} \left( \mathbf{X}'_i \mathbf{B}_{i(t)} \mathbf{X}_i + \Psi_{(t)}^{-1} \right)^{-1} \right\},$$

donde  $\mathbf{B}_{i(t)} = \left( \sigma_{B(t)}^2 \mathbf{Q}_i + \sigma_{\epsilon(t)}^2 \mathbf{I}_i \right)^{-1}$  (la inversión se puede hacer con el algoritmo iterativo que se encuentra en Miller (1981) o de forma directa por medio del Teorema 2 de Peng y Hsu (2012)).

Sustituyendo esto en (2.21) se obtiene que

$$\begin{aligned}
Q \left( \boldsymbol{\eta}, \boldsymbol{\theta} | \boldsymbol{\eta}^{(t)}, \boldsymbol{\theta}^{(t)} \right) &= - \left( m + \frac{np}{2} \right) \log(2\pi) - m \log(\sigma_\epsilon) - m \log(\sigma_B) - \\
&\quad \frac{\sigma_\epsilon^{-2}}{2} \left[ \sum_{i=1}^n \hat{\mathbf{e}}'_i \hat{\mathbf{e}}_i + \sum_{i=1}^n \text{tr} \left( \mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)} \right) \right] - \\
&\quad \frac{\sigma_B^{-2}}{2} \left\{ \sum_{i=1}^n \hat{\mathbf{b}}'_{2i} \hat{\mathbf{b}}_{2i} + \sigma_{B(t)}^2 \left[ m - \sigma_{B(t)}^2 \sum_{i=1}^n \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \right) \right] \right\} - \\
&\quad \frac{n}{2} \left\{ \log |\Psi| + \text{tr} \left\{ \Psi^{-1} \frac{1}{n} \sum_{i=1}^n \left( \left( \mathbf{X}'_i \mathbf{B}_{i(t)} \mathbf{X}_i + \Psi_{(t)}^{-1} \right)^{-1} + \hat{\mathbf{b}}_{1i} \hat{\mathbf{b}}'_{1i} \right) \right\} \right\},
\end{aligned} \tag{2.24}$$

y de aquí

$$\begin{aligned}\frac{\partial Q(\boldsymbol{\eta}, \boldsymbol{\theta} | \boldsymbol{\eta}^{(t)}, \boldsymbol{\theta}^{(t)})}{\partial \boldsymbol{\eta}} &= -\sigma_\epsilon^{-2} \sum_{i=1}^n \mathbf{X}'_i \mathbf{X}_i \boldsymbol{\eta} + \sigma_\epsilon^{-2} \sum_{i=1}^n \mathbf{X}'_i [\mathbf{y}_i - \mathbf{Z}_i \widehat{\mathbf{b}}_i], \\ \frac{\partial Q(\boldsymbol{\eta}, \boldsymbol{\theta} | \boldsymbol{\eta}^{(t)}, \boldsymbol{\theta}^{(t)})}{\partial \sigma_\epsilon} &= -\frac{m}{\sigma_\epsilon} + \frac{1}{\sigma_\epsilon^3} \left[ \sum_{i=1}^n \widehat{\mathbf{e}}'_i \widehat{\mathbf{e}}_i + \sum_{i=1}^n \text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)}) \right], \\ \frac{\partial Q(\boldsymbol{\eta}, \boldsymbol{\theta} | \boldsymbol{\eta}^{(t)}, \boldsymbol{\theta}^{(t)})}{\partial \sigma_B} &= -\frac{m}{\sigma_B} + \frac{1}{\sigma_B^3} \left\{ \sum_{i=1}^n \widehat{\mathbf{b}}'_{2i} \widehat{\mathbf{b}}_{2i} + \sigma_{B(t)}^2 \left[ m - \sigma_{B(t)}^2 \sum_{i=1}^n \text{tr}(\mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i) \right] \right\}.\end{aligned}$$

Luego como éstas expresiones no dependen de  $\boldsymbol{\Psi}$ , igualando a cero las derivadas anteriores y resolviendo las ecuaciones resultantes, los valores de  $\boldsymbol{\eta}$ ,  $\sigma_B$  y  $\sigma_\epsilon$  que maximizan (2.24) para cualquier valor de  $\boldsymbol{\Psi}$ , están dados por

$$\boldsymbol{\eta}_{(t+1)} = \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i [\mathbf{y}_i - \mathbf{Z}_i \widehat{\mathbf{b}}_i], \quad (2.25)$$

$$\sigma_{\epsilon(t+1)} = \sqrt{\frac{1}{m} \left[ \sum_{i=1}^n \widehat{\mathbf{e}}'_{i(t)} \widehat{\mathbf{e}}_{i(t)} + \sum_{i=1}^n \text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)}) \right]}, \quad (2.26)$$

$$\sigma_{B(t+1)} = \sqrt{\frac{1}{m} \left\{ \sum_{i=1}^n \widehat{\mathbf{b}}'_{2i} \widehat{\mathbf{b}}_{2i} + \sigma_{B(t)}^2 \left[ m - \sigma_{B(t)}^2 \sum_{i=1}^n \text{tr}(\mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i) \right] \right\}}, \quad (2.27)$$

donde ahora  $\widehat{\mathbf{e}}_{i(t)} = \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t+1)} - \mathbf{Z}_i \widehat{\mathbf{b}}_i$ . Para finalizar con el paso M, asumiendo que  $\frac{1}{n} \sum_{i=1}^n (\mathbf{X}'_i \mathbf{B}_{i(t)} \mathbf{X}_i + \boldsymbol{\Psi}_{(t)}^{-1})^{-1} + \frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{b}}_{1i} \widehat{\mathbf{b}}'_{1i}$  es una matriz positiva definida, por resultado Seber (2004, A 7.1 página 523) se tiene que para cualesquiera valores dados de los restantes parámetros, (2.24) como función de  $\boldsymbol{\Psi}$  alcanza su máximo en

$$\boldsymbol{\Psi}_{(t+1)} = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}'_i \mathbf{B}_{i(t)} \mathbf{X}_i + \boldsymbol{\Psi}_{(t)}^{-1})^{-1} + \frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{b}}_{1i} \widehat{\mathbf{b}}'_{1i}. \quad (2.28)$$

Así, en resumen la estimación de los parámetros de modelo por medio del algoritmo EM consiste en iterar hasta convergencia las ecuaciones de (2.25-2.28).

## 2.B.1. Cálculos para la implementación

1. Cálculo de  $\text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)})$ :

Notemos que

$$\begin{aligned}\text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)}) &= \sigma_{\epsilon(t)}^2 \text{tr} \left[ \left( \sigma_{\epsilon(t)}^{-2} \mathbf{Z}'_i \mathbf{Z}_i + \boldsymbol{\Psi}_{i(t)}^{*-1} - \boldsymbol{\Psi}_{i(t)}^{*-1} \right) \mathbf{F}_{i(t)} \right] \\ &= (m_i + p) \sigma_{\epsilon(t)}^2 - \sigma_{\epsilon(t)}^2 \text{tr} \left( \boldsymbol{\Psi}_{i(t)}^{*-1} \mathbf{F}_{i(t)} \right),\end{aligned}$$

y por (2.22) y por definición de  $V_{i(t)}$  (2.23) esto se reduce a

$$\begin{aligned}\text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)}) &= (m_i + p) \sigma_{\epsilon(t)}^2 - \sigma_{\epsilon(t)}^2 \text{tr} \left( \mathbf{I}_{m_i+p} - \mathbf{Z}'_i \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \right) \\ &= \sigma_{\epsilon(t)}^2 \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i \right) \\ &= \sigma_{\epsilon(t)}^2 \left\{ \text{tr} \left[ \mathbf{V}_{i(t)}^{-1} \left( \mathbf{X}_i \boldsymbol{\Psi}_{i(t)} \mathbf{X}'_i + \sigma_{B(t)}^2 \mathbf{Q}_i \right) \right] \right\} \\ &= \sigma_{\epsilon(t)}^2 \left[ m_i - \sigma_{\epsilon(t)}^2 \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \right) \right],\end{aligned}$$

y así

$$\sum_{i=1}^n \text{tr}(\mathbf{Z}'_i \mathbf{Z}_i \mathbf{F}_{i(t)}) = \sigma_{\epsilon(t)}^2 \left[ m - \sigma_{\epsilon(t)}^2 \sum_{i=1}^n \text{tr} \left( \mathbf{V}_{i(t)}^{-1} \right) \right].$$

2. Cálculo de “las estimaciones” de los efectos aleatorios:

Primero observemos que usando (2.22)

$$\begin{aligned}\widehat{\mathbf{b}}_i &= \sigma_{\epsilon(t)}^{-2} \mathbf{F}_{i(t)} \mathbf{Z}'_i (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}) \\ &= \sigma_{\epsilon(t)}^{-2} \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i \left[ \mathbf{I}_i - \mathbf{V}_{i(t)}^{-1} \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i \right] (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}),\end{aligned}$$

pero como  $\mathbf{Z}_i \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i = \mathbf{Z}_i \boldsymbol{\Psi}_{i(t)} \mathbf{Z}'_i + \sigma_{B(t)}^2 \mathbf{Q}_i = \mathbf{V}_{i(t)} - \sigma_{\epsilon(t)}^2 \mathbf{I}_i$ ,

$$\begin{aligned}\widehat{\mathbf{b}}_i &= \boldsymbol{\Psi}_{i(t)}^* \mathbf{Z}'_i \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}) \\ &= \begin{bmatrix} \boldsymbol{\Psi}_{i(t)} \mathbf{X}'_i \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}) \\ \sigma_{B(t)}^2 \mathbf{Q}_i^{1/2} \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}) \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{b}}_{1i} \\ \widehat{\mathbf{b}}_{2i} \end{bmatrix}.\end{aligned}$$

Así,

$$\begin{aligned}\widehat{\mathbf{b}}'_{1i} \widehat{\mathbf{b}}_{1i} &= (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)})' \mathbf{V}_{i(t)}^{-1} \mathbf{X}_i \boldsymbol{\Psi}_{i(t)}^2 \mathbf{X}'_i \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}), \\ \widehat{\mathbf{b}}'_{2i} \widehat{\mathbf{b}}_{2i} &= \sigma_{B(t)}^4 (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)})' \mathbf{V}_{i(t)}^{-1} \mathbf{Q}_i \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}), \\ \widehat{\mathbf{b}}'_{1i} \widehat{\mathbf{b}}'_{1i} &= \boldsymbol{\Psi}_{i(t)} \mathbf{X}'_i \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)}) (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{i(t)})' \mathbf{V}_{i(t)}^{-1} \mathbf{X}_i \boldsymbol{\Psi}_{i(t)}.\end{aligned}$$

Para calcular el producto punto de los residuos primero observemos que

$$\begin{aligned}
\widehat{\mathbf{e}}_{i(t)} &= \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)} - \mathbf{Z}_i \widehat{\mathbf{b}}_i \\
&= \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)} - [\mathbf{X}_i \boldsymbol{\Psi}_{(t)} \mathbf{X}_i' + \sigma_{B(t)}^2 \mathbf{Q}_i] \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}) \\
&= -\mathbf{X}_i \boldsymbol{\eta}_{(t)} + \mathbf{X}_i \boldsymbol{\eta}_{(t)} + \sigma_{\epsilon(t)}^2 \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}) \\
&= \sigma_{\epsilon(t)}^2 \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}),
\end{aligned}$$

y así

$$\widetilde{\mathbf{e}}_{i(t)}' \widehat{\mathbf{e}}_{i(t)} = \sigma_{\epsilon(t)}^4 (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)})' \mathbf{V}_{i(t)}^{-2} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}).$$

3. Usando (2.29) se obtiene una forma computacional más conveniente para  $\mathbf{V}_i^{-1}$ :

$$\mathbf{V}_{i(t)}^{-1} = \mathbf{B}_{i(t)} - \mathbf{B}_{i(t)} \mathbf{X}_i \left( \mathbf{X}_i' \mathbf{B}_{i(t)} \mathbf{X}_i + \boldsymbol{\Psi}_{i(t)}^{-1} \right)^{-1} \mathbf{X}_i' \mathbf{B}_{i(t)},$$

donde  $\mathbf{B}_{i(t)} = \left( \sigma_{B(t)}^2 \mathbf{Q}_i + \sigma_{\epsilon(t)}^2 \mathbf{I}_i \right)^{-1}$  puede obtenerse de forma analítica o iterativamente usando Teorema 2 de Peng y Hsu (2012) ó Miller (1981), respectivamente.

4. Para calcular  $\boldsymbol{\eta}_{(t+1)}$ :

$$\begin{aligned}
\mathbf{Z}_i \mathbf{b}_i &= [\mathbf{X}_i \boldsymbol{\Psi}_{(t)} \mathbf{X}_i' + \sigma_{B(t)}^2 \mathbf{Q}_i] \mathbf{V}_{i(t)}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}) \\
&= \left( \mathbf{I}_i - \sigma_{\epsilon(t)}^2 \mathbf{V}_{i(t)}^{-1} \right) (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\eta}_{(t)}).
\end{aligned}$$

Notemos que en ningún lugar es necesario calcular  $\mathbf{Q}_i^{1/2}$ .

**Lema 2.B.1.** (Zacks, 1971, página 133) Si  $Z \sim N(0, 1)$  y  $a, b \in \mathbb{R}$ , entonces  $E[\Phi(a + bZ)] = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right)$ .

## 2.B.2. Resultados de matrices

$$(\mathbf{A}\mathbf{B}\mathbf{A}' + \mathbf{C})^{-1} = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{C}^{-1}\mathbf{A} + \mathbf{B}^{-1})^{-1}\mathbf{A}'\mathbf{C}^{-1}. \quad (2.29)$$



- Resultado Seber (2004, A 7.1, página 523): Si  $f$  es una función matricial dada por

$$f(\Sigma) = \log |\Sigma| + \text{tr}(\mathbf{A}\Sigma),$$

y  $\mathbf{A} > 0$ , es decir,  $\mathbf{A}$  es una matriz positiva definida, entonces sujeto a que  $\Sigma > 0$ ,  $f$  de forma única alcanza su mínimo en  $\Sigma = \mathbf{A}$ .

- Resultados Miller (1981):
  - Lema: Si  $\mathbf{A}$  y  $(\mathbf{A} + \mathbf{B})$  son matrices no singulares, donde  $\mathbf{B}$  es una matriz de rango 1, entonces

$$(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \frac{1}{1 + g} \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1},$$

donde  $g = \text{tr}(\mathbf{B}\mathbf{A}^{-1}) \neq -1$ .

- Teorema: Si  $\mathbf{A}$  y  $\mathbf{A} + \mathbf{B}$  son matrices no singulares y además  $\mathbf{B}$  es una matriz de rango  $r$  talque  $\mathbf{B} = \sum_{i=1}^r \mathbf{B}_i$ , donde  $\mathbf{B}_i, i = 1, \dots, r$ , son matrices de rango uno, entonces definiendo  $\mathbf{C}_1 = \mathbf{A}$  y  $\mathbf{C}_k = \mathbf{A} + \sum_{i=1}^{k-1} \mathbf{B}_i$  se tiene que para  $k = 2, \dots, r$ ,

$$\mathbf{C}_k^{-1} = \mathbf{C}_{k-1}^{-1} - v_{k-1} \mathbf{C}_{k-1}^{-1} \mathbf{B}_{k-1} \mathbf{C}_{k-1}^{-1},$$

donde  $v_{k-1} = \frac{1}{1 + \text{tr}(\mathbf{C}_{k-1}^{-1} \mathbf{B}_{k-1})}$ .

## 2.C. Matriz de varianzas y covarianzas del estimador de máxima verosimilitud del intercepto y pendiente del Browniano

Como el estimador de máxima verosimilitud del parámetro  $\boldsymbol{\eta}$  en el modelo  $Y(t) = \eta_0 + \eta_1 t + \sigma_B B(t)$  está dado por

$$\hat{\boldsymbol{\eta}} = \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{y}_i \right),$$

entonces

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\eta}}) &= \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \text{Var}(\mathbf{y}_i) \mathbf{Q}_i^{-1} \mathbf{X}_i \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \\
&= \sigma_B^2 \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{Q}_i \mathbf{Q}_i^{-1} \mathbf{X}_i \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \\
&= \sigma_B^2 \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1}.
\end{aligned}$$

Pero como en Peng y Tseng (2009) se tiene que

$$\mathbf{Q}_i^{-1} = \begin{bmatrix} a_{i1} + a_{i2} & -a_{i2} & 0 & \dots & 0 \\ -a_{i2} & a_{i2} + a_{i3} & -a_{i3} & \ddots & \vdots \\ 0 & -a_{i3} & \ddots & \ddots & 0 \\ \vdots & \ddots & -a_{i,m_i-1} & a_{i,m_i-1} + a_{im_i} & -a_{im_i} \\ 0 & \dots & 0 & -a_{im_i} & a_{im_i} \end{bmatrix},$$

donde  $a_{ij} = 1/(t_{ij} - t_{i,j-1})$ ,  $t_{i,0} = 0$ ,  $j = 1, \dots, m_i$ , entonces

$$\begin{aligned}
\mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i &= \begin{bmatrix} a_{i1} & a_{i1} t_{i1} \\ a_{i1} t_{i1} & t_{i,m_i} \end{bmatrix} \\
&= \begin{bmatrix} 1/t_{i1} & 1 \\ 1 & t_{i,m_i} \end{bmatrix},
\end{aligned}$$

ya que  $a_{i1} = 1/t_{i1}$ . Así, la varianza de  $\hat{\boldsymbol{\eta}}$  puede calcularse como

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\eta}}) &= \sigma_B^2 \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{Q}_i^{-1} \mathbf{X}_i \right)^{-1} \\
&= \sigma_B^2 \begin{bmatrix} \sum_{i=1}^n 1/t_{i1} & n \\ n & \sum_{i=1}^n t_{i,m_i} \end{bmatrix}^{-1} \\
&= \frac{\sigma_B^2}{(\sum_{i=1}^n 1/t_{i1})(\sum_{i=1}^n t_{i,m_i}) - n^2} \begin{bmatrix} \sum_{i=1}^n t_{i,m_i} & -n \\ -n & \sum_{i=1}^n 1/t_{i1} \end{bmatrix}.
\end{aligned}$$

De aquí la varianza de  $\hat{\boldsymbol{\eta}}_0$  es  $\text{Var}(\hat{\boldsymbol{\eta}}_0) = \frac{\sigma_B^2 \sum_{i=1}^n t_{i,m_i}}{(\sum_{i=1}^n 1/t_{i1})(\sum_{i=1}^n t_{i,m_i}) - n^2}$ , y cuando el primer y último tiempo de inspección son los mismos en todas las unidades, es decir, cuando  $t_{1i} = t_1$  y  $t_{i,m_i} = t_m$ ,  $\forall i = 1, \dots, n$ , ésta se reduce a  $\text{Var}(\hat{\boldsymbol{\eta}}_0) = \frac{\sigma_B^2 n t_m}{n^2 (t_m/t_1 - 1)} = \frac{\sigma_B^2 t_m}{n (t_m/t_1 - 1)}$ .

# Capítulo 3

## Mezcla de procesos Gaussianos inversos

### 3.1. Introducción

La mezcla de modelos provee una forma de modelar heterogeneidad presente en la población de estudio. En particular la mezcla finita de distribuciones surge en una gran cantidad de aplicaciones que ayudan en la descripción de sistemas complejos en astronomía, biología, genética, medicina, psiquiatría, economía, ingeniería y mercadotecnia (Everitt y Hand, 1981; Titterington et al., 1985; McLachlan y Peel, 2000) por mencionar algunas. Específicamente en el análisis tradicional de datos de vida, el uso de éste tipo de modelos ha sido popular para modelar datos de vida heterogéneos, dónde se han considerando desde distribuciones de la misma familia en las componentes de la mezcla (por ejemplo exponenciales) hasta diferentes distribuciones como componentes de ésta (Mendenhall y Hader (1958); Kleye y Dahiya (1975); Jiang y Kececioglu (1992); Meecker y Escobar (1998); Scallan (1999); Bucar et al. (2004); Klutke et al. (2003); Marín et al. (2005); Hanson (2006); Noor y Aslam (2013)).

Los procesos Wiener, Gamma y Gaussiano inverso (PGI) son los tres procesos estocásticos más usados en el análisis de degradación. Estos procesos han sido ampliamente estudiados. Los primeros dos han sido estudiados por ejemplo en Singpurwalla (1995), Whitmore (1995), Whitmore y Schenkelberg (1997), Lawless y Crowder (2004), Peng y Tseng (2009), Wang (2010), Tsai et al. (2011, 2012) y Ye et al. (2013, 2014).

Respecto al tercer modelo, estudios recientes han reportado mayor flexibilidad en algunos casos de aplicación respecto a los dos anteriores (Wang y Xu, 2010; Peng, 2015). Sin embargo, considerando el proceso Gaussiano inverso estándar, en muchas situaciones con datos de degradación sucede que el PGI no es suficiente para capturar la variabilidad entre las unidades ya que tanto  $\mu$  como  $\lambda$  varían de unidad a unidad. Ésta heterogeneidad puede

deberse a la variabilidad de la materia prima, factores ambientales no controlables (temperatura, presión), horas de manufactura, entre otras. Para tal propósito Wang y Xu (2010) proponen una clase de procesos Gaussianos inversos, y solamente consideran que el parámetro  $\lambda$  varía de unidad a unidad de acuerdo a una distribución Gamma. Como extensión del modelo anterior, Peng (2015) propone un modelo llamado proceso Gaussiano inverso Normal-Gamma, donde asume que conjuntamente  $1/\mu$  y  $\lambda$  varían de unidad a unidad de acuerdo a la distribución Normal-Gamma (Whitmore, 1986). Para cuando la inserción de efectos aleatorios no captura parte de la heterogeneidad presente, recientemente Yuan y Ji (2015) proponen la formulación general de degradación dada por Meeker y Escobar (1998) pero donde los efectos aleatorios provienen no de una distribución Gaussiana, sino de una mezcla finita de éstas.

En vez de usar una mezcla continua de procesos Gaussianos inversos, en ésta Capítulo proponemos una mezcla finita de procesos Gaussianos inversos para modelar la heterogeneidad de las trayectorias de degradación. Específicamente, el desarrollo es abordado asumiendo que el proceso de generación de las trayectorias proviene de cualquiera de dos procesos Gaussianos inversos con sus propios parámetros. La estimación del modelo es abordado desde una perspectiva clásica y los estimadores de máxima verosimilitud son obtenidos iterativamente empleando el algoritmo EM.

El contenido del escrito está desplegado de la siguiente manera: En la Sección 3.2 se describe el proceso Gaussiano inverso estándar, en la Sección 3.3 se describe el proceso de degradación de interés como una mezcla de dos procesos Gaussianos inversos (MPGI) estándar, la función de distribución y el tiempo medio del tiempo de cruce que tal modelo induce. En la Sección 3.4 se describe el algoritmo EM para la estimación de los parámetros correspondientes y en seguida se presenta una ilustración con datos reales donde se comparan con algunas extensiones del proceso Gaussiano inverso mencionadas anteriormente. Por último en la Sección 3.5 se presentan las conclusiones.

## 3.2. Proceso Gaussiano inverso

Un proceso Gaussiano inverso  $\{Y(t); t \geq 0\}$  tiene las siguientes propiedades: (1)  $Y(0) = 0$  con probabilidad 1; (2)  $Y(t)$  tiene incrementos independientes: para todo par de intervalos disjuntos  $(t_1, t_2)$  y  $(t_3, t_4)$  con  $t_1 < t_2 < t_3 < t_4$ , las variables aleatorias  $Y(t_2) - Y(t_1)$  y  $Y(t_4) - Y(t_3)$  son independientes; (3) para  $0 \leq s < t$ , cada incremento  $Y(t) - Y(s)$  tiene distribución Gaussiana inversa, es decir,  $Y(t) - Y(s) \sim IG(\mu\Delta\Lambda, \lambda\Delta\Lambda^2)$ , donde  $\Delta\Lambda = \Lambda(t) - \Lambda(s)$ ,  $\Lambda(t)$  es una función monótona creciente del tiempo  $t$  conocida tal que  $\Lambda(0) = 0$ , y la densidad de una distribución Gaussiana inversa con parámetros  $\mu$  y  $\lambda$ ,

$IG(\mu, \lambda)$  está dada por

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right), \quad x > 0.$$

### 3.3. Mezcla de procesos Gaussianos inversos

Consideremos que las trayectorias de degradación provienen de una mezcla de dos procesos Gaussianos inversos, es decir,

$$Y(t) \sim \begin{cases} IG P_1(\mu_1 \Lambda(t), \lambda_1 \Lambda(t)^2) & \text{si } Z = 1 \\ IG P_2(\mu_2 \Lambda(t), \lambda_2 \Lambda(t)^2) & \text{si } Z = 0, \end{cases}$$

donde  $Z$  es una variable aleatoria Bernoulli de parámetro  $p$ .

Como las unidades fallan cuando el valor de degradación alcanza el umbral  $D_f$  y las trayectorias son monótonas, entonces por probabilidad total la función de distribución de los tiempos a la falla que genera el proceso anterior es

$$\begin{aligned} F_T(t) &= P(T \leq t | Z = 1) P(Z = 1) + P(T \leq t | Z = 0) P(Z = 0) \\ &= p P(Y_1(t) \geq D_f) + (1 - p) P(Y_2(t) \geq D_f) \\ &= \sum_{k=1}^2 p_k \left\{ \Phi \left[ \sqrt{\frac{\lambda_k \Lambda(t)^2}{D_f}} \left( 1 - \frac{D_f}{\mu_k \Lambda(t)} \right) \right] - \right. \\ &\quad \left. \exp\left(\frac{2\lambda_k}{\mu_k} \Lambda(t)\right) \Phi \left[ -\sqrt{\frac{\lambda_k \Lambda(t)^2}{D_f}} \left( \frac{D_f}{\mu_k \Lambda(t)} + 1 \right) \right] \right\}, \end{aligned}$$

donde  $p_1 = p$ ,  $p_2 = 1 - p$  y  $\Phi(\cdot)$  es la función de distribución de una variable aleatoria normal estándar.

El tiempo medio de cruce bajo el proceso anterior es

$$\begin{aligned} E(T) &= \int_0^\infty [1 - F_T(t)] dt \\ &= \sum_{k=1}^2 p_k \int_0^\infty P(Y_k(t) < D_f) dt, \end{aligned}$$

para el cual se requieren calcular de forma numérica las integrales involucradas, pero en el caso de trayectorias lineales,  $\Lambda(t) = t$ , éste se reduce a (Peng, 2015)

$$E(T) = \sum_{k=1}^2 p_k \left[ \left( \frac{D_f}{\mu_k} + \frac{\mu_k}{\lambda_k} \right) \Phi \left( \frac{\sqrt{\lambda_k D_f}}{\mu_k} \right) + \sqrt{\frac{D_f}{\lambda_k}} \phi \left( \sqrt{\frac{\lambda_k D_f}{\mu_k}} \right) - \frac{\mu_k}{2\lambda_k} \right]. \quad (3.1)$$

## 3.4. Estimación

### 3.4.1. Máxima verosimilitud

Para la estimación de los parámetros utilizamos el método de máxima verosimilitud. Supongamos que  $n$  unidades son puestas a experimentación y cada una es inspeccionada en los tiempos  $t_{i1}, \dots, t_{im_i}$  para  $i = 1, \dots, n$ . Sea  $\mathbf{y}_i = [y_{i1}, \dots, y_{im_i}]'$  el vector de valores de degradación registrados para la unidad  $i$ ,  $i = 1, \dots, n$ , y  $\Delta\mathbf{y}_i = [\Delta y_{i1}, \dots, \Delta y_{im_i}]'$  el vector de incrementos correspondientes ( $\Delta y_{ij} = y_{ij} - y_{i,j-1}$ ). Notemos que la log-verosimilitud de los parámetros bajo el modelo anterior está dada por

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \log \left[ \prod_{i=1}^n \left( p f_{\Delta\mathbf{Y}_i}^{(1)}(\Delta\mathbf{y}_i) + (1-p) f_{\Delta\mathbf{Y}_i}^{(2)}(\Delta\mathbf{y}_i) \right) \right], \\ &= \sum_{i=1}^n \log \left( p f_{\Delta\mathbf{Y}_i}^{(1)}(\Delta\mathbf{y}_i) + (1-p) f_{\Delta\mathbf{Y}_i}^{(2)}(\Delta\mathbf{y}_i) \right), \end{aligned} \quad (3.2)$$

donde  $f_{\Delta\mathbf{Y}}^{(1)}(\cdot)$  y  $f_{\Delta\mathbf{Y}}^{(2)}(\cdot)$  denotan las densidades conjuntas de los incrementos bajo los procesos Gaussianos inversos 1 y 2 respectivamente.

El estimador de máxima verosimilitud de  $\boldsymbol{\theta}$  no tiene forma cerrada por lo que para encontrarlo se requieren métodos numéricos para maximizar la log-verosimilitud. Por la naturaleza (mezcla finita) del modelo, una alternativa más simple para estimar  $\boldsymbol{\theta}$  es el algoritmo EM (Dempster et al., 1977; Wu, 1983; McLachlan y Krishnan, 2008). Antes es necesario mencionar que a diferencia de otros modelos de mezcla, por ejemplo mezcla de distribuciones normales (Quandt y Ramsey, 1978), cuando cada trayectoria es medida en al menos dos tiempos de inspección diferentes, puede garantizarse que la log-verosimilitud de la mezcla definida en (3.2) es acotada (ver Apéndice 3.B).

Para la implementación del algoritmo EM, notemos que el modelo anterior puede representarse de forma jerárquica usando una variable latente  $Z$  con distribución Bernoulli de parámetro  $p$ :  $Z = 1$  si la trayectoria de degradación proviene del primer proceso y  $Z = 0$  si proviene del segundo. Es decir, si  $Z_1, \dots, Z_n$  son los valores de las variables latentes correspondientes a cada una de las  $n$  unidades en prueba, entonces  $Y_i(t) | Z_i \sim IGP(\mu_{Z_i^*} \Lambda(t), \lambda_{Z_i^*} \Lambda(t)^2)$ , donde  $Z_i^* = 2 - I_{\{Z_i=1\}}$  y  $Z_1, \dots, Z_n \stackrel{iid}{\sim} Ber(p)$ .

Notemos que la distribución conjunta de los incrementos  $\Delta\mathbf{Y}_i$  y la variable latente  $Z_i$

está dada por

$$\begin{aligned}
f_{\Delta \mathbf{Y}_i, Z_i}(\Delta \mathbf{y}_i, z_i) &= f_{\Delta \mathbf{Y}_i | Z_i}(\Delta \mathbf{y}_i | z) f_Z(z) \\
&= f_{\Delta \mathbf{Y}_i | Z_i}(\Delta \mathbf{y}_i | z_i) p^{z_i} (1-p)^{1-z_i} \\
&= \left[ f_{\Delta \mathbf{Y}_i}^{(1)}(\Delta \mathbf{y}_i) p \right]^{z_i} \left[ f_{\Delta \mathbf{Y}_i}^{(2)}(\Delta \mathbf{y}_i) (1-p) \right]^{1-z_i}. \tag{3.3}
\end{aligned}$$

Sean  $\Delta \mathbf{y} = [\mathbf{y}'_1, \dots, \mathbf{y}'_n]'$  y  $\mathbf{Z} = [Z_1, \dots, Z_n]'$ , entonces la log-verosimilitud de los parámetros basada en los datos completos ( $[\Delta \mathbf{y}', \mathbf{Z}']'$ ) está dada por

$$\begin{aligned}
\ell_c(\mu_1, \lambda_1, \mu_2, \lambda_2, p | \Delta \mathbf{y}, \mathbf{z}) &= \sum_{i=1}^n Z_i \log \left( f_{\Delta \mathbf{Y}_i}^{(1)}(\Delta \mathbf{y}_i | \mu_1, \lambda_1) \right) + \\
&\quad \sum_{i=1}^n (1 - Z_i) \log \left( f_{\Delta \mathbf{Y}_i}^{(2)}(\Delta \mathbf{y}_i | \mu_2, \lambda_2) \right) + \\
&\quad \sum_{i=1}^n Z_i \log(p) + \sum_{i=1}^n (1 - Z_i) \log(1 - p) \\
&= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \log \left[ \frac{(\Delta t_{ij})^2}{2\pi (\Delta y_{ij})^3} \right] + \frac{1}{2} \sum_{i=1}^n m_i Z_i \log(\lambda_1) - \\
&\quad \frac{\lambda_1}{2} \sum_{i=1}^n Z_i \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_1)^2}{\Delta y_{ij} \mu_1^2} + \frac{1}{2} \sum_{i=1}^n m_i (1 - Z_i) \log(\lambda_2) - \\
&\quad \frac{\lambda_2}{2} \sum_{i=1}^n (1 - Z_i) \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_2)^2}{\Delta y_{ij} \mu_2^2} + \\
&\quad \sum_{i=1}^n Z_i \log(p) + \sum_{i=1}^n (1 - Z_i) \log(1 - p). \tag{3.4}
\end{aligned}$$

De (3.3) podemos ver que las  $Z_i$ 's son condicionalmente independientes (dado  $\Delta \mathbf{y}$ ), cada una con distribución Bernoulli con parámetro  $p_i$  dado por

$$p_i = P(Z_i = 1 | \Delta \mathbf{y}, \boldsymbol{\theta}) = \frac{p f_1(\Delta \mathbf{y} | \mu_1, \lambda_1)}{p f_1(\Delta \mathbf{y} | \mu_1, \lambda_1) + (1-p) f_2(\Delta \mathbf{y} | \mu_2, \lambda_2)},$$

donde  $\boldsymbol{\theta} = [\mu_1, \lambda_1, \mu_2, \lambda_2, p]'$  es el vector de parámetros a estimar. Luego obteniendo el valor esperado de (3.4) con respecto a la distribución de  $\mathbf{Z}$  dado  $\Delta \mathbf{y}$  y  $\boldsymbol{\theta}^{(t)} =$

$[\mu_1^{(t)}, \lambda_1^{(t)}, \mu_2^{(t)}, \lambda_2^{(t)}, p^{(t)}]'$ , el paso E del algoritmo EM está dado por

$$\begin{aligned}
Q(\mu_1, \lambda_1, \mu_2, \lambda_2, p | \boldsymbol{\theta}^{(t)}) &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \log \left[ \frac{(\Delta t_{ij})^2}{2\pi (\Delta y_{ij})^3} \right] + \frac{1}{2} \sum_{i=1}^n m_i E_i \log(\lambda_1) - \\
&\quad \frac{\lambda_1}{2} \sum_{i=1}^n E_i \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_1)^2}{\Delta y_{ij} \mu_1^2} + \frac{1}{2} \sum_{i=1}^n m_i (1 - E_i) \log(\lambda_2) - \\
&\quad \frac{\lambda_2}{2} \sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_2)^2}{\Delta y_{ij} \mu_2^2} + \\
&\quad \sum_{i=1}^n E_i \log(p) + \sum_{i=1}^n (1 - E_i) \log(1 - p),
\end{aligned}$$

donde

$$\begin{aligned}
E_i &= E(Z_i | \Delta \mathbf{y}_i, \boldsymbol{\theta}^{(t)}) = P(Z_i = 1 | \Delta \mathbf{y}_i, \boldsymbol{\theta}^{(t)}) \\
&= \frac{p^{(t)} f^{(1)}(\Delta \mathbf{y}_i | \mu_1^{(t)}, \lambda_1^{(t)})}{p^{(t)} f^{(1)}(\Delta \mathbf{y}_i | \mu_1^{(t)}, \lambda_1^{(t)}) + (1 - p^{(t)}) f^{(2)}(\Delta \mathbf{y}_i | \mu_2^{(t)}, \lambda_2^{(t)})},
\end{aligned}$$

$i = 1, \dots, n$ .

