



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

Coeficiente de Correlación de Concordancia
bajo la distribución t multivariada

Por

Carla Gislaine Leal Kaymalyz

Tesis presentada a la Facultad de Matemáticas de la
Pontificia Universidad Católica de Chile,
como un requisito para optar al grado de
Doctor en Estadística.

Profesor guía: **Dr. Manuel Galea Rojas**

Junio de 2016
Santiago, Chile



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

Coeficiente de Correlación de Concordancia
bajo la distribución t multivariada

Por

Carla Gislaine Leal Kaymalyz

Comisión informante:

Dr. Manuel Galea Rojas

Dr. Reinaldo Arellano Valle

Dr. Wilfredo Palma Manríquez

Dr. Felipe Osorio Salgado

Junio de 2016

Santiago, Chile

Índice general

1. Introducción	1
1.1. Formulación del problema	1
1.2. Justificación	9
1.3. Objetivos	11
1.4. Resumen de los contenidos	12
1.5. Principales contribuciones de la tesis	13
2. Distribución t multivariada	14
2.1. Definición	15
2.2. Propiedades básicas	15
2.2.1. Representación estocástica de la distribución t	15
2.2.2. Momentos	16
2.2.3. Curtosis	16
2.2.4. Transformación lineal	17
2.2.5. Distribución marginal	17
2.2.6. Distribución condicional	17
2.2.7. Distribución de forma cuadrática	19
2.3. Función de verosimilitud	20
2.4. Estimación	21
2.4.1. Antecedentes	21
2.4.2. Estimación de máxima verosimilitud	21
2.4.3. Matriz de información de Fisher	23

2.5.	Relación con la distribución normal	24
2.5.1.	Distribución límite	25
2.5.2.	Bondad de ajuste	29
2.6.	Comentarios finales	31
3.	Coefficiente de correlación de concordancia	32
3.1.	Concordancia entre dos medidas	32
3.2.	Concordancia entre dos o más medidas	36
3.3.	Inferencia	39
3.3.1.	Estimación máximo verosímil	39
3.3.2.	Distribución asintótica	39
3.3.3.	Intervalo de confianza asintótico	40
3.3.4.	Test de razón de verosimilitud	41
3.3.5.	Prueba de Wald	43
3.4.	Comentarios finales	43
4.	Diagnóstico de influencia sobre el CCC	44
4.1.	Conceptos geométricos	44
4.2.	Análisis de influencia local	48
4.2.1.	Perturbación	50
4.2.2.	Curvatura normal para el desplazamiento de verosimilitud	51
4.2.3.	Conformal curvatura normal	54
4.2.4.	Ponderación de casos	55
4.2.5.	Perturbación de la respuesta	56
4.3.	Influencia local generalizada	59
4.3.1.	Ponderación de casos	62
4.3.2.	Perturbación de la respuesta	63
4.4.	Medidas de influencia	63
4.4.1.	Selección del esquema de perturbación	65
4.4.2.	Medidas de influencia de primer y segundo orden	65
4.4.3.	Ponderación de casos	68
4.4.4.	Perturbación de la respuesta	71
4.5.	Comentarios finales	72

5. Aplicación	74
5.1. Aplicación 1: Estudio del sueño	75
5.1.1. Descripción de los datos	75
5.1.2. Modelo estadístico y ajuste	79
5.1.3. Diagnóstico	80
5.1.4. Conclusión	88
5.2. Aplicación 2: Estudio del sueño con tres métodos	93
5.2.1. Descripción de los datos	93
5.2.2. Modelo estadístico y ajuste	98
5.2.3. Diagnóstico	100
5.2.4. Conclusión	105
6. Conclusión general	110
Bibliografía	112
A. Distribuciones y derivación de algunos resultados	118
A.1. Distribuciones de probabilidad utilizadas	118
A.1.1. Distribución Normal	118
A.1.2. Distribución Gamma	119
A.1.3. Distribución Chi-cuadrado	119
A.1.4. Distribución t multivariada	120
A.2. Algoritmo EM	121
A.3. Vector score y matriz de información de Fisher	124
A.3.1. Derivación del vector score	124
A.3.2. Derivación de la matriz de información observada	125
A.3.3. Derivación de la matriz de información de Fisher	127
A.4. Derivada del CCC	131
A.4.1. Derivada para CCC bivariado	132
A.4.2. Derivada para el CCC global	132
A.5. Derivación del Test de Razón de Verosimilitud	134
A.5.1. Estimador MV bajo todo el espacio paramétrico	134
A.5.2. Estimador MV bajo la hipótesis nula	134

A.5.3. Estadístico de prueba	136
A.6. Derivadas de influencia	136
A.6.1. Ponderación de casos	136
A.6.2. Perturbación de la respuesta	139
B. Estudio de simulación	141
B.1. Escenarios de simulación	141
B.2. Cobertura y consistencia	143
B.3. Potencia empírica para el TRV y Wald	144
B.4. Simulación para las medidas de influencia	146

Índice de tablas

4.1. Comparación entre las medidas de influencia según la derivada de la función objetivo sobre un punto crítico	67
5.1. Estadística descriptiva univariada para los registros log(LPS) generados por el método manual y automático.	78
5.2. Prueba de normalidad Shapiro-Wilk para los registros log(LPS).	78
5.3. Prueba de normalidad multivariada para los registros log(LPS) y prueba de normalidad para la distancia transformada, asumiendo el modelo \mathcal{N} y el modelo \mathcal{T}	78
5.4. Estimador del CCC, error estándar de estimación (eee), intervalo de confianza asintótico (ICA), precisión y sesgo.	81
5.5. Resumen de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCC respecto al esquema de perturbación, medida de influencia y modelo estadístico.	84
5.6. Valores, diferencia y estado de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCC.	84
5.7. Razón de de cambio para el estimador del CCC y medidas de calidad del ajuste al eliminar casos potencialmente influyentes.	86
5.8. Razón de cambio para los estimadores MV del modelo \mathcal{N} y \mathcal{T} al eliminar casos potencialmente influyentes.	86
5.9. Estimador puntual, error estándar de estimación e intervalo de confianza asintótico/confianza para el CCC basado en distintos métodos.	87

5.10. Estimador puntual e intervalo de credibilidad/confianza para del CCC basado en distintos métodos.	87
5.11. Acuerdo bivariado y global para los registros del logaritmo de LPS generados por los métodos manual, parcial y automático.	95
5.12. Registros obtenidos por tres métodos, media, mediana y medidas de variabilidad.	96
5.13. Estadística descriptiva univariada para los puntajes $\log(\text{LPS})$ obtenidos por los tres métodos.	96
5.14. Prueba de normalidad multivariada para los registros $\log(\text{LPS})$ y prueba de normalidad para la distancia transformada, asumiendo el modelo \mathcal{N} y el modelo \mathcal{T}	98
5.15. Prueba de normalidad multivariada de Mardia basado en los coeficientes de asimetría y curtosis multivariada.	99
5.16. Estimador del CCCG, error estándar de estimación (eee) e intervalo de confianza asintótico (ICA).	99
5.17. Resumen de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCCG respecto al esquema de perturbación, medida de influencia y modelo estadístico.	102
5.18. Razón de cambio para el estimador del CCCG y medidas de calidad del ajuste al eliminar casos potencialmente influyentes.	104
5.19. Razón de cambio para los estimadores MV del modelo \mathcal{N} y \mathcal{T} al eliminar casos potencialmente influyentes.	104
5.20. Estimador puntual, error estándar de estimación e intervalo de confianza asintótico/confianza para el CCCG basado en distintos métodos.	105
5.21. Estimador puntual e intervalo de credibilidad/confianza para el CCCG basado en distintos métodos.	105
B.1. Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCC simulando datos desde una distribución t	147

B.2. Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCC simulando datos desde una distribución normal	148
B.3. Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCCG simulando datos desde una distribución t	148
B.4. Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCCG simulando datos desde una distribución normal	149
B.5. Poder empírico para el TRV y tamaño del test para el CCC simulando datos desde una distribución t	152
B.6. Poder empírico para el TRV y tamaño del test para el CCC simulando datos desde una distribución normal	153
B.7. Poder empírico del TRV y tamaño del test para el CCCG simulando datos desde una distribución t	154
B.8. Poder empírico para el TRV y tamaño del test para el CCCG simulando datos desde una distribución normal	154
B.9. Poder empírico de la prueba de Wald y tamaño del test para el CCC simulando datos desde una distribución t	155
B.10. Poder empírico para la prueba de Wald y tamaño del test para el CCC simulando datos desde una distribución normal	156
B.11. Poder empírico para la prueba de Wald y tamaño del test para el CCCG simulando datos desde una distribución t	157
B.12. Poder empírico para la prueba de Wald y tamaño del test para el CCCG simulando datos desde una distribución normal	158
B.13. Porcentaje de detección del dato perturbado según medida de influencia bajo el esquema de ponderación de casos simulando datos desde una distribución normal bivariada	159

B.14. Porcentaje de detección del dato perturbado según medida de influencia bajo el esquema de perturbación de la respuesta simulando datos desde una distribución normal bivariada 159

Índice de figuras

1.1. Ejemplos donde el coeficiente de correlación de Pearson no mide grado de acuerdo, y la prueba t y F fallan al detectar bajo grado de acuerdo.	4
1.2. Ejemplos donde la prueba t y F fallan al detectar alto grado de acuerdo.	5
1.3. Grado de acuerdo para datos simulados desde una distribución normal bivariada basado en los escenarios de Lin (1989)	7
2.1. Elipses del 95% de masa de probabilidad para datos simulados desde una distribución t basado en los escenarios de Lin (1989)	28
2.2. QQ plots para la transformación Wilson-Hilferty de la distancia de Mahalanobis considerando datos simulados bajo los escenarios de Lin (1989)	30
2.3. QQ plots para la transformación Wilson-Hilferty de la distancia de Mahalanobis considerando datos reales	31
4.1. Comparación de estimadores	49
5.1. Dispersión entre los registros $\log(\text{LPS})$ obtenidos con el proceso Manual y Automático.	77
5.2. Límites de acuerdo de Altman para los registros $\log(\text{LPS})$	77
5.3. QQ-plots para las distancia transformada, asumiendo el modelo \mathcal{N} (a) y asumiendo el modelo \mathcal{T} (b), considerando el método M y A.	78

