



**VNiVERSiDAD  
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

VALORACIÓN DE LA ESTABILIDAD Y EL  
RENDIMIENTO DE MODELOS BAYESIANOS  
EN META-ANÁLISIS DE PRUEBAS  
DIAGNÓSTICAS, BASADA EN ESTUDIOS DE  
SIMULACIÓN

Autor: Sergio Alex Bauz Olvera

Directoras:

Dra. M<sup>a</sup> Purificación Galindo Villardón

Dra. Ana Belén Nieto Librero

Doctorado en Estadística Multivariante Aplicada

Tesis Doctoral

Departamento de Estadística  
UNIVERSIDAD DE SALAMANCA  
España

Diciembre 2019

*Dedicado a*

A DIOS,  
a mi familia  
y a nuestros amigos.

# VALORACIÓN DE LA ESTABILIDAD Y EL RENDIMIENTO DE MODELOS BAYESIANOS EN META-ANÁLISIS DE PRUEBAS DIAGNÓSTICAS, BASADA EN ESTUDIOS DE SIMULACIÓN

Bauz-Olvera Sergio A.

## Resumen

En meta-análisis para pruebas de precisión diagnósticas (DTA) se utilizan medidas resumen como la sensibilidad, especificidad y odds ratio. Sin embargo, estas medidas pueden no ser adecuadas para integrar estudios con baja prevalencia. Los modelos jerárquicos son importantes, ya que modelizan la heterogeneidad estadística causada por distintos puntos de corte que existen en cada estudio y que se incluyen en un meta-análisis. Para evaluar el rendimiento de los modelos jerárquicos, se simuló 1200 meta-análisis empleando el paquete R, donde cada meta-análisis generado contiene aleatoriamente entre 5 y 35 estudios, cada estudio contiene aleatoriamente entre 200 hasta 2000 pacientes y prevalencias entre 0 y 0.25 distribuidos uniformemente. Las estimaciones de los parámetros e hiper-parámetros que identifican a los modelos bivalente y HSROC se obtuvieron con el Procedimiento NLMIXED de SAS. Estos resultados algorítmicos son almacenados en base de datos para investigar el rendimiento de cinco modelos jerárquicos (incluido el bivalente), los cuales son particularizaciones del modelo HSROC. La identificabilidad de los modelos se analiza de acuerdo a los criterios de convergencia de, Akaike (AIC), Akaike mejorado (AICC) y bayesiano (BIC). Del 100 % (1200) de meta-análisis simulados el 90 % se ajustaron a un modelo bivalente y 70 % a un modelo HSROC. Evidenciando que, la covarianza (logit de sensibilidad y especificidad) y la precisión diagnóstica (log DOR) se ajustaron a una normal. Este resultado indica que para valores de covarianzas no negativas entre la tasa de verdaderos y falsos positivos es plausible sugerir una medida resumen (sensibilidad y especificidad) y curva ROC.

## Abstract

In meta-analysis for diagnostic precision tests (DTA) summary measures such as sensitivity, specificity and odds ratio are used. However, these measures may not be adequate to integrate studies with low prevalence. Hierarchical models are important because they model the statistical heterogeneity caused by different cut-off points that exist in each study and that is included in a meta-analysis. To evaluate the performance of the hierarchical models, 1200 meta-analyses were simulated using the R package, where each generated meta-analysis randomly contains between 5 and 35 studies, in each study it randomly contains between 200 and 2000 patients and prevalence between 0 and 0.25 evenly distributed. The estimates of the parameters and hyper-parameters that identify the bivariate and HSROC models were obtained with the NLMIXED Procedure of SAS. These algorithmic results are stored in a database to investigate the performance of five hierarchical models (including the bivariate), which are particularizations of the HSROC model. The identifiability of the models is analyzed according to the convergence criteria of Akaike (AIC), Enhanced Akaike (AICC) and Bayesian (BIC). Of the 1200 simulated meta-analyses the (0.90 , 0.70) meta-analyses were adjusted to a model (bivariate, HSROC) respectively. Evidence that covariance (logit of sensitivity and specificity) and diagnostic precision (log DOR) were adjusted to a normal. This result indicates that for non-negative covariance values between the true and false positive rate it is plausible to suggest a summary measure (sensitivity and specificity) and ROC curve.



# VNiVERSiDAD D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

**Dra. M<sup>a</sup> Purificación Galindo Villardón**

Catedrática de la Universidad de Salamanca en el Área de Estadística e Investigación  
Operativa

y

**Dra. Ana Belén Nieto Librero**

Profesora Ayudante Doctora del Departamento de Estadística de la Universidad de  
Salamanca

CERTIFICAN:

Que **Don Sergio Alex Bauz Olvera** ha realizado en el Departamento de Estadística de la Universidad de Salamanca, bajo su dirección, el trabajo para optar al Grado de Doctor en Estadística Multivariante Aplicada, que presenta con el título **VALORACIÓN DE LA ESTABILIDAD Y EL RENDIMIENTO DE MODELOS BAYESIANOS EN META-ANÁLISIS DE PRUEBAS DIAGNÓSTICAS, BASADA EN ESTUDIOS DE SIMULACIÓN**, autorizando expresamente su lectura y defensa.

Y para que conste, firman el presente certificado en Salamanca el 10 de diciembre de 2019.

Dra. M<sup>a</sup> Purificación Galindo Villardón

Dra. Ana Belén Nieto Librero

# Agradecimientos

En primer lugar, a las autoridades de la Escuela Superior Politécnica del Litoral - Espol, a la señora Rectora, Decano, Subdecano y Profesores de las Facultades de Ciencias de la Vida y Ciencias Naturales y Matemáticas por su confianza y apoyo.

A Puri, Directora del Departamento de Estadística de la Universidad de Salamanca, por su guía personal como también científica, Ana Belén Nieto por su orientación y participación activa en el desarrollo de esta investigación y Ana Belén Sánchez que ha sido un motor fundamental para el presente trabajo. A los profesores y colaboradores del Departamento de Estadística de la Usal por sus enseñanzas, tiempo y colaboración prestada para la realización de este Trabajo de Doctorado.

A mi padres Noemí y Abel por sus sabios consejos, valores, que me transmitieron, los cuales me ha permitido ser una persona capaz y útil a la sociedad. A mi esposa Ivonn, por todo los esfuerzos y sacrificios que ha realizado. No hay palabras para describir, cuan agradecido estoy por su apoyo incondicional recibido durante el Doctorado.

A mis hijos Emily, Alejandro y Valentina, todavía son muy pequeños, pero con el pasar del tiempo, se darán cuenta que nos transmiten la fuerza necesaria aun en los momentos más adversos. A mis hermanos por su incondicional apoyo y predisposición para ayudarme cuando estaba fuera del país, siempre les estaré agradecido.

# Índice general

|  |          |
|--|----------|
| Resumen  | III      |
| Declaración  | V        |
| Agradecimientos  | VI       |
| <b>1. Introducción y Objetivos</b>   | <b>1</b> |
| 1.1. Introducción . . . . .  | 1        |
| 1.2. Objetivo General . . . . .  | 3        |
| 1.3. Objetivos Específicos . . . . .   | 4        |
| <b>2. Metodología Estadística</b>  | <b>5</b> |
| 2.1. Introducción . . . . .  | 5        |
| 2.2. Meta-análisis para DTA . . . . .  | 6        |
| 2.2.1. Objetivos . . . . .   | 6        |
| 2.2.2. Tipos de datos y el análisis de precisión de la prueba . . . . .          | 7        |
| 2.2.3. Sensibilidad y especificidad . . . . .                                    | 8        |
| 2.2.4. Valores predictivos . . . . .   | 9        |
| 2.2.5. Ratio de verosimilitud . . . . .  | 10       |
| 2.2.6. Odds ratios de Diagnóstico (DOR) . . . . .                                | 11       |
| 2.2.7. Umbrales de positividad y relación sensibilidad especificidad . . . . .   | 11       |
| 2.3. Presentación gráfica y tabular . . . . .                                    | 13       |
| 2.3.1. Curvas ROC . . . . .  | 13       |
| 2.3.2. Relaciones entre las curvas ROC, los odds-ratios de diagnóstico . . . . . | 14       |
| 2.3.3. Resumen de las curvas ROC . . . . .                                       | 15       |
| 2.3.4. Forest plots . . . . .  | 16       |
| 2.4. Resumen Meta-analítico . . . . .  | 16       |

---

|           |  |           |
|-----------|--|-----------|
| 2.4.1.    | Estimación de una curva SROC o un punto de consenso . . . . .            | 16        |
| 2.4.2.    | Métodos meta-analíticos no utilizados de forma rutinaria . . . . .       | 19        |
| 2.5.      | Heterogeneidad . . . . .   | 19        |
| 2.6.      | Ajuste del modelo . . . . .  | 21        |
| 2.6.1.    | Curvas SROC de Moses Littenberg . . . . .                                | 21        |
| 2.6.2.    | Propiedades de la Curva SROC de Moses Littenberg . . . . .               | 22        |
| 2.6.3.    | Elección de los pesos . . . . .  | 23        |
| 2.7.      | Modelos jerárquicos . . . . .  | 24        |
| 2.7.1.    | El modelo Bivariante . . . . .   | 25        |
| 2.7.2.    | El modelo HSROC de Rutter y Gatsonis . . . . .                           | 27        |
| 2.7.3.    | Inferencia Bayesiana . . . . .   | 29        |
| 2.7.4.    | Modelo Bayesiano para combinar datos de pruebas de diagnóstico . . . . . | 34        |
| 2.8.      | Modelos Mixtos No Lineales . . . . .                                     | 37        |
| 2.8.1.    | Introducción . . . . .   | 37        |
| 2.8.2.    | Estado del arte . . . . .  | 38        |
| 2.8.3.    | Supuestos de modelado y notación . . . . .                               | 39        |
| 2.8.4.    | Aproximaciones integrales . . . . .                                      | 41        |
| 2.8.5.    | Aproximaciones por diferencias finitas de derivadas . . . . .            | 44        |
| 2.8.6.    | Escala Hessiana . . . . .  | 46        |
| 2.8.7.    | Métodos de conjunto activo . . . . .                                     | 47        |
| 2.8.8.    | Métodos de búsqueda por líneas . . . . .                                 | 49        |
| 2.8.9.    | Restricción de la longitud del paso . . . . .                            | 50        |
| 2.8.10.   | Funciones del log de verosimilitud integradas . . . . .                  | 52        |
| 2.8.11.   | Especificación jerárquica de modelos . . . . .                           | 54        |
| 2.9.      | Resumen de los Modelos para DTA . . . . .                                | 58        |
| 2.10.     | Criterios para la selección del modelo . . . . .                         | 60        |
| <b>3.</b> | <b>Metodología Computacional</b> . . . . .                               | <b>61</b> |
| 3.1.      | Introducción . . . . .   | 61        |
| 3.2.      | Evaluación del Software disponible para DTA . . . . .                    | 62        |
| 3.2.1.    | Programa estadístico Stata . . . . .                                     | 62        |
| 3.2.2.    | Programa estadístico Sas . . . . .                                       | 63        |
| 3.2.3.    | Programa estadístico R . . . . .   | 64        |
| 3.3.      | Procedimiento NLMIXED . . . . .  | 66        |



|  |            |
|--|------------|
| 3.3.1. PROC NLMIXED Comparado con otros procedimientos y macros de SAS . . . . . | 66         |
| 3.3.2. Algoritmos de optimización . . . . .                                      | 67         |
| 3.3.3. Problemas de Computacionales . . . . .                                    | 75         |
| 3.3.4. Matriz de Covarianza . . . . .  | 81         |
| 3.3.5. Predicción . . . . .  | 82         |
| 3.3.6. Visualización de Salidas . . . . .  | 83         |
| 3.3.7. Nombre de Tablas ODS . . . . .  | 87         |
| 3.4. Aplicaciones del Procedimiento NLMIXED . . . . .                            | 88         |
| 3.4.1. Curvas de crecimiento no lineal con datos gaussianos . . . . .            | 88         |
| 3.4.2. Modelo Logístico-Normal con Datos Binomiales . . . . .                    | 92         |
| 3.4.3. Modelo de un compartimento con datos farmacocinéticos . . . . .           | 97         |
| 3.4.4. Modelo Probit-Normal con datos binomiales . . . . .                       | 102        |
| 3.4.5. Modelo Probit-Normal con Datos Ordinales . . . . .                        | 106        |
| 3.4.6. Modelo Poisson-Normal con datos de conteo . . . . .                       | 112        |
| <b>4. Estudio de Simulación</b>  | <b>116</b> |
| 4.1. Ejemplo motivador . . . . .   | 116        |
| 4.1.1. Meta-análisis para DTA Datos Reales . . . . .                             | 116        |
| 4.1.2. Meta-análisis para DTA Datos Simulados . . . . .                          | 119        |
| 4.1.3. Análisis para los dos Meta-análisis para DTA . . . . .                    | 120        |
| 4.2. Método de Simulación . . . . .  | 123        |
| 4.3. Generación de los Datos Simulados . . . . .                                 | 123        |
| 4.4. Modelos jerárquicos ajustados a los datos simulados . . . . .               | 126        |
| 4.5. Facilitando la Convergencia de los Modelos Jerárquicos . . . . .            | 127        |
| 4.6. Evaluación, Convergencia y Estabilidad del Modelo . . . . .                 | 127        |
| 4.7. Evaluación del Rendimiento de Modelos Meta-analíticos . . . . .             | 128        |
| <b>5. Resultados</b>   | <b>129</b> |
| 5.1. Meta-Análisis Simulados . . . . .   | 130        |
| 5.2. Tarea de Simulación . . . . .   | 131        |
| 5.3. Modelo Bivariante . . . . .   | 132        |
| 5.4. Modelo HSROC . . . . .  | 133        |
| 5.5. Estimación . . . . .  | 134        |

---

|                                   |     |
|-----------------------------------|-----|
| Conclusiones                      | 139 |
| Bibliografía                      | 141 |
| Apéndice                          | 151 |
| A. Código fuente de la simulación | 151 |
| A.1. Programa en SAS . . . . .    | 151 |
| A.2. Programa en R . . . . .      | 154 |

# Índice de figuras

|  |     |
|--|-----|
| 2.1. Umbral de prueba e impacto en la precisión del diagnóstico . . . . .  | 8   |
| 2.2. Aplicación de curvas ROC e Índice de Youden . . . . .   | 9   |
| 2.3. Relación Sen: 0.99 y Esp: 0.69 . . . . .  | 12  |
| 2.4. Relación Sen: 0.93 y Esp: 0.93 . . . . .  | 12  |
| 2.5. Relación Sen: 0.84 y Esp: 0.98 . . . . .  | 12  |
| 2.6. Relación Sen: 0.69 y Esp: 0.99 . . . . .  | 12  |
| 2.7. Umbrales explícitos para la curva ROC . . . . .   | 13  |
| 2.8. Relación entre las curvas ROC y DOR . . . . .   | 15  |
| 2.9. Gráfica de forest plot que muestra sensibilidad y especificidad específicas del estudio y una medida resumen de sensibilidad y especificidad con las correspondientes estadísticas de heterogeneidad. . . . . | 17  |
| 2.10. Curvas SROC para valores alternativos de los parámetros del modelo . . . . .   | 23  |
| 2.11. Gráfico acíclico dirigido (DAG) del modelo estructural bivalente para meta-análisis de pruebas de diagnóstico . . . . .  | 37  |
| 4.1. Forest Plot para sensibilidad y especificidad para datos AuditC . . . . .   | 117 |
| 4.2. Gráficos crosshair y rocellipse para datos AuditC . . . . .   | 118 |
| 4.3. Gráficos crosshair y rocellipse para datos AuditC . . . . .   | 118 |
| 4.4. Forest Plot para sensibilidad y especificidad para datos Simulados . . . . .  | 120 |
| 4.5. Gráficos crosshair y rocellipse para datos Simulados . . . . .  | 121 |
| 4.6. Gráficos crosshair y rocellipse para datos AuditC . . . . .   | 121 |
| 4.7. Factores que intervienen en la simulación modelo HSROC . . . . .  | 123 |
| 4.8. Esquema de Generación de las tablas 2x2 para DTA . . . . .  | 124 |
| 4.9. Parámetros para Generación de tablas 2x2 para DTA . . . . .   | 125 |
| 4.10. Meta-análisis Generados (dataSet), espacio ROC . . . . .   | 125 |
| 4.11. Parámetros de los modelos HSROC y Bivalente . . . . .  | 126 |

---

|   |     |
|---|-----|
| 5.1. Esquema general de trabajo y software requerido para la simulación de dataset, almacenamiento en base de datos y análisis estadístico para la presentación de resultados . . . . . | 129 |
| 5.2. Box Plot para sensibilidad y especificidad para datos AuditC . . . . .   | 130 |
| 5.3. Curvas SROC seleccionadas aleatoriamente de acuerdo a los rangos de prevalencia [0.05 , 0.10), [0.10 , 0.15) y [0.15 , 0.20) respectivamente . . . . .                             | 135 |
| 5.4. Curvas SROC seleccionadas aleatoriamente de acuerdo a los rangos de prevalencia [0.05 , 0.10), [0.10 , 0.15) y [0.15 , 0.20) respectivamente . . . . .                             | 136 |
| 5.5. Ajuste normal para la covarianza del Logit (Sensibilidad y Especificidad), y el log (DOR) para el modelo Bivariante. . . . .   | 137 |
| 5.6. Covarianza de Sensibilidad y Estabilidad vs. Ajuste (-2 Probabilidad Logarítmica) Número de Estudios en Meta-análisis . . . . .  | 138 |

# Índice de tablas

|  |    |
|--|----|
| 1.1. Estructura de los modelos ajustados a los datos simulados . . . . .   | 3  |
| 2.1. Prueba diagnóstica y el estado de la enfermedad en Tabla 2x2 . . . . .  | 8  |
| 2.2. Modelo anidado de tres niveles . . . . .  | 54 |
| 2.3. Comparación de Meta-Análisis de Estudios Terapéuticos/Intervencionales<br>y Estudios de Precisión de Pruebas de Diagnóstico . . . . . | 58 |
| 2.4. Métodos estadísticos para el resumen meta-analítico de los estudios de pre-<br>cisión en pruebas de diagnóstico . . . . .             | 59 |
| 2.5. Implementación de métodos y software para el meta-análisis de estudios de<br>precisión de pruebas diagnósticas . . . . .              | 60 |
| 3.1. Criterio e implicación de los semáforos . . . . .   | 62 |
| 3.2. Evaluación de paquetes para el meta-análisis de la precisión diagnóstica . .  | 66 |
| 3.3. Técnicas de optimización . . . . .  | 68 |
| 3.4. Derivadas por Técnicas de optimización . . . . .  | 69 |
| 3.5. Tablas de ODS producidas por PROC NLMIXED . . . . .   | 87 |
| 3.6. Especificaciones del modelo . . . . .   | 89 |
| 3.7. Dimensiones para el modelo de curva de crecimiento . . . . .  | 90 |
| 3.8. Valores iniciales de las estimaciones de parámetros y probabilidad de regis-<br>tro negativo . . . . .                                | 90 |
| 3.9. Historia de la iteración para el modelo de curva de crecimiento . . . . .   | 91 |
| 3.10. Estadísticas de ajuste para el modelo de curva de crecimiento . . . . .  | 91 |
| 3.11. Estimaciones de parámetros en la convergencia . . . . .  | 92 |
| 3.12. Información del Modelo Logístico-Normal . . . . .  | 94 |
| 3.13. Dimensiones para el Modelo Logístico-Normal . . . . .  | 95 |
| 3.14. Valores iniciales paravestimaciones de parámetros . . . . .  | 95 |
| 3.15. Historial de iteración y estadísticas de ajuste para el modelo Logístico-Normal  | 96 |

|  |     |
|--|-----|
| 3.16. Estadísticas de ajuste para el modelo Logístico-Normal con Datos Binomiales  | 96  |
| 3.17. Estimación de Parámetros para el Modelo Logístico-Normal . . . . .   | 96  |
| 3.18. Tabla de Estimaciones Adicionales . . . . .  | 96  |
| 3.19. Especificación del modelo para un compartimento . . . . .  | 99  |
| 3.20. Tabla de dimensiones para el modelo de un compartimento . . . . .  | 99  |
| 3.21. Valores iniciales e historial de iteraciones . . . . .   | 99  |
| 3.22. Valores iniciales e historial de iteraciones . . . . .   | 100 |
| 3.23. Estadísticas de ajuste para el modelo de un solo compartimento . . . . .   | 100 |
| 3.24. Cálculo de parámetros para el modelo de un solo compartimento . . . . .  | 101 |
| 3.25. Especificaciones . . . . .   | 104 |
| 3.26. Dimensiones . . . . .  | 104 |
| 3.27. Valores iniciales . . . . .  | 104 |
| 3.28. Historial de iteración para el modelo Probit-Normal . . . . .  | 105 |
| 3.29. Estadísticas de ajuste para el modelo Probit-Normal . . . . .  | 105 |
| 3.30. Estimaciones de parámetros para el modelo Probit-Normal . . . . .  | 106 |
| 3.31. Estimaciones adicionales . . . . .   | 106 |
| 3.32. Especificaciones para el modelo Ordinal Data . . . . .   | 109 |
| 3.33. Tabla de dimensiones para el modelo de datos ordinales . . . . .   | 110 |
| 3.34. Parámetro Valores iniciales y Log Negativo de verosimilitud . . . . .  | 110 |
| 3.35. Historia de la iteración . . . . .   | 111 |
| 3.36. Estadísticas de ajuste para el modelo de datos ordinales . . . . .   | 111 |
| 3.37. Estimaciones de parámetros en la convergencia . . . . .  | 111 |
| 3.38. Estimaciones de Correlación de Umbrales e Intraclase . . . . .   | 112 |
| 3.39. Tabla de dimensiones para el modelo Poisson-Normal . . . . .   | 113 |
| 3.40. Historial de iteración para el modelo Poisson-Normal . . . . .   | 114 |
| 3.41. Estadísticas de ajuste para el modelo Poisson-Normal . . . . .   | 114 |
| 3.42. Estimaciones de parámetros . . . . .   | 115 |
| 3.43. Estimaciones adicionales . . . . .   | 115 |
| 4.1. Estimaciones de la sensibilidad (Sen.), especificidad (Esp.), valor predictivo<br>positivo y negativo (VPP, VPN), razón de verosimilitud positiva y negativa<br>(LR+,LR-) y odds ratio de diagnóstico (DOR), <b>datos AUDIT-C</b> . . . . . | 117 |

---

|      |  |     |
|------|--|-----|
| 4.2. | Estimaciones de la sensibilidad (Sen.), especificidad (Esp.), valor predictivo positivo y negativo (VPP, VPN), razón de verosimilitud positiva y negativa (LR+,LR-) y diagnostic odds ratio (DOR), <b>datos simulados</b> . . . . .              | 119 |
| 4.3. | Estimaciones resúmenes para los datos, AUDIT-C y simulados mediante el modelo HSROC con el programa SAS . . . . .  | 122 |
| 4.4. | Estructura de los modelos ajustados a los datos simulados . . . . .  | 126 |
| 5.1. | Resumen de dataset simulados por número de estudios y rangos de prevalencias . . . . .   | 130 |
| 5.2. | Resumen de dataset simulados por número de estudios y celdas con ceros en las tablas tetracóricas . . . . .  | 131 |
| 5.3. | Estadística de bondad de ajuste para la convergencia, estimación de los parámetros a posteriori y sus intervalos de confianza para el modelo Bivariante incorporando o nó una covariable . . . . .   | 132 |
| 5.4. | Estadística de bondad de ajuste para la convergencia, estimación de los parámetros a posteriori y sus intervalos de confianza para el modelo HSROC incorporando o nó una covariable (en el umbral, la precisión y la forma de la curva). . . . . | 133 |
| 5.5. | meta-análisis que resultaron no estables para el procedimiento NLMIXED, según rango de prevalencia, número de estudios y ajuste . . . . .  | 134 |

# Capítulo 1

## Introducción y Objetivos

### 1.1. Introducción

El crecimiento exponencial de la literatura médica y el uso cada vez más generalizado de las tecnologías de la información y la comunicación, junto con la dispersión de los conocimientos científicos la literatura, dificultan a los investigadores, y sobre todo a los profesionales de la salud, el acceder a la información pertinente. El meta-análisis es una revisión sistemática que incorpora una estrategia estadística para integrar los resultados de varios estudios en una sola estimación. La información de partida es una tabla tetracórica en la que se miden medidas como Verdaderos Positivos (VP), Falsos Positivos (FP), Falsos Negativos (FN) y Verdaderos Negativos (VN).

Los métodos estadísticos para el meta-análisis de las pruebas de precisión diagnósticas (DTA) se han centrado en combinar y contrastar las medidas del rendimiento de la prueba a través de múltiples estudios analizando las diferencias como reales, probablemente explicadas por el azar, o explicables por diferencias conocidas en las características del estudio [95]. Tiene como objetivo cuantificar y comparar estas estadísticas para una o más pruebas diagnósticas para describir qué tan bien cada prueba clasifica a los individuos, y estimar y comparar las tasas de error probables (diagnósticos falsos positivos y falsos negativos) que se pueden encontrar.

Para el meta-análisis de una sola prueba diagnóstica, se han propuesto varios modelos para explicar conjuntamente la sensibilidad y especificidad de la prueba, donde la sensi-



bilidad se define como la probabilidad de que la prueba identifique correctamente a un paciente que tiene la enfermedad y la especificidad como la probabilidad de que la prueba identifique correctamente a un paciente sin enfermedad.

Estos meta-análisis tienen características importantes: (1) la sensibilidad y la especificidad estimadas de una prueba de diagnóstico suelen estar correlacionadas negativamente debido al equilibrio entre estas mediciones de precisión; (2) puede haber heterogeneidad sustancial entre los estudios debido a la variación clínica o metodológica entre los estudios, por ejemplo, diferencias en los umbrales de positividad o variación en los participantes; (3) la continuidad de los rasgos medibles utilizados para determinar el estado de la enfermedad puede imponer correlaciones entre la sensibilidad, la especificidad y la prevalencia de la enfermedad [96] y (4) si la prueba de referencia no es una prueba de referencia oro, es necesario considerar su naturaleza imperfecta. Estas características deben incorporarse en los meta-análisis de las pruebas de diagnóstico.

Para evaluar la precisión diagnóstica de una prueba y evaluar el equilibrio entre sensibilidad y especificidad en un metanálisis, se han propuesto varios modelos de efectos fijos y aleatorios. Para los meta-análisis en los que la prueba de referencia de cada estudio es un estándar de oro, se ha desarrollado la curva de la característica operativa del receptor resumida (SROC) ([97]; [98]; [99]) que modela la relación entre la sensibilidad y la especificidad en los estudios mediante modelos de regresión de efectos fijos.

El modelo de efectos aleatorios BIVARIANTE [100], parte de la asunción de que los logit de sensibilidad y especificidad siguen una distribución normal bivalente. El modelo jerárquico de características operativas del receptor (HSROC) captura la heterogeneidad entre los estudios [101], es el que combina estimaciones de sensibilidad y especificidad específicas de cada estudio mediante un modelo de efectos aleatorios, en ausencia de una prueba oro estandar.

Para investigar el rendimiento de estos modelos utilizamos la simulación estadística, generando diversos escenarios de prevalencia, analizando los criterios de convergencia del algoritmo y la identificabilidad de los parámetros e hiperparámetros.

Simulamos 1200 meta-análisis utilizando el paquete HSROC de R, cada meta-análisis

generado asigna aleatoriamente el número de estudios, el número de pacientes y los rangos de prevalencia (entre 5 y 35), (entre 200 y 2000), (entre 0 y 0.25) respectivamente, las estimaciones de los parámetros e hiperparámetros que identifican los modelos se obtuvieron con la ejecución del Procedimiento NLMIXED (modelo mixto no lineal) de SAS. En ambos modelos hemos incorporado o no una covariable, pero sólo en HSROC variamos los parámetros de precisión, umbral y la forma de la curva para investigar cada modelo por separado ver Tabla 1.1.

| No. | Modelos                            | Covariable |    |
|-----|------------------------------------|------------|----|
|     |                                    | Si         | No |
| 1   | BIVARIANTE                         | ✓          | ✓  |
| 2   | HSROC Completo                     | ✓          |    |
| 3   | HSROC simétrico                    | ✓          |    |
| 4   | HSROC con umbrales fijos           | ✓          |    |
| 5   | HSROC con precisión fija           | ✓          |    |
| 6   | HSROC con precisión y umbral fijos | ✓          |    |

Tabla 1.1: Estructura de los modelos ajustados a los datos simulados

La identificabilidad de los modelos se analiza en función de los criterios de convergencia de Akaike (AIC), Akaike Mejorada (AICC) y Bayesiana (BIC). 90 % y 70 % de estos estudios se ajustaron satisfactoriamente a los modelos bivariante y HSROC respectivamente.

Las pruebas demostraron que la covarianza (entre sensibilidad y especificidad logit) y la precisión diagnóstica (log DOR) fueron ajustados a una distribución normal. Este resultado indica que para valores de covarianza no negativos entre la tasa de verdaderos y falsos positivos es plausible sugerir una medida resumida de sensibilidad, especificidad y curva ROC. Por el contrario, en general, podemos decir que los meta-análisis que no se ajustaron satisfactoriamente a un modelo bivariante estaban conformados por pocos estudios (entre 5 y 8), pacientes (entre 200 y 1000) y una prevalencia entre (0 y 0.12).

## 1.2. Objetivo General

Investigar el rendimiento de los modelos jerárquicos que se emplean en el estudio de meta-análisis para pruebas diagnósticas, generando diversos escenarios de simulación para la modelización estadística, que permita conocer los criterios de convergencia del algoritmo y la identificabilidad de los parámetros e hiper-parámetros del modelo.

## 1.3. Objetivos Específicos

Los objetivos específicos de este trabajo son los siguientes:

1. Realizar una revisión bibliográfica para conocer el estado del arte en relación a la evaluación de las pruebas diagnósticas.
2. Conocer las particularidades que presenta la integración de resultados de pruebas diagnósticas cuando los estudios son heterogéneos.
3. Realizar un estudio crítico de los modelos bivalente y jerárquico HSROC.
4. Investigar el rendimiento de los distintos modelos jerárquicos que se generan en la simulación.
5. Implementar interfaces en C#.NET, Base de Datos y R para facilitar la tarea de simulación.
6. Realizar escenarios críticos para el manejo de la variabilidad dentro y entre estudios, con los datos almacenados de los distintos modelos jerárquicos simulados.

Se puede decir que la prueba diagnóstica es un meta-análisis diferente, por tal razón es necesario conocer algunos conceptos importantes que encierra el meta-análisis antes de abordar el estudio de pruebas diagnósticas, en el Capítulo 2 se enuncian algunos puntos importantes del meta-análisis.

# Capítulo 2

## Metodología Estadística

### 2.1. Introducción

Los profesionales de la salud (principalmente médicos) utilizan pruebas diagnósticas para determinar si un individuo (generalmente un paciente) tiene o no una enfermedad o condición en particular. Las revisiones de la precisión de las pruebas diagnósticas (DTA) proporcionan información sobre qué tan bien las pruebas distinguen a los pacientes con la enfermedad de los que no la padecen. La mayoría de las pruebas son imperfectas y se producen errores. Por lo tanto, los métodos estadísticos se centran en dos medidas estadísticas de la precisión diagnóstica, la sensibilidad de la prueba, que mide la proporción de **individuos enfermos** correctamente diagnosticados, es decir, es la probabilidad de que la **prueba sea positiva** en presencia de la enfermedad y la especificidad de la prueba, que mide la proporción de resultados negativos en **individuos que no padecen la enfermedad**, es decir, mide la probabilidad de que la **prueba sea negativa** en individuos sanos. Una revisión DTA tiene como objetivo cuantificar y comparar estas estadísticas para una o más pruebas diagnósticas para describir qué tan bien cada prueba clasifica a los individuos, estimar y comparar las tasas de error probables (diagnósticos falsos positivos y falsos negativos) que se pueden encontrar.

## 2.2. Meta-análisis para DTA

### 2.2.1. Objetivos

El meta-análisis de los estudios de precisión de las pruebas diagnósticas proporciona resúmenes de los resultados de los estudios pertinentes incluidos: proporciona una estimación de la precisión diagnóstica media de una prueba o pruebas, la incertidumbre de esta media y la variabilidad de los resultados de los estudios en torno a las estimaciones. Los modelos de regresión meta-analíticos pueden comparar estadísticamente la precisión de dos o más pruebas diagnósticas diferentes y describir cómo varía la precisión de la prueba con los umbrales de la prueba y otras características del estudio.

El meta-análisis ayuda a dar sentido a resultados de estudios aparentemente contradictorios, ya que identifica qué diferencias son probables de ser reales, cuáles son explicables por casualidad y cuáles pueden ser explicadas por diferencias conocidas en las características del estudio. A medida que la precisión de las estimaciones aumenta con la cantidad de datos, el meta-análisis puede tener más poder para detectar diferencias reales en la precisión de la prueba que los estudios individuales, y puede producir estimaciones más precisas de la sensibilidad y especificidad esperadas.

Al cuantificar la variabilidad de la precisión de la prueba en muchos entornos, el meta-análisis puede proporcionar información sobre la consistencia de los resultados de las pruebas. Los modelos de meta-análisis también proporcionan un marco para comparar la precisión de las pruebas que no han sido comparadas directamente en las pruebas individuales de los estudios.

La heterogeneidad es de esperar en los resultados de los estudios de precisión de las pruebas, por lo tanto, los modelos de efectos aleatorios, son necesarios para describir la variabilidad de la precisión de la prueba en todos los estudios.

Los métodos para realizar análisis que tengan en cuenta tanto la sensibilidad como la especificidad, y la relación entre ellos, y la heterogeneidad en la precisión de la prueba, requieren un ajuste jerárquico. Los modelos de efectos aleatorios, que están más allá de las capacidades analíticas de los modelos de meta-análisis clásicos, deben llevarse a cabo en

paquetes de software especializados y de programación estadística sofisticada como SAS, Stata, S-Plus, R, MLwiN o winBUGS/OpenBUGS.

### 2.2.2. Tipos de datos y el análisis de precisión de la prueba

Las revisiones sistemáticas de la precisión de las pruebas de diagnóstico incluyen los resultados de las pruebas de uno o más de los tres tipos de datos siguientes:

1. Binario (dicotómico), en el cual el resultado de la prueba se reporta como sí o no, positivo o negativo.
2. Ordinal, en el cual el resultado de la prueba se reporta en un conjunto de categorías ordenadas, a menudo con descriptores verbales, tales como: 1=definidamente normal, 2=presumiblemente normal, 3=equivocado, 4=presumiblemente anormal, 5=definitivamente anormal.
3. Continuo o Conteo, en el cual el resultado de la prueba se reporta en una escala continua o en un conteo, como la concentración de una sustancia o el número de características observadas.

Para ser incluidos en un meta-análisis, los resultados de una prueba ordinal, de conteo o continua deben ser re-clasificados como binario seleccionando un umbral y presentando los datos como una tabla 2x2.

Habiendo elegido un umbral particular para la positividad de la prueba, los datos de un estudio primario pueden ser presentado en una tabla 2x2 que muestra la clasificación cruzada de la situación de la enfermedad (resultado de la referencia) y el resultado de la prueba (resultado de la prueba de índice) como en la Tabla 2.1

La información de partida es una tabla tetracórica en la que aparecen medidas como Verdadero Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) y Verdadero Negativo (VN) y el estado de la enfermedad, esto muestra la concordancia entre el resultado de la prueba en forma binaria y la información sobre la enfermedad Figura 2.1.

| Resultado       | Estado de la Enfermedad |              |                   |
|-----------------|-------------------------|--------------|-------------------|
|                 | Enfermos                | No Enfermos  | Total             |
| Prueba positiva | VP (a)                  | FP (b)       | (a+b)             |
| Prueba negativa | FN (c)                  | VN (d)       | (c+d)             |
| <b>Total</b>    | <b>(a+c)</b>            | <b>(b+d)</b> | <b>N(a+b+c+d)</b> |

Tabla 2.1: Prueba diagnóstica y el estado de la enfermedad en Tabla 2x2

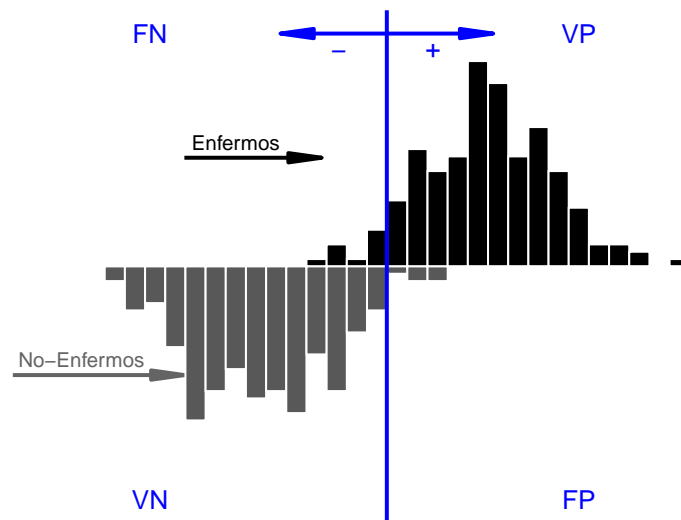


Figura 2.1: Umbral de prueba e impacto en la precisión del diagnóstico

### 2.2.3. Sensibilidad y especificidad

La sensibilidad de una prueba se define como la probabilidad de que el resultado de la prueba sea positivo en un caso de enfermedad. Formalmente, la sensibilidad =  $P(\text{Prueba}+ | \text{Enfermedad}+)$  y se estima usando los números de la tabla como  $a/(a+c)$ . La sensibilidad se conoce a veces como Tasa de detección (TD), Tasa de verdaderos positivos (TVP) o Tasa de fracción positiva (TFP). Se expresa como una proporción o un porcentaje.

La especificidad de una prueba se define como la probabilidad de que el resultado de la prueba sea negativo en un caso de no enfermo. Formalmente, especificidad =  $P(\text{Prueba}- | \text{No Enfermo}-)$  y se estima usando los números de la tabla como  $d/(b+d)$ . Ocasionalmente se hace referencia a la especificidad como la Tasa de verdaderos negativos (TVN) o la Fracción de verdaderos negativos (FVN). Con mayor frecuencia, los términos Tasa de falsos positivos (TFP) y Fracción de falsos positivos (FFP) son utilizados para el complemento

de la especificidad (calculado como  $1 - \text{especificidad}$  o  $b/(b+d)$ ). Una vez más, ambos se utilizan como proporciones o porcentajes.

Aunque los términos fracción positiva verdadera y fracción positiva falsa son técnicamente más correctos porque la sensibilidad y la especificidad son fracciones y no tasas, la tasa de verdaderos positivos y la tasa de falsos positivos son los términos más comunes. Los valores de sensibilidad y especificidad se combinan ocasionalmente en una medida conocida como índice de Youden (Figura 2.2) calculado como  $\text{sensibilidad} + \text{especificidad} - 1$ . El índice de Youden no tiene probabilidades directas pero proporciona un índice general de precisión de la prueba que da igual peso a los errores de la prueba (falsos positivos y falsos negativos). Los valores cercanos a 1 indican alta precisión; un valor de cero indica que una prueba no tiene valor diagnóstico.

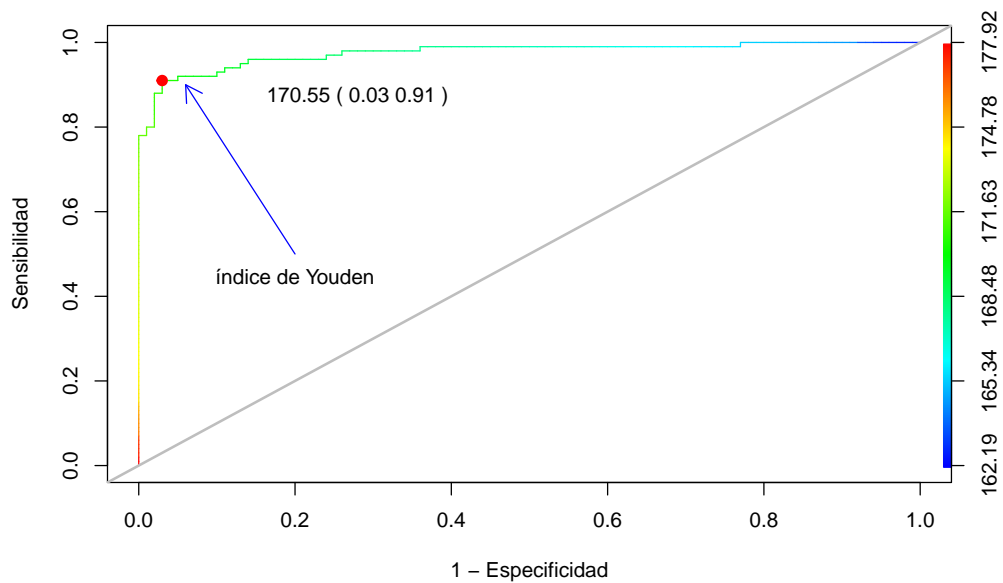


Figura 2.2: Aplicación de curvas ROC e Índice de Youden

#### 2.2.4. Valores predictivos

Los valores predictivos son medidas definidas condicionadas a los resultados de las pruebas, ya que se calculan como proporciones del total con resultados positivos y negativos en las pruebas.

El valor predictivo positivo de una prueba se define como la probabilidad de que un



caso con un resultado positivo en la prueba esté enfermo. Formalmente, el valor predictivo positivo= $P(\text{Enfermo} + | \text{Prueba} +)$  y se estima usando los números de la tabla como  $a/(a+b)$ . Una vez más, los valores predictivos positivos se reportan como proporciones o porcentajes.

El valor predictivo negativo de una prueba se define como la probabilidad de que un caso con un resultado negativo en la prueba no esté enfermo. Formalmente, el valor predictivo negativo= $P(\text{No Enfermo} - | \text{Prueba} -)$  y se estima utilizando los números de la tabla como  $d/(c+d)$ . Una vez más, los valores predictivos negativos se informan como proporciones o porcentajes.

### 2.2.5. Ratio de verosimilitud

Los ratios de verosimilitudes de probabilidad pueden utilizarse para actualizar la probabilidad de la enfermedad antes de la prueba utilizando el teorema de Bayes, una vez que se conozca el resultado de la prueba. La probabilidad actualizada se denomina probabilidad a posterior al ensayo. Para una prueba que sea informativa, la probabilidad a posterior a la prueba debe ser superior a la probabilidad previa a la prueba si el resultado de la prueba es positivo, mientras que la probabilidad después de la prueba debe ser inferior a la probabilidad previa de la prueba si el resultado de la prueba es negativo.

El ratio de verosimilitud positiva describe cuántas veces más probables fueron los resultados positivos de la prueba en el grupo enfermo en comparación con el grupo no enfermo. El cociente positivo de probabilidad, que debería ser mayor que 1 si el valor informativo de la prueba se define como:

$$LR+ = \frac{P(\text{Prueba} + | \text{Enfermo} +)}{P(\text{Prueba} + | \text{NoEnfermo} -)} = \frac{\text{sens}}{(1 - \text{spec})} \quad (2.2.1)$$

y se estima como  $(a/(a + c))/(b/(b + d))$

El ratio de verosimilitud negativa describe cuántos veces menos probables fueron los resultados negativos de la prueba en el grupo enfermo en comparación con el grupo no enfermo. El cociente de probabilidad negativo, que debería ser menos de 1 si el valor informativo de la prueba, se define como:

$$LR- = \frac{P(\text{Prueba} - | \text{Enfermo}+)}{P(\text{Prueba} - | \text{NoEnfermo}-)} = \frac{(1 - \text{sens})}{\text{spec}} \quad (2.2.2)$$

y se estima como  $(c/(a + c))/(d/(b + d))$

### 2.2.6. Odds ratios de Diagnóstico (DOR)

El odds ratios de diagnóstico (DOR) resume la precisión diagnóstica de la prueba de índice como un solo número que describe cuántas veces más probabilidades hay de obtener un resultado positivo en un individuo enfermo que en un individuo no enfermo. El hecho de que resuma la precisión de la prueba en un solo número facilita el uso de esta medida para el meta-análisis, pero el hecho de expresar la precisión en términos de proporciones de probabilidades significa que la medida tiene poca relevancia clínica directa, y rara vez se utiliza como estadística de resumen en estudios primarios. De hecho, el médico suele estar interesado en la suma del número de falsos negativos y falsos positivos, mientras que el DOR refleja su producto. Sin embargo, el DOR sigue siendo un elemento importante en la construcción de modelos meta-analíticos. Se define formalmente como:

$$DOR = \frac{\left(\frac{\text{sensibilidad}}{1 - \text{sensibilidad}}\right)}{\left(\frac{1 - \text{especificidad}}{\text{especificidad}}\right)} = \frac{LR+}{LR-} \quad (2.2.3)$$

y se estima como  $\frac{ad}{bc}$

### 2.2.7. Umbrales de positividad y relación sensibilidad especificidad

Los resultados de las pruebas binarias se definen sobre la base de un umbral para la positividad de la prueba y el cambio si se altera el umbral. Esta dependencia del umbral es un aspecto fundamental en la evaluación de la prueba de diagnóstico. En el caso de la sensibilidad y especificidad, la dependencia induce a una compensación entre las dos cantidades, un valor aumenta mientras que el otro disminuye cuando el umbral para la positividad se mueve.

Esto se ilustra en las figuras 2.3-2.4-2.5 y 2.6, cada una de las cuales muestra las mismas distribuciones hipotéticas de los resultados de las pruebas para personas enfermas y sanas en escala continua. Los paneles varían en el valor numérico del umbral de enfermedad

utilizado para definir el resultado positivo de la prueba. En cada umbral, la sensibilidad de la prueba se mide en función de la proporción de la superficie sometida a la prueba. Curva 'Enfermos' a la derecha del umbral. Del mismo modo, la especificidad se mide por la proporción del área bajo la curva "Sanos" a la izquierda del umbral.