Las derivadas parciales de  $Q$  con respecto a  $\mu_1, \lambda_1, \mu_2, \lambda_2$  y  $p$  están dadas por

$$\frac{\partial Q}{\partial \mu_1} = \frac{\lambda_1}{\mu_1^3} \sum_{i=1}^n E_i \sum_{j=1}^{m_i} \Delta y_{ij} - \frac{\lambda_1}{\mu_1^2} \sum_{i=1}^n E_i \sum_{j=1}^{m_i} \Delta t_{ij} \quad (3.5)$$

$$\frac{\partial Q}{\partial \lambda_1} = \frac{1}{2\lambda_1} \sum_{i=1}^n m_i E_i - \frac{1}{2} \sum_{i=1}^n E_i \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_1)^2}{\Delta y_{ij} \mu_1^2} \quad (3.6)$$

$$\frac{\partial Q}{\partial \mu_2} = \frac{\lambda_2}{\mu_2^3} \sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \Delta y_{ij} - \frac{\lambda_2}{\mu_2^2} \sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \Delta t_{ij} \quad (3.7)$$

$$\frac{\partial Q}{\partial \lambda_2} = \frac{1}{2\lambda_2} \sum_{i=1}^n m_i (1 - E_i) - \frac{1}{2} \sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \mu_2)^2}{\Delta y_{ij} \mu_2^2} \quad (3.8)$$

$$\frac{\partial Q}{\partial p} = \frac{\sum_{i=1}^n E_i}{p} - \frac{\sum_{i=1}^n (1 - E_i)}{1 - p}. \quad (3.9)$$

Igualando a cero éstas derivadas, de las correspondientes a (3.9), (3.5) y (3.7) se tiene



que

$$p^{(t+1)} = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( Z_i | \Delta \mathbf{y}_i, \boldsymbol{\theta}^{(t)} \right) \quad (3.10)$$

$$\mu_1^{(t+1)} = \frac{\sum_{i=1}^n E_i \sum_{j=1}^{m_i} \Delta y_{ij}}{\sum_{i=1}^n E_i \sum_{j=1}^{m_i} \Delta t_{ij}} \quad (3.11)$$

$$\mu_2^{(t+1)} = \frac{\sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \Delta y_{ij}}{\sum_{i=1}^n (1 - E_i) \sum_{j=1}^{m_i} \Delta t_{ij}}, \quad (3.12)$$

y sustituyendo (3.11) y (3.12) en las derivadas (3.6) y (3.8) igualadas a cero se obtiene que

$$\lambda_1^{(t+1)} = \frac{\sum_{i=1}^n m_i E_i}{\sum_{i=1}^n \sum_{j=1}^{m_i} \left( \frac{(\Delta t_{ij})^2}{\Delta y_{ij}} - \frac{E_i \sum_{j=1}^{m_i} \Delta t_{ij}}{\mu_1^{(t+1)}} \right)} \quad (3.13)$$

$$\lambda_2^{(t+1)} = \frac{\sum_{i=1}^n m_i (1 - E_i)}{\sum_{i=1}^n \sum_{j=1}^{m_i} \left( \frac{(\Delta t_{ij})^2}{\Delta y_{ij}} - \frac{(1 - E_i) \sum_{j=1}^{m_i} \Delta t_{ij}}{\mu_2^{(t+1)}} \right)}. \quad (3.14)$$

Así, estimaciones de los parámetros se pueden obtener iterando las ecuaciones (3.10)-(3.14) hasta convergencia.

Una vez obtenidos los estimadores de máxima verosimilitud de los parámetros, estimaciones de máxima verosimilitud de otras cantidades de interés pueden obtenerse sustituyendo  $\hat{\boldsymbol{\theta}}$  en sus correspondientes expresiones. Por ejemplo, sustituyendo  $\hat{\boldsymbol{\theta}}$  en (3.1) se obtiene el estimador de máxima verosimilitud del tiempo medio en el caso  $\Lambda(t) = t$ . Dado que del EM no se obtienen directamente estimaciones de las varianzas y covarianzas de los parámetros, estimaciones por intervalo pueden obtenerse por medio de la matriz de información observada correspondiente a la verosimilitud original  $\ell(\boldsymbol{\theta})$  (Efron y Hinkley, 1978). Sin embargo dado que en confiabilidad muchas veces es de interés estimar ciertos cuantiles y estos bajo el modelo anterior no tienen formas explícitas, una mejor alternativa es el uso de bootstrap (Efron y Tibshirani, 1993; Davison y Hinkley, 1997; Meeker y Escobar, 1998).

### 3.4.2. Ejemplo

Para ilustrar el uso del modelo anterior consideremos los datos de láser de Meeker y Escobar (1998, página 462). Los datos consisten de trayectorias de degradación que resultan de medir repetidamente cada 250 horas a 15 dispositivos durante un periodo de tiempo de 4000 horas. La característica medida es el incremento (%) en la corriente de

operación para mantener aproximadamente constante la intensidad de la luz de salida de cada dispositivo a través del tiempo. La falla del dispositivo se define cuando la corriente de operación excede cierto umbral (10 %) (en porcentajes relativo a la corriente de operación inicial). Los datos se encuentran en la Figura 3.1, donde se puede ver que éstas son aproximadamente lineales.

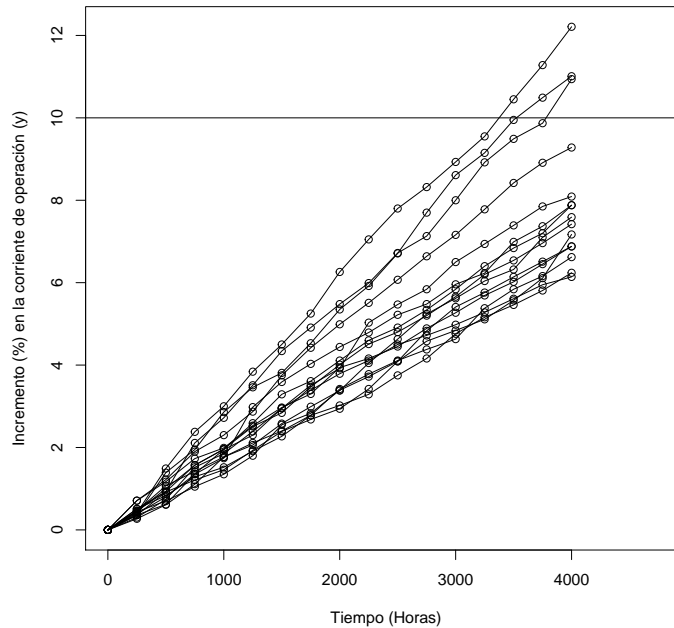


Figura 3.1: Trayectorias de degradación de datos de láser.

Aplicando el modelo anterior a éstos datos se obtienen las estimaciones que se muestran en el Cuadro 3.1, donde también se presentan las estimaciones del proceso Gaussiano inverso Normal-Gamma (PGING) (ver Apéndice 3.A) recientemente propuesto por Peng (2015), como extensión del proceso Gaussiano inverso propuesto por Wang y Xu (2010) quienes solamente consideran aleatorio al parámetro de forma  $\lambda$ . El AIC favorece al modelo aquí propuesto en relación al modelo proceso Gaussiano inverso Normal-Gamma aún cuando en éste último para tener mayor flexibilidad los parámetros  $1/\mu$  y  $\lambda$  se consideran que varían de unidad de acuerdo a una distribución Normal-Gamma.

MPGI		PGING	
Parámetro	Estimación	Parámetro	Estimación
$\mu_1 \times 10^{-3}$	2.72	$\delta_0 \times 10^2$	5.04
$\lambda_1 \times 10^{-4}$	1.18	$\sigma_\delta \times 10^{-1}$	7.85
$\mu_2 \times 10^{-3}$	1.79	$r_2 \times 10^5$	3.12
$\lambda_2 \times 10^{-5}$	6.17	$b_2 \times 10$	2.35
$p$	0.26		
loglik	102.12	loglik	95.61
AIC	-194.25	AIC	-183.21

Cuadro 3.1: Estimaciones de los parámetros de la mezcla de dos procesos Gaussianos inversos (MPGI) y del proceso Gaussiano inverso Normal-Gamma (PGING) (ver modelo en el Apéndice 3.A).

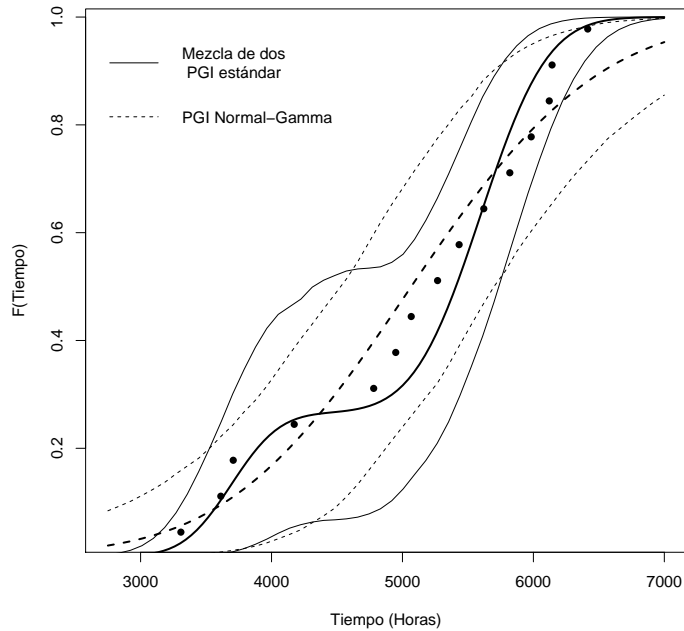


Figura 3.2: Distribución empírica de los pseudotiempo de falla (puntos negros), estimación de la función de distribución de los tiempos de cruce bajo la mezcla de dos procesos Gaussianos inversos (MPGI) y bajo el proceso Gaussiano inverso Normal-Gamma (PGING), junto con sus correspondientes intervalos bootstrap (percentil corregido por sesgo) de 95 % de confianza .

En la Figura 3.2 se presenta la distribución empírica de los pseudotiempo de falla (puntos negros), junto con la estimación por máxima verosimilitud de la función de distribución de la mezcla de dos procesos Gaussianos inversos estándar y bajo el modelo propuesto por Peng (2015). También se encuentran las bandas de confianza correspondientes bajo cada modelo, donde podemos apreciar que el ajuste del modelo propuesto es comparable con el recién propuesto por Peng (2015), dando menor incertidumbre en las colas de la distribución, no así en la parte donde se mezclan las funciones de distribución de los dos procesos Gaussianos inversos considerados. Ésto se debe a la marcada separación de dos posibles grupos de trayectorias de degradación (Figura 3.1) y al número reducido de unidades en la muestra.

Incrementando el número de componentes en la mezcla (considerando más procesos Gaussianos inversos), la log-verosimilitud aumenta como es de esperarse, sin embargo de acuerdo al AIC tal incremento no es suficiente y el modelo ajustado anteriormente con dos componentes es el que mejor ajusta a los datos.

Para hacer más parsimonioso el modelo, podríamos averiguar si el modelo se puede reducir haciendo que las dos componentes de la mezcla compartan el mismo segundo parámetro, es decir, probar si  $H_0 : \lambda_1 = \lambda_2 = \lambda$ . De acuerdo a la prueba de razón de verosimilitudes, ésta hipótesis se rechaza ya que el p-valor es 0.00886. Bajo modelos regulares el estadístico de razón de verosimilitud asintóticamente se comporta de acuerdo a una distribución Ji-Cuadrada, sin embargo a pesar de que aquí el tamaño de muestra no es grande aquí tal supuesto no es problema ya que se cumple bastante bien como se puede apreciar en la Figura 3.C.1, donde se muestra la distribución del estadístico obtenida por simulación junto a la distribución Ji-Cuadrada con un grado de libertad.

### 3.5. Conclusiones

Para tomar en cuenta la variabilidad de las unidades más allá de lo que un proceso Gaussiano inverso estándar puede capturar, a diferencia de las propuestas anteriores para mejorar ésta clase de procesos, aquí proponemos una mezcla discreta de éstos para cuando las trayectorias de degradación pertenecen a subpoblaciones coexistentes en el proceso. Para la estimación del modelo el algoritmo EM fue implementado eficientemente en la estimación de los parámetros. Bajo el modelo se pueden obtener cantidades de interés de forma sencilla, además de que se puede estimar la proporción de la población que pertenece a cada subpoblación (modo de falla por ejemplo), y con ello tomar mejores decisiones en la concentración de esfuerzos para el rediseño o la mejora del sistema de producción (Mendenhall y Hader, 1958).

Para cuando la existencia de subpoblaciones está presente pero éstas a la vez presentan heterogeneidad dentro de ellas, el uso de efectos aleatorios puede ser una alternativa para considerar ésta heterogeneidad dentro de clases. Específicamente, una vía es como la sugerida por Yuan y Ji (2015) en el contexto Bayesiano a través de los modelos de trayectoria de degradación general (general degradation path model). Por supuesto, lo mismo se puede hacer para los otros procesos estocásticos usados en degradación.

# Apéndices

## 3.A. Proceso Gaussiano inverso Normal-Gamma

Peng (2015) propone el siguiente modelo

$$\begin{aligned} Y(t)|\mu, \lambda &\sim IGP(\mu\Lambda(t), \lambda\Lambda(t)^2) \\ \delta = \mu^{-1}|\lambda &\sim N(\delta_0, \sigma_\delta^2/\lambda) \\ \lambda &\sim G(r_2, b_2), \end{aligned}$$

llamado la mezcla normal-Gamma de procesos Gaussianos inversos. La densidad marginal del vector de incrementos  $\Delta\mathbf{Y}$  está dada por

$$\begin{aligned} f_{\Delta\mathbf{Y}}(\Delta\mathbf{y}) &= \left( \prod_{j=1}^m \sqrt{\frac{(\Delta\Lambda_j)^2}{2\pi(\Delta y_j)^3}} \right) \sqrt{\frac{1}{\sigma_\delta^2(y_m - y_0 + \sigma_\delta^{-2})}} \frac{\Gamma(m/2 + b_2)}{\Gamma(b_2)} \\ &\times \frac{r_2^{b_2}}{\left[ r_2 + \frac{1}{2} \left( \Delta\Lambda' D \Delta\Lambda + \frac{\delta_0^2}{\sigma_\delta^2} - \frac{\left( \Lambda(t_m) - \Lambda(t_0) + \frac{\delta_0}{\sigma_\delta^2} \right)^2}{y_m - y_0 + \sigma_\delta^{-2}} \right) \right]^{m/2 + b_2}}, \end{aligned}$$

donde  $\Delta\Lambda = [\Delta\Lambda_1, \Delta\Lambda_m]'$  y  $D = \text{Diag}\left(\frac{1}{\Delta y_1}, \dots, \frac{1}{\Delta y_m}\right)$ .

## 3.B. La log-verosimilitud de la mezcla finita de procesos Gaussianos inversos es acotada

Bajo el proceso Gaussiano inverso definido en la Sección 3.2, la densidad conjunta de los incrementos para una trayectoria particular, está dada por

$$\begin{aligned}
f_{\Delta \mathbf{Y}|\mu, \lambda}(\Delta \mathbf{y}|\mu, \lambda) &= \prod_{j=1}^m \sqrt{\frac{\lambda (\Delta \Lambda_j)^2}{2\pi (\Delta y_j)^3}} \exp \left\{ -\frac{\lambda}{2\Delta y_j} \left( \frac{\Delta y_j}{\mu} - \Delta \Lambda_j \right)^2 \right\} \\
&= \left( \frac{\lambda}{2\pi} \right)^{m/2} \prod_{j=1}^m \sqrt{\frac{(\Delta \Lambda_j)^2}{(\Delta y_j)^3}} \exp \left\{ -\sum_{j=1}^m \frac{\lambda}{2\Delta y_j} \left( \frac{\Delta y_j}{\mu} - \Delta \Lambda_j \right)^2 \right\},
\end{aligned}$$

donde  $\Delta \mathbf{Y} = [\Delta Y_1, \dots, \Delta Y_m]'$  y  $\Delta \mathbf{y} = [\Delta y_1, \dots, \Delta y_m]'$ .

Para ver que la log-verosimilitud (3.2) correspondiente a la mezcla finita de procesos Gaussianos inversos definida en la Sección 3.3, es acotada, observemos que si cada trayectoria es medida en al menos dos tiempos de inspección diferentes

$$f_{\Delta \mathbf{Y}_i|\mu, \lambda}(\Delta \mathbf{y}_i|\mu_k, \lambda_k) \leq f_{\Delta \mathbf{Y}_i|\mu, \lambda}(\Delta \mathbf{y}_i|\hat{\mu}_{ik}, \hat{\lambda}_{ik}) = \left( \frac{\hat{\lambda}_{ik}}{2\pi} \right)^{m/2} \prod_{j=1}^m \sqrt{\frac{(\Delta \Lambda_j)^2}{(\Delta y_j)^3}} \exp \left( -\frac{m_i}{2} \right),$$

donde  $\hat{\mu}_{ik} = \frac{\sum_{j=1}^{m_i} \Delta y_{ij}}{\sum_{j=1}^{m_i} \Delta t_{ij}}$  y  $\hat{\lambda}_{ik} = \frac{m_i}{\sum_{j=1}^{m_i} \frac{(\Delta y_{ij} - \Delta t_{ij} \hat{\mu}_i)^2}{\Delta y_{ij} (\hat{\mu}_i)^2}}$  son los valores de  $\mu_k$  y  $\lambda_k$  que

maximizan  $f_{\Delta \mathbf{Y}_i|\mu, \lambda}(\Delta \mathbf{y}_i|\mu_k, \lambda_k)$ , y además como  $\log(x) \leq x \forall x > 0$ , entonces

$$\begin{aligned}
\ell(\theta|\mathbf{y}) &= \sum_{i=1}^n \log \left( \sum_{k=1}^K p_k f_k(\Delta \mathbf{y}_i; \mu_k, \lambda_k) \right) \\
&\leq \sum_{i=1}^n \sum_{k=1}^K p_k f_k(\Delta \mathbf{y}_i; \mu_k, \lambda_k) \\
&\leq \sum_{i=1}^n \sum_{k=1}^K p_k f_k(\Delta \mathbf{y}_i; \hat{\mu}_{ik}, \hat{\lambda}_{ik}) \\
&\leq \sum_{i=1}^n \sum_{k=1}^K f_k(\Delta \mathbf{y}_i; \hat{\mu}_{ik}, \hat{\lambda}_{ik}) \\
&\leq K \sum_{i=1}^n f_1(\Delta \mathbf{y}_i; \hat{\mu}_{i1}, \hat{\lambda}_{i1}),
\end{aligned}$$

y por lo tanto la log-verosimilitud dada en la expresión (3.2) es acotada.

### 3.C. Distribución empírica de la prueba de razón de verosimilitud para contrastar $H_0 : \lambda_1 = \lambda_2$

El estadístico de razón de verosimilitud para contrastar  $H_0 : \lambda_1 = \lambda_2$  vs  $H_1 : \lambda_1 \neq \lambda_2$  está dado por

$$\Lambda(\mathbf{y}, \mathbf{t}) = \frac{L(\hat{\boldsymbol{\theta}})}{L(\hat{\boldsymbol{\theta}}_0)},$$

donde  $L(\hat{\boldsymbol{\theta}})$  y  $L(\hat{\boldsymbol{\theta}}_0)$  denotan la verosimilitud evaluada en el estimador de máxima verosimilitud de  $\boldsymbol{\theta}$  sin restricción ( $\hat{\boldsymbol{\theta}}$ ) y bajo la restricción de  $H_0(\hat{\boldsymbol{\theta}}_0)$ , respectivamente.

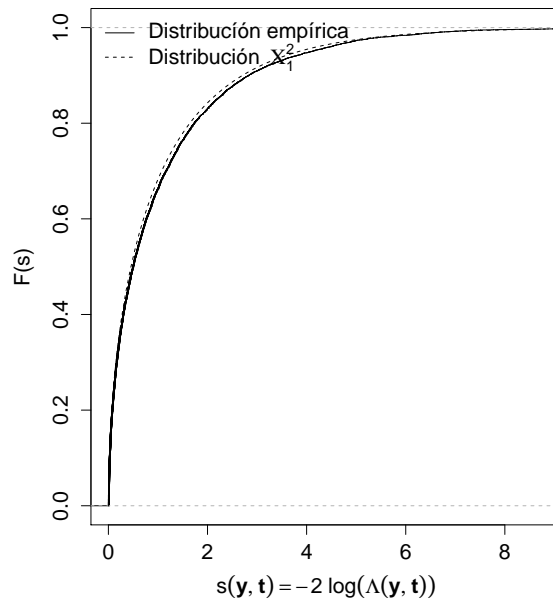


Figura 3.C.1: Distribución por simulación del estadístico de razón de verosimilitud bajo  $H_0 : \lambda_1 = \lambda_2$  junto con la distribución  $\chi_1^2$ .



# Capítulo 4

## Un método alternativo para la estimación de vida de anaquel

### 4.1. Introducción

El estudio de tiempos de vida de anaquel es de gran utilidad para administrar el movimiento de los productos desde su manufactura hasta su consumo. En estos estudios se evalúa la evolución del proceso de aceptación-rechazo del producto por parte de los consumidores.

Para el análisis de datos de vida de anaquel se han utilizado los enfoques de punto de corte y de análisis de supervivencia. De acuerdo al análisis de supervivencia, el tiempo que transcurre desde que un producto alimenticio se fabrica hasta que este se rechaza por un consumidor (que denominamos tiempo de rechazo), se considera como una variable aleatoria que se modela mediante una distribución que con frecuencia es de tipo Weibull o Lognormal (Hough et al., 2003; Hough y Garitta, 2012).

En los estudios de vida de anaquel no observamos tiempos de rechazo exactos, sino datos censurados, ya que los participantes en el estudio reciben en orden aleatorio solo un número finito de muestras (usualmente seis muestras) del producto alimenticio que tienen diferentes tiempos de almacenamiento ( $t_1, t_2, \dots, t_6$ ) y en cada caso aceptan o rechazan el producto. De estos datos, se deducen los tiempos de rechazo censurados. Si resulta que el tiempo de rechazo se desconoce, pero sabemos que se encuentra en un intervalo  $(t_r, t_l)$ , decimos que es un dato censurado por intervalo. Cuando resulta que el tiempo de rechazo (desconocido) es mayor que un tiempo  $t_r$ , tenemos una censura por la derecha. Finalmente, si concluimos que el tiempo de rechazo es menor que un cierto tiempo  $t_l$ , tenemos un dato censurado por la izquierda.

La información que aportan los consumidores en un estudio sensorial de vida de anaquel es muy compleja, debido a que no están entrenados y les es muy difícil distinguir muestras con pocos días de almacenamiento. Las respuestas de los consumidores tienen una gran variabilidad que se refleja frecuentemente en la longitud de los intervalos de censura de los tiempos de rechazo. Por esta razón, los datos de este tipo de estudios, están constituidos por un gran número de intervalos de censura amplios. En los datos que surgen de cada participante en el estudio, tenemos dos piezas de información que caracterizan su respuesta: El tiempo de rechazo y la consistencia con que responde a las preguntas de aceptación o rechazo de las muestras que recibe. Si denotamos las respuestas del consumidor por ceros y unos (0=rechazo y 1=aceptación), una sucesión de decisiones hechas por un consumidor consistente sería (1,1,1,0,0,0) (al ordenar las respuestas en orden creciente respecto al tiempo de almacenamiento de la muestra), en cambio, si la sucesión de respuestas de un consumidor es (1,0,1,0,0,0), diríamos que este consumidor es inconsistente. En el primer caso el tiempo de rechazo se encuentra entre  $t_3$  y  $t_4$ , y en cambio, en el caso del consumidor inconsistente diríamos que el tiempo de rechazo se encuentra en un intervalo de mayor amplitud, entre  $t_1$  y  $t_4$ . En general los consumidores inconsistentes darán respuestas con una mayor incertidumbre.

En este trabajo, proponemos un método alternativo para modelar y analizar los datos de un estudio sensorial de vida de anaquel, que nos permite expresar y cuantificar el tiempo de vida de anaquel y la consistencia para cada participante en el estudio. Inicialmente para cada consumidor estimamos su función de probabilidad de aceptación, de tal forma que queda parametrizada por dos parámetros, uno que “corresponda” a el tiempo de rechazo y el otro a una medida de consistencia. Debido a la gran variabilidad que existe entre los consumidores, no tiene sentido construir una función de aceptación para todos, sino una para cada uno. Por esta razón, modelamos estas funciones utilizando una función decreciente con parámetros aleatorios que expresan la variación entre individuos. La estimación del modelo propuesto de efectos mixtos no es simple, ya que la verosimilitud tiene involucradas integrales que no tienen una forma explícita. Para evaluar la verosimilitud utilizamos una aproximación numérica (cuadratura Gaussiana). Para estimar la distribución del tiempo de vida de anaquel, utilizamos las distribuciones propuestas para los parámetros aleatorios que consideramos inicialmente y la estimación de tales parámetros. La estimación del tiempo de vida de anaquel se compara con la obtenida con el enfoque de análisis de supervivencia convencional propuesto por Hough et al. (2003).

El modelo propuesto puede considerarse como un modelo de degradación binario, en donde la característica que observamos a través del tiempo, es la aceptación del producto por cada consumidor. El modelo que ajustamos a estas observaciones es una función de

probabilidad de aceptación, que inicialmente toma el valor uno y posteriormente va disminuyendo y eventualmente va a cero. El valor umbral de la probabilidad de aceptación que tomamos para definir el tiempo de falla o rechazo, es  $p_f$ . Por la gran variabilidad que hay entre consumidores, cada consumidor tiene una función de aceptación diferente, así que no tenemos una función de aceptación para todos los consumidores, sino una familia de funciones, modelada por una función de aceptación, con parámetros aleatorios.

El resto del trabajo está organizado de la siguiente manera. En la Sección 4.2, construimos el modelo de la probabilidad de aceptación del producto por parte de los consumidores, utilizando modelos mixtos no lineales generalizados. En la Sección 4.3 se obtiene la función de confiabilidad del tiempo de rechazo. En la Sección 4.4, estimamos los parámetros del modelo de efectos mixtos, utilizando el enfoque de verosimilitud. Además obtenemos la distribución del tiempo de rechazo. En la Sección 4.5, exponemos dos ejemplos con datos reales para ilustrar la metodología propuesta. Basando en las estimaciones de los parámetros del segundo ejemplo, en la Sección 4.6 se presenta un pequeño estudio de simulación para ilustrar el efecto de ignorar la inconsistencia de los consumidores. En la Sección 4.7 presentamos las conclusiones del trabajo.

## 4.2. Modelación de la función de probabilidad de aceptación

Modelamos el proceso de aceptación-rechazo del producto, por parte de los consumidores, como una función de probabilidad de aceptación decreciente al aumentar el tiempo de almacenamiento. Suponemos que la decisión del  $i$ -ésimo consumidor al tiempo de almacenamiento  $t_j$  ( $Y_{ij}$ ), es una variable Bernoulli con parámetro

$$p_i(t_j; \mu_i, \sigma_i) = P(Y_{ij} = 1 | \mu_i, \sigma_i) = 1 - F\left(\frac{\log(t_j) - \mu_i}{\sigma_i}\right), \quad i = 1, \dots, n, \quad j = 1, \dots, m, \quad (4.1)$$

donde  $F$  es una función de distribución de localización y escala estándar,  $\mu_i$  y  $\sigma_i$  son parámetros específicos de cada individuo, que varían de acuerdo a cierta distribución para permitir diferentes patrones de aceptación-rechazo entre estos, además de inducir correlación entre las respuestas correspondientes a un mismo consumidor. El efecto del tiempo sobre la probabilidad de aceptación, se tomó como  $\log(t_j)$  solamente por conveniencia para que las curvas de aceptación comenzarán a decrecer de 1, aunque otras formas de éste podrían adoptarse que mejor se ajusten al caso de estudio.

Éste modelo es un miembro de los llamados modelos lineales mixtos generalizados, con una parametrización diferente a la usual, donde el predictor  $(\log(t_j) - \mu_i) / \sigma_i$  así expresado queda como  $\beta_{0i} + \beta_{1i} \log(t_j)$ , donde  $\beta_{0i} = -\mu_i / \sigma_i$ ,  $\beta_{1i} = 1 / \sigma_i$  y el vector de parámetros  $(\beta_{0i}, \beta_{1i})'$  varía entre los individuos de acuerdo a una distribución normal bivariada. En la parametrización dada en la ecuación (4.1), los parámetros tienen una interpretación clara. El primero ( $\mu_i$ ) representa la localización de la curva de aceptación y nos indica la ubicación de tiempo de rechazo, mientras que el segundo ( $\sigma_i$ ) es un parámetro de escala, que dependiendo de su magnitud nos indica el grado de consistencia del individuo en sus respuestas de aceptación y rechazo del producto. Valores grandes de  $\sigma_i$  indican una consistencia baja en las respuestas del consumidor (gran incertidumbre), mientras que con valores pequeño ocurre lo contrario, esto es, son señal de un buen nivel de consistencia en las respuestas. La rapidez con que disminuye la curva de aceptación es inversamente proporcional al valor de  $\sigma_i$ , como se muestra en la Figura 4.1.

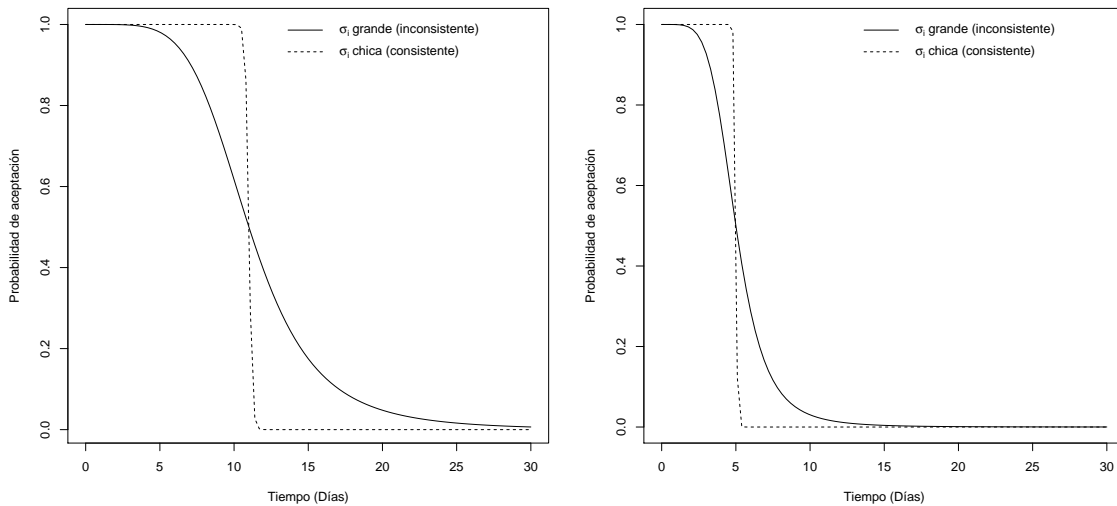


Figura 4.1: Curva de aceptación de dos consumidores: Derecha con tiempo de rechazo asociado medio e izquierda con tiempo de rechazo pequeño.

### 4.3. Tiempos de rechazo

Como mencionamos anteriormente en la Sección 4.1 , en los estudios de vida de anaquel los consumidores participantes no evalúan el producto alimenticio de manera continua a

través del tiempo de almacenamiento, para poder determinar en un momento dado su rechazo, si no que solo reciben un número finito de muestras con diferentes tiempos de almacenamiento y manifiestan en cada tiempo una respuesta binaria (aceptación y rechazo). Como ya mencionamos antes, modelamos este proceso de aceptación-rechazo utilizando una función de probabilidad que decrece con el tiempo de almacenamiento. Tomamos como tiempo de rechazo del producto al tiempo en que se alcanza una probabilidad de aceptación  $p_f$ . Consideramos que para tiempos de almacenamiento pequeños el consumidor aceptará el producto con alta probabilidad y para tiempos muy largos el producto tendrá una probabilidad de aceptación muy baja. Supondremos que en un tiempo intermedio se da el rechazo, cuando la probabilidad de aceptación se reduce  $p_f = 0.5$ , que podemos interpretar como el tiempo en que el consumidor le es difícil distinguir entre una muestra fresca y una envejecida y por lo tanto la decisión de aceptación la haría al azar, con probabilidad 0.5. Podemos notar que éste estudio de vida de anaquel es similar a un proceso de degradación en donde la respuesta es binaria (aceptación-rechazo) y la trayectoria de degradación es la función latente  $p_i(t)$ , el valor esperado de las respuestas (rechazo=0 y aceptación=1) . En éste proceso se declara la falla cuando ésta alcanza el valor de 0.5.

De acuerdo a la definición del tiempo de rechazo anterior, si  $T$  denota el tiempo de rechazo de un individuo seleccionado aleatoriamente, entonces

$$\begin{aligned}
 P(T > t) &= P(p_i(t) > p_f) \\
 &= P\left[\frac{\log(t) - \mu_i}{\sigma_i} < F^{-1}(1 - p_f)\right] \\
 &= P\{\exp[\mu_i + \sigma_i F^{-1}(1 - p_f)] > t\}. \tag{4.2}
 \end{aligned}$$

De la segunda igualdad, éste tiempo ( $T$ ) podría relacionarse a un proceso de degradación creciente continuo subyacente con trayectoria media el predictor  $\log(t_j)/\sigma_i - \mu_i/\sigma_i$  y donde la falla se declara cuando éste alcanza el valor  $F^{-1}(1 - p_f)$ . Notemos que si  $F$  es una función de distribución simétrica y la falla se declara cuando la curva de aceptación alcanza el valor  $p_f = 0.5$ , entonces la distribución de  $T$  solamente depende de  $\mu_i$ , y en tal caso si  $\mu_i$  es de la familia de localización y escala, la distribución de  $T$  es de la familia de log localización y escala. Por ejemplo Lognormal, Weibull o log-logística si  $\mu_i$  es normal, Gumbel o logística, respectivamente.

## 4.4. Estimación

Para la estimación de los parámetros involucrados en el modelo (4.1), considerando la parametrización  $\beta_{0i} + \beta_{1i} \log(t_j)$  anteriormente comentada, junto con el supuesto de que

$(\beta_{0i}, \beta_{1i})'$  varían entre los consumidores de acuerdo a una distribución normal bivariada, en la literatura existen diferentes técnicas para aproximar las integrales involucradas en la verosimilitud correspondiente. Por ejemplo, la aproximación de Laplace la cual se encuentra implementada en diferentes softwares, por ejemplo en SAS en el macro GLIMMIX (SAS-Institute-Inc., 2004), y en el paquete lme4 de R (Bates et al., 2015). Otro método numérico frecuentemente encontrado en estadística para aproximar integrales, es el uso de cuadratura adaptativa Gaussiana-Hermitiana (Liu y Pierce, 1994; Pinheiro y Bates, 1995) implementada por ejemplo en el PROC NLMIXED de SAS y en el paquete lme4 de R Team (2015), aunque en éste último solamente para cuando se tiene un solo efecto aleatorio. Ésta aproximación se reduce a la primera cuando el número de puntos de cuadratura es 1.