5.4. Elipse del 95 % basada en el vector de medias estimado y matriz de varianza-covarianza estimada, asumiendo el modelo \mathcal{N} (línea roja) y el modelo \mathcal{T} (línea azul).	80
5.5. Observaciones detectadas como potencialmente influyentes a través de la distribución de los pesos versus la distancia de Mahalanobis, asumiendo el modelo \mathcal{T}	81
5.6. Comparación de los intervalos de confianza y credibilidad basado en diferentes métodos.	87
5.7. Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.	88
5.8. Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos. . . .	89
5.9. Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos. . . .	89
5.10. Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local (o medida de primer orden) asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.	90
5.11. Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.	90
5.12. Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.	91
5.13. Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.	91

5.14. Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local (o medida de primer orden) asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.	92
5.15. Gráfico de influencia de \mathbf{h}_{max} para la medida de influencia de segundo orden asumiendo que los datos siguen una distribución normal bajo ambos esquemas de perturbación.	92
5.16. Dispersión entre los registros log(LPS) obtenidos con el proceso Manual (M), Parcial (P) y Automático (A).	94
5.17. Media versus desviación estándar (de) (a) y mediana versus desviación media (dm) (b) de los registros obtenidos por los tres métodos.	95
5.18. QQ-plots para las distancia transformada, asumiendo el modelo \mathcal{N} (a) y asumiendo el modelo \mathcal{T} (b), considerando el método M,P y A.	97
5.19. Observaciones detectadas como potencialmente influyentes a través de la distribución de los pesos versus la distancia de Mahalanobis.	100
5.20. Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos. . . .	106
5.21. Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.	106
5.22. Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.	107
5.23. Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.	107

5.24. Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.	108
5.25. Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.	108
5.26. Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.	109
5.27. Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.	109
B.1. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 1 . . .	149
B.2. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 2 . . .	150
B.3. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 3 . . .	150
B.4. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 1	151
B.5. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 2	151

B.6. Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 3	152
B.7. Dispersión de datos simulados (a), \mathbf{h}_{max} de la conformal curvatura normal (b), \mathbf{h}_{max} de la curvatura (c), \mathbf{h}_{max} de la medida de primer orden (d) y \mathbf{h}_{max} de la medida de segundo orden (e), bajo el esquema de ponderación de casos	160
B.8. Dispersión de datos simulados (a), \mathbf{h}_{max} de la conformal curvatura normal (b), \mathbf{h}_{seg} de la curvatura (c), \mathbf{h}_{max} de la medida de primer orden (d) y \mathbf{h}_{max} de la medida de segundo orden (e), bajo el esquema de perturbación de la respuesta	161

AGRADECIMIENTOS

Quiero agradecer al Doctor Manuel Galea Rojas por su infinita paciencia y confianza. Sin su orientación, compasión y ayuda este trabajo no hubiera sido posible de la manera que lo es. Gracias por su tiempo y conocimiento.

También quiero agradecer a mi querido amigo Jaime Arrue por su amistad incondicional, por estar siempre apoyándome. Gracias amigo por ser quien eres, una persona excepcional y maravillosa. A mi amiga Luz Polo, eres una mujer digna de imitar, por tu valentía y fuerza frente a la vida, cualidades que me has enseñado en este largo caminar. A mi familia y compañeros por hacer que el día a día sea una brisa en este camino, el cual necesita de mucha entereza para ser recorrido.

Además, quiero agradecer a los profesores de la Pontificia Universidad Católica de Chile, que han entregado su conocimiento a través de sus trabajos y clases, sin duda un aporte a las nuevas generaciones.

Finalmente, agradecer al Doctor Felipe Osorio por su colaboración a través del desarrollo de la librería en R llamada MVT y al Proyecto Fondecyt 1150325 por el financiamiento del presente trabajo.

“La buena suerte no es casual, es producto del trabajo; así la sonrisa de la fortuna tiene que ganarse a pulso”.

Emily Dickinson.

Coeficiente de correlación de concordancia bajo la distribución t multivariada

Carla G. Leal Kaymalyz

Profesor guía: Dr. Manuel Galea Rojas

Resumen

El Coeficiente de Correlación de Concordancia (CCC), propuesto por Lin (1989), permite evaluar el grado de acuerdo entre medidas generadas por dos instrumentos de medición. Una extensión es el CCC global, propuesto por Barnhart *et al.* (2007), que permite evaluar el grado de acuerdo entre medidas generadas por dos o más instrumentos.

Es bien sabido que la inferencia estadística basada en la distribución normal es susceptible a la presencia de datos atípicos y/o extremos, y por ende las estimaciones del CCC pueden verse afectadas. Una alternativa es el uso de la distribución t , la cual permite realizar inferencia estadística más robusta entregando estimaciones del CCC y del CCC global menos sensible a la presencia de datos atípicos y/o extremos.

El método de diagnóstico de influencia local, propuesto por Cook (1986), es usado para detectar observaciones potencialmente influyentes sobre el estimador máximo verosímil, y por ende sobre el estimador del CCC. Una extensión del método es la influencia local generalizada y las medidas de influencia de primer y segundo orden, que permiten detectar observaciones potencialmente influyentes sobre el estimador del CCC.

Finalmente, se ilustra con datos reales que, bajo la distribución t , los estimadores del CCC y del CCC global son menos sensibles que los estimadores obtenidos bajo el supuesto de normalidad.

Introducción

Este capítulo introduce el concepto de acuerdo entre medidas generadas por diferentes instrumentos de medición y el uso del Coeficiente de Correlación de Concordancia (CCC) para evaluar el grado de similitud entre medidas. Además presenta la importancia de obtener estimaciones robustas del CCC y se justifica el diagnóstico de influencia local sobre el estimador Máximo Verosímil (MV) y el estimador del CCC.

Por otro lado, plantea el objetivo general, los objetivos específicos, el resumen de los contenidos de los capítulos posteriores y destaca la contribución del trabajo de tesis.

1.1. Formulación del problema

Todo proceso de medición debe asegurar que las medidas generadas por distintos instrumentos o métodos concuerden entre si o concuerden con la verdadera medida. Sin embargo, en la práctica es poco viable considerar una medida como verdadera pues, por un lado, se acepta que la medida considere un error tolerable por el investigador y, por el otro la medida verdadera no siempre se conoce o está disponible. Bajo este contexto, se entenderá como medida verdadera a la medida de referencia si es que está disponible.

Considerando que las medidas sean continuas, Barnhart *et al.* (2007) definen varios conceptos usados en diferentes áreas de estudio para referirse a la concordancia entre medidas. Conceptos como exactitud y precisión, repetibilidad y reproducibilidad, validez y confiabilidad, dependibilidad, generabilidad y acuerdo están estrechamente relacionados con la idea de cuantificar la “cercanía entre las medidas” considerando o no una medida de referencia.

En este trabajo el concepto de “acuerdo entre medidas” se refiere a la concordancia de las mismas, el cual involucra los conceptos de exactitud y precisión.

Barnhart *et al.* (2007) y Lin *et al.* (2004) se basan en la FDA (Food and Drug Administration, por sus siglas en inglés) para definir exactitud y precisión. El primer concepto se define como la desviación de la media respecto al valor de referencia (sesgo sistemático) y precisión como el grado de variabilidad entre las medidas (grado de dispersión).

Por otro lado, Choudhary y Nagaraja (2015) definen tres tipos de estudios para evaluar el acuerdo entre medidas, el primero es llamado **estudio de comparación**, el segundo el **estudio de calibración** y el tercero el **estudio de conversión**.

El estudio de comparación consiste en evaluar si el nuevo método genera medidas concordantes con el método de referencia, a menudo llamado Gold Standard, de modo que el nuevo método pueda reemplazar al antiguo. El estudio de calibración compara un método aproximado con uno que es conocido, preciso y exacto cuyo error de medición es marginal. El estudio de conversión compara métodos que miden la característica aproximada en diferentes unidades de medida.

En cuanto a la estructura de los datos para el proceso de comparación entre medidas, se considera: (1) uno o más de un método sin réplica; (2) dos o más métodos con réplicas; (3) uno o más de un método tratado como referencia fija o aleatoria; (4) datos longitudinales donde el método es tomar las medidas sobre el tiempo

y (5) datos donde las covariables están disponibles para evaluar el impacto de varios factores sobre las medidas de acuerdo.

Para cualquier estudio de comparación y estructura de datos a considerar, en la práctica, se han usado diferentes enfoques para evaluar el grado de acuerdo entre medidas. Barnhart *et al.* (2007) los resume en: métodos descriptivos, por ejemplo la dispersión de los pares de datos alrededor de la línea de 45 grados; índices no escalados, por ejemplo los límites de acuerdo propuesto por Bland y Altman (1999) y los índices escalados, que toman valores entre -1 y 1, por ejemplo el Coeficiente de Correlación Intra-clase (CCI) y el Coeficiente de Correlación de Concordancia (CCC) propuesto por Lin (1989) entre otros.

Por otro lado, también se han usado técnicas estadísticas para evaluar el acuerdo entre medidas; por ejemplo el Coeficiente de Correlación de Pearson (CCP), la prueba t para datos pareados, la prueba F y el Coeficiente de Variación (CV). No obstante, estas técnicas presentan inconvenientes al detectar realmente acuerdo entre las medidas, ya que, detectan acuerdo significativo cuando en realidad existe bajo acuerdo. Por lo tanto, ninguno de estos métodos por sí sólo puede evaluar el grado de acuerdo entre medidas.