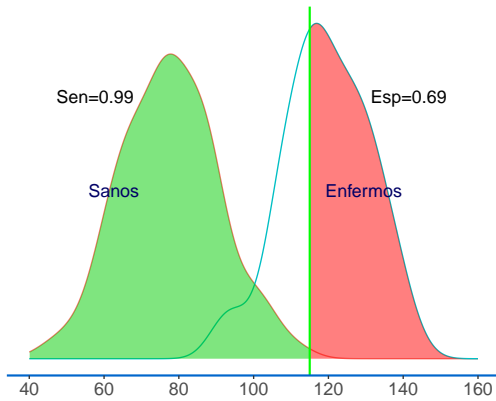


Figura 2.3: Relación Sen: 0.99 y Esp: 0.69

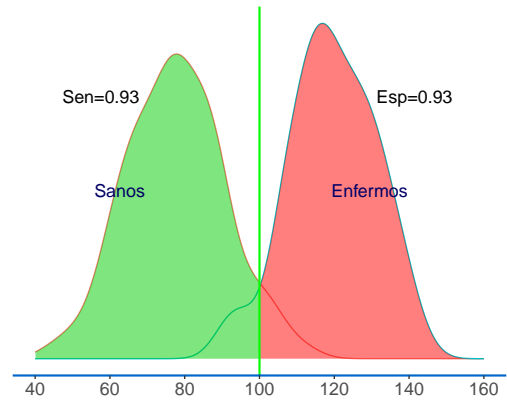


Figura 2.4: Relación Sen: 0.93 y Esp: 0.93

A medida que el umbral disminuye desde los paneles de las figuras 2.3-2.4-2.5 y 2.6, la proporción de personas enfermas que se encuentran por encima del umbral, por lo tanto tienen una prueba positiva, aumenta del 0.69 al 0.99. Estas cifras indican la sensibilidad de la prueba. Al mismo tiempo, la proporción de personas sanas que se encuentran por debajo del umbral, por lo tanto tienen un resultado negativo en la prueba, disminuye del 0.99 al 0.69. Estas cifras indican la especificidad de la prueba.

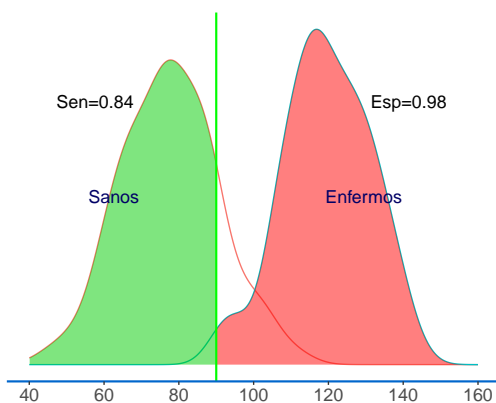


Figura 2.5: Relación Sen: 0.84 y Esp: 0.98

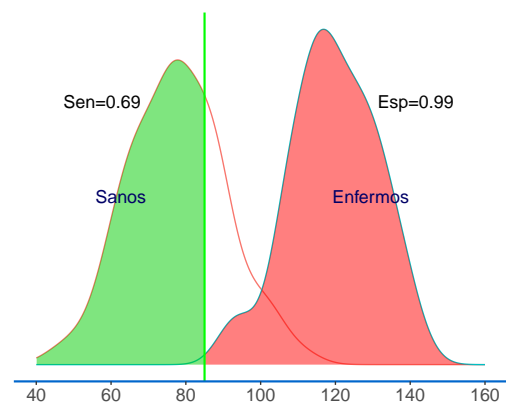


Figura 2.6: Relación Sen: 0.69 y Esp: 0.99

## 2.3. Presentación gráfica y tabular

### 2.3.1. Curvas ROC

Los estudios primarios que evalúan una prueba en varios umbrales a veces presentan resultados como curvas ROC. La curva ROC de una prueba es el gráfico de los valores de sensibilidad y especificidad que se obtienen variando el umbral de positividad entre todos los valores posibles. El gráfico muestra la sensibilidad (positivo verdadero) frente a 1-especificidad (tasa de falsos positivos). La curva de cualquier prueba se mueve desde el punto donde sensibilidad y 1-especificidad son 1 (la esquina superior derecha), lo que se logra para un umbral del extremo inferior de su rango (clasificando a todos los participantes como positivos en la prueba, de modo que no haya falsos negativos pero muchos falsos positivos) hasta un punto en el que tanto la sensibilidad como 1-especificidad son cero (la parte inferior izquierda que se consigue cuando el umbral se desplaza al extremo superior de su rango (y a todos los participantes se clasifican como negativos en la prueba, sin dar falsos positivos, pero con muchos falsos negativos)). La forma de la curva entre estos dos puntos fijos depende de la capacidad discriminatoria de la prueba.

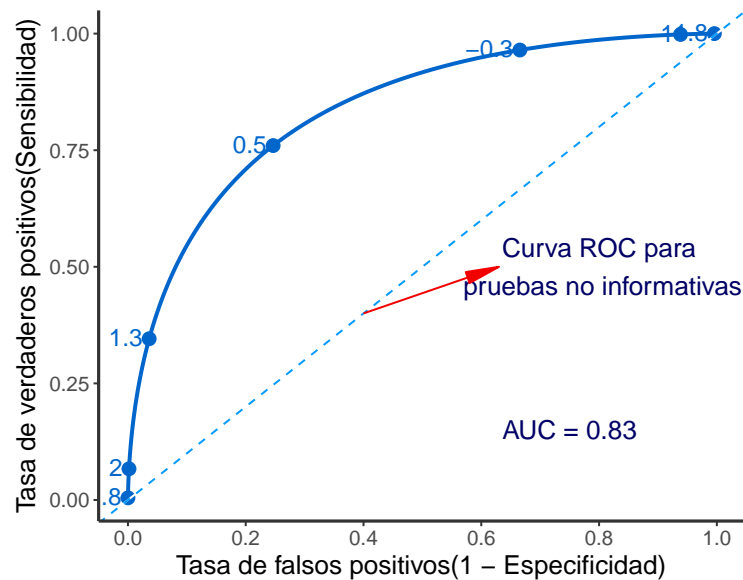


Figura 2.7: Umbrales explícitos para la curva ROC

La posición de la curva ROC depende del grado de superposición de las distribuciones de la prueba, en personas enfermas y no enfermas. Cuando una prueba discrimina claramente entre los enfermos y sanos de tal manera que no hay o hay poca superposición de

distribuciones, la curva ROC indicará que la alta sensibilidad se consigue con una alta especificidad, es decir, que la curva se aproxima a la parte superior izquierda en la que la sensibilidad es 1 y la especificidad es 1 (Figura 2.7). Si las distribuciones de los resultados de la prueba en los enfermos y en los sanos coinciden, la prueba sería completamente desinformativa y su curva ROC sería la diagonal ascendente del cuadrado (Figura 2.7).

La comparación de los ensayos sobre la base de sus curvas ROC tiene en cuenta su precisión a través de una serie de umbrales, y se ve favorecido por estadísticas de resumen únicas. Varias de estas medidas están en la literatura. El más utilizado entre ellas es el área bajo la curva (AUC), que es igual a 1 para una prueba perfecta y 0.5 para una prueba completamente desinformativa. Las AUC son igual a la probabilidad de que si un par de individuos enfermos y sanos son seleccionados al azar, el individuo enfermo tendrá un resultado de prueba más alto que el individuo sano. Las AUC puede interpretarse también como una sensibilidad media de la prueba, que abarca todos los valores de especificidad (o bien igual que la especificidad media sobre todos los valores de sensibilidad). Otros resúmenes incluyen áreas parciales bajo de la curva, los valores de sensibilidad correspondientes a los valores de especificidad seleccionados (y viceversa), y puntos de funcionamiento óptimos, definidos de acuerdo con criterios específicos.

### 2.3.2. Relaciones entre las curvas ROC, los odds-ratios de diagnóstico

Existe un vínculo útil entre las curvas ROC y los odds-ratios diagnósticos que es importante apreciar para entender la forma en que se construyen los modelos meta-analíticos. Para las curvas ROC simétricas mostrada en la figura 2.8, todos los puntos de cada curva tienen un odds ratio de diagnóstico común. Esta propiedad surge cuando los resultados de la prueba en los grupos enfermos y no enfermos tienen una distribución matemática particular conocida como distribución logística con igual varianza en ambos grupos. Por ejemplo, una curva ROC con un odds ratio diagnóstico de 21 pasaría por los puntos (sensibilidad, especificidad) de (0.70, 0.90), (0.82, 0.82) y (0.90, 0.70). Por lo tanto, una forma de resumir una curva ROC simétrica es por el valor del odds ratio de diagnóstico. Cuando las curvas ROC son asimétricas, la relación de probabilidad diagnóstica no es constante en toda la longitud de la curva, sino que aumenta (o disminuye) sistemáticamente.

te al aumentar el umbral, y la curva puede ser matemáticamente descrita al anotar cómo cambia el odds ratio de diagnóstico con el umbral, o un cantidad relacionada con el umbral.

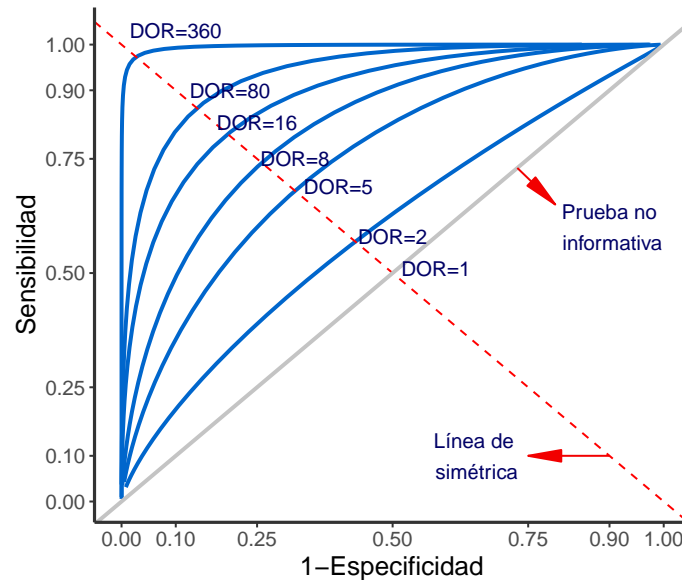


Figura 2.8: Relación entre las curvas ROC y DOR

### 2.3.3. Resumen de las curvas ROC

Las gráficas de resumen ROC muestran los resultados de los estudios individuales en el espacio ROC. Cada estudio se representa en forma de gráfica, punto de sensibilidad-especificidad único. El tamaño de los puntos puede ser controlado para representar la precisión de la estimación (típicamente escalada según el inverso del error estándar del  $\text{logit}(\text{sensibilidad})$  y  $\text{logit}(\text{especificidad})$ ) o según el tamaño de sus muestras.

Los gráficos de resumen ROC muestran la dispersión de los resultados del estudio. Ocasionalmente se añaden “ retículas ” a cada punto de estudio para indicar los límites de confianza para la sensibilidad y la especificidad, pero esto puede hacer que la trama quede muy desordenada si hay muchos estudios.

Se pueden añadir dos tipos de resumen meta-analítico al gráfico: curvas de resumen ROC (SROC) y resumen de los puntos de sensibilidad y especificidad. Regiones de confianza para la sensibilidad de resumen, pueden incluirse puntos de especificidad así como

regiones de predicción que dan una indicación de heterogeneidad entre los estudios.

### 2.3.4. Forest plots

Los forest plots para la precisión de las pruebas de diagnóstico informan el número de verdaderos positivos y falsos negativos en individuos enfermos, falsos positivos y verdaderos negativos en individuos no enfermos en cada estudio, y la sensibilidad y especificidad estimadas, junto con intervalos de confianza Figura 2.9. Los forest plot se conocen como forest plot acoplados, ya que contienen dos secciones gráficas: una que representa la sensibilidad, y otra que describe la especificidad. El criterio de ordenación de los estudios puede ser a menudo por valores de sensibilidad, o agrupados por tipo de prueba o valores de covariables. Si bien es posible observar la heterogeneidad en sensibilidad y especificidad individualmente en dichas gráficas, no es tan fácil visualizar si existen relaciones de tipo umbral. Las estadísticas de resumen calculadas a partir de meta-análisis rara vez se agregan a forest plot acopladas.

## 2.4. Resumen Meta-analítico

En una revisión sistemática es probable que los datos recopilados sean una mezcla de diferentes umbrales de positividad. Mientras que para algunas pruebas hay consenso sobre el valor que debería tener el umbral de positividad, más a menudo las pruebas se evalúan en diferentes umbrales y en diferentes estudios. También se encuentra la presentación de los resultados en múltiples umbrales dentro de un mismo estudio, con algunos estudios que presentan estimaciones de las curvas ROC que representan la precisión de la prueba en todos los umbrales posibles. Además, la notificación selectiva de los umbrales identificados para optimizar la precisión de la prueba puede introducir sesgos si se seleccionan en función de los datos.

### 2.4.1. Estimación de una curva SROC o un punto de consenso

Un principio clave que subyace a la elección del resumen estadístico en el meta-análisis de la precisión de la prueba es que la sensibilidad y la especificidad de una prueba variarán,

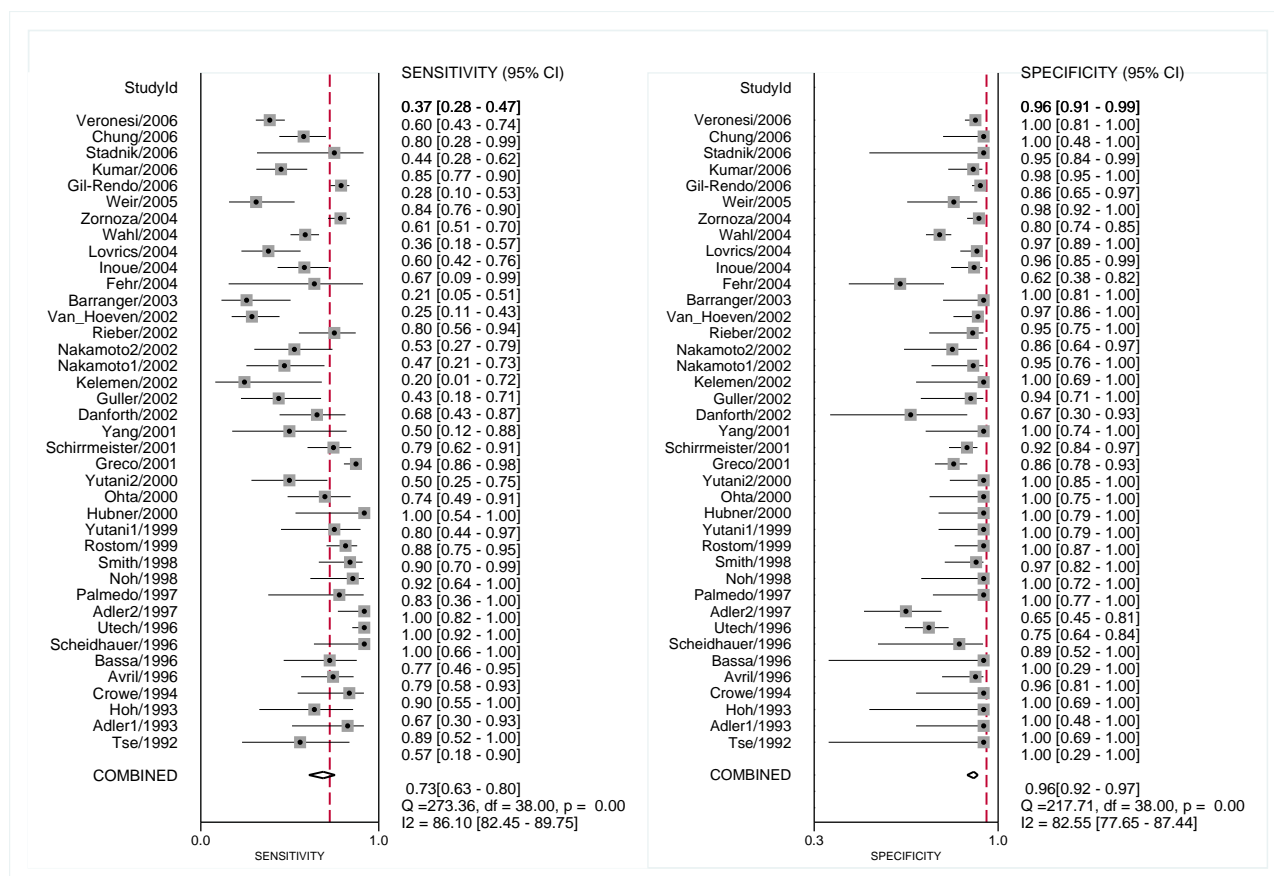


Figura 2.9: Gráfica de forest plot que muestra sensibilidad y especificidad específicas del estudio y una medida resumen de sensibilidad y especificidad con las correspondientes estadísticas de heterogeneidad.

como los umbrales de positividad varíen, esto se representa gráficamente utilizando una curva ROC.

Es importante señalar que los modelos jerárquicos recomendados para el metanálisis en DTA tienen en cuenta la correlación entre la sensibilidad y la especificidad observada en todos los estudios, lo que se debe a la relación funcional entre la sensibilidad y la especificidad, ya que el umbral varía dentro de cada estudio. Esto ocurre independientemente de si la salida elegida es una curva de resumen ROC o un punto de consenso. Un revisor debe decidir si utilizará todos los estudios disponibles para estimar la curva (en cuyo caso el metanálisis estimará la curva ROC resumida) o para estimar un punto de sensibilidad y especificidad resumida en esta curva en un umbral elegido. La estimación de la sensibilidad y especificidad resumida mediante la combinación de estudios que mezclan umbrales



producirá una estimación que se relaciona con algún promedio no especificado teórico de los umbrales que se dan en los estudios incluidos, lo cual es clínicamente poco útil y debe evitarse.

La variación en el umbral es muy probable cuando no hay un punto de corte numérico explícito y definiciones de un resultado positivo se basan en el juicio más que en la medición. Pero incluso cuando es posible definir un punto de corte común sobre la base de un valor numérico o de un punto en una escala de valoración, debe ser reconocido que seguirá existiendo cierta variabilidad en el umbral real entre los estudios a través de diferencias de calibración entre equipos, diferencias entre calificadores u observadores, como por ejemplo así como la variación en la implementación de las pruebas. La consecuencia de esta variabilidad será heterogeneidad adicional en los resultados de las pruebas observados en el punto de corte común. La sensibilidad sumaria y el punto de especificidad reflejará la precisión media observada, mientras que la región de predicción reflejará la heterogeneidad en la forma en que se aplica.

Por lo tanto, las dos estrategias principales para manejar umbrales mixtos y variables en un análisis son:

1. Estimación de la sensibilidad y especificidad resumida de la prueba para un umbral común, o en cada uno de varios umbrales comunes diferentes. Cada estudio puede contribuir a uno o más análisis dependiendo de los umbrales de los que se informe. Se excluyen los estudios que no presentan informes en ninguno de los umbrales seleccionados.
2. Estimación de la curva ROC subyacente que describe cómo la sensibilidad y la especificidad se compensan entre sí cuando los umbrales varían. En este caso, se selecciona un umbral por estudio para incluirlo en el análisis.

La elección del enfoque analítico dependerá de la variación de los umbrales en los estudios disponibles. Por ejemplo, si hay poca coherencia en los umbrales utilizados, los meta-análisis que se limitan a umbrales comunes contendrán muy pocos datos, y puede ser preferible estimar una curva ROC de resumen. Si hay poca variación en el umbral entre los estudios que intentan ajustarse a una curva ROC resumida, será difícil, ya que

es probable que los puntos estén demasiado agrupados en el espacio ROC.

Es razonable estimar tanto las curvas del SROC como los puntos medios de funcionamiento en una revisión, ya que pueden complementarse entre sí para proporcionar resúmenes clínicamente útiles y poderosas formas de detectar efectos. Por ejemplo, se pueden utilizar análisis separados de los datos de las pruebas en diferentes umbrales para proporcionar estimaciones clínicamente informativas de la sensibilidad y la especificidad, mientras que se incluyen todos los estudios para estimar cómo las curvas ROC de resumen dependen de las covariables o tipo de prueba será la manera más poderosa de probar e investigar la heterogeneidad.

### 2.4.2. Métodos meta-analíticos no utilizados de forma rutinaria

Los métodos que no se incluyen rutinariamente en las revisiones clásicas, se encuentran comúnmente en la literatura para el metanálisis diagnóstico. La combinación por separado de las estimaciones de sensibilidad y especificidad no tiene en cuenta el equilibrio entre sensibilidad y especificidad, lo que puede llevar a subestimar la precisión de la prueba ([103]). De manera similar, la agrupación separada de los cocientes de probabilidad ignora las correlaciones entre los cocientes de probabilidad positivos y negativos, y teóricamente puede producir estimaciones que son imposibles ([104]). La agrupación de los valores predictivos es posible utilizando el método bivariante, pero no se recomienda, ya que se sabe que los valores predictivos dependen de la prevalencia, que es probable que varíe de un estudio a otro. Las consecuencias de esto son dobles:

1. En primer lugar, que la variación de la prevalencia entre los estudios puede inducir una heterogeneidad mayor que la observada para la sensibilidad y la especificidad.
2. En segundo lugar, que los valores predictivos medios se relacionarán con el uso de la prueba con una prevalencia media, aunque desconocida.

## 2.5. Heterogeneidad

La heterogeneidad es de esperar en los meta-análisis de la precisión de las pruebas diagnósticas. Una consecuencia de esto es que los meta-análisis de los estudios de preci-

sión de las pruebas tienden a centrarse en el cálculo de la media más que en los efectos comunes. En las revisiones sistemáticas de las intervenciones, a veces se observa que las estimaciones del efecto de la intervención en los diferentes estudios son muy similares, siendo las diferencias entre ellas lo suficientemente pequeñas como para ser explicables por casualidad. En tales situaciones, es apropiado utilizar un meta-análisis de enfoque en efectos fijos que estima el efecto común subyacente (y se interpreta como el efecto real de la intervención).

En las revisiones de la precisión de las pruebas diagnósticas se observan con frecuencia grandes diferencias entre los estudios, demasiado grandes para ser explicadas por la casualidad, lo que indica que la precisión real de la prueba varía entre los estudios incluidos, o que hay heterogeneidad en la precisión de la prueba. Se recomiendan los métodos de meta-análisis de efectos aleatorios cuando los datos son heterogéneos, que se centran en proporcionar una estimación de la precisión media de la prueba y en describir la variabilidad de este efecto. En las revisiones DTA, se presume que existe heterogeneidad y los modelos de efectos aleatorios se ajustan por defecto, sólo se simplifican a modelos de efectos fijos cuando hay muy pocos estudios para estimar la variabilidad entre los estudios, o el análisis demuestra que los efectos fijos son apropiados.

Las pruebas univariadas para la heterogeneidad en la sensibilidad y especificidad y las estimaciones de la estadística  $I^2$  ([14]) no se utilizan sistemáticamente en las revisiones para DTA, ya que no tienen en cuenta la heterogeneidad explicada por fenómenos como los efectos del umbral de positividad. Si en un meta-análisis hay variación en el umbral, lo importante es el grado en que los resultados observados del estudio se encuentran cerca de la curva ROC resumida, no cuán dispersos están en el espacio ROC. La magnitud de la heterogeneidad observada se representa mejor gráficamente cuando tales relaciones pueden ser observadas por la dispersión de puntos y de la elipse de predicción. Las estimaciones numéricas de los términos de efectos aleatorios en los modelos jerárquicos cuantifican la cantidad de heterogeneidad observada, pero no se interpretan fácilmente ya que representan variaciones en los parámetros expresados en las probabilidades de registro.

## 2.6. Ajuste del modelo

### 2.6.1. Curvas SROC de Moses Littenberg

El método Moses-Littenberg ([97]) ([106]) proporciona un modelo sencillo para derivar un SROC. Fue uno de los primeros modelos propuestos y se ha utilizado ampliamente en meta-análisis de la precisión de las pruebas de diagnóstico. Es más parecido a un modelo de efecto fijo que de efectos aleatorios, ya que no proporciona estimaciones de la heterogeneidad entre los estudios. A pesar de que ha sido reemplazado por modelos jerárquicos más complejos que permiten modelizar los efectos aleatorios, el modelo Moses-Littenberg se utiliza para proporcionar a los revisores la posibilidad de realizar análisis puramente exploratorios basados en curvas SROC sin necesidad de exportar datos a otros software especializados. Debido a las limitaciones del método Moses-Littenberg, algunos programas no proporcionan estimaciones de parámetros o errores estándar de este modelo, ya que las inferencias deben ser basadas en modelos jerárquicos que tienen en cuenta por separado el error de muestreo dentro del estudio y heterogeneidad adicional inexplicable entre los estudios.

A continuación se ofrece una breve descripción del método Moses-Littenberg para explicar cómo se derivan las curvas SROC. El método se desarrolla en tres pasos:

- i) Los pares de estimaciones de sensibilidad y especificidad de cada estudio se transforman en el logaritmo probabilidades (logit) para calcular,

$$\begin{aligned} D &= \text{logit}(\text{sensibilidad}) - \text{logit}(1 - \text{especificidad}) \\ S &= \text{logit}(\text{sensibilidad}) + \text{logit}(1 - \text{especificidad}) \end{aligned} \tag{2.6.4}$$

donde  $D$  es el logaritmo natural del odds ratio diagnóstico ( $\ln\text{DOR}$ ) y  $S$  es una cantidad relacionada con la proporción total de resultados positivos de la prueba.  $S$  puede considerarse un indicador aproximado del umbral de la prueba, ya que  $S$  aumentará a medida que aumente la proporción global de positivos en las pruebas, tanto en los grupos de enfermos como en los de no enfermos. Se espera que la relación entre  $D$  y  $S$  sea lineal.

- ii) El modelo de regresión lineal simple  $D = \alpha + \beta S + \text{error}$  caracteriza la forma en que la precisión de la prueba, medida por el odds ratio de registro de diagnóstico ( $D$ ), varía con  $S$ , un indicador del umbral de positividad en todos los estudios.

- iii) Las estimaciones de alfa y beta se utilizan para obtener la sensibilidad estimada a través de un rango escogido de posibles valores de especificidad usando

$$E(\text{sensibilidad}) = 1/[1 + \exp(-[\alpha + (1 + \beta)\text{logit}(1 - \text{especificidad})]/(1 - \beta))] \quad (2.6.5)$$

Esto proporcionará la curva SROC estimada en las coordenadas ROC originales. El rango de especificidades sobre las que se dibuja la curva suele limitarse al rango observado en los datos para evitar la extrapolación.

### 2.6.2. Propiedades de la Curva SROC de Moses Littenberg

La Figura 2.10 ilustra tres posibles curvas SROC que podrían surgir del modelo Moses-Littenberg. Todos comparten el mismo valor de  $\alpha$  (tomado como 3 para cada curva), pero con variaciones  $\beta$  (tomado como -0.35, 0 y 0.35). El punto de intersección de las tres curvas se encuentra en la diagonal donde sensibilidad=especificidad ( $S=0$ ). La sensibilidad y especificidad de la prueba en este punto también se refiere como  $Q^*$ . Cuando  $\beta = 0$ , la curva es simétrica respecto a la línea diagonal dada por  $S=0$ . La función  $\ln DOR$  es el mismo (e igual alfa) en cada punto de esta curva simétrica ya que no hay asociación entre la precisión ( $D$ ) y el umbral, ( $S$ ) en el modelo. Sin embargo, cuando  $\beta \neq 0$ , la curva no es simétrica y la precisión esperada ( $\ln DOR$ ) aumenta (o disminuye) con el umbral.

Es posible que en algunos conjuntos de datos el valor estimado de  $\beta$  puede llevar a curvas SROC incorrectas que pasen por las esquinas inferior izquierda (sensibilidad=0, especificidad=1) y superior derecha (sensibilidad=1, especificidad=0) del gráfico SROC. Si  $\beta \geq 1$  o  $\beta \leq -1$  la curva SROC estimada tiene la propiedad no intuitiva de que la sensibilidad disminuye a medida que aumenta 1-especificidad (tasa de falsos positivos). Tales situaciones pueden surgir si existen estudios externos que influyen en la determinación de la pendiente de la línea de regresión.

La exclusión del estudio de los valores atípicos permite evaluar su influencia en la curva SROC ajustada. Los valores extremos de  $\beta$  también pueden resultar si hay heterogeneidad en la precisión de la prueba entre los subgrupos de estudios. Esta heterogeneidad puede explorarse a través de análisis de subgrupos cuando hay suficientes estudios que lo permitan.

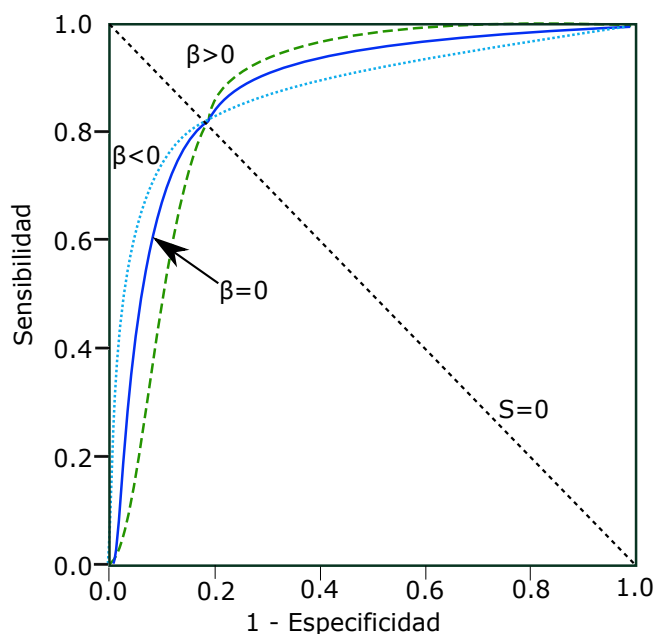


Figura 2.10: Curvas SROC para valores alternativos de los parámetros del modelo

### 2.6.3. Elección de los pesos

La línea de regresión puede ajustarse utilizando el método de mínimos cuadrados ponderados (WLS) para tener en cuenta diferencias en el error de muestreo en  $D$  entre los estudios al ponderar cada estudio por el inverso de la varianza del  $\ln DOR$  para ese estudio (estimada como  $\text{var}(\ln DOR) = 1/a + 1/b + 1/c + 1/d$ , donde  $a$ ,  $b$ ,  $c$  y  $d$  representan las celdas de la tabla  $2 \times 2$  que se muestra en la Tabla 2.1). Un enfoque alternativo es asignar el mismo peso a todos los estudios sobre la base de que es probable que la heterogeneidad inexplicable en la precisión de la prueba entre los estudios sea grande en comparación con la variabilidad debida al error de muestreo ([97]) ([98]). Tanto los mínimos cuadrados ponderados como los no ponderados (igualmente ponderados) se implementan en software estadísticos. En la práctica, ambos esquemas de ponderación conducen a menudo a curvas similares.

Ninguno de los dos enfoques aborda el tema del error de muestreo en la variable explicativa ( $S$ ) (violando una suposición básica de regresión lineal) y no se ocupa adecuadamente de la heterogeneidad adicional no explicada en  $D$ . Por consiguiente, el método Moses-Littenberg para el análisis SROC descrito anteriormente se utiliza sólo para los análisis

exploratorios preliminares y no debe utilizarse para calcular los intervalos de confianza para las estimaciones resumidas de la exactitud de las pruebas, ni para establecer si las diferencias entre subgrupos se encuentran dentro de los límites de lo que se espera que se pueda ver por casualidad únicamente.

## 2.7. Modelos jerárquicos

Se han propuesto enfoques más rigurosos desde el punto de vista estadístico, basados en modelos jerárquicos que superan las limitaciones del método Moses-Littenberg. En esta sección se describen y discuten el modelo bivalente ([100]) y el modelo jerárquico SROC (HSROC) de Rutter y Gatsonis ([101]). Ambos modelos jerárquicos implican distribuciones estadísticas a dos niveles.

En el nivel inferior, modelan las frecuencias absolutas de las celdas en las tablas 2x2 extraídas de cada estudio usando distribuciones binomiales y transformaciones logísticas (*log – odds*) de proporciones.

En el nivel superior, se supone que los efectos aleatorios de los estudios tienen en cuenta la heterogeneidad en la precisión de las pruebas de diagnóstico entre los estudios, más allá de lo que representa la variabilidad del muestreo en el nivel inferior. El modelo Bivalente y el modelo HSROC de Rutter y Gatsonis son matemáticamente equivalentes cuando no hay covariables ([107]), ([108]), pero difieren en sus parametrizaciones. La parametrización Bivalente modela la sensibilidad, la especificidad y la correlación entre ellas directamente, mientras que la parametrización HSROC de Rutter y Gatsonis modela las funciones de sensibilidad y especificidad para definir una curva ROC resumida (SROC).

Las estimaciones de parámetros del modelo Bivalente o del modelo HSROC de Rutter y Gatsonis se pueden introducir en software estadísticos para producir:

- i) El resumen de la curva ROC.
- ii) El resumen del punto de operación (es decir, los valores resumidos de sensibilidad especificidad).
- iii) Una región de confianza del 95 % alrededor del punto de operación resumido.
- iv) Una región de predicción del 95 %.

Esta región de predicción es una forma de ilustrar el alcance de la heterogeneidad estadística al representar una región dentro de la cual, suponiendo que el modelo sea correcto, tenemos una confianza del 95 % de que la verdadera sensibilidad y especificidad de un estudio futuro debería estar presente ([107]).

A partir de la curva ROC resumida, se puede calcular la sensibilidad esperada a un valor dado de especificidad (o viceversa). Además, también se pueden derivar valores de resumen e intervalos de confianza para los cocientes de probabilidad positivos y negativos o para los odds ratio de diagnóstico en el punto de resumen.

No todas estas posibles medidas resumidas serán pertinentes o apropiadas para un análisis determinado. La elección de la(s) medida(s) resumida(s) debe basarse en la pregunta de la investigación y también en la variabilidad de los umbrales utilizados en los estudios para definir la positividad de la prueba.

La motivación para elegir uno de estos dos modelos jerárquicos alternativos queda clara cuando se añaden covariables para explorar la heterogeneidad en la precisión de la prueba. En última instancia, la elección del método vendrá determinada por el enfoque que se desee adoptar y cuál de los dos aborda directamente la cuestión de la investigación.

Ambos modelos requieren el uso de software estadístico externo, ya que su adaptación requiere métodos demasiado complejos. Sin embargo, la salida gráfica de las estimaciones de parámetros de cualquiera de los dos modelos, añada resúmenes de modelos a gráficas como curvas SROC.

### 2.7.1. El modelo Bivariante

El método Bivariante modela directamente la sensibilidad y la especificidad. Se puede considerar que el modelo tiene dos niveles que corresponden a la variación dentro de los estudios y entre ellos. En el primer nivel, se supone que la variabilidad dentro del estudio, tanto para la sensibilidad como para la especificidad, sigue una distribución binomial.

Se denota  $A$  y  $B$  para la sensibilidad y especificidad respectivamente, el número de verdaderos positivos por  $y_{Ai}$  en el estudio  $i$ , donde  $n_{Ai}$  denotan el número total de indivi-



duos enfermos y  $\pi_{Ai}$  representa la probabilidad de que la prueba dé un resultado positivo en el estudio  $i$ . Similarmente,  $y_{Bi}$  el número de verdaderos negativos, donde  $n_{Bi}$  denota el total de individuos sanos y  $\pi_{Bi}$  es la probabilidad de que la prueba dé un resultado negativo en el estudio  $i$ , es decir:

$$y_{Ai} \sim \text{Binomial}(n_{Ai}, \pi_{Ai}), \quad y_{Bi} \sim \text{Binomial}(n_{Bi}, \pi_{Bi})$$

El par sensibilidad-especificidad para cada estudio debe modelarse conjuntamente dentro del estudio en el nivel inferior del análisis, ya que están vinculados por características compartidas del estudio, incluido el umbral de positividad.

En el nivel superior, se supone que las transformaciones logit para la sensibilidades tienen una distribución normal con la media  $\mu_A$  y varianza  $\sigma_A^2$ , mientras que las transformaciones logit para la especificidad tienen una distribución normal con la media  $\mu_B$  y varianza  $\sigma_B^2$ . Su correlación se incluye mediante la modelización de ambas a la vez mediante una única distribución normal bivalente:

$$\begin{pmatrix} \mu_{Ai} \\ \mu_{Bi} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \Sigma \right), \quad \Sigma = \begin{pmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{pmatrix} \quad (2.7.6)$$

donde  $\sigma_A^2$  y  $\sigma_B^2$  denotan la variabilidad entre las transformaciones logit de sensibilidad y especificidad de logit, respectivamente, y  $\sigma_{AB}$  denota covarianza entre las transformaciones logit de sensibilidad y especificidad. El modelo también puede ser parametrizado usando la correlación  $\rho_{AB} = \sigma_{AB} / (\sigma_A \sigma_B)$  que puede ser más interpretable que la covarianza. Por lo tanto, el modelo bivalente tiene cinco parámetros cuando no se incluyen covariables:  $\mu_A$ ,  $\mu_B$ ,  $\sigma_A^2$ ,  $\sigma_B^2$  and  $\sigma_{AB}$  (o  $\rho_{AB}$ ).

La inclusión de un parámetro de correlación en el modelo permite la compensación esperada en sensibilidad y especificidad, ya que el umbral de positividad de la prueba varía entre los estudios. Cuando la variación entre los estudios surge a través de tal compensación, se espera que esta correlación sea negativa, pero la correlación puede ser positiva si hay otras fuentes de heterogeneidad.

Reitsma ([100]) propuso originalmente ajustar estos modelos mediante la aproximación de las distribuciones binomiales dentro del estudio por distribuciones normales. Aunque

esto permite que el modelo se adapte a una gama ligeramente mayor de software (por ejemplo, el procedimiento MIXED en SAS), Chu ([115]) demostró posteriormente que la aproximación puede funcionar mal y recomendó que se utilice software que pueda modelar explícitamente las distribuciones binomiales dentro del estudio.

### 2.7.2. El modelo HSROC de Rutter y Gatsonis

El modelo HSROC propuesto por Rutter y Gatsonis ([109]), ([101]) se basa en un modelo de regresión logística a escala latente ([110]), ([111]). El modelo HSROC asume que hay una curva ROC subyacente en cada estudio con parámetros  $\alpha$  y  $\beta$  que caracterizan la exactitud y asimetría de la curva, de forma similar (aunque técnicamente distinta) a los parámetros  $\alpha$  y  $\beta$  del método de regresión lineal de Moses y Littenberg. A diferencia del modelo Moses-Littenberg, el modelo Rutter y Gatsonis está limitado para proporcionar una curva ROC donde la sensibilidad no puede disminuir a medida que aumenta la especificidad.

La precisión, definida en términos de ln DOR, determina la posición de la curva resumen con respecto a la esquina superior izquierda de los ejes ROC. Al igual que con el método de regresión SROC, cada estudio aporta al análisis datos en un solo umbral. La tabla 2x2 para cada estudio surge entonces de la dicotomía en un umbral de positividad denotado por  $\theta$ . Los parámetros  $\alpha$  y  $\theta$  se supone que varían entre los estudios: se supone que ambos tienen distribuciones normales como en el meta-análisis convencional de efectos aleatorios.

También puede considerarse que el modelo HSROC tiene dos niveles que corresponden a la variación dentro de los estudios y entre ellos. En el primer nivel, el número de personas enfermas que dan positivo es denotado por  $y_{i1}$  para el  $i$ -ésimo estudio, y el número correspondiente de personas no enfermas que dan positivo se denota por  $y_{i2}$ . Para cada estudio ( $i$ ), el número de pruebas positivas en cada grupo de enfermos ( $j$ ) se supone que sigue una distribución binomial tal que  $y_{ij} \sim \text{Binomial}(n_{ij}, \pi_{ij})$ ,  $j=1,2$  donde  $n_{ij}$  y  $\pi_{ij}$  respectivamente, representan el número total de pruebas y la probabilidad de un resultado positivo. El número de pruebas positivas en cada par enfermo y no enfermo se analiza conjuntamente dentro de cada estudio en el nivel uno del análisis. El modelo toma

la forma de:

$$\text{logit}(\pi_{ij}) = (\theta_i + \alpha_i \text{dis}_{ij}) \exp(-\beta \text{dis}_{ij}) \quad (2.7.7)$$

Donde  $\text{dis}_{ij}$  representa la situación “ verdadera ” de la enfermedad (codificada como -0.5 para los no enfermos y 0.5 para los enfermos), teniendo en cuenta así la variabilidad dentro del estudio en el nivel uno.

Usando la terminología habitual para este modelo, generalmente nos referimos a  $\theta_i$  como el umbral de positividad calculado como la media log odds del resultado de prueba positiva para un enfermo y los log odds del resultado de prueba positivo para el grupo no enfermo en el estudio  $i$  (equivalente a  $S_i/2$  en el modelo de Moses-Littenberg).  $\alpha_i$  (ln DOR para el estudio  $i$ ) representa una medida de la precisión diagnóstica en el  $i$ -ésimo estudio que incorpora tanto la sensibilidad como la especificidad para ese estudio. El parámetro de escala ( $\beta$ ) proporciona asimetría en el SROC al permitir que la precisión varíe con el umbral. Desde que cada estudio aporta sólo una estimación de sensibilidad y especificidad en un único umbral, es necesario suponer que la forma de la curva ROC subyacente real en cada estudio es la misma y, por lo tanto,  $\beta$  se ajusta como un efecto fijo.

El umbral y la precisión diagnóstica para cada estudio se especifican como efectos aleatorios y se supone que son independientes (no correlacionados) y normalmente distribuidos. El parámetro de precisión tiene media  $\Lambda$  (lambda mayúscula) y varianza  $\sigma_\alpha^2$ , mientras que el parámetro de positividad (umbral) tiene media  $\Theta$  (teta mayúscula) y varianza  $\sigma_\theta^2$ . El parámetro de forma ( $\beta$ ) se estima utilizando datos de los estudios considerados conjuntamente, suponiendo efectos aleatorios normalmente distribuidos para la precisión de la prueba. Cuando no se incluyen covariables, el modelo HRSOC también tiene cinco parámetros:  $\Lambda$ ,  $\Theta$ ,  $\beta$ ,  $\sigma_\alpha^2$  y  $\sigma_\theta^2$ .

Se puede construir una curva ROC resumida a partir del modelo HSROC seleccionando un rango de valores de 1-especificidad y utilizando el estimado del parámetro de ubicación medio ( $\Lambda$ ) y el parámetro de escala ( $\beta$ ) para calcular los valores correspondientes de sensibilidad. La sensibilidad esperada para una fracción de falso positivo seleccionada (1-especificidad) viene dada por:

$$\text{sensibilidad} = 1/[1 + \exp(-(\Lambda e^{-0,5\beta + \text{logit}(1-\text{especificidad})e^{-\beta}}))]. \quad (2.7.8)$$

Cuando  $\beta = 0$ , la precisión de la prueba puede resumirse por  $\Lambda$  que representa la precisión esperada (log DOR), y la curva de resumen resultante será simétrica.

### 2.7.3. Inferencia Bayesiana

Hay al menos dos razones por las que las estadísticas bayesianas son particularmente importantes hoy en día. Una de ellas es la capacidad de combinar múltiples fuentes de información en un mismo sistema (síntesis). Los modelos bayesianos son especialmente adecuados para esta tarea, que cuenta con una gran impacto en el meta-análisis moderno, los modelos multinivel y el modelado jerárquico en general. La otra razón es la revolución computacional producida por el redescubrimiento de Cadenas de Markov y Monte Carlo (MCMC), técnicas en estadística. Como resultado, podemos construir modelos estadísticos complejos arbitrarios que reflejan la complejidad de los fenómenos de interés.

#### Teorema de Bayes e inferencia estadística

Supongamos que  $y^T = (y_1, y_2, \dots, y_n)$  es un vector de  $n$  observaciones cuya distribución de probabilidad  $p(y|\theta)$  depende de los valores de  $k$  cantidades desconocidas  $\theta^T = (\theta_1, \dots, \theta_k)$ . En la estadística clásica  $\theta$  es una cantidad fija desconocida, en la estadística bayesiana es una cantidad incierta. Esta incertidumbre sobre  $\theta$  se modela con una distribución de probabilidad  $p(\theta)$ . Entonces

$$p(y|\theta)p(\theta) = p(y, \theta) = p(\theta|y)p(y) \quad (2.7.9)$$

Dados los datos observados, el teorema de Bayes dice que la distribución condicional de  $\theta$  es

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (2.7.10)$$

que se deriva claramente del punto (2.7.9). Ahora  $p(y) = c^{-1}$  es sólo una constante normalizadora para asegurar que la distribución  $p(\theta|y)$  se integra. Entonces, el teorema de Bayes a veces se afirma como

$$p(\theta|y) = cp(y|\theta)p(\theta) \quad (2.7.11)$$

o en su forma corta

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (2.7.12)$$

donde  $\propto$  denota proporcional a. La expresión (2.7.12) suele denominarse distribución posterior no normalizada.

La distribución de probabilidad  $p(\theta)$ , nos dice lo que se sabe sobre  $\theta$  de forma independiente de los datos y, se llama la distribución previa de  $\theta$ , o la distribución de  $\theta$  a priori. La distribución de probabilidad  $p(\theta|y)$ , nos dice lo que se sabe sobre  $\theta$  dado el conocimiento de los datos, se llama la distribución posterior de  $\theta$  dada, o la distribución de  $\theta$  a posteriori. También se puede referir como la distribución previa y la distribución posterior simplemente como “anterior” y “posterior”, respectivamente.

Dados los datos  $y$ , la distribución de probabilidad  $p(y|\theta)$  puede ser considerada no como una función de  $y$  sino de  $\theta$ . De esta manera la llamamos la función de probabilidad de  $\theta$ . Podemos ver claramente que el teorema de Bayes nos dice que el posterior de  $\theta$  es proporcional al producto del anterior y a la probabilidad, es decir,

distribución posterior  $\propto$  verosimilitud  $\times$  distribución anterior

Una vez que el paradigma Bayesiano es aceptado para inferencia, se usa la distribución posterior  $p(\theta|y)$  para inferir sobre  $\theta$  y no se requieren otros conceptos. Por ejemplo, no juegan ningún papel en las estadísticas bayesianas conceptos clásicos como muestreo repetido, sesgo, consistencia, suficiencia, etc. La mayoría de las dificultades se centran en una especificación adecuada de la distribución anterior  $p(\theta)$  y el cálculo de  $p(\theta|y)$ , que podrían ser extremadamente difícil en la práctica. El uso revolucionario de las técnicas de MCMC y su facilidad de uso implementación en varios softwares estadísticos (por ejemplo, WinBUGS, R, SAS, MLWin, etc.) ha popularizado el uso de técnicas bayesianas, pero la especificación de la distribución previa, particularmente en situaciones multivariantes, sigue siendo una parte difícil de este enfoque.