Para la estimación del modelo de probabilidad de aceptación descrito en la Sección 4.2, utilizamos el enfoque de máxima verosimilitud. Para ello aquí asumimos que los parámetros  $\mu_i$  y  $\sigma_i$  son independientes con distribuciones  $\exp(\mu_i) \sim \text{WEI}(\beta_W, \alpha)$  (forma y escala) y  $\sigma_i \sim \text{G}(\beta_G, \lambda)$  (forma y razón), donde WEI y G denotan las distribuciones Weibull y Gamma respectivamente. Además, supongamos que a cada uno de  $n$  individuos reclutados de forma aleatoria se le dieron  $m$  muestras de un producto correspondientes a  $t_1, t_2, \dots, t_m$ , días de almacenamiento, y en cada una de ellas se les preguntó si consumiría ( $Y_{ij} = 1$ ) o no ( $Y_{ij} = 0$ ) el producto. Bajo el modelo anterior dados los efectos aleatorios de individuo,  $(\mu_i, \sigma_i)$ , las respuestas dentro de un individuo son variables aleatorias independientes Bernoulli con parámetro dado en (4.1), y así la verosimilitud los parámetros  $\theta = (\beta_W, \alpha, \beta_G, \lambda)'$  del modelo descrito en la Sección 4.2, está dada por

$$\begin{aligned}
 L(\theta; \mathbf{y}) &= \prod_{i=1}^n L_i \\
 &= \prod_{i=1}^n \int_{-\infty}^{\infty} \int_0^{\infty} \prod_{j=1}^m [p_i(t_j; \mu_i, \sigma_i)]^{y_{ij}} [1 - p_i(t_j; \mu_i, \sigma_i)]^{1-y_{ij}} \times \\
 &\quad f_{\mu_i}(\mu_i) f_{\sigma_i}(\sigma_i) d\mu_i d\sigma_i,
 \end{aligned} \tag{4.3}$$

la cual no es sencilla de calcular porque las integrales involucradas no tienen forma explícita. De forma similar que para la parametrización comentada al inicio de ésta sección, podemos usar la aproximación de Laplace a la verosimilitud, sin embargo a pesar de su simplicidad en algunas situaciones no provee buenas aproximaciones sobre todo cuando el número de medidas repetidas es pequeño y la variabilidad entre sujetos es grande (Goldstein, 1991; Breslow y Clayton, 1993; Wolfinger, 1993; Rodriguez y Goldman, 1995;

German Rodriguez, 2001; Goldstein y Rasbash, 1996), lo cual puede ser el caso en el contexto que se está describiendo. Así, para aproximar la verosimilitud aquí optamos por la cuadratura adaptativa Gaussiana (Liu y Pierce, 1994; Pinheiro y Bates, 1995), aunque antes es conveniente hacer el siguiente cambio de variable en las integrales de (4.3) para que el soporte de éstas sea todo  $\mathbb{R} \times \mathbb{R}$  y así obtener mejores resultados:

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{i=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{j=1}^m [p_i(t_j; \mu_i, \exp(l_{\sigma_i}))]^{y_{ij}} [1 - p_i(t_j; \mu_i, \exp(l_{\sigma_i}))]^{1-y_{ij}} \times f_{\mu_i}(\mu_i) f_{\sigma_i}(\exp(l_{\sigma_i})) \exp(l_{\sigma_i}) d\mu_i dl_{\sigma_i}. \quad (4.4)$$

La aproximación de la verosimilitud (4.4) por cuadratura Gaussiana Hermitiana adaptativa (Pinheiro y Bates, 1995) está dada por

$$L(\boldsymbol{\theta}; \mathbf{y}) \approx \sum_{i_1=1}^{N_{CG}} \sum_{i_2=1}^{N_{CG}} \prod_{k=1}^2 w_{j_k} \exp \left[ g(\widehat{b}_i + \sqrt{2} H_i^{-1/2} z_{j_1, j_2}^*; \mathbf{y}_i) \right], \quad (4.5)$$

donde  $b_i = (\mu_i, l_{\sigma_i})'$ ,

$$g(b_i; \mathbf{y}_i) = \sum_{j=1}^m \log \{ [p_i(t_j; \mu_i, \exp(l_{\sigma_i}))]^{y_{ij}} [1 - p_i(t_j; \mu_i, \exp(l_{\sigma_i}))]^{1-y_{ij}} \} + \log [f_{\mu_i}(\mu_i) f_{\sigma_i}(\exp(l_{\sigma_i})) \exp(l_{\sigma_i})],$$

$\widehat{b}_i = (\widehat{\mu}_i, \widehat{l}_{\sigma_i})'$  es el vector  $b_i$  que minimiza  $g(b_i; \mathbf{y}_i)$ ,  $H_i^{-1/2}$  es la inversa de la matriz raíz cuadrada del hessiano de  $g(b_i; \mathbf{y}_i)$ ,  $H = \frac{\partial^2}{\partial b_i \partial b_i'} g(b_i; \mathbf{y}_i)$ ,  $z_{j_1, j_2}^* = [z_{j_1}, z_{j_2}]'$ ,  $z_j$  y  $w_j$ ,  $j = 1, \dots, N_{CG}$ , son los nodos y los pesos respectivamente de la cuadratura Gaussiana-Hermitiana (raíces de polinomios ortogonales respecto a la distribución normal estándar (Severini, 2005, página 305)), los cuales pueden obtenerse por ejemplo, mediante la función *gauss.quad* del paquete **statmod** (Smyth et al., 2016) del software estadístico R. Luego aproximaciones de los estimadores de máxima verosimilitud de los parámetros pueden obtenerse maximizando (4.5), por medio de algún método numérico, por decir, el método de Powell implementado en el paquete **powell** de R (Dorai-Raj y original Fortran from Mike Powell, 2015).

Una alternativa reciente, aunque computacionalmente demandante, es obtener los estimadores de máxima verosimilitud del modelo (4.1) usando el procedimiento data cloning el cual implícitamente aproxima las integrales por medio de herramientas bayesianas (Lele et al., 2007, 2010), y que cuando el número de clones es 1, ésta se reduce a la aproximación bayesiana donde aprioris de los parámetros deben especificarse y pueden llegar a tener influencia.

## 4.5. Ejemplos

Para ilustrar el procedimiento aquí propuesto para modelar y analizar los datos de un estudio de vida de anaquel, se muestran dos ejemplos con datos reales, donde un grupo de consumidores evalúa un conjunto de seis muestras de un producto alimenticio con diferentes tiempos de almacenamiento, y en cada evaluación, el consumidor responde (Si ó No) a la pregunta ¿Consumiría éste producto?. En cada uno de los ejemplos participan 50 consumidores no entrenados que reciben en un orden aleatorio las muestras del producto alimenticio (tostadas y pan blanco) con diferente tiempo de almacenamiento.

Las respuestas de los consumidores en los diferentes tiempos de inspección (tiempos de almacenamiento del producto) son binarias, estos es, aceptan o rechazan el producto. En la metodología de análisis de supervivencia propuesta por Hough et al. (2003), éstos datos se utilizan para obtener los datos censurados por intervalo que se analizan con técnicas de supervivencia o confiabilidad. En cambio, en la metodología de modelos mixtos generalizados (MDB) propuesta en la Sección 4.2, los datos entran en el análisis directamente, como datos binarios.

En cada ejemplo, se estima la función de supervivencia con el método MDB propuesto, así como también con el enfoque de análisis de supervivencia (Hough et al., 2003) y con un enfoque no paramétrico Turnbull (1976). Estas estimaciones (puntuales y por intervalo) se presentan y comparan gráficamente. En el caso del MDB, en donde consideramos efectos (los parámetros) aleatorios, evaluamos en cada caso la aleatoriedad del parámetro de escala que representa la consistencia de las respuestas de cada consumidor participante en el estudio.

### 4.5.1. Vida útil de tostadas

Para estimar la vida útil de tostadas, producidas por un fabricante, se llevo a cabo un estudio de vida de anaquel en el que participaron 50 consumidores no entrenados, que evaluaron el producto. En orden aleatorio cada participante recibió seis muestras que estuvieron previamente almacenadas durante 2, 7, 11, 15, 22 y 29 días. Para tener información del tiempo en que cada individuo rechazaría el producto, en cada muestra se les pidió que contestaran la pregunta ¿Consumiría éste producto? Si ó No. Los datos colectados en 6 consumidores se muestran en el Cuadro 4.1, donde el 1 representa que si consumiría el producto y cero que no lo consumiría.



Id	Días de almacenamiento					
	2	7	11	15	22	29
1	1	0	0	0	0	0
2	1	1	1	0	0	0
3	1	1	1	0	0	0
4	1	0	0	0	0	0
5	1	1	1	0	0	0
6	1	1	0	1	0	1

Cuadro 4.1: Datos de 6 consumidores de tostadas charras.

Para el análisis de los datos, asumimos que la distribución de los tiempos de rechazo subyacentes es una distribución Weibull bajo el análisis de supervivencia sugerido por Hough et al. (2003). Para aplicar la metodología aquí propuesta, asumimos que la distribución  $F$  involucrada en el modelo (4.1), es una distribución logística. Suponemos además que el tiempo de rechazo del producto por cada consumidor, corresponde al tiempo que su función de probabilidad de aceptación alcanza el umbral  $p_f = 0.5$ .

En el proceso de estimación de la función de confiabilidad, expuesto en la sección anterior, se utilizó una aproximación de la función de confiabilidad, siguiendo un algoritmo de cuadratura Gaussian Hermitiana adaptativa (AGHQ, por sus siglas en inglés) (Pinheiro y Bates, 1995), con 20 puntos de cuadratura, tomados directamente del paquete **statmod** de R (Smyth et al., 2016).

Con el fin de investigar el patrón de variabilidad del parámetro de escala  $\sigma_i$ , se estima el modelo (4.1) en su forma reducida, considerando fijo el parámetro de escala, esto es, tomando a  $\sigma_i$  una constante ( $\sigma$ ) a través de los consumidores (MDBR).

En la Figura 4.1, se muestran las estimaciones puntuales y por intervalo de la función de confiabilidad (función de aceptación) obtenidas con los enfoques mencionados: 1) Análisis de supervivencia de Hough et al. (2003), 2) No paramétrico, 3) el MDB (con  $\sigma_i$  aleatorio) aquí propuesto, y 4) el modelo propuesto con inconsistencia constante entre los individuos,  $\sigma_i = \sigma$  (MDBR). Observamos que las estimaciones de la función de confiabilidad, bajo los diferentes enfoques mencionados, muestran una diferencia muy pequeña, y en particular, los datos no muestran evidencia de una diferencia entre los modelos completo y reducido ya que las curvas obtenidas son casi indistinguibles, así que podemos tomar el modelo reducido ( $\sigma_i$  fijo). Esto nos lleva a plantear que los datos no muestran una diferencia significativa en la consistencia de las respuestas de los participantes. Esta conclusión también es apoyada por el hecho de que el el AIC (302.0483) para el modelo completo es muy cercano al

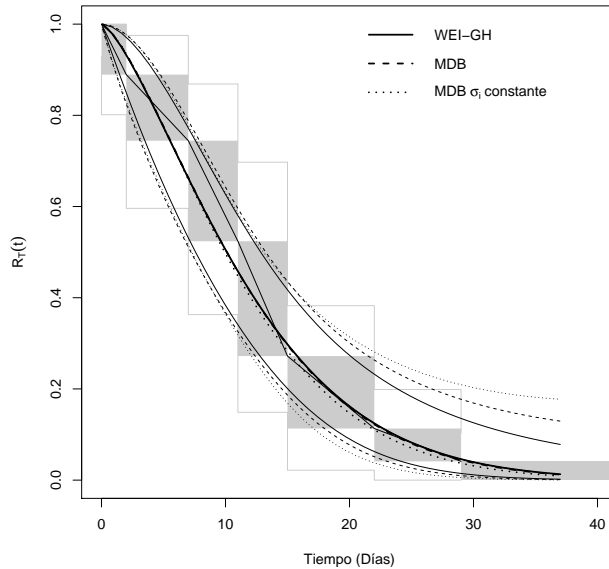


Figura 4.1: Función de confiabilidad estimada con la metodología de Hough et al. (2003) (WEI-GH) bajo el modelo Weibull, el MDB (4.1) y el MDB (4.1) con consistencia constante ( $\sigma_i = \sigma$ ) (MDBR), todas con intervalo de confianza asintóticos (método delta).

AIC (302.1549) del modelo reducido. También se puede obtener una conclusión equivalente utilizando una prueba de razón de verosimilitudes.

Para el modelo propuesto completo (MDB) y también para el reducido (MDBR) se obtuvieron intervalos de confianza asintóticos de la confiabilidad usando el método delta a partir de la matriz de varianzas y covarianzas estimada de los parámetros.

Para la evaluación del ajuste de un modelo de regresión con respuesta continua, frecuentemente una gráfica de los residuos es de utilidad, pero esto no es así cuando tenemos un modelo cuya respuesta es categórica, como en nuestro caso. Como los datos están agrupados y las observaciones dentro de cada grupo no son idénticamente distribuidos, para tener una idea del ajuste del modelo, comparamos las proporciones de aceptación (empírica) en cada tiempo de inspección ( $t_j$ ) con las obtenidas bajo el modelo logístico ajustado (proporción muestral de aceptación), como se muestra en la Figura 4.2. Aquí se puede ver que el modelo es concordante a los datos observados, ya que la distribución (cuantiles 2.5% y 97.5%) de las proporciones muestrales bajo el modelo ajustado contienen a las observadas, por lo que el ajuste del modelo parece razonable. Un comportamiento muy similar se obtiene con el modelo reducido donde se considera que  $\sigma_i$  es constante entre los consumidores, lo que también soporta al modelo reducido como suficiente para representar

la variabilidad de los datos.

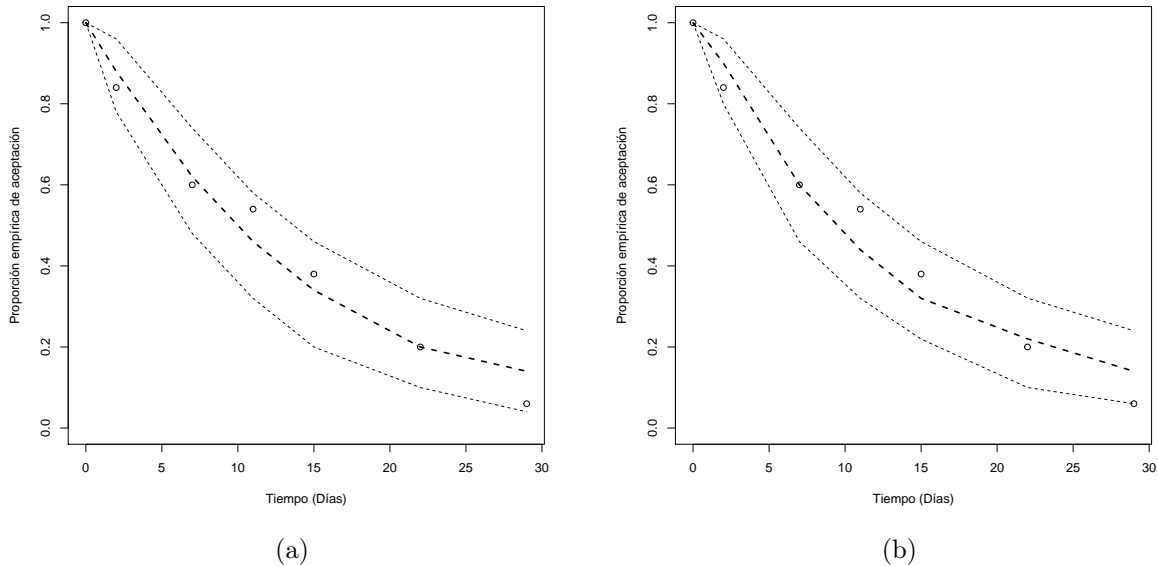


Figura 4.2: Cuantiles 2.5% y 97.5% de la distribución de las proporciones muestrales de aceptación en cada uno de los tiempos de inspección con  $n = 50$  consumidores (datos de tostadas) bajo el modelo MDB (a) y el modelo MDBR (b).

#### 4.5.2. Vida útil de pan

Para estimar la vida útil de un tipo de pan blanco, se realizó un estudio de vida de anaquel, de manera similar al que se describió en el ejemplo anterior. Se reclutaron 50 consumidores regulares de este tipo de pan, que recibieron en un orden aleatorio muestras de un tipo de pan que tenían 3, 8, 13, 21, 23 y 25 días de almacenamiento. Después, para obtener información de los tiempos de rechazo del producto, se les preguntó a los participantes en el estudio si consumirían el producto.

Aquí también asumimos nuevamente que bajo la metodología (de análisis de supervivencia) de Hough et al. (2003), la distribución de los tiempos de rechazo subyacentes, es una distribución Weibull. Aquí asumimos nuevamente, que la distribución  $F$  del modelo (4.1), es una distribución logística y que el tiempo de rechazo de cada consumidor corresponde al tiempo en que la curva de aceptación (4.1) alcanza el valor  $p_f = 0.5$ . El proceso de estimación de la función de confiabilidad es igual al que se siguió en el ejemplo anterior.

En la Figura 4.3, se muestran las estimaciones puntual y por intervalo de la función de confiabilidad (aceptación), obtenidas con los enfoques: 1) Análisis de supervivencia

de Hough et al. (2003), 2) No paramétrico, 3) Modelo (con  $\sigma_i$  aleatorio) aquí propuesto (MDB), y 4) el modelo propuesto en forma reducida, esto es, con  $\sigma_i = \sigma$  (MDBR). En este ejemplo, a diferencia del anterior, las gráficas muestran una discrepancia entre los diferentes modelos considerados, sobre todo en las colas de la distribución. Respecto al modelo propuesto (4.1), en su versión completa ( $\sigma_i$  aleatorio) y reducida ( $\sigma_i = \sigma$  fijo), para valores pequeños del tiempo, la diferencia tanto puntual como por intervalo, de las funciones de confiabilidad no es tan grande, pero para tiempos grandes si hay una diferencia mayor y más aún en el límite superior del intervalo de confianza, lo que puede sugerir que en este caso el modelo completo representa mejor la variabilidad presente en los datos. Esta diferencia también se refleja en los valores del AIC para ambos modelos (AIC=338.412 para el modelo completo y AIC=345.915). Esta diferencia también se observa en las gráficas de la Figura 4.4, donde se presentan los cuantiles 2.5 % y 97.5 % de la distribución de las proporciones muestrales de aceptación en cada uno de los tiempos de inspección simulada bajo los modelos completo y reducido.

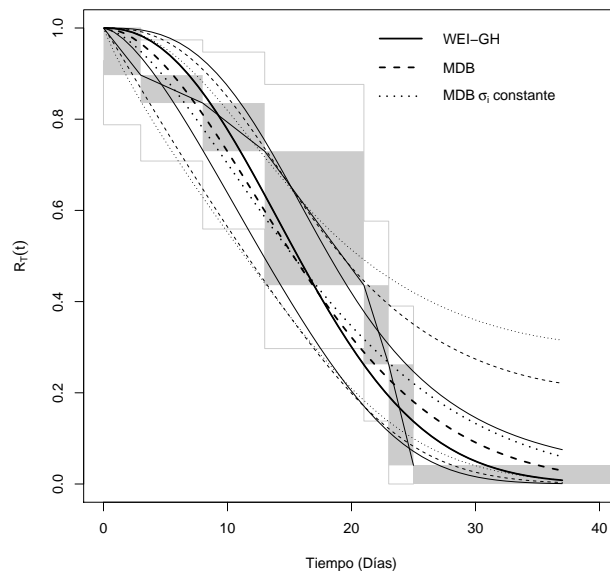


Figura 4.3: Función de confiabilidad estimada: a) metodología de Hough et al. (2003) (WEI-GH) bajo el modelo Weibull, MDB (4.1) y MDB (4.1) con consistencia constante ( $\sigma_i = \sigma$ ) (MDBR), todas con intervalos de confianza asintóticos (método delta).

En este ejemplo, a diferencia del anterior, observamos que el modelo completo ajusta mejor que el reducido, y en tal caso, tenemos evidencia de la existencia de una variación en la consistencia de las respuestas entre los diferentes consumidores que participan en

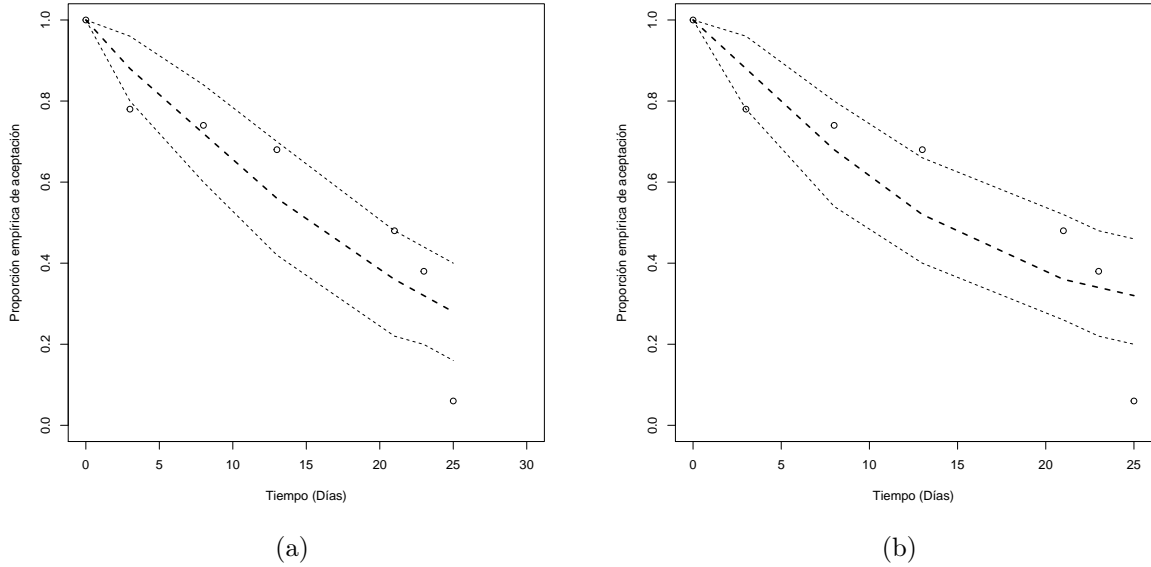


Figura 4.4: Cuantiles 2.5% y 97.5% de la distribución de las proporciones muestrales de aceptación en cada uno de los tiempos de inspección con  $n = 50$  consumidores bajo el MDB (a) y el MDBR (b).

la muestra. A partir de las estimaciones de máxima verosimilitud de los parámetros, se pueden obtener “estimaciones” de los efectos individuales del modelo. Un procedimiento de estimación consiste en obtener la moda posterior, es decir,

$$\begin{aligned}
 (\hat{\mu}_i, \hat{\sigma}_i) &= \arg \max_{\mu_i, \sigma_i} f(\mu_i, \sigma_i | \mathbf{y}_i) = \arg \max_{\mu_i, \sigma_i} f(\mathbf{y}_i | \mu_i, \sigma_i) f(\mu_i; \beta, \alpha) f(\sigma_i; \beta_G, \lambda) \\
 &= \arg \max_{\mu_i, \sigma_i} \prod_{j=1}^m [p_i(t_j; \mu_i, \sigma_i)]^{y_{ij}} [1 - p_i(t_j; \mu_i, \sigma_i)]^{1-y_{ij}} f_{\mu_i}(\mu_i; \hat{\beta}, \hat{\alpha}) f_{\sigma_i}(\sigma_i; \hat{\beta}_G, \hat{\lambda})
 \end{aligned}$$

donde  $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\alpha}, \hat{\beta}_G, \hat{\lambda})'$  es el estimador de máxima verosimilitud de los parámetros  $\boldsymbol{\theta}$ . Estas predicciones que se muestran en el la Figura 4.5, nos dan información del patrón que sigue la consistencia de las respuestas entre los diferentes participantes en el estudio. Podemos observar que aunque algunos valores de  $\hat{\sigma}_i$  son muy pequeños, algunos consumidores muestran valores mayores de inconsistencia. Además, observemos que la localización del log-tiempo de rechazo ( $\hat{\mu}_i$ ) y el parámetro  $\hat{\sigma}_i$  no muestran evidencia clara de estar correlacionados, por lo que el supuesto hecho en el modelo (4.1) de que la localización del log-tiempo de los consumidores no depende de la consistencia, tiene sentido. Es interesante que cuando en un panel de consumidores que participan en un estudio de vida de anaquel,

existan diferentes niveles de inconsistencia, esta se pueda estimar. Esta estimación es posible en el caso de los modelos lineales generalizados de efectos mixtos, como el que aquí se propone.

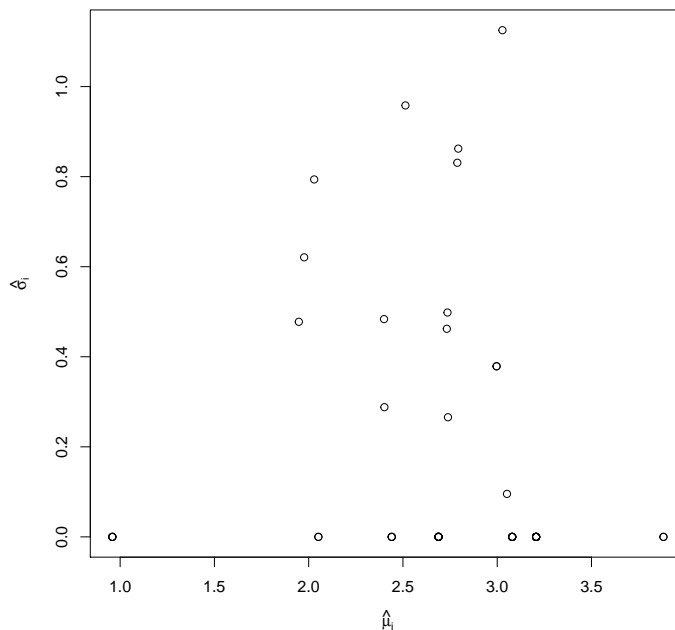


Figura 4.5: “Estimación de efectos” individuales.

Todos los cálculos se realizaron utilizando el lenguaje de cómputo estadístico R (Team, 2015).

## 4.6. Simulación

Para ver si la diferencia observada en la estimación de la función de confiabilidad se mantiene en muestras repetidas, realizamos un pequeño estudio de simulación, pero ahora en la estimación de algunos cuantiles de ésta.

Específicamente el pequeño estudio de simulación que se presenta es para ver el efecto que tienen las inconsistencias (en las respuestas) de los participantes en la estimación de la función de distribución (o confiabilidad) que resulta con el método de Hough et al. (2003) y con el método MDB propuesto. En este estudio, asumimos que el mecanismo generador de datos es el modelo (4.1) (MDB) descrito en la Sección 4.2.

Los parámetros utilizados en este estudio de simulación, se tomaron del segundo ejemplo, que fue donde resultó una mayor inconsistencia de los consumidores participantes y donde se obtuvo una mayor diferencia entre la metodología de Hough et al. (2003) y la aquí propuesta.

Se simularon  $M = 1000$  conjuntos de datos, cada uno con 50 consumidores y tiempos de inspección: 0, 3, 8, 13, 21, 23 y 25, bajo el modelo (4.1), con los parámetros:  $\log(\beta_W) = -0.6120$ ,  $\log(\alpha) = 2.9272$ ,  $\log(\mu_G) = -0.4700$  y  $\log(\lambda) = 0.1760$ . Para cada conjunto de datos simulados, se estimaron, siguiendo ambas metodologías, los cuantiles 0.05, 0.10, 0.50, 0.90 y 0.95 % de la distribución del tiempo de rechazo, y luego de todos los conjuntos de datos simulados se calculó, el promedio, el sesgo relativo y el error estándar de cada uno de éstos cuantiles de interés, bajo cada metodología.

$p$	$t_p$	Hough et al. (2003)			MDB		
		Media	EE	SR	Media	EE	SR
0.05	3.73	4.2630	1.2126	14.2786	3.9069	1.3291	4.7324
0.10	5.51	6.0301	1.3746	9.4066	5.6561	1.5121	2.6214
0.50	15.31	15.1975	1.5134	-0.7286	15.3307	1.8491	0.1413
0.90	29.36	27.7559	2.7425	-5.4494	29.6130	4.6911	0.8768
0.95	33.86	31.7187	3.5859	-6.3200	34.2961	6.2402	1.2922

Cuadro 4.1: Promedio (Media), error estándar (EE) y sesgo relativo (SR) en la estimación de cuantiles  $p = 0.05, 0.1, 0.5, 0.9$  y  $0.95$  bajo la metodología de Hough et al. (2003) y la metodología propuesta en la Sección 4.2 (MDB). Para aproximar la verosimilitud en (4.5) se usaron 20 puntos de cuadratura.

De los resultados que se presentan en el Cuadro 4.1, observemos que con el modelo de Hough et al. (2003) (WEI-H) los primeros dos cuantiles estudiados son sobreestimados, y el resto subestimados, como lo observado en el Ejemplo 2, mientras que bajo el MDB, todos los cuantiles son sobreestimados. Un patrón general común en los dos métodos de estimación es que el sesgo relativo de ambos decrece en magnitud hasta llegar al cuantil 50 %, pero aumentan en el cuantil más grande estudiado.

Valores grandes del sesgo relativo para los cuantiles 0.05 y 0.10, puede deberse a que en ésta región (de forma relativa a la magnitud de lo que se quiere estimar) los tiempos de inspección son los más espaciados (0, 3, 8, 13, 21, 23, 25). Los sesgos relativos bajo WEI-GH son aproximadamente tres veces más grandes que los obtenidos con el MDB en los primeros dos cuantiles, y la diferencia es poco más grande en los cuantiles 0.5, 0.90 y 9.95. Esto último puede deberse a que la metodología MDB aprovecha implícitamente

la curva media de rechazo y en cambio, el procedimiento WEI-H toma como censurados por la derecha a los individuos que rechazan después del tiempo 25. Un comportamiento similar pero contrario se obtiene al considerar el error estándar de los estimadores, aunque la diferencia en los tres primeros cuantiles es ligera, y moderada en los últimos dos, debido a que la segunda metodología da aproximadamente errores estándar 42% más grandes.

Así de lo anterior, vemos que aunque bajo el modelo propuesto hay sesgo en la estimación de los cuantiles, bajo la metodología de Hough et al. (2003) éstos llegan a ser más grandes sobre todo en los cuantiles pequeños, que son los de mayor interés cuando se utilizan para definir la vida útil de los productos que se quiere lleguen al consumidor como productos frescos, apenas distinguibles del producto nuevo, como ocurre con el pan. Esta diferencia se debe a que la metodología WEI-H que se basa en los intervalos de censura no toma en cuenta la incertidumbre presente en éstos como consecuencia de la inconsistencia de las respuestas de los consumidores.

## 4.7. Conclusiones

La metodología que aquí se propone para el análisis de estudios de datos de vida de anaquel, es un procedimiento alternativo al análisis de supervivencia o confiabilidad propuesto por Hough et al. (2003). Una ventaja del método propuesto es que involucra explícitamente el nivel de consistencia de las respuestas de los consumidores participantes en el estudio. Hemos mostrado dos ejemplos con datos reales de vida de anaquel. En el primer ejemplo, en donde la variación de las inconsistencias fue no significativa, las funciones de confiabilidad obtenidas con ambas metodologías son comparables. En cambio, en el segundo, en donde los datos provienen de una población de consumidores con una inconsistencia más variable, los resultados en ambas metodologías muestran mayores diferencias.

El modelo propuesto también puede considerarse como un modelo de degradación binario, en donde la característica que observamos a través del tiempo, es la aceptación o rechazo del producto por cada consumidor y la falla (rechazo) del producto se declara cuando la función de probabilidad de aceptación  $p_i(t)$  alcanza un nivel de probabilidad  $p_f$ .

Aunque en los ejemplos se asume que la distribución  $F$  corresponde a la distribución logística, hay que notar que se puede elegir cualquier distribución de localización y escala, y en particular cuando se tenga evidencia de asimetría de la curva de aceptación, se puede optar por una distribución asimétrica, como puede ser la distribución Gumbel, normal asimétrica o logística asimétrica, por mencionar algunas. En ambos ejemplos también se asumió que en el tiempo de rechazo (subyacente) el consumidor tenía el mayor grado de confusión para decidir aceptar o no el producto ( $p_f = 0.5$ ). Sin embargo, éste valor se



puede cambiar cuando se requiera ser conservador en la estimación de la distribución de los tiempos de rechazo.

Aquí asumimos que los parámetros que varían entre los consumidores  $(\mu_i, \sigma_i)$  son independientes con distribuciones  $\exp(\mu_i) \sim \text{WEI}(\beta_W, \alpha)$  (forma y escala) y  $\sigma_i \sim \text{G}(\beta_G, \lambda)$  (forma y razón), pero se podrían considerar otras distribuciones.

Dada la inconsistencia en las respuestas de los consumidores, una pregunta natural es, ¿Cómo deben elegirse los tiempos de almacenamientos de forma óptima?, y más generalmente, ¿Cómo diseñar el experimento para obtener la mayor información posible con los recursos disponibles?. Estas preguntas pueden explorarse como trabajo a futuro, así como la incorporación de información previa de experimentos similares para mejorar la precisión de las cantidades de interés.



# Capítulo 5

## Comparación de algunos criterios de selección de modelos con datos de vida censurados

### 5.1. Introducción

El análisis de datos de confiabilidad involucra la identificación y estimación de parámetros, siendo el primero muchas veces el paso más importante y difícil, todavía aún más cuando se tienen datos censurados. Entre los métodos más tradicionales para la elección de distribuciones se encuentran las gráficas de probabilidad, pruebas de hipótesis, pruebas de bondad de ajuste, y criterios de selección. El AIC (criterio de información de Akaike) y BIC (criterio de información bayesiana) son dos criterios de uso frecuente para la selección de modelos, los cuales difieren solamente en los términos de penalización aunque sus orígenes provienen de avenidas completamente distintas.

El AIC fue propuesto por Akaike (1974) como un estimador insesgado asintótico de la información de Kullback-Leibler esperada, entre un modelo candidato ajustado y el modelo verdadero, mientras que el BIC fue derivado por Schwarz (1978) como una aproximación a una transformación de la probabilidad posterior de un modelo candidato.

A través del tiempo el uso de ambos criterios para la selección de modelos ha crecido significativamente. Entre algunas de las primeras aplicaciones del AIC sugeridas por el autor se encuentran análisis factorial, análisis de componentes principales, regresión múltiple y series de tiempo. En general, el área de aplicaciones de éste índice es diverso, por

ejemplo en ecología (Johnson y Omland, 2004; Dennis et al., 2006; Hoeting et al., 2006), bioinformática (Edwards et al., 2010), astrofísica (Liddle, 2007), por mencionar algunas.

En el intento de averiguar que criterios resultan mejores en diferentes situaciones, existen varios estudios teóricos y empíricos alrededor de esto. Por ejemplo, Akaike (1977) mostró que el desempeño del BIC es mejor que el AIC en la elección del orden de modelos polinomiales y el orden de modelos autoregresivos. Shibata (1976) y Hannan (1982) muestran la inconsistencia del AIC en la elección de modelos autorregresivos y modelos ARMA, respectivamente. Geweke y Meese (1981) muestran, además de la consistencia del BIC en modelos de regresión, que el AIC es más probable a elegir un modelo más grande en comparación que el BIC.

Koehler y Murphree (1988) comparan el AIC y BIC en la elección del orden de modelos de series de tiempo en datos reales, donde encuentran que el orden identificado por el AIC excede por dos o más el orden del modelo elegido por el BIC en más de la mitad de los casos. Sin embargo, en estos casos los modelos seleccionados por el segundo criterio, producen predicciones que son tan buenas como las producidas por los modelos seleccionados por el AIC.

Volinsky y Raftery (2000) proponen una modificación del BIC inspirados en el hecho de que bajo ciertas condiciones de regularidad, se tiene una buena aproximación de éste al factor de Bayes ( $2 \log(BF)$ ). Ellos proponen definir la penalización del BIC en términos del número de fallas en vez del número de observaciones. Basados en la corrección del AIC hecha por Hurvich y Tsai (1989) (AICc), Liang y Zou (2008) proponen la siguiente versión mejorada de AIC en el contexto de modelos de supervivencia:

$$AIC_{SUR} = AIC + \frac{2(p+2)(p+3)}{n-p-3}$$

donde  $p$  es el número de variables explicativas incluidas en el modelo. La racionalidad de ésta propuesta es demostrada por medio de simulación al compararla con un estimador insesgado aproximado de la información de Kullback-Leibler entre el modelo verdadero y un modelo ajustado, en el caso de que la distribución de los tiempos de supervivencia es la exponencial y el tiempo de censura es una constante. También por medio de simulación, reportan que ésta modificación resultó superior que el AIC y BIC en otros escenario estudiados con las distribuciones lognormal, Weibull y log-logística. Sin embargo, no abordan explícitamente el efecto de la fracción de censura en el desempeño de tales criterios.