Basado en Lin *et al.* (2004), se presentan algunos ejemplos para ilustrar el inconveniente de no poder cuantificar realmente el acuerdo entre las medidas, usando las técnicas estadísticas anteriormente mencionadas.

En la Figura 1.1 los gráficos (a), (b) y (c) presentan un CCP igual a uno. No obstante, el CCC va disminuyendo considerablemente a medida que los pares de datos se alejan de la recta de 45 grados (línea de concordancia). Esta situación refleja que el CCP no mide grado de acuerdo entre las medidas, sino que tan sólo mide el grado de relación lineal. Otro ejemplo es presentado en las gráficas (d) y (e), donde la prueba t , de igualdad de medias, falla al no rechazar la hipótesis $H_0 : \mu_1 = \mu_2$, ya que, debería ser rechazada, pues existe un bajo grado de acuerdo. Esta situación se repite en la prueba F presentada en la gráfica (f).

En la Figura 1.2, gráfico (a), la prueba t rechaza la hipótesis de igualdad de medias en presencia de un alto grado de acuerdo, esto implica que la prueba no es capaz de detectar alta concordancia entre las medidas, y ocurre lo mismo con la prueba F , gráfica (b).

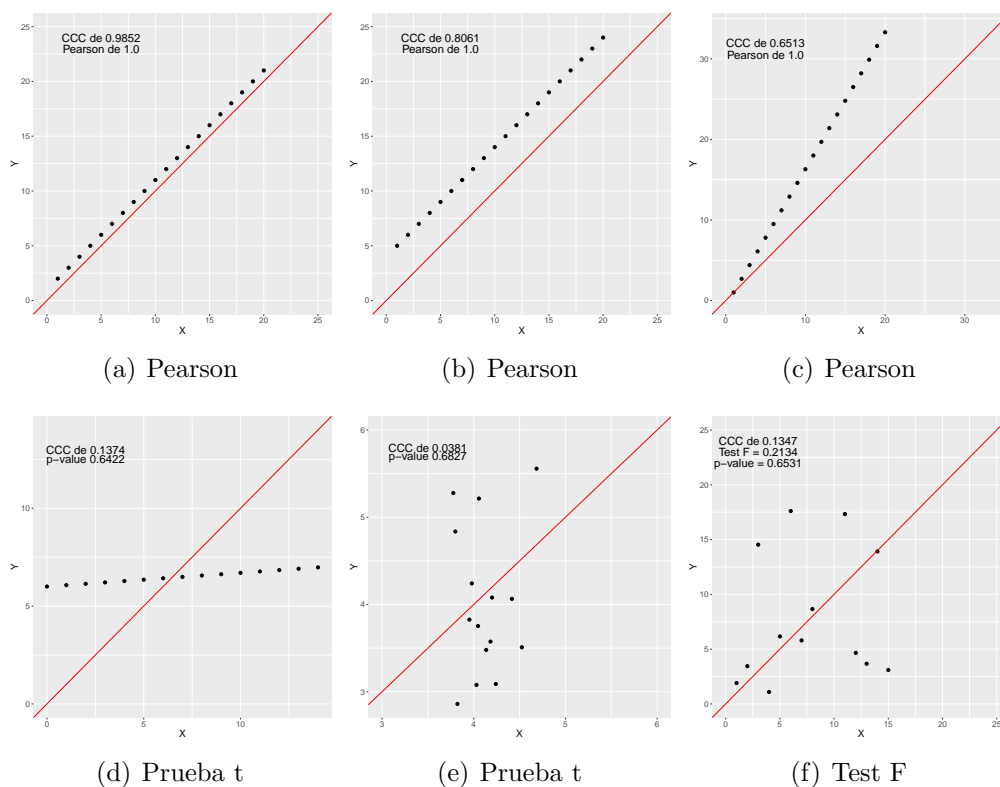


Figura 1.1: Ejemplos donde el coeficiente de correlación de Pearson no mide grado de acuerdo, y la prueba t y F fallan al detectar bajo grado de acuerdo.

Por lo anterior, el CCC es un índice que permite medir el grado de acuerdo entre medidas, ya sea en el caso de considerar o no un método de referencia.

Sea $(X_1, X_2)^T$ seleccionadas independientemente de una población con media $(\mu_1, \mu_2)^T$ y matriz de varianza-covarianza

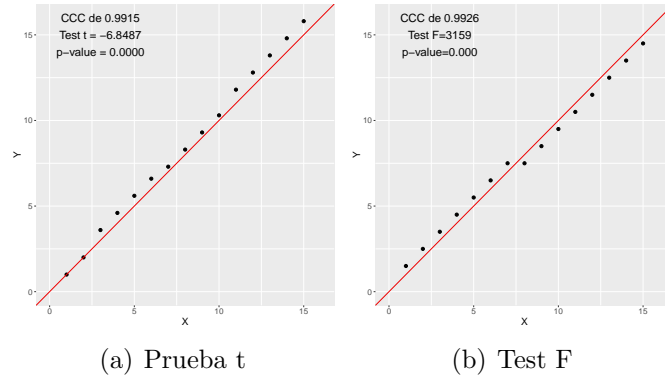


Figura 1.2: Ejemplos donde la prueba t y F fallan al detectar alto grado de acuerdo.

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

se define el CCC, sin considerar un método de referencia, como

$$\rho_c = \frac{2\sigma_{12}}{\sigma_{11} + \sigma_{22} + (\mu_1 - \mu_2)^2} = \rho_{12}C_{12}$$

donde

- $\rho_{12} = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$ es el coeficiente de precisión (correlación de Pearson).
- $C_{12} = 2[b + b^{-1} + a^2]^{-1}$ es el coeficiente de exactitud (sesgo).

con $b = \sqrt{\sigma_{11}/\sigma_{22}}$ y $a = (\mu_1 - \mu_2)/(\sigma_{11}\sigma_{22})^{1/4}$.

El CCC depende de la **precisión** de los datos, en el sentido de que tan cerca están los pares de datos a la mejor línea ajustada y de la **exactitud** de los datos, en el sentido de que tan cerca está la mejor línea ajustada de la línea de concordancia.

Para ilustrar cómo el acuerdo entre medidas depende de estos dos últimos conceptos, se presentan a continuación tres conjuntos de datos simulados desde una distribución normal, en base a los escenarios dados en Lin (1989).

En la Figura 1.3 se muestra el CCC bajo estos escenarios, considerando una muestra de tamaño $n = 150$. El escenario (a) considera un CCC de 0.95 con vector de media de $(0, 0)^T$ y matriz de varianza-covarianza

$$\begin{pmatrix} 1 & 0.95 \\ 0.95 & 1 \end{pmatrix},$$

el escenario (b) considera un CCC de 0.75 con un vector de media $(-\sqrt{0.1}/2, \sqrt{0.1}/2)^T$ y matriz de varianza-covarianza

$$\begin{pmatrix} 0.90^2 & 0.80 \times 0.90 \times 1.1 \\ 0.80 \times 0.90 \times 1.1 & 1.1^2 \end{pmatrix}$$

y el escenario (c) considera un CCC de 0.36 con vector de media $(-\sqrt{0.25}/2, \sqrt{0.25}/2)^T$ y matriz de varianza-covarianza

$$\begin{pmatrix} (4/3)^2 & 0.5 \times (4/3) \times (2/3) \\ 0.5 \times (4/3) \times (2/3) & (2/3)^2 \end{pmatrix}.$$

El escenario (a) no presenta diferencia en el vector de localización ni en la matriz de escala, el escenario (b) presenta una leve diferencia en el vector de locación y en la matriz de escala, y el escenario (c) presenta una gran diferencia, tanto en el vector de localización como en la matriz de escala.

En la Figura 1.3 se puede observar el grado de acuerdo entre las medidas, en base a los escenarios antes definidos. Bajo el caso que exista alta concordancia, gráfico (a), los pares de medidas están cerca de la línea de concordancia y a medida que ésta disminuye, los pares de medidas se van alejando de la línea de concordancia, gráfico (b) y (c). Si el CCC pierde precisión, la gráfica tiende a ser una nube de puntos, mientras que si pierde exactitud, los pares de puntos se van alejando de la línea de concordancia. Por lo tanto, la forma que tome el gráfico depende claramente de la medida de precisión y de la medida de exactitud.

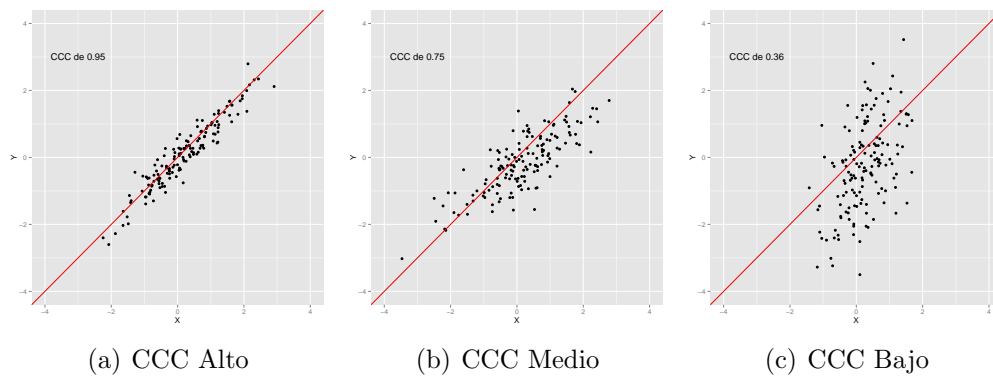


Figura 1.3: Grado de acuerdo para datos simulados desde una distribución normal bivariada basado en los escenarios de Lin (1989)

Otro punto de interés, en este contexto, es conocer los métodos de inferencia estadística usados sobre el CCC, como el método de estimación, el de construcción de intervalos de confianza y el de contraste de hipótesis estadística, este último poco desarrollado. Por ejemplo, Lin (1989) usa el Método de los Momentos (MM) y, asumiendo que los pares de datos siguen una distribución normal bivariada, obtiene el estimador de Máximo Verosímil (MV). King y Chinchilli (2001) estiman el CCC usando los Estadísticos U (EU), mientras que Barnhart y Williamson (2001) estiman el CCC usando las ecuaciones de estimación generalizada (GEE por su sigla en inglés) y en Barnhart *et al.* (2002) extiende la metodología GEE para más de dos instrumentos. Por otro lado, Carrasco y Jover (2003) proponen un modelo de Componente de Varianza (CV) con efectos mixtos, asumiendo que los datos siguen una distribución normal, para estimar el CCC y en Carrasco *et al.* (2007) comparan las estimaciones del CCC obtenidas a través del MM, CV, EU y GEE en datos asimétricos.