### Naturaleza secuencial del teorema de Bayes

Otro resultado importante de las estadísticas bayesianas es la naturaleza secuencial del Teorema de Bayes. Supongamos que tenemos una muestra inicial de observaciones  $y_1$ , entonces la fórmula de Bayes da

$$p(\theta|y_1) \propto p(y_1|\theta)p(\theta) \tag{2.7.13}$$

Ahora, supongamos que tenemos una segunda muestra de observaciones  $y_2$  distribuida independientemente de la primera muestra, entonces,

$$p(\theta|y_2, y_1) \propto p(y_2|\theta)p(y_1|\theta)p(\theta) \quad (2.7.14)$$

$$\propto p(y_2|\theta)p(\theta|y_1) \quad (2.7.15)$$

La última expresión muestra que la distribución posterior  $p(\theta|y_1)$  juega el papel de una distribución previa de la segunda muestra  $y_2$ . Obviamente, este proceso puede repetirse cualquier número de veces. Por lo tanto, el teorema de Bayes describe el proceso de aprendizaje a partir de la experiencia, y muestra cómo el conocimiento sobre  $\theta$  se modifica continuamente a medida que se dispone de nuevos datos.

### Cantidades desconocidas, predicciones y comprobación de modelos

En la inferencia bayesiana la cantidad desconocida  $\theta$  es genérica, puede denotar un vector de parámetros, datos faltantes, valores predictivos, datos mal medidos, etc. Este enfoque ofrece una gran flexibilidad a la hora de construir modelos estadísticos que reflejen mejor las complejidades encontradas en la práctica. Por ejemplo, si estamos interesados en hacer predicciones, podemos escribir  $\theta = (y^{Pred}, \eta)$ , después de observar  $y$  y tenemos

$$\begin{aligned} p(y^{Pred}|y) &= \int p(y^{Pred}, \eta|y) d\eta \\ &= \int p(y^{Pred}|\eta, y) p(\eta|y) d\eta \\ &= \int p(y^{Pred}|\eta) p(\eta|y) d\eta \end{aligned}$$

La segunda línea muestra claramente que la distribución predictiva posterior  $p(y^{Pred}|y)$  es un promedio de las predicciones condicionales,  $p(y^{Pred}|\eta)$  ponderadas por la distribución posterior de  $\eta$ . Por supuesto, la integral se calcula sobre el espacio de parámetros de  $\eta$  y se sustituye por sumas si el espacio de parámetros es discreto.

Las predicciones posteriores juegan un papel muy importante en la comprobación de los supuestos del modelo. El punto básico aquí es que, como de costumbre en la práctica, no podemos garantizar que el modelo sea correcto. Así que en el modelado bayesiano analizamos el modelo ajustado mediante la comprobación posterior del modelo predictivo, que consiste en simular datos del modelo ajustado y comparando estas cantidades con

las observadas. Este nos permite entender los déficits del modelo ajustado y corregirlos mediante un modelo actualizado.

Matemáticamente, la fórmula de Bayes es una declaración válida de probabilidad condicional. Lo que ha estado en debate durante mucho tiempo es su aplicabilidad para la inferencia estadística general. Es el uso de probabilidades para la inferencia inductiva. Las dificultades se centran en el significado de probabilidad y en las estadísticas bayesianas la elección y necesidad de las distribuciones previas.

### Intercambiabilidad

Como antes supongamos que  $y^T = (y_1, y_2, \dots, y_n)$  es un vector de  $n$  observaciones cuya distribución de probabilidad es  $p(y|\theta)$ . La intercambiabilidad es una expresión formal de la idea de que no encontramos ninguna razón sistemática para distinguir entre las variables individuales  $Y_1, \dots, Y_n$  que han producido el resultado  $y$ . Decimos que estas variables son similares pero no idénticas.

La intercambiabilidad representa un componente de nuestro conocimiento de los datos. Por ejemplo, si sabemos que un conjunto de datos resulta de una mezcla de estudios con un diseño diferente, no sería razonable considerar estos datos como intercambiables. Aunque, los resultados numéricos pueden ser similares entre los estudios, la calidad de los datos es diferente si un estudio se realizó con un diseño retrospectivo o prospectivo. Técnicamente la intercambiabilidad implicaba la invariabilidad de  $p(y|\theta)$  bajo la permutación de los índices. Tenga en cuenta que no queremos decir que  $Y_1, \dots, Y_n$  son independientes e idénticos distribuido (i.i.d) con una función de distribución  $p(y|\theta)$ , esto significaría que  $p(y|\theta) = p(y_1|\theta) \dots p(y_n|\theta)$

### A priori

¿Quién puede negar que existen creencias previas cuantificables en medicina y otras áreas científicas? Cualquier persona que haya estado involucrada en la planificación de un ensayo clínico o de otro tipo, sabe que hay un cuerpo claro de información antes de que se realice un experimento. Para ser más precisos, en casi todos los análisis de datos hay cantidades sustanciales de información, que son externas a los datos que deseamos analizar. El enfoque bayesiano formaliza un procedimiento para combinar diferentes fuen-

tes de información. Sin embargo, la cuestión principal es cómo traducir este conjunto de conocimientos en una función de distribución formal. Dado que el análisis de los datos bayesianos es impulsado por los antecedentes y vale la pena señalar algunos malentendidos con respecto a distribuciones anteriores, a continuación se mencionan algunos puntos.

- El nombre anterior sugiere una relación temporal, sin embargo, esto es engañoso. La distribución anterior modela la incertidumbre dada por la evidencia externa.
- A priori no es necesariamente único, en un artículo reciente, [112] analizan el uso de 13 prioridades diferentes para el parámetro de varianza entre estudios en el metaanálisis de efectos aleatorios. No existe tal cosa como el “correcto” previo. Se considera que el análisis bayesiano transforma la opinión previa en posterior, en lugar de producir “la” distribución posterior.
- Apriori puede ser sobreparametrizado. A veces intencionalmente sobreparametrizamos los antecedentes para acelerar la convergencia de los métodos de simulación, ver [113] y [114].
- La inferencia puede basarse sólo en las distribuciones anteriores. Hay situaciones en las que no hay datos adicionales disponibles para combinar con nuestros anteriores o no hay intención de actualizar los antecedentes. Este es el caso típico del análisis de riesgo, determinación del tamaño de la muestra en experimentos, simulación de procesos complejos, etc.

### Análisis moderno de datos bayesianos

La construcción del modelo estadístico es la combinación coherente de técnicas estadísticas y probabilísticas:

- El uso de modelos jerárquicos para reflejar múltiples fuentes de incertidumbre en el análisis de datos.
- El uso extensivo del método MCMC como un enfoque basado en la simulación para calcular las distribuciones posteriores marginales de las cantidades de interés.
- El uso sistemático de modelos gráficos para dar una descripción esquemática de las cantidades del modelo y su interrelación. Esta técnica se suele combinar con algoritmos algebraicos automáticos para factorizar distribuciones de articulaciones com-



plejas en sus distribuciones marginales condicionales. Esta factorización automática simplifica el uso de técnicas de MCMC.

#### 2.7.4. Modelo Bayesiano para combinar datos de pruebas de diagnóstico

En esta sección se presenta el modelo bayesiano jerárquico para combinar los datos de las pruebas de diagnóstico. La construcción del modelo comienza especificando un modelo de datos,  $p(y|\theta)$ , donde  $y = (y_1, \dots, y_N)$  denota los datos disponibles y  $\theta = (\theta_1, \dots, \theta_N)$  todas las cantidades desconocidas específicas del estudio (por ejemplo, efectos aleatorios). El conjunto de parámetros específicos del estudio  $\theta_i$  está modelado por la distribución estructural  $p(\theta_i|\phi)$ . La incertidumbre acerca de  $\phi$  está modelada por una distribución previa  $p(\phi)$ . Las afirmaciones inferenciales sobre  $\theta$  y  $\phi$  se basan en las afirmaciones posteriores.

$$p(\theta, \phi|y) \propto p(\phi)p(\theta|\phi)p(y|\theta) \quad (2.7.16)$$

La comprobación del modelo se realiza simulando cantidades desconocidas de (2.7.16), estas simulaciones se comparan con los valores observados mediante inspección visual y más formalmente calculando las cantidades de desviación del modelo.

#### Un Modelo Bayesiano

Sean  $tp_i$  y  $fp_i$  los resultados de verdaderos positivos y falsos positivos para el estudio  $i$  ( $i = 1, \dots, N$ ). Acondicionando en  $n_{i,1}$ , que es el número de pacientes con enfermedad, y en  $n_{i,2}$ , el número de pacientes sin enfermedad, nuestro modelo de datos es una distribución binomial con

$$tp_i \sim Bin(TPR_i, n_{i,1}), fp_i \sim Bin(FPR_i, n_{i,2}) \quad (2.7.17)$$

donde  $TPR_i$  y  $FPR_i$  son las probabilidades de observar un resultado positivo en la prueba en la población de enfermos y no enfermos, respectivamente.

Los  $N$  pares de probabilidades  $TPR_i$  y  $FPR_i$  se transforman mediante una función de enlace  $g$  a una escala en la que se definen en el rango  $(-\infty, \infty)$ . La función de enlace canónico para datos binomiales es la función de enlace logística, pero hay otros enlaces

alternativos, por ejemplo, la función complementaria de enlace log-log (clog). Elegir un enlace adecuado para los datos es importante.

### Distribución estructural

Se modela la variabilidad entre los estudios con una distribución normal bivariante sobre la base de las diferencias

$$D_i = g(TPR_i) - g(FPR_i)$$

y las sumas con

$$S_i = g(TPR_i) + g(FPR_i)$$

con

$$(D_i, S_i) \sim Normal_2(\mu, \Lambda), i = 1, 2, \dots, N. \quad (2.7.18)$$

Donde  $\mu$  es el valor medio de  $(D_i, S_i)$  y  $\Lambda$  su matriz de precisión, es decir,  $\Lambda = \Sigma^{-1}$  con  $\Sigma$  la matriz de covarianza de varianza. Esta parametrización es conveniente en una configuración Bayesiana. Claramente, cuando  $g(\cdot)$  es la función de enlace logístico, los efectos aleatorios  $D_i$  y  $S_i$  se pueden relacionar con la curva SROC clásica. Por lo tanto,  $D_i$  es el efecto del estudio relacionado con la potencia discriminatoria diagnóstica y  $S_i$  es el efecto asociado con el valor del umbral de diagnóstico.

El modelado  $(D_i, S_i)$  es similar al modelado directo  $(g(TPR_i), g(FPR_i))$ , sin embargo, la transformación lineal debe dejar  $(D_i, S_i)$  aproximadamente independiente haciendo nuestra inferencia menos sensible a la distribución previa de  $\Lambda$ .

### Distribuciones Apriori

Para el enfoque de modelado bayesiano, necesitamos especificar distribuciones previas en los parámetros del modelo. Asumimos distribuciones normales previas independientes para los componentes de  $\mu = (\mu_D, \mu_S)^T$

$$\mu_D \sim N(m_D, v_D), \mu_S \sim N(m_S, v_S)$$

Para la matriz de precisión  $\Lambda$  damos una distribución previa de Wishart (es una generalización multivariante de la distribución gamma) con la matriz de escala  $R$  y  $k$  grados de libertad:

$$\Lambda \sim Wishart(R, k)$$

### Distribución Aposteriori

Dada esta especificación del modelo, toda la inferencia se basa en la distribución a posterior.

$$\begin{aligned} p(\theta, \phi|y) &\propto \prod_{i=1}^N \left[ \binom{n_{i,1}}{tp_i} TPR_i^{tp_i} (1 - TPR_i)^{(n_{i,1}-tp_i)} \binom{n_{i,2}}{fp_i} FPR_i^{fp_i} (1 - FPR_i)^{(n_{i,1}-fp_i)} \right] \\ &\times \prod_{i=1}^N \exp[-1/2(D_i - \mu_D, S_i - \mu_S)^T \Lambda (D_i - \mu_D, S_i - \mu_S)] \times |\Lambda|^{\frac{N}{2}} \\ &\times [-1/2(v_D(\mu_D - m_D)^2 + v_S(\mu_S - m_S)^2)] \\ &\times \frac{|\Lambda|^{(N-K-2)/2} \exp[-1/2 \text{trace}(\Lambda R^{-1})]}{2^{k(N-1)/2} \pi^{k(k-1)/4} |\Lambda|^{(N-1)/2} \prod_{i=1}^N \Gamma(1/2(N-i))} \end{aligned}$$

### Cálculos de MCMC

BUGS significa análisis bayesiano Using Gibbs Sampling, un nombre que refleja los métodos computacionales implementados originalmente. Es interesante que BUGS es anterior al documento seminal del muestreo de Gibbs en la inferencia bayesiana por [87] y tiene sus raíces en el modelado de procesos de decisión complejos con sistemas de inteligencia artificial durante los años 80 [88].

Este origen ha influido en la arquitectura del sistema, que separa claramente base de conocimientos (por ejemplo, uso de programación declarativa, modelado gráfico, etc.) y motor de inferencia (es decir, sacar conclusiones en una situación específica). En otras palabras, la filosofía detrás de BUGS es: primero construir un modelo probabilístico para el problema en cuestión, y luego usar los datos observados para actualizar el modelo. Si no se observan datos, el modelo puede para la inferencia predictiva previa.

En BUGS suponemos que la distribución posterior de la articulación puede ser representada por un gráfico acíclico dirigido (DAG). Un DAG describe todos los componentes

del modelo como nodos y sus relaciones mediante el comando vínculos entre ellos. Los nodos ovalados representan cantidades estocásticas y los nodos cuadrados parámetros conocidos. Las flechas simples definen las distribuciones de probabilidad, por ejemplo,  $tp_i \sim Bin(TPR_i, n_{1,i})$ , mientras que las flechas dobles representan relaciones lógicas o funcionales, por ejemplo  $TPR_i = g^{-1}((D_i + S_i)/2)$ . El gran cuadrado que contiene varios nodos y flechas se llama una placa, la cual representa estructuras sub-modelo replicadas sobre un índice, por ejemplo,  $i = 1, \dots, N$ .

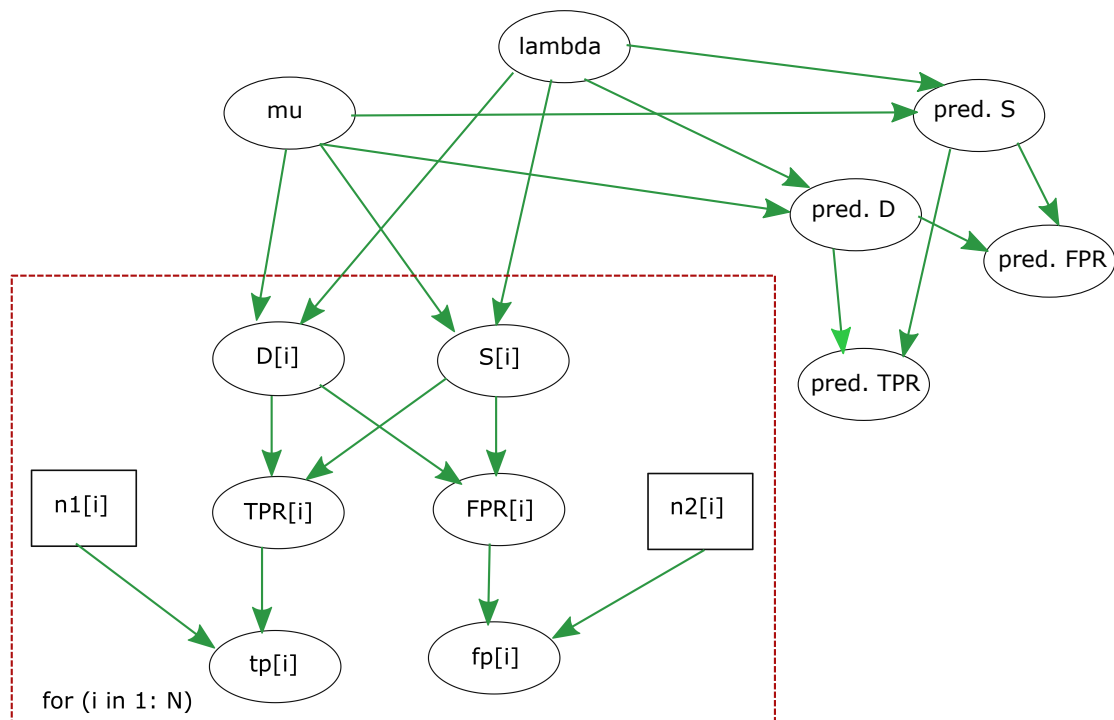


Figura 2.11: Gráfico acíclico dirigido (DAG) del modelo estructural bivariante para meta-análisis de pruebas de diagnóstico

## 2.8. Modelos Mixtos No Lineales

### 2.8.1. Introducción

El procedimiento NLMIXED (PROC NLMIXED) se ajusta a modelos mixtos no lineales, es decir, modelos en los que tanto los efectos fijos como los aleatorios se introducen de forma no lineal. Estos modelos tienen una amplia variedad de aplicaciones, dos de las más comunes son la farmacocinética y los datos binomiales sobredispersados. PROC NLMIXED le permite especificar una distribución condicional para sus datos (dados los efectos

aleatorios) que tenga una forma estándar (normal, binomial, Poisson) o una distribución general que codifique utilizando sentencias de programación SAS.

PROC NLMIXED se ajusta a modelos mixtos no lineales maximizando una aproximación al modelo de probabilidad integrada sobre los efectos aleatorios. Están disponibles diferentes aproximaciones integrales, siendo las principales la cuadratura gaussiana adaptativa y una aproximación de primer orden de la serie de Taylor. Una variedad de técnicas alternativas de optimización están disponibles para llevar a cabo la maximización; el valor por defecto es un algoritmo dual cuasi-Newton.

La convergencia exitosa del problema de optimización resulta en estimaciones de parámetros junto con sus errores estándar aproximados basados en la segunda matriz derivada de la función de verosimilitud. PROC NLMIXED le permite utilizar el modelo estimado para construir predicciones de funciones arbitrarias utilizando estimaciones empíricas de Bayes de los efectos aleatorios. También puede estimar funciones arbitrarias de los parámetros no aleatorios, y PROC NLMIXED calcula sus errores estándar aproximados utilizando el método delta.

### 2.8.2. Estado del arte

[15] y [53] ofrecen buenas perspectivas, así como desarrollos teóricos generales y ejemplos de modelos mixtos no lineales. [57] es una referencia primaria para la teoría y las técnicas computacionales de PROC NLMIXED. Describen y comparan varias aproximaciones de verosimilitud integradas y proporcionan evidencia de que la cuadratura gaussiana adaptativa es uno de los mejores métodos. [18] también utilizan la cuadratura gaussiana para modelos mixtos no lineales, aunque la suave densidad no paramétrica que defienden para los efectos aleatorios no está en la actualidad disponible en PROC NLMIXED.

Los enfoques tradicionales para ajustar modelos mixtos no lineales incluyen expansiones de la serie de Taylor, que se expanden en torno a cero o a las mejores predicciones empíricas lineales insesgadas de los efectos aleatorios. La primera es la base del conocido método de primer orden ([19] y [20]; [21]), y está disponible opcionalmente en PROC NLMIXED. Este último es la base para el método de estimación de [56], y no está disponible en PROC NLMIXED. Sin embargo, la aproximación laplaciana estrechamente relacionada

es una opción; es equivalente a la cuadratura gaussiana adaptativa con sólo un punto de cuadratura. La aproximación laplaciana y su relación con el método de Lindstrom-Bates se discuten mediante: [23]; [52]; ([25], [26]); [53]; [28].

Existe una literatura paralela en el área de los modelos mixtos lineales generalizados, en los que los efectos aleatorios aparecen como parte del predictor lineal dentro de una función de enlace. Los métodos de Taylor-series similares a los que acabamos de describir se discuten en artículos como: [29]; [30]; [31]; [32]; [33]; [34]; [51]; [52]; [37], pero tales métodos no han sido implementados en PROC NLMIXED porque pueden producir resultados sesgados en ciertas situaciones de datos binarios ([38]; [39]). En cambio, en PROC NLMIXED se dispone de un enfoque de cuadratura numérica, como se discute en: [40]; [41]; [42]; [43]; [44]; [45]; [46]; [47].

Los modelos mixtos no lineales tienen aplicaciones importantes en farmacocinética, y [48] proporciona una comparación de gran alcance de muchas técnicas populares. [49] proporcionan una extensa bibliografía sobre modelos mixtos no lineales y su uso en farmacocinética.

### 2.8.3. Supuestos de modelado y notación

PROC NLMIXED opera bajo el siguiente marco general para modelos mixtos no lineales. Supongamos que se tiene un vector de datos observado  $y_i$  para cada uno de los individuos  $i=1, \dots, s$ . Se supone que los  $y_i$  son independientes a través de  $i$ , pero es probable que exista una covarianza dentro de los individuos porque cada uno de los elementos de  $y_i$  se mide sobre el mismo individuo. Como mecanismo estadístico para modelar esta covarianza dentro del individuo, se asume que existen vectores latentes de efectos aleatorios  $u_i$  de baja dimensión (típicamente uno o dos) que también son independientes a través de  $i$ . Supongamos también que existe un modelo apropiado que vincula  $y_i$  y  $u_i$ , lo que conduce a la unión función de densidad de probabilidad.

$$p(y_i|X_i, \phi, u_i)q(u_i|\xi)$$

donde  $X_i$  es una matriz de variables explicativas observadas y  $\phi$  y  $\xi$  son vectores de parámetros desconocidos.

Decimos que  $\theta = [\phi, \xi]$  y asumamos que es de dimensión  $n$ . Entonces las inferencias sobre  $\theta$  se basan en la función de verosimilitud marginal.

$$m(\theta) = \prod_{i=1}^s \int p(y_i | X_i, \phi, u_i) q(u_i | \xi) du_i \quad (2.8.19)$$

En particular, la función

$$f(\theta) = -\log m(\theta)$$

es minimizada sobre  $\theta$  numéricamente para poder estimar  $\theta$ , y la inversa de la matriz hessiana (segunda derivada) en las estimaciones proporciona una matriz de varianza-covarianza aproximada para la estimación de  $\theta$ . La función  $f(\theta)$  se conoce tanto como la función del log de verosimilitud de negativo y como la función objetiva para la optimización.

Como ejemplo del marco general anterior, considere el ejemplo de la curva de crecimiento no lineal. Aquí, la distribución condicional  $p(y_i | X_i, \phi, u_i)$  es normal con media:

$$\frac{b_1 + u_{i1}}{1 + \exp[-(d_{ij} - b_2)/b_3]}$$

y varianza  $\sigma_e^2$  para  $\phi = [b_1, b_2, b_3, \sigma_e^2]$ . Además,  $u_i$  es un escalar y  $q(u_i | \xi)$  es normal con media 0 y varianza  $\sigma_u^2$  por lo tanto  $\xi = \sigma_u^2$ .

La cantidad  $\theta^{(k)}$  se refiere al parámetro vector en la  $k$ -ésima iteración, el vector  $g(\theta)$  se refiere al vector de gradiente  $\nabla f(\theta)$ , y la matriz  $H(\theta)$  se refiere al Hessiano  $\nabla^2 f(\theta)$ .

### Modelos mixtos multinivel anidados no lineales

El marco general para los modelos mixtos no lineales multinivel anidados en el caso de dos niveles puede explicarse de la siguiente manera. Que  $y_{j(i)}$  sea el vector de respuesta observado sobre el individuo  $j$  que está anidado dentro del individuo  $i$ , donde  $j$  es comúnmente referido como el individuo de segundo nivel e  $i$  es el individuo de primer

nivel. Hay  $s$  primer-nivel de individuos, y cada una de las  $s_i$  tienen segundo-nivel que están anidados en su interior.

Un ejemplo es  $y_{j(i)}$ , que son las alturas de los estudiantes en la clase  $j$  de la escuela  $i$ , donde  $j = 1, \dots, s_i$ ; si para cada  $i$  e  $i=1, \dots, s$ . Supongamos que existen vectores latentes de efectos aleatorios  $v_{j(i)}$  y  $v_i$  de pequeñas dimensiones para modelar dentro de la covarianza de los individuos.

Supongamos también que existe un modelo apropiado que vincule  $y_{j(i)}$  y  $(v_{j(i)}, v_i)$  y si se utiliza la notación  $y_i = (y_{1(i)}, \dots, y_{s_i(i)})$ ,  $u_i = (v_i, v_{1(i)}, \dots, v_{s_i(i)})$ , y  $\xi = (\xi_1, \xi_2)$ , la función de densidad conjunta en términos de individuos de primer-nivel puede expresarse como:

$$p(y_i|X_i, \phi, u_i)q(u_i|\xi) = \left( \prod_{j=1}^{s_i} p(y_{j(i)}|X_i, \phi, v_i, v_{j(i)})q_2(v_{j(i)}|\xi_2) \right) q_1(v_i|\xi_1) \quad (2.8.20)$$

Como se definió en la sección anterior, la función de verosimilitud marginal donde  $\theta = [\phi, \xi]$  es:

$$m(\theta) = \prod_{i=1}^s \int p(y_i|X_i, \phi, u_i)q(u_i|\xi)du_i \quad (2.8.21)$$

Una vez más, la función

$$f(\theta) = -\log m(\theta)$$

es minimizada sobre  $\theta$  con la finalidad de estimar numéricamente  $\theta$ . Los modelos que tienen más de dos niveles siguen una notación similar.

#### 2.8.4. Aproximaciones integrales

Una parte importante del método de máxima verosimilitud marginal descrito anteriormente es el cálculo de la integral sobre los efectos aleatorios. El método por defecto en PROC NLMIXED para calcular esta integral es la cuadratura gaussiana adaptativa como se describe en [57]. Otro método de aproximación es el método de primer orden de ([19]; [20]). A continuación se describen estos dos métodos.



### Cuadratura Gaussiana Adaptativa

Un método de cuadratura aproxima una integral dada por una suma ponderada sobre abscisas predefinidas para los efectos aleatorios. Una buena aproximación puede obtenerse normalmente con un número adecuado de puntos de cuadratura, así como con un centrado y escalado adecuados de las abscisas. La cuadratura gaussiana adaptativa para la integral sobre  $u_i$  centra la integral en la estimación empírica de Bayes de  $u_i$ , definida como el vector  $\hat{u}_i$  que minimiza.

$$-\log[p(y_i|X_i, \phi, u_i)q(u_i|\xi)]$$

con  $\phi$  y  $\xi$  sus estimaciones actuales. La matriz Hessiana final de esta optimización puede ser usada para escalar las abscisas de cuadratura.

Supongamos  $(z_j, w_j; j = 1, \dots, p)$  denota las abscisas y pesos estándar de Gauss-Hermite ([61]; [62]). La aproximación integral de la cuadratura gaussiana adaptativa es la siguiente:

$$\int p(y_i|X_i, \phi, u_i)q(u_i|\xi)du_i \approx 2^{r/2}|\Gamma(X_i, \theta)|^{-1/2} \sum_{j_1=1}^p \dots \sum_{j_r=1}^p \left[ p(y_i|X_i, \phi, a_{j_1, \dots, j_r})q(a_{j_1, \dots, j_r}|\xi) \prod_{k=1}^r w_{j_k} \exp z_{j_k}^2 \right]$$

Donde  $r$  es la dimensión de  $u_i$ ,  $\Gamma(X_i, \theta)$  es la matriz hessiana de la minimización empírica de Bayes,  $z_{j_1, \dots, j_r}$  es un vector con elementos  $(z_{j_1, \dots, j_r})$  y

$$a_{j_1, \dots, j_r} = \hat{u}_i + 2^{1/2}\Gamma(X_i, \theta)^{-1/2}z_{j_1, \dots, j_r}$$

PROC NLMIXED selecciona el número de puntos de cuadratura de forma adaptativa evaluando la función del log de verosimilitud en los valores iniciales de los parámetros hasta que dos evaluaciones sucesivas tengan una diferencia relativa menor que el valor de la tolerancia.

### Método de primer orden

Otra aproximación integral disponible en PROC NLMIXED es el método de primer orden de Beal y Sheiner ([19]; [20]) y [21]. Esta aproximación se utiliza sólo en el caso

de que  $p(y_i|X_i, \phi, u_i)$  es normal, es decir.

$$p(y_i|X_i, \phi, u_i) = (2\pi)^{-n_i/2} |R_i(X_i, \phi)|^{-1/2} \exp\left\{-(1/2)[y_i - m_i(X_i, \phi, u_i)]' R_i(X_i, \phi)^{-1} [y_i - m_i(X_i, \phi, u_i)]\right\}$$

Donde  $n_i$  es la dimensión de  $y_i$ ,  $R_i$  es la diagonal de la matriz de varianza, y  $m_i$  es el vector de media condicional de  $y_i$ .

La aproximación de primer orden se obtiene mediante la expansión de la  $m_i(X_i, \phi, u_i)$  con una expansión de la serie de Taylor de un término sobre  $u_i = 0$ , resultando en la aproximación:

$$p(y_i|X_i, \phi, u_i) \approx (2\pi)^{-n_i/2} |R_i(X_i, \phi)|^{-1/2} \exp\left\{-(1/2)[y_i - m_i(X_i, \phi, 0) - Z_i(X_i, \phi)u_i]' R_i(X_i, \phi)^{-1} [y_i - m_i(X_i, \phi, 0) - Z_i(X_i, \phi)u_i]\right\}$$

donde  $Z_i(X_i, \phi)$  es la matriz jacobiana  $\partial m_i(X_i, \phi, u_i)/\partial u_i$  evaluada en  $u_i = 0$ .

Asumiendo que  $q(u_i|\xi)$  es normal con media 0 y matriz de varianza  $G(\xi)$ , la aproximación integral de primer-orden es calculable en forma cerrada después de completar el cuadrado:

$$\int p(y_i|X_i, \phi, u_i)q(u_i|\xi)du_i \approx (2\pi)^{-n_i/2} |V_i(X_i, \theta)|^{-1/2} \exp\left\{-(1/2)[y_i - m_i(X_i, \phi, 0)]' V_i(X_i, \theta)^{-1} [y_i - m_i(X_i, \phi, 0)]\right\}$$

donde  $V_i(X_i, \theta) = Z_i(X_i, \phi)G(\xi)Z_i(X_i, \phi)' + R_i(X_i, \phi)$ .

La aproximación resultante para  $f(\theta)$  es entonces minimizada sobre  $\theta = [\phi, \xi]$ ; para obtener las estimaciones de primer-orden. PROC NLMIXED utiliza derivados de diferencias finitas de la aproximación integral de primer orden cuando lleva a cabo la optimización predeterminada dual cuasi-newton.

### 2.8.5. Aproximaciones por diferencias finitas de derivadas

Las opciones FD y FDHESSIAN especifican el uso de aproximaciones de diferencias finitas de las derivadas. La opción FD especifica que todas las derivadas son aproximados usando evaluaciones de función, y la opción FDHESSIAN especifica que las derivadas de segundo orden se aproximan utilizando evaluaciones de gradiente.

El cálculo de los derivados mediante aproximaciones por diferencias finitas puede llevar mucho tiempo, especialmente para derivadas de segundo orden basadas únicamente en los valores de la función objetivo (FD = opción). Si las derivadas analíticas son difíciles de obtener (por ejemplo, si una función es calculada por un proceso iterativo), se puede considerar una de las técnicas de optimización que utilizan únicamente derivadas de primer orden (QUANEW, DBLDOG o CONGRA). En las expresiones que siguen,  $\theta$  denota el parámetro vector,  $h_i$  denota el tamaño del paso para el  $i$ -ésimo parámetro, y  $e_i$  es un vector de ceros con un 1 en la  $i$ -ésima posición.

#### 1. Aproximaciones por diferencias hacia adelante

Las aproximaciones derivadas por diferencia hacia adelante consumen menos tiempo de la computadora, pero usualmente no son precisas como aproximaciones que utilizan fórmulas de diferencia central.

- Para derivados de primer orden, se requieren  $n$  llamadas de función adicionales:

$$g_i = \frac{\partial f}{\partial \theta_i} \approx \frac{f(\theta + h_i e_i) - f(\theta)}{h_i}$$

- Para derivados de segundo orden basados sólo en llamadas de función ([70], p. 80),  $n + n^2/2$  llamadas de función adicionales son necesarias para el Hessiano:

$$\frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \approx \frac{f(\theta + h_i e_i + h_j e_j) - f(\theta + h_i e_i) - f(\theta + h_j e_j) + f(\theta)}{h_i h_j}$$

- Para las derivadas de segundo orden basados en llamadas de gradiente ([70], p. 103), se requieren  $n$  llamadas de gradiente adicionales:

$$\frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \approx \frac{g_i(\theta + h_j e_j) - g_i(\theta)}{2h_j} + \frac{g_j(\theta + h_i e_i) - g_j(\theta)}{2h_i}$$

## 2. Aproximaciones de diferencia central

Las aproximaciones de diferencia central suelen ser más precisas, pero consumen más tiempo de computadora que las aproximaciones que utilizan fórmulas derivadas de diferencias hacia adelante.

- Para derivados de primer orden, requieren  $2n$  llamadas de función adicionales:

$$g_i = \frac{\partial f}{\partial \theta_i} \approx \frac{f(\theta + h_i e_i) - f(\theta - h_i e_i)}{2h_i}$$

- Para las derivadas de segundo orden basados únicamente en llamadas de función ([62], p. 884), se requieren  $2n+4n^2/2$  llamadas de función adicionales.

$$\frac{\partial^2 f}{\partial \theta_i^2} \approx \frac{-f(\theta + 2h_i e_i) + 16f(\theta + h_i e_i) - 30f(\theta) + 16f(\theta - h_i e_i) - f(\theta - 2h_i e_i)}{12h_i^2}$$

$$\frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \approx \frac{f(\theta + h_i e_i + h_j e_j) - f(\theta + h_i e_i - h_j e_j) - f(\theta - h_i e_i + h_j e_j) + f(\theta - h_i e_i - h_j e_j)}{4h_i h_j}$$

- Para las derivadas de segundo orden basados en llamadas de gradiente, se requieren  $2n$  llamadas graduales adicionales:

$$\frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \approx \frac{g_i(\theta + h_j e_j) - g_i(\theta - h_j e_j)}{4h_j} + \frac{g_j(\theta + h_i e_i) - g_j(\theta - h_i e_i)}{4h_i}$$

Puede utilizar la opción FDIGITS para especificar el número de dígitos exactos en la evaluación de la función objetivo. Esta especificación es útil para determinar un tamaño de intervalo apropiado  $h$  que se utilizará en las fórmulas de diferencia finita.

Los tamaños de paso  $h_j$ ,  $j=1, \dots, n$  se definen de la siguiente manera:

- Para la aproximación por diferencias hacia adelante de las derivadas de primer orden que utilizan llamadas de función y de las derivadas segundo orden que utilizan llamadas de gradiente,  $h_j = \sqrt[3]{n}(1 + |\theta_j|)$ .

- Para la aproximación de diferencias hacia adelante de derivados de segundo orden que utilizan sólo llamadas de función y todas las fórmulas de diferencia central,  $h_j = \sqrt[3]{n}(1 + |\theta_j|)$ .

El valor de n está definido por la opción FDIGITS:

- Si especifica el número de dígitos exactos utilizando FDIGITS=r, n se fija en  $10^{-r}$ .
- Si no especifica la opción FDIGITS, n se ajusta a la precisión de la máquina  $\epsilon$ .

### 2.8.6. Escala Hessiana

Las filas y columnas de la matriz Hessiana se pueden escalar cuando se utiliza la región de confianza, Newton-Raphson, y técnicas de optimización de double-dogleg. Cada elemento  $H_{i,j}$ ,  $i, j = 1, \dots, n$  se divide por la escala  $d_i d_j$ , donde el vector de escala  $d = (d_1, \dots, d_n)$  se actualiza de forma iterativa de la forma especificada por la directiva HESCAL=i, según se indica:

i=0: No se realiza ningún escalado (equivalente a  $d_i=1$ ).

i  $\neq$  0: La primera iteración y cada iteración de reinicio se establece:

$$d_i^{(0)} = \sqrt{\max(|H_{i,i}^{(0)}|, \epsilon)}$$

i=1: Ver [72]:

$$d_i^{(k+1)} = \max \left[ d_i^{(k)}, \sqrt{\max(|H_{i,i}^{(k)}|, \epsilon)} \right]$$

i=2: Ver [63]:

$$d_i^{(k+1)} = \max \left[ 0, 6d_i^{(k)}, \sqrt{\max(|H_{i,i}^{(k)}|, \epsilon)} \right]$$

i=3:  $d_i$  se reajusta en cada iteración:

$$d_i^{(k+1)} = \sqrt{\max(|H_{i,i}^{(k)}|, \epsilon)}$$

En las ecuaciones anteriores,  $\epsilon$  es la precisión relativa de la máquina o, equivalente, la mayor precisión doble que, sumado a 1, da como resultado 1.

### 2.8.7. Métodos de conjunto activo

El parámetro vector  $\theta \in R^n$  puede estar sujeto a un conjunto de  $m$  restricciones lineales de igualdad y desigualdad:

$$\sum_{j=1}^n a_{ij}\theta_j = b_i \quad \text{para } i = 1, \dots, m_e$$

$$\sum_{j=1}^n a_{ij}\theta_j \geq b_i \quad \text{para } i = m_e + 1, \dots, m$$

Los coeficientes  $a_{ij}$  y derecha  $b_i$  de las restricciones de igualdad y desigualdad se recogen en la matriz  $A$  de  $m \times n$  y en el vector  $b$ .

Las restricciones lineales  $m$  definen una región viable  $G$  en  $R^n$  que debe contener el punto  $\theta^*$  que minimiza el problema. Si la región  $G$  está vacía, no existe ninguna solución al problema de optimización.

En PROC NLMIXED, todas las técnicas de optimización utilizan métodos de conjuntos activos. La iteración comienza con un punto factible  $\theta^{(0)}$ , que se puede proporcionar o que se puede calcular mediante el algoritmo de [71] implementado en PROC NLMIXED. El algoritmo pasa entonces de un punto factible  $\theta^{(k-1)}$  a un punto más factible  $\theta^{(k)}$  a lo largo de una dirección de búsqueda factible  $s^{(k)}$ ,

$$\theta^{(k)} = \theta^{(k-1)} + \alpha^{(k)} s^{(k)}, \quad \alpha^{(k)} > 0$$

Teóricamente, la trayectoria de los puntos  $\theta^{(k)}$  nunca deja la región factible  $G$  del problema de optimización, pero puede alcanzar sus límites. El conjunto activo  $A^{(k)}$  del punto  $\theta^{(k)}$  se define como el conjunto de índices de todas las restricciones lineales de igualdad y las restricciones de desigualdad que se satisfacen en  $\theta^{(k)}$ . Si no hay ninguna restricción activa  $\theta^{(k)}$ , el punto se encuentra en el interior de  $G$ , y el conjunto activo  $A^{(k)} = \emptyset$ ; está vacío. Si el punto  $\theta^{(k)}$  en iteración  $k$  golpea el límite de la restricción de la desigualdad  $i$ , esta restricción  $i$  se activa y se añade a  $A^{(k)}$ . Cada igualdad y cada restricción de desigual-

dad activa reduce la dimensión (grados de libertad) del problema de optimización.

En la práctica, las restricciones activas sólo pueden satisfacerse con una precisión finita. La opción `LCEPSILON=r` especifica el rango de las restricciones lineales activas y violadas. Si el punto  $\theta^{(k)}$  cumple la condición.

$$\left| \sum_{j=1}^n a_{ij} \theta_j^{(k)} - b_i \right| \leq t$$

donde  $t = r(|b_i| + 1)$ , la restricción  $i$  se reconoce como una restricción activa. De lo contrario, la restricción  $i$  es una desigualdad inactiva o una desigualdad violada o una restricción de igualdad. Debido a errores de redondeo en el cálculo de la dirección de búsqueda proyectada, el error puede acumularse de modo que una iteración  $\theta^{(k)}$  se salga de la región factible.

En esos casos, PROC NLMIXED podría tratar de tirar de la iteración  $\theta^{(k)}$  de vuelta a la región factible. Sin embargo, en algunos casos el algoritmo necesita aumentar la región factible aumentando el valor de `LCEPSILON=r`. Si esto sucede, se visualiza un mensaje en la salida del log.

Si el algoritmo no puede mejorar el valor de la función objetivo al pasar de una restricción activa al interior de la región factible, hace que esta restricción de desigualdad sea una restricción de igualdad en la siguiente iteración. Esto significa que el conjunto activo  $A^{k+1}$  todavía contiene la restricción  $i$ . De lo contrario, libera la restricción de desigualdad activa y aumenta la dimensión del problema de optimización en la siguiente iteración.

Si la solución  $\theta^*$  está sujeta a  $n_{act}$  restricciones de igualdad lineal o de desigualdad activa, la descomposición QR de la matriz  $n \times n_{act}$   $\hat{A}'$  de las restricciones lineales se calcula mediante  $\hat{A}' = QR$ , donde  $Q$  es una matriz ortogonal de  $n \times n$  y  $R$  es una matriz triangular superior de  $n \times n_{act}$ . Las  $n$  columnas de la matriz  $Q$  pueden separarse en dos matrices,  $Q = [Y, Z]$ , donde  $Y$  contiene las primeras  $n_{act}$  columnas ortogonales de  $Q$  y  $Z$  contiene las últimas  $n - n_{act}$  columnas ortogonales de  $Q$ . La  $n \times (n - n_{act})$  columna-ortogonal de la matriz  $Z$  también se denomina matriz de espacio nulos de las restricciones lineales activas  $\hat{A}'$ . Las  $n - n_{act}$  columnas de  $n \times (n - n_{act})$  de la matriz  $Z$  forman una base ortogonal a las filas de  $n_{act} \times n$  de la matriz  $\hat{A}'$ .

Al final de la iteración, PROC NLMIXED calcula el gradiente proyectado  $g_Z$ ,  
 $g_Z = Z'g$ .

En el caso de la optimización restringida por límites, los elementos del gradiente proyectado corresponden a los elementos de gradiente de parámetros libres. Una condición necesaria para ser  $\theta^*$  un mínimo local del problema de optimización es  
 $g_Z(\theta^*) = Z'g(\theta^*) = 0$ .

La matriz simétrica de  $n_{act} \times n_{act}$   $G_Z$ ,

$$G_Z = Z'GZ$$

se llama matriz hessiana proyectada. Una condición necesaria de segundo orden para ser un minimizador local requiere que la matriz hessiana proyectada sea semidefinida positiva.

Aquellos elementos del  $n_{act}$  vector  $\lambda$  de las estimaciones de primer orden de los multiplicadores de Lagrange,

$$\lambda = (\hat{A}\hat{A}')^{-1}\hat{A}ZZ'g$$

que corresponden a restricciones de desigualdad activa, indican si se puede obtener una mejora de la función objetivo liberando esta restricción activa. Para la minimización, un multiplicador negativo significativo de Lagrange indica que se puede lograr una posible reducción de la función objetivo liberando esta restricción lineal activa. La opción LCDEACT=r especifica un umbral r para el multiplicador de Lagrange que determina si una restricción de desigualdad activa permanece activa o puede ser desactivada.

### 2.8.8. Métodos de búsqueda por líneas

En cada iteración k, las técnicas de minimización (dual) cuasi-Newton, gradiente conjugado y Newton-Raphson usan algoritmos iterativos de búsqueda por línea que intentan



optimizar una aproximación lineal, cuadrática o cúbica de  $f$  a lo largo de un descenso factible dirección de búsqueda  $s^{(k)}$ ,

$$\theta^{(k+1)} = \theta^{(k)} + \alpha^{(k)} s^{(k)}, \quad \alpha^{(k)} > 0$$

calculando un escalar aproximadamente óptimo  $\alpha^{(k)}$ .

Por lo tanto, un algoritmo de búsqueda por línea es un proceso iterativo que optimiza una función no lineal  $f(\alpha)$  de un parámetro ( $\alpha$ ) dentro de cada iteración  $k$  de la técnica de optimización. Dado que el proceso de iteración externa es basado únicamente en la aproximación de la función objetivo, la iteración interna del algoritmo de búsqueda por línea no tiene que ser perfecto. Por lo general, es satisfactorio que la elección de  $\alpha$  reduce significativamente (en una minimización) la función objetivo. Los criterios que se utilizan a menudo para la terminación de los algoritmos de búsqueda por línea son los siguientes Condiciones de Goldstein (ver [66]).

### 2.8.9. Restricción de la longitud del paso

Casi todos los algoritmos de búsqueda por línea utilizan técnicas de extrapolación iterativa que pueden llevarlos fácilmente a puntos (factibles) en los que la función objetivo  $f$  ya no está definida o es difícil de calcular. Por lo tanto, PROC NLMIXED proporciona opciones que restringen la longitud del paso  $\alpha$  o el radio de la región de confianza  $\Delta$ , especialmente durante las primeras iteraciones principales.

El producto interior  $g's$  de la pendiente  $g$  y la dirección de búsqueda  $s$  es la pendiente de  $f(\alpha) = f(\theta + \alpha s)$  a lo largo de la dirección de búsqueda  $s$ . El valor inicial predeterminado  $\alpha^{(0)} = \alpha^{(k,0)}$  en cada algoritmo de búsqueda por línea ( $\min_{\alpha>0} f(\theta + \alpha s)$ ) durante la iteración principal  $k$  se calcula en tres pasos:

1. El primer paso utiliza la diferencia  $|\Delta f| = |f^{(k)} - f^{(k-1)}|$  de los valores de la función durante las dos últimas iteraciones consecutivas o el valor final del tamaño del paso  $\alpha$  de la última iteración  $k-1$  para calcular un primer valor de  $\alpha_1^{(0)}$ .