Más recientemente Lee y Ghosh (2009) evaluaron el desempeño del AIC, AICc y BIC con varios modelos de covarianza espaciales, donde sus resultados de simulación revelaron que éstos criterios se desempeñan bien para algunos modelos y mal para otros. En la mayoría de los casos estudiados obtienen un mejor desempeño del AIC o AICc en comparación que el BIC. Por otro lado, Emiliano et al. (2014) evalúan el desempeño de estos mismos criterios en la selección de modelos estándar, modelos de crecimiento biológicos y modelos de series de tiempo, donde concluyen que el uso indiscriminado de tales criterios puede llevar a resultados imprecisos y que el desempeño de estos depende del modelo.

Así, aunque han existido muchos estudios en la comparación del AIC y BIC, poca atención se le ha puesto y mucho uso se le ha dado a estos, en el caso de muestras censuradas, una situación en la cual no fueron derivados. Por ejemplo, el software JMP (JMP, 2013) tiene implementado estos criterios además de una modificación del AIC, no obstante, hasta hoy no se tiene un respaldo teórico ni empírico para éstas situaciones. Motivado por esto, en el presente trabajo se presenta un estudio de simulación para ver el efecto de la fracción de censura en el desempeño de estos criterios.

El trabajo ésta organizado de la siguiente manera: en la Sección 5.2 presentamos una breve descripción de los criterios a estudiar, en la Sección 5.3 se presenta el estudio de simulación y la comparación de los criterios. Conclusiones y discusiones de los resultados obtenidos son dados en la Sección 5.4.

## 5.2. Criterios de selección de modelos

La estimación de la información de Kullback-Leibler es el punto clave en la derivación del criterio de información AIC, el cual ha sido ampliamente usado para la selección de modelos estadísticos. La información de Kullback-Leibler es considerada como una medida de bondad de ajuste del modelo propuesto  $f(x)$  hacia el modelo verdadero  $g(x)$  (Konishi y Kitagawa, 2008). Sin embargo, en la práctica es bastante limitada ya que casi siempre se desconoce la verdadera distribución que genera los datos, lo cual impide calcularla. De esta manera Akaike (1974) derivó el AIC como un estimador insesgado asintótico del negativo de la información de Kullback-Leibler esperada relativa  $(-E[\log(f(X))])$  donde  $X$  tiene densidad  $g(\cdot)$ , entre el modelo candidato ajustado  $f(x)$  y el verdadero  $g(x)$ . El AIC está definido como

$$\text{AIC} = -2\ell(\hat{\theta}) + 2p \tag{5.1}$$

donde  $\ell(\hat{\theta}) = \log[L(\mathbf{x}|\hat{\theta})]$  es la log-verosimilitud del modelo considerado evaluado en el estimador de máxima verosimilitud de  $\theta$  y  $p$  es la dimensión de  $\theta$ .

Para reducir el sesgo del AIC en modelos de regresión lineal, una versión corregida del AIC fue sugerida por Sugiura (1978), y expandida por Hurvich y Tsai (1989, 1991) para modelos de regresión no lineales y modelos autorregresivos. Ésta nueva versión, conocida en la literatura como criterio de información de Akaike corregida (AIC<sub>c</sub>), está definida por

$$\begin{aligned} \text{AIC}_c &= -2\ell(\hat{\theta}) + \frac{2pn}{n-p-1} \\ &= \text{AIC} + \frac{2p(p+1)}{n-p-1} \end{aligned} \quad (5.2)$$

Con ésta versión se obtiene una penalización más fuerte para modelos complejos en comparación con la versión estándar del AIC, sin embargo, para  $p$  fijo, conforme  $n$  crece éste converge al AIC. Así, a menos que el tamaño de muestra sea grande respecto al número de parámetros, Burnham y Anderson (2002, página 67) recomiendan el uso del AIC<sub>c</sub> (por ejemplo, cuando la razón  $n/p$  es menor que 40 cuando  $p$  es grande).

Otro método popular usado para la selección de modelos es el criterio de información Bayesiana (BIC) propuesto por Schwarz (1978). Este fue derivado como una aproximación a una transformación de la probabilidad posterior de cada uno de los modelos candidatos, asumiendo que estos a prior son igualmente probables. El BIC está definido como

$$\text{BIC} = -2\ell(\hat{\theta}) + p \log(n) \quad (5.3)$$

donde  $\ell(\hat{\theta}) = \log[L(\mathbf{x}|\hat{\theta})]$  es la log-verosimilitud del modelo considerado evaluado en el estimador de máxima verosimilitud de  $\theta$ ,  $p$  es la dimensión de  $\theta$  y  $n$  es el tamaño de muestra. En el estudio también consideramos la modificación de BIC propuesta por Volinsky y Raftery (2000) que usa el número observado de fallas ( $n_f$ ) en vez del tamaño de muestra:

$$\text{BIC}_M = -2\ell(\hat{\theta}) + p \log(n_f). \quad (5.4)$$

Estos criterios sirven para la elección de un modelo dentro de un conjunto de candidatos, haciendo un balance entre el ajuste del modelo (primer término de los tres) y la complejidad de este (segundo término). Para un conjunto de datos dado, estos criterios trabajan bajo la misma lógica; estos se calculan para cada modelo y se elige como el más apropiado al que tenga menor AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub>, respectivamente. La penalización del modelo en el BIC es mayor al de AIC por lo que los modelos elegidos por el BIC tienden a ser más parsimoniosos que los elegidos por el AIC (Claeskens y Hjort, 2008).

### 5.3. Estudio de simulación

En ésta sección se presenta el estudio de simulación para evaluar el desempeño de los criterios antes descritos en la identificación del verdadero modelo. De especial interés se consideran algunas distribuciones de tiempo de vida más comunes (exponencial, Rayleigh y Weibull), censura aleatoria por la derecha, diferentes fracciones de censura y diferentes números nominales de fallas.

Para cada simulación, se generan 10,000 muestras de tamaño  $n = \lfloor r/(1 - \gamma) \rfloor$  ( $\lfloor \cdot \rfloor$  es la función máximo entero) de una de las distribuciones, la cual es tomada como la distribución verdadera, como más abajo se detalla. Para cada muestra generada, por máxima verosimilitud se estiman los parámetros de cada una de las distribuciones candidatas (incluida la que genera los datos), y posteriormente se calculan cada uno de los 4 criterios (AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub>). Finalmente, se calcula la proporción de veces que cada uno de los modelos candidatos es favorecido como el mejor en cada uno de los 4 criterios.

#### 5.3.1. Implementación del mecanismo de censura aleatoria

El estudio de simulación es hecho considerando censura aleatoria por la derecha, donde asumimos que el tiempo a la falla ( $T$ ) y el tiempo de censura ( $C$ ) son variables aleatorias independientes. Si  $F_T(\cdot)$  es la función de distribución de  $T$  y  $F_T^{-1}(\cdot)$  la función cuantil asociada, entonces para generar muestras con una fracción de censura  $\gamma$  deseada, se propone el siguiente algoritmo:

1. Generar  $U_{k_\gamma} \sim U(0, k_\gamma)$  y  $U \sim U(0, 1)$  independientes.
2. Si  $U_{k_\gamma} < U$ , definir  $\delta = 0$  y tomar  $T^* \sim F_T^{-1}(U_{k_\gamma})$ , mientras que si  $U < U_{k_\gamma}$ ,  $\delta = 1$  y tomar  $T^* \sim F_T^{-1}(U)$ .
3. Repetir los pasos 1 y 2 tantas veces como el tamaño de muestra que se requiera.

donde  $k_\gamma = \begin{cases} 2(1 - \gamma) & \text{si } 0.5 \leq \gamma \leq 1 \\ \frac{1}{2\gamma} & \text{si } 0 \leq \gamma < 0.5 \end{cases}$ , dependiendo de la fracción de censura  $\gamma$  deseada.

Notemos que bajo éste algoritmo no se garantiza que cada muestra  $T_1^*, \dots, T_n^*$  extraída contenga exactamente  $100(1 - \gamma)\%$  observaciones no censuradas, si no que el número esperado

de observaciones censuradas es  $\gamma$ . Más detalles de éste mecanismo de censura pueden verse en el Apéndice 5.A, donde también puede consultarse un algoritmo alternativo para la generación de muestras con censura aleatoria.

### 5.3.2. Resultados de simulación

El estudio se hace considerando las distribuciones exponencial (EXP), Rayleigh (RAY) y la distribución Weibull (WEI) como distribuciones candidatas. Las fracciones de censura experimentadas son  $\gamma = 0, 0.01, 0.05, 0.1, 0.25, 0.40, 0.5, 0.60, 0.75, 0.95$  y  $0.99$ , mientras que  $4, 6, 8, 10, 15, 20, 25, 50, 75$  y  $100$  son los números nominal de fallas aleatorias ( $r$ ) utilizados.

Puede verse que bajo los criterios de selección de modelos aquí examinados, los resultados no dependen del parámetro de escala de la distribución de la cual provienen los datos, y tampoco dependen de cual de los dos mecanismos anteriores se use para simularlos. Esto significa que los parámetros de escala de las distribuciones no necesitan variarse, y por ello sin pérdida de generalidad estos son tomados igual a  $50$  en cada una de las distribuciones, y únicamente el parámetro de forma en la distribución WEI es el que se varía:  $0.5, 1.5, 2.5$  y  $3.5$ .

En los Cuadros 5.B.1-5.B.3 se presentan los resultados de la simulación cuando los datos son generados de la distribución EXP con parámetro de escala  $\eta = 50$ . Para un valor de  $\gamma$  fijo, como era de esperarse, en general se observa un mejoramiento de los criterios en la selección del modelo EXP a medida que el número esperado de fallas aumenta. Para un valor fijo de  $r$  de los experimentados menores que  $10$  (Cuadro 5.B.1), observemos que en todos los criterios, sorprendentemente, la probabilidad de selección correcta (PSC) muestra una tendencia creciente a medida que la fracción de censura crece. En éste rango el BIC es el que tiene el mayor incremento pero el  $AIC_c$  es el que comienza (en  $\gamma = 0$ ) con la mayor probabilidad de selección correcta, y se tiene que el modelo RAY es su mayor competidor de acuerdo a todos los criterios. También puede observarse que para  $r$  fijo, las probabilidades de selección correcta del  $AIC_c$  son ligeramente mayores que las obtenidas con el AIC para valores pequeños de la fracción de censura, pero tal diferencia es cada vez más despreciable para valores grandes de ésta. Lo anterior lo explica el hecho de que para un número esperado de fallas  $r$  y una fracción de censura grande (por decir del  $99\%$ ), se requiere un tamaño de muestra grande (relativo al número de parámetro), y como consecuencia el segundo término en la definición del  $AIC_c$  (5.2) es casi despreciable.

Para este rango de valores los criterios  $AIC_c$  y BIC son los que tienen el mejor desempeño, pero ninguno de estos dos es uniformemente mejor dado que depende del valor de



$r$  y  $\gamma$ ; en  $r = 4$  el  $AIC_c$  resulta mejor para  $\gamma \leq 0.60$ , mientras que el BIC para  $\gamma \geq 0.75$ ; en  $r = 6$  el  $AIC_c$  para  $\gamma \leq 0.5$  y el BIC para  $\gamma \geq 0.60$ ; en  $r = 8$  el  $AIC_c$  para  $\gamma \leq 0.40$  y el BIC para  $\gamma \geq 0.40$ ; y para  $r = 10$  el  $AIC_c$  para  $\gamma \leq 0.25$ , mientras que el BIC para el resto de valores de  $\gamma$ , donde a partir de  $\gamma = 0.75$  el  $BIC_M$  es el que ocupa el segundo mejor lugar. También podemos observar que tampoco el peor criterio se tiene de forma uniforme ya que el criterio  $BIC_M$  resultó ser el peor en  $r = 4, 6, y 8$ , aunque en el último solamente para los dos últimos valores de  $\gamma$ , y el AIC resultó el peor en el resto de las combinaciones.

Del Cuadro 5.B.2 observemos que para valores de  $r$  mayores que 15 la situación cambia un poco. De acuerdo al AIC el modelo WEI es el mayor competidor, lo mismo con el  $AIC_c$  con excepción en los valores de  $r = 15$  y  $\gamma$  menores que 0.25, mientras que con el  $BIC_M$  para valores mayores que 20. Para el BIC esto depende de la combinación de  $r$  y  $\gamma$ : para  $r = 15$  el modelo RAY es aún el más competidor; para  $\gamma = 20$  solamente para valores de  $\gamma$  menores o iguales que 0.5; si  $r = 25$  para  $\gamma$  menor o igual que 0.60; si  $r = 50$  en todos los valores de  $\gamma$  excepto en 0.95; mientras que a partir de 75 en todos los valores de  $\gamma$ . Para  $r$  fijo, las probabilidades de selección correcta para el AIC y  $BIC_M$  casi no se ven afectadas por los valores de  $\gamma$ , donde ahora el desempeño del segundo comienza a hacer mejor y tal diferencia crece si  $r$  aumenta. La probabilidad de selección correcta del  $AIC_c$  tiene un sentido contrario a lo observado en los valores de  $r$  menores que 10, ya que como puede observarse, ahora la probabilidad de selección correcta disminuye aproximadamente un 3% al pasar del 0% de censura al 99% en los valores  $r = 15, 20$  y 25, donde éste criterio tiene un comportamiento muy similar.

Aquí la diferencia entre el AIC y el  $AIC_c$  es menos despreciable, y cada vez menos a medida que se tienen un  $r$  más grande. A partir de  $r = 50$  las probabilidades de selección correcta del  $AIC_c$  comienzan a estabilizarse alrededor del 84% para todos los valores de  $r$  y  $\gamma$ , comportamiento que también lo comporte el AIC. Ésta similitud en gran medida se debe al hecho de que los términos de penalización tienden a ser muy parecidos para tamaños de muestras mayores que 40, como se expresó en la introducción. Para el BIC la tendencia sigue siendo la misma dado que para  $r$  fijo se observa un incremento en la probabilidad de selección correcta conforme la fracción de censura lo hace, aunque a partir de  $r = 15$  el desempeño de éste comienza a ser claramente superior de forma uniforme al resto de los criterios, y más aún en valores más grandes de  $r$ . Por ejemplo, en  $r = 50$  se obtienen probabilidades de selección correcta mayores de 0.95 que crecen conforme se tiene más censura (Cuadro 5.B.3). Una representación gráfica de estos resultados de simulación para 6 de los 11 valores estudiados de la fracción de censura ( $\gamma = 0, 0.25, 0.50, 0.75, 0.95$  y 0.99), se muestran en la Figura 5.1.

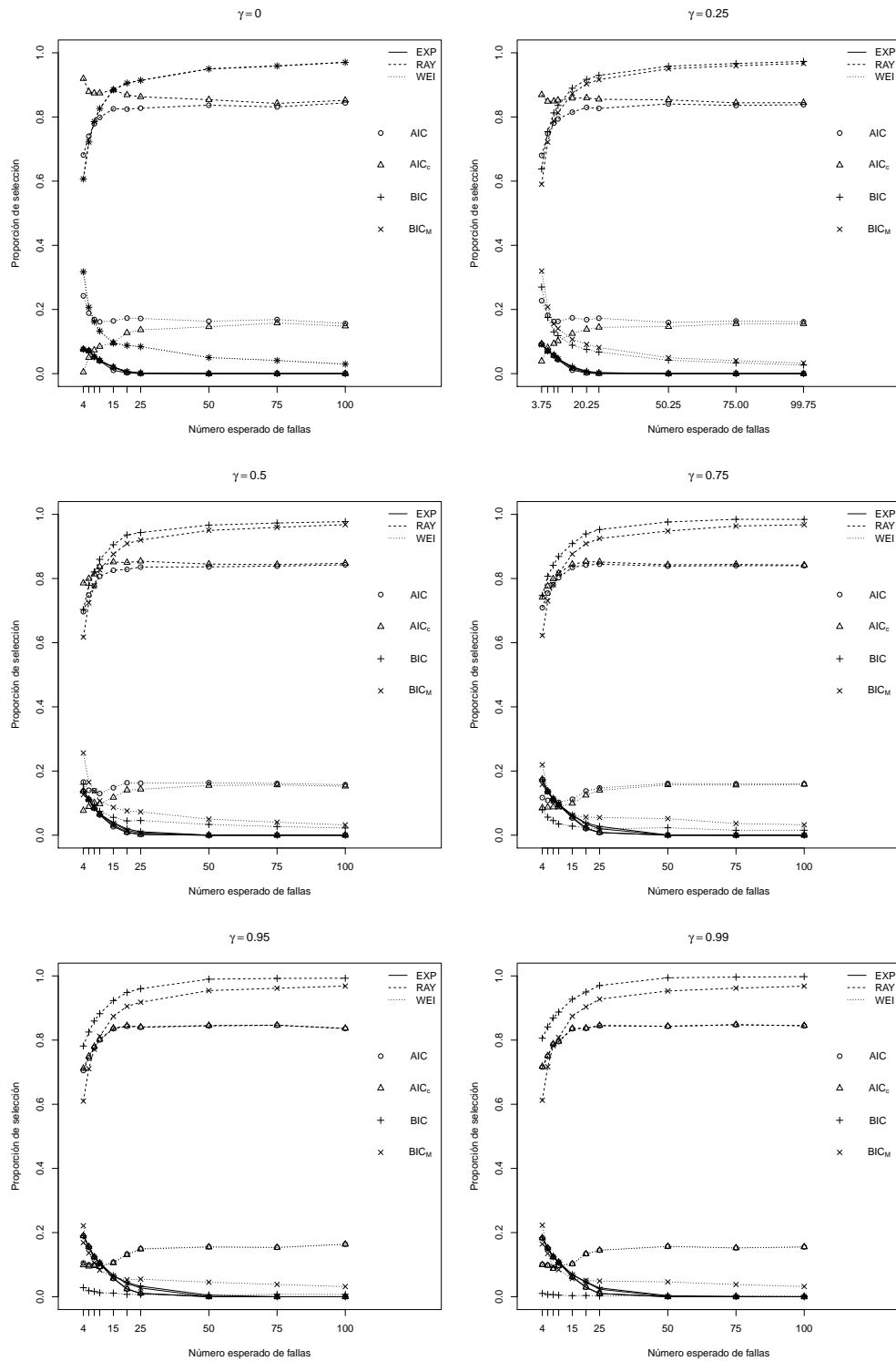


Figura 5.1: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución EXP.

Los resultados de simulación cuando los datos son generados del modelo RAY con parámetro 50 se muestran en los Cuadros 5.B.4-5.B.6.

De forma similar al caso anterior donde los datos provienen del modelo EXP, para un valor de  $\gamma$  fijo, en general se observa un mejor desempeño de los criterios en la probabilidad de selección correcta a medida que el número esperado de fallas  $r$  aumenta. En comparación a cuando los datos provienen de la distribución EXP, la mayor diferencia se observa en los primeros 4 valores de  $r$  considerados, debido a que aquí las probabilidades de selección correcta son mayores, y más aún en el criterio  $AIC_c$ . Respecto al AIC, en el primero de estos ( $r = 4$ ), aquí la probabilidad de selección correcta comienza a crecer (en  $\gamma = 0$ ) a partir de un valor (0.6815) moderadamente más grande, mientras que a partir del segundo ( $r = 6$ ) el efecto de la fracción de censura sigue siendo positivo aunque cada vez menos hasta llegar a  $r = 50$ , donde al igual que el en caso anterior (EXP), las probabilidades se comienzan a estabilizar alrededor de 0.84.

En éste mismo rango de  $r$  (los primeros 4 valores,  $r = 4, 6, 8$  y  $10$ ), respecto al  $AIC_c$  observemos que para  $r$  fijo, el efecto de la fracción de censura es opuesto al obtenido en la situación anterior (EXP) debido a que conforme ésta aumenta la probabilidad de selección correcta decrece (Cuadro 5.B.4), pero tal decrecimiento va siendo cada vez menor a medida que  $r$  es más grande. En cuanto al BIC observemos que se tiene que para  $r$  fijo, también la probabilidad de selección correcta aumentan conforme la fracción de censura es mayor. Sin embargo, aquí el desempeño de los criterios depende de muchos más combinaciones de  $r$  y  $\gamma$ , ya que por ejemplo, para  $r = 4, 6, 8$  y  $10$ , el  $AIC_c$  es el que resulta mejor para valores de  $\gamma$  menores que 0.60, 0.50, 0.40 y 0.25 respectivamente, y para el complemento de los valores de  $\gamma$  anteriores, el BIC es el que resulta mejor.

A partir de  $r = 15$  (Cuadros 5.B.5-5.B.6) observemos que las probabilidades de selección correcta son muy similares a cuando los datos son generados del modelo EXP. En particular, el desempeño del BIC también comienza a ser claramente superior de forma uniforme al resto de los criterios, y mucho más en valores grandes de  $r$ , por ejemplo, en  $r = 50$  se obtienen probabilidades de selección correcta superiores a 0.95 que aumentan si se tiene más censura (Cuadro 5.B.6). En la Figura 5.2 se muestra una representación gráfica de estos resultados de simulación para 6 de los 11 valores experimentados de la fracción de censura ( $\gamma = 0, 0.25, 0.50, 0.75, 0.95$  y  $0.99$ ).

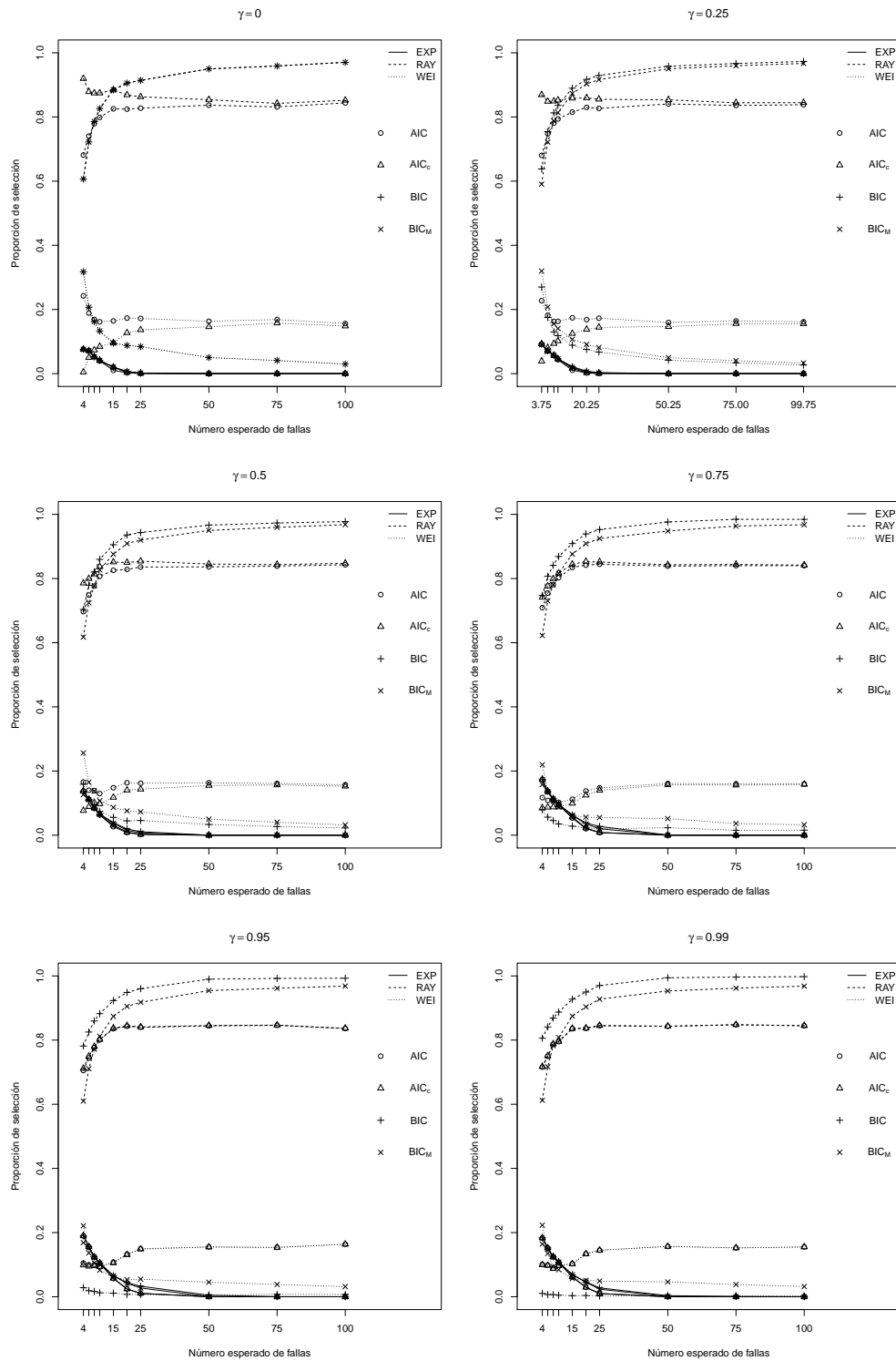


Figura 5.2: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución RAY.

En los Cuadros 5.B.7-5.B.9 se presentan los resultados de la simulación cuando los datos son generados de la distribución WEI con parámetro de escala 50 y de forma 0.5.

Al igual que antes se observa que, para una fracción de censura específica, el desempeño en la selección de los criterios mejora a medida que el número de fallas crece. De estos Cuadros también se puede observar que ninguno de los criterios considerados resulta uniformemente mejor que el resto. Por ejemplo, para  $r = 4$  y  $6$  (Cuadro 5.B.7), el  $BIC_M$  es el que resulta mejor en todos los valores de  $\gamma$ , mientras que el segundo mejor comportamiento está dividido dado que le corresponde al BIC para valores de  $\gamma$  menores que 0.40, y al AIC para el resto de tales valores (Cuadro 5.B.7). El peor comportamiento también es dependiente de los valores de  $\gamma$ , ya que para antes de 0.60 lo posee el  $AIC_c$ , mientras que el BIC para el complemento.

En comparación al caso en que los datos fueron generados de la distribución EXP, en general para cada uno de los primeros dos valores de  $r$  (4 y 6), no se observa un comportamiento con tendencia creciente en la probabilidad de selección correcta a medida que la fracción de censura aumenta, pero si se tienen algunas tendencias. Específicamente, para el BIC la fracción de censura muestra tener una tendencia negativa en la probabilidad de selección correcta en ambos valores de  $r$  (4 y 6), mientras que en el resto de los criterios generalmente se tiene una tendencia positiva. Por ejemplo, en  $r = 4$  el  $BIC_M$  tiene un incremento de aproximadamente de 0.20 en la PSC al pasar de 0.4877, cuando no se tiene censura, a 0.6766, cuando se alcanza el 99% de censura, y para el BIC la PSC disminuye de 0.4877 a 0.2495.

Para  $r = 8$ , el comportamiento de éstos criterios es diferente ya que ahora el  $AIC$  y el  $BIC_M$  resultan mejores en las 5 primeras y 6 últimas fracciones de censura respectivamente, y los que resultan peores son los criterios  $AIC_c$  y BIC en las 6 primeras y 5 últimas fracciones de censura, respectivamente.

Para  $r = 10$  el AIC es el que resulta mejor con probabilidades de selección correcta que van de 0.82 a 0.8711. Este es seguido por el  $BIC_M$  en los primeros 9 valores de  $\gamma$  y por el  $AIC_c$  en los últimos dos. El  $AIC_c$  es el que resulta el peor en los valores de  $\gamma$  menores o iguales a 0.25, y el BIC en valores mayores o iguales a 0.40

La superioridad del AIC se sigue manteniendo en el resto de los valores de  $r$ , donde en  $r = 15$  las probabilidades de selección correcta resultan altas variando entre 0.9330 y 0.9555 para los diferentes valores de  $\gamma$  (Cuadro 5.B.8). Es hasta  $r = 20$  donde el  $AIC_c$  ocupa la segunda mejor posición en la selección correcta del modelo que generó los datos, y además donde el efecto de la fracción de censura en la PSC es apenas ligero, con excepción del BIC donde ésta disminuye de 0.9651 a 0.8078. Éste efecto negativo de  $\gamma$  en el BIC se sigue observando en  $r = 25$ , pero a partir de  $r = 50$ , este efecto es casi despreciable y de hecho el

comportamiento de todos los criterios comienza a ser casi indistinguible con probabilidades de selección correcta muy cercanas a 1. En éste escenario, en todos los criterios el modelo EXP es el que resultó el mayor competidor del modelo WEI, lo cual se debe al valor del parámetro de forma elegido. Gráficamente los resultados para este caso (WEI con parámetro de forma 0.5) en 6 valores de la fracción de censura ( $\gamma = 0, 0.25, 0.50, 0.75, 0.95$  y  $0.99$ ) se muestran en la Figura 5.3.

Un escenario que resulta mucho más complicado para los criterios es cuando simulamos del modelo WEI con parámetro de forma igual a 1.5, un valor que se encuentra entre los correspondientes valores del parámetro de forma en la distribución WEI que definen a los modelos EXP y RAY, respectivamente. Para cada valor de  $\gamma$ , en los primeros 4 valores de  $r$  (Cuadros 5.B.7-5.B.12), se observa un efecto negativo de éste ( $r$ ) en el desempeño de los criterios. En los primeros dos (4 y 6), el  $BIC_M$  es el que resulta con mejor desempeño aunque las probabilidades de selección correcta son bastante bajas, alrededor del 0.15 en  $r = 4$  y alrededor de 0.07 en  $r = 6$ , para todos los valores de  $\gamma$ . El segundo mejor comportamiento se tiene con el BIC pero para los valores de  $\gamma$  menores o iguales a 0.40, mientras que con el AIC para el resto.

En  $r = 8$  el AIC comienza a ser más competitivo ya que éste registra el mejor desempeño en todos excepto en los últimos dos valores de  $\gamma$ , donde con poca ventaja (sobre el AIC), el  $BIC_M$  continua desempeñándose mejor. El peor desempeño lo tiene el  $AIC_c$  hasta  $\gamma = 0.40$ , y en valores superiores a éste, el BIC.

A partir de  $r = 10$ , el AIC comienza a tener un comportamiento superior sobre el resto, el efecto de  $r$  ya tiene un efecto positivo sobre la PSC, pero también se sigue observando un efecto negativo de la fracción de censura (y mucho más en el BIC) ya que para  $r$  fija, la PSC disminuye conforme  $\gamma$  aumenta. Este efecto se hace cada vez menos conforme  $r$  crece. El segundo mejor comportamiento se tiene con el  $AIC_c$  a partir de  $r = 15$ , aunque en éste valor las probabilidades de selección correcta son mucho más bajas que las del AIC a medida que la fracción de censura es más chica. Como era de esperarse, en  $r = 50$  las PSC con el  $AIC_c$  empiezan a ser muy similares a las del AIC, y es a partir de donde estos dos criterios comienzan a identificar al modelo correcto con probabilidades mayores que 0.76, bastante mejores que las correspondientes al BIC y al  $BIC_M$ , donde de éstos dos últimos el  $BIC_M$  es el menos afectado por  $\gamma$ .

A medida que  $r$  aumenta, la diferencia entre los criterios es cada vez menos. En  $r = 100$  las probabilidades de selección correcta en todos los criterios son mayores que 0.90 excepto en el BIC y el  $BIC_M$  en fracciones de censura mayores que 0.60 y 0.75, respectivamente. A pesar de tener un número esperado de fallas bastante grande, el efecto de  $\gamma$  en el BIC continua siendo grande debido a que la PSC pasa de 0.9305 a 0.5283. Gráficamente los

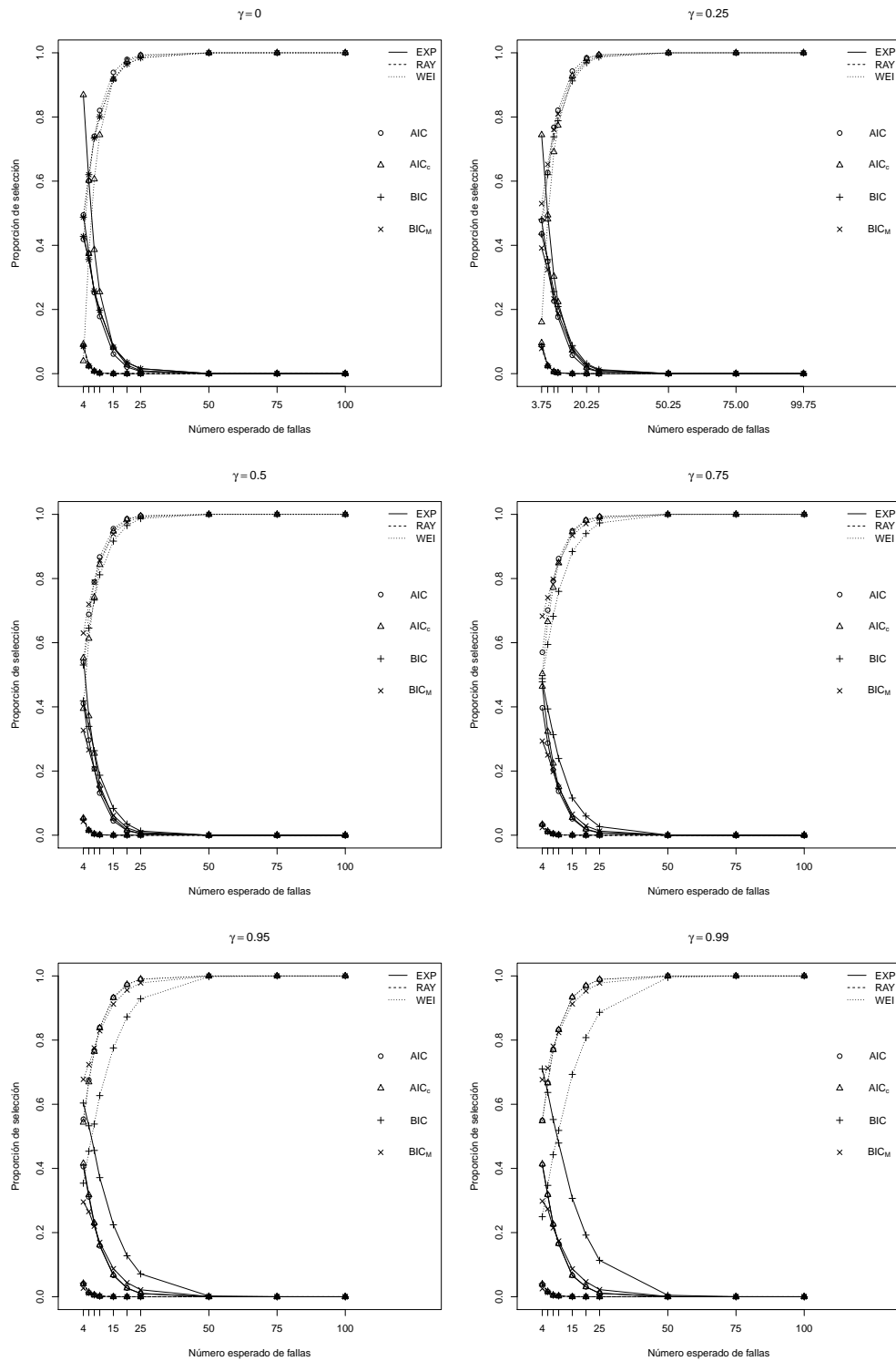


Figura 5.3: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5.

resultados de los Cuadros 5.B.7-5.B.12 para 6 valores de la fracción de censura ( $\gamma = 0, 0.25, 0.50, 0.75, 0.95$  y  $0.99$ ) se muestran en la Figura 5.4.