En cuanto a la construcción de intervalos de confianza para el CCC, Lin (1989) encuentra la distribución asintótica del CCC en base al método delta y la Z-transformación de Fisher para su construcción y están implementados en la librería Agreement de R. Desde la perspectiva Bayesiana, en Feng *et al.* (015a) se presentan resultados de intervalos de credibilidad que están implementados en

la librería `agRee` de R. En general, los intervalos de confianza y credibilidad son comparados con los intervalos Bootstrap y Jakknife.

Casi toda la inferencia estadística, de estimación e intervalos de confianza, es basada sobre el supuesto de normalidad de datos. Sin embargo, es bien sabido que la inferencia estadística, asumiendo esta distribución, es vulnerable a la presencia de datos extremos y/o atípicos, ya que, es imposible ajustar la curtosis de estas observaciones debido a que el cuarto momento de la distribución normal es determinado por el momento de primer y segundo orden.

La distribución t ha sido ampliamente usada para modelar conjuntos de datos con colas más pesadas respecto a la distribución normal y por ende, proporciona un enfoque más robusto a la inferencia estadística. Por ejemplo, Lange *et al.* (1989) estudian el modelamiento estadístico basado en la distribución t obteniendo estimadores robustos a través del algoritmo EM y McLachlan y Peel (2000) proporcionan un método de estimación basado en el algoritmo EM para modelos de mezcla t multivariadas. En el caso del CCC, Feng *et al.* (015a) propone un estimador para el CCC basado en el enfoque Bayesiano asumiendo que los datos siguen una distribución t y al ser comparado con otros métodos, basado en la distribución normal, este entrega estimaciones robustas del CCC en presencia de observaciones atípicas y/o extremas, y en el caso que los datos sean asimétricos, Feng *et al.* (015b), proponen un método robusto bayesiano bajo este escenario.

Bajo el planteamiento del problema que se ha discutido, a continuación, se presenta la justificación de realizar inferencia estadística para el CCC bajo el supuesto distribucional t y la aplicación de un método de diagnóstico para detectar observaciones que potencialmente puedan influir en la estimación del CCC, intervalos de confianza e hipótesis estadística.

1.2. Justificación

Desde el enfoque clásico no existen trabajos que estimen el CCC desde una perspectiva robusta. Por lo tanto, se propone estimar el CCC basado en la distribución t y como es importante evaluar la sensibilidad del estimador del CCC en presencia de observaciones atípicas y/o extremas se justifica el desarrollo y aplicación de un método de diagnóstico que permita detectar observaciones potencialmente influyentes sobre el estimador del CCC.

Existen varias técnicas de diagnóstico para estudiar la sensibilidad sobre el estimador MV en presencia de valores extremos y/o atípicos. Una de ellas es el método de **Influencia Local** propuesto por Cook (1986), el cual ha sido aplicado por varios autores; por ejemplo Paula (1993) aplica la metodología en modelos de regresión restringidos, Galea *et al.* (1997) en modelos de regresión lineal elípticos, Lesaffre y Verbeke (1998) en modelos lineales mixtos con medidas repetidas, Galea *et al.* (2000) en el modelo de regresión lineal elíptico, Diaz *et al.* (2003) en el modelo de regresión lineal elíptico multivariado, Galea *et al.* (2005) en modelos de calibración comparativa bajo distribución t elíptica, Osorio *et al.* (2007) en el modelo lineal elíptico con estructura longitudinal y Bastiani *et al.* (2014) en el modelo lineal espacial elíptico, entre otros.

El objetivo del método de diagnóstico es evaluar la sensibilidad de los estimadores MV a pequeñas perturbaciones en los datos y/o los supuestos del modelo estadístico.

Para evaluar la influencia local, que producen pequeñas perturbaciones introducidas al modelo estadístico, se debe estudiar el comportamiento de la superficie generada por el Desplazamiento de Verosimilitudes (DV) y ver cómo ésta se desvía de su plano tangente desde un punto crítico. Esto se puede realizar estudiando las curvaturas de secciones normales sobre dicho punto crítico, de modo que se pueda encontrar la dirección donde se produce la mayor curvatura y si en esa dirección su i -ésimo elemento es relativamente grande respecto a un punto de

corte, se concluye que el i -ésimo caso es un dato potencialmente influyente sobre los estimadores MV.

En vez de considerar el DV como la función objetivo que genera una determinada superficie, Wu y Luo (1993) proponen el estimador MV y estudian su superficie. Este análisis es llamado **enfoque de segundo orden** a la influencia local y es complementario a la influencia propuesta por Cook (1986). Por otro lado, la influencia local de **primer orden** se basa en propiedades de la pendiente, y según Wu y Luo (1993) no necesariamente son adecuadas en el sentido que no siempre detecta casos con influencia importante. Esta situación es la motivación para estudiar el comportamiento de la superficie generada por el estimador MV o una función del estimador.

Bajo el enfoque de la influencia local de segundo orden, autores como Lawrance (1988), Thomas y Cook (1990) y Cadigan y Farrel (2002) han propuesto otras funciones objetivos diferentes al del DV y del estimador MV.

En particular, Cadigan y Farrel (2002), destacan dos dificultades del DV como función objetivo; (1) la dirección donde se produce la mayor curvatura local sobre la superficie del DV (gráfico de influencia), es importante como método de diagnóstico, sin embargo, en altas dimensiones encontrar la dirección en dónde se produce la mayor curvatura es difícil desde el punto de vista computacional y (2) el método de influencia local, considerando el DV como función objetivo, se enfoca directamente sobre los estimadores MV y en muchas situaciones el interés es evaluar la sensibilidad de una función de los estimadores MV, como es el caso del estimador del CCC.

Además Poon y Poon (1999) muestran que la curvatura del DV no es invariante bajo transformación de escala uniforme, lo cual puede conllevar a conclusiones ambiguas respecto al estudio de su superficie. Los autores proponen la **Conformal curvatura normal**, la cual es invariante bajo transformación de escala uniforme cuando la primera derivada de la función objetivo, evaluada en un punto

crítico, es cero. Esta curvatura proporciona una medida de la influencia acotada entre 0 y 1 permitiendo comparar curvaturas bajo distintos modelos estadísticos.

Por otro lado, Fung y Kwan (1997) también destacan que para algunas funciones objetivos (DV, estimador MV o un test estadístico), bajo el escenario donde su primera derivada evaluada en un punto crítico no es cero, la curvatura normal y la conformal curvatura normal son invariantes bajo transformación de escala uniforme.

Bajo el contexto anterior, Zhu *et al.* (2007) proponen nuevas medidas de influencia, las cuales, son invariantes bajo transformación de escala uniforme, llamadas **Medidas de influencia de primer y segundo orden**. Este enfoque permite estudiar el comportamiento de la superficie generada por cualquier función objetivo diferenciable, sin importar que su primera derivada, evaluada en un punto crítico, sea o no cero.

1.3. Objetivos

General

Realizar inferencia estadística, desarrollar e implementar el método diagnóstico y analizar la sensibilidad del estimador del CCC asumiendo que los datos siguen una distribución t multivariada, y comparar los resultados obtenidos al asumir normalidad.

Específicos

- Estimar el CCC, construir un intervalo de confianza asintótico basado en las propiedades del estimador MV y el método delta, y desarrollar estadísticos de prueba asintóticos.
- Desarrollar e implementar el método de diagnóstico de influencia local, influencia local generalizada y medidas de influencia de primer y segundo orden para el estimador del CCC.

- Estudiar la cobertura de intervalos de confianza asintóticos, estudiar el desempeño de estadísticos de prueba y determinar la proporción de observaciones potencialmente influyentes sobre el estimador del CCC.
- Aplicar el método de diagnóstico con datos reales y estudiar la sensibilidad del estimador del CCC suponiendo una distribución t .
- Comparar todos los resultados de inferencia y diagnóstico obtenidos al asumir que los datos reales siguen una distribución t con los resultados obtenidos al asumir que los datos reales siguen una distribución normal.

1.4. Resumen de los contenidos

El resumen de los contenidos de este trabajo se detalla en los siguientes capítulos.

En el capítulo dos se define la distribución t multivariada, sus propiedades básicas, el método de estimación y la relación con la distribución normal.

El capítulo tres presenta la definición del CCC para medidas generadas por dos y más de dos instrumentos o métodos. Además, contiene la inferencia estadística; como la estimación del CCC, el intervalo de confianza asintótico y estadístico de prueba asintótico.

El capítulo cuatro contiene algunos conceptos de geometría diferencial y el desarrollo del método de diagnóstico de influencia local, influencia local generalizada y construcción de las medidas de influencia de primer y segundo orden para el caso t y normal.

El capítulo cinco es la aplicación con datos reales del trabajo de tesis y por último el capítulo seis contiene conclusiones generales.