- Si no se utiliza la opción DAMPSTEP,

$$\alpha_1^{(0)} = \begin{cases} \textit{paso} & \textit{si } 0.1 \leq \textit{paso} \leq 10 \\ 10 & \textit{si } \textit{paso} > 10 \\ 0.1 & \textit{si } \textit{paso} < 0.1 \end{cases}$$

con

$$\textit{paso} = \begin{cases} |\Delta f|/|g's| & \textit{si } |g's| \geq \epsilon \max(100 \times |\Delta f|, 1) \\ 1 & \textit{en otro caso} \end{cases}$$

Este valor de  $\alpha_1^{(0)}$  puede ser demasiado grande y puede llevar a una evaluación de funciones difícil o imposible, especialmente para funciones altamente no lineales como la función EXP.

- Si se utiliza la opción DAMPSTEP=r,

$$\alpha_1^{(0)} = \min(1, r\alpha_-)$$

El valor inicial para la nueva longitud de paso no puede ser mayor que r veces la longitud de paso final  $\alpha_-$  de la iteración anterior. El valor por defecto es  $r = 2$ .

2. Durante las primeras cinco iteraciones, el segundo paso le permite reducir  $\alpha_1^{(0)}$  a un valor inicial menor  $\alpha_2^{(0)}$  utilizando la opción INSTEP=r:

$$\alpha_2^{(0)} = \min(\alpha_1^{(0)}, r)$$

Después de más de cinco iteraciones,  $\alpha_2^{(0)}$  se ajusta a  $\alpha_1^{(0)}$ .

3. El tercer paso puede reducir aún más la longitud del paso en

$$\alpha_3^{(0)} = \min(\alpha_2^{(0)}, \min(10, u))$$

donde u es la longitud máxima de un escalón dentro de la región factible.

### 2.8.10. Funciones del log de verosimilitud integradas

Esta sección muestra las fórmulas básicas utilizadas por el procedimiento NLMIXED para calcular las funciones del log de verosimilitud de los datos dados los efectos aleatorios. Tenga en cuenta, sin embargo, que además de estas ecuaciones básicas, el procedimiento NLMIXED emplea una serie de verificaciones de los valores que faltan y de la aritmética de coma flotante.

Puede ver el programa completo utilizado por el procedimiento NLMIXED para calcular el condicional de las funciones del log de verosimilitud  $l(\phi; y)$  añadiendo la opción de depuración LIST a la sentencia PROC NLMIXED.

$$Y \sim normal(m, v)$$

$$l(m, v; y) = -\frac{1}{2} \left( \log \{2\pi\} + \frac{(y - m)^2}{v} + \log \{v\} \right)$$

$$E[Y] = m$$

$$Var[Y] = v$$

$$v > 0$$

$$Y \sim binary(p)$$

$$l_1(p; y) = \begin{cases} y \log \{p\} & y > 0 \\ 0 & \text{en otro caso} \end{cases}$$

$$l_2(p; y) = \begin{cases} (1 - y) \log \{1 - p\} & y < 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$l(p; y) = l_1(p; y) + l_2(p; y)$$

$$E[Y] = p$$

$$Var[Y] = p(1 - p)$$

$$0 < p < 1$$

$Y \sim \text{binomial}(n, p)$

$$l_c = \log \{\Gamma(n+1)\} - \log \{\Gamma(y+1)\} - \log \{\Gamma(n-y+1)\}$$

$$l_1(n, p; y) = \begin{cases} y \log \{p\} & y > 0 \\ 0 & \text{en otro caso} \end{cases}$$

$$l_2(n, p; y) = \begin{cases} (n-y) \log \{1-p\} & n-y > 0 \\ 0 & \text{en otro caso} \end{cases}$$

$$l(n, p; y) = l_c + l_1(n, p; y) + l_2(n, p; y)$$

$$E[Y] = np$$

$$\text{Var}[Y] = np(1-p)$$

$$0 < p < 1$$

$Y \sim \text{gamma}(a, b)$

$$l(a, b; y) = -a \log\{b\} - \log\{\Gamma(a)\} + (a-1)\log\{y\} - y/b$$

$$E[Y] = ab$$

$$\text{Var}[Y] = ab^2$$

$$a > 0$$

$$b > 0$$

$Y \sim \text{negbin}(n, p)$

$$l(n, p; y) = \log\{\Gamma(n+y)\} - \log\{\Gamma(n)\} - \log\{\Gamma(y+1)\} + n \log\{p\} + y \log\{1-p\}$$

$$E[Y] = nP = n \left( \frac{1-p}{p} \right)$$

$$\text{Var}[Y] = nP(1-P) = n \left( \frac{1-p}{p} \right) \frac{1}{p}$$

$$n \geq 0$$

$$0 < p < 1$$

Esta forma de distribución binomial negativa es una de las muchas parametrizaciones en las que aparece la función de masa o la función de logaritmo de verosimilitud. Otra parametrización común utiliza.

$$l(n, p; y) = \log\{\Gamma(n + y)\} - \log\{\Gamma(n)\} - \log\{\Gamma(y + 1)\} + n \log\{1 - P/(1 + P)\} + y \log\{P/(1 + P)\}$$

Con  $P = (1 - p)/p, P > 0$ .

$$Y \sim \text{Poisson}(m)$$

$$l(m; y) = y \log\{m\} - m - \log\{\Gamma(y + 1)\}$$

$$E[Y] = m$$

$$\text{Var}[Y] = m$$

$$m > 0$$

### 2.8.11. Especificación jerárquica de modelos

PROC NLMIXED soporta múltiples sentencias RANDOM para acomodar modelos mixtos no lineales multinivel anidados. Si utilizan varias sentencias RANDOM, PROC NLMIXED asume que la variable SUBJECT de cada sentencia RANDOM forma una jerarquía de contención. En la jerarquía de contención, cada SUBJECT es contenida por otra SUBJECT, y la variable SUBJECT= variable que es contenida por todas las variables SUBJECT es considerada “ LA ” variable SUBJECT. Por ejemplo, considere el siguiente modelo anidado de tres niveles que tiene tres variables SUBJECT.

| A | B | C |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 2 |
| 1 | 2 | 1 |
| 1 | 2 | 2 |
| 2 | 1 | 1 |
| 2 | 1 | 2 |
| 2 | 1 | 3 |
| 2 | 2 | 1 |
| 2 | 2 | 2 |

Tabla 2.2: Modelo anidado de tres niveles

Suponga que especifica un modelo anidado de tres niveles utilizando las siguientes tres sentencias RANDOM:

```
random r11 ~ normal(0, sd1) subject = A;
random r21 ~ normal(0, sd2) subject = B(A);
random r31 ~ normal(0, sd3) subject = C(A*B);
```

Entonces PROC NL MIXED asume la jerarquía de contención de la siguiente manera. La jerarquía de primer nivel se define usando SUBJECT=A. De manera similar, la jerarquía de segundo nivel se define usando SUBJECT=B, que está anidada dentro del primer nivel. Finalmente, SUBJECT=C define la jerarquía de tercer nivel que está anidada tanto en el primer como en el segundo nivel.

En resumen, el SUBJECT=A es SUBJECT=B está contenido en A, y C está contenido tanto en A como en B. En este ejemplo, hay dos asignaturas de primer nivel que se determinan utilizando SUBJECT=A.

Basado en la especificación de la jerarquía anterior, la indexación de PROC NL MIXED de sujetos anidados para cada individuo de primer-nivel puede visualizarse de la siguiente manera:

$$\begin{array}{ccccccc}
 & & A_1 & & & A_2 & \\
 \overbrace{\hspace{10em}} & & & & & & \\
 B_{1(1)} & & B_{2(1)} & & B_{1(2)} & & B_{2(2)} \\
 \overbrace{\hspace{10em}} & & & & \overbrace{\hspace{10em}} & & \\
 C_{1(11)} & C_{2(11)} & C_{1(12)} & C_{2(12)} & C_{1(21)} & C_{2(21)} & C_{3(21)} & C_{1(22)} & C_{2(22)}
 \end{array}$$

El  $i$ -ésimo sujeto del primer-nivel se denomina  $A_i$ , y los sujetos anidados del segundo-nivel se denominan  $B_{j(i)}$ , y los sujetos anidados de tercer-nivel se denominan  $C_{k(ij)}$ .

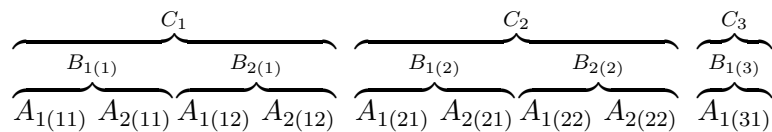
Puede especificar cualquier estructura anidada utilizando la sintaxis SUBJECT=sintaxis en PROC NL MIXED. Por ejemplo, utilizando las siguientes tres sentencias RANDOM, PROC NL MIXED se ajusta a un modelo diferente:

```
random r11 ~ normal(0, sd1) subject = C;
random r21 ~ normal(0, sd2) subject = B(C);
random r31 ~ normal(0, sd3) subject = A(C*B);
```

En este caso, PROC NL MIXED procesa los individuos utilizando SUBJECT=C como “ the ” SUBJECT=variable, y la jerarquía de contención se modifica de la siguiente manera:

|          |          |          |   |          |          |          |
|----------|----------|----------|---|----------|----------|----------|
| <b>A</b> | <b>B</b> | <b>C</b> |   | <b>A</b> | <b>B</b> | <b>C</b> |
| 1        | 1        | 1        |   | 1        | 1        | 1        |
| 1        | 1        | 2        |   | 1        | 1        | 2        |
| 1        | 2        | 1        |   | 1        | 2        | 1        |
| 1        | 2        | 2        | ⇒ | 1        | 2        | 2        |
| 2        | 1        | 1        |   | 2        | 1        | 1        |
| 2        | 1        | 2        |   | 2        | 1        | 2        |
| 2        | 1        | 3        |   | 2        | 2        | 1        |
| 2        | 2        | 1        |   | 2        | 2        | 2        |
| 2        | 2        | 2        |   | 3        | 1        | 2        |

Una vez más, la indexación de PROC NL MIXED de los individuos anidados en esta jerarquía de contención puede visualizarse de la siguiente manera:



Aquí, PROC NL MIXED asume que C es SUBJECT. B está contenido en C, y A está contenido en C y B. En este caso, hay tres individuos de primer-nivel que se determinan utilizando SUBJECT=C. Como se ha explicado anteriormente, en este caso, el i-ésimo individuo del primer-nivel se denomina  $C_i$ , los individuos anidados del segundo-nivel se denominan  $B_{j(i)}$  y para el tercer-nivel, los individuos anidados se denominan  $A_{k(ij)}$ . Tenga en cuenta que la jerarquía de contención podría crear potencialmente más temas que el número único de individuos. Por ejemplo, considere la siguiente tabla, en la que A es SUBJECT y B está anidado dentro del individuo A:

|          |          |
|----------|----------|
| <b>A</b> | <b>B</b> |
| a        | 1        |
| a        | 2        |
| b        | 1        |
| b        | 2        |
| c        | 1        |
| c        | 2        |

Aunque la variable SUBJECT=B tiene sólo dos individuos únicos (1 y 2), cuando la jerarquía de contención que se especifica junto con B está anidada dentro de A, PROC NLMIXED crea seis individuos B anidados. Estos sujetos anidados se pueden denominar como 1(a), 2(a), 1(b), 2(b), 1(c) y 2(c).

PROC NLMIXED no soporta modelos de jerarquía de no contención (o no anidados). Por ejemplo, las siguientes instrucciones no son compatibles, porque el individuo C no está anidado dentro de B y A:

```
random r11 ~ normal(0, sd1) subject = A;  
random r21 ~ normal(0, sd2) subject = B(A);  
random r31 ~ normal(0, sd3) subject = C;
```



## 2.9. Resumen de los Modelos para DTA

|   | <b>Estudio de Intervenciones</b>  | <b>Prueba de Precisión Diagnóstica</b>  |
|---|---|---|
| Número de variables de resultado  | Resultado único   | Pareja de resultados, sensibilidad y especificidad con correlación inversa  |
| Análisis de la heterogeneidad entre los estudios  | Prueba de ji cuadrado (Q): $p < 0.1$ generalmente indica heterogeneidad significativa. Estadística de Higgins en la $I^2$ : guía general para la interpretación es la siguiente:<br>0.0 a 0.25, no importante<br>0.25 a 0.50, baja<br>0.50 a 0.75, moderada<br>0.75 a 1.0, alta | Las estadísticas de Cochran Q o Higgins $I^2$ por si solas pueden no ser informativas, ya que no tienen en cuenta el efecto umbral. Evaluación visual de la parcela forestal acoplada o parcela SROC para encontrar el efecto umbral<br>Análisis de correlación de Spearman entre la sensibilidad y la tasa de falsos positivos: $r \geq 0.6$ generalmente indica un efecto umbral considerable |
| Resumen analítico   | Meta- Punto de resumen y su IC del 0.95 obtenido con Modelo de efectos fijos: cuando no existe heterogeneidad de los estudios<br>Modelo de efectos aleatorios: cuando se sospecha la existencia de heterogeneidad del estudio   | Punto de resumen Sensibilidad y especificidad y su IC del 0.95 obtenido con el modelo bivariable: recomendado Gráfico resumen (curva SROC)<br>Modelo Moses-Littenberg: no recomendado<br>Curva HSROC: recomendada   |
| <p>IC = intervalo de confianza, HSROC = característica operativa del receptor sumario jerárquico<br/>SROC = característica operativa del receptor sumario</p> |   |   |

Tabla 2.3: Comparación de Meta-Análisis de Estudios Terapéuticos/Intervencionales y Estudios de Precisión de Pruebas de Diagnóstico

| Método                                 | Medidas Resumen  | Ponderación                 | Observaciones   |
|--|--|-----------------------------|---|
| Resumen individual                     | Resumen de sensibilidad, especificidad, LR+, LR- y DOR       | Efectos fijos o aleatorios  | <b>No recomendado:</b><br>Realiza metanálisis separados para cada punto de resumen<br>Ignora el efecto umbral, así como la correlación entre la sensibilidad y especificidad      |
| Métodos jerárquicos (bivariante/HSROC) | Resumen de sensibilidad, especificidad, LR+, LR- y DOR       | Efectos aleatorios          | <b>Recomendado:</b><br>El modelo bivariante se utiliza para calcular los puntos de resumen, mientras que el modelo HSROC para construir la curva SROC.                            |
| Modelo Moses-Littenberg                | curva SROC, AUC, y Q*  | Similar a los efectos fijos | <b>No recomendado:</b><br>No tiene en cuenta la variabilidad entre los estudios<br>No pondera óptimamente los estudios e ignora la correlación entre sensibilidad y especificidad |
| Modelo jerárquico                      | Curva HSROC, AUC, región de confianza y región de predicción | Efectos aleatorios          | <b>Recomendado:</b><br>Explica la heterogeneidad dentro y entre los estudios.<br>Toma en cuenta la correlación entre la sensibilidad y la especificidad.                          |

AUC = área bajo la curva ROC, DOR = odds ratio de diagnóstico, HSROC = característica de operación del receptor sumario jerárquico, SROC = característica de operación del receptor sumario

Tabla 2.4: Métodos estadísticos para el resumen meta-analítico de los estudios de precisión en pruebas de diagnóstico

## 2.10. Criterios para la selección del modelo

En este apartado analizamos la selección del modelo a utilizar, tomando en cuenta si en la estrategia de modelización jerárquica incluimos o no una covariable. Esta decisión de mantener una covariable disminuye el espacio disponible de los programas estadísticos que aceptan esta parametrización. A continuación en la Tabla 2.5 se muestra un resumen de los métodos y el software disponible:

| meta-análisis de precisión de pruebas diagnósticas |                      |       |            |            |            |
|--|----------------------|-------|------------|------------|------------|
| Software   | Moses-<br>Littenburg | HSROC | Bivariante | Covariable | Resolución |
| R  | ✓                    | ✓     | ✓          | ✓          | MCMC       |
| Stata  | ✓                    | ✓     | ✓          | X          | X          |
| SAS  | ✓                    | ✓     | ✓          | ✓          | NLMIXED    |
| WinBUGS  | ✓                    | ✓     | ✓          | ✓          | MCMC       |

Tabla 2.5: Implementación de métodos y software para el meta-análisis de estudios de precisión de pruebas diagnósticas

La resolución de los modelos jerárquico con la restricción de una covariable en el modelo, tiene dos vertientes. La primera está basada en un método de resolución mediante el uso de Cadena Markov y simulación de Monte Carlo (conocida por las siglas MCMC). La estimación del MCMC requiere programación, simulación, evaluación de la convergencia y adecuación del modelo, y síntesis de los resultados de la simulación. Implicará tareas de análisis no triviales para adecuación de las distribuciones previas y la evaluación de las estimaciones de las distribuciones posteriores. La segunda está basada en un método de resolución mediante la implementación de los modelos mixtos no lineales (NLMIXED). El software estadístico SAS utiliza la implementación NLMIXED para adaptar estos modelos y proporcionan medidas de ajuste, como el criterio de información de Akaike (AIC), Akaike mejorado (AICC) y el criterio de información bayesiano (BIC). En la tarea de simulación de los modelos jerárquicos (HSROC y Bivariante) implementamos NLMIXED de SAS para la resolución de los modelos con una covariable en la estimación de los parámetros.

# Capítulo 3

## Metodología Computacional

### 3.1. Introducción

Han pasado casi 30 años desde la publicación del primer meta-análisis de la precisión de las pruebas de diagnóstico (DTA) [1]. Desde entonces, los métodos estadísticos han evolucionado desde simplemente seguir los enfoques utilizados para los meta-análisis de intervención hasta el modelo de resumen ROC (SROC) (también conocido como modelo Moses-Littenberg), que tiene en cuenta el efecto umbral [2], y luego a modelos más avanzados, es decir, el modelo bivariante [3] y el modelo HSROC [4], que también permiten la variabilidad dentro de los estudios y entre ellos.

Aunque los modelos HSROC y bivariante son comúnmente, se consideran las mejores prácticas y se recomiendan en la Manual Cochrane [95], los métodos subóptimos son todavía frecuentemente utilizado en estudios que realizan meta-análisis de precisión diagnóstica, ya que son ampliamente implementados en software y paquetes y más accesible que los referidos. En el reciente libro de Meta- Análisis de diagnóstico, Doebler et al. dieron una visión general de los paquetes disponible para el meta-análisis de DTA y discutieron brevemente sus fortalezas y debilidades [6]. Cualquier software estadístico con modelo general mixto puede ser empleado para ajustar los modelos avanzados de meta-análisis. Sin embargo, los investigadores prefieren paquetes diseñados para análisis específicos y sin trabajo de codificación adicional.

## 3.2. Evaluación del Software disponible para DTA

En esta revisión, haremos un inventario de los paquetes más populares proporcionando meta-análisis de precisión diagnóstica, y los evaluaremos por su accesibilidad, usabilidad, y funciones y metodología con un enfoque de semáforo (Tabla 3.1). Al salvar la brecha metodológica entre los estadísticos y los médicos, permitimos a los autores de la revisión comprender lo que los paquetes pueden hacer por ellos y elegir los métodos correctos.




| Criterio  |  |  |  |
|---|---|---|---|
| <b>Accesibilidad</b>  | No accesible  | Software comercial  | Software libre  |
| <b>Usabilidad</b>   | Muy difícil de usar   | Difícil de usar   | Fácil de usar   |
| <b>Funciones para el meta-análisis y solidez metodológica</b> | No se prefieren los métodos   | Las funciones no están completas  | Los métodos preferidos están disponibles  |
| <b>En general</b>   | No recomendado  | Neutral   | Recomendado   |

Tabla 3.1: Criterio e implicación de los semáforos




### 3.2.1. Programa estadístico Stata

**Accesibilidad**   

Stata es un paquete estadístico ampliamente utilizado en los campos de la biomedicina y la epidemiología, así como en economía, sociología y ciencias políticas. Stata es un software comercial y es necesario comprar una licencia.

**Usabilidad**   

Stata es considerado como un paquete de software integrado “Rápido, Preciso y Fácil de usar”. Tiene una interfaz de línea de comandos y menús y cuadros de diálogo para todos los comandos incorporados. Stata permite comandos escritos por el usuario, que pueden ser instalados directamente en Stata (“`ssc install name-command`”) o descargados de Internet (ado-files).

**Funciones para el meta-análisis y solidez metodológica**   

Hay dos comandos escritos por el usuario en Stata, es decir, `midas` y `metandi`, que implementan algunos métodos estadísticos contemporáneos para el meta-análisis de pruebas

diagnósticas binarias.

En Midas, la sensibilidad y la especificidad del resumen se derivan de del modelo bivariante, y puede representarse en SROC con las regiones de confianza y predicción (opcional). Midas también puede facilitar la exploración de la heterogeneidad y sesgo de publicación. Sin embargo, debido al efecto umbral,  $I^2$  (en la gráfica de forestplot) y embudo de tronco (DOR) no son tan significativos como en el meta-análisis de la intervención. El las principales ventajas de midas son la incorporación de meta-regresión y la comprobación de modelos gráficos, que suelen ser que faltan en otros paquetes. Metandi usa el mismo método bivariante que Midas, pero tiene menos análisis sobre la heterogeneidad (de nuevo, lo que no es malo). La mayor contribución de metandi es que proporciona el Estimaciones de parámetros en el modelo HSROC  $(\Lambda, \Theta, \beta, \sigma_\alpha^2, \sigma_\theta^2)$  y modelo bivariante  $(\mu_A, \mu_B, \sigma_A^2, \sigma_B^2, \rho_{AB})$ .

**En general** 

Stata puede proporcionar estimaciones de parámetros de modelos HSROC y bivariante, así como otros análisis. Dado que Stata es familiar para los investigadores clínicos y la función de hacer clic para correr, le damos nuestra recomendación.

### 3.2.2. Programa estadístico Sas

**Accesibilidad** 

SAS es la abreviatura de su nombre anterior “ Sistema de análisis estadístico ”. SAS gana su popularidad en las industrias farmacéutica y financiera con su gran fuerza en la gestión de conjuntos de datos complejos. SAS también es extremadamente caro, pero la buena noticia es que SAS lanzó una edición universitaria que es gratuita para uso académico y no comercial.

**Usabilidad** 

SAS ofrece una extraordinaria gama de herramientas de análisis de datos para los usuarios, pero es difícil de utilizar y aprender y se considera la elección de los usuarios avanzados. Hay pocos análisis estadísticos que no se puedan hacer en SAS, pero primero hay que manejar el uso de SAS y la curva de aprendizaje es pronunciada. El algoritmo diseñado por el usuario puede ser codificado como SAS Macro, que es similar al archivo ado en el

paquete Stata o R. El objetivo de SAS Macro es hacer que los algoritmos complejos sean fáciles de usar, pero al final descubrirá que no lo son en absoluto.

Para el meta-análisis DTA, se desarrolló una macro SAS denominada MetaDAS para automatizar la adaptación de modelos bivariados y HSROC [7]. Los usuarios poco familiarizados con SAS disponen de un gremio de usuarios detallado que se logra mediante una referencia rápida y un ejemplo práctico.

### Funciones para el meta-análisis y solidez metodológica

MetaDAS puede ajustar modelos jerárquicos para el análisis de una sola prueba, comparaciones de pruebas (incluyendo el tipo de prueba como una covariable), investigaciones de heterogeneidad y meta regresión con covariables a nivel de estudio. Se estiman los parámetros del modelo HSROC  $(\Lambda, \Theta, \beta, \sigma_a^2, \sigma_b^2)$  y modelo bivalente  $(\mu_A, \mu_B, \sigma_A^2, \sigma_B^2, \rho_{AB})$  y el modelo bivalente  $(\mu_A, \mu_B, \sigma_A^2, \sigma_B^2, \rho_{AB})$  y se proporcionan también estimaciones resumidas de las medidas de precisión de la prueba (sensibilidad, especificidad, DOR, LR+ y LR-) de las pruebas individuales y de las medidas relativas de precisión de la prueba y comparación de pruebas. Se pueden añadir covariables a ambos modelos, y se permite más de una covariable, que es la principal ventaja de SAS macro MetaDAS en comparación con otras paquetes como midas en Stata y mada en R.

### En general

SAS es una herramienta poderosa para casi todos los análisis estadísticos, y la macro de MetaDAS está escrita en alta calidad. Sin embargo, SAS es caro y no todos los institutos tienen una licencia SAS. Esto también lleva a que muchos investigadores clínicos no tengan experiencia con el uso de SAS.

### 3.2.3. Programa estadístico R

#### Accesibilidad

R es un entorno de software libre para el análisis estadístico y es el más popular entre los estadísticos. R puede descargarse de CRAN (Comprehensive R Archive Network) con espejos en todo el mundo. Hay varios paquetes R que pueden facilitar el meta-análisis para DTA, que también pueden ser encontrados en CRAN o instalados directamente desde R usando el comando “install”.

**Usabilidad** ○ ● ○

Hablando con franqueza, R no es muy fácil de usar. El comando puede ser bastante desalentador para un principiante de R. Algunas interfaces gráficas de usuario para R, por ejemplo, RStudio, pueden ayudar un poco, pero todavía no son tan convenientes como esos programas de ejecución por clic. La mayoría de los paquetes R proporcionan un documento explicativo detallado con ejemplos de todas las funciones del paquete.

**Funciones para el meta-análisis y solidez metodológica** ○ ● ●

Una ventaja de R es que los nuevos métodos pueden ser implementados muy rápidamente en paquetes R escritos por el usuario. La desventaja puede ser que estos métodos y paquetes no fueron ampliamente discutidos o validados antes de estar en línea.

Se desarrollaron muchos paquetes para estudios de meta-análisis para DTA, incluyendo mada [8], paquetes usando el enfoque Bayesiano[HSROC [9], bamdit [10], meta4diag [11]], y paquetes para situaciones especiales[DiagMeta [12] para varias de un estudio, Metatrón [14], para el estudio de la norma de referencia]. En esta evaluación, nos centramos en mada, que proporciona el enfoque estándar y es más comparable a otro software.

Mada se ajusta al modelo bivariante y luego transforma el parámetros de modelos bivariantes para las estimaciones de HSROC. También calcula el área bajo la curva de resumen ROC como una sola medida de precisión, sin embargo, se muestra en la reciente investigación, esta medida puede conducir a resultados sesgados. Meta-regresión se puede realizar una variable a la vez. Mada también ofrece otros enfoques no muy utilizados, p. ej., agrupación univariada de DOR y riesgos proporcionales enfoque modelo.

**En general** ○ ● ●

La amplia variedad de paquetes en R le da al investigador más opciones al realizar el análisis, a la vez que trae consigo algunas dificultades en la toma de decisiones sobre qué paquete usar. Recomendamos mada a los usuarios básicos, y otros paquetes pueden ser interesantes para los estadísticos que están haciendo investigación metodológica.

La mayoría de los revisores clínicos confían en el software cuando implementan métodos



estadísticos, por lo tanto, la calidad metodológica de la revisión depende en gran medida del paquete que elijan. La Tabla 3.2 muestra el resumen de los resultados de la evaluación. R, Stata y SAS que son los paquetes recomendados por uso de la metodología para DTA.

| Criterio   | STATA | SAS | R |
|--|-------|-----|---|
| Accesibilidad  | ●     | ●   | ● |
| Usabilidad   | ●     | ●   | ● |
| Funciones para el meta-análisis y solidez metodológica | ●     | ●   | ● |
| En general   | ●     | ●   | ● |

Tabla 3.2: Evaluación de paquetes para el meta-análisis de la precisión diagnóstica

### 3.3. Procedimiento NLMIXED

#### 3.3.1. PROC NLMIXED Comparado con otros procedimientos y macros de SAS

Los modelos ajustados por PROC NLMIXED pueden ser vistos como generalizaciones de los modelos de coeficientes aleatorios ajustados por el procedimiento MIXED. Esta generalización permite que los coeficientes aleatorios entren en el modelo no linealmente, mientras que en PROC MIXED entran linealmente. Con PROC MIXED puede realizar tanto la estimación de máxima verosimilitud como la de máxima verosimilitud restringida (REML), mientras que PROC NLMIXED implementa sólo la de máxima verosimilitud. Esto se debe a que la analogía con el método REML en PROC NLMIXED implicaría una integral de alta dimensión sobre todos los parámetros de efectos fijos, y esta integral normalmente no está disponible en forma cerrada. Finalmente, PROC MIXED asume que los datos se distribuyen normalmente, mientras que PROC NLMIXED le permite analizar datos que son normales, binomiales o de Poisson o que tienen alguna probabilidad. programable con sentencias SAS.

PROC NLMIXED no implementa las mismas técnicas de estimación disponibles con la macro NLINMIX o el método de estimación por defecto del procedimiento GLIMMIX. Éstos se basan en los métodos de estimación de: [56] ; [51]; [52], y encajan iterativamente

en un conjunto de ecuaciones de estimación generalizadas. En contraste, PROC NLMIXED maximiza directamente una probabilidad integrada aproximada. Esta observación se aplica también a las macros SAS/IML MIXNLIN ( [53] ) y NLMEM ( [54] ).

El procedimiento GLIMMIX también se ajusta a modelos mixtos para datos no normales con no linealidad en la función de media condicional. A diferencia del procedimiento NLMIXED, PROC GLIMMIX asume que el modelo contiene un predictor lineal que vincula las covariables a la media condicional de la respuesta. El procedimiento NLMIXED está diseñado para manejar funciones de media condicional general, tanto si contienen una componente lineal como si no. Como se mencionó anteriormente, el procedimiento GLIMMIX por defecto estima los parámetros en modelos lineales mixtos generalizados mediante técnicas de pseudo-verosimilitud, mientras que PROC NLMIXED realiza por defecto la estimación de la máxima verosimilitud por cuadratura adaptativa de Gauss-Hermite. Este método de estimación también está disponible con el procedimiento GLIMMIX (METHOD=QUAD en la sentencia PROC GLIMMIX).

PROC NLMIXED está estrechamente vinculado al procedimiento de PNL en el software SAS/OR. PROC NLMIXED utiliza un subconjunto del código de optimización subyacente a PROC NLP y tiene muchas de las mismas opciones basadas en la optimización. Además, la funcionalidad de la sentencia de programación utilizada por PROC NLMIXED es la misma que la utilizada por PROC NLP y el procedimiento MODELO en el software SAS/ETS.

### 3.3.2. Algoritmos de optimización

Existen varias técnicas de optimización disponibles en PROC NLMIXED. Puede elegir un optimizador particular con la opción TECH= en la sentencia PROC NLMIXED.

No existe ningún algoritmo para optimizar las funciones no lineales generales, para que siempre encuentre el óptimo global en un problema de minimización no lineal general, y en un tiempo razonable. Dado que ninguna técnica de optimización es invariablemente superior a otras, PROC NLMIXED proporciona una variedad de técnicas de optimización que funcionan bien en diversas circunstancias. Sin embargo, usted puede idear problemas para los cuales ninguna de las técnicas de PROC NLMIXED puede encontrar la solu-

| Algoritmo                                       | Técnica |
|---|---------|
| Método de la región de confianza                | TRUREG  |
| Método Newton-Raphson con búsqueda por línea    | NEWRAP  |
| Método Newton-Raphson con reborde               | NRRIDG  |
| Métodos quasi-Newton (DBFGS, DDFP, BFGS, DFP)   | QUANEW  |
| Método double-dogleg (DBFGS, DDFP)              | DBLDOG  |
| Métodos de gradiente conjugado (PB, FR, PR, CD) | CONGRA  |
| Método Nelder-Mead simplex                      | NMSIMP  |

Tabla 3.3: Técnicas de optimización

ción correcta. Además, la optimización no lineal puede ser computacionalmente costosa en términos de tiempo y memoria, por lo que debe tener cuidado al comparar un algoritmo con un problema.

Todas las técnicas de optimización en PROC NLMIXED utilizan memoria  $O(n^2)$ . Excepto los métodos de gradiente conjugado, que utilizan sólo  $O(n)$  de memoria y están diseñados para optimizar problemas con muchos parámetros. Son iterativas, y requieren el cómputo repetido de lo siguiente:

1. El valor de la función (criterio de optimización).
2. El vector de gradiente (derivados parciales de primer orden).
3. Para algunas técnicas, la matriz Hessiana (aproximada) (derivados parciales de segundo orden).

Sin embargo, dado que cada uno de los optimizadores requiere diferentes derivadas, se pueden obtener algunas eficiencias computacionales. La siguiente Tabla 3.4 muestra, para cada técnica de optimización, qué derivadas son necesarias (FOD: derivados de primer orden; SOD: derivados de segundo-orden).

Cada método de optimización emplea uno o más criterios de convergencia que determinan cuándo ha convergido.

Se considera que un algoritmo ha convergido cuando se cumple cualquiera de los criterios de convergencia. Por ejemplo, bajo los ajustes por defecto, el algoritmo QUANEW convergerá si  $ABSGCONV < 1E - 5$ ;  $FCONV < 10^{-FDIGITS}$ ; o  $GCONV < 1E - 8$ .

| Algoritmo | 1era. Der. | 2da. Der. |
|-----------|------------|-----------|
| TRUREG    | ✓          | ✓         |
| NEWRAP    | ✓          | ✓         |
| NRRIDG    | ✓          | ✓         |
| QUANEW    | ✓          | -         |
| DBLDOG    | ✓          | -         |
| CONGRA    | ✓          | -         |
| NMSIMP    | -          | -         |

Tabla 3.4: Derivadas por Técnicas de optimización

### Selección de un algoritmo de optimización

Los factores que entran en la elección de una técnica de optimización para un problema en particular, son complejos y pueden implicar ensayo y error.

Para muchos problemas de optimización, calcular el gradiente toma más tiempo de computadora que calcular el valor de la función, y calcular el Hessiano a veces toma mucho más tiempo de computadora y la memoria que calcular el gradiente, especialmente cuando hay muchas variables de decisión. Desafortunadamente, las técnicas de optimización que no utilizan algún tipo de aproximación hessiana suelen requerir muchas más iteraciones que las técnicas que utilizan una matriz hessiana, y como resultado el tiempo total de ejecución de estas técnicas es a menudo mayor. Las técnicas que no usan el Hessiano también tienden a ser menos confiables. Por ejemplo, pueden más fácilmente terminar en puntos estacionarios en lugar de puntos óptimos globales.

A continuación se presentan algunas observaciones generales sobre las diferentes técnicas de optimización:

- Los métodos TRUREG (región de confianza), NEWRAP (Newton-Raphson con búsqueda por línea) y NRRRIDG (Newton-Raphson con reborde) son mejores para problemas pequeños donde la matriz Hessiana no es costosa de calcular. A veces el algoritmo NRRRIDG puede ser más rápido que el algoritmo TRUREG, pero TRUREG puede ser más estable. El algoritmo NRRRIDG requiere sólo una matriz con  $n(n+1)/2$  palabras dobles; TRUREG y NEWRAP requieren dos de estas matrices.

- Las primeras derivadas de los métodos QUANEW y DBLDOG son los mejores para problemas de tamaño-medio en los que la función objetivo y el gradiente son mucho más rápidos de evaluar que el Hessiano. Los algoritmos QUANEW y DBLDOG, en general, requieren más iteraciones que TRUREG, NRRRIDG y NEWRAP, pero cada iteración puede ser mucho más rápida. Los algoritmos QUANEW y DBLDOG requieren sólo el gradiente para actualizar un Hessiano aproximado, y requieren un poco menos de memoria que TRUREG o NEWRAP (esencialmente una matriz con  $n(n + 1)/2$  palabras dobles). QUANEW es el método de optimización por defecto.
  
- La primera derivada del método CONGRA es el mejor para problemas grandes donde la función objetivo y el gradiente pueden ser calculados mucho más rápido que el Hessiano y donde se requiere demasiada memoria para almacenar el Hessiano (aproximado). El algoritmo CONGRA, en general, requiere más iteraciones que QUANEW o DBLDOG, pero puede que cada iteración pueda ser más rápida. Dado que CONGRA sólo requiere un factor de memoria  $n$  double-word (es una unidad de datos que es dos veces el tamaño de una palabra), muchas aplicaciones grandes de PROC NLMIXED sólo pueden ser resueltas por CONGRA.
  
- El método NMSIMP sin-derivados es mejor para problemas pequeños en los que las derivadas no son continuas o son muy difíciles de calcular.

### Descripciones de algoritmos

#### 1. Optimización de la región de confianza (TRUREG)

El método de la región de confianza utiliza el gradiente  $g(\theta^{(k)})$  y la matriz Hessiana  $H(\theta^{(k)})$ ; por lo tanto, requiere que la función objetivo  $f(\theta)$  tenga derivadas continuas de primer y segundo orden dentro de la región factible.

El método de región de confianza optimiza iterativamente una aproximación cuadrática a la función de objetivo no lineal dentro de una región de confianza hiperelíptica con un radio  $\Delta$  que limita el tamaño del paso correspondiente a la calidad de la aproximación cuadrática. El método de la región de confianza se implementa usando: [63]; [64]; [65].

El método de región de confianza funciona bien para problemas pequeños y medianos, y no necesita muchas llamadas de función, gradiente y Hessiano. Sin embargo, si el cálculo de la matriz Hessiana es costoso, uno de los algoritmos de gradiente (dual) cuasi-Newton o conjugado podría ser más eficiente.

## 2. Optimización de Newton-Raphson con búsqueda de línea (NEWRAP)

La técnica NEWRAP utiliza el gradiente  $g(\theta^{(k)})$  y la matriz hessiana  $H(\theta^{(k)})$ ; por lo tanto, requiere que la función objetivo tenga derivados continuos de primer y segundo orden dentro de la región viable. Si los derivados de segundo orden se calculan de manera eficiente y precisa, el método NEWRAP puede funcionar bien para problemas de tamaño mediano a grande, y no necesita muchas llamadas de función, gradiente y Hessiano.

Este algoritmo utiliza un paso de Newton puro cuando el Hessiano es positivo definitivo y cuando el paso de Newton reduce el valor de la función objetivo con éxito. De lo contrario, una combinación de búsqueda de surcos y líneas se realiza para calcular los pasos exitosos. Si el Hessiano no es positivo definitivo, un múltiplo de la matriz de identidad se añade a la matriz Hessiana para hacerla definitiva (Eskow y Schnabel 1991).

## 3. Optimización de Newton-Raphson Ridge (NRRRIDG) La técnica NRRRIDG utiliza el gradiente $g(\theta^{(k)})$ y la matriz hessiana $H(\theta^{(k)})$ ; por lo tanto, requiere que la función objetivo tenga derivados continuos de primer y segundo orden dentro de la región factible.

Este algoritmo utiliza un paso de Newton puro cuando el Hessiano es positivo definitivo y cuando el paso de Newton reduce el valor de la función objetivo con éxito. Si no se cumple al menos una de estas dos condiciones, se añade un múltiplo de la matriz de identidad a la matriz de Hessiana.

El método NRRRIDG funciona bien para problemas pequeños y medianos, y no re-

quiere muchas llamadas de función, gradiente y Hessiano. Sin embargo, si el cálculo de la matriz Hessiana es costosa, uno de los algoritmos de gradiente (dual) cuasi-Newton o conjugado podría ser más eficiente.

Dado que la técnica NRRRIDG utiliza una descomposición ortogonal del Hessiano aproximado, cada iteración de NRRRIDG puede ser más lenta que la de la técnica NEWRAP, que funciona con la descomposición de Cholesky. Sin embargo, por lo general, NRRRIDG requiere menos iteraciones que NEWRAP.

4. Optimización Quasi-Newton (QUANEW) El método (dual) cuasi-Newton utiliza el gradiente  $g(\theta^{(k)})$ , y no necesita calcular derivadas de segundo-orden, ya que son aproximados. Funciona bien para problemas de optimización medianos a moderadamente grandes donde la función objetivo y el gradiente son mucho más rápidos de calcular que el Hessiano; pero, en general, requiere más iteraciones que las técnicas de TRUREG, NEWRAP, y NRRRIDG, que calculan derivados de segundo orden. QUANEW es el algoritmo de optimización por defecto porque proporciona un equilibrio adecuado entre la velocidad y la estabilidad requeridas para la mayoría de las aplicaciones de modelos mixtos no lineales.

La técnica QUANEW es una de las siguientes, dependiendo del valor de la opción UPDATE:

- El algoritmo cuasi-Newton original, que actualiza una aproximación del algoritmo inverso de Hessiano.
- El algoritmo dual cuasi-Newton, que actualiza el factor Cholesky del Hessiano aproximado (por defecto).

Puede especificar cuatro fórmulas de actualización con la opción UPDATE.

- DBFGS realiza la actualización dual de Broyden, Fletcher, Goldfarb y Shanno (BFGS) del factor Cholesky de la matriz Hessiana. Este es el valor por defecto.
- El DDFP realiza la actualización dual de Davidon, Fletcher y Powell (DFP) del factor Cholesky de la Matriz de Hessiana.

- BFGS realiza la actualización original BFGS de la matriz hessiana inversa.
- DFP realiza la actualización DFP original de la matriz hessiana inversa.

En cada iteración, se realiza una búsqueda de línea a lo largo de la dirección de búsqueda para encontrar un óptimo aproximado. El método de búsqueda por línea por defecto utiliza interpolación cuadrática y extrapolación cúbica para obtener un tamaño de paso  $\alpha$  satisfaciendo las condiciones de Goldstein. Una de las condiciones de Goldstein puede ser violada si la región factible define un límite superior del tamaño del paso. La violación de la condición Goldstein del lado izquierdo puede afectar el resultado positivo. En ese caso, o bien se omite la actualización o bien se omiten las iteraciones. reiniciado con una matriz de identidad, resultando en la dirección de búsqueda de descenso o ascenso más pronunciada. Puede especificar algoritmos de búsqueda por línea que no sean los predeterminados con la opción LINESEARCH.

#### 5. Optimización Double-Dogleg (DBLDOG)

El método de optimización Double-Dogleg combina las ideas de los métodos cuasi-Newton y de la región de confianza. En cada iteración, el algoritmo de Double-Dogleg calcula el paso  $s^{(k)}$  como la combinación lineal de la dirección de búsqueda de descenso o ascenso más pronunciada  $s_1^{(k)}$  y una dirección de búsqueda cuasi-Newton  $s_2^{(k)}$ :

$$s^{(k)} = \alpha_1 s_1^{(k)} + \alpha_2 s_2^{(k)}$$

Se solicita que el paso se mantenga dentro de un radio de región de confianza preespecificado; véase [66]. Así, la subrutina DBLDOG utiliza la actualización dual cuasi-Newton pero no realiza una búsqueda por línea. Usted puede especificar dos fórmulas de actualización con la opción UPDATE:

- DBFGS realiza la actualización dual de Broyden, Fletcher, Goldfarb y Shanno del factor Cholesky de la matriz de Hessiana. Este es el valor por defecto.
- El DDFP realiza la actualización dual de Davidon, Fletcher y Powell del factor Cholesky de la matriz Hessiana.



La técnica de optimización de double-dogleg funciona bien para problemas de optimización de tamaño medio a moderado, donde la función objetivo y el gradiente son mucho más rápidos de calcular que el Hessiano. La implementación se basa en [67] y [64], pero se amplía para tratar con la frontera y las restricciones lineales. La técnica DBLDOG generalmente requiere más iteraciones que la técnica TRUREG, NEWRAP, y NRRRIDG, que requieren derivados de segundo orden; sin embargo, cada una de las iteraciones de DBLDOG es computacionalmente barato. Además, la técnica DBLDOG sólo requiere llamadas de gradiente para la actualización, del factor Cholesky de un Hessiano aproximado.

#### 6. Optimización del gradiente conjugado (CONGRA)

Los derivados de segundo orden no son requeridos por el algoritmo CONGRA y ni siquiera son aproximados. El algoritmo CONGRA puede ser costoso en llamadas de función y gradiente, pero sólo requiere  $O(n)$  memoria para optimización sin restricciones. En general, se requieren muchas iteraciones para obtener una solución precisa, pero cada una de las iteraciones de CONGRA es computacionalmente barato. Puede especificar cuatro fórmulas de actualización diferentes para generar las direcciones conjugadas usando la opción UPDATE:

- PB realiza el método de actualización de reinicio automático de [68] y [69]. Este es el valor por defecto.
- FR realiza la actualización de Fletcher-Reeves [66].
- PR realiza la actualización de Polak-Ribiere [66].
- CD realiza una actualización conjugado-descendiente de [66].

El valor predeterminado, UPDATE=PB, se comportó mejor en la mayoría de los ejemplos de pruebas. Se aconseja evitar la opción UPDATE=CD, que se comportó peor en la mayoría de los ejemplos de pruebas.

La subrutina CONGRA debe ser usada para problemas de optimización con  $n$  grande. CONGRA sólo requiere  $O(n)$  bytes de memoria de trabajo, mientras que todos los otros los métodos de optimización requieren  $O(n^2)$  bytes de orden de la memoria

de trabajo. Durante  $n$  iteraciones sucesivas, ininterrumpido por reinicios o cambios en el equipo de trabajo, el algoritmo de gradiente conjugado calcula un ciclo de  $n$  direcciones de búsqueda conjugadas. En cada iteración, se realiza una búsqueda de línea a lo largo de la dirección de búsqueda para encontrar un óptimo aproximado de la función objetivo. El método de búsqueda por línea por defecto utiliza el método cuadrático interpolación y extrapolación cúbica para obtener un tamaño de paso  $\alpha$  que satisfaga las condiciones de Goldstein. Una de las condiciones de Goldstein puede ser violadas si la región factible define un límite superior para el tamaño del paso. Otros algoritmos de búsqueda por línea se pueden especificar con la opción LINESEARCH.

#### 7. Optimización de Nelder-Mead Simplex (NMSIMP)

El método Nelder-Mead simplex no utiliza ningún derivado y no asume que la función objetivo tiene derivados continuos. La función de objetivo debe ser continua. Esta técnica es bastante cara en el número de llamadas de función, y podría ser incapaz de generar resultados precisos para  $n \gg 40$ .

El algoritmo original de Nelder-Mead simplex se implementa y se extiende a las restricciones de límites. Este no calcula el objetivo de los puntos no factibles, pero cambia la forma del simplex adaptando a la no linealidad de la función objetivo, lo que contribuye a una mayor velocidad de convergencia. Se utiliza un criterio de terminación especial.