Otro escenario considerado es cuando simulamos del modelo WEI con parámetro de forma igual a 2.5. Por la cercanía a 2, de antemano se espera que el modelo que más le compita al modelo WEI sea el modelo RAY. Los resultados se muestran en los Cuadros 5.B.13-5.B.15. En general, en cuanto al orden en el desempeño de los criterios, se observa un comportamiento similar al comentado en el escenario anterior. Sin embargo, aquí las probabilidades de selección correcta son mucho más grandes, y tal diferencia se hace más chica a medida que  $r$  crece hasta el punto en que en  $r = 50$ , las cosas se invierten. Por ejemplo, la PSC en el AIC decrece de 0.7633 a 0.6252 al pasar de  $\gamma = 0$  a la fracción de censura 0.99, mientras que en el escenario anterior ocurre el decrecimiento de 0.8435 a 0.7663, lo que sugiere que la consistencia de la selección correcta de los criterios es mucho más lenta en éste escenario. Gráficamente estos resultados se muestran en la Figura 5.5 para 6 de los 11 valores de  $\gamma$  estudiados.

Al considerar  $\beta = 3.5$  en la distribución WEI, comportamientos similares pueden observarse en el desempeño de los criterios (Cuadros 5.B.16-5.B.18), solo que aquí por la mayor separación de las distribuciones que éste valor del parámetro implica, las probabilidades de selección tienen un mucho mejor comportamiento que los escenarios anteriores. Por ejemplo, en  $r = 10$  el AIC tiene un desempeño bastante bueno ya que la PSC es 0.8601 con muestras completas, aunque ésta se degrada a 0.6813 al alcanzar la fracción de censura más grande considerada (0.99). El segundo mejor posicionado es el  $BIC_M$  con excepción en los últimos dos valores de  $\gamma$  donde el  $AIC_c$  es el que domina. A partir de  $r = 15$  la jerarquía en el desempeño de éstos criterios ya es uniforme, con el AIC en la primera posición, el  $AIC_c$  en la segunda, y el  $BIC_M$  y BIC en la tercera y cuarta respectivamente. Aunque estas PSC son bastante altas, observemos que hasta  $r = 25$  el efecto de la fracción de censura aun se manifiesta y más aún en el BIC, donde la PSC decrece de 0.9738 a 0.5195. En  $r = 50$  el efecto de la censura ya no es tan fuerte, y mucho menos en los valores  $r$  más grandes donde todos los criterios identifican al modelo WEI correctamente con probabilidades muy cercanas a 1. Para 6 de los 11 valores de  $\gamma$  estudiados, en la Figura 5.6 se presenta el comportamiento de estos criterios bajo éste escenario.

Estos resultados también pueden explicarse asintóticamente. Si en vez de considerar los tres modelos solamente consideramos dos de ellos, digamos el modelo WEI y del que se generaron los datos, EXP (o RAY), la comparación (o selección) de estos modelos, en caso del AIC equivale a la prueba de razón de verosimilitud sobre la hipótesis  $H_0 : \beta = 1$  (o  $\beta = 2$ ) con error tipo I igual a 0.16. Sin embargo, al considerar los tres modelos candidatos simultáneamente, una consecuencia como la anterior no es directa, aunque también se



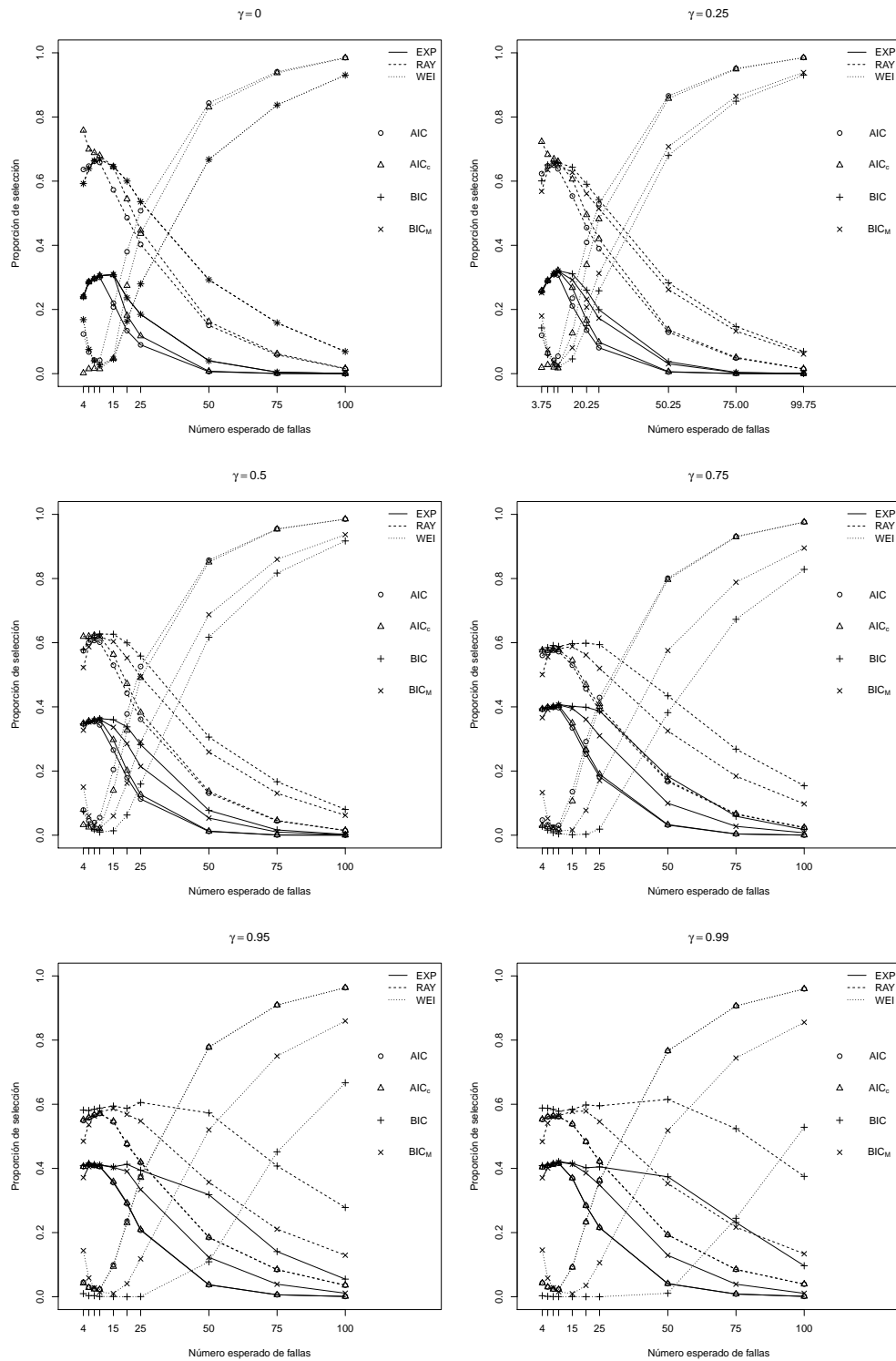


Figura 5.4: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 1.5.

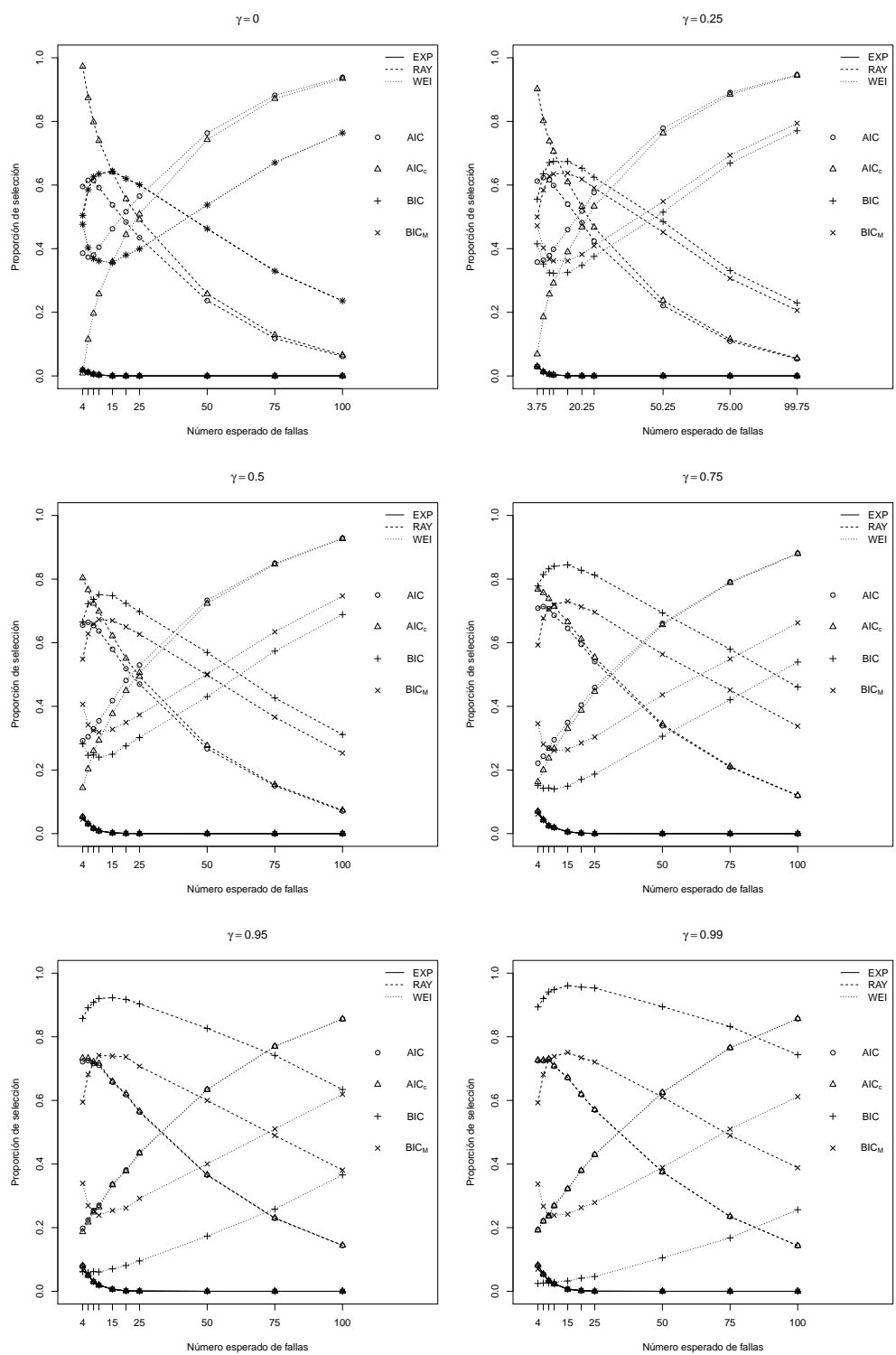


Figura 5.5: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 2.5.

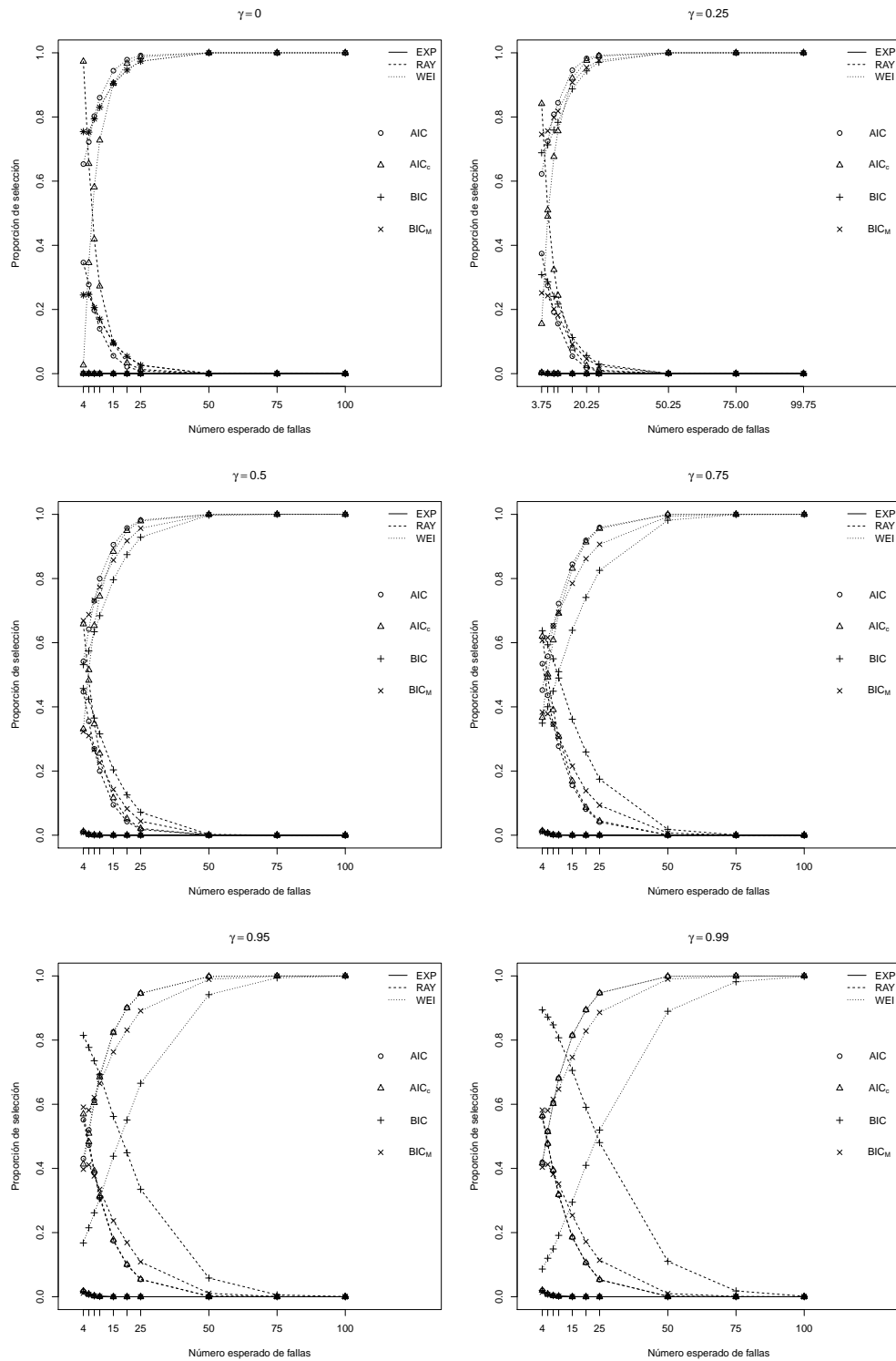


Figura 5.6: Proporción de veces que el AIC,  $AIC_c$ , BIC y  $BIC_M$  (indicado por el tipo de símbolo usado) eligen a cada una de las tres distribuciones candidatas (indicadas por el tipo de línea usada), siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 3.5.

tiene como en seguida se explica.

Consideremos una muestra con censura aleatoria de la distribución  $\text{EXP}(\eta)$ ,  $X_1, \dots, X_n \stackrel{iid}{\sim} \text{EXP}(\eta)$ , y denotemos por  $\text{AIC}_1$ ,  $\text{AIC}_2$  y  $\text{AIC}_3$  el AIC de los modelos EXP, RAY y WEI, respectivamente. Notemos que el AIC elije al modelo EXP como el mejor si y sólo si

$$\{\text{AIC}_1 < \text{AIC}_2, \text{AIC}_1 < \text{AIC}_3\} = \left\{ \ell_1(\hat{\eta}|\mathbf{t}) > \ell_2(\hat{\sigma}|\mathbf{t}), -2 \log \left( \frac{L_1(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})} \right) < 2 \right\}.$$

Sin embargo, por Lema 5.C.3  $\lim_{n \rightarrow \infty} P(\ell_1(\hat{\eta}|\mathbf{t}) > \ell_2(\hat{\sigma}|\mathbf{t})) = 1$ , entonces por Lema 5.C.1

$$\lim_{n \rightarrow \infty} P(\text{AIC}_1 < \text{AIC}_2, \text{AIC}_1 < \text{AIC}_3) = \lim_{n \rightarrow \infty} P \left[ -2 \log \left( \frac{L_1(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})} \right) < 2 \right],$$

y así, asintóticamente la probabilidad de que el AIC elija al modelo EXP como el mejor es equivalente a que la prueba de razón de verosimilitud no rechace la hipótesis nula  $H_0 : \beta = 1$  vs  $H_1 : \beta \neq 1$ , con error tipo I igual a  $0.1573 = 1 - 0.8427 = 1 - P(\chi_1^2 < 2)$ . De aquí, haciendo los ajustes correspondientes, para algunos casos se puede derivar lo que Collett (2003, página 81) sugiere en la modificación del término de penalización en el AIC para tener un nivel deseado de la PSC. Usando el mismo Lema 5.C.3, lo mismo se puede obtener si los datos son generados bajo el modelo RAY. Para cuando los datos son generados del modelo WEI con  $\beta$  diferente de 1 y 2, como sugerido en los resultados de simulación anteriores, también puede verse que en todos los criterios, la probabilidad de selección correcta converge a 1. Las tendencias de los restantes criterios también pueden verse asintóticamente haciendo pequeñas modificaciones a los argumentos usados en el AIC. Éste mismo comportamiento asintótico también se puede ver (Lema 5.C.4) bajo el mecanismo de censura II que se encuentra en la Sección 5.A.2.

## 5.4. Conclusiones

Con un conjunto de distribuciones pequeño pero frecuentemente usados en el análisis de datos de vida, por medio de simulación se estudió el desempeño de 4 criterios de selección de modelos basados en verosimilitud. Específicamente éste desempeño fue evaluado usando a las distribuciones exponencial, Rayleigh y Weibull como el conjunto de distribuciones candidatas, diferentes valores del número nominal de fallas en una muestra y diferentes fracciones de censura aleatoria por la derecha fueron estudiadas.

En general, en todos los criterios se observó un mejoramiento en la PSC a medida que el número nominal de fallas crecía. Además, se observó que no existe un criterio que muestre un comportamiento uniformemente mejor que el resto ya que bajo las distribuciones consideradas, éste depende de la distribución que genera los datos y de la combinación del número nominal de fallas ( $r$ ) y de la fracción de censura presente ( $\gamma$ ).

Cuando los datos fueron generados del modelo EXP o RAY, cada uno de los criterios mostró un comportamiento muy similar, pero cuando los datos fueron generados de éste último (RAY), en los primeros cuatro valores de  $r$  experimentados ( $r = 4, 6, 8$  y  $10$ ), la diferencia entre el que mejor se desempeña con el que se encuentra en segunda posición, resultó mucho mayor que cuando los datos fueron generados del modelo EXP. En éstos primeros 4 valores de  $r$  el  $AIC_c$  y el BIC son los que mejor se desempeñaron en regiones de valores pequeños y valores grandes de  $\gamma$ , respectivamente. Aquí el efecto de la fracción de censura fue moderadamente positivo en la PSC en todos criterios, aunque fue siendo cada vez más ligero a medida que  $r$  crecía.

A partir de  $r = 15$  el BIC comenzó a mostrar un comportamiento uniformemente mejor que el resto, seguido por el  $BIC_M$ , el cual difirió cada vez menos respecto al primero (BIC) a medida que la fracción de censura era pequeña y  $r$  crecía. Estos últimos dos criterios tendieron a elegir el modelo correctamente con probabilidades cada más cercanas a 1 conforme  $r$  aumentaba, mientras que las probabilidades de selección correcta del AIC y  $AIC_c$  tendieron a estabilizarse alrededor de 0.84. Estos comportamientos también fueron explicados asintóticamente para dos mecanismos de censura aleatoria por la derecha.

Cuando los datos fueron simulados del modelo WEI con parámetro de forma  $\beta = 0.5$  el comportamiento de los criterios fue un poco diferente. Aquí para los valores más chicos de  $r$  (4 y 6), el  $BIC_M$  fue el que mostró un mejor desempeño que el resto pero con probabilidades de selección correcta menores que 0.5 para valores de  $\gamma \leq 0.10$  en  $r = 4$ , aunque éstas aumentaron a medida que  $\gamma$  lo hacía. A partir de  $r = 10$  el AIC es el que tuvo mayores PSC, pero no fueron muy superiores a las obtenidas en el resto (en particular con el  $BIC_M$  en los primeros 9 valores de  $\gamma$  y con el  $AIC_c$  en el resto) de los criterios con excepción del  $AIC_c$  y BIC donde tal superioridad fue cada vez mayor conforme  $\gamma$  decrecía y decrecía, respectivamente, aunque aún más en éste último.

Cuando se simularon los datos del modelo WEI con parámetro de forma 1.5, todos los criterios tuvieron un mal desempeño en los primeros 7 valores de  $r$  con probabilidades de selección correcta que fueron creciendo a valores menores o iguales a 0.50, donde en los primeros dos el  $BIC_M$  fue el que tuvo el mejor desempeño, mientras que a partir del tercer valor el AIC se comenzó a desempeñarse mejor de forma uniforme sobre el resto. En los últimos dos valores de  $r$  se obtuvo un desempeño bastante bueno de los criterios (pero

mucho más con el AIC y  $AIC_c$ ) aunque no tanto con los criterios  $BIC_M$  y BIC, donde éste último continuo siendo afectado por la fracción de censura, ya que por ejemplo, con  $r = 100$ , la PSC baja de 0.93 a solo 0.5283 cuando se alcanzó un valor de  $\gamma = 0.99$ .

Al simular los datos de la distribución WEI pero con el parámetro de forma igual a 2.5, el desempeño de los criterios mejoró hasta el valor de  $r = 25$ , pero para valores de  $r$  iguales o mayores de 50 se obtuvieron probabilidades de selección correcta más pequeñas que las correspondientes al escenario anterior ( $\beta = 1.5$ ).

Al tener una mayor separación en las distribuciones candidatas cuando los datos fueron simulados con  $\beta = 3.5$ , el comportamiento de todos los criterios mejoró, y aún más a medida que  $r$  aumentaba. También se observó un efecto negativo en la PSC de  $\gamma$  en todos los criterios aunque el más afectado fue el BIC, donde por ejemplo en  $r = 20$ , se tenía una PSC igual a 0.96464 en  $\gamma = 0$  y al alcanzar la mayor fracción de censura (0.99) la PSC disminuyó a un valor menor que la mitad (0.4096). En los último valores de  $r$  considerados el efecto de  $\gamma$  resultó casi despreciable, y las probabilidades de selección correcta fueron muy cercanas a 1 en todos los criterios.

Así, de los resultados aquí obtenidos no es es fácil elegir el mejor criterio ya que esto puede depender de las distribuciones candidatas, el número nominal de fallas ( $r$ ), y también de la fracción de censura presente. A pesar de esto, en general al parecer el  $BIC_M$  es el que resultó tener un comportamiento bastante regular después de  $r = 15$  con excepción donde los datos fueron generados de la distribución WEI con parámetro de forma 1.5 y 2.5. Sin embargo, se deben tener precauciones en el uso de estos índices como criterios de selección de modelos de datos de vida con presencia de censura y un número de fallas pequeño en la muestras, ya que algunas veces nos pueden llevar a elegir un modelo incorrecto y en consecuencia obtener conclusiones erróneas. Así, estos criterios pueden y debería usarse solamente como una herramienta más en la elección del modelo, donde otros aspectos de la naturaleza del fenómeno debería también tomarse en cuenta.

La comparación anterior podría extenderse a familias de distribuciones candidatas más grande, por ejemplo, a la familia de la distribución gama generalizada donde además de ésta se incluirían el modelo lognormal, Weibull, gama, exponencial y Rayleigh. Sin embargo, de acuerdo a lo observado en este estudio, obtener una conclusión a cerca de cual de los criterios resulta en un comportamiento regular en comparación al resto sería más complicado, además de que se deberían usar valores del número de fallas nominal superiores a los aquí usados. También podría extenderse el estudio en la elección de variables explicativas en un modelo de regresión de confiabilidad.

# Apéndices

## 5.A. Mecanismos de generación de censura aleatoria por la derecha

### 5.A.1. Mecanismo I

Para ver que el mecanismo de generación de datos utilizado en la Sección 5.3.1 en efecto genera una fracción de censura  $\gamma$  esperada, sean  $U$  y  $U_k$  variables aleatorias independientes con distribución  $U(0, 1)$  y  $U(0, k)$ , respectivamente. Observemos que

$$\begin{aligned}\mathbb{P}(U_k < U) &= \int_0^k \mathbb{P}(U > u) f_{U_k}(u) du = \int_0^k \mathbb{P}(U > u) \frac{1}{k} du \\ &= \begin{cases} \frac{1}{k} \int_0^k (1-u) du & \text{si } k \leq 1, \\ \frac{1}{k} \int_0^1 (1-u) du & \text{si } k > 1, \end{cases} \\ &= \begin{cases} 1 - \frac{k}{2} & \text{si } k \leq 1, \\ \frac{1}{2k} & \text{si } k > 1. \end{cases}\end{aligned}$$

De aquí, si  $U$  denota el tiempo de vida de una unidad y  $U_k$  su tiempo de censura, entonces el valor de  $k$  que satisface  $\mathbb{P}(U_k \leq U) = 1 - \mathbb{P}(U_k > U) = \gamma$ , es decir, el  $k$  con el cual se obtiene una muestra con  $100\gamma\%$  de censura aleatoria está dado por

$$k_\gamma = \begin{cases} 2(1 - \gamma) & \text{si } 0.5 \leq \gamma \leq 1, \\ \frac{1}{2\gamma} & \text{si } 0 \leq \gamma < 0.5. \end{cases}$$

Si  $T$  es el tiempo a la falla de interés con función de distribución  $F_T$  y función cuantil  $F_T^{-1}(\cdot)$ , como  $T \stackrel{d}{=} F_T^{-1}(U)$ ,  $\mathbb{P}[F_T^{-1}(U_{k_\gamma}) \leq F_T^{-1}(U)] = \mathbb{P}[U_{k_\gamma} \leq U] = \gamma$  por ser  $F_T^{-1}$  una función no decreciente, y  $F_T^{-1}(U_{k_\gamma})$  y  $F_T^{-1}(U)$  son independientes ya que  $U_{k_\gamma}$  y  $U$

lo son, entonces a partir de muestras con censura aleatoria por la derecha uniformes se pueden generar muestras también con censura aleatoria por la derecha de cualquier variable aleatoria continua  $T$ . El algoritmo es el siguiente:

1. Simular  $U_{k_\gamma} \sim U(0, k_\gamma)$  y  $U \sim U(0, 1)$ .
2. Si  $U_{k_\gamma} < U$ ,  $\delta = 0$  y tomar  $T^* = F_T^{-1}(U_{k_\gamma})$ , pero si  $U < U_{k_\gamma}$ ,  $\delta = 1$  y tomar  $T^* = F_T^{-1}(U)$ , es decir,  $T^* = \min\{F_T^{-1}(U_{k_\gamma}), F_T^{-1}(U)\}$ . donde  $C := F_T^{-1}(U_{k_\gamma})$  es el tiempo de censura.

### 5.A.2. Mecanismo II

Otra forma de generar los datos con una fracción de censura deseada  $\gamma$  es la propuesta por Koziol y Green (1976) y que a continuación se describe.

Si  $T$  denota el tiempo a la falla de interés con función de distribución  $F_T(\cdot)$ , entonces ésta propuesta consiste en tomar al tiempo de censura  $C$  independiente con función de distribución  $F_C(c) = 1 - (1 - F_T(c))^p$ , donde  $p$  es una constante positiva conocida que está relacionada directamente a la fracción de censura  $\gamma$  deseada. Esto por que si  $\delta = I_{(T \leq C)}$  es el indicador de censura entonces

$$\begin{aligned}
 E(\delta) &= P(T \leq C) = \int_{-\infty}^{\infty} P(T \leq c) dF_C(c) \\
 &= p \int_{-\infty}^{\infty} F_T(c) [1 - F_T(c)]^{p-1} dF_T(c) \\
 &= p \int_0^1 (1 - u)u^{p-1} du = p \left( \frac{u^p}{p} - \frac{u^{p+1}}{p+1} \right) \Big|_0^1 \\
 &= \frac{1}{p+1},
 \end{aligned}$$

de donde  $p = \frac{\gamma}{1 - \gamma}$ .

### 5.B. Cuadros de resultados de simulación



$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
4	AIC	EXP	0.5481	0.5381	0.5293	0.5185	0.5561	0.6167	0.6380	0.6583	0.6769	0.6757	0.6697
		RAY	0.3873	0.3919	0.4019	0.3990	0.3668	0.3194	0.2920	0.2744	0.2582	0.2554	0.2597
		WEI	0.0646	0.0700	0.0688	0.0825	0.0771	0.0639	0.0699	0.0673	0.0649	0.0689	0.0706
	AIC <sub>c</sub>	EXP	0.5705	0.5643	0.5533	0.5453	0.5863	0.6471	0.6698	0.6859	0.6962	0.6805	0.6704
		RAY	0.4293	0.4352	0.4459	0.4532	0.4056	0.3363	0.3042	0.2827	0.2614	0.2565	0.2599
		WEI	0.0002	0.0005	0.0008	0.0015	0.0080	0.0166	0.0259	0.0314	0.0424	0.0630	0.0697
	BIC	EXP	0.5323	0.5232	0.5131	0.5042	0.5488	0.6157	0.6406	0.6678	0.6991	0.7211	0.7237
		RAY	0.3689	0.3734	0.3838	0.3788	0.3574	0.3187	0.2929	0.2766	0.2620	0.2642	0.2700
		WEI	0.0988	0.1034	0.1031	0.1170	0.0938	0.0656	0.0665	0.0556	0.0389	0.0147	0.0063
	BIC <sub>M</sub>	EXP	0.5323	0.5228	0.5108	0.4987	0.5329	0.5823	0.5894	0.6000	0.6121	0.6017	0.5992
		RAY	0.3689	0.3729	0.3804	0.3696	0.3389	0.2983	0.2675	0.2504	0.2298	0.2195	0.2205
		WEI	0.0988	0.1043	0.1088	0.1317	0.1282	0.1194	0.1431	0.1496	0.1580	0.1789	0.1803
6	AIC	EXP	0.6647	0.6584	0.6520	0.6801	0.6836	0.7040	0.7211	0.7326	0.7472	0.7346	0.7317
		RAY	0.2902	0.2998	0.2997	0.2737	0.2695	0.2411	0.2187	0.2066	0.1906	0.2025	0.2027
		WEI	0.0451	0.0418	0.0483	0.0462	0.0469	0.0549	0.0601	0.0608	0.0621	0.0629	0.0657
	AIC <sub>c</sub>	EXP	0.6934	0.6847	0.6811	0.7072	0.7069	0.7279	0.7442	0.7558	0.7623	0.7369	0.7326
		RAY	0.3005	0.3093	0.3127	0.2810	0.2779	0.2447	0.2222	0.2088	0.1920	0.2026	0.2027
		WEI	0.0061	0.0060	0.0062	0.0118	0.0152	0.0274	0.0335	0.0354	0.0456	0.0605	0.0647
	BIC	EXP	0.6585	0.6530	0.6470	0.6794	0.6852	0.7130	0.7352	0.7566	0.7790	0.7829	0.7888
		RAY	0.2890	0.2980	0.2971	0.2734	0.2701	0.2422	0.2208	0.2088	0.1931	0.2051	0.2060
		WEI	0.0525	0.0490	0.0559	0.0472	0.0447	0.0448	0.0439	0.0346	0.0278	0.0120	0.0052
	BIC <sub>M</sub>	EXP	0.6585	0.6528	0.6455	0.6762	0.6771	0.6960	0.7086	0.7134	0.7283	0.7107	0.7093
		RAY	0.2890	0.2980	0.2956	0.2717	0.2659	0.2366	0.2157	0.2019	0.1852	0.1954	0.1929
		WEI	0.0525	0.0492	0.0589	0.0521	0.0570	0.0674	0.0756	0.0847	0.0865	0.0940	0.0977
8	AIC	EXP	0.7487	0.7411	0.7288	0.7373	0.7566	0.7598	0.7712	0.7883	0.7856	0.7758	0.7800
		RAY	0.2111	0.2165	0.2326	0.2209	0.1951	0.1932	0.1709	0.1486	0.1513	0.1606	0.1538
		WEI	0.0402	0.0424	0.0386	0.0418	0.0483	0.0470	0.0579	0.0631	0.0631	0.0636	0.0661
	AIC <sub>c</sub>	EXP	0.7731	0.7682	0.7542	0.7606	0.7817	0.7789	0.7894	0.8052	0.7976	0.7773	0.7801
		RAY	0.2146	0.2194	0.2362	0.2233	0.1981	0.1958	0.1740	0.1505	0.1524	0.1607	0.1539
		WEI	0.0123	0.0124	0.0096	0.0161	0.0202	0.0253	0.0366	0.0443	0.0500	0.0620	0.0659
	BIC	EXP	0.7510	0.7422	0.7311	0.7415	0.7691	0.7744	0.7923	0.8173	0.8252	0.8270	0.8409
		RAY	0.2111	0.2166	0.2330	0.2213	0.1974	0.1955	0.1740	0.1509	0.1531	0.1637	0.1553
		WEI	0.0379	0.0412	0.0359	0.0372	0.0335	0.0301	0.0337	0.0318	0.0217	0.0093	0.0038
	BIC <sub>M</sub>	EXP	0.7510	0.7418	0.7299	0.7394	0.7611	0.7615	0.7729	0.7891	0.7862	0.7721	0.7769
		RAY	0.2111	0.2166	0.2323	0.2208	0.1968	0.1927	0.1731	0.1494	0.1508	0.1594	0.1521
		WEI	0.0379	0.0416	0.0378	0.0398	0.0421	0.0458	0.0540	0.0615	0.0630	0.0685	0.0709
10	AIC	EXP	0.7881	0.7888	0.8073	0.7868	0.7834	0.7999	0.8079	0.8135	0.8087	0.8072	0.8014
		RAY	0.1671	0.1658	0.1456	0.1634	0.1611	0.1338	0.1173	0.1114	0.1135	0.1219	0.1258
		WEI	0.0448	0.0454	0.0471	0.0498	0.0555	0.0663	0.0748	0.0751	0.0778	0.0709	0.0728
	AIC <sub>c</sub>	EXP	0.8145	0.8133	0.8311	0.8132	0.8092	0.8285	0.8347	0.8333	0.8209	0.8099	0.8019
		RAY	0.1700	0.1686	0.1502	0.1674	0.1662	0.1415	0.1227	0.1159	0.1169	0.1221	0.1259
		WEI	0.0155	0.0181	0.0187	0.0194	0.0246	0.0300	0.0426	0.0508	0.0622	0.0680	0.0722
	BIC	EXP	0.7998	0.7984	0.8212	0.8022	0.8039	0.8333	0.8477	0.8570	0.8583	0.8689	0.8686
		RAY	0.1689	0.1680	0.1497	0.1666	0.1659	0.1417	0.1243	0.1182	0.1192	0.1241	0.1291
		WEI	0.0313	0.0336	0.0291	0.0312	0.0302	0.0250	0.0280	0.0248	0.0225	0.0070	0.0023
	BIC <sub>M</sub>	EXP	0.7998	0.7984	0.8208	0.7999	0.7975	0.8205	0.8282	0.8314	0.8244	0.8189	0.8131
		RAY	0.1689	0.1680	0.1496	0.1661	0.1654	0.1407	0.1228	0.1171	0.1185	0.1234	0.1274
		WEI	0.0313	0.0336	0.0296	0.0340	0.0371	0.0388	0.0490	0.0515	0.0571	0.0577	0.0595