1.5. Principales contribuciones de la tesis

En este trabajo se desarrollan contribuciones a la inferencia estadística y al método de diagnóstico, estas son:

1. Al asumir que los datos siguen una distribución t multivariada, el proceso de inferencia para el CCC es robusto, en el sentido que en presencia de observaciones atípicas y/o extremas, las estimaciones no se ven afectadas.
2. El desarrollo e implementación del método de diagnóstico de influencia local generalizada (enfoque de primer y segundo orden) sobre el CCC, asumiendo normalidad de datos, y una extensión del método es desarrollado asumiendo que los datos siguen una distribución t , para el enfoque de primer orden.
3. Se desarrollan las medidas de influencia invariantes bajo transformación de escala uniforme, de primer y segundo orden, asumiendo normalidad de datos, para el diagnóstico de influencia local sobre el CCC y se verifica si el esquema de perturbación es adecuado.

Distribución t multivariada

La distribución t multivariada es una generalización natural de la distribución t -Student. Una primera versión fue derivada por Cornish (1954) y también de manera independiente por Dunnett y Sobel (1954).

El interés de la distribución t multivariada sigue creciendo debido a que ofrece una alternativa para modelar datos con colas más pesadas que la distribución normal. Esto se debe a que esta distribución proporciona un parámetro de forma que permite estimar la curtosis de los datos, por lo tanto estos pueden ser adecuadamente modelados. Por otro lado, también su uso proporciona procedimientos de estimación robusta contra posible observaciones extremas y/o atípicas. Por ejemplo, autores como Lange *et al.* (1989), Arellano (1994), Galea (1995), Bolfarine y Galea (1996) y Osorio y Galea (2016) han usado la distribución t para realizar inferencia sobre el vector de media y la matriz de varianza-covarianza.

Una caracterización más general de la distribución t , dentro de la familia elíptica, es dada en Arellano y Bolfarine (1995) y la representación más usada de la distribución t multivariada se encuentra en Nadarajah y Kotz (2004) con vector de localización, matriz de escala y parámetro de forma, sin embargo en Sutradhar (1993) se presenta una reparametrización de la distribución t que permite compararla directamente con la distribución normal multivariada a través del vector de medias y la matriz de varianza-covarianza. Además Lange *et al.* (1989) sugieren

una reparametrización del parámetro de forma.

2.1. Definición

A continuación se presenta la definición formal de la función de densidad de probabilidad (fdp) de la distribución t multivariada reparametrizada.

Definición 2.1.1. Un vector aleatorio $\mathbf{X} = (X_1, \dots, X_d)^T$ que toma valores en \mathbb{R}^d sigue una distribución t multivariada si su fdp tiene la siguiente forma

$$p(\mathbf{x}) = K_d(\eta) |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)} \quad (2.1.1)$$

con

$$K_d(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} \text{ y } c(\eta) = \frac{\eta}{(1 - 2\eta)}$$

y donde $\delta^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$ es la distancia de Mahalanobis, $\eta \in (0, 1/2)$ parámetro de forma, $\boldsymbol{\mu} \in \mathbb{R}^d$ vector de media, $\boldsymbol{\Sigma}_{d \times d}$ matriz de varianza-covarianza simétrica defina positiva.

Notación: Se denota como $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ a la variable aleatoria d dimensional con distribución t multivariada y se denota como $\mathbf{X}_1, \dots, \mathbf{X}_n$ a la muestra de vectores aleatorios independientes e idénticamente distribuidos (iid).

2.2. Propiedades básicas

Las siguientes propiedades de la fdp (2.1.1) son similares a las propiedades de la distribución normal.

2.2.1. Representación estocástica de la distribución t

Sea $U \sim \Gamma(1/2\eta, 1/2c(\eta))$ independiente de $\mathbf{Y} \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$ de modo que se considera la siguiente representación estocástica

$$\mathbf{X} = \boldsymbol{\mu} + \frac{\mathbf{Y}}{\sqrt{U}} \quad (2.2.1)$$

con vector $\boldsymbol{\mu}$ de dimensión $(d \times 1)$ y admite una fdp (2.1.1) (ver Apéndice A.1).

2.2.2. Momentos

Si $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. Entonces

$$E(\mathbf{X}) = \boldsymbol{\mu}, \quad (2.2.2)$$

$$E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T] = \boldsymbol{\Sigma}, \quad (2.2.3)$$

$$E[(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})] = d \quad (2.2.4)$$

Demostración: Basta considerar la representación (2.2.1) y aplicar el operador esperanza a \mathbf{X} , $(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T$ y $(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$, lo que demuestra (2.2.2), (2.2.3) y (2.2.4) respectivamente.

2.2.3. Curtosis

Sea $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. Entonces el coeficiente de curtosis multivariado está dado por

$$E[\{(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})\}^2] = d(d+2)(1 + \kappa) \quad (2.2.5)$$

con $\kappa = 2\eta/(1 - 4\eta)$ para todo $\eta < 1/4$.

Demostración: Por la representación (2.2.1) se tiene que

$$E[\{(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})\}^2] = E[U^{-2}]E[(\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y})^2].$$

Como $U \sim \Gamma(1/2\eta, 1/2c(\eta))$ se tiene que $E(U^{-2}) = (1 - 2\eta)/(1 - 4\eta)$ y como $\mathbf{Y} \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$ entonces $\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y} \sim \chi^2(d)$ de modo que $E[(\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y})^2] = 2d + d^2$, luego (2.2.5) queda demostrado.

2.2.4. Transformación lineal

Si $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ e $\mathbf{Y} = \mathbf{A}\mathbf{X}$ con \mathbf{A} matriz de dimensión $(q \times d)$ no singular ($q \leq d$), entonces

$$\mathbf{Y} \sim \mathcal{T}_d(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T, \eta). \quad (2.2.6)$$

Demostración: Esta propiedad sigue de la representación (2.2.1) con distribución t multivariada y de la propiedad de transformación lineal de la distribución normal multivariante. Es decir,

$$E(\mathbf{Y}) = E(\mathbf{A}\mathbf{X}) = \mathbf{A}\boldsymbol{\mu} \text{ y } V(\mathbf{Y}) = V(\mathbf{A}\mathbf{X}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T.$$

Luego, (2.2.6) queda demostrada.

2.2.5. Distribución marginal

Si $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ y $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^T$ es una partición donde \mathbf{X}_1 es de dimensión $(q_1 \times 1)$ con

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \text{ y } q_1 \leq d.$$

Se tiene que la distribución marginal para cualquier elemento de la partición de \mathbf{X} también sigue una distribución t . Es decir,

$$\mathbf{X}_1 \sim \mathcal{T}_{q_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \eta). \quad (2.2.7)$$

Demostración: Para obtener (2.2.7), basta tomar $\mathbf{A} = [\mathbf{I}_{q_1} \mathbf{0}]$ y aplicar una transformación lineal $\mathbf{A}\mathbf{X}$ de modo que $E(\mathbf{A}\mathbf{X}) = \boldsymbol{\mu}_1$ y $V(\mathbf{A}\mathbf{X}) = \boldsymbol{\Sigma}_{11}$.

2.2.6. Distribución condicional

Sin pérdida de generalidad, considerando la partición anterior de modo que \mathbf{X}_1 de dimensión $(q_1 \times 1)$ y \mathbf{X}_2 de dimensión $(q_2 \times 1)$. La distribución condicional \mathbf{X}_1

dato \mathbf{X}_2 es

$$\mathbf{X}_1|\mathbf{X}_2 \sim \mathcal{T}_{q_1}(\boldsymbol{\mu}_{1|2}, \boldsymbol{\Sigma}_{1|2}, 1/\eta + q_2), \quad (2.2.8)$$

donde

$$\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$$

y

$$\boldsymbol{\Sigma}_{1|2} = \frac{1/c(\eta) + (\mathbf{X}_2 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)}{1/\eta + q_2} (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}).$$

Demostración:

Por definición se tiene $p(\mathbf{x}_1|\mathbf{x}_2) = p(\mathbf{x}_1, \mathbf{x}_2)/p(\mathbf{x}_2)$, de modo que

$$\begin{aligned} p(\mathbf{x}_1|\mathbf{x}_2) &= \left(\frac{c(\eta)}{\pi}\right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta)\delta^2)^{-(1/2\eta + d/2)} \\ &\times \left(\frac{c(\eta)}{\pi}\right)^{-q_2/2} \frac{\Gamma(1/2\eta)}{\Gamma(1/2\eta + q_2/2)} |\boldsymbol{\Sigma}_{22}|^{1/2} (1 + c(\eta)\delta_2^2)^{(1/2\eta + q_2/2)} \\ &= \left(\frac{c(\eta)}{\pi}\right)^{q_1/2} \frac{\Gamma(1/2\eta + q_2/2 + q_1/2)}{\Gamma(1/2\eta + q_2/2)} |\boldsymbol{\Sigma}|^{-1/2} |\boldsymbol{\Sigma}_{22}|^{1/2} \\ &\times (1 + c(\eta)\delta^2)^{-(1/2\eta + d/2)} (1 + c(\eta)\delta_2^2)^{(1/2\eta + q_2/2)} \end{aligned}$$

donde $\delta_2^2 = (\mathbf{x}_2 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ y $d = q_1 + q_2$.

Se tiene que

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1} & -(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ & \boldsymbol{\Sigma}_{22}^{-1} + \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12} \end{bmatrix},$$

de modo que $|\boldsymbol{\Sigma}|^{-1/2}|\boldsymbol{\Sigma}_{22}|^{1/2} = |\mathbf{B}_{11}|^{1/2}$ y δ^2 se puede reescribir como

$$\begin{aligned} \delta^2 &= \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix}^T \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix} \\ &= (\mathbf{x}_1 - \boldsymbol{\mu}_{1|2})^T \mathbf{B}_{11}(\mathbf{x}_1 - \boldsymbol{\mu}_{1|2}) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ &= \delta_{1|2}^2 + \delta_2^2 \end{aligned}$$

donde $\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 - \mathbf{B}_{11}^{-1} \mathbf{B}_{12}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$.