### 3.3.3. Problemas de Computacionales

#### 1. Errores y desbordamientos de punto flotante

La optimización numérica de una función integrada numéricamente es una tarea difícil, y el cálculo de sus derivadas pueden dar lugar a excepciones y desbordamientos aritméticos. Una causa típica de estos problemas son parámetros con escalas muy variadas. Si la escala de sus parámetros varía más que unos pocos órdenes de magnitud, la estabilidad numérica del problema de optimización puede ser seriamente reducida y puede resultar en dificultades computacionales. Un remedio simple es reescalar cada parámetro para que su valor final tenga una magnitud cercana a 1.

Si el reajuste de parámetros no ayuda, considere:

- Especificar la opción ITDETAILS en la sentencia PROC NLMIXED para obtener información más detallada sobre cuándo y dónde está ocurriendo el problema.
- Proporcionar diferentes valores iniciales o intente una búsqueda en cuadrícula de valores.
- Utilizar las restricciones de los límites para evitar la región en la que pueden producirse desbordamientos.
- Suprimir las observaciones o los temas alejados de los datos de entrada, si es razonable.
- Cambiar el algoritmo (especificado en las instrucciones de programación) que calcula la función objetivo.

## 2. Tiempos de funcionamiento largos

PROC NLMIXED puede tardar mucho tiempo en ejecutarse para problemas que involucren modelos complejos, muchos parámetros, o grandes conjuntos de datos de entrada. Aunque las técnicas de optimización utilizadas por PROC NLMIXED son algunas de las mejores disponibles, no se garantiza que converjan rápidamente para todos los problemas. Modelos mal colocados o mal especificados pueden hacer que los algoritmos usen cálculos más extensos diseñados para lograr la convergencia, y esto puede resultar en tiempos de ejecución más largos. Así que primero hay que asegurarse que el modelo este especificado correctamente, que los parámetros se escalan para que sean del mismo orden de magnitud, y que los datos coincidan razonablemente con el modelo que se está contemplando.

Si está utilizando el algoritmo de cuadratura gaussiana adaptativa por defecto y no se está imprimiendo ningún historial de iteración, entonces PROC NLMIXED podría estar empantanado tratando de determinar el número de puntos de la cuadratura en la primera serie de valores iniciales. Especificar la opción QPOINTS evitará esta etapa y procederá directamente; sin embargo, tenga en cuenta que la aproximación de probabilidad podría no ser exacta si hay muy pocos puntos de cuadratura.

PROC NLMIXED también puede tener dificultades para determinar el número de puntos de cuadratura si los valores iniciales están lejos de los valores óptimos. Para obtener valores iniciales más precisos para los parámetros del modelo, un método sencillo es ajustar un modelo sin sentencia RANDOM. A continuación, puede utilizar estas estimaciones como inicio aunque todavía tendrá que especificar el valor para la distribución de efectos aleatorios para modelos normal-normal, otra estrategia es utilizar METHOD=FIRO. Si pueden obtener estimaciones utilizando este método aproximado entonces se pueden usar como valores iniciales para obtener aproximaciones de verosimilitud más precisas.

Si está ejecutando PROC NLMIXED varias veces, es probable que desee incluir una sentencia como la siguiente en su programa:

```
ods output ParameterEstimates=pe;
```

Esta sentencia crea un conjunto de datos SAS llamado PE al finalizar la ejecución.

En su próxima invocación de PROC NLMIXED, puede especificar

```
parms / data=pe;
```

para leer en las estimaciones anteriores como valores iniciales.

Para acelerar los cálculos generales, debe verificar sus instrucciones de programación para minimizar el número de operaciones en coma flotante. El uso de variables auxiliares y el factoraje de expresiones amigables pueden ser cambios útiles a este respecto.

### 3. Problemas de evaluación del código para la función objetivo

El punto de partida  $\theta^{(0)}$  debe ser un punto en el que puedan evaluarse las instrucciones de programación. Sin embargo, durante la optimización, el optimizador puede iterar a un punto  $\theta^{(k)}$ , donde la función objetivo o sus derivados no pueden ser evaluados. En algunos casos, la especificación de los límites de los parámetros puede evitar tales situaciones. En muchos otros casos, se puede indicar que el punto  $\theta^{(0)}$  es un punto malo simplemente devolviendo un valor de gran valor para la función objetivo. En estos casos, el algoritmo de optimización reduce la longitud del paso y se acerca más al punto que ha sido evaluado con éxito en la iteración anterior.

### 4. No hay convergencia

Hay una serie de cosas que se pueden intentar si el optimizador no converge.

- Cambiar los valores iniciales utilizando una especificación de búsqueda en cuadrícula para obtener un conjunto de buenos valores iniciales factibles.
- Cambiar o modificar la técnica de actualización o el algoritmo de búsqueda por línea.

Este método se aplica sólo a TECH=QUANEW y TECH=CONGRA. Por ejemplo, si utiliza la función y el algoritmo de búsqueda de línea predeterminado.

- Cambiar la técnica de optimización. Por ejemplo, si utiliza la opción por defecto, TECH=QUANEW, puede probar una de las siguientes derivadas si su problema es pequeño o el método del gradiente conjugado si es grande.
- Ajustar los derivados de diferencias finitas. Las derivadas de diferencias son especificadas con la opción FD o FDHESSIAN pueden no ser lo suficientemente precisas para satisfacer los fuertes criterios de terminación de gradiente. Es posible que tenga que especificar más costosas fórmulas de diferencia central. Los intervalos de diferencia finita pueden ser demasiado pequeños o demasiado

grandes, y los derivados de diferencias finitas podrían ser erróneos.

- Comprobar la entrada de datos y la especificación del programa.

### 5. Convergencia con el punto estacionario

El gradiente en un punto estacionario es el vector nulo, que siempre conduce a una dirección de búsqueda cero. Este punto satisface el criterio de terminación de primer orden. Las direcciones de búsqueda que se basan en el gradiente son cero, por lo que el algoritmo termina. Hay dos maneras de evitar esta situación:

- Utilizar la sentencia PARMs para especificar una cuadrícula de puntos iniciales factibles.
- Utilizar la opción OPTCHECK=r para evitar terminar en el punto estacionario.

Los signos de los valores propios de la matriz (reducida) Hessiana contienen la siguiente información relativa a un punto estacionario:

- Si todos los valores propios son positivos, la matriz Hessiana es positiva definida, y el punto es un mínimo.
- Si algunos de los valores propios son positivos y todos los valores propios restantes son cero, la matriz Hessiana es semidefinida positiva, y el punto es un punto mínimo o punto de silla.
- Si todos los valores propios son negativos, la matriz Hessiana es negativa definida, y el punto es un máximo.
- Si algunos de los valores propios son negativos y todos los valores propios restantes son cero, la matriz Hessiana es semidefinida negativa, y el punto es un punto máximo o punto de silla.

- Si todos los valores propios son cero, el punto puede ser un mínimo, máximo o de silla.

#### 6. Precisión de la solución

En algunas aplicaciones, PROC NLMIXED puede dar lugar a valores de parámetros que no son lo suficientemente precisos. Normalmente, esto significa que el procedimiento terminó en un punto demasiado lejos del punto óptimo. Cualquier punto dentro de esta región puede ser aceptado para terminar el proceso de optimización. Los valores propuestos de los criterios de terminación se fijan en satisfacer un compromiso razonable entre el esfuerzo computacional (tiempo de computadora) y la precisión de estimaciones calculadas para las aplicaciones más comunes. Sin embargo, hay una serie de circunstancias en que los valores propuestos de los criterios de terminación especifican una región que es demasiado grande o demasiado pequeña.

Si la región de terminación es demasiado grande, entonces puede contener puntos con baja precisión. En tales casos, usted debe determinar qué criterio de terminación detuvo el proceso de optimización. En muchas aplicaciones, puede obtener una solución con mayor precisión simplemente utilizando los cálculos de parámetros antiguos como valores iniciales en una ejecución posterior en la que se especifica un valor menor para el criterio de terminación que se cumplió en la ejecución anterior.

#### 7. Precisión de la solución

En algunas aplicaciones, PROC NLMIXED puede dar lugar a valores de parámetros que no son lo suficientemente precisos. Normalmente, esto significa que el procedimiento terminó en un punto demasiado lejos del punto óptimo. Los criterios de terminación, definir el tamaño de la región de terminación alrededor del punto óptimo. Cualquier punto dentro de esta región puede ser aceptado para terminar el proceso de optimización. Los valores propuestos de los criterios de terminación se fijan en satisfacer un compromiso razonable entre el esfuerzo computacional (tiempo de computadora) y la precisión de estimaciones calculadas para las aplicaciones más comunes. Sin embargo, hay una serie de circunstancias en que los valores propuestos de los criterios de terminación especifican una región que es demasiado grande o demasiado pequeña.

Si la región de terminación es demasiado grande, entonces puede contener puntos con baja precisión. En tales casos, debe determinar qué criterio de terminación detuvo el proceso de optimización. En muchas aplicaciones, puede obtener una solución con mayor precisión simplemente utilizando los cálculos de parámetros antiguos como valores iniciales, en una ejecución posterior en la que se especifica un valor menor para el criterio de terminación que se cumplió en la ejecución anterior.

### 3.3.4. Matriz de Covarianza

La matriz de covarianza estimada de las estimaciones de los parámetros se calcula como la matriz hessiana inversa, y para problemas no restringidos debería ser positivo definitivo. Si las estimaciones de los parámetros finales están sujetas a  $n_{act} > 0$  limitaciones de desigualdad lineal activa, las fórmulas de las matrices de covarianza se modifican de forma similar a [73] y [74] y adicionalmente generalizado para aplicaciones con matrices singulares.

Hay varios pasos disponibles que le permiten ajustar los cálculos de rango de la matriz de covarianza.

1. Puede utilizar las opciones ASINGULAR, MSINGULAR, y VSINGULAR para establecer tres criterios para la inversión de la matriz hessiana H. El criterio de singularidad utilizado para la inversión es:

$$|d_{j,j}| \leq \max(ASING, VSING * |H_{j,j}|, MISING * \max(|H_{1,1}|, \dots, |H_{n,n}|))$$

donde  $d_{j,j}$  es el pivote diagonal de la matriz H, y ASING, VSING, y MSING son los especificados de las opciones ASINGULAR, VSINGULAR, y MSINGULAR, respectivamente. Los valores por defecto son los siguientes:

- ASING: la raíz cuadrada del valor positivo de doble precisión más pequeño
- MSING: 1E-12 si no se especifica la opción SINGHESS y  $\max(10\epsilon, 1E-4xSINGHESS)$  en caso contrario, dónde  $\epsilon$  es la precisión de la máquina.
- VSING: 1E-8 si no se especifica la opción SINGHESS y el valor de SINGHESS de lo contrario



Nótese que, en muchos casos, se descompone una matriz normalizada  $D^{-1}AD^{-1}$ , y los criterios de singularidad se modifican en consecuencia.

2. Si la matriz H es singular en el primer paso, se calcula un inverso generalizado. Dependiendo de la opción G4, se calcula la inversa generalizada que satisface las cuatro condiciones de Moore-Penrose (un g4-inverso) o una inversa generalizada que satisfaga sólo dos condiciones de Moore-Penrose es calculada (a g2-inverso, [75]). Si el número de parámetros n de la aplicación es menor que o igual a G4=i, se calcula un g4-inverso; de lo contrario, sólo se calcula un g2-inverso. El g4-inverso es calculado por la (muy caro pero numéricamente estable) descomposición en valores propios, y el g2-inverso es calculado por la transformación de Gauss. El g4-inverso se calcula utilizando la descomposición del valor propio  $A = Z\Lambda Z'$ , donde Z es la matriz ortogonal de los vectores propios y es la matriz diagonal de los valores propios,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . El g4-inverso de H está ajustado a

$$A^- = Z\Lambda^-Z'$$

donde la matriz diagonal  $\Lambda^- = \text{diag}(\lambda_1^-, \dots, \lambda_n^-)$  se define mediante la opción COVSING:

$$\lambda_i^- = \begin{cases} 1/\lambda_i & \text{si } |\lambda_i| > \text{COVSING} \\ 0 & \text{si } |\lambda_i| \leq \text{COVSING} \end{cases}$$

Si no especifica la opción COVSING=, los valores propios más pequeños se ponen a cero, donde nr es el número de deficiencias de rango encontradas en el primer paso.

Para las técnicas de optimización que no utilizan derivadas de segundo orden, la matriz de covarianza se calcula utilizando aproximaciones de diferencias finitas de las derivadas.

### 3.3.5. Predicción

El modelo mixto no lineal es una herramienta útil para la predicción estadística. Suponiendo que se deba hacer una predicción con respecto al i-ésimo sujeto, supongamos que  $f(\theta, u_i)$  es una función diferenciable que predice alguna cantidad de interés. Recordar

que  $\theta$  denota el vector de parámetros desconocidos y  $u_i$  denota el vector de efectos aleatorios para el  $i$ -ésimo individuo. Una predicción de punto natural es  $f(\hat{\theta}, \hat{u}_i)$ , donde está la estimación de máxima verosimilitud de  $\theta$  y  $\hat{u}_i$  es la estimación empírica de Bayes de la  $u_i$ .

Una matriz de predicción de varianza aproximada para  $(\hat{\theta}, \hat{u}_i)$  es

$$P = \begin{bmatrix} \hat{H}^{-1} & \hat{H}^{-1} \left( \frac{\partial \hat{u}_i}{\partial \theta} \right)' \\ \left( \frac{\partial \hat{u}_i}{\partial \theta} \right) \hat{H}^{-1} & \hat{\Gamma}^{-1} + \left( \frac{\partial \hat{u}_i}{\partial \theta} \right) \hat{H}^{-1} \left( \frac{\partial \hat{u}_i}{\partial \theta} \right)' \end{bmatrix}$$

donde  $\hat{H}$  es la matriz hessiana aproximada de la optimización para  $\hat{\theta}$ ,  $\hat{\Gamma}$  es la matriz hessiana aproximada de la optimización para  $\hat{u}_i$ , y  $(\partial \hat{u}_i / \partial \theta)$  es la derivada de  $\hat{u}_i$  con respecto a  $\theta$ , evaluado en  $(\hat{\theta}, \hat{u}_i)$ . La matriz de varianza aproximada para  $\hat{\theta}$  que es el estándar, y  $\hat{u}_i$  es una aproximación al error de predicción de la media cuadrática condicional descrita por [76].

La varianza de predicción para una función escalar general  $f(\theta, u_i)$  se define como la diferencia cuadrada esperada  $E[f(\hat{\theta}, \hat{u}_i) - f(\theta, u_i)]^2$ . PROC NLMIXED calcula una aproximación a ella de la siguiente manera. La derivada de  $f(\theta, u_i)$  se calcula con respecto a cada elemento de  $(\theta, u_i)$  y se evalúa  $(\hat{\theta}, \hat{u}_i)$ . Si  $a_i$  es el vector resultante, entonces la varianza de predicción aproximada es  $a_i' P a_i$ . Esta aproximación se conoce como el método delta ([59]; [60]).

### 3.3.6. Visualización de Salidas

#### 1. Especificaciones

El procedimiento NLMIXED muestra primero la tabla “Especificaciones”, que contiene información básica sobre el modelo mixto no lineal que ha especificado. Incluye las principales variables y métodos de estimación.

#### 2. Dimensiones

La tabla “Dimensiones” enumera los recuentos de cantidades importantes en su modelo mixto no lineal, incluyendo la tabla número de observaciones, sujetos, parámetros y puntos de cuadratura.

### 3. Parámetros

La tabla “ Parámetros ” muestra la información proporcionada con la sentencia PARMs y el valor de la función negativa de verosimilitud logarítmica evaluada en los valores iniciales.

### 4. Gradiente de inicio y Hessiano

La opción START de la sentencia PROC NLMIXED muestra el gradiente del logaritmos de verosimilitud negativa en los valores iniciales de los parámetros. Si también especifica la opción HESS, entonces también se muestra el Hessiano.

### 5. Iteraciones

El historial de iteración consiste en una línea de salida para cada iteración en el proceso de optimización. La iteración se muestra de forma predeterminada porque es importante que compruebe si hay posibles problemas de convergencia. El historial de iteración predeterminado incluye las siguientes variables y su definición:

- Iter, el número de iteración
- Calls, el número de llamadas de función.
- NegLogLike, el valor de la función objetivo.
- Diff, la diferencia entre los valores de las funciones adyacentes
- MaxGrad, el máximo de las componentes (excepto NMSIMP) de gradiente absoluto (proyectado).
- Slope, la pendiente  $g'$ s de la dirección de búsqueda  $s$  en el parámetro actual iterate  $\theta^{(k)}$  (sólo QUANEW).
- Rho, la relación entre los valores alcanzados y predichos de Diff (sólo NRRRIDG).
- Radius, el radio de la región de confianza (sólo TRUREG).
- StdDev, la desviación estándar de los valores del simplex (sólo NMSIMP).
- Delta, la longitud del vértice del simplex (sólo NMSIMP).
- Size, el tamaño del simplex (sólo NMSIMP).

Para el método QUANEW, el valor de Slope debe ser significativamente negativo. De lo contrario, la función de búsqueda por línea tiene dificultad para reducir suficientemente el valor de la función. Si se encuentra esta dificultad, un asterisco (\*)

aparece después del número de iteración. Si hay una tilde ( $\sim$ ) después del número de iteración, la actualización del BFGS es y se producen valores muy altos de la función Lagrange. Una barra invertida ( $\backslash$ ) después de la iteración indica que se utiliza la corrección de Powell para la actualización de BFGS.

En el caso de los métodos que utilizan segundas derivadas, un asterisco (\*) después del número de iteración significa que la aproximación hessiana era singular, regularización (ridged) y con un valor positivo.

Para el método NMSIMP, sólo se visualiza una línea para varias iteraciones internas. Esta técnica se salta las salidas para algunas iteraciones porque algunas de las pruebas de terminación (StdDev y Size) son bastante lentas en comparación con las operaciones simplex, y se realizan sólo cada cinco operaciones simplex.

La opción ITDETAILS en la sentencia PROC NLMIXED proporciona un historial de iteración más detallada. Además de enumerar los valores actuales de los parámetros y sus gradientes, la opción ITDETAILS proporciona los siguientes valores, además de la salida por defecto:

- Restart, el número de iteraciones de reinicia.
- Active, el número de restricciones activas.
- Lambda, el valor del multiplicador de Lagrange (sólo TRUREG y DBLDOG).
- Ridge, el valor de la ridge (sólo NRRRIDG).
- Alpha, el tamaño del paso de búsqueda por línea (sólo QUANEW).

Un apóstrofe (') que sigue al número de restricciones activas indica que al menos una de las restricciones activas fue liberada del conjunto activo debido a un multiplicador de Lagrange significativo.

## 6. Estado de Convergencia

La tabla “ Convergence Status ” contiene un mensaje de estado que describe el motivo de la finalización de la optimización. El nombre ODS de esta tabla es ConvergenceStatus, y puede consultar la variable numérica Status para comprobar si la

optimización se ha realizado correctamente. Esto es útil en el tratamiento de lotes o en el tratamiento de grupos BY, por ejemplo, en simulaciones. La convergencia exitosa se indica con Status=0.

### 7. Información de ajuste

La tabla “ Fitting Information ” enumera el valor final minimizado de -2 veces el logaritmo de verosimilitud, así como los criterios de información de Akaike (AIC) y Schwarz (BIC), y una versión corregida de muestra-finita de AIC (AICC). Los criterios se calculan de la siguiente manera:

$$\begin{aligned} AIC &= 2f(\hat{\theta}) + 2p \\ AICC &= 2f(\hat{\theta}) + 2pn/(n - p - 1) \\ BIC &= 2f(\hat{\theta}) + p \log(s) \end{aligned}$$

donde  $f()$  es el negativo de la función marginal de verosimilitud logarítmica,  $\hat{\theta}$  es el vector de las estimaciones de parámetros,  $p$  es el número de parámetros,  $n$  es el número de observaciones y  $s$  es el número de sujetos. Ver [77] y [78] para más detalles.

### 8. Estimaciones de parámetros

La tabla Estimaciones de parámetros “ Parameter Estimates ” enumera las estimaciones de los valores de los parámetros tras una convergencia satisfactoria de el problema de optimización o los valores finales de los parámetros en caso de no convergencia. Si el problema convergió, los errores estándar se calculan a partir de la matriz Hessiana final.

La relación de estimación con su error estándar produce un valor  $t$ , con grados de libertad aproximados como el número de sujetos menos el número de efectos aleatorios. Finalmente, se muestra el gradiente de la función del logaritmo de verosimilitud negativa para cada parámetro, y debe verificar que cada uno de ellos sea lo suficientemente pequeño para los parámetros no restringidos.

### 9. Matrices de Covarianza y Correlación

Siguiendo la teoría estándar de máxima verosimilitud (por ejemplo, [79] ), la matriz asintótica de varianza-covarianza de los cálculos del parámetro es igual a la inversa de la matriz hessiana. Puede visualizar esta matriz con la opción COV en la sentencia

PROC NLMIXED. La forma de correlación correspondiente está disponible con la opción CORR.

#### 10. Estimaciones adicionales

La tabla “ Additional Estimates ” muestra los resultados de todas las sentencias ESTIMATE que se especifiquen, con la opción mismas columnas que la tabla “ Estimación de parámetros ”. Las opciones ECOV y ECORR en el PROC NLMIXED producen tablas que muestran las matrices de covarianza y correlación aproximadas de variables de las estimaciones adicionales. Se calculan utilizando el método delta ( [59]; [60] ). La opción EDER produce una tabla que muestra las derivadas.

### 3.3.7. Nombre de Tablas ODS

PROC NLMIXED asigna un nombre a cada tabla que crea. Puede utilizar estos nombres para hacer referencia a la tabla cuando utilice el Sistema de entrega de salida (ODS) para seleccionar tablas y crear conjuntos de datos de salida. Estos nombres se enumeran en la Tabla 3.5

| Tabla (ODS)        | Descripción   | Opción     |
|--------------------|---|------------|
| AdicionalEstimates | Resultados de sentencias ESTIMATE                     | ESTIMATE   |
| Contrasts          | Resultados de las declaraciones CONTRAST              | CONTRAST   |
| ConvergenceStatus  | Estado de la convergencia                             | default    |
| CorrMatAddEst      | Matriz de correlación de las estimaciones adicionales | ECORR      |
| CorrMatParmEst     | Matriz de correlación de parámetros estimados         | CORR       |
| CovMatAddEst       | Matriz de covarianza de estimaciones adicionales      | ECOV       |
| CovMatParmEst      | Matriz de covarianza de parámetros estimados          | COV        |
| DerAddEst          | Derivados de estimaciones adicionales                 | EDER       |
| Dimensions         | Dimensiones del problema                              | default    |
| FitStatistics      | ajuste Estadístico                                    | default    |
| Hessian            | Segunda matriz derivada                               | HESS       |
| IterHistory        | Historial de iteración                                | default    |
| Parameters         | Parámetros iniciales                                  | default    |
| ParameterEstimates | Parámetros estimados                                  | default    |
| Specifications     | Especificaciones del modelo                           | default    |
| StartingHessian    | Inicio Matriz Hessiana                                | START HESS |
| StartingValues     | Valores iniciales y gradiente                         | START      |

Tabla 3.5: Tablas de ODS producidas por PROC NLMIXED

## 3.4. Aplicaciones del Procedimiento NLMIXED

### 3.4.1. Curvas de crecimiento no lineal con datos gaussianos

Como ejemplo introductorio, consideremos los datos de los naranjos de [94]. Estos datos consisten en siete mediciones de la circunferencia del tronco (en milímetros) en cada uno de los cinco naranjos. Puede introducir estos datos en un conjunto de datos SAS como se indica a continuación:

```
data tree;
input tree day y;
datalines;
1 118 30
1 484 58
1 664 87
... more lines ...
5 1582 177;
```

[56] y [57] proponen el siguiente modelo logístico mixto no lineal para estos datos:

$$y_{ij} = \frac{b_1 + u_{i1}}{1 + \exp[-(d_{ij} - b_2)/b_3]} + e_{ij} \quad (3.4.1)$$

Donde,  $y_{ij}$  representa la  $j$ -ésima medida en el  $i$ -ésimo árbol ( $i=1,\dots,5$ ;  $j=1,\dots,7$ ),  $d_{ij}$  es la medida correspondiente a los días,  $b_1$ ,  $b_2$ ,  $b_3$  son los parámetros de efectos fijos,  $u_{i1}$  son los parámetros de efectos aleatorios que se supone que son iid  $N(0, \sigma_u^2)$ , y  $e_{ij}$  son los errores residuales que se supone que son iid  $N(0, \sigma_e^2)$  independiente de la  $u_{i1}$ . Este modelo tiene una forma logística, y los parámetros de efectos aleatorios  $u_{i1}$  entran en el modelo linealmente. Las sentencias para la estimación de este modelo mixto no lineal son los siguientes:

```
proc nlmixed data=Tree;
parms b1=190 b2=700 b3=350 s2u=1000 s2e=60;
num = b1+u1;
ex = exp(-(day-b2)/b3);
den = 1 + ex;
model y ~ normal(num/den, s2e);
```

```
random u1 ~ normal(0,s2u) subject=tree;
run;
```

La sentencia PROC NLMIXED invoca el procedimiento e introduce el conjunto de datos “Tree”. La sentencia PARMS identifica los parámetros desconocidos y sus valores iniciales. Aquí hay tres parámetros de efectos fijos (b1,b2, b3) y dos componentes de varianza (s2u, s2e).

Las tres siguientes son instrucciones de programación de SAS que especifican el modelo logístico mixto. Se incluye una nueva variable u1 para identificar el efecto aleatorio. Estas declaraciones se evalúan para cada observación en el conjunto de datos cuando el procedimiento NLMIXED calcula la función del log. de verosimilitud y sus derivadas.

La sentencia MODEL define la variable dependiente y su distribución condicional dados los efectos aleatorios. Aquí se especifica una distribución condicional normal (gaussiana) con la media num/den y la varianza s2e.

La sentencia RANDOM define el efecto aleatorio único como u1, y especifica que sigue una distribución normal con media 0 y varianza s2u. El argumento SUBJECT= en la sentencia RANDOM define una variable que indica cuando el efecto aleatorio obtiene nuevas realizaciones; en este caso, cambia de acuerdo a los valores de la variable Tree. PROC NLMIXED asume que el conjunto de datos de entrada está agrupado de acuerdo con los niveles de la variable de Tree; es decir, todas las observaciones del mismo Tree ocurren secuencialmente en el conjunto de datos de entrada. El resultado de este análisis es el siguiente.

|  | <b>Specifications</b>        |
|--|------------------------------|
| <b>Data Set</b>                            | WORK.TREE                    |
| <b>Dependent Variable</b>                  | y                            |
| <b>Distribution for Dependent Variable</b> | Normal                       |
| <b>Random Effects</b>                      | u1                           |
| <b>Distribution for Random Effects</b>     | Normal                       |
| <b>Subject Variable</b>                    | tree                         |
| <b>Optimization Technique</b>              | Dual Quasi–Newton            |
| <b>Integration Method</b>                  | Adaptive Gaussian Quadrature |

Tabla 3.6: Especificaciones del modelo



La tabla “ Specifications ” contiene información básica sobre el modelo mixto no lineal que ha especificado (Tabla 3.6). Se incluye el conjunto de datos de entrada, las variables dependientes y sujetas, los efectos aleatorios, las distribuciones relevantes y el tipo de optimización. La tabla “ Dimensions ” enumera varios conteos relacionados con el modelo, incluyendo el número de observaciones, temas y parámetros (Tabla 3.7).

Estas cantidades son útiles para verificar que se ha especificado el conjunto de datos y el modelo correctamente. También se enumera el número de puntos de cuadratura que PROC NLMIXED ha seleccionado basándose en la evaluación del log de verosimilitud en los valores iniciales de los parámetros. Aquí sólo se necesita un punto de cuadratura porque los parámetros de efectos aleatorios  $u_1$  entran en el modelo de forma lineal. (La cuadratura de Gauss-Hermite con un solo punto de cuadratura resulta en la aproximación de Laplace del log De verosimilitud).

|                              | <b>Dimensions</b> |
|------------------------------|-------------------|
| <b>Observations Used</b>     | 35                |
| <b>Observations Not Used</b> | 0                 |
| <b>Total Observations</b>    | 35                |
| <b>Subjects</b>              | 5                 |
| <b>Max Obs per Subject</b>   | 7                 |
| <b>Parameters</b>            | 5                 |
| <b>Quadrature Points</b>     | 1                 |

Tabla 3.7: Dimensiones para el modelo de curva de crecimiento

La tabla “ Parameters ” (parámetros) enumera los parámetros que deben estimarse, sus valores iniciales y la probabilidad de registro negativo evaluada en los valores iniciales (Tabla 3.8).

| Initial Parameters |     |     |      |     |                         |
|--------------------|-----|-----|------|-----|-------------------------|
| b1                 | b2  | b3  | s2u  | s2e | Negative Log Likelihood |
| 190                | 700 | 350 | 1000 | 60  | 132.491787              |

Tabla 3.8: Valores iniciales de las estimaciones de parámetros y probabilidad de registro negativo

| Iteration | Calls | Iteration History       |            |                  |          |
|-----------|-------|-------------------------|------------|------------------|----------|
|           |       | Negative Log Likelihood | Difference | Maximum Gradient | Slope    |
| 1         | 8     | 131.6867                | 0.805045   | 0.010269         | -0.63300 |
| 2         | 12    | 131.6447                | 0.042082   | 0.014783         | -0.01820 |
| 3         | 16    | 131.6141                | 0.030583   | 0.009809         | -0.02796 |
| 4         | 20    | 131.5725                | 0.041555   | 0.001186         | -0.01344 |
| 5         | 22    | 131.5719                | 0.000627   | 0.000200         | -0.00121 |
| 6         | 25    | 131.5719                | 5.549E-6   | 0.000092         | -7.68E-6 |
| 7         | 28    | 131.5719                | 1.096E-6   | 6.097E-6         | -1.29E-6 |

Tabla 3.9: Historia de la iteración para el modelo de curva de crecimiento

La tabla “ Iteration History ” registra el historial de minimización del log de verosimilitud negativo (Tabla 3.9). Para cada iteración de la optimización cuasi-Newton, se enumeran los valores para el número de llamadas de función, el valor del log de verosimilitud negativo, la diferencia con respecto a la iteración anterior, el valor absoluto de la mayor pendiente y la pendiente de la dirección de búsqueda. La nota al final de la tabla indica que el algoritmo ha convergido exitosamente de acuerdo con el criterio de convergencia de GCONV, un criterio estándar calculado usando una forma cuadrática en el gradiente y el Hessian inverso.

El valor final maximizado de la probabilidad de registro, así como el criterio de información de Akaike (AIC), su versión corregida de sesgo de muestra pequeña (AICC) y el criterio de información bayesiano (BIC) en la forma “ **lo más pequeño es lo mejor** ” aparecen en la tabla “ Estadísticas de ajuste ” (Tabla 3.10). Estas estadísticas pueden utilizarse para comparar diferentes modelos mixtos no lineales.

| Fit Statistics           |       |
|--------------------------|-------|
| -2 Log Likelihood        | 263.1 |
| AIC (smaller is better)  | 273.1 |
| AICC (smaller is better) | 275.2 |
| BIC (smaller is better)  | 271.2 |

Tabla 3.10: Estadísticas de ajuste para el modelo de curva de crecimiento

| Parameter | Estimate | Standard<br>Error | Parameter Estimates |         |            | 0.95 Conf. Lim. | Gradient |
|-----------|----------|-------------------|---------------------|---------|------------|-----------------|----------|
|           |          |                   | DF                  | t Value | $Pr >  t $ |                 |          |
| b1        | 192.05   | 15.6473           | 4                   | 12.27   | 0.0003     | 148.61 235.50   | 1.154E-6 |
| b2        | 727.90   | 35.2474           | 4                   | 20.65   | <.0001     | 630.04 825.76   | 5.289E-6 |
| b3        | 348.07   | 27.0793           | 4                   | 12.85   | 0.0002     | 272.88 423.25   | -6.1E-6  |
| s2u       | 999.88   | 647.44            | 4                   | 1.54    | 0.1974     | -797.71 2797.46 | -3.84E-6 |
| s2e       | 61.5139  | 15.8832           | 4                   | 3.87    | 0.0179     | 17.4150 105.61  | 2.892E-6 |

Tabla 3.11: Estimaciones de parámetros en la convergencia

Las estimaciones de máxima verosimilitud de los cinco parámetros y sus errores estándar aproximados calculados utilizando la matriz hessiana final se muestran en la tabla “Parameter Estimates” (Estimaciones de parámetros) (Tabla 3.11).

También se proporcionan valores t aproximados y límites de confianza tipo Wald, con grados de libertad iguales al número de sujetos menos el número de efectos aleatorios. Se deben interpretar estas estadísticas con cautela para los parámetros de varianza como s2u y s2e. La última columna de la salida muestra el vector de gradiente en la solución de optimización. Cada elemento parece ser lo suficientemente pequeño para indicar un punto estacionario.

### 3.4.2. Modelo Logístico-Normal con Datos Binomiales

Este ejemplo analiza los datos de [58], que representan los resultados de un ensayo clínico multicéntrico que investiga la eficacia de dos tratamientos tópicos con crema (medicamento activo, control) para curar una infección. Para cada una de las ocho clínicas, se registra el número de ensayos y curas favorables para cada tratamiento. El conjunto de datos SAS es el siguiente:

```
data infection;
input clinic t x n;
datalines;
1 1 11 36
```

```

1 0 10 37
2 1 16 20
2 0 22 32
3 1 14 19
3 0 7 19
4 1 2 16
4 0 1 17
5 1 6 17
5 0 0 12
6 1 1 11
6 0 0 10
7 1 1 5
7 0 1 9
8 1 4 6
8 0 6 7
;

```

Supongamos que  $n_{ij}$  denota el número de ensayos para la  $i$ -ésima clínica y el tratamiento  $j$ -ésimo ( $i=1,\dots,8;j=0,1$ ), y  $x_{ij}$  denota el número correspondiente de curas favorables. Entonces un modelo razonable para los datos precedentes es el siguiente modelo logístico con efectos aleatorios:

$$x_{ij}|u_i \sim \text{Binomial}(n_{ij}, p_{ij})$$

y la ecuación:

$$n_{ij} = \log \left( \frac{p_{ij}}{1 - p_{ij}} \right) = \beta_0 + \beta_1 t_j + u_i \quad (3.4.2)$$

La notación  $t_j$  indica el tratamiento  $j$ -ésimo, y se supone que la  $u_i$  es iid  $N(0, \sigma_u^2)$

Las instrucciones PROC NLMIXED para el ajuste del modelo son las siguientes:

```

proc nlmixed data=infection;
parms beta0=-1 beta1=1 s2u=2;
eta = beta0 + beta1*t + u;
expeta = exp(eta);

```

```

p = expeta/(1+expeta);
model x ~ binomial(n,p);
random u ~ normal(0,s2u) subject=clinic;
predict eta out=eta;
estimate '1/beta1' 1/beta1;
run;

```

La sentencia PROC NLMIXED invoca el procedimiento, y la sentencia PARMS define los parámetros y sus valores iniciales. Las siguientes tres expresiones definen  $p_{ij}$ , y la expresión MODELO define la distribución condicional de  $x_{ij}$  como binomial. La sentencia RANDOM define  $u$  como el efecto aleatorio con sujetos definidos por la variable clínica.

La declaración PREDICT construye predicciones para cada observación en el conjunto de datos de entrada. Para este ejemplo, las predicciones de  $n_{ij}$  y los errores estándar aproximados de predicción se envían a un conjunto de datos llamado eta. Estas predicciones incluyen estimaciones empíricas de Bayes de los efectos aleatorios  $u_i$ .

La sentencia ESTIMATE solicita una estimación del recíproco de  $\beta_1$ .

El resultado de este modelo es el siguiente:

|  | Specifications               |
|--|------------------------------|
| <b>Data Set</b>                            | WORK.INFECTION               |
| <b>Dependent Variable</b>                  | x                            |
| <b>Distribution for Dependent Variable</b> | Binomial                     |
| <b>Random Effects</b>                      | u                            |
| <b>Distribution for Random Effects</b>     | Normal                       |
| <b>Subject Variable</b>                    | clinic                       |
| <b>Optimization Technique</b>              | Dual Quasi-Newton            |
| <b>Integration Method</b>                  | Adaptive Gaussian Quadrature |

Tabla 3.12: Información del Modelo Logístico-Normal

|                              | Dimensions |
|------------------------------|------------|
| <b>Observations Used</b>     | 16         |
| <b>Observations Not Used</b> | 0          |
| <b>Total Observations</b>    | 16         |
| <b>Subjects</b>              | 8          |
| <b>Max Obs per Subject</b>   | 2          |
| <b>Parameters</b>            | 3          |
| <b>Quadrature Points</b>     | 5          |

Tabla 3.13: Dimensiones para el Modelo Logístico-Normal

La tabla “ Specifications ” proporciona información básica sobre el modelo mixto no lineal (Tabla 3.12). Por ejemplo, la distribución de la variable de respuesta, condicionada a efectos aleatorios normalmente distribuidos, es binomial. La tabla “ Dimensions ” proporciona el recuento de varias variables. Debe comprobar esta tabla para asegurarse de que el conjunto de datos y el modelo se han introducido correctamente. PROC NLMIXED selecciona cinco puntos de cuadratura para lograr la precisión por defecto en los cálculos de verosimilitud.

| Initial Parameters |       |     |                         |
|--------------------|-------|-----|-------------------------|
| beta0              | beta1 | s2u | Negative Log Likelihood |
| -1                 | 1     | 2   | 37.5945925              |

Tabla 3.14: Valores iniciales paravestimaciones de parámetros

La tabla “ Parameters ” (Parámetros) enumera el punto inicial de la optimización del log de verosimilitud negativo en sus valores iniciales (Tabla 3.14).

| Iteration History |       |                         |            |                  |          |
|-------------------|-------|-------------------------|------------|------------------|----------|
| Iteration         | Calls | Negative Log Likelihood | Difference | Maximum Gradient | Slope    |
| 1                 | 4     | 37.3622692              | 0.232323   | 2.88208          | -19.3762 |
| 2                 | 6     | 37.1460375              | 0.216232   | 0.92193          | -0.82852 |
| 3                 | 9     | 37.0300936              | 0.115944   | 0.31590          | -0.59175 |
| 4                 | 11    | 37.0223017              | 0.007792   | 0.019060         | -0.01615 |
| 5                 | 13    | 37.0222472              | 0.000054   | 0.001743         | -0.00011 |
| 6                 | 16    | 37.0222466              | 6.57E-7    | 0.000091         | -1.28E-6 |
| 7                 | 19    | 37.0222466              | 5.38E-10   | 2.078E-6         | -1.1E-9  |

Tabla 3.15: Historial de iteración y estadísticas de ajuste para el modelo Logístico-Normal

| Fit Statistics           |      |
|--------------------------|------|
| -2 Log Likelihood        | 74.0 |
| AIC (smaller is better)  | 80.0 |
| AICC (smaller is better) | 82.0 |
| BIC (smaller is better)  | 80.3 |

Tabla 3.16: Estadísticas de ajuste para el modelo Logístico-Normal con Datos Binomiales

La tabla “ Iteration History ” indica una convergencia exitosa en siete iteraciones (Tabla 3.15). La tabla “ Fit Statistics ” enumera algunas estadísticas útiles basadas en el valor maximizado del log de la verosimilitud.

| Parameter | Estimate | Parameter Estimates |    |         |            | 95 % C. L.     | Gradient |
|-----------|----------|---------------------|----|---------|------------|----------------|----------|
|           |          | Standard Error      | DF | t Value | $Pr >  t $ |                |          |
| beta0     | -1.1974  | 0.5561              | 7  | -2.15   | 0.0683     | -2.5123 0.1175 | -3.1E-7  |
| beta1     | 0.7385   | 0.3004              | 7  | 2.46    | 0.0436     | 0.02806 1.4488 | -2.08E-6 |
| s2u       | 1.9591   | 1.1903              | 7  | 1.65    | 0.1438     | -0.8555 4.7737 | -2.48E-7 |

Tabla 3.17: Estimación de Parámetros para el Modelo Logístico-Normal

La tabla “ Parameter Estimates ” indica la importancia marginal de los dos parámetros de efectos fijos (Tabla 3.17). El valor positivo de la estimación de  $\beta_1$  indica que el tratamiento aumenta significativamente la posibilidad de una cura favorable.

| Label   | Estimate | Additional Estimates |    |         |            | Alpha Lower | Upper          |
|---------|----------|----------------------|----|---------|------------|-------------|----------------|
|         |          | Standard Error       | DF | t Value | $Pr >  t $ |             |                |
| 1/beta1 | 1.3542   | 0.5509               | 7  | 2.46    | 0.0436     | 0.05        | 0.05146 2.6569 |

Tabla 3.18: Tabla de Estimaciones Adicionales

La tabla “ Additional Estimates ” muestra los resultados de la sentencia ESTIMATE (Tabla 3.18). La estimación de  $1/\beta_1$  es igual a  $1/0,7385 = 1,3542$  y su error estándar

es igual a  $0,3004/0,7385^2 = 0,5509$  por el método delta ([59]; [60]). Nótese que esta aproximación en particular produce una t-estadística idéntica a la de la estimación de  $\beta_1$ . No se muestra el conjunto de datos eta, que contiene las 16 observaciones y predicciones originales del  $n_{ij}$ .

### 3.4.3. Modelo de un compartimento con datos farmacocinéticos

Una aplicación popular de los modelos mixtos no lineales es en el campo de la farmacocinética, que estudia cómo se dispersa un medicamento en un individuo vivo. Este ejemplo considera los datos de teofilina de [57]. Las concentraciones séricas del medicamento teofilina se miden en 12 sujetos durante un período de 25 horas después de la administración oral. Los datos son los siguientes:

```
data theoph;
input subject time conc dose wt;
datalines;
1 0.00 0.74 4.02 79.6
1 0.25 2.84 4.02 79.6
1 0.57 6.57 4.02 79.6
1 1.12 10.50 4.02 79.6
1 2.02 9.66 4.02 79.6
1 3.82 8.58 4.02 79.6
1 5.10 8.36 4.02 79.6
1 7.03 7.47 4.02 79.6
... more lines ...
12 24.15 1.17 5.30 60.5;
```

[57] considere el siguiente modelo de compartimento de primer orden para estos datos:

$$C_{it} = \frac{Dk_{ei}k_{ai}}{Cl_i(k_{ai} - k_{ei})} [\exp(-k_{ei}t) - \exp(-k_{ait})] + e_{it}$$

donde  $C_{it}$  es la concentración observada del  $i$ -ésimo sujeto en el tiempo  $t$ ,  $D$  es la dosis de teofilina,  $k_{ei}$  es la constante de la tasa de eliminación para el sujeto  $i$ ,  $k_{ai}$  es la constante de la tasa de absorción para el sujeto  $i$ ,  $Cl_i$  es la depuración para el sujeto  $i$ , y  $e_{it}$  son



errores normales. Para permitir la variabilidad aleatoria entre sujetos, ellos asumen.

$$Cl_i = \exp(\beta_1 + b_{i1})$$

$$k_{ai} = \exp(\beta_2 + b_{i2})$$

$$k_{ei} = \exp(\beta_3)$$

donde los  $\beta_s$  denota parámetros de efectos fijos y  $b_{is}$  denota parámetros de efectos aleatorios con un valor desconocido para la matriz de covarianza. Las instrucciones PROC NLMIXED para ajustarse a este modelo son las siguientes:

```
proc nlmixed data=theoph;
parms beta1=-3.22 beta2=0.47 beta3=-2.45
s2b1 =0.03 cb12 =0 s2b2 =0.4 s2=0.5;
cl = exp(beta1 + b1);
ka = exp(beta2 + b2);
ke = exp(beta3);
pred = dose*ke*ka*(exp(-ke*time)-exp(-ka*time))/cl/(ka-ke);
model conc ~ normal(pred, s2);
random b1 b2 ~ normal([0,0],[s2b1,cb12,s2b2]) subject=subject;
run;
```

La sentencia PARMs especifica valores iniciales para los tres  $\beta_s$  cuatro parámetros de covarianza-varianzas. Las constantes de depuración y tasa se definen utilizando instrucciones de programación SAS, y el modelo condicional para los datos se define como normal con la media pred y la varianza s2. Los dos efectos aleatorios son b1 y b2, y su distribución conjunta se define en la declaración RANDOM. Los paréntesis se utilizan para definir su vector de medias (dos ceros) y la matriz triangular inferior de varianza-covarianza (una matriz general de 2x2). Los resultados de este análisis son los siguientes.

|                                     | Specifications |
|-------------------------------------|----------------|
| Data Set                            | WORK.THEOPH    |
| Dependent Variable                  | conc           |
| Distribution for Dependent Variable | Normal         |
| Random Effects                      | b1 b2          |
| Distribution for Random Effects     | Normal         |
| Subject Variable                    | subject        |

|                        |                        |
|------------------------|------------------------|
| Optimization Technique | Dual Quasi–Newton      |
| Integration Method     | Adaptive      Gaussian |
|                        | Quadrature             |

Tabla 3.19: Especificación del modelo para un compartimento

En la tabla “ Specifications ” se indica la configuración del modelo (Tabla 3.19). La tabla “ Dimensions ” indica que hay 132 observaciones, 12 sujetos y 7 parámetros. PROC NLMIXED selecciona 5 puntos de cuadratura para cada efecto aleatorio, produciendo una cuadrícula total de 25 (Tabla 3.20).

|                       | Dimensions |
|-----------------------|------------|
| Observations Used     | 132        |
| Observations Not Used | 0          |
| Total Observations    | 132        |
| Subjects              | 12         |
| Max Obs per Subject   | 11         |
| Parameters            | 7          |
| Quadrature Points     | 5          |

Tabla 3.20: Tabla de dimensiones para el modelo de un compartimento

La tabla “ Parameters ” enumera los 7 parámetros, sus valores iniciales y la evaluación inicial de los valores negativos del log likelihood usando cuadratura gaussiana adaptativa (Tabla 3.21). La tabla “ Iteration History ” indica que se requieren 10 pasos para que el algoritmo dual cuasi-Newton logre la convergencia.

| Initial Parameters |       |       |      |      |      |     |                         |
|--------------------|-------|-------|------|------|------|-----|-------------------------|
| beta1              | beta2 | beta3 | s2b1 | cb12 | s2b2 | s2  | Negative Log Likelihood |
| −3.22              | 0.47  | −2.45 | 0.03 | 0    | 0.4  | 0.5 | 177.789945              |

Tabla 3.21: Valores iniciales e historial de iteraciones

| Iteration History |       |                         |            |                  |          |
|-------------------|-------|-------------------------|------------|------------------|----------|
| Iteration         | Calls | Negative Log Likelihood | Difference | Maximum Gradient | Slope    |
| 1                 | 7     | 177.7762                | 0.013697   | 2.87337          | -63.0744 |
| 2                 | 11    | 177.7643                | 0.011948   | 1.69814          | -4.75239 |
| 3                 | 14    | 177.7573                | 0.007036   | 1.29744          | -1.97311 |
| 4                 | 17    | 177.7557                | 0.001576   | 1.44141          | -0.49772 |
| 5                 | 20    | 177.7467                | 0.008988   | 1.13228          | -0.82230 |
| 6                 | 24    | 177.7464                | 0.000299   | 0.83129          | -0.00244 |
| 7                 | 27    | 177.7463                | 0.000083   | 0.72420          | -0.00789 |
| 8                 | 31    | 177.7457                | 0.000578   | 0.18002          | -0.00583 |
| 9                 | 34    | 177.7457                | 3.88E-6    | 0.017958         | -8.25E-6 |
| 10                | 37    | 177.7457                | 3.222E-8   | 0.000143         | -6.51E-8 |

Tabla 3.22: Valores iniciales e historial de iteraciones

| Fit Statistics           |       |
|--------------------------|-------|
| -2 Log Likelihood        | 355.5 |
| AIC (smaller is better)  | 369.5 |
| AICC (smaller is better) | 370.4 |
| BIC (smaller is better)  | 372.9 |

Tabla 3.23: Estadísticas de ajuste para el modelo de un solo compartimento

La tabla “ Fit Statistics ” enumera los valores finales optimizados de la función de registro de la verosimilitud y los criterios de información en la forma ”lo más pequeño es mejor” (Tabla 3.23).