Cuadro 5.B.1: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
15	AIC	EXP	0.8254	0.8117	0.8150	0.8219	0.8152	0.8223	0.8266	0.8351	0.8385	0.8379	0.8395
		RAY	0.0609	0.0669	0.0606	0.0550	0.0570	0.0476	0.0439	0.0428	0.0490	0.0617	0.0596
		WEI	0.1137	0.1214	0.1244	0.1231	0.1278	0.1301	0.1295	0.1221	0.1125	0.1004	0.1009
	AIC <sub>c</sub>	EXP	0.8829	0.8686	0.8722	0.8725	0.8583	0.8581	0.8538	0.8552	0.8491	0.8397	0.8403
		RAY	0.0805	0.0880	0.0771	0.0708	0.0715	0.0566	0.0516	0.0473	0.0518	0.0625	0.0596
		WEI	0.0366	0.0434	0.0507	0.0567	0.0702	0.0853	0.0946	0.0975	0.0991	0.0978	0.1001
	BIC	EXP	0.8834	0.8697	0.8809	0.8872	0.8853	0.9022	0.9054	0.9176	0.9186	0.9182	0.9261
		RAY	0.0807	0.0883	0.0803	0.0761	0.0799	0.0698	0.0671	0.0604	0.0621	0.0748	0.0714
		WEI	0.0359	0.0420	0.0388	0.0367	0.0348	0.0280	0.0275	0.0220	0.0193	0.0070	0.0025
	BIC <sub>M</sub>	EXP	0.8834	0.8694	0.8778	0.8826	0.8729	0.8797	0.8804	0.8891	0.8873	0.8817	0.8818
		RAY	0.0807	0.0882	0.0794	0.0738	0.0756	0.0635	0.0598	0.0557	0.0597	0.0733	0.0703
		WEI	0.0359	0.0424	0.0428	0.0436	0.0515	0.0568	0.0598	0.0552	0.0530	0.0450	0.0479
20	AIC	EXP	0.8244	0.8269	0.8233	0.8265	0.8298	0.8271	0.8285	0.8339	0.8425	0.8436	0.8371
		RAY	0.0204	0.0226	0.0203	0.0194	0.0155	0.0173	0.0140	0.0135	0.0173	0.0261	0.0297
		WEI	0.1552	0.1505	0.1564	0.1541	0.1547	0.1556	0.1575	0.1526	0.1402	0.1303	0.1332
	AIC <sub>c</sub>	EXP	0.8688	0.8703	0.8661	0.8654	0.8594	0.8520	0.8498	0.8510	0.8535	0.8453	0.8375
		RAY	0.0266	0.0290	0.0266	0.0255	0.0195	0.0195	0.0159	0.0148	0.0187	0.0262	0.0297
		WEI	0.1046	0.1007	0.1073	0.1091	0.1211	0.1285	0.1343	0.1342	0.1278	0.1285	0.1328
	BIC	EXP	0.9055	0.9031	0.9072	0.9107	0.9174	0.9271	0.9369	0.9419	0.9480	0.9494	0.9518
		RAY	0.0349	0.0355	0.0346	0.0360	0.0307	0.0351	0.0309	0.0324	0.0362	0.0457	0.0456
		WEI	0.0596	0.0614	0.0582	0.0533	0.0519	0.0378	0.0322	0.0257	0.0158	0.0049	0.0026
	BIC <sub>M</sub>	EXP	0.9055	0.9022	0.9044	0.9031	0.9032	0.9034	0.9094	0.9087	0.9128	0.9098	0.9097
		RAY	0.0349	0.0353	0.0338	0.0331	0.0268	0.0273	0.0227	0.0234	0.0281	0.0410	0.0429
		WEI	0.0596	0.0625	0.0618	0.0638	0.0700	0.0693	0.0679	0.0679	0.0591	0.0492	0.0474
25	AIC	EXP	0.8275	0.8281	0.8323	0.8304	0.8267	0.8296	0.8358	0.8431	0.8446	0.8399	0.8452
		RAY	0.0067	0.0077	0.0068	0.0059	0.0062	0.0049	0.0044	0.0048	0.0069	0.0099	0.0108
		WEI	0.1658	0.1642	0.1609	0.1637	0.1671	0.1655	0.1598	0.1521	0.1485	0.1502	0.1440
	AIC <sub>c</sub>	EXP	0.8628	0.8616	0.8655	0.8605	0.8551	0.8524	0.8541	0.8566	0.8517	0.8410	0.8455
		RAY	0.0088	0.0096	0.0095	0.0076	0.0066	0.0057	0.0051	0.0054	0.0077	0.0101	0.0108
		WEI	0.1284	0.1288	0.1250	0.1319	0.1383	0.1419	0.1408	0.1380	0.1406	0.1489	0.1437
	BIC	EXP	0.9142	0.9128	0.9223	0.9271	0.9297	0.9426	0.9441	0.9570	0.9624	0.9656	0.9704
		RAY	0.0155	0.0147	0.0141	0.0133	0.0129	0.0116	0.0128	0.0144	0.0192	0.0286	0.0282
		WEI	0.0703	0.0725	0.0636	0.0596	0.0574	0.0458	0.0431	0.0286	0.0184	0.0058	0.0014
	BIC <sub>M</sub>	EXP	0.9142	0.9126	0.9203	0.9218	0.9163	0.9215	0.9199	0.9261	0.9252	0.9188	0.9296
		RAY	0.0155	0.0147	0.0139	0.0126	0.0110	0.0086	0.0084	0.0093	0.0129	0.0214	0.0215
		WEI	0.0703	0.0727	0.0658	0.0656	0.0727	0.0699	0.0717	0.0646	0.0619	0.0598	0.0489
50	AIC	EXP	0.8367	0.8363	0.8321	0.8357	0.8406	0.8369	0.8364	0.8385	0.8388	0.8444	0.8430
		RAY	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		WEI	0.1633	0.1637	0.1679	0.1643	0.1593	0.1631	0.1636	0.1615	0.1612	0.1556	0.1569
	AIC <sub>c</sub>	EXP	0.8539	0.8526	0.8458	0.8494	0.8533	0.8484	0.8451	0.8455	0.8429	0.8453	0.8432
		RAY	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		WEI	0.1461	0.1474	0.1542	0.1506	0.1466	0.1516	0.1549	0.1545	0.1571	0.1547	0.1567
	BIC	EXP	0.9499	0.9484	0.9481	0.9494	0.9582	0.9618	0.9665	0.9696	0.9763	0.9908	0.9953
		RAY	0.0002	0.0000	0.0002	0.0002	0.0001	0.0000	0.0000	0.0001	0.0004	0.0020	0.0030
		WEI	0.0499	0.0516	0.0517	0.0504	0.0417	0.0382	0.0335	0.0303	0.0233	0.0072	0.0017
	BIC <sub>M</sub>	EXP	0.9499	0.9475	0.9469	0.9455	0.9502	0.9483	0.9500	0.9494	0.9481	0.9543	0.9531
		RAY	0.0002	0.0000	0.0002	0.0002	0.0001	0.0000	0.0000	0.0000	0.0001	0.0007	0.0006
		WEI	0.0499	0.0525	0.0529	0.0543	0.0497	0.0517	0.0500	0.0506	0.0518	0.0450	0.0463

Cuadro 5.B.2: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
75	AIC	EXP	0.8314	0.8330	0.8454	0.8329	0.8358	0.8460	0.8387	0.8386	0.8402	0.8460	0.8478
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.1686	0.1670	0.1546	0.1671	0.1642	0.1540	0.1613	0.1614	0.1598	0.1540	0.1522
	AIC <sub>c</sub>	EXP	0.8423	0.8443	0.8567	0.8437	0.8442	0.8525	0.8434	0.8435	0.8440	0.8466	0.8479
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.1577	0.1557	0.1433	0.1563	0.1558	0.1475	0.1566	0.1565	0.1560	0.1534	0.1521
	BIC	EXP	0.9590	0.9598	0.9615	0.9635	0.9666	0.9736	0.9731	0.9760	0.9846	0.9923	0.9966
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
		WEI	0.0410	0.0402	0.0385	0.0365	0.0334	0.0264	0.0269	0.0240	0.0154	0.0076	0.0033
	BIC <sub>M</sub>	EXP	0.9590	0.9594	0.9609	0.9611	0.9596	0.9636	0.9595	0.9602	0.9636	0.9616	0.9619
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.0410	0.0406	0.0391	0.0389	0.0404	0.0364	0.0405	0.0398	0.0364	0.0384	0.0381
100	AIC	EXP	0.8438	0.8381	0.8331	0.8342	0.8383	0.8330	0.8428	0.8378	0.8401	0.8361	0.8447
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.1562	0.1619	0.1669	0.1658	0.1617	0.1670	0.1572	0.1622	0.1599	0.1639	0.1553
	AIC <sub>c</sub>	EXP	0.8518	0.8462	0.8427	0.8417	0.8446	0.8382	0.8474	0.8413	0.8418	0.8366	0.8449
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.1482	0.1538	0.1573	0.1583	0.1554	0.1618	0.1526	0.1587	0.1582	0.1634	0.1551
	BIC	EXP	0.9703	0.9689	0.9667	0.9679	0.9731	0.9731	0.9775	0.9795	0.9844	0.9929	0.9978
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.0297	0.0311	0.0333	0.0321	0.0269	0.0269	0.0225	0.0205	0.0156	0.0071	0.0022
	BIC <sub>M</sub>	EXP	0.9703	0.9687	0.9648	0.9653	0.9669	0.9634	0.9680	0.9656	0.9676	0.9684	0.9681
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.0297	0.0313	0.0352	0.0347	0.0331	0.0366	0.0320	0.0344	0.0324	0.0316	0.0319

Cuadro 5.B.3: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución EXP.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
4	AIC	EXP	0.0759	0.0746	0.0817	0.0829	0.0922	0.1131	0.1376	0.1614	0.1736	0.1897	0.1831
		RAY	0.6815	0.6731	0.6714	0.6670	0.6804	0.6966	0.6967	0.6974	0.7093	0.7051	0.7158
		WEI	0.2426	0.2523	0.2469	0.2502	0.2274	0.1903	0.1657	0.1413	0.1171	0.1052	0.1011
	AIC <sub>c</sub>	EXP	0.0759	0.0746	0.0817	0.0829	0.0923	0.1132	0.1376	0.1615	0.1736	0.1897	0.1831
		RAY	0.9196	0.9202	0.9126	0.9099	0.8691	0.8119	0.7858	0.7620	0.7414	0.7114	0.7174
		WEI	0.0045	0.0052	0.0057	0.0072	0.0386	0.0749	0.0766	0.0766	0.0850	0.0989	0.0995
	BIC	EXP	0.0758	0.0746	0.0816	0.0826	0.0922	0.1131	0.1376	0.1614	0.1736	0.1901	0.1836
		RAY	0.6068	0.6040	0.6007	0.5948	0.6380	0.6920	0.7035	0.7181	0.7471	0.7815	0.8059
		WEI	0.3174	0.3214	0.3176	0.3226	0.2699	0.1950	0.1588	0.1206	0.0793	0.0285	0.0105
	BIC <sub>M</sub>	EXP	0.0758	0.0746	0.0816	0.0824	0.0899	0.1088	0.1262	0.1490	0.1581	0.1688	0.1645
		RAY	0.6068	0.6023	0.5916	0.5745	0.5905	0.6246	0.6175	0.6118	0.6223	0.6099	0.6124
		WEI	0.3174	0.3231	0.3267	0.3431	0.3197	0.2666	0.2563	0.2392	0.2195	0.2213	0.2231
6	AIC	EXP	0.0711	0.0655	0.0676	0.0676	0.0706	0.0853	0.1114	0.1271	0.1367	0.1559	0.1518
		RAY	0.7402	0.7390	0.7322	0.7463	0.7478	0.7508	0.7483	0.7465	0.7548	0.7445	0.7494
		WEI	0.1887	0.1955	0.2002	0.1861	0.1815	0.1639	0.1402	0.1264	0.1085	0.0996	0.0987
	AIC <sub>c</sub>	EXP	0.0711	0.0655	0.0676	0.0676	0.0707	0.0853	0.1117	0.1271	0.1367	0.1559	0.1518
		RAY	0.8786	0.8828	0.8779	0.8647	0.8486	0.8261	0.7997	0.7836	0.7761	0.7492	0.7508
		WEI	0.0503	0.0517	0.0545	0.0677	0.0807	0.0885	0.0886	0.0893	0.0872	0.0949	0.0973
	BIC	EXP	0.0711	0.0655	0.0676	0.0676	0.0706	0.0853	0.1117	0.1271	0.1367	0.1560	0.1520
		RAY	0.7220	0.7168	0.7128	0.7417	0.7543	0.7779	0.7792	0.7845	0.8071	0.8251	0.8411
		WEI	0.2069	0.2177	0.2196	0.1907	0.1751	0.1368	0.1091	0.0884	0.0562	0.0189	0.0068
	BIC <sub>M</sub>	EXP	0.0711	0.0655	0.0676	0.0676	0.0706	0.0849	0.1106	0.1251	0.1341	0.1526	0.1482
		RAY	0.7220	0.7154	0.7061	0.7278	0.7219	0.7253	0.7243	0.7236	0.7303	0.7109	0.7168
		WEI	0.2069	0.2191	0.2263	0.2046	0.2074	0.1898	0.1650	0.1513	0.1356	0.1366	0.1349
8	AIC	EXP	0.0522	0.0568	0.0545	0.0529	0.0563	0.0644	0.0843	0.1038	0.1131	0.1239	0.1243
		RAY	0.7789	0.7766	0.7692	0.7705	0.7811	0.7802	0.7766	0.7804	0.7807	0.7764	0.7867
		WEI	0.1689	0.1666	0.1763	0.1766	0.1626	0.1554	0.1391	0.1158	0.1062	0.0997	0.0889
	AIC <sub>c</sub>	EXP	0.0522	0.0568	0.0545	0.0531	0.0571	0.0650	0.0847	0.1042	0.1132	0.1239	0.1243
		RAY	0.8744	0.8732	0.8651	0.8607	0.8492	0.8320	0.8138	0.8076	0.7991	0.7798	0.7878
		WEI	0.0734	0.0700	0.0804	0.0862	0.0937	0.1030	0.1015	0.0882	0.0877	0.0963	0.0878
	BIC	EXP	0.0522	0.0568	0.0545	0.0531	0.0571	0.0650	0.0847	0.1042	0.1132	0.1240	0.1244
		RAY	0.7856	0.7821	0.7750	0.7904	0.8128	0.8193	0.8206	0.8293	0.8412	0.8595	0.8690
		WEI	0.1622	0.1611	0.1705	0.1565	0.1301	0.1157	0.0947	0.0665	0.0456	0.0165	0.0066
	BIC <sub>M</sub>	EXP	0.0522	0.0568	0.0545	0.0531	0.0571	0.0650	0.0847	0.1040	0.1128	0.1230	0.1239
		RAY	0.7856	0.7814	0.7713	0.7790	0.7875	0.7807	0.7778	0.7835	0.7808	0.7716	0.7827
		WEI	0.1622	0.1618	0.1742	0.1679	0.1554	0.1543	0.1375	0.1125	0.1064	0.1055	0.0933
10	AIC	EXP	0.0398	0.0439	0.0347	0.0425	0.0433	0.0436	0.0630	0.0816	0.0948	0.1043	0.1072
		RAY	0.7986	0.7984	0.8133	0.7982	0.7938	0.8027	0.8072	0.8025	0.8028	0.8005	0.7952
		WEI	0.1616	0.1577	0.1520	0.1593	0.1629	0.1537	0.1298	0.1159	0.1024	0.0952	0.0976
	AIC <sub>c</sub>	EXP	0.0410	0.0455	0.0377	0.0443	0.0458	0.0470	0.0657	0.0828	0.0953	0.1046	0.1073
		RAY	0.8745	0.8723	0.8767	0.8630	0.8522	0.8482	0.8367	0.8261	0.8168	0.8020	0.7960
		WEI	0.0845	0.0822	0.0856	0.0927	0.1020	0.1048	0.0976	0.0911	0.0879	0.0934	0.0967
	BIC	EXP	0.0410	0.0453	0.0375	0.0441	0.0458	0.0472	0.0659	0.0830	0.0962	0.1053	0.1080
		RAY	0.8261	0.8205	0.8424	0.8289	0.8360	0.8611	0.8598	0.8631	0.8688	0.8823	0.8872
		WEI	0.1329	0.1342	0.1201	0.1270	0.1182	0.0917	0.0743	0.0539	0.0350	0.0124	0.0048
	BIC <sub>M</sub>	EXP	0.0410	0.0453	0.0374	0.0440	0.0456	0.0468	0.0658	0.0830	0.0959	0.1052	0.1077
		RAY	0.8261	0.8197	0.8384	0.8204	0.8137	0.8268	0.8270	0.8193	0.8162	0.8120	0.8080
		WEI	0.1329	0.1350	0.1242	0.1356	0.1407	0.1264	0.1072	0.0977	0.0879	0.0828	0.0843

Cuadro 5.B.4: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
15	AIC	EXP	0.0104	0.0106	0.0089	0.0086	0.0109	0.0166	0.0264	0.0371	0.0535	0.0568	0.0616
		RAY	0.8254	0.8117	0.8150	0.8219	0.8151	0.8213	0.8256	0.8335	0.8346	0.8354	0.8354
		WEI	0.1642	0.1777	0.1761	0.1695	0.1740	0.1621	0.1480	0.1294	0.1119	0.1078	0.1030
	AIC <sub>c</sub>	EXP	0.0200	0.0232	0.0180	0.0166	0.0165	0.0220	0.0308	0.0401	0.0564	0.0575	0.0617
		RAY	0.8839	0.8712	0.8732	0.8727	0.8585	0.8569	0.8514	0.8526	0.8443	0.8370	0.8361
		WEI	0.0961	0.1056	0.1088	0.1107	0.1250	0.1211	0.1178	0.1073	0.0993	0.1055	0.1022
	BIC	EXP	0.0201	0.0233	0.0199	0.0198	0.0218	0.0291	0.0390	0.0475	0.0629	0.0656	0.0691
		RAY	0.8848	0.8728	0.8831	0.8887	0.8897	0.9061	0.9054	0.9101	0.9089	0.9231	0.9275
		WEI	0.0951	0.1039	0.0970	0.0915	0.0885	0.0648	0.0556	0.0424	0.0282	0.0113	0.0034
	BIC <sub>M</sub>	EXP	0.0201	0.0232	0.0194	0.0191	0.0199	0.0264	0.0374	0.0468	0.0624	0.0655	0.0687
		RAY	0.8848	0.8725	0.8790	0.8833	0.8733	0.8778	0.8760	0.8827	0.8762	0.8739	0.8746
		WEI	0.0951	0.1043	0.1016	0.0976	0.1068	0.0958	0.0866	0.0705	0.0614	0.0606	0.0567
20	AIC	EXP	0.0024	0.0023	0.0017	0.0018	0.0019	0.0042	0.0081	0.0140	0.0203	0.0241	0.0292
		RAY	0.8244	0.8269	0.8233	0.8265	0.8298	0.8271	0.8285	0.8337	0.8419	0.8432	0.8367
		WEI	0.1732	0.1708	0.1750	0.1717	0.1683	0.1687	0.1634	0.1523	0.1378	0.1327	0.1341
	AIC <sub>c</sub>	EXP	0.0038	0.0038	0.0031	0.0038	0.0029	0.0047	0.0099	0.0162	0.0225	0.0243	0.0296
		RAY	0.8688	0.8703	0.8661	0.8654	0.8594	0.8520	0.8498	0.8505	0.8528	0.8448	0.8370
		WEI	0.1274	0.1259	0.1308	0.1308	0.1377	0.1433	0.1403	0.1333	0.1247	0.1309	0.1334
	BIC	EXP	0.0067	0.0069	0.0071	0.0070	0.0075	0.0120	0.0194	0.0303	0.0378	0.0440	0.0464
		RAY	0.9055	0.9031	0.9072	0.9108	0.9171	0.9282	0.9360	0.9361	0.9392	0.9485	0.9497
		WEI	0.0878	0.0900	0.0857	0.0822	0.0754	0.0598	0.0446	0.0336	0.0230	0.0075	0.0039
	BIC <sub>M</sub>	EXP	0.0067	0.0069	0.0066	0.0066	0.0056	0.0087	0.0151	0.0261	0.0350	0.0417	0.0455
		RAY	0.9055	0.9022	0.9044	0.9032	0.9032	0.9033	0.9092	0.9066	0.9087	0.9051	0.9035
		WEI	0.0878	0.0909	0.0890	0.0902	0.0912	0.0880	0.0757	0.0673	0.0563	0.0532	0.0510
25	AIC	EXP	0.0005	0.0002	0.0003	0.0002	0.0005	0.0008	0.0021	0.0050	0.0080	0.0104	0.0100
		RAY	0.8275	0.8281	0.8323	0.8304	0.8267	0.8296	0.8358	0.8431	0.8446	0.8399	0.8449
		WEI	0.1720	0.1717	0.1674	0.1694	0.1728	0.1696	0.1621	0.1519	0.1474	0.1497	0.1451
	AIC <sub>c</sub>	EXP	0.0008	0.0003	0.0003	0.0006	0.0005	0.0010	0.0026	0.0057	0.0085	0.0105	0.0100
		RAY	0.8628	0.8616	0.8655	0.8605	0.8551	0.8524	0.8541	0.8566	0.8517	0.8410	0.8452
		WEI	0.1364	0.1381	0.1342	0.1389	0.1444	0.1466	0.1433	0.1377	0.1398	0.1485	0.1448
	BIC	EXP	0.0017	0.0020	0.0014	0.0018	0.0032	0.0040	0.0108	0.0158	0.0273	0.0328	0.0276
		RAY	0.9142	0.9128	0.9223	0.9271	0.9296	0.9422	0.9434	0.9546	0.9527	0.9602	0.9701
		WEI	0.0841	0.0852	0.0763	0.0711	0.0672	0.0538	0.0458	0.0296	0.0200	0.0070	0.0023
	BIC <sub>M</sub>	EXP	0.0017	0.0017	0.0013	0.0015	0.0023	0.0020	0.0072	0.0123	0.0201	0.0273	0.0235
		RAY	0.9142	0.9126	0.9203	0.9218	0.9163	0.9215	0.9199	0.9260	0.9248	0.9179	0.9276
		WEI	0.0841	0.0857	0.0784	0.0767	0.0814	0.0765	0.0729	0.0617	0.0551	0.0548	0.0489
50	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.8367	0.8363	0.8321	0.8357	0.8406	0.8369	0.8364	0.8385	0.8388	0.8444	0.8430
		WEI	0.1633	0.1637	0.1679	0.1643	0.1594	0.1631	0.1636	0.1615	0.1612	0.1556	0.1570
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.8539	0.8526	0.8458	0.8494	0.8533	0.8484	0.8451	0.8455	0.8429	0.8453	0.8432
		WEI	0.1461	0.1474	0.1542	0.1506	0.1467	0.1516	0.1549	0.1545	0.1571	0.1547	0.1568
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0005	0.0048	0.0034
		RAY	0.9499	0.9484	0.9481	0.9494	0.9582	0.9618	0.9665	0.9696	0.9763	0.9898	0.9943
		WEI	0.0501	0.0516	0.0519	0.0506	0.0418	0.0382	0.0334	0.0300	0.0232	0.0054	0.0023
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0004	0.0006
		RAY	0.9499	0.9475	0.9469	0.9455	0.9502	0.9483	0.9500	0.9494	0.9481	0.9543	0.9531
		WEI	0.0501	0.0525	0.0531	0.0545	0.0498	0.0517	0.0500	0.0506	0.0517	0.0453	0.0463

Cuadro 5.B.5: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY.

			Fracción de censura $\gamma$											
$r$	Criterio	Distr.	0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99	
75	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.8314	0.8330	0.8454	0.8329	0.8358	0.8460	0.8387	0.8386	0.8402	0.8460	0.8478	
		WEI	0.1686	0.1670	0.1546	0.1671	0.1642	0.1540	0.1613	0.1614	0.1598	0.1540	0.1522	
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.8423	0.8443	0.8567	0.8437	0.8442	0.8525	0.8434	0.8435	0.8440	0.8466	0.8479	
		WEI	0.1577	0.1557	0.1433	0.1563	0.1558	0.1475	0.1566	0.1565	0.1560	0.1534	0.1521	
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	
		RAY	0.9590	0.9598	0.9615	0.9635	0.9666	0.9736	0.9731	0.9760	0.9846	0.9923	0.9964	
		WEI	0.0410	0.0402	0.0385	0.0365	0.0334	0.0264	0.0269	0.0240	0.0154	0.0076	0.0029	
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.9590	0.9594	0.9609	0.9611	0.9596	0.9636	0.9595	0.9602	0.9636	0.9616	0.9619	
		WEI	0.0410	0.0406	0.0391	0.0389	0.0404	0.0364	0.0405	0.0398	0.0364	0.0384	0.0381	
100	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.8438	0.8381	0.8331	0.8342	0.8383	0.8330	0.8428	0.8378	0.8401	0.8361	0.8447	
		WEI	0.1562	0.1619	0.1669	0.1658	0.1617	0.1670	0.1572	0.1622	0.1599	0.1639	0.1553	
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.8518	0.8462	0.8427	0.8417	0.8446	0.8382	0.8474	0.8413	0.8418	0.8366	0.8449	
		WEI	0.1482	0.1538	0.1573	0.1583	0.1554	0.1618	0.1526	0.1587	0.1582	0.1634	0.1551	
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.9703	0.9689	0.9667	0.9679	0.9731	0.9731	0.9775	0.9795	0.9844	0.9929	0.9978	
		WEI	0.0297	0.0311	0.0333	0.0321	0.0269	0.0269	0.0225	0.0205	0.0156	0.0071	0.0022	
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.9703	0.9687	0.9648	0.9653	0.9669	0.9634	0.9680	0.9656	0.9676	0.9684	0.9681	
		WEI	0.0297	0.0313	0.0352	0.0347	0.0331	0.0366	0.0320	0.0344	0.0324	0.0316	0.0319	

Cuadro 5.B.6: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución RAY.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
4	AIC	EXP	0.4946	0.5042	0.5001	0.5022	0.4775	0.4291	0.4102	0.4018	0.3971	0.4058	0.4113
		RAY	0.0867	0.0897	0.0960	0.1087	0.0870	0.0560	0.0505	0.0406	0.0329	0.0410	0.0390
		WEI	0.4187	0.4061	0.4039	0.3891	0.4355	0.5148	0.5394	0.5576	0.5700	0.5532	0.5496
	AIC <sub>c</sub>	EXP	0.8688	0.8632	0.8551	0.8409	0.7441	0.5910	0.5524	0.5085	0.4637	0.4155	0.4134
		RAY	0.0916	0.0952	0.1042	0.1201	0.0954	0.0585	0.0531	0.0419	0.0334	0.0410	0.0390
		WEI	0.0396	0.0416	0.0406	0.0389	0.1605	0.3505	0.3946	0.4496	0.5029	0.5435	0.5476
	BIC	EXP	0.4275	0.4386	0.4339	0.4387	0.4345	0.4238	0.4191	0.4305	0.4782	0.6038	0.7102
		RAY	0.0848	0.0871	0.0918	0.1044	0.0850	0.0559	0.0506	0.0410	0.0336	0.0422	0.0403
		WEI	0.4877	0.4743	0.4743	0.4569	0.4805	0.5203	0.5304	0.5286	0.4881	0.3540	0.2495
	BIC <sub>M</sub>	EXP	0.4275	0.4369	0.4287	0.4246	0.3916	0.3584	0.3269	0.3076	0.2934	0.2954	0.2981
		RAY	0.0848	0.0868	0.0908	0.1016	0.0784	0.0512	0.0430	0.0329	0.0239	0.0268	0.0253
		WEI	0.4877	0.4763	0.4805	0.4738	0.5300	0.5904	0.6301	0.6595	0.6826	0.6777	0.6766
6	AIC	EXP	0.3751	0.3866	0.3847	0.3457	0.3493	0.3167	0.2966	0.2919	0.2870	0.3112	0.3170
		RAY	0.0239	0.0221	0.0320	0.0186	0.0242	0.0200	0.0152	0.0121	0.0115	0.0132	0.0158
		WEI	0.6010	0.5913	0.5833	0.6357	0.6265	0.6633	0.6882	0.6960	0.7015	0.6756	0.6671
	AIC <sub>c</sub>	EXP	0.6020	0.6229	0.6193	0.5244	0.4940	0.4195	0.3716	0.3542	0.3231	0.3176	0.3182
		RAY	0.0242	0.0226	0.0328	0.0191	0.0246	0.0204	0.0154	0.0122	0.0115	0.0132	0.0158
		WEI	0.3738	0.3545	0.3479	0.4565	0.4814	0.5601	0.6130	0.6336	0.6654	0.6692	0.6659
	BIC	EXP	0.3559	0.3673	0.3634	0.3416	0.3555	0.3437	0.3392	0.3556	0.3937	0.5329	0.6368
		RAY	0.0239	0.0219	0.0318	0.0185	0.0242	0.0201	0.0154	0.0122	0.0118	0.0135	0.0160
		WEI	0.6202	0.6108	0.6048	0.6399	0.6203	0.6362	0.6454	0.6322	0.5945	0.4536	0.3472
	BIC <sub>M</sub>	EXP	0.3559	0.3656	0.3569	0.3313	0.3240	0.2878	0.2660	0.2597	0.2500	0.2655	0.2738
		RAY	0.0239	0.0219	0.0317	0.0182	0.0237	0.0194	0.0145	0.0114	0.0097	0.0109	0.0135
		WEI	0.6202	0.6125	0.6114	0.6505	0.6523	0.6928	0.7195	0.7289	0.7403	0.7236	0.7126
8	AIC	EXP	0.2524	0.2626	0.2793	0.2578	0.2262	0.2324	0.2071	0.1901	0.2036	0.2264	0.2249
		RAY	0.0076	0.0066	0.0095	0.0069	0.0060	0.0068	0.0038	0.0023	0.0043	0.0056	0.0047
		WEI	0.7400	0.7308	0.7112	0.7353	0.7678	0.7608	0.7891	0.8076	0.7921	0.7680	0.7704
	AIC <sub>c</sub>	EXP	0.3862	0.3965	0.4146	0.3659	0.3030	0.2928	0.2547	0.2248	0.2239	0.2303	0.2258
		RAY	0.0076	0.0066	0.0095	0.0069	0.0060	0.0069	0.0039	0.0023	0.0043	0.0056	0.0047
		WEI	0.6062	0.5969	0.5759	0.6272	0.6910	0.7003	0.7414	0.7729	0.7718	0.7641	0.7695
	BIC	EXP	0.2580	0.2695	0.2861	0.2727	0.2557	0.2733	0.2635	0.2613	0.3135	0.4561	0.5525
		RAY	0.0076	0.0066	0.0095	0.0069	0.0060	0.0069	0.0039	0.0024	0.0043	0.0056	0.0049
		WEI	0.7344	0.7239	0.7044	0.7204	0.7383	0.7198	0.7326	0.7363	0.6822	0.5383	0.4426
	BIC <sub>M</sub>	EXP	0.2580	0.2684	0.2808	0.2627	0.2331	0.2300	0.2069	0.1877	0.1982	0.2200	0.2148
		RAY	0.0076	0.0066	0.0095	0.0069	0.0060	0.0067	0.0039	0.0023	0.0041	0.0046	0.0044
		WEI	0.7344	0.7250	0.7097	0.7304	0.7609	0.7633	0.7892	0.8100	0.7977	0.7754	0.7808
10	AIC	EXP	0.1780	0.1769	0.1529	0.1744	0.1762	0.1445	0.1313	0.1277	0.1370	0.1586	0.1651
		RAY	0.0016	0.0015	0.0013	0.0026	0.0026	0.0010	0.0013	0.0012	0.0009	0.0016	0.0025
		WEI	0.8204	0.8216	0.8458	0.8230	0.8212	0.8545	0.8674	0.8711	0.8621	0.8398	0.8324
	AIC <sub>c</sub>	EXP	0.2542	0.2586	0.2173	0.2384	0.2237	0.1743	0.1555	0.1492	0.1508	0.1609	0.1655
		RAY	0.0016	0.0015	0.0013	0.0026	0.0026	0.0010	0.0013	0.0012	0.0009	0.0016	0.0025
		WEI	0.7442	0.7399	0.7814	0.7590	0.7737	0.8247	0.8432	0.8496	0.8483	0.8375	0.8320
	BIC	EXP	0.1976	0.1977	0.1746	0.2007	0.2093	0.1874	0.1869	0.1974	0.2390	0.3714	0.4791
		RAY	0.0016	0.0015	0.0013	0.0026	0.0026	0.0010	0.0013	0.0012	0.0011	0.0016	0.0025
		WEI	0.8008	0.8008	0.8241	0.7967	0.7881	0.8116	0.8118	0.8014	0.7599	0.6270	0.5184
	BIC <sub>M</sub>	EXP	0.1976	0.1969	0.1713	0.1921	0.1880	0.1563	0.1416	0.1397	0.1491	0.1694	0.1737
		RAY	0.0016	0.0015	0.0013	0.0026	0.0026	0.0010	0.0012	0.0011	0.0011	0.0015	0.0024
		WEI	0.8008	0.8016	0.8274	0.8053	0.8094	0.8427	0.8572	0.8592	0.8498	0.8291	0.8239

Cuadro 5.B.7: Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
15	AIC	EXP	0.0609	0.0669	0.0606	0.0550	0.0570	0.0479	0.0445	0.0435	0.0511	0.0666	0.0663
		RAY	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
		WEI	0.9390	0.9330	0.9393	0.9450	0.9430	0.9521	0.9555	0.9565	0.9488	0.9332	0.9334
	AIC <sub>c</sub>	EXP	0.0817	0.0912	0.0785	0.0725	0.0725	0.0584	0.0527	0.0490	0.0548	0.0680	0.0663
		RAY	0.0001	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
		WEI	0.9182	0.9085	0.9214	0.9275	0.9275	0.9416	0.9473	0.9510	0.9451	0.9318	0.9334
	BIC	EXP	0.0824	0.0919	0.0839	0.0796	0.0876	0.0842	0.0835	0.0846	0.1159	0.2242	0.3067
		RAY	0.0001	0.0003	0.0001	0.0001	0.0000	0.0001	0.0000	0.0000	0.0001	0.0002	0.0003
		WEI	0.9175	0.9078	0.9160	0.9203	0.9124	0.9157	0.9165	0.9154	0.8840	0.7756	0.6930
	BIC <sub>M</sub>	EXP	0.0824	0.0914	0.0811	0.0755	0.0766	0.0656	0.0613	0.0581	0.0654	0.0874	0.0869
		RAY	0.0001	0.0003	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001	0.0002	0.0003
		WEI	0.9175	0.9083	0.9188	0.9245	0.9234	0.9343	0.9387	0.9419	0.9345	0.9124	0.9128
20	AIC	EXP	0.0204	0.0226	0.0203	0.0194	0.0155	0.0173	0.0140	0.0136	0.0173	0.0269	0.0307
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9796	0.9774	0.9797	0.9806	0.9845	0.9827	0.9860	0.9864	0.9827	0.9731	0.9693
	AIC <sub>c</sub>	EXP	0.0266	0.0290	0.0266	0.0255	0.0195	0.0195	0.0159	0.0149	0.0189	0.0271	0.0307
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9734	0.9710	0.9734	0.9745	0.9805	0.9805	0.9841	0.9851	0.9811	0.9729	0.9693
	BIC	EXP	0.0349	0.0355	0.0346	0.0361	0.0309	0.0370	0.0347	0.0391	0.0598	0.1276	0.1922
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9651	0.9645	0.9654	0.9639	0.9691	0.9630	0.9653	0.9609	0.9402	0.8724	0.8078
	BIC <sub>M</sub>	EXP	0.0349	0.0353	0.0338	0.0331	0.0268	0.0273	0.0228	0.0236	0.0289	0.0440	0.0469
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9651	0.9647	0.9662	0.9669	0.9732	0.9727	0.9772	0.9764	0.9711	0.9560	0.9531
25	AIC	EXP	0.0067	0.0077	0.0068	0.0059	0.0062	0.0049	0.0044	0.0048	0.0069	0.0100	0.0108
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9933	0.9923	0.9932	0.9941	0.9938	0.9951	0.9956	0.9952	0.9931	0.9900	0.9892
	AIC <sub>c</sub>	EXP	0.0088	0.0096	0.0095	0.0076	0.0066	0.0057	0.0051	0.0054	0.0077	0.0102	0.0108
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9912	0.9904	0.9905	0.9924	0.9934	0.9943	0.9949	0.9946	0.9923	0.9898	0.9892
	BIC	EXP	0.0155	0.0147	0.0141	0.0133	0.0129	0.0116	0.0130	0.0154	0.0267	0.0712	0.1135
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9845	0.9853	0.9859	0.9867	0.9871	0.9884	0.9870	0.9846	0.9733	0.9288	0.8865
	BIC <sub>M</sub>	EXP	0.0155	0.0147	0.0139	0.0126	0.0110	0.0086	0.0084	0.0093	0.0130	0.0218	0.0223
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9845	0.9853	0.9861	0.9874	0.9890	0.9914	0.9916	0.9907	0.9870	0.9782	0.9777
50	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	BIC	EXP	0.0002	0.0000	0.0002	0.0002	0.0001	0.0000	0.0000	0.0001	0.0004	0.0023	0.0048
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9998	1.0000	0.9998	0.9998	0.9999	1.0000	1.0000	0.9999	0.9996	0.9977	0.9952
	BIC <sub>M</sub>	EXP	0.0002	0.0000	0.0002	0.0002	0.0001	0.0000	0.0000	0.0000	0.0001	0.0007	0.0006
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	0.9998	1.0000	0.9998	0.9998	0.9999	1.0000	1.0000	1.0000	0.9999	0.9993	0.9994

Cuadro 5.B.8: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5.