Luego, como $\delta^2 = \delta_{1|2}^2 + \delta_2^2$, se tiene que $(1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)} (1 + c(\eta)\delta_2^2)^{(1/2\eta+q_2/2)}$ es igual a $\left(1 + c(\eta)\delta_{1|2}^2 + c(\eta)\delta_2^2\right)^{-(1/2\eta+d/2)} (1 + c(\eta)\delta_2^2)^{(1/2\eta+q_2/2)}$ y multiplicando por $(1 + c(\eta)\delta_2^2)^{-(1/2\eta+d/2)}/(1 + c(\eta)\delta_2^2)^{-(1/2\eta+d/2)}$ la expresión queda

$$\left(1 + \frac{c(\eta)\delta_{1|2}^2}{1 + c(\eta)\delta_2^2}\right)^{-(1/2\eta+d/2)} (1 + c(\eta)\delta_2^2)^{-q_1/2}.$$

Por lo tanto, la fdp condicional está dada por

$$p(\mathbf{x}_1|\mathbf{x}_2) = \left(\frac{c^*(\eta)}{\pi}\right)^{q_1/2} \frac{\Gamma((1/2\eta + q_2/2) + q_1/2)}{\Gamma(1/2\eta + q_2/2)} |\mathbf{B}_{11}|^{1/2} (1 + c^*(\eta)\delta_{1|2}^2)^{-(1/2\eta+q_2/2)-q_1/2}$$

con $c^*(\eta) = c(\eta)/(1 + c(\eta)\delta_2^2)$.

Además notar que

$$\begin{aligned} \mathbf{X}_{1|2} &= \boldsymbol{\mu}_{1|2} + \frac{\mathbf{Z}_{1|2}}{U_{1|2}}, \mathbf{Z}_{1|2} \sim N_{q_1}(\mathbf{0}, \mathbf{B}_{11}^{-1}) \text{ y } U_{1|2} \sim \Gamma(\alpha_{1|2}, \beta_{1|2}) \text{ con } \alpha_{1|2} = (1/\eta + q_2) \text{ y} \\ \beta_{1|2} &= 1/c(\eta) + \delta_2^2 \text{ entonces } \boldsymbol{\mu}_{1|2} = E(\mathbf{X}_{1|2}) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{11} \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2) \text{ y} \\ \boldsymbol{\Sigma}_{1|2} &= V(\mathbf{X}_{1|2}) = E\left(\frac{\mathbf{Z}_{1|2} \mathbf{Z}_{1|2}^T}{U_{1|2}}\right) = \frac{\beta_{1|2}}{\alpha_{1|2}} \mathbf{B}_{11}^{-1}. \end{aligned}$$

Finalmente, así se obtiene la distribución de \mathbf{X}_1 dado \mathbf{X}_2 dado en (2.2.8).

2.2.7. Distribución de forma cuadrática

Si $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ y $\delta^2 = (\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})^T$ la distancia de Mahalanobis.

Se tiene que

$$F = \frac{\delta^2}{(1 - 2\eta)d} \sim F(d, 1/\eta). \quad (2.2.9)$$

Demostración:

Considerando la representación (2.2.1) de modo que

$$\delta^2 = (\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) = \frac{[\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y}]}{U^*} \quad (2.2.10)$$

donde $[\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y}] \sim \chi^2(d)$ y $U^* = [U/c(\eta)] \sim \Gamma(1/2\eta, 1/2)$ o equivalentemente $U^* \sim \chi^2(1/\eta)$ se tiene que al dividir el numerador y denominador de (2.2.10) por sus respectivos grados de libertad se obtiene (2.2.9).

2.3. Función de verosimilitud

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios de modo que $\mathbf{X}_i = (X_{i1}, \dots, X_{id})^T$ para todo $i = 1, \dots, n$ sigue una distribución $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. La función de verosimilitud se denota por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n p(\mathbf{x}_i) = \prod_{i=1}^n \left\{ K_d(\eta) |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta) \delta_i^2)^{-(1/2\eta + d/2)} \right\} \quad (2.3.1)$$

con $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \eta)^T \in \Theta \subseteq \mathbb{R}^{d+q+1}$ es el vector de parámetros y $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ es el vector columna que contiene los diferentes elementos de $\boldsymbol{\Sigma}$. Notar que $d = \dim(\boldsymbol{\mu})$ y $q = \dim(\boldsymbol{\phi})$.

El operador $\text{vech}(\cdot) = \mathbf{D}_d^T \text{vec}(\cdot)$ donde \mathbf{D}_d matriz de duplicación de dimensión $d^2 \times (d/2)(d+1)$ y $\text{vec}(\cdot)$ es el operador que vectoriza una matriz apilando sus columnas (Ver Magnus y Neudecker (1988)). Luego, el logaritmo de (2.3.1) es

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \log(K_d(\eta)) - \frac{1}{2} \log(|\boldsymbol{\Sigma}|) - \frac{1}{2} \left(\frac{1}{\eta} + d \right) \log(1 + c(\eta) \delta_i^2) \right\} \\ &= n \log(K_d(\eta)) - \frac{n}{2} \log(|\boldsymbol{\Sigma}|) - \frac{1}{2} \left(\frac{1}{\eta} + d \right) \sum_{i=1}^n \log(1 + c(\eta) \delta_i^2) \end{aligned} \quad (2.3.2)$$

donde

$$\log(K_d(\eta)) = \frac{d}{2} \log\left(\frac{c(\eta)}{\pi}\right) + \log\left(\Gamma\left(\frac{1}{2\eta} + \frac{d}{2}\right)\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right).$$

2.4. Estimación

2.4.1. Antecedentes

Dempster *et al.* (1977) utilizan el algoritmo EM para encontrar los estimadores MV considerando datos completos y fijando los grados de libertad. Rubin (1983) muestra este resultado para el caso de la distribución t multivariada, adicionalmente Little y Rubin (1980) lo extienden para el problema de datos perdidos. Por otro lado, Lange *et al.* (1989) presentan un caso más general al considerar los grados de libertad desconocidos y en conjunto estimar todos los parámetros. Osorio y Galea (2016) describen una extensión para estimar los parámetros de una distribución t , la cual es implementada en la librería MVT de R.

2.4.2. Estimación de máxima verosimilitud

Sea $\mathbf{U}(\boldsymbol{\theta}) = \left(\mathbf{U}(\boldsymbol{\mu})^T, \mathbf{U}(\boldsymbol{\phi})^T, \mathbf{U}(\eta)^T \right)^T$ el vector score, la primera derivada del logaritmo de la función verosimilitud definida en (2.3.2), de modo que

$$\begin{aligned} \mathbf{U}(\boldsymbol{\mu}) &= \sum_{i=1}^n u_i \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \\ \mathbf{U}(\boldsymbol{\phi}) &= \frac{n}{2} \mathbf{D}_d^T \text{vec} \left(\boldsymbol{\Sigma}^{-1} \left\{ \frac{1}{n} \sum_{i=1}^n u_i (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T \right\} \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \right) \\ \mathbf{U}(\eta) &= \frac{1}{2\eta^2} \sum_{i=1}^n \{ c(\eta)(d - u_i \delta_i^2) - \psi(1/2\eta + d/2) + \psi(1/2\eta) + \log(1 + c(\eta)\delta_i^2) \} \end{aligned}$$

con $u_i = (1/\eta + d)/(1/c(\eta) + \delta_i^2)$ y $\psi'(z) = d \log(\Gamma(z))/dz$ es la derivada del logaritmo de la función gamma $\Gamma(z)$.

Además $\mathbf{N}_d = \frac{1}{2}(\mathbf{I}_{d^2} + \mathbf{K}_d)$ donde \mathbf{I}_{d^2} matriz identidad de dimensión d^2 y \mathbf{K}_d la matriz de conmutación dadas en Magnus y Neudecker (1988).

Para η fijo e igualando $\mathbf{U}(\boldsymbol{\mu})$ y $\mathbf{U}(\boldsymbol{\phi})$ a cero, se obtienen las ecuaciones de estimación para $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, dadas por

$$\boldsymbol{\mu} = \frac{\sum_{i=1}^n u_i \mathbf{x}_i}{\sum_{i=1}^n u_i} \text{ y } \boldsymbol{\Sigma} = \frac{1}{n} \sum_{i=1}^n u_i (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T$$

con $\mathbf{x}_i = (x_{i1}, \dots, x_{id})^T$, para todo $i = 1, \dots, n$.

Sin embargo, como los estimadores no pueden ser expresados explícitamente, es necesario recurrir a métodos iterativos para obtener las raíces de la función score. El algoritmo EM, implementado por Osorio y Galea (2016) para la distribución t multivariada, es utilizado para obtener dichas estimaciones.

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios desde una distribución $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. En el marco del algoritmo de esperanza y maximización, conocido como algoritmo EM, se considera un vector de datos observados \mathbf{x}_{obs} y un vector de datos perdidos (u_1, \dots, u_n) , de modo que se tiene un vector de datos completo $\mathbf{x}_c = (\mathbf{x}_{obs}, u_1, \dots, u_n)$.

Considerando $\mathbf{X}_i | U_i = u_i \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}/u_i)$ y $U_i \sim \Gamma(1/2\eta, 1/2c(\eta))$ para todo $i = 1, \dots, n$, se puede obtener la función de verosimilitud para datos completos denotada por $L_c(\boldsymbol{\theta}) = \prod_{i=1}^n p(\mathbf{x}_i, u_i)$, que puede ser factorizada como $\prod_{i=1}^n p(\mathbf{x}_i | u_i) p(u_i)$.

De acuerdo a los datos completos el logaritmo de la función de verosimilitud queda $\ell_c(\boldsymbol{\theta}) = \sum_{i=1}^n \log p(\mathbf{x}_i | u_i) + \sum_{i=1}^n \log p(u_i)$.