La tabla “ Parameter Estimates ” contiene las estimaciones de máxima verosimilitud de los parámetros (Tabla 3.24). Tanto s2b1 como s2b2 son marginalmente significativos, lo que indica una variabilidad entre sujetos en las tolerancias y las constantes de la tasa de absorción, respectivamente. No parece haber una covarianza significativa entre ellos, como se observa en la estimación de cb12.

| Parameter | Estimate | Parameter Estimates |    |         |            |                  |         |          |
|-----------|----------|---------------------|----|---------|------------|------------------|---------|----------|
|           |          | Standard Error      | DF | t Value | $Pr >  t $ | 95 % C.I. Limits |         | Gradient |
| beta1     | -3.2268  | 0.05950             | 10 | -54.23  | <.0001     | -3.3594          | -3.0942 | -0.00009 |
| beta2     | 0.4806   | 0.1989              | 10 | 2.42    | 0.0363     | 0.03745          | 0.9238  | 3.645E-7 |
| beta3     | -2.4592  | 0.05126             | 10 | -47.97  | <.0001     | -2.5734          | -2.3449 | 0.000039 |
| s2b1      | 0.02803  | 0.01221             | 10 | 2.30    | 0.0446     | 0.000828         | 0.05524 | -0.00014 |
| cb12      | -0.00127 | 0.03404             | 10 | -0.04   | 0.9710     | -0.07712         | 0.07458 | -0.00007 |
| s2b2      | 0.4331   | 0.2005              | 10 | 2.16    | 0.0561     | -0.01354         | 0.8798  | -6.98E-6 |
| s2        | 0.5016   | 0.06837             | 10 | 7.34    | <.0001     | 0.3493           | 0.6540  | 6.133E-6 |

Tabla 3.24: Cálculo de parámetros para el modelo de un solo compartimento

Las estimaciones de  $\beta_1$ ,  $\beta_2$  y  $\beta_3$  se acercan a las estimaciones de la cuadratura adaptativa enumeradas de [57]. Sin embargo, Pinheiro y Bates utilizan una parametrización Cholesky-root para la matriz de varianza de efectos aleatorios y una parametrización logarítmica para la varianza residual. Las instrucciones PROC NLMIXED que utilizan su parametrización son las siguientes, y los resultados son similares.

```
proc nlmixed data=theoph;
parms l11=-1.5 l2=0 l13=-0.1 beta1=-3 beta2=0.5 beta3=-2.5 ls2=-0.7;
s2 = exp(ls2);
l1 = exp(l11);
l3 = exp(l13);
s2b1 = l1*l1*s2;
cb12 = l2*l1*s2;
s2b2 = (l2*l2 + l3*l3)*s2;
cl = exp(beta1 + b1);
ka = exp(beta2 + b2);
ke = exp(beta3);
pred = dose*ke*ka*(exp(-ke*time)-exp(-ka*time))/cl/(ka-ke);
model conc ~ normal(pred, s2);
random b1 b2 ~ normal([0, 0], [s2b1, cb12, s2b2]) subject=subject;
```

```
run;
```

### 3.4.4. Modelo Probit-Normal con datos binomiales

Considere los datos de [80], también estudiados por: [81]; [82]; [45]. En este experimento 16 ratas preñadas reciben una dieta de control y 16 reciben una dieta tratada químicamente, y el tamaño de la camada para cada rata se registra después de 4 y 21 días. A continuación se presenta el conjunto de datos SAS:

```
data rats;
input trt m * 4 ;
if (trt='c') then do;
x1 = 1;
x2 = 0;
end;
else do;
x1 = 0;
x2 = 1;
end;
litter = _n_;
datalines;
c 13 13 c 12 12 c 9 9 c 9 9 c 8 8 c 8 8 c 13 12 c 12 11
c 10 9 c 10 9 c 9 8 c 13 11 c 5 4 c 7 5 c 10 7 c 10 7
t 12 12 t 11 11 t 10 10 t 9 9 t 11 10 t 10 9 t 10 9 t 9 8
t 9 8 t 5 4 t 9 7 t 7 4 t 10 5 t 6 3 t 10 3 t 7 0
;
```

Aquí,  $m$  representa el tamaño de la camada después de 4 días, y  $x$  representa el tamaño de la camada después de 21 días. Además, se construyen variables indicadoras  $x_1$  y  $x_2$  para los dos niveles de tratamiento.

Siguiendo a [45], se asume un modelo de supervivencia latente de la forma

$$y_{ijk} = t_i + \alpha_1 i j + e_{ijk}$$

Donde los índices  $i$ ,  $j$  y  $k$  corresponden al tratamiento, camada y ratas recién nacidas dentro de una camada, respectivamente. El  $t_i$  representa la media de los tratamientos,  $\alpha_{ij}$  representa los efectos aleatorios de la camada que se supone que son iid  $N(0, s_i^2)$ , y  $e_{ijk}$  representa los errores residuales, todos en la escala latente.

En lugar de observar los tiempos de supervivencia  $y_{ijk}$ , asuma que sólo se observa la variable binaria que indica si  $y_{ijk}$  excede 0. Si  $x_{ij}$  denota la suma de estas variables binarias para el  $i$ -ésimo tratamiento y la camada  $j$ , entonces los supuestos precedentes conducen al siguiente modelo mixto lineal generalizado:

$$x_{ij} | \alpha_{ij} \sim \text{Binomial}(m_{ij}, p_{ij})$$

donde  $m_{ij}$  es el tamaño de cada camada después de 4 días y

$$p_{ij} = \Phi(t_i + \alpha_{ij})$$

Las instrucciones PROC NLMIXED para ajustarse a este modelo son las siguientes:

```
proc nlmixed data=rats ;
parms t1=1 t2=1 s1=.05 s2=1;
eta = x1*t1 + x2*t2 + alpha ;
p = probnorm(eta );
model x ~ binomial(m,p);
random alpha ~ normal(0,x1*s1*s1+x2*s2*s2) subject=litter ;
estimate 'gamma2' t2/sqrt(1+s2*s2);
predict p out=p;
run ;
```

La sentencia PROC NLMIXED invoca el procedimiento y la sentencia PARMS define los parámetros. Los parámetros son los dos medias de tratamiento  $t1$  y  $t2$  y las dos desviaciones estándar de efectos aleatorios,  $s1$  y  $s2$ .

Las variables indicadoras  $x1$  y  $x2$  se utilizan en el programa para asignar la media adecuada a cada observación en el conjunto de datos de entrada, así como la varianza adecuada de los efectos aleatorios. Tenga en cuenta que las expresiones de programación son permitidas dentro de las especificaciones de distribución. El resultado de este modelo

es el siguiente:

### Especificaciones, dimensiones y valores iniciales

|   | Specifications                  |
|---|---------------------------------|
| Conjunto de datos                         | WORK.RATS                       |
| Variable dependiente                      | x                               |
| Distribución para la variable dependiente | Binomial                        |
| Efectos aleatorios                        | alpha                           |
| Distribución de efectos aleatorios        | Normal                          |
| Variable objetivo                         | litter                          |
| Técnica de optimización                   | Dual Quasi–Newton               |
| Método de integración                     | Adaptive Gaussian<br>Quadrature |

Tabla 3.25: Especificaciones

|                       | Dimensions |
|-----------------------|------------|
| Observations Used     | 32         |
| Observations Not Used | 0          |
| Total Observations    | 32         |
| Subjects              | 32         |
| Max Obs per Subject   | 1          |
| Parameters            | 4          |
| Quadrature Points     | 7          |

Tabla 3.26: Dimensiones

| Initial Parameters |    |      |    |                         |
|--------------------|----|------|----|-------------------------|
| t1                 | t2 | s1   | s2 | Negative Log Likelihood |
| 1                  | 1  | 0.05 | 1  | 54.9362323              |

Tabla 3.27: Valores iniciales

La tabla “ Specifications ” proporciona información básica sobre este modelo mixto

no lineal (Tabla 3.25), el recuento de varias variables. Obsérvese que cada observación de los datos comprende un tema distinto. Utilizando los valores iniciales de la tabla de “Parameters”, PROC NLMIXED determina que el parámetro la función logarítmica de la verosimilitud puede ser aproximada con suficiente precisión con una regla de cuadratura de siete puntos.

| Iteration History |       |                         |            |                  |          |
|-------------------|-------|-------------------------|------------|------------------|----------|
| Iteration         | Calls | Negative Log Likelihood | Difference | Maximum Gradient | Slope    |
| 1                 | 4     | 53.9933934              | 0.942839   | 11.0326          | -81.9428 |
| 2                 | 6     | 52.8753530              | 1.11804    | 2.14895          | -2.86277 |
| 3                 | 9     | 52.6350386              | 0.240314   | 0.32996          | -1.05049 |
| 4                 | 11    | 52.6319939              | 0.003045   | 0.12293          | -0.00672 |
| 5                 | 14    | 52.6313583              | 0.000636   | 0.028246         | -0.00352 |
| 6                 | 18    | 52.6313174              | 0.000041   | 0.013551         | -0.00023 |
| 7                 | 21    | 52.6313115              | 5.839E-6   | 0.000603         | -0.00001 |
| 8                 | 24    | 52.6313115              | 9.45E-9    | 0.000022         | -1.68E-8 |

Tabla 3.28: Historial de iteración para el modelo Probit-Normal

La tabla “Iteration History” indica una convergencia exitosa en 8 iteraciones (Tabla 3.28).

| Fit Statistics           |       |
|--------------------------|-------|
| -2 Log Likelihood        | 105.3 |
| AIC (smaller is better)  | 113.3 |
| AICC (smaller is better) | 114.7 |
| BIC (smaller is better)  | 119.1 |

Tabla 3.29: Estadísticas de ajuste para el modelo Probit-Normal

La tabla “Fit Statistics” enumera estadísticas útiles basadas en el valor maximizado de la probabilidad de registro (Tabla 3.29).



| Parameter Estimates |          |                |    |                |                    |                 |        |          |
|---------------------|----------|----------------|----|----------------|--------------------|-----------------|--------|----------|
| Parameter           | Estimate | Standard Error | DF | <i>t Value</i> | <i>Pr &gt;  t </i> | 95% C.I. Limits |        | Gradient |
| t1                  | 1.3063   | 0.1685         | 31 | 7.75           | <.0001             | 0.9626          | 1.6499 | -0.00002 |
| t2                  | 0.9475   | 0.3055         | 31 | 3.10           | 0.0041             | 0.3244          | 1.5705 | 9.283E-6 |
| s1                  | 0.2403   | 0.3015         | 31 | 0.80           | 0.4315             | -0.3746         | 0.8552 | 0.000014 |
| s2                  | 1.0292   | 0.2988         | 31 | 3.44           | 0.0017             | 0.4198          | 1.6386 | -3.16E-6 |

Tabla 3.30: Estimaciones de parámetros para el modelo Probit-Normal

La tabla “Parameter Estimates” indica el significado de todos los parámetros excepto S1 (Tabla 3.30).

| Additional Estimates |          |                |    |                |                    |       |        |        |
|----------------------|----------|----------------|----|----------------|--------------------|-------|--------|--------|
| Label                | Estimate | Standard Error | DF | <i>t Value</i> | <i>Pr &gt;  t </i> | Alpha | Lower  | Upper  |
| gamma2               | 0.6603   | 0.2165         | 31 | 3.05           | 0.0047             | 0.05  | 0.2186 | 1.1019 |

Tabla 3.31: Estimaciones adicionales

La tabla “Additional Estimates” muestra los resultados de la sentencia ESTIMATE (Tabla 3.31). La estimación de  $\gamma_2$  es igual a 0.66, coincidiendo con el obtenido por [45]. El error estándar 0.22 se calcula utilizando el método delta ([59]; [60]).

### 3.4.5. Modelo Probit-Normal con Datos Ordinales

Los datos son de [83], quienes describen un experimento cruzado en dos grupos de pacientes que usan dos dispositivos inhaladores diferentes (A y B). Los pacientes del grupo 1 usaron el dispositivo A durante una semana y luego el dispositivo B durante otra semana. Los pacientes del grupo 2 utilizaron los dispositivos en orden inverso. Los datos introducidos como un conjunto de datos SAS son los siguientes:

```
data inhaler;
```

```

input clarity group time;
gt = group*time;
sub = floor(( _n_+1)/2);
datalines;
1 0 0 59 1 0 1 59 1 0 0 35 2 0 1 35 1 0 0 3 3 0 1 3 1 0 0 2
4 0 1 2 2 0 0 11 1 0 1 11 2 0 0 27 2 0 1 27 2 0 0 2 3 0 1 2
2 0 0 1 4 0 1 1 4 0 0 1 1 0 1 1 4 0 0 1 2 0 1 1 1 1 0 63
1 1 1 63 1 1 0 13 2 1 1 13 2 1 0 40 1 1 1 40 2 1 0 15 2 1 1 15
3 1 0 7 1 1 1 7 3 1 0 2 2 1 1 2 3 1 0 1 3 1 1 1 4 1 0 2
1 1 1 2 4 1 0 1 3 1 1 1
;

```

La medida de la respuesta, la claridad, es la evaluación de los pacientes sobre la claridad de las instrucciones del folleto para los dispositivos. La variable de claridad está en una escala ordinal, con 1=fácil, 2=sólo claro después de la relectura, 3=no muy claro y 4=confuso. La variable de grupo indica el grupo de tratamiento, y la variable de tiempo indica el tiempo de medición. La variable freq indica el número de pacientes con exactamente las mismas respuestas. Se crea una variable gt para indicar una interacción de grupo a tiempo, y una variable subordinada para indicar a los pacientes.

$$y_{ij} = \beta_0 + \beta_1 g_i + \beta_2 t_j + \beta_3 g_i t_j + u_i + e_{ij}$$

donde el índice  $i$  es para el paciente y  $j$  es para el período de tiempo,  $g_i$  indica grupos,  $t_j$  indica tiempo,  $u_i$  es un efecto aleatorio normal a nivel de paciente, y  $e_{ij}$  son errores iid normales. Los  $\beta_s$  son coeficientes desconocidos para ser estimados.

En lugar de observar  $y_{ij}$ , sin embargo, sólo se observa si cae en uno de los cuatro intervalos:  $(-\infty, 0)$ ,  $(0, I_1)$ ,  $(I_1, I_1 + I_2)$ , o  $(I_1 + I_2, \infty)$ , donde  $I_1$  y  $I_2$  son positivos. La categoría resultante es el valor asignado a la variable de claridad.

El siguiente código configura y se ajusta a este modelo probit ordinal:

```

proc nlmixed data=inhaler corr ecorr;
parms b0=0 b1=0 b2=0 b3=0 sd=1 i1=1 i2=1;
bounds i1 > 0, i2 > 0;

```

```

eta = b0 + b1*group + b2*time + b3*gt + u;
if (clarity=1) then p = probnorm(-eta);
else if (clarity=2) then
p = probnorm(i1-eta) - probnorm(-eta);
else if (clarity=3) then
p = probnorm(i1+i2-eta) - probnorm(i1-eta);
else p = 1 - probnorm(i1+i2-eta);
if (p > 1e-8) then ll = log(p);
else ll = -1e20;
model clarity ~ general(ll);
random u ~ normal(0,sd*sd) subject=sub;
replicate freq;
estimate 'thresh2' i1;
estimate 'thresh3' i1 + i2;
estimate 'icc' sd*sd/(1+sd*sd);
run;

```

La sentencia PROC NLMIXED especifica el conjunto de datos de entrada y solicita correlaciones tanto para el parámetro (opción CORR) y para las estimaciones adicionales especificadas en las declaraciones ESTIMATE (ECORR opción).

Los parámetros definidos en la sentencia PARMs son los siguientes: b0 (intercepción global), b1 (grupo principal) efecto principal de tiempo), B2 (efecto principal de tiempo), b3 (interacción de tiempo con grupo), sd (desviación estándar del efecto aleatorio), i1 (incremento entre el primer y el segundo umbral) e i2 (incremento entre el segundo y el tercer umbral). La sentencia BOUNDS restringe i1 e i2 para que sean positivas.

Las sentencias de programación de SAS comienzan por definir el predictor lineal eta, que es una combinación lineal de los parámetros b y un único efecto aleatorio u. Las siguientes sentencias definen la verosimilitud ordinal de acuerdo con la variable de claridad, eta y las variables de incremento. Se incluye una trampa de errores en caso de que la probabilidad sea demasiado pequeña.

En la declaración MODELO se utiliza una especificación general de la verosimilitud

logarítmica, y la declaración RANDOM define el efecto aleatorio  $u$  tiene una desviación estándar  $sd$  y una variable subordinada. La sentencia REPLICATE indica que los datos de cada sujeto deben ser replicados de acuerdo con la variable `freq`.

Las sentencias ESTIMATE especifican el segundo y tercer umbral en términos de las variables de incremento (se supone que el primer umbral es igual a cero para la identificabilidad del modelo). También se calcula la correlación intraclase. La salida es la siguiente.

|   | Specifications                       |
|---|--------------------------------------|
| Conjunto de datos                         | WORK.INHALER                         |
| Variable dependiente                      | clarity                              |
| Distribución para la variable dependiente | General                              |
| Efectos aleatorios                        | $u$                                  |
| Distribución de efectos aleatorios        | Normal                               |
| Variable objetivo                         | sub                                  |
| Replicar variable                         | freq                                 |
| Técnica de optimización                   | Dual Quasi–Newton                    |
| Método de integración                     | Adaptive      Gaussian<br>Quadrature |

Tabla 3.32: Especificaciones para el modelo Ordinal Data

La tabla de “ Specifications ” hace eco de alguna información primaria especificada para este modelo mixto no lineal (Tabla 3.32). Debido a que la función de logaritmo de la verosimilitud se expresaba con las instrucciones de programación de SAS, la distribución se muestra como General en la tabla de “ Especificaciones ”.

La tabla “ Dimensions ” revela un total de 286 sujetos, que es la suma de los valores de la variable `FREQ` por segundo año consecutivo. Se seleccionan cinco puntos de cuadratura para la evaluación de la logaritmo de la verosimilitud (Tabla 3.33).

|                       | Dimensions |
|-----------------------|------------|
| Observations Used     | 38         |
| Observations Not Used | 0          |
| Total Observations    | 38         |
| Subjects              | 286        |
| Max Obs per Subject   | 2          |
| Parameters            | 7          |
| Quadrature Points     | 5          |

Tabla 3.33: Tabla de dimensiones para el modelo de datos ordinales

| Initial Parameters |    |    |    |    |    |    |            |     |
|--------------------|----|----|----|----|----|----|------------|-----|
| b0                 | b1 | b2 | b3 | sd | i1 | i2 | Negative   | Log |
|                    |    |    |    |    |    |    | Likelihood |     |
| 0                  | 0  | 0  | 0  | 1  | 1  | 1  | 538.484276 |     |

Tabla 3.34: Parámetro Valores iniciales y Log Negativo de verosimilitud

En la tabla “ Parameters ” se indican los valores iniciales simples para este problema (Tabla 3.34). La tabla “ Iteration History ” indica una convergencia exitosa en 13 iteraciones (Tabla 3.35).

| Iteration History |       |              |            |         |          |
|-------------------|-------|--------------|------------|---------|----------|
| Iteration         | Calls | Negative Log | Difference | Maximum | Slope    |
|                   |       | Likelihood   | Gradient   |         |          |
| 1                 | 4     | 476.3825     | 62.10176   | 43.7506 | -1431.40 |
| 2                 | 7     | 463.2282     | 13.15431   | 14.2465 | -106.753 |
| 3                 | 9     | 458.5281     | 4.70008    | 48.3132 | -33.0389 |
| 4                 | 11    | 450.9757     | 7.552383   | 22.6010 | -40.9954 |
| 5                 | 14    | 448.0127     | 2.963033   | 14.8688 | -16.7453 |
| 6                 | 17    | 447.2452     | 0.767549   | 7.77419 | -2.26743 |
| 7                 | 19    | 446.7277     | 0.517483   | 3.79353 | -1.59278 |
| 8                 | 22    | 446.5183     | 0.209396   | 0.86864 | -0.37801 |
| 9                 | 26    | 446.5145     | 0.003745   | 0.32857 | -0.02356 |

|    |    |          |          |          |          |
|----|----|----------|----------|----------|----------|
| 10 | 29 | 446.5133 | 0.001187 | 0.056778 | -0.00183 |
| 11 | 32 | 446.5133 | 0.000027 | 0.010785 | -0.00004 |
| 12 | 35 | 446.5133 | 3.956E-6 | 0.004922 | -5.41E-6 |
| 13 | 38 | 446.5133 | 1.989E-7 | 0.000470 | -4E-7    |

Tabla 3.35: Historia de la iteración

|                          | Fit Statistics |
|--------------------------|----------------|
| -2 Log Likelihood        | 893.0          |
| AIC (smaller is better)  | 907.0          |
| AICC (smaller is better) | 910.8          |
| BIC (smaller is better)  | 932.6          |

Tabla 3.36: Estadísticas de ajuste para el modelo de datos ordinales

La tabla “ Fit Statistics ” enumera los criterios de información y verosimilitud de los registros para las comparaciones de los modelos (Tabla 3.36).

| Parameter Estimate | Parameter Estimates |        |                |                    |                 |         |          |          |
|--------------------|---------------------|--------|----------------|--------------------|-----------------|---------|----------|----------|
|                    | Standard Error      | DF     | <i>t Value</i> | <i>Pr &gt;  t </i> | 95% C.I. Limits |         | Gradient |          |
| b0                 | -0.6364             | 0.1342 | 285            | -4.74              | <.0001          | -0.9006 | -0.3722  | 0.000470 |
| b1                 | 0.6007              | 0.1770 | 285            | 3.39               | 0.0008          | 0.2523  | 0.9491   | 0.000265 |
| b2                 | 0.6015              | 0.1582 | 285            | 3.80               | 0.0002          | 0.2900  | 0.9129   | 0.000080 |
| b3                 | -1.4817             | 0.2385 | 285            | -6.21              | <.0001          | -1.9512 | -1.0122  | 0.000102 |
| sd                 | 0.6599              | 0.1312 | 285            | 5.03               | <.0001          | 0.4017  | 0.9181   | -0.00009 |
| i1                 | 1.7450              | 0.1474 | 285            | 11.84              | <.0001          | 1.4548  | 2.0352   | 0.000202 |
| i2                 | 0.5985              | 0.1427 | 285            | 4.19               | <.0001          | 0.3176  | 0.8794   | 0.000087 |

Tabla 3.37: Estimaciones de parámetros en la convergencia

La tabla “ Parameter Estimates ” indica el significado de todos los parámetros (Tabla 3.37).

| Label   | Estimate | Additional Estimates |     |                |                        |       |        |        |
|---------|----------|----------------------|-----|----------------|------------------------|-------|--------|--------|
|         |          | Standard Error       | DF  | <i>t Value</i> | <i>Pr</i> >   <i>t</i> | Alpha | Lower  | Upper  |
| thresh2 | 1.7450   | 0.1474               | 285 | 11.84          | <.0001                 | 0.05  | 1.4548 | 2.0352 |
| thresh3 | 2.3435   | 0.2073               | 285 | 11.31          | <.0001                 | 0.05  | 1.9355 | 2.7516 |
| icc     | 0.3034   | 0.08402              | 285 | 3.61           | 0.0004                 | 0.05  | 0.1380 | 0.4687 |

Tabla 3.38: Estimaciones de Correlación de Umbrales e Intraclase

La tabla “ Estimaciones adicionales ” muestra los resultados de las declaraciones ESTIMATE (Resultado 3.38).

### 3.4.6. Modelo Poisson-Normal con datos de conteo

Utilizamos los datos de falla de la bomba de [84]. El número de fallos y el tiempo de funcionamiento se registran para 10 bombas. Cada una de las bombas se clasifica en uno de los dos grupos correspondientes al funcionamiento continuo o intermitente. Los datos son los siguientes:

```

data pump;
input y t group;
pump = _n_;
logtstd = log(t) - 2.4564900;
datalines;
5 94.320 1
1 15.720 2
5 62.880 1
14 125.760 1
3 5.240 2
19 31.440 1
1 1.048 2
1 1.048 2
4 2.096 2
22 10.480 2;

```

Cada fila denota datos para una sola bomba, y la variable `logtstd` contiene los tiempos de operación centrados. Dejando que  $y_{ij}$  denote el número de fallas para la bomba  $j$ -ésima en el grupo  $i$ -ésimo, [85] considera el siguiente modelo jerárquico para estos datos:

$$\begin{aligned} y_{ij} | \lambda_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\ \log \lambda_{ij} &= \alpha_i + \beta_i (\log t_{ij} - \overline{\log t}) + e_{ij} \\ e_{ij} | \sigma^2 &\sim \text{Normal}(0, \sigma^2) \end{aligned}$$

El modelo especifica diferentes intercepciones y pendientes para cada grupo, y el efecto aleatorio es un mecanismo para contabilizar la sobredispersión.

Las declaraciones correspondientes de PROC NLMIXED son las siguientes:

```
proc nlmixed data=pump;
parms logsig 0 beta1 1 beta2 1 alpha1 1 alpha2 1;
if (group = 1) then eta = alpha1 + beta1*logtstd + e;
else eta = alpha2 + beta2*logtstd + e;
lambda = exp(eta);
model y ~ poisson(lambda);
random e ~ normal(0,exp(2*logsig)) subject=pump;
estimate 'alpha1-alpha2' alpha1-alpha2;
estimate 'beta1-beta2' beta1-beta2;
run;
```

La salida seleccionada es la siguiente:

|                       | Dimensions |
|-----------------------|------------|
| Observations Used     | 10         |
| Observations Not Used | 0          |
| Total Observations    | 10         |
| Subjects              | 10         |
| Max Obs per Subject   | 1          |
| Parameters            | 5          |
| Quadrature Points     | 5          |

Tabla 3.39: Tabla de dimensiones para el modelo Poisson-Normal



La tabla “Dimensions” indica que se utilizan los datos de 10 bombas con una observación para cada una (Tabla 3.39).

| Iteration History |       |                         |            |                  |          |
|-------------------|-------|-------------------------|------------|------------------|----------|
| Iteration         | Calls | Negative Log Likelihood | Difference | Maximum Gradient | Slope    |
| 1                 | 4     | 30.6986932              | 2.162768   | 5.10725          | -91.6020 |
| 2                 | 9     | 30.0255468              | 0.673146   | 2.76174          | -11.0489 |
| 3                 | 12    | 29.7263250              | 0.299222   | 2.99040          | -2.36048 |
| 4                 | 16    | 28.7390263              | 0.987299   | 2.07443          | -3.93678 |
| 5                 | 18    | 28.3161933              | 0.422833   | 0.61253          | -0.63084 |
| 6                 | 21    | 28.0956400              | 0.220553   | 0.46216          | -0.52684 |
| 7                 | 24    | 28.0438024              | 0.051838   | 0.40505          | -0.10018 |
| 8                 | 27    | 28.0357134              | 0.008089   | 0.13506          | -0.01875 |
| 9                 | 30    | 28.0339250              | 0.001788   | 0.026279         | -0.00514 |
| 10                | 33    | 28.0338744              | 0.000051   | 0.004020         | -0.00012 |
| 11                | 36    | 28.0338727              | 1.681E-6   | 0.002864         | -5.09E-6 |
| 12                | 39    | 28.0338724              | 3.199E-7   | 0.000147         | -6.87E-7 |
| 13                | 42    | 28.0338724              | 2.532E-9   | 0.000017         | -5.75E-9 |

Tabla 3.40: Historial de iteración para el modelo Poisson-Normal

La tabla “Iteration History” indica una convergencia exitosa en 13 iteraciones (Tabla 3.40).

|                          | Fit Statistics |
|--------------------------|----------------|
| -2 Log Likelihood        | 56.1           |
| AIC (smaller is better)  | 66.1           |
| AICC (smaller is better) | 81.1           |
| BIC (smaller is better)  | 67.6           |

Tabla 3.41: Estadísticas de ajuste para el modelo Poisson-Normal

La tabla “Fit Statistics” enumera la probabilidad final de registro y los criterios de información asociados (Resultado 3.41).

| Parameter Estimates |          |                |    |                |                    |                 |        |          |
|---------------------|----------|----------------|----|----------------|--------------------|-----------------|--------|----------|
| Parameter           | Estimate | Standard Error | DF | <i>t Value</i> | <i>Pr &gt;  t </i> | 95% C.I. Limits |        | Gradient |
| logsig              | -0.3161  | 0.3213         | 9  | -0.98          | 0.3508             | -1.0429         | 0.4107 | -0.00002 |
| beta1               | -0.4256  | 0.7473         | 9  | -0.57          | 0.5829             | -2.1162         | 1.2649 | -0.00002 |
| beta2               | 0.6097   | 0.3814         | 9  | 1.60           | 0.1443             | -0.2530         | 1.4725 | -1.61E-6 |
| alpha1              | 2.9644   | 1.3826         | 9  | 2.14           | 0.0606             | -0.1632         | 6.0921 | -5.25E-6 |
| alpha2              | 1.7992   | 0.5492         | 9  | 3.28           | 0.0096             | 0.5568          | 3.0415 | -5.73E-6 |

Tabla 3.42: Estimaciones de parámetros

| Additional Estimates |          |                |    |                |                    |       |         |        |
|----------------------|----------|----------------|----|----------------|--------------------|-------|---------|--------|
| Label                | Estimate | Standard Error | DF | <i>t Value</i> | <i>Pr &gt;  t </i> | Alpha | Lower   | Upper  |
| alpha1-alpha2        | 1.1653   | 1.4855         | 9  | 0.78           | 0.4529             | 0.05  | -2.1952 | 4.5257 |
| beta1-beta2          | -1.0354  | 0.8389         | 9  | -1.23          | 0.2484             | 0.05  | -2.9331 | 0.8623 |

Tabla 3.43: Estimaciones adicionales

Las tablas “ Estimaciones de parámetros ” y “ Estimaciones adicionales ” enumeran las estimaciones de máxima verosimilitud para cada uno de los siguientes casos de los parámetros y dos diferencias (Tabla 3.42). Las estimaciones puntuales de los parámetros medios concuerdan bastante cerca de las medias posteriores bayesianas reportadas por [94]; sin embargo, las medias basadas en la probabilidad son aproximadamente la mitad de las desviaciones estándar posteriores bayesianas. Esto se debe probablemente al hecho de que que las desviaciones típicas bayesianas explican la incertidumbre en la estimación  $\sigma^2$ , mientras que la probabilidad se conectan a su valor estimado. Este sesgo a la baja puede corregirse en cierta medida utilizando la distribución  $t_9$  que se muestra aquí.

# Capítulo 4

## Estudio de Simulación

### 4.1. Ejemplo motivador

#### 4.1.1. Meta-análisis para DTA Datos Reales

[89], realizaron pruebas para detectar problemas por el consumo de alcohol, mediante un cuestionario que tenía 3 ítems en su versión abreviada (AUDIT-C). Catorce estudios, que evaluaron 18 332 individuos, de los cuales 2 071 tenían problemas con el alcohol, fueron incluidos en la revisión. La prevalencia en los estudios oscilo entre 5 % y 37 %, con una mediana de 11 %.

La Tabla 4.1 muestra la variación de los estudios en sus sensibilidades y especificidades. Aunque la sensibilidad fue perfecta (100 %) en dos estudios, esta particularidad se ve reflejada en otros índices con valores de 100 % y cero para el valor predictivo negativo (VPN) y la razón de verosimilitud negativa (LR-) respectivamente.

Para odds ratio de diagnóstico (DOR) se obtuvieron indeterminaciones (división por cero). Esta variabilidad en los estudios se puede observar de manera gráfica, en las Figuras 4.1, 4.2 y 4.3, donde se muestra las gráficas, forest plot, crosshair, rocellipse y curva SROC.

| Estudio | VP  | FN | FP  | VN   | Prev | Sen. | Esp. | VPP  | VPN  | LR+   | LR-  | DOR   |
|---------|-----|----|-----|------|------|------|------|------|------|-------|------|-------|
| 1       | 47  | 9  | 101 | 738  | 0.06 | 0.84 | 0.88 | 0.32 | 0.99 | 6.97  | 0.18 | 38.16 |
| 2       | 126 | 51 | 272 | 1543 | 0.09 | 0.71 | 0.85 | 0.32 | 0.97 | 4.75  | 0.34 | 14.02 |
| 3       | 19  | 10 | 12  | 192  | 0.12 | 0.66 | 0.94 | 0.61 | 0.95 | 11.14 | 0.37 | 30.4  |

|    |     |    |      |      |      |      |      |      |      |      |      |       |
|----|-----|----|------|------|------|------|------|------|------|------|------|-------|
| 4  | 36  | 3  | 78   | 276  | 0.1  | 0.92 | 0.78 | 0.32 | 0.99 | 4.19 | 0.1  | 42.46 |
| 5  | 130 | 19 | 211  | 959  | 0.11 | 0.87 | 0.82 | 0.38 | 0.98 | 4.84 | 0.16 | 31.1  |
| 6  | 84  | 2  | 68   | 89   | 0.35 | 0.98 | 0.57 | 0.55 | 0.98 | 2.26 | 0.04 | 54.97 |
| 7  | 68  | 0  | 112  | 423  | 0.11 | 1    | 0.79 | 0.38 | 1    | 4.78 | 0    | -     |
| 8  | 752 | 0  | 3226 | 2977 | 0.11 | 1    | 0.48 | 0.19 | 1    | 1.92 | 0    | -     |
| 9  | 59  | 5  | 55   | 136  | 0.25 | 0.92 | 0.71 | 0.52 | 0.96 | 3.2  | 0.11 | 29.18 |
| 10 | 142 | 50 | 571  | 2788 | 0.05 | 0.74 | 0.83 | 0.2  | 0.98 | 4.35 | 0.31 | 13.87 |
| 11 | 137 | 24 | 107  | 358  | 0.26 | 0.85 | 0.77 | 0.56 | 0.94 | 3.7  | 0.19 | 19.1  |
| 12 | 57  | 3  | 103  | 437  | 0.1  | 0.95 | 0.81 | 0.36 | 0.99 | 4.98 | 0.06 | 80.61 |
| 13 | 34  | 1  | 21   | 56   | 0.31 | 0.97 | 0.73 | 0.62 | 0.98 | 3.56 | 0.04 | 90.67 |
| 14 | 152 | 51 | 88   | 264  | 0.37 | 0.75 | 0.75 | 0.63 | 0.84 | 3    | 0.33 | 8.94  |

Tabla 4.1: Estimaciones de la sensibilidad (Sen.), especificidad (Esp.), valor predictivo positivo y negativo (VPP, VPN), razón de verosimilitud positiva y negativa (LR+,LR-) y odds ratio de diagnóstico (DOR), **datos AUDIT-C**

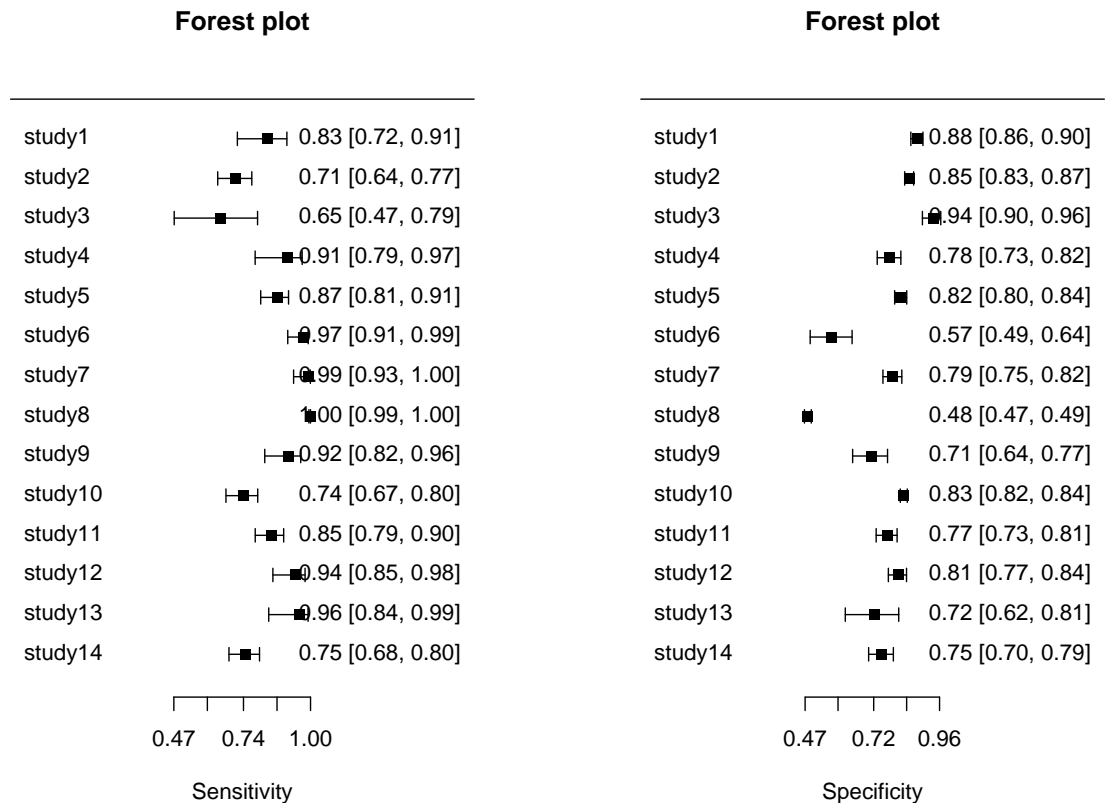


Figura 4.1: Forest Plot para sensibilidad y especificidad para datos AuditC

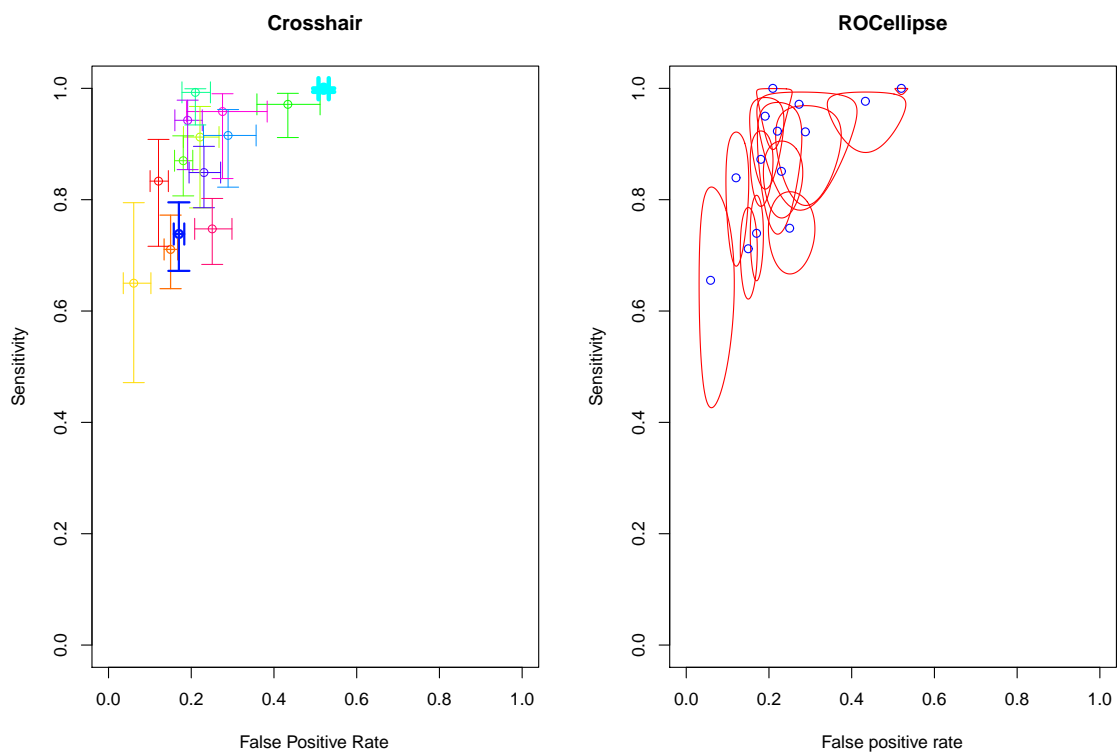


Figura 4.2: Gráficos crosshair y rocellipse para datos AuditC

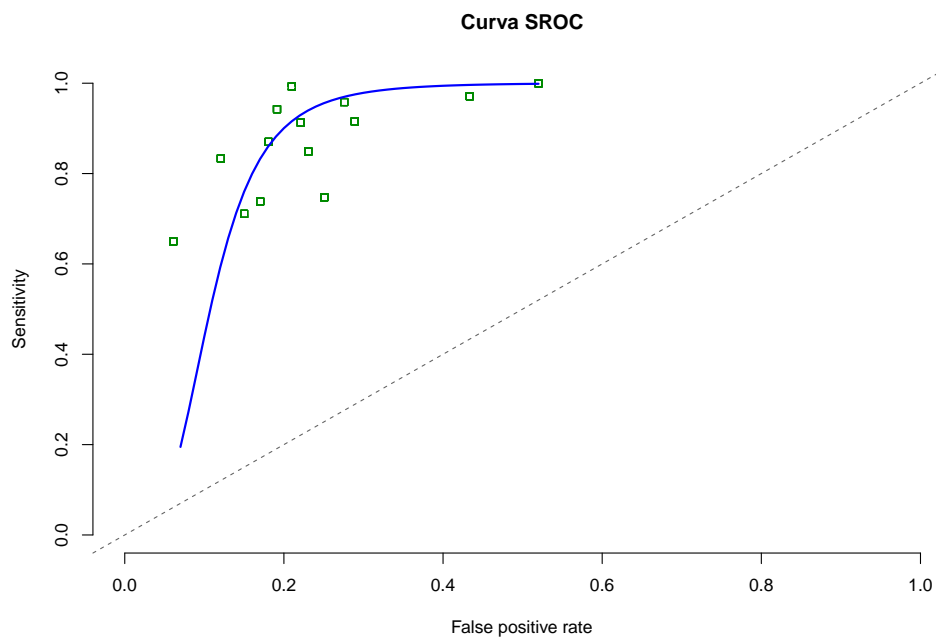


Figura 4.3: Gráficos crosshair y rocellipse para datos AuditC

### 4.1.2. Meta-análisis para DTA Datos Simulados

Empleando la función `simdata` del paquete `HSROC` de R, se simuló un meta-análisis con baja prevalencia. Se generaron 16 estudios que evaluaron 139 169 individuos, de los cuales 932 tenían la alteración o la enfermedad. La prevalencia en los estudios oscilo entre 0.18 % y 3.08 %, con una mediana de 0.93 %.