$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
75	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
100	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Cuadro 5.B.9: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma igual a 0.5.





$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
75	AIC	EXP	0.0004	0.0003	0.0008	0.0003	0.0002	0.0007	0.0005	0.0022	0.0037	0.0061	0.0084
		RAY	0.0586	0.0549	0.0506	0.0528	0.0483	0.0388	0.0443	0.0525	0.0650	0.0845	0.0850
		WEI	0.9410	0.9448	0.9486	0.9469	0.9515	0.9605	0.9552	0.9453	0.9313	0.9094	0.9066
	AIC <sub>c</sub>	EXP	0.0007	0.0007	0.0009	0.0003	0.0002	0.0007	0.0006	0.0022	0.0040	0.0061	0.0084
		RAY	0.0621	0.0592	0.0541	0.0557	0.0507	0.0401	0.0458	0.0545	0.0666	0.0848	0.0851
		WEI	0.9372	0.9401	0.9450	0.9440	0.9491	0.9592	0.9536	0.9433	0.9294	0.9091	0.9065
	BIC	EXP	0.0046	0.0057	0.0038	0.0056	0.0040	0.0072	0.0164	0.0302	0.0590	0.1409	0.2316
		RAY	0.1583	0.1515	0.1478	0.1522	0.1469	0.1416	0.1667	0.2010	0.2682	0.4076	0.5238
		WEI	0.8371	0.8428	0.8484	0.8422	0.8491	0.8512	0.8169	0.7688	0.6728	0.4515	0.2446
	BIC <sub>M</sub>	EXP	0.0046	0.0057	0.0038	0.0048	0.0031	0.0051	0.0098	0.0173	0.0276	0.0391	0.0391
		RAY	0.1583	0.1511	0.1452	0.1464	0.1326	0.1194	0.1306	0.1499	0.1840	0.2107	0.2167
		WEI	0.8371	0.8432	0.8510	0.8488	0.8643	0.8755	0.8596	0.8328	0.7884	0.7502	0.7442
100	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0003	0.0002	0.0009	0.0011
		RAY	0.0151	0.0185	0.0187	0.0180	0.0147	0.0161	0.0146	0.0185	0.0235	0.0358	0.0392
		WEI	0.9849	0.9815	0.9813	0.9820	0.9853	0.9838	0.9853	0.9812	0.9763	0.9633	0.9597
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0003	0.0002	0.0009	0.0011
		RAY	0.0161	0.0194	0.0195	0.0189	0.0152	0.0161	0.0148	0.0189	0.0239	0.0358	0.0392
		WEI	0.9839	0.9806	0.9805	0.9811	0.9848	0.9838	0.9851	0.9808	0.9759	0.9633	0.9597
	BIC	EXP	0.0006	0.0004	0.0004	0.0008	0.0004	0.0008	0.0025	0.0075	0.0175	0.0549	0.0966
		RAY	0.0689	0.0709	0.0760	0.0742	0.0690	0.0712	0.0803	0.1069	0.1540	0.2783	0.3751
		WEI	0.9305	0.9287	0.9236	0.9250	0.9306	0.9280	0.9172	0.8856	0.8285	0.6668	0.5283
	BIC <sub>M</sub>	EXP	0.0006	0.0004	0.0004	0.0003	0.0004	0.0005	0.0013	0.0036	0.0074	0.0111	0.0107
		RAY	0.0689	0.0706	0.0741	0.0713	0.0611	0.0593	0.0620	0.0753	0.0976	0.1294	0.1339
		WEI	0.9305	0.9290	0.9255	0.9284	0.9385	0.9402	0.9367	0.9211	0.8950	0.8595	0.8554

Cuadro 5.B.12: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 1.5.



$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
15	AIC	EXP	0.0000	0.0002	0.0001	0.0002	0.0005	0.0008	0.0022	0.0032	0.0052	0.0064	0.0064
		RAY	0.5376	0.5484	0.5360	0.5322	0.5399	0.5467	0.5796	0.6084	0.6454	0.6572	0.6712
		WEI	0.4624	0.4514	0.4639	0.4676	0.4596	0.4525	0.4182	0.3884	0.3494	0.3364	0.3224
	AIC <sub>c</sub>	EXP	0.0004	0.0004	0.0002	0.0003	0.0007	0.0008	0.0022	0.0033	0.0053	0.0064	0.0064
		RAY	0.6411	0.6401	0.6222	0.6171	0.6098	0.6020	0.6210	0.6388	0.6651	0.6597	0.6717
		WEI	0.3585	0.3595	0.3776	0.3826	0.3895	0.3972	0.3768	0.3579	0.3296	0.3339	0.3219
	BIC	EXP	0.0004	0.0004	0.0002	0.0005	0.0010	0.0011	0.0027	0.0036	0.0057	0.0066	0.0070
		RAY	0.6437	0.6414	0.6428	0.6497	0.6739	0.7037	0.7478	0.7890	0.8451	0.9229	0.9606
		WEI	0.3559	0.3582	0.3570	0.3498	0.3251	0.2952	0.2495	0.2074	0.1492	0.0705	0.0324
	BIC <sub>M</sub>	EXP	0.0004	0.0004	0.0002	0.0005	0.0010	0.0011	0.0026	0.0036	0.0057	0.0066	0.0070
		RAY	0.6437	0.6401	0.6345	0.6356	0.6374	0.6421	0.6697	0.6970	0.7304	0.7395	0.7507
		WEI	0.3559	0.3595	0.3653	0.3639	0.3616	0.3568	0.3277	0.2994	0.2639	0.2539	0.2423
20	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0005	0.0011	0.0012	0.0017
		RAY	0.4838	0.4823	0.4895	0.4945	0.4819	0.4951	0.5183	0.5640	0.5949	0.6178	0.6184
		WEI	0.5162	0.5177	0.5105	0.5055	0.5181	0.5048	0.4814	0.4355	0.4040	0.3810	0.3799
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0007	0.0013	0.0012	0.0017
		RAY	0.5562	0.5511	0.5549	0.5608	0.5327	0.5352	0.5504	0.5863	0.6114	0.6208	0.6192
		WEI	0.4438	0.4489	0.4451	0.4392	0.4673	0.4647	0.4491	0.4130	0.3873	0.3780	0.3791
	BIC	EXP	0.0000	0.0001	0.0001	0.0000	0.0000	0.0001	0.0008	0.0011	0.0020	0.0017	0.0027
		RAY	0.6201	0.6193	0.6230	0.6472	0.6527	0.6878	0.7235	0.7693	0.8278	0.9172	0.9561
		WEI	0.3799	0.3806	0.3769	0.3528	0.3473	0.3121	0.2757	0.2296	0.1702	0.0811	0.0412
	BIC <sub>M</sub>	EXP	0.0000	0.0001	0.0001	0.0000	0.0000	0.0001	0.0008	0.0010	0.0020	0.0016	0.0027
		RAY	0.6201	0.6179	0.6181	0.6322	0.6182	0.6276	0.6499	0.6801	0.7130	0.7368	0.7344
		WEI	0.3799	0.3820	0.3818	0.3678	0.3818	0.3723	0.3493	0.3189	0.2850	0.2616	0.2629
25	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0004	0.0001
		RAY	0.4346	0.4289	0.4309	0.4196	0.4236	0.4432	0.4700	0.5068	0.5406	0.5630	0.5699
		WEI	0.5654	0.5711	0.5691	0.5804	0.5764	0.5568	0.5299	0.4932	0.4594	0.4366	0.4300
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0004	0.0001
		RAY	0.4915	0.4814	0.4843	0.4712	0.4669	0.4744	0.4940	0.5313	0.5537	0.5662	0.5706
		WEI	0.5085	0.5186	0.5157	0.5288	0.5331	0.5256	0.5059	0.4687	0.4463	0.4334	0.4293
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003	0.0005	0.0009	0.0004
		RAY	0.6009	0.5926	0.6042	0.6031	0.6244	0.6638	0.6978	0.7513	0.8125	0.9036	0.9532
		WEI	0.3991	0.4074	0.3958	0.3969	0.3756	0.3361	0.3020	0.2484	0.1870	0.0955	0.0464
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0002	0.0004	0.0008	0.0004
		RAY	0.6009	0.5914	0.5992	0.5892	0.5909	0.6072	0.6265	0.6642	0.6962	0.7072	0.7207
		WEI	0.3991	0.4086	0.4008	0.4108	0.4091	0.3928	0.3733	0.3356	0.3034	0.2920	0.2789
50	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.2367	0.2395	0.2310	0.2241	0.2211	0.2335	0.2666	0.2994	0.3391	0.3652	0.3748
		WEI	0.7633	0.7605	0.7690	0.7759	0.7789	0.7665	0.7334	0.7006	0.6609	0.6348	0.6252
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.2577	0.2574	0.2538	0.2409	0.2368	0.2465	0.2766	0.3088	0.3439	0.3664	0.3750
		WEI	0.7423	0.7426	0.7462	0.7591	0.7632	0.7535	0.7234	0.6912	0.6561	0.6336	0.6250
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.4627	0.4611	0.4691	0.4582	0.4852	0.5123	0.5695	0.6181	0.6941	0.8264	0.8947
		WEI	0.5373	0.5389	0.5309	0.5418	0.5148	0.4877	0.4305	0.3819	0.3059	0.1736	0.1053
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.4627	0.4602	0.4634	0.4455	0.4517	0.4588	0.4985	0.5284	0.5636	0.5997	0.6108
		WEI	0.5373	0.5398	0.5366	0.5545	0.5483	0.5412	0.5015	0.4716	0.4364	0.4003	0.3892

Cuadro 5.B.14: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 2.5.

$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
75	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.1178	0.1176	0.1098	0.1173	0.1090	0.1152	0.1502	0.1720	0.2085	0.2295	0.2349
		WEI	0.8822	0.8824	0.8902	0.8827	0.8910	0.8848	0.8498	0.8280	0.7915	0.7705	0.7651
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.1284	0.1262	0.1175	0.1236	0.1151	0.1200	0.1535	0.1770	0.2113	0.2304	0.2349
		WEI	0.8716	0.8738	0.8825	0.8764	0.8849	0.8800	0.8465	0.8230	0.7887	0.7696	0.7651
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.3298	0.3295	0.3359	0.3400	0.3314	0.3699	0.4264	0.4897	0.5794	0.7416	0.8322
		WEI	0.6702	0.6705	0.6641	0.6600	0.6686	0.6301	0.5736	0.5103	0.4206	0.2584	0.1678
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.3298	0.3290	0.3319	0.3284	0.3064	0.3240	0.3659	0.4001	0.4511	0.4898	0.4901
		WEI	0.6702	0.6710	0.6681	0.6716	0.6936	0.6760	0.6341	0.5999	0.5489	0.5102	0.5099
100	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0614	0.0608	0.0603	0.0566	0.0534	0.0579	0.0710	0.0992	0.1189	0.1440	0.1433
		WEI	0.9386	0.9392	0.9397	0.9434	0.9466	0.9421	0.9290	0.9008	0.8811	0.8560	0.8567
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0654	0.0649	0.0642	0.0598	0.0555	0.0602	0.0736	0.1006	0.1202	0.1443	0.1433
		WEI	0.9346	0.9351	0.9358	0.9402	0.9445	0.9398	0.9264	0.8994	0.8798	0.8557	0.8567
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.2361	0.2392	0.2239	0.2336	0.2294	0.2516	0.3115	0.3760	0.4608	0.6342	0.7436
		WEI	0.7639	0.7608	0.7761	0.7664	0.7706	0.7484	0.6885	0.6240	0.5392	0.3658	0.2564
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.2361	0.2384	0.2203	0.2255	0.2059	0.2131	0.2531	0.2947	0.3375	0.3809	0.3884
		WEI	0.7639	0.7616	0.7797	0.7745	0.7941	0.7869	0.7469	0.7053	0.6625	0.6191	0.6116

Cuadro 5.B.15: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 2.5.





$r$	Criterio	Distr.	Fracción de censura $\gamma$										
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99
15	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0557	0.0620	0.0535	0.0490	0.0542	0.0680	0.0946	0.1168	0.1555	0.1743	0.1851
		WEI	0.9443	0.9380	0.9465	0.9510	0.9458	0.9320	0.9054	0.8832	0.8445	0.8257	0.8148
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0949	0.1059	0.0892	0.0804	0.0793	0.0890	0.1162	0.1362	0.1686	0.1773	0.1859
		WEI	0.9051	0.8941	0.9108	0.9196	0.9207	0.9110	0.8838	0.8638	0.8314	0.8227	0.8140
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0956	0.1075	0.0988	0.0971	0.1127	0.1484	0.2037	0.2553	0.3614	0.5620	0.7053
		WEI	0.9044	0.8925	0.9012	0.9029	0.8873	0.8516	0.7963	0.7447	0.6386	0.4380	0.2946
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		RAY	0.0956	0.1071	0.0938	0.0893	0.0921	0.1072	0.1426	0.1691	0.2152	0.2368	0.2535
		WEI	0.9044	0.8929	0.9062	0.9107	0.9079	0.8928	0.8574	0.8309	0.7848	0.7632	0.7464
20	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0214	0.0235	0.0192	0.0211	0.0174	0.0251	0.0424	0.0610	0.0810	0.0992	0.1060
		WEI	0.9786	0.9765	0.9808	0.9789	0.9826	0.9749	0.9576	0.9390	0.9190	0.9008	0.8940
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0333	0.0368	0.0294	0.0324	0.0241	0.0339	0.0509	0.0696	0.0865	0.1003	0.1061
		WEI	0.9667	0.9632	0.9706	0.9676	0.9759	0.9661	0.9491	0.9304	0.9135	0.8997	0.8939
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0536	0.0544	0.0518	0.0558	0.0563	0.0834	0.1254	0.1774	0.2590	0.4489	0.5904
		WEI	0.9464	0.9456	0.9482	0.9442	0.9437	0.9166	0.8746	0.8226	0.7410	0.5511	0.4096
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0536	0.0538	0.0493	0.0516	0.0457	0.0576	0.0825	0.1120	0.1385	0.1689	0.1720
		WEI	0.9464	0.9462	0.9507	0.9484	0.9543	0.9424	0.9175	0.8880	0.8615	0.8311	0.8280
25	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0083	0.0087	0.0068	0.0058	0.0077	0.0098	0.0178	0.0254	0.0408	0.0540	0.0529
		WEI	0.9917	0.9913	0.9932	0.9942	0.9923	0.9902	0.9822	0.9746	0.9592	0.9460	0.9471
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0132	0.0135	0.0100	0.0087	0.0106	0.0118	0.0210	0.0288	0.0443	0.0544	0.0533
		WEI	0.9868	0.9865	0.9900	0.9913	0.9894	0.9882	0.9790	0.9712	0.9557	0.9456	0.9467
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0262	0.0274	0.0267	0.0225	0.0294	0.0430	0.0715	0.1098	0.1746	0.3346	0.4805
		WEI	0.9738	0.9726	0.9733	0.9775	0.9706	0.9570	0.9285	0.8902	0.8254	0.6654	0.5195
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0262	0.0272	0.0257	0.0209	0.0237	0.0275	0.0434	0.0601	0.0933	0.1088	0.1137
		WEI	0.9738	0.9728	0.9743	0.9791	0.9763	0.9725	0.9566	0.9399	0.9067	0.8912	0.8863
50	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0001	0.0002	0.0000	0.0000	0.0000	0.0002	0.0005	0.0004	0.0018	0.0015
		WEI	1.0000	0.9999	0.9998	1.0000	1.0000	1.0000	0.9998	0.9995	0.9996	0.9982	0.9985
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0001	0.0003	0.0000	0.0000	0.0000	0.0002	0.0005	0.0005	0.0018	0.0015
		WEI	1.0000	0.9999	0.9997	1.0000	1.0000	1.0000	0.9998	0.9995	0.9995	0.9982	0.9985
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0001	0.0003	0.0005	0.0003	0.0002	0.0007	0.0026	0.0077	0.0183	0.0583	0.1103
		WEI	0.9999	0.9997	0.9995	0.9997	0.9998	0.9993	0.9974	0.9923	0.9817	0.9417	0.8897
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0001	0.0003	0.0005	0.0001	0.0000	0.0003	0.0010	0.0028	0.0067	0.0107	0.0099
		WEI	0.9999	0.9997	0.9995	0.9999	1.0000	0.9997	0.9990	0.9972	0.9933	0.9893	0.9901

Cuadro 5.B.17: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 3.5.

$r$	Criterio	Distr.	Fracción de censura $\gamma$											
			0	0.01	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.950	0.99	
75	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0007	0.0059	0.0182
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9993	0.9941	0.9818
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0011
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9989
100	AIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	AIC <sub>c</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	BIC	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0016
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9984
	BIC <sub>M</sub>	EXP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		RAY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		WEI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Cuadro 5.B.18: (Continuación). Proporción de veces que el AIC, AIC<sub>c</sub>, BIC y BIC<sub>M</sub> eligen a cada una de las tres candidatas siendo que los datos se generan de la distribución WEI con parámetro de forma 3.5.

## 5.C. Comportamiento asintótico del AIC

**Lema 5.C.1.** Si  $\lim_{n \rightarrow \infty} P(A_n) = 1$  y  $\lim_{n \rightarrow \infty} P(B_n) = c$ , entonces  $\lim_{n \rightarrow \infty} P(A_n \cap B_n) = c$ .

*Demostración.*  $\lim_{n \rightarrow \infty} P(A_n \cap B_n) = 1 - \lim_{n \rightarrow \infty} P(A_n^c \cup B_n^c) \geq 1 - \lim_{n \rightarrow \infty} P(A_n^c) - \lim_{n \rightarrow \infty} P(B_n^c) = c$ .  $\square$

**Lema 5.C.2.** Si  $Z_n \xrightarrow{P} c > 0$ , entonces  $\lim_{n \rightarrow \infty} P(Z_n > 0) = 1$ .

*Demostración.* Tomando  $\epsilon = c$ ,

$$\lim_{n \rightarrow \infty} P(|Z_n - c| < \epsilon) = 1 \Rightarrow 1 = \lim_{n \rightarrow \infty} P(0 < Z_n < 2c) \leq \lim_{n \rightarrow \infty} P(0 < Z_n),$$

y así  $\lim_{n \rightarrow \infty} P(Z_n > 0) = 1$ .  $\square$

**Lema 5.C.3.** (Datos censurados) Bajo el esquema de simulación descrito en la Sección 5.3.1, si  $T \sim EXP(\eta)$  se tiene que  $\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t}) > 0) = 1, \forall \gamma \in [0, 1)$ , mientras que si  $T \sim RAY(\sigma)$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t}) > 0) = 1, \forall \gamma \in [0, 1)$ .

*Demostración.* La demostración la haremos considerando tres casos: cuando no se tiene censura, cuando la censura es menor que 0.5 y cuando la censura es mayor que 0.5 pero menor que 1.

1). Consideremos primero en caso en que no se tiene censura,  $\gamma = 0$ . Notemos que

$$\begin{aligned} \frac{1}{n} (\ell_1(\hat{\eta}|\mathbf{x}) - \ell_2(\hat{\sigma}|\mathbf{x})) &= \frac{1}{n} \left[ -n \log(\bar{X}_n) - n - \sum_{i=1}^n \log(X_i) + n \log \left( \sum_{i=1}^n X_i^2 / (2n) \right) + n \right] \\ &= \frac{1}{n} \left[ -n \log(\bar{X}_n) - \sum_{i=1}^n \log(X_i) + n \log \left( \sum_{i=1}^n X_i^2 / (2n) \right) \right]. \end{aligned}$$

Así, por un lado, si  $X_1, \dots, X_n \stackrel{iid}{\sim} \text{EXP}(\eta)$ ,  $\bar{X}_n \xrightarrow{p} \eta$ ,  $-\frac{1}{n} \sum_{i=1}^n \log(X_i) \xrightarrow{p} \mathbb{E}(-\log(X)) = -\log(\eta) + 0.5772$  y  $\log(\sum_{i=1}^n X_i^2 / (2n)) \xrightarrow{p} 2\log(\eta)$ , y entonces por Slustky

$$\frac{1}{n} (\ell_1(\hat{\eta}|\mathbf{x}) - \ell_2(\hat{\sigma}|\mathbf{x})) \xrightarrow{p} -\log(\eta) - \log(\eta) + 0.5772 + 2\log(\eta) = 0.5772 > 0,$$

de donde  $\lim_{n \rightarrow \infty} P(\ell_1(\hat{\eta}|\mathbf{x}) > \ell_2(\hat{\sigma}|\mathbf{x})) = 1$ , por Lema 5.C.2.

Por otro, si  $X_1, \dots, X_n \stackrel{iid}{\sim} \text{RAY}(\sigma)$ ,  $\bar{X}_n \xrightarrow{p} \sigma\sqrt{\frac{\pi}{2}}$ ,  $-\frac{1}{n} \sum_{i=1}^n \log(X_i) \xrightarrow{p} \mathbb{E}(-\log(X)) = -\log(\sqrt{2}\sigma) + 0.5772$ , y  $\log(\sum_{i=1}^n X_i^2 / (2n)) \xrightarrow{p} 2\log(\sigma)$ , y por Slustky nuevamente

$$\frac{1}{n} (\ell_2(\hat{\eta}|\mathbf{x}) - \ell_1(\hat{\sigma}|\mathbf{x})) \xrightarrow{p} \log\left(\sigma\sqrt{\frac{\pi}{2}}\right) + \log(\sqrt{2}\sigma) - 0.5772 - 2\log(\sigma) = \frac{\log(\pi) - 0.5772}{2} > 0,$$

de donde  $\lim_{n \rightarrow \infty} P(\ell_2(\hat{\sigma}|\mathbf{x}) > \ell_1(\hat{\eta}|\mathbf{x})) = 1$ , por Lema 5.C.2.

2). Ahora consideremos el caso en que hay una fracción de censura  $\gamma \in (0, 0.5)$ , es decir,  $k \geq 1$ . Notemos que

$$\begin{aligned} \mathbb{P}(T^* \leq t, \delta = 1) &= F_T(t) - \frac{1}{2k} [F_T(t)]^2 \\ &= F_T(t) - \frac{1}{2k} [F_T(t)]^2, \end{aligned}$$

y

$$\mathbb{P}(T^* \leq t, \delta = 0) = \frac{F_T(t)}{k} - \frac{[F_T(t)]^2}{2k}.$$

Por lo tanto,

$$\begin{aligned} \mathbb{P}(T^* \leq t) &= \mathbb{P}(T^* \leq t, \delta = 1) + \mathbb{P}(T^* \leq t, \delta = 0) \\ &= F_T(t) \left( 1 + \frac{1}{k} \right) - \frac{[F_T(t)]^2}{k}, \end{aligned}$$

de donde

$$f_{T^*}(t) = f_T(t) \left(1 + \frac{1}{k}\right) - \frac{2F_T(t)}{k} f_T(t),$$

y así

$$E(T^{*s}) = E(T^s) \left(1 + \frac{1}{k}\right) - \frac{2}{k} \int_0^\infty t^s F_T(t) f_T(t) dt.$$

Si  $T \sim \text{EXP}(\eta)$ , notemos que

$$E(T^*) = \eta \left(1 - \frac{1}{k}\right) + \frac{\eta}{2k} = \eta \left(1 - \frac{1}{2k}\right),$$

y

$$E(T^{*2}) = 2\eta^2 \left(1 - \frac{1}{k}\right) + \frac{2\eta^2/4}{k} = 2\eta^2 \left(1 - \frac{3}{4k}\right).$$

Además, como  $\mathbb{P}(T^* = t, \delta = 1) = f_T(t) - \frac{1}{k} f_T(t) F_T(t)$ ,

$$E(\delta \log(T^*)) = - \left(1 - \frac{1}{k}\right) E(Z_1) - \frac{1}{2k} E(Z_2),$$

donde  $Z_1 \sim \text{Gumbel}(-\log(\eta), 1)$  y  $Z_2 \sim \text{Gumbel}(-\log(\eta/2), 1)$ . Luego

$$E(\delta \log(T^*)) = \left(1 - \frac{1}{2k}\right) (\log(\eta) - 0.5272) - \frac{\log(2)}{2k}.$$

Así,

$$\frac{1}{n} (\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t})) \xrightarrow{p} (1 - \gamma) \left( \log \left( \frac{1 - \frac{3\gamma}{2}}{1 - \gamma} \right) + 0.5772 \right) + \gamma \log(2) = h(\gamma),$$

el cuál es positivo (ver Figura 5.C.1), es decir,  $\frac{1}{n} (\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t})) \xrightarrow{p} h(\gamma) > 0, \forall \gamma \in (0, 0.5]$ . Por lo tanto,

$\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t}) > 0) = 1, \forall \gamma \in (0, 0.5]$ .

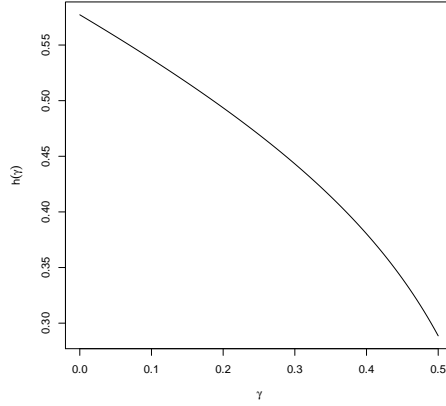


Figura 5.C.1:  $h(\gamma)$  en la distribución EXP,  $0 \leq \gamma \leq 0.5$ .

Ahora si  $T \sim \text{RAY}(\sigma)$  se tiene que

$$E(T^*) = \sigma \sqrt{\pi/2} \left( 1 - \frac{1}{k} + \frac{1}{\sqrt{2k}} \right),$$

y

$$E(T^{*2}) = 2\sigma^2 \left( 1 - \frac{1}{k} + \frac{1}{2k} \right) = 2\sigma^2 \left( 1 - \frac{1}{2k} \right).$$

Además, como  $\mathbb{P}(T^* = t, \delta = 1) = f_T(t) - \frac{1}{k} f_T(t) F_T(t)$ ,

$$E(\delta \log(T^*)) = - \left( 1 - \frac{1}{k} \right) E(Z_1) - \frac{1}{2k} E(Z_2),$$

donde  $Z_1 \sim \text{Gumbel}(-\log(\sqrt{2}\sigma), 1/2)$  y  $Z_2 \sim \text{Gumbel}(-\log(\sigma), 1/2)$ . Luego

$$E(\delta \log(T^*)) = \left( 1 - \frac{1}{2k} \right) \left( \log(\sigma) - \frac{0.5272}{2} \right) + \frac{\log(2)}{2} \left( 1 - \frac{1}{k} \right).$$

Así,

$$\frac{1}{n} (\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t})) \xrightarrow{p} h(\gamma),$$

donde

$$h(\gamma) = (1 - \gamma) \left[ -\frac{0.5272}{2} + \frac{1}{2} \log(\pi) + \log \left( 1 - 2\gamma + \sqrt{2}\gamma \right) - \log(1 - \gamma) \right] - \frac{\gamma \log(2)}{2}.$$

Notemos que

$$\frac{\partial h(\gamma)}{\partial \gamma} = \frac{0.5272 - \log(\pi)}{2} - \frac{1}{1 - \gamma(2 - \sqrt{2})} + \log\left(\frac{1}{1 - \gamma(2 - \sqrt{2})}\right) - \frac{\log(2)}{2} + \log(1 - \gamma) < 0,$$

ya que  $\log\left(\frac{1}{1 - \gamma(2 - \sqrt{2})}\right) < \frac{1}{1 - \gamma(2 - \sqrt{2})}$ , y así todos los términos involucrados agrupados son negativos. Así,  $h(\gamma)$  es decreciente en  $(0, 0.5)$ , y  $\lim_{\gamma \rightarrow 0} h(\gamma) = \frac{\log(\pi) - 0.5272}{2} = 0.2837$ ,  $\lim_{\gamma \rightarrow 0.5} h(\gamma) = 0.1418$ . Por lo tanto,  $h(\gamma)$  es siempre positiva (ver también Figura 5.C.2) para todo  $\gamma \in (0, 0.5)$ , y así  $\frac{1}{n} (\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t})) \xrightarrow{p} h(\gamma) > 0$ , de donde por Lema 5.C.2

$$\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t}) > 0) = 1, \forall \gamma \in (0, 0.5].$$

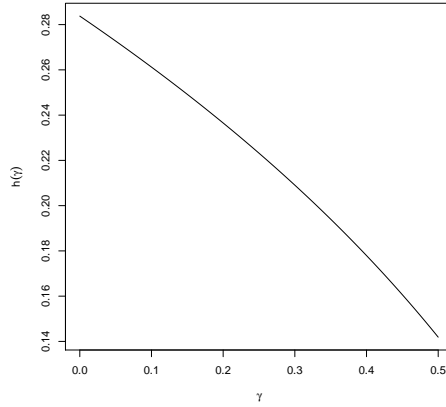


Figura 5.C.2:  $h(\gamma)$  en la distribución RAY,  $0 \leq \gamma \leq 0.5$ .

3). Por último consideremos el caso en que la fracción de censura es mayor que 0.5, es decir,  $k < 1$ . Notemos que

$$\mathbb{P}(T^* \leq t, \delta = 1) = \begin{cases} \frac{k}{2} & F_T(t) \geq k, \\ F_T(t) - \frac{1}{2k} [F_T(t)]^2 & F_T(t) < k, \end{cases}$$

y

$$\mathbb{P}(T^* \leq t, \delta = 0) = \begin{cases} 1 - \frac{k}{2} & F_T(t) \geq k, \\ \frac{F_T(t)}{k} - \frac{[F_T(t)]^2}{2k} & F_T(t) < k. \end{cases}$$

Por lo tanto,

$$\mathbb{P}(T^* \leq t) = \begin{cases} 1 & F_T(t) \geq k, \\ F_T(t) \left(1 + \frac{1}{k}\right) - \frac{[F_T(t)]^2}{k} & F_T(t) < k, \end{cases}$$

de donde

$$f_{T^*}(t) = \begin{cases} f_T(t) \left(1 + \frac{1}{k}\right) - \frac{2F_T(t)}{k} f_T(t) & t < F_T^{-1}(k), \\ 0 & t \geq F_T^{-1}(k), \end{cases}$$

y así

$$\mathbb{E}(T^{*s}) = \left(1 + \frac{1}{k}\right) \int_0^{t_k} t^s f_T(t) dt - \frac{2}{k} \int_0^{t_k} t^s F_T(t) f_T(t) dt.$$

Si  $T \sim \text{EXP}(\eta)$  se tiene que

$$\begin{aligned} \mathbb{E}(T^*) &= C_1 \eta, \\ \mathbb{E}(T^{*2}) &= C_2 \eta^2, \end{aligned}$$

donde

$$\begin{aligned} C_1(\gamma) &= \left(1 - \frac{1}{2k}\right) - \left(1 - \frac{1}{k}\right) [1 - \log(1 - k)] (1 - k) - \frac{1}{k} \left(\frac{1}{2} - \log(1 - k)\right) (1 - k)^2 \\ C_2(\gamma) &= \frac{4k - 3}{2k} - \left(1 - \frac{1}{k}\right) (2 - 2 \log(1 - k) + \log^2(1 - k)) (1 - k) - \\ &\quad \frac{1}{k} \left(\frac{1}{2} - \log(1 - k) + \log^2(1 - k)\right) (1 - k)^2, \end{aligned}$$

y además, como  $\mathbb{P}(T^* = t, \delta = 1) = f_T(t) - \frac{1}{k} f_T(t) F_T(t)$ ,

$$\mathbb{E}(\delta \log(T^*)) = C_3(\gamma) + \frac{1}{2} k \log(\eta),$$

donde

$$\begin{aligned} C_3(\gamma) &= \left(1 - \frac{1}{k}\right) \int_0^{-\log(1-k)} \log(u) \exp(-u) du + \frac{1}{2k} \int_0^{-2\log(1-k)} \log(u) \exp(-u) du - \\ &\quad \frac{1}{2k} \log(2) (1 - (1 - k)^2). \end{aligned}$$

Así,

$$\frac{1}{n} (\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t})) \xrightarrow{P} h(\gamma),$$

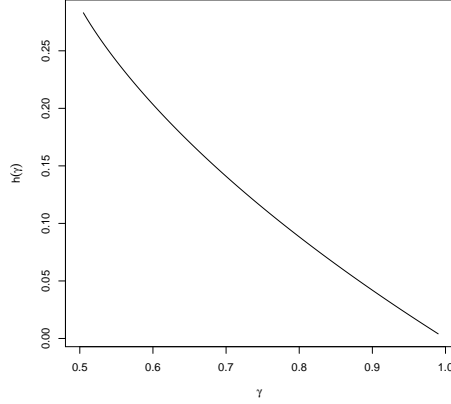


Figura 5.C.3:  $h(\gamma)$  en la distribución EXP,  $0.5 \leq \gamma < 1$ .

donde

$$h(\gamma) = (1 - \gamma) (\log(C_2(\gamma)/2) - \log(C_1(\gamma))) - C_3(\gamma),$$

la cual no depende de  $\eta$  y es mayor que cero  $\forall \gamma \in [0, 1)$ , ver Figura 5.C.3. Así por Lema 5.C.2

$$\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\eta}|\mathbf{t}) - \ell(\hat{\sigma}|\mathbf{t}) > 0) = 1, \forall \gamma \in [0.5, 1).$$

Si  $T \sim \text{RAY}(\sigma)$  se tiene que

$$\begin{aligned} \mathbb{E}(T^*) &= \sigma \left[ \left(1 - \frac{1}{k}\right) \sqrt{2}C_{11}(\gamma) + \frac{1}{k}C_{12}(\gamma) \right], \\ \mathbb{E}(T^{*2}) &= 2\sigma^2 \left( C_{21}(\gamma) - \frac{1}{k} \left( C_{21}(\gamma) - \frac{C_{22}(\gamma)}{2} \right) \right), \end{aligned}$$

donde

$$\begin{aligned} C_{11}(\gamma) &= \sqrt{-\log(1-k)}(1-k) + \sqrt{\pi} \left[ \Phi \left( \sqrt{-2\log(1-k)} \right) - \frac{1}{2} \right] \\ C_{12}(\gamma) &= \left[ \sqrt{-2\log(1-k)}(1-k)^2 + \sqrt{\pi} \left[ \Phi \left( 2\sqrt{-\log(1-k)} \right) - \frac{1}{2} \right] \right] \\ C_{21}(\gamma) &= k - (1-k) \log(1-k) \\ C_{22}(\gamma) &= [1 - (1-k)^2 - 2(1-k)^2 \log(1-k)], \end{aligned}$$

y además, como  $\mathbb{P}(T^* = t, \delta = 1) = f_T(t) - \frac{1}{k}f_T(t)F_T(t)$ ,

$$\mathbb{E}(\delta \log(T^*)) = \frac{k}{2} \log(\sigma) + \left(1 - \frac{1}{k}\right) \left( \log(\sqrt{2})k + C_{311}(\gamma) \right) + \frac{C_{321}(\gamma)}{2k},$$



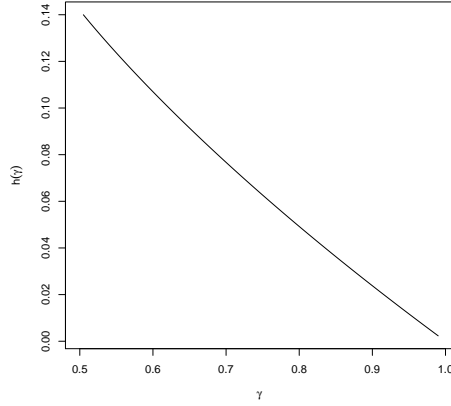


Figura 5.C.4:  $h(\gamma)$  en la distribución RAY,  $0.5 \leq \gamma < 1$ .

donde

$$C_{311}(\gamma) = -\frac{1}{2} \int_{-\log(-\log(1-k))}^{\infty} v \exp(-v - \exp(-v)) dv,$$

$$C_{321}(\gamma) = -\frac{1}{2} \int_{-\log(-2\log(1-k))}^{\infty} v \exp(-v - \exp(-v)) dv.$$

Así,

$$\frac{1}{n} (\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t})) \xrightarrow{p} h(\gamma),$$

donde

$$h(\gamma) = (1 - \gamma) \left[ \log \left( \left( 1 - \frac{1}{k} \right) \sqrt{2} C_{11}(\gamma) + \frac{1}{k} C_{12}(\gamma) \right) - \log \left( C_{21}(\gamma) - \frac{1}{k} \left( C_{21}(\gamma) - \frac{C_{22}(\gamma)}{2} \right) \right) \right] + \left( 1 - \frac{1}{k} \right) \left( \log(\sqrt{2})k + C_{311}(\gamma) \right) + \frac{C_{321}(\gamma)}{k},$$

la cual no depende de  $\sigma$  y es positiva  $\forall \gamma \in [0.5, 1)$ , ver Figura 5.C.4. Así, nuevamente por Lema 5.C.2,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\ell(\hat{\sigma}|\mathbf{t}) - \ell(\hat{\eta}|\mathbf{t}) > 0) = 1, \forall \gamma \in [0.5, 1).$$

□

**Lema 5.C.4.** *Supongamos que los datos son generados bajo el mecanismo II de censura aleatoria por la derecha y que los únicos modelos candidatos son el EXP, RAY y WEI. Entonces*

1. Si los datos provienen de la distribución  $WEI(\alpha, \beta)$  con  $\beta \neq 1, 2$ , la probabilidad asintótica de que el AIC elija al modelo WEI es igual a 1, es decir,

$$\lim_{n \rightarrow \infty} P(AIC_3 < AIC_1, AIC_3 < AIC_2) = 1, \forall \alpha > 0, \beta \in \mathbb{R}_+ \setminus \{1, 2\}. \quad (5.5)$$

donde  $AIC_1$ ,  $AIC_2$  y  $AIC_3$  denotan el AIC de los modelos EXP, RAY y WEI, respectivamente.