El algoritmo procede a calcular la esperanza de $\ell_c(\boldsymbol{\theta})$ condicionada a los datos observados, conocido como paso E (Esperanza), para encontrar un $\boldsymbol{\theta}^{(k+1)}$ que maximiza la esperanza condicionada de $\ell_c(\boldsymbol{\theta})$ conocido como paso M (Maximización). Estos pasos se resumen a continuación y el detalle está en el Apéndice A.2.

Paso E:

Para obtener la $(k+1)$ -ésima iteración se necesita calcular la esperanza de la función de verosimilitud de los datos completos condicionado a los datos observados en la iteración actual (k) , es decir $\mathbf{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}) = E(\ell_c(\boldsymbol{\theta}) | \mathbf{x}_{obs}, \boldsymbol{\theta}^{(k)})$.

Paso M:

Determinar un $\boldsymbol{\theta}^{(k+1)}$ de modo que $\mathbf{Q}(\boldsymbol{\theta}^{(k+1)}, \boldsymbol{\theta}^{(k)}) \geq \mathbf{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)})$ para todo $\boldsymbol{\theta} \in \Theta$.

Finalmente, los estimadores de máxima verosimilitud para la $(k+1)$ iteración, están dados por

$$\boldsymbol{\mu}^{(k+1)} = \frac{\sum_{i=1}^n u_i^{(k)} \mathbf{x}_i}{\sum_{i=1}^n u_i^{(k)}} \quad \text{y} \quad \boldsymbol{\Sigma}^{(k+1)} = \frac{1}{n} \sum_{i=1}^n u_i^{(k)} (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)}) (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})^T$$
$$\boldsymbol{\phi}^{(k+1)} = \text{vech} \left(\boldsymbol{\Sigma}^{(k+1)} \right) \quad \text{donde} \quad u_i^{(k)} = \frac{(1/\eta^{(k)} + d)}{(1/c(\eta^{(k)}) + \delta_i^2(\boldsymbol{\tau}^{(k)}))} \quad \text{y} \quad \boldsymbol{\tau}^{(k)} = (\boldsymbol{\mu}^{(k)}, \boldsymbol{\phi}^{(k)}).$$

Por otro lado, el parámetro de forma actualizado $\eta^{(k+1)}$ se obtiene al maximizar $\mathbf{Q}(\eta, \boldsymbol{\theta}^{(k)})$ dada en Apéndice A.2.

Notación: Si el algoritmo EM encuentra convergencia en la iteración $(k+1)$, entonces se denotarán los estimadores MV como $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}^{(k+1)}$, $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{(k+1)}$ y $\hat{\eta} = \eta^{(k+1)}$.

2.4.3. Matriz de información de Fisher

Sean $\mathbf{X}_1, \dots, \mathbf{X}_n$ vectores aleatorios independiente e idénticamente distribuidos, de modo que $\mathbf{X}_i \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ para todo $i = 1, \dots, n$. Se define la matriz de información de Fisher basado en (2.3.2) como

$$\mathcal{I}(\boldsymbol{\theta}) = V(\mathbf{U}(\boldsymbol{\theta})) = \sum_{i=1}^n E_{\boldsymbol{\theta}}(U_i(\boldsymbol{\theta})U_i^T(\boldsymbol{\theta})) = nE_{\boldsymbol{\theta}}(U_1(\boldsymbol{\theta})U_1^T(\boldsymbol{\theta})) = n\mathbf{I}(\boldsymbol{\theta})$$

y su estructura matricial es

$$\mathcal{I}(\boldsymbol{\theta}) = n \begin{pmatrix} \mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) & \mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{I}_{\phi\eta}^T(\boldsymbol{\theta}) & \mathbf{I}_{\eta\eta}(\boldsymbol{\theta}) \end{pmatrix} \quad (2.4.1)$$

con elementos dador por

$$\begin{aligned} \mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) &= c_\mu(\eta)\boldsymbol{\Sigma}^{-1} \\ \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) &= \frac{1}{4}\mathbf{D}_d^T\{2c_\phi(\eta)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\mathbf{N}_d + (c_\phi(\eta) - 1)\text{vec}(\boldsymbol{\Sigma}^{-1})\text{vec}^T(\boldsymbol{\Sigma}^{-1})\}\mathbf{D}_d \\ \mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) &= -\frac{c(\eta)(d+2)}{(1+\eta(d+2))(1+\eta d)}\mathbf{D}_d^T\text{vec}(\boldsymbol{\Sigma}^{-1}) \\ \mathbf{I}_{\eta\eta}(\boldsymbol{\theta}) &= -\frac{1}{2\eta^2} \left\{ \frac{d}{(1-2\eta)^2} \left[\frac{1+\eta d(1-4\eta)-8\eta^2}{(1+\eta d)(1+(d+2)\eta)} \right] - \frac{d\beta(\eta)}{d\eta} \right\} \end{aligned}$$

con $\mathbf{N}_d = \frac{1}{2}(\mathbf{I}_{d^2} + \mathbf{K}_d)$ donde \mathbf{I}_{d^2} matriz identidad y \mathbf{K}_d la matriz de conmutación de dimensión $d^2 \times d^2$ (ver Magnus y Neudecker (1988)) y

$$\begin{aligned} u_i &= \frac{(1/\eta + d)}{(1/c(\eta) + \delta_i^2)}, \quad c_\mu(\eta) = \frac{c_\phi(\eta)}{(1-2\eta)}, \quad c_\phi(\eta) = \frac{(1+2\eta)}{(1+(d+2)\eta)} \text{ y} \\ \frac{d\beta(\eta)}{d\eta} &= -\frac{1}{2\eta^2} \left[\psi' \left(\frac{1}{2\eta} + \frac{d}{2} \right) - \psi' \left(\frac{1}{2\eta} \right) \right] \text{ con } \psi'(z) = \frac{d \log(\Gamma(z))}{dz}, \end{aligned}$$

elementos dados en Osorio y Galea (2016) y calculados en el Apéndice A.3.

Luego para obtener la matriz de información de Fisher estimada, basta reemplazar los estimadores MV obtenidos con el algoritmo EM.

2.5. Relación con la distribución normal

La distribución normal puede ser vista convenientemente como un caso especial de la distribución t , pues cuando $\eta \rightarrow 0$ o equivalentemente $1/\eta \rightarrow \infty$, la distribución t converge a una distribución normal.

Este hecho se puede ver intuitivamente en la representación (2.2.1) la cual es influenciada por la fdp de U , pues $\mathbf{X}|U = u$ tiene fdp normal con matriz de

varianza-covarianza escalada, es decir $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}/u)$.

A continuación se presenta una propiedad y corolario que justifican la distribución límite de un vector aleatorio con fdp $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$.

2.5.1. Distribución límite

Propiedad 2.5.1. Si $\mathbf{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ y $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denota la fdp de una distribución normal. Se tiene que, si η tiende a cero, entonces

$$\lim_{1/\eta \rightarrow \infty} \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta) = N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ para todo } \mathbf{x} \in \mathbb{R}^d. \quad (2.5.1)$$

El siguiente Lema dado en Hogg y Craig (1978) y el Corolario dado en Tong (1990) son usados para demostrar la Propiedad 2.5.1.

Lema 2.5.1. Sean b y c constantes, y a un entero positivo, se tiene que

(1) Si $\psi(n) \rightarrow 0$ cuando $n \rightarrow \infty$ entonces

$$\lim_{n \rightarrow \infty} \left[1 + \frac{b}{n} + \frac{\psi(n)}{n} \right]^{cn} = \lim_{n \rightarrow \infty} \left[1 + \frac{b}{n} \right]^{cn} = e^{bc}$$

(2) Si $\Gamma(\cdot)$ es la función gamma, entonces

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+a)}{\Gamma(n)n^a} = 1.$$

Corolario 2.5.1. Sea \mathbf{X} un vector aleatorio con densidad definida en (2.1.1) y sea \mathbf{Y} un vector aleatorio con densidad $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Entonces

$$\lim_{1/\eta \rightarrow \infty} \mathbb{P}(\mathbf{X} \in A) = \mathbb{P}(\mathbf{Y} \in A)$$

para un conjunto borel-medible $A \in \mathbb{R}^d$. Es decir, \mathbf{X} converge en distribución a \mathbf{Y} cuando $1/\eta \rightarrow \infty$ o equivalentemente $\eta \rightarrow 0$.

Demostración:

Aplicando el Lema 2.5.1, se tiene que

$$\begin{aligned} (1 + c(\eta)\delta^2)^{-1/2\eta+d/2} &= \left(1 + \frac{\delta^2}{1/\eta} - \frac{\delta^2}{1/\eta} + c(\eta)\delta^2\right)^{-1/2\eta+d/2} \\ &= \left(1 + \frac{\delta^2}{1/\eta} + \frac{\psi(1/\eta)}{1/\eta}\right)^{-1/2\eta} \times \left(1 + \frac{\delta^2}{1/\eta} + \frac{\psi(1/\eta)}{1/\eta}\right)^{-d/2}. \end{aligned}$$

con $\psi(1/\eta) = 2\delta^2/(1/\eta - 2)$.

Luego, si $1/\eta \rightarrow \infty$ entonces,

$$\lim_{1/\eta \rightarrow \infty} (1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)} = e^{-\delta^2/2}$$

pues $\psi(1/\eta) \rightarrow 0$ cuando $1/\eta \rightarrow \infty$.

Por otro lado,

$$\lim_{1/\eta \rightarrow \infty} K_d(\eta)|\Sigma|^{-1/2} = (2\pi)^{-d/2}|\Sigma|^{-1/2}$$

pues

$$\lim_{1/\eta \rightarrow \infty} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)(1/2\eta)^{d/2}} \times \lim_{1/\eta \rightarrow \infty} \frac{\delta^2(1/\eta)}{(1/\eta - 2)} = 1.$$

Finalmente, por el teorema de convergencia de Scheffé dado en Resnick (1999), pág. 253, y el Corolario 2.5.1 dado en Tong (1990), pág. 212, la Propiedad 2.5.1 queda demostrada.