La Tabla 4.2 muestra la variación de los estudios en sus sensibilidades y especificidades. Aunque la sensibilidad fue perfecta (100 %) en un estudio, esta particularidad se ve reflejada en otros índices con valores de 100 % y cero para el valor predictivo negativo (VPN) y la razón de verosimilitud negativa (LR-) respectivamente. Para diagnostic odds ratio (DOR) se obtuvieron indeterminaciones (división por cero). Esta variabilidad en los estudios se puede observar de manera gráfica, en las Figuras 4.4, 4.5 y 4.6, donde se muestra las gráficas, forest plot, crosshair, rocellipse y curva sroc.

| Estudio | VP  | FN | FP  | VN    | Prev   | Sen. | Esp. | VPP  | VPN  | LR+   | LR-  | DOR   |
|---------|-----|----|-----|-------|--------|------|------|------|------|-------|------|-------|
| 1       | 30  | 6  | 3   | 3939  | 0.009  | 0.83 | 1    | 0.91 | 1    | 1095  | 0.17 | 6565  |
| 2       | 33  | 10 | 3   | 3939  | 0.0108 | 0.77 | 1    | 0.92 | 1    | 1008  | 0.23 | 4333  |
| 3       | 43  | 2  | 5   | 3935  | 0.0113 | 0.96 | 1    | 0.9  | 1    | 753   | 0.04 | 16920 |
| 4       | 20  | 74 | 14  | 16127 | 0.0058 | 0.21 | 1    | 0.59 | 1    | 245   | 0.79 | 311   |
| 5       | 27  | 3  | 17  | 3803  | 0.0078 | 0.9  | 1    | 0.61 | 1    | 202   | 0.1  | 2013  |
| 6       | 28  | 2  | 20  | 2978  | 0.0099 | 0.93 | 0.99 | 0.58 | 1    | 140   | 0.07 | 2085  |
| 7       | 20  | 22 | 24  | 1661  | 0.0243 | 0.48 | 0.99 | 0.45 | 0.99 | 33    | 0.53 | 63    |
| 8       | 92  | 6  | 79  | 18269 | 0.0053 | 0.94 | 1    | 0.54 | 1    | 218   | 0.06 | 3546  |
| 9       | 6   | 0  | 25  | 2024  | 0.0029 | 1    | 0.99 | 0.19 | 1    | 82    | 0    | -     |
| 10      | 170 | 30 | 41  | 20521 | 0.0096 | 0.85 | 1    | 0.81 | 1    | 426   | 0.15 | 2836  |
| 11      | 20  | 11 | 207 | 929   | 0.0266 | 0.65 | 0.82 | 0.09 | 0.99 | 4     | 0.43 | 8     |
| 12      | 20  | 16 | 84  | 1047  | 0.0308 | 0.56 | 0.93 | 0.19 | 0.98 | 7     | 0.48 | 16    |
| 13      | 28  | 13 | 265 | 6502  | 0.006  | 0.68 | 0.96 | 0.1  | 1    | 17    | 0.33 | 53    |
| 14      | 10  | 2  | 17  | 638   | 0.018  | 0.83 | 0.97 | 0.37 | 1    | 32    | 0.17 | 188   |
| 15      | 105 | 18 | 116 | 15373 | 0.0079 | 0.85 | 0.99 | 0.48 | 1    | 114   | 0.15 | 773   |
| 16      | 55  | 10 | 3   | 35629 | 0.0018 | 0.85 | 1    | 0.95 | 1    | 10050 | 0.15 | 65320 |

Tabla 4.2: Estimaciones de la sensibilidad (Sen.), especificidad (Esp.), valor predictivo positivo y negativo (VPP, VPN), razón de verosimilitud positiva y negativa (LR+,LR-) y diagnostic odds ratio (DOR), **datos simulados**

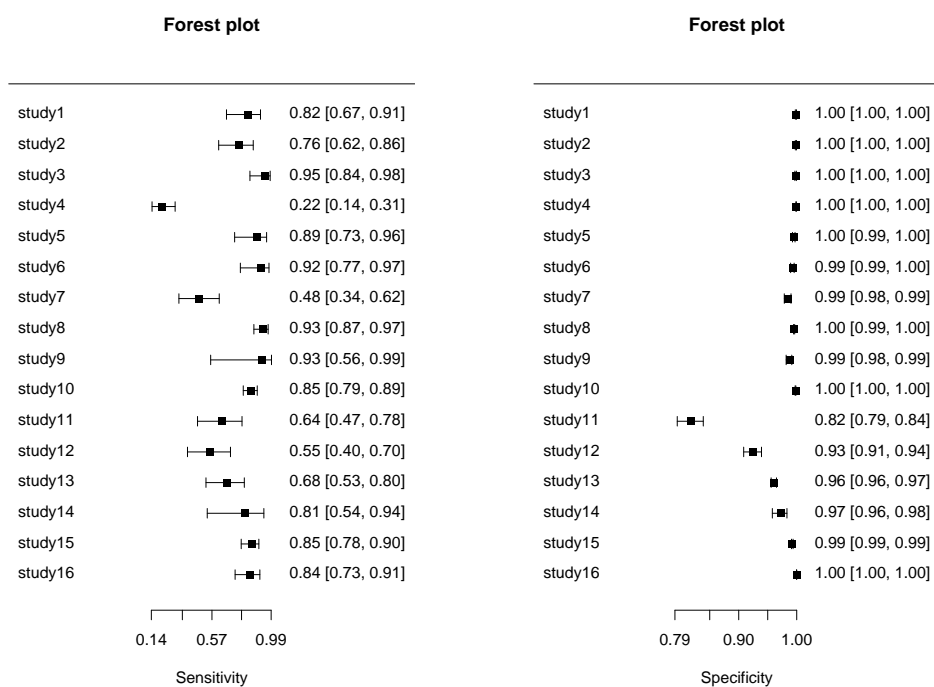


Figura 4.4: Forest Plot para sensibilidad y especificidad para datos Simulados

### 4.1.3. Análisis para los dos Meta-análisis para DTA

Se analizaron los dos conjuntos de datos mediante los modelos, Bivariante y HSROC, utilizando el programa SAS. Medidas de resumen adicionales tales como relaciones de verosimilitud y DORs pueden ser estimadas utilizando dicho procedimiento [90].

Para el conjunto de datos, simulado y AUDIT-C, los modelos Bivariante y HSROC convergieron exitosamente produciendo estimaciones confiables. Para el caso bivariante (datos simulados) los valores de sensibilidad y especificidad son de 0.80 (con un IC al 95 %: 0.6986-0.9056) y 0.99 (con un IC al 95 %: 0.9892-1.003) respectivamente. Las estimaciones de sensibilidad y especificidad para los datos AUDIT-C del modelo bivariante son de 0.9249 (con un IC al 95 %: 0.8551-0.9948) y 0.7818 (con un IC al 95 %: 0.7190-0.8446) respectivamente. Para mayor detalle de las estimaciones del modelo HSROC vea Tabla 4.3.

Los dos ejemplos ilustran el problema de las estimaciones de los modelos, por lo que se necesita un esquema de cuando poder utilizar ambos enfoques. Se encontraron diferencias sutiles en las estimaciones de la sensibilidad tanto para el conjunto de datos AUDIT-C y los datos simulados, pero no así para las estimaciones de las especificidades.

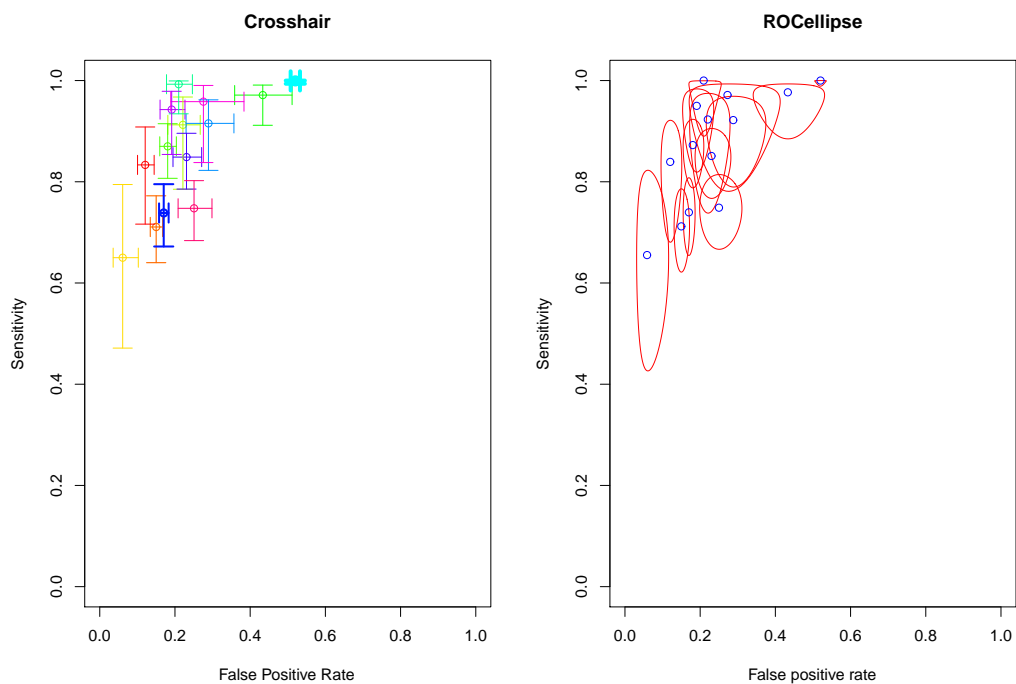


Figura 4.5: Gráficos crosshair y rocellipse para datos Simulados

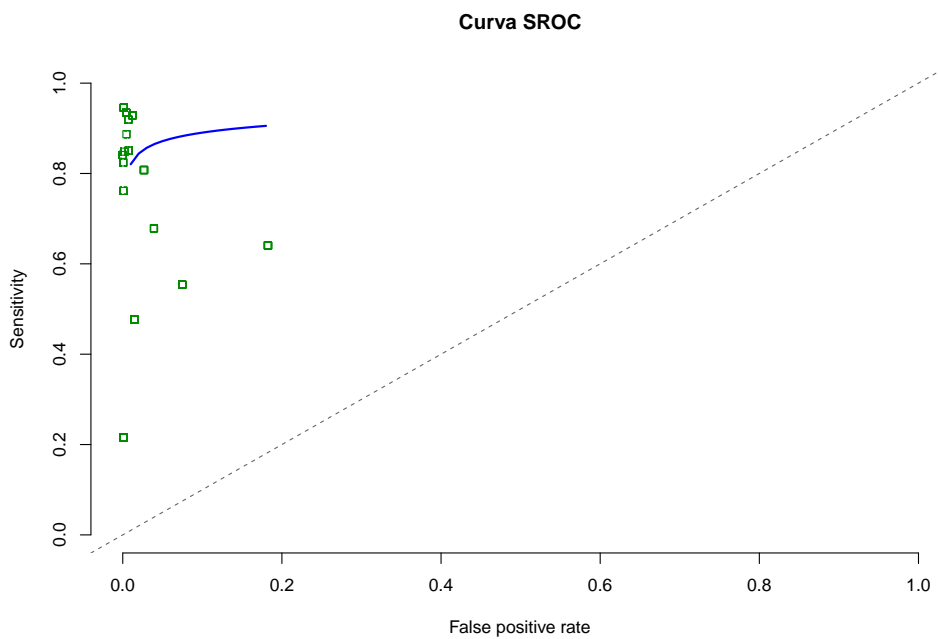


Figura 4.6: Gráficos crosshair y rocellipse para datos AuditC



Estos ejemplos muestran que los resultados pueden diferir entre modelos, y las diferencias pueden no ser insignificantes. Por lo tanto, la identificación, de que enfoque aplicar en situaciones de baja prevalencia es de suma importancia en las pruebas diagnósticas, para así poder dar respuestas válidas en situaciones donde pueden fallar ciertos modelos.

| <b>Medidas</b>                 | <b>Datos<br/>simulados</b> | <b>Datos<br/>AuditC</b> |
|--------------------------------|----------------------------|-------------------------|
| Sensibilidad                   | 0.6303                     | 0.7311                  |
| Intervalo de Confianza<br>(IC) | (0.5921-<br>0.6685)        | (0.7077-<br>0.7544)     |
| Especificidad                  | 0.986                      | 0.7312                  |
| Intervalo de Confianza<br>(IC) | (0.9848-<br>0.9872)        | (0.7159-<br>0.7465)     |

Tabla 4.3: Estimaciones resúmenes para los datos, AUDIT-C y simulados mediante el modelo HSROC con el programa SAS

Los dos meta-análisis para DTA ( datos de AUDIT-C y datos simulados) podemos observar que ambos meta-análisis contienen estudios con ceros en las celdas de falsos positivos (FP) o falsos negativos (FN), pocos números de estudios (14 AUDIT-C y 16 datos simulados) y prevalencia menor al 20 %. Ilustran el problema de convergencia de los algoritmos tradicionales, estos aspectos realistas de los modelos bivariante y HSROC para meta-análisis en DTA, lo abordamos en el estudio de simulación.

## 4.2. Método de Simulación

Se realizó un estudio de simulación para comparar el modelo HSROC con diversas simplificaciones (mediante la eliminación de los parámetros del modelo). Hemos elegido el modelo HSROC porque tiene una mayor flexibilidad a la hora de elegir los parámetros [89]. Las especificaciones de los escenarios fueron diseñadas para reproducir situaciones realistas encontradas en el meta-análisis de estudios de precisión diagnóstica. Se investigó el efecto de estos factores:

- Número de estudios
- Prevalencia de la enfermedad
- Variabilidad entre los estudios en exactitud y umbral
- Asimetría en la curva SROC. Sólo investigamos los métodos que utilizan una distribución binomial

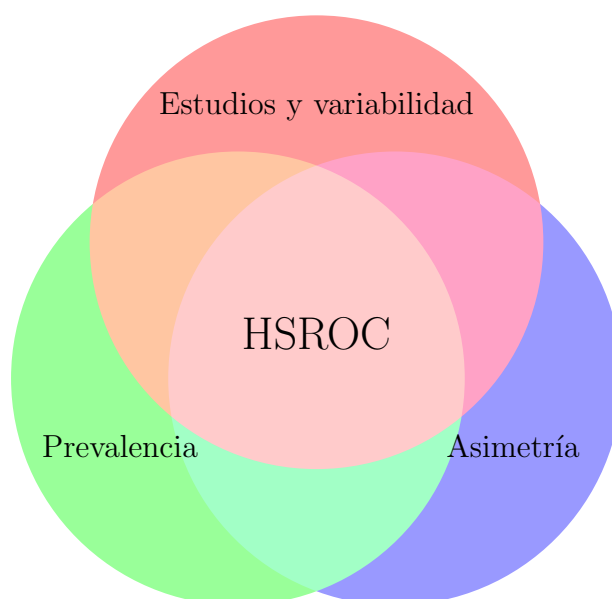


Figura 4.7: Factores que intervienen en la simulación modelo HSROC

## 4.3. Generación de los Datos Simulados

Se investigaron meta-análisis con diferentes números de estudios [86] de manera aleatoria ( $k = 5, 10, 20, 35$ ). El tamaño de un estudio en cada meta-análisis  $n_j$ , fue muestreado

aleatoriamente de una distribución uniforme,  $U(200;2.000)$  dada una prevalencia subyacente  $p$ , los individuos dentro de cada estudio fueron clasificados aleatoriamente como enfermos o no diagnosticados, y se asignó un valor de resultado de prueba continuo,  $x$ , que fue muestreado aleatoriamente, [91].

Para determinar las tablas tetracóricas correspondientes, se utilizó el paquete HSROC del programa estadístico R, mediante el empleo de la función `simdata`. Para crear la tabla de  $2 \times 2$  para cada estudio, los individuos fueron clasificados como verdaderos positivos, falsos negativos, falsos positivos o verdaderos negativos basados en el resultado de la prueba y el estado de la enfermedad. Las prevalencias de los estudios se ajustaron mediante una distribución uniforme con parámetros 0 y 0.25  $U(0;0.25)$ , es decir, la prevalencia de dichos estudios tomó valores menores o iguales a 25 %.

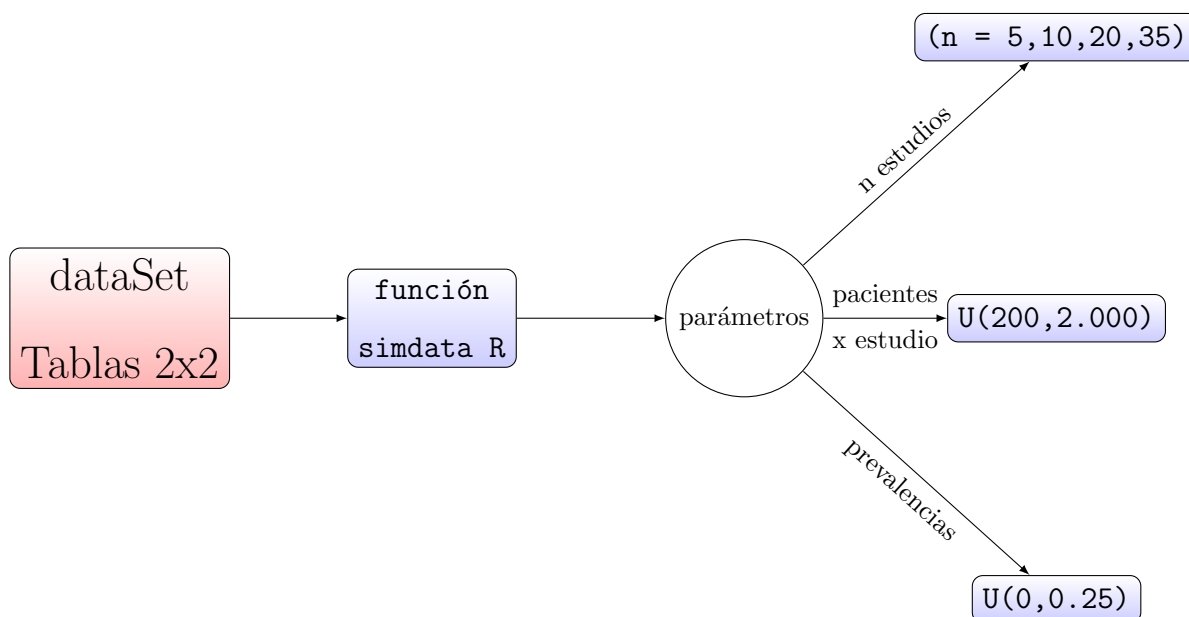


Figura 4.8: Esquema de Generación de las tablas 2x2 para DTA

Generamos 23.110 estudios en 1.200 datasets de meta-análisis independientes para permitir una estimación precisa del rendimiento del modelo, incluso si una gran proporción de modelos no convergen. Para poder ejecutar la función `simdata`, es necesario especificar los parámetros de entradas de la función, es decir es necesario elegir las distribuciones a priori de la función `simdata`, la parametrización inicial se realizó utilizando las siguientes condiciones iniciales, [92]:

$$\text{Dataset} \left\{ \begin{array}{l} N \sim \text{Uniform} [5, 35] \\ n \sim \text{Uniform} [200, 2000] \\ p \sim \text{Uniform} [0,0001, 0,2499] \\ \beta \sim \text{Uniform} [-0,75, 0,75] \\ \Lambda \sim \text{Uniform} [-0,75, 0,75] \\ \sigma_{\alpha}^2 \sim \text{Uniform} [0,0001, 2] \\ \Theta \sim \text{Uniform} [-1,5; 1,5] \\ \sigma_{\theta}^2 \sim \text{Uniform} [0,0001, 2] \end{array} \right.$$

Figura 4.9: Parámetros para Generación de tablas 2x2 para DTA

Adicionalmente, en cada meta-análisis generado, se incluyó una covariable con categorías desde 1 hasta 3, mediante una distribución uniforme  $U(1;3)$ . La Figura 4.10 muestra los conjuntos de datos generados (1200 meta-análisis) en el espacio ROC.

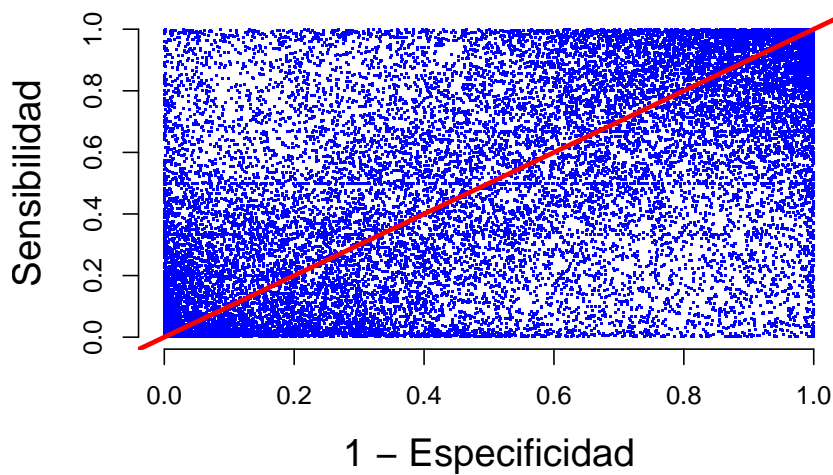


Figura 4.10: Meta-análisis Generados (dataSet), espacio ROC

## 4.4. Modelos jerárquicos ajustados a los datos simulados

A continuación en la Figura 4.11 se presentan los parámetros de los modelos HSROC y Bivariante en el espacio ROC.

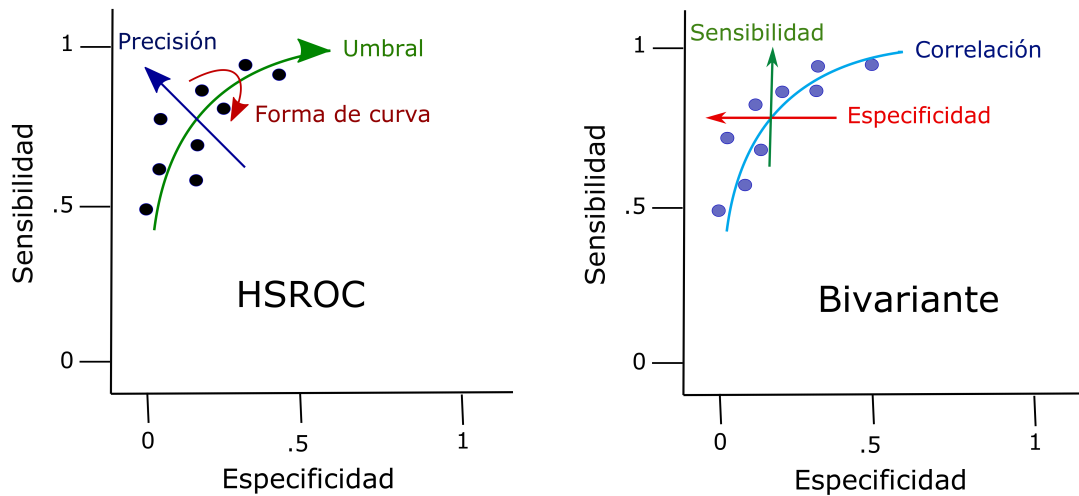


Figura 4.11: Parámetros de los modelos HSROC y Bivariante

En el resto de este trabajo, nos referimos a un modelo HSROC que contiene cinco parámetros. Se ajustaron los siguientes cinco modelos a los 1.200 conjuntos de datos generados (ver Tabla 4.4)

| Num. | Models                                    | $\Lambda$ | $\Theta$ | $\beta$ | $\sigma_{\alpha}^2$ | $\sigma_{\theta}^2$ |
|------|---|-----------|----------|---------|---------------------|---------------------|
| 1    | Modelo completo HSROC                     | ✓         | ✓        | ✓       | ✓                   | ✓                   |
| 2    | Modelo HSROC simétrico                    | ✓         | ✓        |         | ✓                   | ✓                   |
| 3    | Modelo HSROC con umbrales fijos           | ✓         | ✓        | ✓       | ✓                   |                     |
| 4    | Modelo HSROC con precisión fija           | ✓         | ✓        | ✓       |                     | ✓                   |
| 5    | Modelo HSROC con precisión y umbral fijos | ✓         | ✓        | ✓       |                     |                     |

Tabla 4.4: Estructura de los modelos ajustados a los datos simulados

De la Tabla 4.4, se deduce que para un modelo HSROC con precisión y umbrales fijos (modelo 5), solo es necesario identificar a priori las distribuciones para  $\Lambda$ ,  $\Theta$  y  $\beta$ . De manera similar se entiende los cuatro modelos restantes.

Tenga en cuenta que debido a la relación matemática entre el modelo bivariante y HSROC, es posible encontrar las estimaciones de los cinco parámetros del modelo HSROC mediante la función `reitsma` del paquete `mada` de R [93].

## 4.5. Facilitando la Convergencia de los Modelos Jerárquicos

La inferencia del modelo HSROC se fundamenta en una distribución a posteriori (previa) de los parámetros del modelo. Debido a que los modelos que se consideran no son conjugados, es decir, son expresiones de forma cerradas, las distribuciones a posteriori no existen. Luego estos valores a posteriori son estimados mediante una simulación de la distribución a posteriori, utilizando cadenas de Markov y Monte Carlo (MCMC), [87]. Estos valores simulados de las distribuciones a posteriori se utilizan para estimar la distribución marginal de interés, tal como distribuciones a posteriori de los parámetros particulares.

La estrategia es emplear distribuciones (a priori) para la mayoría de los parámetros. Las previas para  $\Lambda$ ,  $\Theta$  y  $\beta$  son seleccionadas de tal manera que la distribución marginal resultante del “**pooled**” de la sensibilidad y especificidad sea aproximadamente normal  $\sim N(0, 1)$  [92].

## 4.6. Evaluación, Convergencia y Estabilidad del Modelo

Debido a que un modelo que cumple un criterio de convergencia puede ser inestable por problemas con la identificabilidad del modelo, evaluamos la convergencia en dos etapas. En primer lugar, verificamos si se cumplió el criterio de convergencia y también si se produjeron las estimaciones correctas para baja prevalencia. En segundo lugar, observan-

do la solución de optimización del vector de gradientes de los parámetros estimados, es decir, cada elemento debe ser lo suficientemente pequeño para indicar que existe un punto estacionario.

Por lo tanto, para que la convergencia sea considerada exitosa, el modelo tuvo que cumplir el criterio de convergencia, producir estimaciones correctas de baja prevalencia, y tener un vector de gradiente muy pequeño, es decir valores cercanos a cero.

## 4.7. Evaluación del Rendimiento de Modelos Meta-analíticos

Se evaluó el rendimiento de los métodos mediante el examen de las estimaciones de las siguientes medidas de precisión diagnóstica: para el modelo Bivariante se calculó el logit de sensibilidad y especificidad, variabilidad de logit de sensibilidad y especificidad, covarianza entre las transformaciones logit de la sensibilidad y especificidad. Para el modelo HSROC según la incorporación de una covariable o no, hemos ajustado cinco modelos HSROC con distintas especificaciones para umbral, precisión y forma de curva.

Para cada escenario usamos solamente los resultados de los meta-análisis que convergieron exitosamente de acuerdo a los criterios anteriormente definidos, la presentación de los resultados es mediante la cobertura de los CI (intervalos de confianza) del 95% y el valor medio en cada escenario.

# Capítulo 5

## Resultados

En este capítulo se muestran los resultados de las simulaciones de los meta-análisis generados. La plataforma tecnológica (software estadísticos y base de datos ) se muestra en el siguiente diagrama general para la ejecución de las rutinas NLMIXED (Modelo jerárquico HSROC y Bivariante).

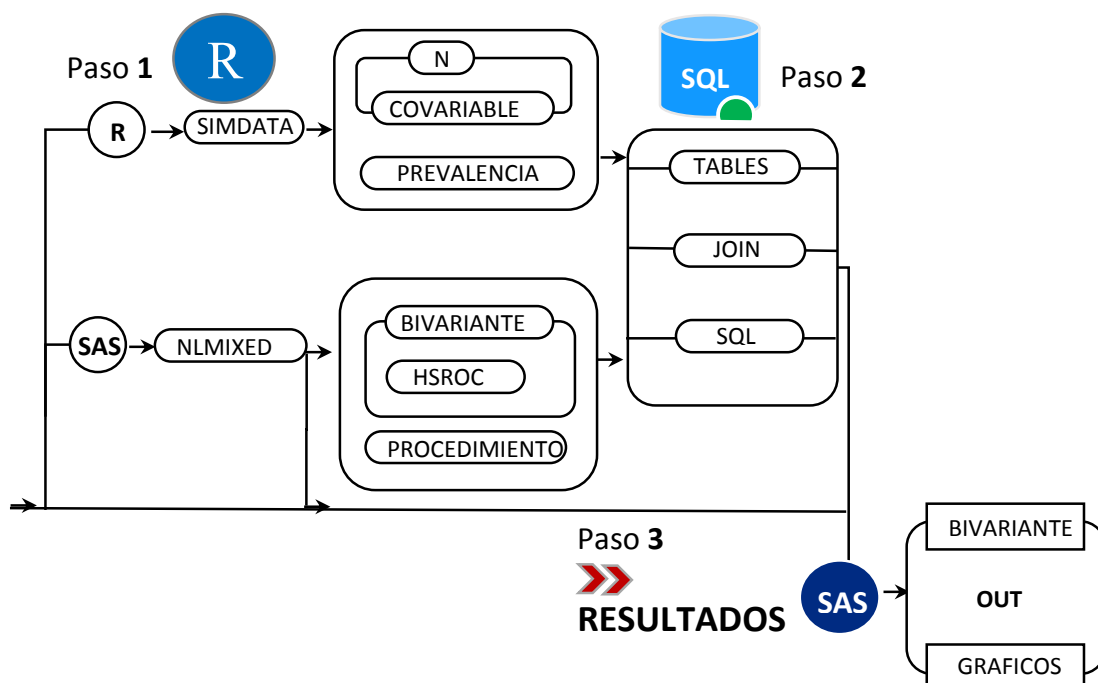


Figura 5.1: Esquema general de trabajo y software requerido para la simulación de dataset, almacenamiento en base de datos y análisis estadístico para la presentación de resultados



## 5.1. Meta-Análisis Simulados

La Tabla 5.1 muestra la distribución de los 1200 meta-análisis generados por rangos de prevalencias y número de estudios en cada meta-análisis. Las prevalencias se eligieron uniformemente con una amplitud del 5%.

| Rangos  | Prevalencias |             |             |             |             | Total |
|---------|--------------|-------------|-------------|-------------|-------------|-------|
|         | [0,0.05)     | [0.05,0.10) | [0.10,0.15) | [0.15,0.20) | [0.20,0.25] |       |
| [5,8)   | 133          | 131         | 123         | 143         | 151         | 681   |
| [8,11)  | 252          | 246         | 244         | 222         | 247         | 1211  |
| [11,14) | 329          | 338         | 315         | 314         | 301         | 1597  |
| [14,17) | 377          | 376         | 355         | 364         | 403         | 1875  |
| [17,20) | 437          | 445         | 417         | 421         | 431         | 2151  |
| [20,23) | 498          | 489         | 514         | 483         | 491         | 2475  |
| [23,26) | 513          | 554         | 500         | 494         | 524         | 2585  |
| [26,29) | 661          | 665         | 627         | 600         | 659         | 3212  |
| [29,32) | 661          | 714         | 729         | 747         | 768         | 3619  |
| [32,35] | 768          | 755         | 719         | 754         | 708         | 3704  |

Tabla 5.1: Resumen de dataset simulados por número de estudios y rangos de prevalencias

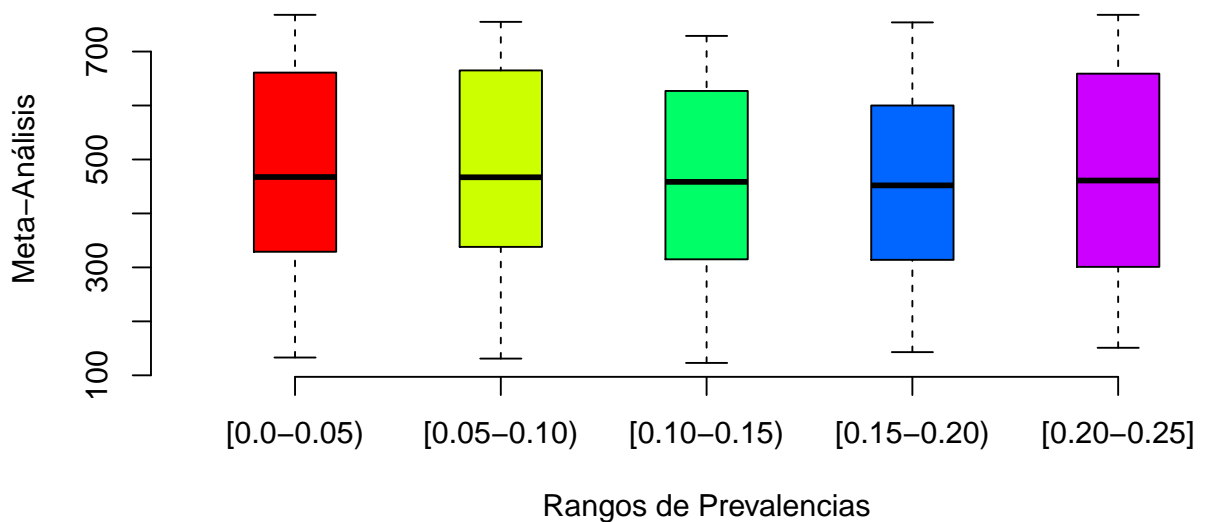


Figura 5.2: Box Plot para sensibilidad y especificidad para datos AuditC

| Estudios | Prevalencias    |                 |                 |                 |                 | Total<br>Estudios | Celdas<br>con Cero | %  |
|----------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|--------------------|----|
|          | [0.00-<br>0.05) | [0.05-<br>0.10) | [0.10-<br>0.15) | [0.15-<br>0.20) | [0.20-<br>0.25] |                   |                    |    |
| [5-8)    | 0.0257          | 0.0737          | 0.1245          | 0.1749          | 0.2245          | 681               | 172                | 25 |
| [8-11)   | 0.0253          | 0.0743          | 0.1248          | 0.1731          | 0.2262          | 1211              | 279                | 23 |
| [11-14)  | 0.0251          | 0.0749          | 0.1249          | 0.1745          | 0.2229          | 1597              | 381                | 23 |
| [14-17)  | 0.0244          | 0.0757          | 0.1257          | 0.1751          | 0.2241          | 1875              | 406                | 21 |
| [17-20)  | 0.0249          | 0.0744          | 0.1246          | 0.1756          | 0.2254          | 2151              | 561                | 26 |
| [20-23)  | 0.0245          | 0.0756          | 0.1241          | 0.1739          | 0.2254          | 2475              | 587                | 23 |
| [23-26)  | 0.0242          | 0.0760          | 0.1246          | 0.1754          | 0.2252          | 2585              | 638                | 24 |
| [26-29)  | 0.0257          | 0.0764          | 0.1255          | 0.1754          | 0.2247          | 3212              | 675                | 21 |
| [29-32)  | 0.0252          | 0.0751          | 0.1254          | 0.1746          | 0.2253          | 3619              | 783                | 21 |
| [32-35]  | 0.0248          | 0.0746          | 0.1243          | 0.1747          | 0.2250          | 3704              | 780                | 21 |

Tabla 5.2: Resumen de dataset simulados por número de estudios y celdas con ceros en las tablas tetracóricas

## 5.2. Tarea de Simulación

La tarea de simulación se dividió en tres etapas:

1. Generación de dataset
2. Almacenamiento de los datos
3. Ejecución del algoritmo para ajuste de los modelos (Bivariante y HSROC) y análisis estadístico de los resultados de la simulación, ver Figura 5.1

Para la primera etapa, se utilizó el paquete estadístico R y la función SIMDATA para generar tablas de respuestas binarias. Para la etapa dos, cada tabla tetracórica creada, así como sus parámetros de generación se almacenan en el motor de base de datos SQL (debido a la gran cantidad de datos simulados, este almacenamiento logra optimizar el uso de la memoria ram del generador de R).

Finalmente, se utilizó el procedimiento NLMIXED de SAS que mediante su ejecución ajusta los datos simulados a los modelos jerárquicos de la Sección 2.3. Así, con los resultados de las estimaciones de los parámetros de los modelos, se volvió a usar el software

SAS para el análisis final y presentación de los resultados de la simulación.

### 5.3. Modelo Bivariante

La Tabla 5.3, muestra las estimaciones de los parámetros mediante la ejecución del procedimiento NLMIXED, para el modelo Bivariante y la incorporación o no de una covariable. Los resultados de cada estimación de los parámetros se presentan en intervalos de confianza al 95 % para la media.

Es importante destacar que la incorporación de una covariable al modelo proporciona mejor ajuste a la modelización (observar  $-2 \text{ LogLikelihood}$ ). Es importante notar que la media de las estimaciones de los parámetros con covariables es menor que las estimaciones sin covariable, ya que la inclusión de la misma permite distinguir entre los estudios diferentes subgrupos.

| <b>Modelo Bivariante</b>          |                                   |                 |
|-----------------------------------|-----------------------------------|-----------------|
|                                   | Sin Covariable                    | Con Covariable  |
|                                   | Intervalo de Confianza 95 %       |                 |
| <b>Parámetros<br/>Estimados</b>   | $\mu_A$ : (-4.48,-3.56)           | (-4.11,-3.27)   |
|                                   | $\mu_B$ : (-0.12,-0.11)           | (-0.30,0.14)    |
|                                   | $\sigma_A^2$ : (-0.16,-0.07)      | (-0.14,0.10)    |
|                                   | $\sigma_B^2$ : (5.25,-12.85)      | (2.58,3.90)     |
|                                   | $\sigma_{AB}$ : (7.05,-8.39)      | (6.52,7.53)     |
|                                   | Intervalo de Confianza 95 %       |                 |
| <b>Estadísticas<br/>de Ajuste</b> | $-2\text{logl}$ : (372.23,392.42) | (370.29,390.5)  |
|                                   | Aic: (382.23,402.42)              | (384.29,404.5)  |
|                                   | Aicc: (385.07,405.03)             | (390.99,410.57) |
|                                   | Bic: (386.52,407)                 | (390.3,410.9)   |

Tabla 5.3: Estadística de bondad de ajuste para la convergencia, estimación de los parámetros a posteriori y sus intervalos de confianza para el modelo Bivariante incorporando o no una covariable

## 5.4. Modelo HSROC

La Tabla 5.4, denota las estimaciones de los parámetros mediante la ejecución de NLMIXED para el ajuste HSROC (modelo completo) y la incorporación o no de una covariable. Así, a partir del modelo completo, se generaron cinco escenarios dependiendo de la inclusión de la covariable a los parámetros del modelo. Por cada escenario en particular se presentan los resultados en intervalos de confianza al 95 % para la media; es importante destacar que la incorporación de una covariable en el modelo HSROC completo (precisión, umbral y forma de la curva), mejoraron las estimaciones del ajuste y la convergencia.

|                               |                                   | Modelo HSROC                |                |                    |                            |
|-------------------------------|-----------------------------------|-----------------------------|----------------|--------------------|----------------------------|
| Sin Covariable                |                                   | Umbral                      | Con Covariable |                    |                            |
|                               |                                   |                             | Precisión      | Precisión y Umbral | Precisión y Umbral y Forma |
|                               |                                   | Intervalo de Confianza 95 % |                |                    |                            |
| <b>Parámetros Estimados</b>   | $\Lambda$ : (-0.37,0.26)          | (-0.34,0.29)                | (-0.29,0.35)   | (-0.33,0.31)       | (-0.33,0.30)               |
|                               | $\theta$ : (-0.31,0.03)           | (-0.32,0.03)                | (-0.28,0.07)   | (-0.30,0.04)       | (-0.30,0.05)               |
|                               | $\beta$ : (0.99,1.17)             | (0.99,1.17)                 | (1.00,1.18)    | (0.98,1.17)        | (0.97,1.16)                |
|                               | $\sigma_{\alpha}^2$ : (2.49,3.52) | (2.58,3.90)                 | (2.40,3.42)    | (2.50,3.53)        | (2.49,3.52)                |
|                               | $\sigma_{\theta}^2$ : (1.32,1.68) | (1.43,1.83)                 | (1.35,1.75)    | (1.38,1.79)        | (1.34,1.73)                |
|                               |                                   | Intervalo de Confianza 95 % |                |                    |                            |
| <b>Estadísticas de ajuste</b> | -2logl: (1027,1161)               | (1012,1145)                 | (1022,1155)    | (1014,1147)        | (1019,1153)                |
|                               | Aic: (1037,1171)                  | (1024,1157)                 | (1034,1167)    | (1028,1161)        | (1035,1169)                |
|                               | Aicc: (1040,1174)                 | (1029,1161)                 | (1039,1171)    | (1035,1168)        | (1048,1182)                |
|                               | Bic: (1041,1175)                  | (1029,1162)                 | (1039,1172)    | (1033,1166)        | (1042,1176)                |

Tabla 5.4: Estadística de bondad de ajuste para la convergencia, estimación de los parámetros a posteriori y sus intervalos de confianza para el modelo HSROC incorporando o no una covariable (en el umbral, la precisión y la forma de la curva).

## 5.5. Estimación

En la generación de los estudios simulados existen celdas con ceros como se muestra en la Tabla 5.2. Sin embargo, esto no incide mayormente en la convergencia de los modelos. La no convergencia de los modelos HSROC se suscitó con mayor frecuencia cuando la varianza estimada del parámetro  $\alpha$  o  $\theta$  eran muy cercanas a cero. Otro punto importante que influye en la inestabilidad de los ajustes para el modelo HSROC, es cuando la relación de la varianza del parámetro  $\alpha$  es dos veces o más la varianza del parámetro  $\theta$ , especialmente en el modelo HSROC completo.

En la Tabla 5.5, se muestra el total de meta-análisis que resultaron no estables para el procedimiento NLMIXED, según rango de prevalencia, número de estudios y ajuste. En otras palabras, el 90 % de meta-análisis convergieron para el ajuste Bivariante y 70 % para la modelización HSROC. Con relación a la no convergencia se obtuvieron 112 y 369 meta-análisis que resultaron inestables para los modelos Bivariante y HSROC respectivamente.

| Estudios | No Convergencia |       |             |       |             |       |       |       |       |      |   |  |
|----------|-----------------|-------|-------------|-------|-------------|-------|-------|-------|-------|------|---|--|
|          | Prevalencias    |       |             |       |             |       |       |       | Total |      | % |  |
|          | [0.05,0.10)     |       | [0.10,0.15) |       | [0.15,0.20) |       |       |       |       |      |   |  |
| Bvte.    | Hsroc           | Bvte. | Hsroc       | Bvte. | Hsroc       | Bvte. | Hsroc | Bvte. | Hsroc |      |   |  |
| [5-8)    | 7               | 12    | 16          | 28    | 6           | 9     | 29    | 49    | 0.35  | 0.48 |   |  |
| [8-11)   | 2               | 12    | 14          | 25    | 1           | 3     | 17    | 40    | 0.14  | 0.3  |   |  |
| [11-14)  | 1               | 3     | 13          | 43    | 0           | 1     | 14    | 47    | 0.12  | 0.38 |   |  |
| [14-17)  | 3               | 6     | 9           | 25    | 1           | 4     | 13    | 35    | 0.12  | 0.29 |   |  |
| [17-20)  | 0               | 1     | 10          | 34    | 1           | 1     | 11    | 36    | 0.1   | 0.31 |   |  |
| [20-23)  | 0               | 3     | 6           | 26    | 1           | 1     | 7     | 30    | 0.06  | 0.26 |   |  |
| [23-26)  | 1               | 2     | 8           | 29    | 0           | 5     | 9     | 36    | 0.09  | 0.35 |   |  |
| [26-29)  | 0               | 2     | 7           | 28    | 2           | 1     | 9     | 31    | 0.08  | 0.27 |   |  |
| [29-32)  | 0               | 0     | 2           | 35    | 0           | 1     | 2     | 36    | 0.02  | 0.33 |   |  |
| [32-35]  | 1               | 1     | 0           | 26    | 0           | 2     | 1     | 29    | 0.01  | 0.27 |   |  |
|          |                 |       |             |       |             |       | 112   | 369   | 0.1   | 0.32 |   |  |

Tabla 5.5: meta-análisis que resultaron no estables para el procedimiento NLMIXED, según rango de prevalencia, número de estudios y ajuste

La Figura 5.3, muestra tres curvas SROC seleccionadas de manera aleatoria en la simulación. La primera curva tiene una precisión DOR de 13.5452 y una covarianza ( $\sigma_{SenSpec}$ ) de 0.0948, la segunda curva tiene una precisión DOR de 4.0526 y una covarianza ( $\sigma_{SenSpec}$ ) de 4.68E-05 y la tercera curva tiene una precisión DOR de 8.1024 y una covarianza ( $\sigma_{SenSpec}$ ) de 0.0049. De acuerdo al análisis de los valores de los parámetros de precisión y covarianzas, el mejor modelo que describe esta relación en las curvas SROC es el modelo Bivariante.

### MODELO BIVARIANTE

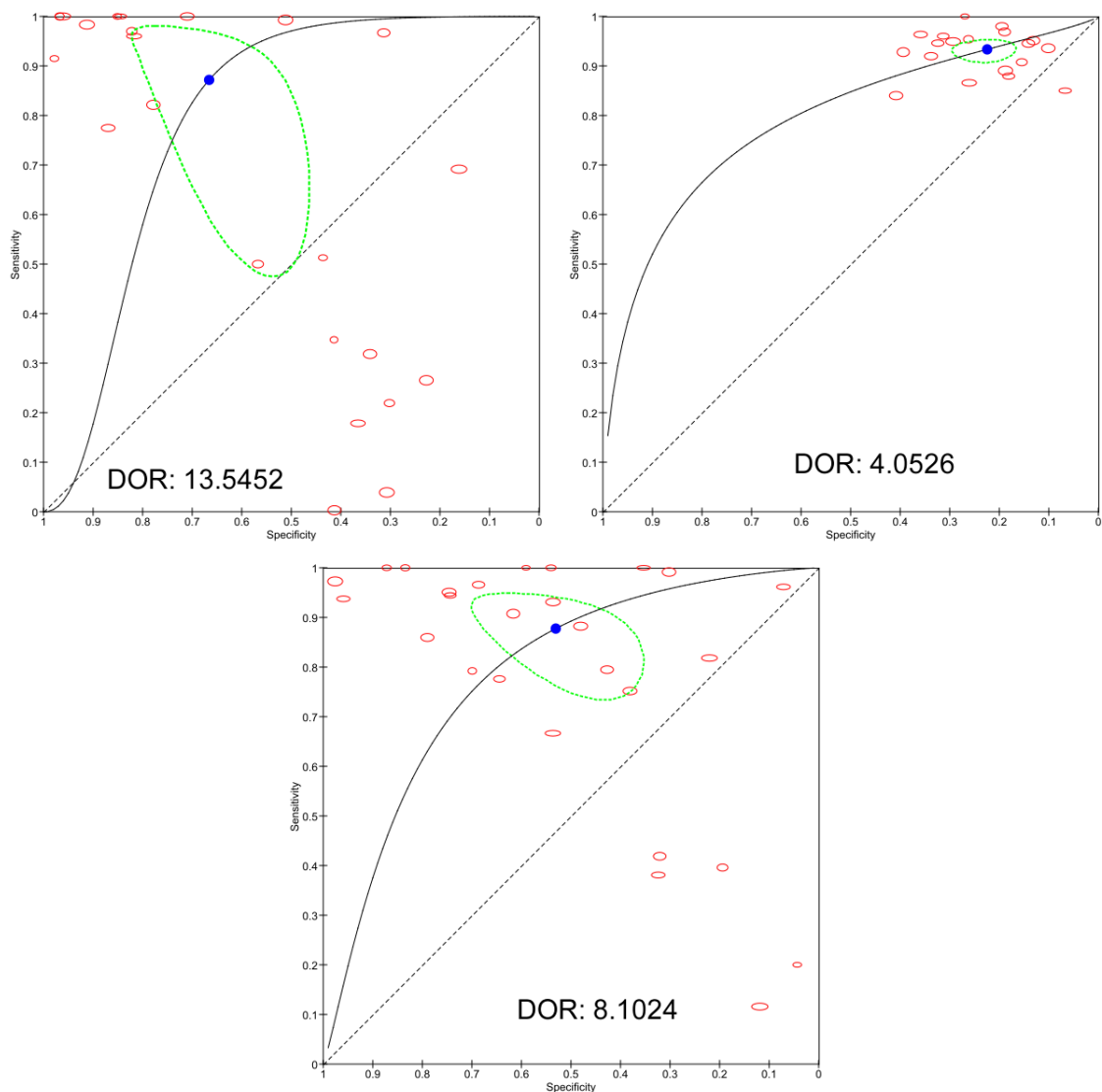


Figura 5.3: Curvas SROC seleccionadas aleatoriamente de acuerdo a los rangos de prevalencia [0.05 , 0.10), [0.10 , 0.15) y [0.15 , 0.20) respectivamente

La Figura 5.4, muestra tres curvas SROC seleccionadas de manera aleatoria en la simulación. Los valores de los parámetros de umbral, precisión y forma de la curva para la primera curva(1) son:  $\Lambda : 0,7368$   $\Theta : 0,4432$   $\beta : 0,8065$  respectivamente. Para la segunda curva(2) son:  $\Lambda : 1,5948$   $\Theta : 0,5343$   $\beta : -0,0659$  respectivamente. Finalmente para la tercera curva(3) son:  $\Lambda : 1,0504$   $\Theta : 0,6215$   $\beta : -0,6055$ . De acuerdo al análisis de los valores de los parámetros de precisión y covarianzas, el mejor modelo que describe esta relación en las curvas SROC es el modelo HSROC.