2. Si los datos provienen de la distribución exponencial con parámetro  $\eta$ , la probabilidad asintótica de que el AIC elija al modelo EXP está dada por

$$\lim_{n \rightarrow \infty} P(AIC_1 < AIC_2, AIC_1 < AIC_3) = \lim_{n \rightarrow \infty} P \left[ -2 \log \left( \frac{L_1(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})} \right) < 2 \right] = 0.84.$$

3. Cuando los datos provienen de la distribución Rayleigh con parámetro  $\eta$ , la probabilidad asintótica de que el AIC elija a éste modelo es igual a

$$\lim_{n \rightarrow \infty} P(AIC_2 < AIC_1, AIC_2 < AIC_3) = \lim_{n \rightarrow \infty} P \left[ -2 \log \left( \frac{L_2(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})} \right) < 2 \right] = 0.84.$$

*Demostración.* Los tres puntos anteriores lo podemos ver primero haciéndolo en el caso de la distribución WEI y después reduciéndolo a los casos particulares. Antes notemos que si  $T^* = \min(T, C)$  es el tiempo observado bajo el mecanismo II de censura aleatoria, entonces la función de distribución de éste esta dada por

$$\begin{aligned} F_{T^*}(t) &= P(\min(T, C) \leq t) \\ &= 1 - P(\min(T, C) > t) \\ &= 1 - [1 - F_T(t)]^{p+1}. \end{aligned}$$

Además

$$\begin{aligned} P(T^* \leq t, \delta = 1) &= P(T \leq t, T \leq C) \\ &= \frac{1}{1+p} \{1 - [1 - F_T(t)]^{1+p}\} \end{aligned}$$

Así,

$$P(T^* = t, \delta = 1) = [1 - F_T(t)]^p dF_T(t). \quad (5.6)$$

También podemos ver que

$$\begin{aligned}
P(T^* \leq t, \delta = 0) &= P(C \leq t, C \leq T) \\
&= \int_{-\infty}^{\infty} P(C \leq t, C \leq u) dF_T(u) \\
&= \int_{-\infty}^t P(C \leq u) dF_T(u) + \int_t^{\infty} P(C \leq t) dF_T(u) \\
&= \frac{p}{1+p} \{1 - [1 - F_T(t)]^{1+p}\},
\end{aligned}$$

y así la distribución del tiempo observado es

$$F_{T^*}(t) = 1 - [1 - F_T(t)]^{1+p}, \quad (5.7)$$

es decir, la distribución del tiempo observado pertenece a la misma familia del tiempo de censura  $C$  pero con  $1+p$  en vez de  $p$ .

1) Supongamos que  $T \sim \text{WEI}(\alpha, \beta)$  entonces de (5.7) se tiene que  $T^* \sim \text{WEI}(\alpha(1+p)^{-1/\beta}, \beta)$  y así

$$\begin{aligned}
E(T^*) &= \alpha(p+1)^{-1/\beta} \Gamma(1+1/\beta), \\
E(T^{*2}) &= \alpha^2(p+1)^{-2/\beta} \Gamma(1+2/\beta), \\
E\left[\left(\frac{T^*}{\alpha}\right)^\beta\right] &= \frac{1}{1+p} E\left[\left(\frac{T^*}{\alpha(1+p)^{-1/\beta}}\right)^\beta\right] = \frac{1}{1+p}, \\
E[\log(T^*)] &= -\log[\alpha(1+p)^{-1/\beta}] - 0.5772/\beta.
\end{aligned}$$

De (5.6) notemos que

$$\begin{aligned}
E[\delta \log(T^*)] &= \int_0^{\infty} \log(t) [1 - F_T(t)]^p dF_T(t) \\
&= \frac{1}{p+1} \int_0^{\infty} \log(t) dF_{T^*}(t) \\
&= \frac{1}{1+p} E[\log(T^*)] \\
&= \frac{1}{1+p} \left( \log(\alpha) - \frac{1}{\beta} (\log(1+p) + 0.5772) \right).
\end{aligned}$$

Luego, asumiendo que  $T_1, \dots, T_n \stackrel{iid}{\sim} \text{WEI}(\alpha, \beta)$ , entonces por ley de los grandes números, teorema del mapeo continuo, teorema de Slutsky y usando los resultados de (Ferguson, 1996, ecuación (5) página 108) tenemos que

$$\begin{aligned} \frac{1}{n} \ell(\widehat{\alpha}, \widehat{\beta}|\mathbf{t}) &\xrightarrow{c.s.} (1-\gamma) [\log(\beta) - \beta \log(\alpha)] + (\beta-1) \frac{\log(\alpha)}{1+p} - \\ &\quad \frac{\beta-1}{\beta(1+p)} (\log(1+p) + 0.5772) - \frac{1}{1+p} \\ &\xrightarrow{c.s.} (1-\gamma) \left[ \log(\beta) - \frac{\beta-1}{\beta} (\log(1+p) + 0.5772) - 1 \right] - (1-\gamma) \log(\alpha), \end{aligned}$$

y nuevamente usando lo mismo tenemos que

$$\frac{1}{n} \ell(\widehat{\eta}|\mathbf{t}) \xrightarrow{c.s.} -(1-\gamma) \left[ \log \left( \frac{\alpha(p+1)^{-1/\beta} \Gamma(1+1/\beta)}{1-\gamma} \right) + 1 \right],$$

y

$$\begin{aligned} \frac{1}{n} \ell(\widehat{\sigma}|\mathbf{t}) &\xrightarrow{c.s.} -(1-\gamma) \log \left[ \frac{\alpha^2(p+1)^{-2/\beta} \Gamma(1+2/\beta)}{2(1-\gamma)} \right] + \\ &\quad \frac{1}{1+p} \left( \log(\alpha) - \frac{1}{\beta} (\log(1+p) + 0.5772) \right) - (1-\gamma) \\ &\xrightarrow{c.s.} (1-\gamma) \left[ -\log(\alpha) + \frac{1}{\beta} (\log(1+p) - 0.5772) - \log \left[ \frac{\Gamma(1+2/\beta)}{2(1-\gamma)} \right] - 1 \right], \end{aligned}$$

entonces por Slutsky nuevamente y ya que  $\gamma = \frac{p}{1+p}$ ,

$$\begin{aligned} \frac{1}{n} \ell(\widehat{\alpha}, \widehat{\beta}|\mathbf{t}) - \frac{1}{n} \ell(\widehat{\eta}|\mathbf{t}) &\xrightarrow{c.s.} (1-\gamma) \left[ \log(\beta) - \log(1+p) - \frac{\beta-1}{\beta} (0.5772) + \log(\Gamma(1+1/\beta)) \right] - \\ &\quad (1-\gamma) \log(1-\gamma) \\ &\xrightarrow{c.s.} (1-\gamma) \left[ -\frac{\beta-1}{\beta} (0.5772) + \log(\Gamma(1/\beta)) \right] = h(\beta), \end{aligned}$$

y

$$\begin{aligned} \frac{1}{n} \ell(\widehat{\alpha}, \widehat{\beta}|\mathbf{t}) - \frac{1}{n} \ell(\widehat{\eta}|\mathbf{t}) &\xrightarrow{c.s.} (1-\gamma) \left[ -\log(1+p) - 0.5772 + \frac{2}{\beta} (\log(1+p) + 0.5772) \right] + \\ &\quad (1-\gamma) \left[ \log(\Gamma(2/\beta)) - \frac{2}{\beta} \log(1+p) - \log(1-\gamma) \right] \\ &\xrightarrow{c.s.} (1-\gamma) \left[ \left( \frac{\beta/2-1}{\beta/2} \right) 0.5772 + \log(\Gamma(2/\beta)) \right] = h(\beta/2), \end{aligned}$$

donde  $h(\beta) = -\frac{\beta-1}{\beta} (0.5772) + \log(\Gamma(1/\beta))$  y no depende de  $\alpha$ . Además

$$\begin{aligned}
h'(\beta) &= (1 - \gamma) \left[ -\frac{1}{\beta^2} (0.5772) - \frac{1}{\beta^2} \psi(1/\beta) \right] \\
&= -(1 - \gamma) \left[ \frac{1}{\beta^2} \sum_{k=0}^{\infty} \left( \frac{1}{k+1} - \frac{1}{k+1/\beta} \right) \right],
\end{aligned}$$

donde  $\psi(\cdot)$  es la función digamma. Como  $h'(\beta)$  es negativo si  $\beta < 1$ , cero si  $\beta = 1$  y positivo si  $\beta > 1$ , entonces  $h(\beta) > h(1) = 0$  y  $h(\beta/2) > h(2/2) = 0$  para todo  $\beta \in \mathbb{R}_+/\{1\}$  y  $\beta \in \mathbb{R}_+/\{2\}$ , respectivamente. Así por Lema 5.C.2

$$\begin{aligned}
\lim_{n \rightarrow \infty} P(\text{AIC}_3 < \text{AIC}_1) &= \lim_{n \rightarrow \infty} P \left( \frac{\ell_{WEI}(\widehat{\alpha}, \widehat{\beta}|\mathbf{t})}{n} - \frac{\ell_{EXP}(\widehat{\eta}|\mathbf{t})}{n} - \frac{1}{n} > 0 \right) \\
&= 1, \forall \alpha > 0, \beta \in \mathbb{R}_+ \setminus \{1\},
\end{aligned} \tag{5.8}$$

y

$$\begin{aligned}
\lim_{n \rightarrow \infty} P(\text{AIC}_3 < \text{AIC}_2) &= \lim_{n \rightarrow \infty} P \left( \frac{\ell_{WEI}(\widehat{\alpha}, \widehat{\beta}|\mathbf{t})}{n} - \frac{\ell_{RAY}(\widehat{\sigma}|\mathbf{t})}{n} - \frac{1}{n} > 0 \right) \\
&= 1, \forall \alpha > 0, \beta \in \mathbb{R}_+ \setminus \{2\}.
\end{aligned} \tag{5.9}$$

de donde se sigue que cuando los datos provienen de la distribución  $WEI(\alpha, \beta)$  con  $\beta \neq 1, 2$  con una fracción de censura  $\gamma$ , la probabilidad asintótica de que el AIC elija al modelo WEI es igual a 1, es decir,

$$\lim_{n \rightarrow \infty} P(\text{AIC}_3 < \text{AIC}_1, \text{AIC}_3 < \text{AIC}_2) = 1, \forall \alpha > 0, \beta \in \mathbb{R}_+ \setminus \{1, 2\}. \tag{5.10}$$

2) Ahora consideremos una muestra con censura aleatoria de la distribución  $EXP(\eta)$ , es decir,  $X_1, \dots, X_n \stackrel{iid}{\sim} EXP(\eta)$ , y denotemos por  $\text{AIC}_1$ ,  $\text{AIC}_2$  y  $\text{AIC}_3$  el AIC de los modelos EXP, RAY y WEI, respectivamente. Notemos que el AIC elije al modelo EXP como el mejor si y sólo si

$$\{\text{AIC}_1 < \text{AIC}_2, \text{AIC}_1 < \text{AIC}_3\} = \left\{ \ell_1(\widehat{\eta}|\mathbf{t}) > \ell_2(\widehat{\sigma}|\mathbf{t}), -2 \log \left( \frac{L_1(\widehat{\eta}|\mathbf{t})}{L_3(\widehat{\alpha}, \widehat{\beta}|\mathbf{t})} \right) < 2 \right\}.$$

Además tenemos que

$$\begin{aligned}
\frac{1}{n} \ell(\widehat{\eta}|\mathbf{t}) &\xrightarrow{c.s.} (1 - \gamma) [-\log(\eta) - 1], \\
\frac{1}{n} \ell(\widehat{\sigma}|\mathbf{t}) &\xrightarrow{c.s.} (1 - \gamma) [-\log(\eta) - 0.5772 - 1],
\end{aligned}$$

y así

$$\frac{1}{n}\ell(\hat{\eta}|\mathbf{t}) - \frac{1}{n}\ell(\hat{\sigma}|\mathbf{t}) \xrightarrow{c.s.} (1-\gamma)0.5772 > 0 \forall \gamma \in [0, 1).$$

Luego por Lema 5.C.3  $\lim_{n \rightarrow \infty} P(\ell_1(\hat{\eta}|\mathbf{t}) > \ell_2(\hat{\sigma}|\mathbf{t})) = 1$ , entonces por Lema 5.C.1 la probabilidad asintótica de que el AIC elija al modelo EXP es igual a

$$\lim_{n \rightarrow \infty} P(\text{AIC}_1 < \text{AIC}_2, \text{AIC}_1 < \text{AIC}_3) = \lim_{n \rightarrow \infty} P\left[-2 \log\left(\frac{L_1(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})}\right) < 2\right] = 0.84.$$

3) Consideremos una muestra con censura aleatoria de la distribución RAY( $\sigma$ ),  $X_1, \dots, X_n \stackrel{iid}{\sim} \text{RAY}(\sigma)$ , y denotemos por  $\text{AIC}_1$ ,  $\text{AIC}_2$  y  $\text{AIC}_3$  el AIC de los modelos EXP, RAY y WEI, respectivamente, entonces

$$\frac{1}{n}\ell(\hat{\eta}|\mathbf{t}) \xrightarrow{c.s.} (1-\gamma) \left[-\log(\sigma) + \frac{1}{2} \log(1-\gamma) - \log(\Gamma(1/2)) + \frac{1}{2} \log(2) - 1\right],$$

y

$$\frac{1}{n}\ell(\hat{\sigma}|\mathbf{t}) \xrightarrow{c.s.} (1-\gamma) \left[-\frac{1}{2} \log(2) - \log(\sigma) + \frac{1}{2} \log(1-\gamma) - \frac{0.5772}{2} - 1\right].$$

Luego

$$\frac{1}{n}\ell(\hat{\sigma}|\mathbf{t}) - \frac{1}{n}\ell(\hat{\eta}|\mathbf{t}) \xrightarrow{c.s.} (1-\gamma) \left[\log(2) + \log(\Gamma(1/2)) - \frac{1}{2}0.5772\right] > 0 \forall \gamma \in [0, 1).$$

Por lo tanto la probabilidad asintótica de que el AIC elija al modelo RAY es igual a

$$\lim_{n \rightarrow \infty} P(\text{AIC}_2 < \text{AIC}_1, \text{AIC}_2 < \text{AIC}_3) = \lim_{n \rightarrow \infty} P\left(-2 \log\left(\frac{L_2(\hat{\eta}|\mathbf{t})}{L_3(\hat{\alpha}, \hat{\beta}|\mathbf{t})}\right) < 2\right) = 0.84.$$

□

# Capítulo 6

## Conclusiones y comentarios generales

En el Capítulo 2 se consideró una extensión del modelo presentado por Peng y Tseng (2009) que permite que las trayectorias de degradación comiencen de un valor diferente de cero, además de que éstas difieran de unidad a unidad, como puede ser el caso en algunos conjuntos de datos. Se obtuvieron expresiones para la densidad y la función de distribución del tiempo de cruce que ésta extensión induce. Se discutió la identificabilidad de éste y la estimación de los parámetros fue abordada por medio de máxima verosimilitud, y como alternativa se derivaron las expresiones correspondientes del algoritmo EM. Para la ilustración del modelo se usó un conjunto de datos reales, y además se ajustaron diferentes submodelos de tal extensión para elegir uno más parsimonioso. Con algunos submodelos se exploraron algunas razones de porque con el criterio AIC se obtuvieron grandes diferencias que no se vieron reflejadas en la predicción de la función de distribución.

En el Capítulo 3 se propuso usar una mezcla finita de procesos Gaussianos inversos para tomar en cuenta la variabilidad de las unidades que uno sólo de éstos no puede capturar. La estimación de los parámetros del modelo fue hecha por máxima verosimilitud a través del algoritmo EM. Una ventaja del modelo es su simplicidad ya que se pueden obtener cantidades de interés de forma simple, además de que se puede estimar la proporción de la población debido a cada modo de falla, y como consecuencia tomar mejores decisiones para concentrar los esfuerzos en el rediseño o mejora del sistema de producción. Para cuando la existencia de subpoblaciones está presente pero éstas a la vez presentan heterogeneidad dentro de ellas, se comentó que una alternativa para tomar ésta heterogeneidad dentro de clases, es el uso de efectos aleatorios como sugerido por Yuan y Ji (2015) en el contexto Bayesiano a través de los modelos de trayectoria de degradación general (general degradation path model).

En el Capítulo 4 se propuso un método para el análisis de estudios de vida de anaquel,

como alternativa al análisis de supervivencia o confiabilidad propuesto por Hough et al. (2003). Esto por que en la metodología anterior no se toma en cuenta la incertidumbre de los intervalos de censura, la cual ésta presente en éste tipo de estudios debido a la inconsistencia en las respuestas de los consumidores. Esta fue ilustrada con dos ejemplos reales, y también se presentó un pequeño estudio de simulación para mostrar el efecto de ignorar la inconsistencia en la estimación de ciertos cuantiles. Además, el modelo propuesto también puede considerarse como un modelo de degradación binario, en donde la característica que observamos a través del tiempo, es la aceptación o rechazo del producto por cada consumidor y la falla (rechazo) del producto se declara cuando la función de probabilidad de aceptación  $p_i(t)$  alcanza un nivel de probabilidad  $p_f$ .

Como trabajo a futuro se plantearon las siguientes preguntas ¿Cómo deben elegirse los tiempos de almacenamiento de forma óptima?, y más generalmente, ¿Cómo diseñar el experimento para obtener la mayor información posible con los recursos disponibles?. También puede explorarse como incorporar información previa de experimentos similares para mejorar la precisión de las cantidades de interés.

En el Capítulo 5, por medio de simulación se estudió el desempeño de 4 criterios de selección de modelos, usando las distribuciones exponencial, Rayleigh y Weibull como el conjunto de distribuciones candidatas, diferentes valores del número nominal de fallas en una muestra y diferentes fracciones de censura aleatoria por la derecha fueron estudiadas. Los resultados aquí obtenidos indican que no es fácil elegir un mejor criterio ya que esto puede depender de las distribuciones candidatas, el número nominal de fallas ( $r$ ), y también de la fracción de censura presente. A pesar de esto, en general al parecer el  $BIC_M$  es el que resultó tener un comportamiento bastante regular y superior al resto después de  $r = 15$ . Sin embargo, se deben tener precauciones en el uso de estos índices como criterios de selección de modelos de datos de vida con presencia de censura y un número de fallas pequeño en la muestras, ya que algunas veces nos pueden llevar a elegir un modelo incorrecto y en consecuencia obtener conclusiones erróneas. Así, estos criterios pueden y deberían usarse solamente como una herramienta más en la elección del modelo, donde otros aspectos de la naturaleza del fenómeno debería también tomarse en cuenta.

Aunque la comparación podría extenderse a familias de distribuciones candidatas más grandes (por ejemplo, a la familia de la distribución gama generalizada y todos los submodelos de ésta), obtener una conclusión a cerca de cual de los criterios resulta en un comportamiento regular en comparación al resto, sería más complicado, además de que requerirían usar valores del número de fallas nominal superiores a los aquí usados y que muchas veces son más allá de los encontrados en casos prácticos.



# Referencias

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–723.
- Bae, S. J. y P. H. Kvam (2004). A nonlinear random-coefficients model for degradation testing. *Technometrics* 46(4), 460–469.
- Bates, D., M. Mächler, B. Bolker, y S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Breslow, N. E. y D. G. Clayton (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88(421), 9–25.
- Bucar, T., M. Nagode, y M. Fajdiga (2004). Reliability approximation using finite Weibull mixture distributions. *Reliability Engineering and System Safety* 84(3), 241–251.
- Burnham, K. P. y D. R. Anderson (2002). *Model Selection and Multimodel Inference: A Practical Information - Theoretic Approach*. Springer-Verlag, New York.
- Casella, G. y R. L. Berger (2002). *Statistical Inference*. Duxbury Press, Pacific Grove, California.
- Cheng, Y. S. y C. Y. Peng (2012). Integrated degradation models in R using iDEMO. *Journal of Statistical Software* 49(2), 1–22.
- Claeskens, G. y N. L. Hjort (2008). *Model Selection and Model Averaging*. Cambridge University Press, Cambridge.
- Collett, D. (2003). *Modelling Survival Data in Medical Research*. Chapman & Hall/CRC, Boca Raton, Florida.
- Davison, A. C. y D. V. Hinkley (1997). *Bootstrap methods and their application*. Cambridge, UK: Cambridge University Press.

- Demidenko, E. (2004). *Mixed Models: Theory and Applications*. Wiley-Interscience, Hoboken, New Jersey.
- Dempster, A. P., N. M. Laird, y D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B: Methodological* 39(1), 1–22.
- Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, , y D. F. Staples (2006). Estimating density dependence process noise and observation error. *Ecological Monographs* 76(3), 323–341.
- Dorai-Raj, S. y original Fortran from Mike Powell (2015). *powell: Powell's UObyQA algorithm*. R package version 1.0-0.
- Edwards, D., G. C. de Abreu, y R. Labouriau (2010). Selecting high-dimensional mixed graphical models using minimal aic or bic forests. *BMC Bioinformatics* 11(18), 1471–2105.
- Efron, B. y D. V. Hinkley (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected fisher information. *Biometrika* 65(3), 457–482.
- Efron, B. y R. Tibshirani (1993). *An introduction to the bootstrap*. London; New York: Chapman & Hall Ltd.
- Elandt, R. C. (1961). The folded normal distribution: Two methods of estimating parameters from moments. *Technometrics* 3(4), 551–562.
- Emiliano, P. C., M. J. Vivanco, y F. S. de Menezes (2014). Information criteria: How do they behave in different models? *Computational Statistics and Data Analysis* 69, 141–153.
- Everitt, B. S. y D. J. Hand (1981). *Finite mixture distributions*. Monographs on applied probability and statistics. London, New York: Chapman & Hall.
- Ferguson, T. S. (1996). *A Course in Large Sample Theory*. Chapman & Hall/CRC, Boca Raton, Florida.
- Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge University Press, Cambridge.

- German Rodriguez, N. G. (2001). Improved estimation procedures for multilevel models with binary response: A case-study. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 164(2), 339–355.
- Gertsbakh, I. B. y K. B. Kordonsky (1969). *Models of Failure. English translation from the Russian version*. New York: Springer-Verlag.
- Geweke, J. y R. Meese (1981). Estimating regression models of finite but unknown order. *International Economic Review* 22(1), 55–70.
- Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. *Biometrika* 78(1), 45–51.
- Goldstein, H. y J. Rasbash (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159(3), 505–513.
- Hamada, M. S., A. G. Wilson, C. S. Reese, y H. F. Martz (2008). *Bayesian Reliability*. Springer-Verlag New York.
- Hannan, E. J. (1982). Testing for autocorrelation and akaike’s criterion. *Journal of Applied Probability* 19, 403–412.
- Hanson, T. E. (2006). Modeling censored lifetime data using a mixture of Gammas baseline. *Bayesian Analysis* 1(3), 575–594.
- Hoeting, J. A., R. A. Davis, A. A. Merton, , y S. E. Thompson (2006). Model selection for geostatistical models. *Ecological Applications* 16(1), 87–98.
- Hough, G. y L. Garitta (2012). Methodology for sensory shelf-life estimation: A review. *Journal of Sensory Studies* 27(3), 137–147.
- Hough, G., K. Langohr, G. Gómez, y A. Curia (2003). Survival analysis applied to sensory shelf life of foods. *Journal of Food Science* 68(1), 359–362.
- Hurvich, C. M. y C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika* 76(2), 297–307.
- Hurvich, C. M. y C.-L. Tsai (1991). Bias of the corrected aic criterion for underfitted regression and time series models. *Biometrika* 78(3), 499–509.

- Jiang, S. y D. Kececioglu (1992). Maximum likelihood estimates, from censored data, for mixed-Weibull distributions. *IEEE Transactions on Reliability* 41(2), 248–255.
- JMP (2013). *Using JMP*. Cary, NC, USA: SAS Institute Inc.
- Johnson, J. B. y K. S. Omland (2004). Model selection in ecology and evolution. *Trends in Ecology and Evolution* 19(2), 101–108.
- Kleyle, R. M. y R. C. Dahiya (1975). Estimation of parameters of mixed failure time distribution from censored data. *Communications in Statistics* 4(9), 873–882.
- Klutke, G.-A., P. C. Kiessler, y M. A. Wortman (2003). A critical look at the bathtub curve. *IEEE Transactions on Reliability* 52(1), 125–129.
- Koehler, A. B. y E. S. Murphree (1988). A comparison of the akaike and schwarz criteria for selecting model order. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 37(2), 187–195.
- Konishi, S. y G. Kitagawa (2008). *Information Criteria and Statistical Modeling*. Springer, New York.
- Koziol, J. A. y S. B. Green (1976). A cramer-von mises statistic for randomly censored data. *Biometrika* 63(3), 465–474.
- Lawless, J. F. y M. Crowder (2004). Covariates and random effects in a Gamma process model with application to degradation and failure. *Lifetime Data Analysis* 10(3), 213–227.
- Lee, H. y S. K. Ghosh (2009). Performance of information criteria for spatial models. *Journal of Statistical Computation and Simulation* 79(1), 93–106.
- Lele, S. R., B. Dennis, y F. Lutscher (2007). Data cloning: easy maximum likelihood estimation for complex ecological models using bayesian markov chain monte carlo methods. *Ecology Letters* 10(7), 551–563.
- Lele, S. R., K. Nadeem, y B. Schmuland (2010). Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association* 105(492), 1617–1625.
- Leone, F. C., L. S. Nelson, y R. B. Nottingham (1961). The folded normal distribution. *Technometrics* 3(4), 543–550.

- Liang, H. y G. Zou (2008). Improved aic selection strategy for survival analysis. *Computational Statistics and Data Analysis* 52(5), 2538–2548.
- Liddle, A. R. (2007). Information criteria for astrophysical model selection. *Monthly Notices of the Royal Astronomical Society: Letters* 377(1), L74–L78.
- Liu, Q. y D. A. Pierce (1994). A note on Gauss-Hermite quadrature. *Biometrika* 81(3), 624–629.
- Lu, C. J. y W. Q. Meeker (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics* 34(2), 161–174.
- Lu, C. J., W. Q. Meeker, y L. A. Escobar (1996). A comparison of degradation and failure-time analysis methods of estimating a time-to-failure distribution. *Statistica Sinica* 6, 531–546.
- Lu, J. C., J. Park, y Q. Yang (1997). Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics* 39(4), 391–400.
- Marín, J. M., M. T. Rodríguez-Bernal, y M. P. Wiper (2005). Using Weibull mixture distributions to model heterogeneous survival data. *Communications in Statistics: Simulation and Computation* 34(3), 673–684.
- McLachlan, G. y D. Peel (2000). *Finite Mixture Models*. John Wiley & Sons, Inc., New York.
- McLachlan, G. J. y T. Krishnan (2008). *The EM Algorithm and Extensions* (Second ed.). Wiley, New York.
- Meeker, W. Q. y L. A. Escobar (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons Inc., New York.
- Mendenhall, W. y R. J. Hader (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data. *Biometrika* 45(3-4), 504–520.
- Miller, K. S. (1981). On the inverse of the sum of matrices. *Mathematics Magazine* 54(2), 67–72.
- Noor, F. y M. Aslam (2013). Bayesian inference of the inverse Weibull mixture distribution using type-i censoring. *Journal of Applied Statistics* 40(3), 1076–1089.

- Pawitan, Y. (2001). *In All Likelihood: modelling and inference using the likelihood*. Oxford University Press.
- Peña de la Mora, E. (2003). Desarrollo de una laca para recubrir acero inoxidable. Master's thesis, Centro de Investigación en Matemáticas A.C. (Guanajuato, Gto.).
- Peng, C.-Y. (2015). Inverse Gaussian processes with random effects and explanatory variables for degradation data. *Technometrics* 57(1), 100–111.
- Peng, C. Y. y S. C. Hsu (2012). A note on a wiener process with measurement error. *Applied Mathematics Letters* 25(4), 729–732.
- Peng, C. Y. y S. T. Tseng (2009). Miss-specification analysis of linear degradation models. *IEEE Trans. Reliability* 58(3), 444–455.
- Pinheiro, J. C. y D. M. Bates (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 4(1), 12–35.
- Quandt, R. E. y J. B. Ramsey (1978). Estimating mixtures of normal distributions and switching regressions. *Journal of the American Statistical Association* 73(364), 730–738.
- Rao, P. S. y N. A. Zaino (2004). Reliability estimation through the linear mixed effects model. *Journal of Statistical Planning and Inference* 129(1-2), 327–337.
- Robinson, M. E. y M. J. Crowder (2000). Bayesian method for a growth-curve degradation model with repeated measures. *Lifetime Data Analysis* 6(4), 357–374.
- Rodríguez-Narciso, S. (2013). *Análisis secuencial bayesiano para pruebas de degradación*. Ph. D. thesis, Centro de Investigación en Matemáticas, A. C.
- Rodriguez, G. y N. Goldman (1995). An assessment of estimation procedures for multilevel models with binary responses. *ournal of the Royal Statistical Society. Series A (Statistics in Society)* 158(1), 73–89.
- Samaniego, F. J. (2010). *A Comparison of the Bayesian and Frequentist Approaches to Estimation*. Springer-Verlag, New York.
- SAS-Institute-Inc. (2004). *SAS/STAT User's Guide, Version 9.1.3*. Cary, NC: SAS Institute Inc.

- Scallan, A. J. (1999). Fitting a mixture distribution to complex censored survival data using generalized linear models. *Journal of Applied Statistics* 26(6), 747–753.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* 6(2), 461–464.
- Seber, G. A. F. (2004). *Multivariate Observations*. John Wiley & Sons, Inc.
- Severini, T. A. (2005). *Elements of Distribution Theory*. Cambridge University Press.
- Shibata, R. (1976). Selection of the order of an autoregressive model by akaike’s information criterion. *Biometrika* 63(1), 117–126.
- Singpurwalla, N. D. (1995). Survival in dynamic environments. *Statistical Science* 10(1), 86–103.
- Smyth, G., Y. Hu, P. Dunn, B. Phipson, y Y. Chen (2016). *statmod: Statistical Modeling*. R package version 1.4.24.
- Sugiura, N. (1978). Further analysts of the data by akaike’s information criterion and the finite corrections. *Communications in Statistics-Theory and Methods* 7(1), 13–26.
- Team, R. C. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Titterton, D. M., A. F. M. Smith, y U. E. Makov (1985). *Statistical Analysis of Finite Mixture Distributions*. New York: John Wiley & Sons.
- Tsai, C.-C., Tamkang, S.-T. Tseng, y N. Balakrishnan (2012). Optimal design for degradation tests based on Gamma processes with random effects. *Reliability, IEEE Transaction* 61(2), 604–613.
- Tsai, C. C., S.-T. Tseng, y N. Balakrishnan (2011). Mis-specification analyses of Gamma and Wiener degradation processes. *Journal of Statistical Planning and Inference* 141(12), 3725–3735.
- Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society Series B (Methodological)* 38(3), 290–295.
- Volinsky, C. T. y A. E. Raftery (2000). Bayesian information criterion for censored survival models. *Biometrics* 56(1), 256–262.

- Wang, X. (2010). Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis* 101(2), 340–351.
- Wang, X. y D. Xu (2010). An inverse Gaussian process model for degradation data. *Technometrics* 52(2), 188–197.
- Weaver, B. P., W. Q. Meeker, L. A. Escobar, y J. Wendelberger (2013). Methods for planning repeated measures degradation studies. *Technometrics* 55(2), 122–134.
- Whitmore, G. A. (1986). Normal-Gamma mixtures of inverse Gaussian distributions. *Scandinavian Journal of Statistics* 13(3), 211–220.
- Whitmore, G. A. (1995). Estimation degradation by a Wiener diffusion process subject to measurement error. *Lifetime Data Analysis* 1(3), 307–319.
- Whitmore, G. A. y F. Schenkelberg (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis* 3(1), 27–45.
- Wolfinger, R. (1993). Laplace’s approximation for nonlinear mixed models. *Biometrika* 80(4), 791–795.
- Wu, C. F. J. (1983). On the convergence properties of the EM algorithm. *The Annals of Statistics* 11(1), 95–103.
- Ye, Z. S., K. L. Tsui, Y. Wang, y M. Pecht (2013). Degradation data analysis using a Wiener processes with measurement errors. *IEEE Transactions on Reliability* 62(4), 772–780.
- Ye, Z.-S., M. Xie, L.-C. Tang, y N. Chen (2014). Semiparametric estimation of Gamma processes for deteriorating products. *Technometrics* 56(4), 504–513.
- Yáñez, S. y R. A. Granada (2006). Comparación de metodologías para el análisis de datos de degradación para trayectorias lineales. *Revista Colombiana de Estadística* 29(2), 133–151.
- Yuan, T. y Y. Ji (2015). A hierarchical bayesian degradation model for heterogeneous data. *EEE Transactions on Reliability* 64(1), 63–70.
- Zacks, S. (1971). *The Theory of Statistical Inference*. New York: JohnWiley.