### MODELO HSROC

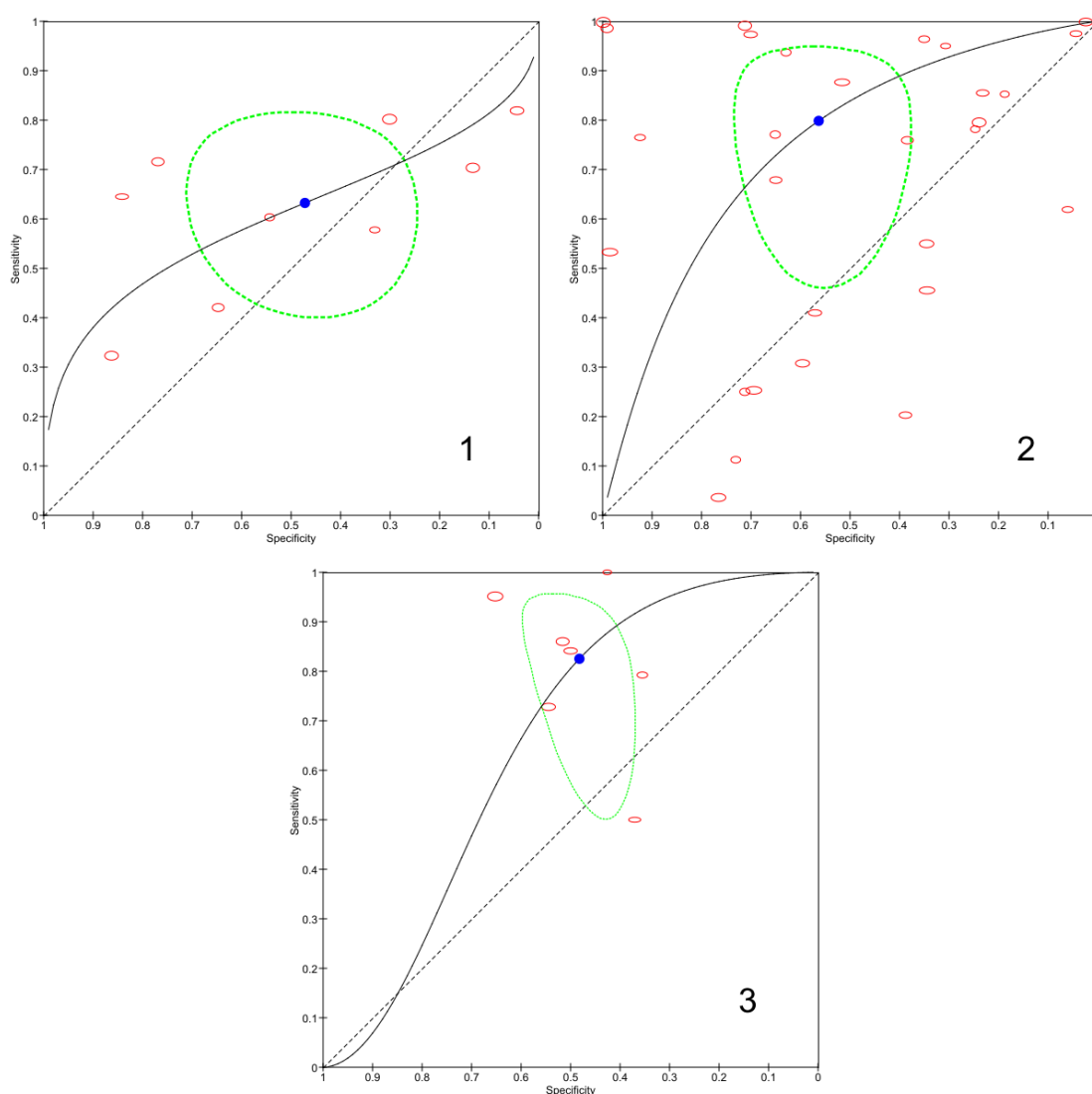


Figura 5.4: Curvas SROC seleccionadas aleatoriamente de acuerdo a los rangos de prevalencia  $[0.05, 0.10)$ ,  $[0.10, 0.15)$  y  $[0.15, 0.20)$  respectivamente

Otro aspecto de no convergencia se dio cuando la varianza de los efectos aleatorios del umbral y la precisión son muy pequeñas o cercanas a cero, esto es entendible pues el modelo HSROC fue diseñado para modelizar los efectos aleatorios.

En la Figura 5.5, se observa como los valores de los logaritmos del DOR y covarianza de sensibilidad y especificidad para el modelo bivalente se ajustan a una distribución normal.

Esto era de esperarse ya que estas medidas representan los efectos aleatorios y en la modelización de dicho ajuste los efectos aleatorios se distribuyen de acuerdo a una normal.

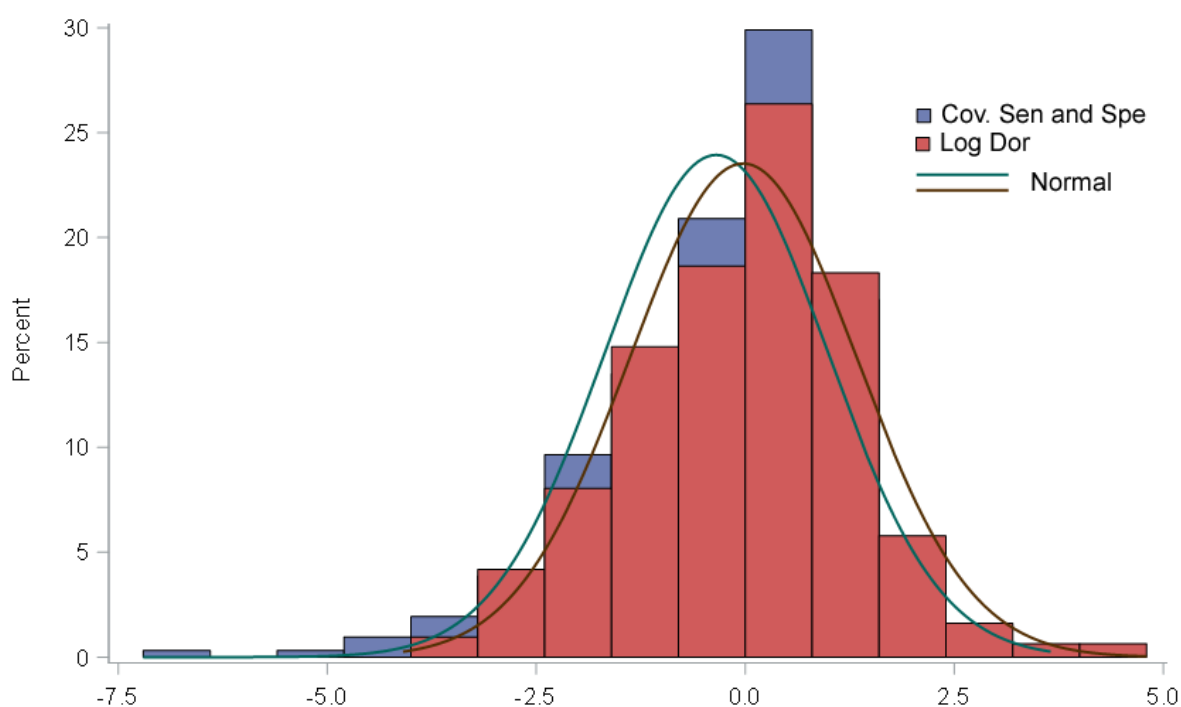


Figura 5.5: Ajuste normal para la covarianza del Logit (Sensibilidad y Especificidad), y el log (DOR) para el modelo Bivalente.

Mediante un gráfico de paneles, se puede establecer que existe una dependencia lineal creciente entre la medida de ajuste ( $-2\text{LogLikelihood}$ ) y número de estudios en el meta-análisis para el enfoque bivalente, ver Figura 5.6.



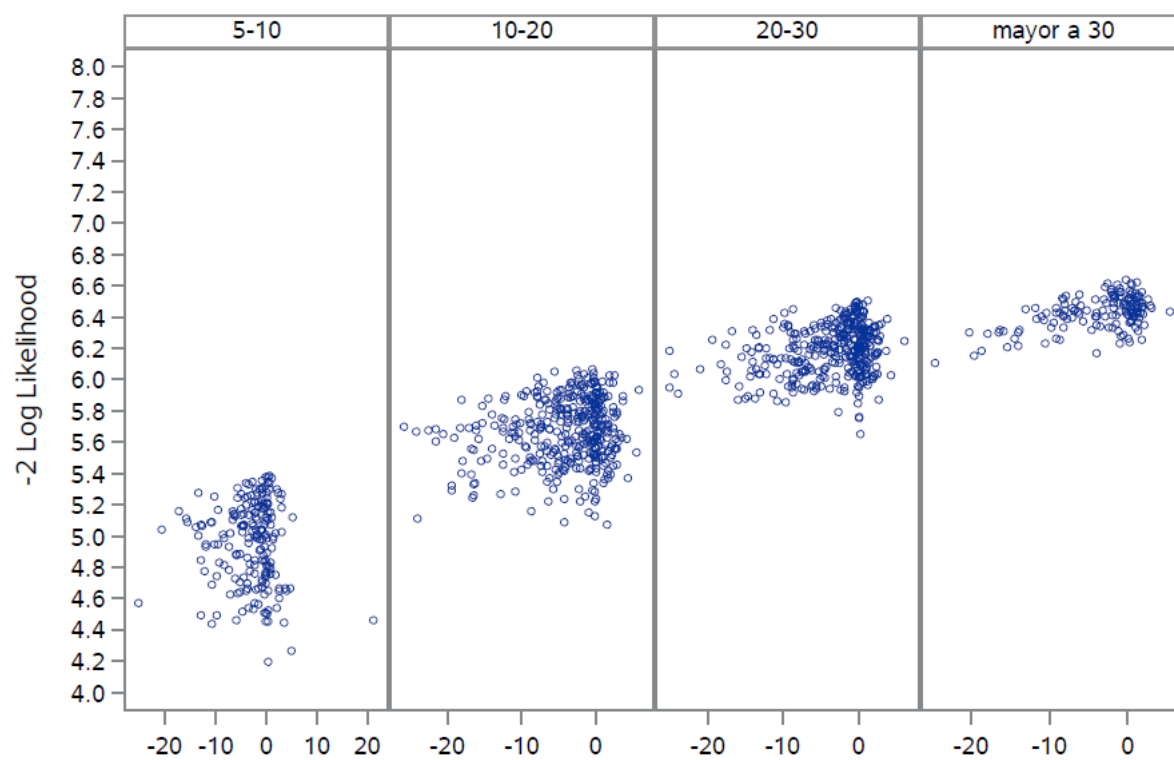


Figura 5.6: Covarianza de Sensibilidad y Estabilidad vs. Ajuste (-2 Probabilidad Logarítmica) Número de Estudios en Meta-análisis

# Conclusiones

En esta sección escribimos las conclusiones referentes al trabajo de valoración de los modelos jerárquicos bayesianos para meta-análisis para pruebas de precisión diagnósticas, mediante el uso de simulaciones.

1. La exhaustiva revisión bibliográfica realizada ha puesto de manifiesto que el modelo jerárquico bivalente, que modeliza los logit de sensibilidad y especificidad y el HS-ROC que modeliza la relación entre sensibilidad y especificidad a través de distintos puntos de corte, no han sido debidamente analizados en las publicaciones científicas, ni en los diferentes programas estadísticos.
2. La eficiencia de los algoritmos meta-analíticos, está determinada no solo por el tiempo de convergencia y el uso de memoria RAM utilizada, sino también por el número de estudios incluidos en el meta-análisis, los verdaderos positivos, los verdaderos negativos, los falsos positivos y los falsos negativos, de las tablas tetracóricas.
3. El análisis de simulación basado en 1200 meta-análisis que tienen en cuenta distintas parametrizaciones de los modelos jerárquicos variando precisión, puntos de corte, forma de la curva ROC, variabilidad entre y dentro de los estudios, covariables, prevalencias y número de estudios, nos ha permitido establecer condiciones de convergencia y eficiencia de los algoritmos meta-analíticos.
4. La plataforma tecnológica empleada en la simulación, consideró un esquema general de integración estadística computacional, que ha permitido detectar las condiciones de eficiencia y convergencia de los algoritmos, mediante generación de los meta-análisis con el software R de código libre y el almacenamiento de los meta-análisis simulados, en una base de datos estructurada. Así mismo nos ha permitido ejecutar el algoritmo meta-analítico, a través del software especializado SAS, que usa la modelización no lineal mixta generalizada mediante el procedimiento NLMIXED.

5. La tasa de convergencia de los meta-análisis simulados fue del 70% en situaciones clínicamente probables; en el resto de los meta-análisis, o bien el gradiente de la función de verosimilitud no encontró una solución óptima, ó bien los resultados de las estimaciones de los parámetros, generaron situaciones clínicamente improbables.
6. Las simulaciones realizadas en esta investigación, ponen de manifiesto, que en condiciones de heterogeneidad moderada el modelo más adecuado es el bivariante, modelo con el que solo se estima un pool de sensibilidad y especificidad; en condiciones de alta heterogeneidad el modelo más adecuado es el HSROC. Este modelo es más complejo pero permite no solo estimar las medidas resumen sino también hacer inferencia en relación a las predicciones.

# Bibliografía

- [1] M. M. Leeflang, J. J. Deeks, Y. Takwoingi, and P. Macaskill, “Cochrane diagnostic test accuracy reviews,” *Systematic Reviews*, vol. 2, no. 1, Oct. 2013.
- [2] L. E. Moses, D. Shapiro, and B. Littenberg, “Combining independent studies of a diagnostic test into a summary roc curve: Data-analytic approaches and some additional considerations,” *Statistics in Medicine*, vol. 12, no. 14, pp. 1293–1316, Jul. 1993.
- [3] J. B. Reitsma, A. S. Glas, A. W. S. Rutjes, R. J. P. M. Scholten, P. M. Bossuyt, and A. H. Zwinderman, “Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews,” *Journal of Clinical Epidemiology*, vol. 58, no. 10, pp. 982–990, Oct. 2005.
- [4] C. M. Rutter and C. A. Gatsonis, “A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations,” *Statistics in Medicine*, vol. 20, no. 19, pp. 2865–2884, 2001.
- [5] P. Macaskill, C. Gatsonis, J. Deeks, R. Harbord, Y. Takwoingi, “Cochrane handbook for systematic reviews of diagnostic test accuracy,” Version 0.9. 0. London: The Cochrane Collaboration, 2010.
- [6] P. Doebler, P.-C. Bürkner, and G. Rucker, “Statistical Packages for Diagnostic Meta-Analysis and Their Application,” *Diagnostic Meta-Analysis*, pp. 161–181, 2018.
- [7] Y. Takwoingi, J. J. Deeks, “METADAS: a SAS macro for meta-analysis of diagnostic accuracy studies,” In *Methods for evaluating medical tests symposium*, University of Birmingham, 2008.
- [8] P. Doebler, H. Holling, “Meta-analysis of diagnostic accuracy with mada.” Retrieved at: <https://cran.rproject.org/web/packages/mada/vignettes/mada.pdf>, 2015.

- 
- [9] I. Schiller, N. Dendukuri, “HSROC: An R package for Bayesian meta-analysis of diagnostic test accuracy.” 2015.
- [10] P. E. Verde, “bamdit: An R Package for Bayesian Meta-Analysis of Diagnostic Test Data,” *Journal of Statistical Software*, vol. 86, no. 10, 2018.
- [11] J. Guo and A. Riebler, “meta4diag: Bayesian Bivariate Meta-Analysis of Diagnostic Test Studies for Routine Practice,” *Journal of Statistical Software*, vol. 83, no. 1, 2018.
- [12] S. Steinhauser, M. Schumacher, and G. Rücker, “Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies,” *BMC Medical Research Methodology*, vol. 16, no. 1, Aug. 2016.
- [13] H. Huang, “Metatron: meta-analysis for classification data and correction to imperfect reference,” R package version 0.1-1.2014, 2018.
- [14] J. Wang, AH. Zwinderman, MM. Leeftang “The area under the ROC curve may be a biased performance measure for metaanalysis of diagnostic accuracy studies. A simulation study,” *Global Evidence Summit, Cape Town, South Africa 2017*.
- [15] M. Davidian, D. M. Giltinan “Nonlinear Models for Repeated Measurement Data,” New York, Chapman & Hall, 1995.
- [16] E. F. Vonesh, V. M. Chinchilli “Linear and Nonlinear Models for the Analysis of Repeated Measurements,” New York, Marcel Dekker, 1997.
- [17] J. C. Pinheiro, D. M. Bates “Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model,” *Journal of Computational and Graphical Statistics* 4:12–35, 1995.
- [18] M. Davidian, A. R. Gallant “The Nonlinear Mixed Effects Model with a Smooth Random Effects Density,” *Biometrika* 80:475–488, 1993.
- [19] S. L. Beal, L. B. Sheiner “Estimating Population Kinetics,” *CRC Critical Reviews in Biomedical, Engineering* 8:195–222, 1982.
- [20] S. L. Beal, L. B. Sheiner “Heteroscedastic Nonlinear Regression,” *Technometrics* 30:327–338, 1988.

- [21] L. B. Sheiner, S. L. Beal “Pharmacokinetic Parameter Estimates from Several Least Squares Procedures: Superiority of Extended Least Squares,” *Journal of Pharmacokinetics and Biopharmaceutics* 13:185–201, 1985.
- [22] M. J. Lindstrom, D. M. Bates “Nonlinear Mixed Effects Models for Repeated Measures Data,” *Biometrics* 46:673–687, 1990.
- [23] S. L. Beal, L. B. Sheiner “NONMEM User’s Guide,” San Francisco: NONMEM Project Group, University of California, San Francisco, 1992.
- [24] R. D. Wolfinger “Laplace’s Approximation for Nonlinear Mixed Models,” *Biometrika* 80:791–795. R. D., 1993.
- [25] E. F. Vonesh “Nonlinear Models for the Analysis of Longitudinal Data,” *Statistics in Medicine* 11:1929–1954, 1992.
- [26] E. F. Vonesh “A Note on Laplace’s Approximation for Nonlinear Mixed-Effects Models,” *Biometrika* 83:447–452, 1996.
- [27] E. F. Vonesh, V. M. Chinchilli “Linear and Nonlinear Models for the Analysis of Repeated Measurements,” New York: Marcel Dekker, 1997.
- [28] R. D. Wolfinger, X. Lin “Two Taylor-Series Approximation Methods for Nonlinear Mixed Models,” *Computational Statistics and Data Analysis* 25:465–490, 1997.
- [29] D. A. Harville, R. W. Mee “A Mixed-Model Procedure for Analyzing Ordered Categorical Data,” *Biometrics* 40:393–408, 1984.
- [30] R. Stiratelli, N. M. Laird, J. H. Ware “Random Effects Models for Serial Observations with Binary Response,” *Biometrics* 40:961–971, 1984.
- [31] A. R. Gilmour, R. D. Anderson, A. L. Rae “The Analysis of Binomial Data by Generalized Linear Mixed Model.” *Biometrika* 72:593–599, 1985.
- [32] H. Goldstein “Nonlinear Multilevel Models, with an Application to Discrete Response Data,” *Biometrika* 78:45–51, 1991.
- [33] R. Schall “Estimation in Generalized Linear Models with Random Effects,” *Biometrika* 78:719–727, 1991.

- [34] B. Engel, A. Keen “A Simple Approach for the Analysis of Generalized Linear Mixed Models,” vol. LWA-92-6. Wageningen, Netherlands: Agricultural Mathematics Group (GLW-DLO), 1992.
- [35] N. E. Breslow, D. G. Clayton “Approximate Inference in Generalized Linear Mixed Models.” *Journal of the American Statistical Association* 88:9–25, 1993.
- [36] R. D. Wolfinger, M. A. O’Connell “Generalized Linear Mixed Models: A Pseudo-likelihood Approach.” *Journal of Statistical Computation and Simulation* 48:233–243, 1993.
- [37] C. A. McGilchrist “Estimation in Generalized Mixed Models.” *Journal of the Royal Statistical Society, Series B* 56:61–69, 1994.
- [38] G. Rodriguez, N. Goldman “An Assessment of Estimation Procedures for Multilevel Models with Binary Response,” *Journal of the Royal Statistical Society, Series A* 158:73–89, 1995.
- [39] X. Lin, N. E. Breslow “Bias Correction in Generalized Linear Mixed Models with Multiple Components of Dispersion,” *Journal of the American Statistical Association* 91:1007–1016, 1996.
- [40] D. A. Pierce, B. R. “Sands Extra-Bernoulli Variation in Binary Data,” Technical Report 46, Department of Statistics, Oregon State University, 1975.
- [41] D. A. Anderson, M. Aitkin “Variance Component Models with Binary Response: Interviewer Variability,” *Journal of the Royal Statistical Society, Series B* 47:203–210, 1985.
- [42] D. Hedeker, R. D. Gibbons “A Random Effects Ordinal Regression Model for Multilevel Analysis.” *Biometrics* 50:933–944, 1994.
- [43] E. A. C. Crouch, D. Spiegelman “The Evaluation of Integrals of the Form  $\int_{-\infty}^{\infty} f(t)\exp(-t^2)dt$ : Application to Logistic-Normal Models,” *Journal of the American Statistical Association* 85:464–469, 1990.
- [44] N. T. Longford “Logistic Regression with Random Coefficients,” *Computational Statistics and Data Analysis* 17:1–15, 1994.

- [45] C. E. McCulloch “Maximum Likelihood Variance Components Estimation for Binary Data,” *Journal of the American Statistical Association* 89:330–335, 1994.
- [46] Q. Liu, D. A. Pierce “A Note on Gauss-Hermite Quadrature,” *Biometrika* 81:624–629, 1994.
- [47] P. J. Diggle, K.-Y. Liang, S. L. Zeger “Analysis of Longitudinal Data,” Oxford: Clarendon Press, 1994.
- [48] D. J. Roe “Comparison of Population Pharmacokinetic Modeling Methods Using Simulated Data: Results from the Population Modeling Workgroup,” *Statistics in Medicine* 16:1241–1262, 1997.
- [49] L. Yuh, S. L. Beal, M. Davidian, F. Harrison, A. Hester, K. Kowalski, E. Vonesh, R. D. Wolfinger “Population Pharmacokinetic/Pharmacodynamic Methodology and Applications: A Bibliography,” *Biometrics* 50:566–575, 1994.
- [50] M. J. Lindstrom, D. M. Bates “Nonlinear Mixed Effects Models for Repeated Measures Data,” *Biometrics* 46:673–687, 1990.
- [51] N. E. Breslow, D. G. Clayton “Approximate Inference in Generalized Linear Mixed Models,” *Journal of the American Statistical Association* 88:9–25, 1993.
- [52] R. D. Wolfinger, M. A. O’Connell “Generalized Linear Mixed Models: A Pseudo-likelihood Approach,” *Journal of Statistical Computation and Simulation* 48:233–243, 1993.
- [53] E. F. Vonesh, V. M. Chinchilli “Linear and Nonlinear Models for the Analysis of Repeated Measurements,” New York: Marcel Dekker, 1997.
- [54] A. T. Galecki “NLMEM: New SAS/IML Macro for Hierarchical Nonlinear Models,” *Computer Methods and Programs in Biomedicine* 55:207–216, 1998.
- [55] N. R. Draper, H. Smith “Applied Regression Analysis,” 2nd ed. New York: John Wiley & Sons, 1981.
- [56] M. J. Lindstrom, D. M. Bates “Nonlinear Mixed Effects Models for Repeated Measures Data,” *Biometrics* 46:673–687, 1990.



- [57] J. C. Pinheiro, D. M. Bates “Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model,” *Journal of Computational and Graphical Statistics* 4:12–35, 1995.
- [58] P. J. Beitler, J. R. Landis “A Mixed-Effects Model for Categorical Data,” *Biometrics* 41:991–1000, 1985.
- [59] P. Billingsley “Probability and Measure,” 2nd ed. New York: John Wiley & Sons, 1986.
- [60] C. Cox “Delta Method,” In *Encyclopedia of Biostatistics*, edited by P. Armitage, and T. Colton. New York: John Wiley & Sons, 1998.
- [61] G. H. Golub, J. H. Welsch “Calculation of Gaussian Quadrature Rules,” *Mathematical Computing* 23:221–230, 1969.
- [62] M. Abramowitz, I. A. Stegun “Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,” 10th printing. New York: Dover, eds. 1972.
- [63] J. E. Dennis, D. M. Gay, R. E. Welsch “An Adaptive Nonlinear Least-Squares Algorithm,” *ACM Transactions on Mathematical Software* 7:348–368, 1981.
- [64] D. M. Gay “Subroutines for Unconstrained Minimization,” *ACM Transactions on Mathematical Software* 9:503–524, 1983.
- [65] J. J. Moré, D. C. Sorensen “Computing a Trust-Region Step,” *SIAM Journal on Scientific and Statistical Computing* 4:553–572, 1983.
- [66] R. Fletcher “Practical Methods of Optimization,” 2nd ed. Chichester, UK: John Wiley & Sons, 1987.
- [67] J. E. Dennis, H. H. W. Mei “Two New Unconstrained Optimization Algorithms Which Use Function and Gradient Values,” *Journal of Optimization Theory and Applications* 28:453–482, 1979.
- [68] M. J. D. Powell “Restart Procedures for the Conjugate Gradient Method,” *Mathematical Programming* 12:241–254, 1977.
- [69] E. M. L. Beale “A Derivation of Conjugate Gradients,” In *Numerical Methods for Nonlinear Optimization*, edited by F. A. Lootsma, 39–43. London: Academic Press, 1972.

- [70] J. E. Dennis, R. B. Schnabel “Numerical Methods for Unconstrained Optimization and Nonlinear Equations,” Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [71] K. Schittkowski, J. Stoer “A Factorization Method for the Solution of Constrained Linear Least Squares Problems Allowing Subsequent Data Changes,” *Numerische Mathematik* 31:431–463, 1979.
- [72] J. J. Moré “The Levenberg-Marquardt Algorithm: Implementation and Theory,” In *Lecture Notes in Mathematics*, vol. 30, edited by G. A. Watson, 105–116. Berlin: Springer-Verlag, 1978.
- [73] A. R. Gallant “Nonlinear Statistical Models,” New York: John Wiley & Sons 1987.
- [74] J. S. Cramer “Econometric Applications of Maximum Likelihood Methods,” Cambridge: Cambridge University Press, 1986
- [75] R. M. Pringle, A. A. Rayner “Generalized Inverse Matrices with Applications to Statistics,” New York: Hafner Publishing, 1971.
- [76] J. G. Booth, J. P. Hobert “Standard Errors of Prediction in Generalized Linear Mixed Models,” *Journal of the American Statistical Association* 93:262–272, 1998.
- [77] C. M. Hurvich, C.-L. Tsai “Regression and Time Series Model Selection in Small Samples,” *Biometrika* 76:297–307, 1989.
- [78] K. P. Burnham, D. R. Anderson “Model Selection and Inference: A Practical Information-Theoretic Approach,” New York: Springer-Verlag, 1998.
- [79] R. J. Serfling “Approximation Theorems of Mathematical Statistics,” New York: John Wiley & Sons, 1980.
- [80] C. S. Weil “Selection of the Valid Number of Sampling Units and Consideration of Their Combination in Toxicological Studies Involving Reproduction, Teratogenesis, or Carcinogenesis,” *Food and Cosmetic Toxicology* 8:177–182, 1970.
- [81] D. A. Williams “The Analysis of Binary Responses from Toxicological Experiments Involving Reproduction and Teratogenicity,” *Biometrics* 31:949–952, 1975.
- [82] Y. Ochi, R. L. Prentice “Likelihood Inference in a Correlated Probit Regression Model,” *Biometrika* 71:531–543, 1984.

- [83] F. Ezzet, J. Whitehead “A Random Effects Model for Ordinal Responses from a Crossover Trial,” *Statistics in Medicine* 10:901–907, 1991.
- [84] D. P. Gaver, I. G. O’Muircheartaigh “Robust Empirical Bayes Analysis of Event Rates,” *Technometrics* 29:1–15, 1987.
- [85] D. Draper “Discussion of the Paper by Lee and Nelder,” *Journal of the Royal Statistical Society, Series B* 58:662–663, 1996.
- [86] L. E. Eberly, G. Casella “Bayesian estimation of the number of unseen studies in a meta-analysis,” *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-*, 15(4), 477-494, 1999.
- [87] A. E. Gelfand, A. F. Smith “Sampling-based approaches to calculating marginal densities,” *Journal of the American statistical association*, 85(410),398-409 1990.
- [88] R. G. Cowell, P. Dawid, S. L. Lauritzen, D. J. Spiegelhalter “Probabilistic networks and expert systems: Exact computational methods for Bayesian networks,” *Springer Science & Business Media*, 2006.
- [89] L. Kriston, L. Hölzel, A. K. Weiser, M. M. Berner, M. Härter “Meta-analysis: are 3 questions enough to detect unhealthy alcohol use?,” *Annals of internal medicine*, 149(12), 879-888, 2008.
- [90] P. Macaskill “Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis,” *Journal of clinical epidemiology*, 57(9), 925-932, 2004.
- [91] L. M. Bachmann, M. A. Puhan, G. ter Riet, P. M. Bossuyt “Sample sizes of studies on diagnostic accuracy: literature survey,” *Bmj*, 332(7550), 1127-1129, 2006.
- [92] N. Dendukuri, I. Schiller, L. Joseph, M. Pai “Bayesian Meta-Analysis of the Accuracy of a Test for Tuberculous Pleuritis in the Absence of a Gold Standard Reference,” *Biometrics*, 68(4), 1285-1293, 2012.
- [93] R. M. Harbord, J. J. Deeks, M. Egger, P. Whiting, J. A. Sterne “A unification of models for meta-analysis of diagnostic accuracy studies,” *Biostatistics*, 8(2), 239-251, 2006.

- [94] N. R. Draper, H. Smith “Applied Regression Analysis,” 2nd ed. New York: John Wiley & Sons, 1981.
- [95] P. Macaskill, C. Gatsonis, J. Deeks, R. Harbord, Y. Takwoingi “Cochrane handbook for systematic reviews of diagnostic test accuracy,” Version 0.9. 0. London: The Cochrane Collaboration, 2010.
- [96] M. M. Leeflang, P. M. Bossuyt, L. Irwig “Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis,” *Journal of clinical epidemiology*, 62(1), 5-12, 2009.
- [97] L. E. Moses, D. Shapiro, B. Littenberg “Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations,” *Statistics in medicine*, 12(14), 1293-1316, 1993.
- [98] L. Irwig, P. Macaskill, P. Glasziou, M. Fahey “Meta-analytic methods for diagnostic test accuracy,” *J Clin Epidemiol*, 48: 119-130. 1995
- [99] V. Dukic, C. Gatsonis “Meta-analysis of diagnostic test accuracy assessment studies with varying number of thresholds,” *Biometrics*, 59: 936-946, 2003.
- [100] J. B. Reitsma, A. S. Glas, A. W. Rutjes, R. J. Scholten, P. M. Bossuyt, A. H. Zwinderman “Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews,” *Journal of clinical epidemiology*, 58(10), 982-990, 2005.
- [101] C. M. Rutter, C. A. Gatsonis “A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations,” *Statistics in medicine*, 20(19), 2865-2884, 2001.
- [102] J.J. Deeks, P. Macaskill, L. Irwig “The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed,” *J Clin Epidemiol*, 58: 882-893, 2005.
- [103] J.J. Deeks “Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests,” *BMJ*, 323: 157-162, 2001.
- [104] A.H. Zwinderman, P.M. Bossuyt “We should not pool diagnostic likelihood ratios in systematic reviews,” *Stat Med*, 27: 687-697, 2008.

- [105] J.P. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman “Measuring inconsistency in meta-analyses,” *BMJ*, 327: 557-560, 2003.
- [106] B. Littenberg, L.E. Moses “Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method,” *Med Decis Making*, 13: 313-321, 1993.
- [107] R.M. Harbord, J.J. Deeks, M. Egger, P. Whiting, J.A. Sterne “A unification of models for meta-analysis of diagnostic accuracy studies,” *Biostatistics*, 8: 239-251, 2007.
- [108] L.R. Arends, T.H. Hamza, J.C. van Houwelingen, M.H. Heijnenbrok-Kal, M.G. Hunink, T. Stijnen “Bivariate random effects meta-analysis of ROC curves,” *Med Decis Making*, 28: 621-638, 2008.
- [109] C.M. Rutter, C.A. Gatsonis “Regression methods for meta-analysis of diagnostic test data,” *Acad Radiol*, 2 Suppl 1: S48-S56, 1995.
- [110] McCullagh “Regression models for ordinal data,” *Journal of the Royal Statistical Society Series B*, 42: 109-142, 1980.
- [111] A.N. Tosteson, C.B. Begg “A general regression methodology for ROC curve estimation,” *Med Decis Making*, 8: 204-215, 1988.
- [112] P.C. LAMBERT, A. J. SUTTON, P.R. BURTON, K.R. ABRAMS, D.R. JONES “How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS,” *Statist. Med.*, 24, 2401-2428, 2005
- [113] D. K. Park, A. Gelman, J. Bafumi “Bayesian multilevel estimation with poststratification: State-level estimates from national polls,” *Political Analysis*, 12(4), 375-385, 2004.
- [114] A. GELMAN, J. HILL “Data Analysis Using Regression and Multilevel/Hierarchical Models,” Cambridge University Press pag. 513–527, 2007.
- [115] H. Chu, S.R. Cole “Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach,” *J Clin Epidemiol*, 59: 1331-1332, 2006.

# Apéndice A

## Código fuente de la simulación

### A.1. Programa en SAS

```
SAS PROGRAM
/*****
READ DATASET FROM THE SAS OLEDB LIBRARY
*****/
libname Conn oledb user=sa password=xxx datasource=
provider=SQLOLEDB.1 properties=('initial_catalog'=meta);

proc sql;
create table Work.dataMeta as
$select test ,study_id ,tp ,fp ,fn ,tn
from Conn.muestraSAS$
—where test in (1,2) and clase in ('C')
order by test;
quit;

data dataMeta;
set dataMeta;
dis=0.5; pos=tp;n=tp+fn; output;
dis=-0.5; pos=fp;n=tn+fp; output;
```

```

data dataMeta;
set dataMeta;
covble=0;
if covariable='V2' then covble=1;
run;

ods select none;
proc nlmixed data=dataMeta cov ecov start alpha=0.05 ;
by test;

/*****
EXECUTION WITH COVARIABLE USING THE NLMIXED PROCEDURE
*****/
parms alpha=2 theta=0 beta=0 s2ua=0 s2ut=0 a1=0 t1=0 b1=0;

logitp=(theta+ut+t1*covble +
(alpha+ua+a1*covble)*dis)*
exp(-(beta+b1*covble)*dis);
p=exp(logitp)/(1+exp(logitp));
model pos ~ binomial(n,p);

$random ut ua ~ normal([0,0],[s2ut,0,s2ua])
subject=study_id out=randeffs;$

estimate 'theta_V2' theta+t1;
estimate 'alpha_V2' alpha+a1;
estimate 'beta_V2' beta+b1;

estimate 'logDOR' log((((exp(exp(-beta/2) * (theta + 0.5 * alpha)))/(1+exp(ex
/(1-(exp(exp(-beta/2) * (theta + 0.5 * alpha)))/
(1+exp(exp(-beta/2) * (theta + 0.5 * alpha))))))
/(((1-(exp(-exp(beta/2) * (theta - 0.5 * alpha)))/
(1+exp(-exp(beta/2) * (theta - 0.5 * alpha))))))

```

```

/(exp(-exp(beta/2) * (theta - 0.5 * alpha)) /
(1+exp(-exp(beta/2) * (theta - 0.5 * alpha)))));
estimate 'logLR+' log((exp(exp(-beta/2) * (theta + 0.5 * alpha)) / (1+exp(exp
/(1-(exp(-exp(beta/2) * (theta - 0.5 * alpha)) /
(1+exp(-exp(beta/2) * (theta - 0.5 * alpha))))));
estimate 'logLR-' log((1-(exp(exp(-beta/2) * (theta + 0.5 * alpha)) / (1+exp
/(exp(-exp(beta/2) * (theta - 0.5 * alpha)) /
(1+exp(-exp(beta/2) * (theta - 0.5 * alpha)))))) ;

/* CREATE DATASETS OF RESULTS */
ods output FitStatistics = estConvergencia;
ods output ParameterEstimates = estParametros;
ods output CovMatParmEst = matCovarianza;
ods output CovMatAddEst = matCovAdicionales;
ods output AdditionalEstimates = EstAdicionales;
run;

/*****
SEND TABLES TO SQLSERVER
*****/
proc sql;
delete from Conn.Estconvergencia;
delete from Conn.Estparametros;
delete from Conn.Matcovadicionales;
delete from Conn.MatcovarianzaH;
delete from Conn.EfectAlea;
delete from Conn.EstAdicionales;

insert into Conn.Estconvergencia (numero, parametro, valor)
select test, Descr, Value from Work.Estconvergencia;
insert into Conn.Estparametros (numero, Parameter, Estimate,
StandardError, DF, tValue, Probt, Alpha, Lower, Upper, Gradient)
select test, Parameter, Estimate, StandardError, DF,

```



```
tValue , Probt , Alpha , Lower , Upper , Gradient from Work.Estparametros ;
insert into Conn.Matcovadicionales ( numero , Row , Label , Cov1 , Cov2 )
select test , Row , Label , Cov1 , Cov2 from Work.Matcovadicionales ;
insert into Conn.MatcovarianzaH ( numero , Row , Parameter , alpha , theta , beta , s2ua
Work.Matcovarianza ; $insert into Conn.EfectAlea ( numero , study_id , Effect , Est
tValue , Probt , Alpha , Lower , Upper ) select test , study_id , Effect , Estimate , StdErr
Probt , Alpha , Lower , Upper from Work.Randeffs ; $insert into Conn.EstAdicionales
quit ;
```

## A.2. Programa en R

```
#####
# R PROGRAM
#####

library(HSROC)
library(RODBC)
rm(list = ls())

options(scipen=999) # Disable scientific notation
dir<-"C:\\Meta\\" # directorio de trabajo
conn<-odbcDriverConnect( 'DSN=DSN_META;UID=sa;pwd=@@123;DATABASE=meta' )

#####
# DATASET, Meta-analysis
#####
muestras<-1200
for (i in 1:muestras){
N<-trunc(runif(1,min=5,max=35)) # N:meta-analysis {LAMBDA min=5,max=35}
n <- seq(200,2000,1) # n:individuals
pi<-runif(N,min=0.0001,max=0.2499) # {Prevalence min=0.0001,max=0.2499}
beta<-runif(1,min=-0.75,max=0.75) # {beta min=0.75,max=0.75}
LAMBDA<-runif(1,min=-0.75,max=0.75) # {LAMBDA min=-0.75,max=0.75}
```

```

sd_alpha<-runif(1,min=0.0001,max=2) #{ sig_alpha min=0.0001,max=2}
THETA<-runif(1,min=-1.5,max=1.5) # {THETA min=-1.5,max=1.5}
sd_theta<-runif(1,min=0.0001,max=2) #{sd_theta min=0,max=2}

sim.data1 <-simdata(N, n,n.random = TRUE, prev=pi, T=THETA,
L=LAMBDA,sd_t=sd_theta, sd_a=sd_alpha, b=beta)

#Covariate in uniform groups
numRef=round(runif(1,min=1,max=3)); prop=round(N/numRef);
if (numRef==1){ coVar=rep('V1',N)}
if (numRef==2){coVar=c(rep('V1',prop),rep('V2',(N-prop)))}
if (numRef==3){coVar=c(rep('V1',prop),rep('V2',prop),rep('V3',
(N-(2*prop))))}

#Insert meta-analysis in database tables
sqlQuery(conn,paste("exec_SPtetra_",
paste(" ",paste(sim.data1$Data[,1],collapse=","),"',""),
paste(" ",paste(sim.data1$Data[,2],collapse=","),"',""),
paste(" ",paste(sim.data1$Data[,3],collapse=","),"',""),
paste(" ",paste(sim.data1$Data[,4],collapse=","),"',""),
i,
paste(" ','",paste(coVar,collapse=",")," '"),
"','C'"),
paste(" ",paste(round(sim.data1$'Within study parameters'[,1],4),
collapse=","),"',""),
paste(" ",paste(round(sim.data1$'Within study parameters'[,2],4),
collapse=","),"',""),
paste(" ",paste(round(sim.data1$'Within study parameters'[,3],4),
collapse=","),"',""),
paste(" ",paste(round(sim.data1$'Within study parameters'[,4],4),
collapse=","),"',""),
paste(" ",paste(round(sim.data1$'Within study parameters'[,5],4),
collapse=",")," '")

```

```

paste(" ", round(sim.data1$'Between study parameters'["THETA"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["sigma_theta"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["LAMBDA"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["sigma_alpha"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["beta"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["Overall_++"], 4)),
paste(" ", round(sim.data1$'Between study parameters'["Overall_--"], 4))
))
print(paste('Tetrachoric Table (Table_2x2): ', i))
sim.data1<-NULL
}
odbcClose(conn)
setwd("C:\\Meta\\")

#####
#           HSROC diagnostic meta-analysis model
#####
rm(list = ls())
dir<-"C:\\Meta\\"
conn1<-odbcDriverConnect('DSN=DSN_META;UID=sa;pwd=@@123;DATABASE=meta')
clase<-"C1"
escenario<-"E1"
strSQL<-paste("select distinct numero from muestraTetra where clase=",
clase, " and escenario=", escenario, " and numero in (73,169)
order by numero")
rst<-sqlQuery(conn1, strSQL)
for (i in 1:length(rst$numero)){
ruta = paste(dir, "Meta_", rst[i,1], sep="")
if(!dir.exists(ruta)){dir.create(ruta, mode = "0777")}
else{unlink(ruta, recursive = TRUE)}
setwd(ruta)
sql<-paste("select [+ +],[+ -],[- +],[- -] from muestraTetra where clase=",
clase, " and escenario=", escenario, " and numero=", rst[i,1])

```

```

datos<-sqlQuery(conn1 , sql)
N<-length(datos$'++')

#HSROC Algorithm
init.alpha<-c(runif(N,min=-3,max=3))
init.theta<-c(runif(N,min=-1.5,max=1.5))
init.s1 = rep(0.9 ,N)
init.c1 = rep(0.9 ,N)
init.pi<-runif(N,min=0,max=1)
init_within = cbind(init.alpha , init.theta , init.s1 , init.c1 , init.pi)
init_within = cbind(init.theta , init.alpha , init.s1 , init.c1 , init.pi)

init.THETA<-runif(1 ,min=-1.5,max=1.5)
init.sig.theta<-runif(1 ,min=0,max=2)
init.LAMBDA<-runif(1 ,min=-3,max=3)
init.sig.alpha<-runif(1 ,min=0,max=2)
init.beta<-runif(1 ,min=-0.75,max=0.75)
init_between = c(init.THETA, init.sig.theta , init.LAMBDA,
init.sig.alpha , init.beta)
init = list(init_within , init_between)

HSROC(data=datos , iter.num=20000, init=init , path=getwd())
HSROCSummary(data = datos , burn_in=10000, Thin=2, print_plot=TRUE)

#Insert Results (HSROC) in database tables
sqlQuery(conn1 , paste("exec _SPtetraResultados" ,
paste(" _" , rst [i ,1] , " , " ) ,
paste(" '" , paste( round(init.alpha ,4) , collapse = " , " ) , "' , " ) ,
paste(" '" , paste( round(init.theta ,4) , collapse = " , " ) , "' , " ) ,
paste(" '" , paste( round(init.s1 ,4) , collapse = " , " ) , "' , " ) ,
paste(" '" , paste( round(init.c1 ,4) , collapse = " , " ) , "' , " ) ,
paste(" '" , paste( round(init.pi ,4) , collapse = " , " ) , "' ' ) ,
paste(" , " , round(init.THETA ,4) ) ,

```

```
paste(" ", round(init.sig.theta,4)),
paste(" ", round(init.LAMBDA,4)),
paste(" ", round(init.sig.alpha,4)),
paste(" ", round(init.beta,4)),
", null",
", null",
paste(" ", "Meta_", rst[i,1], " ")
))

print(paste(' Finalized Meta: ', rst[i,1])) }
odbcClose(conn1)
setwd("../")
```