Los estimadores MV definidos para la distribución $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ convergen a los estimadores MV de una distribución $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ cuando $\eta \rightarrow 0$. Esto sucede pues $u_i \rightarrow 1$ cuando $\eta \rightarrow 0$, entonces

$$\hat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad \text{y} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T.$$

Aplicando el mismo principio a la matriz de información de Fisher asociada a

la distribución t , se obtiene la matriz de información de Fisher asociada a la distribución normal, ya que $c_\phi(\eta) \rightarrow 1$ y $c_\mu(\eta) \rightarrow 1$ cuando $\eta \rightarrow 0$. Por lo tanto, la matriz de información de Fisher es

$$\mathbf{I}(\boldsymbol{\theta}) = n \begin{pmatrix} \mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) \end{pmatrix}$$

donde

$$\mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{-1} \text{ y } \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) = \frac{1}{2} \mathbf{D}_d^T (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{N}_d$$

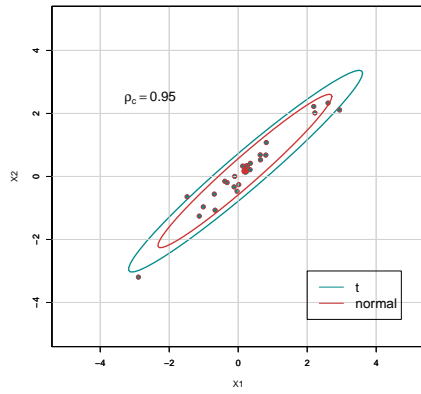
con \mathbf{D}_d y $\mathbf{N}_d = \frac{1}{2}(\mathbf{I}_{d^2} + \mathbf{K}_d)$ dadas en Magnus y Neudecker (1988).

Desde la fdp (2.1.1) se puede ver que el comportamiento de las colas de la distribución t depende del parámetro de forma; cuando η tiende a $1/2$ la distribución presenta colas más pesadas y cuando η tiende a cero converge a la distribución normal (ver Propiedad 2.5.1).

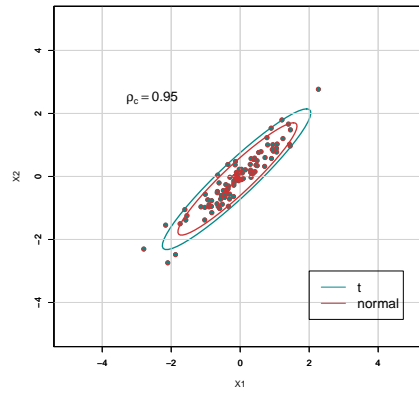
El parámetro de forma está relacionado con la curtosis de los datos, por ejemplo si el estimador de η está lejos de cero esto implica que el conjunto de datos presenta observaciones extremas y/o atípicas, y por ende la variabilidad será mayor que la normal.

Para $\eta = 0.25$ se simularon datos t bivariados de tamaño $n = 25$ y $n = 100$ bajo los escenarios de Lin (1989). La Figura 2.1 muestra la elipse de datos asumiendo que se distribuyen normal y la elipse de datos asumiendo que se distribuyen t . Se observa que en el caso de la elipse normal las gráficas (a), (c) y (d) muestran más observaciones como atípicas y/o extremas respecto al caso de la elipse t . Este comportamiento es más notorio en las gráficas (b), (d) y (f), es decir cuando $n = 100$.

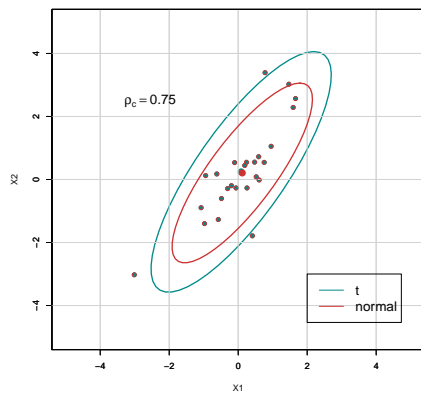
Por lo anterior la distribución t es ampliamente usada en presencia de observaciones atípicas y/o extremas, y también bajo el contexto de muestra pequeñas.



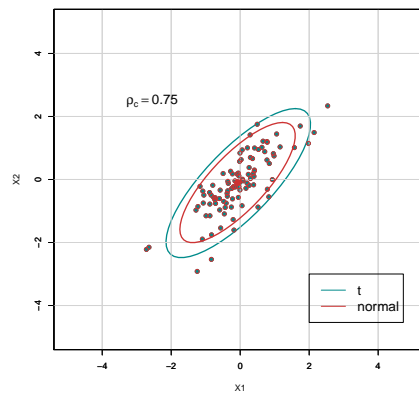
(a) CCC Alto y $n=25$



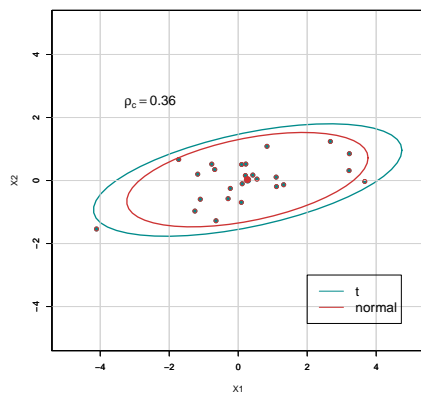
(b) CCC Alto y $n=100$



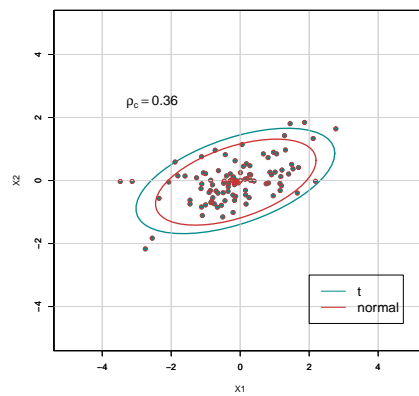
(c) CCC Medio y $n=25$



(d) CCC Medio y $n=100$



(e) CCC Bajo y $n=25$



(f) CCC Bajo y $n=100$

Figura 2.1: Elipses del 95 % de masa de probabilidad para datos simulados desde una distribución t basado en los escenarios de Lin (1989)

2.5.2. Bondad de ajuste

Un procedimiento gráfico usado por Osorio y Galea (2016) consiste en el uso de la distancia de Mahalanobis δ_i^2 para todo $i = 1, \dots, n$. Desde la ecuación (2.2.9) se sabe que si $\mathbf{X}_i \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ entonces $F_i = \delta_i^2 / (1 - 2\eta)d \sim F(d, 1/\eta)$ y cuando η tiende a cero, por la Propiedad (2.5.1), se tiene que $\mathbf{X}_i \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ entonces $F_i = \delta_i^2 \sim \chi^2(d)$.

La transformación de Wilson-Hilferty dada en Johnson *et al.* (1994) para la distancia de Mahalanobis, asumiendo que los datos siguen una distribución t , es

$$z_i = \frac{(1 - \frac{2\eta}{9})F_i^{1/3} - (1 - \frac{2}{9d})}{(\frac{2\eta}{9}F_i^{2/3} + \frac{2}{9d})^{1/2}},$$

mientras que si los datos siguen una distribución normal, es

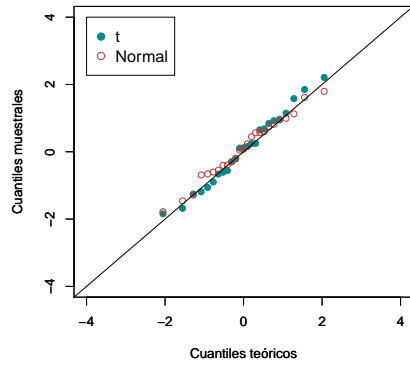
$$z_i = \frac{F_i^{1/3} - (1 - \frac{2}{9d})}{(\frac{2}{9d})^{1/2}}.$$

Como los z_i siguen una distribución aproximadamente normal con media 0 y varianza 1 Lange *et al.* (1989) sugieren evaluar la calidad del ajuste de los datos a través de las gráficas cuantil-cuantil (QQ plot en inglés).

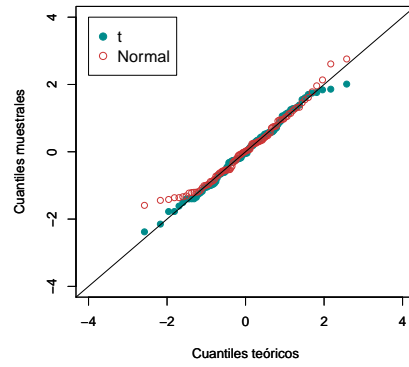
A continuación se presentan algunos ejemplos del uso de la transformación de Wilson-Hilferty en datos simulados y reales.

La Figura 2.2 presenta el QQ plot para la transformación Wilson-Hilferty de la distancia de Mahalanobis para datos simulados desde una distribución t y normal bivariadas, basado en los escenarios de Lin (1989), se puede observar que los datos siguen la distribución de cual fueron generados.

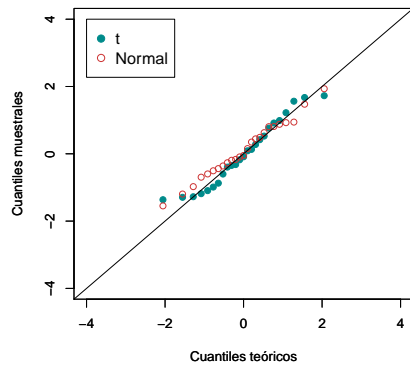
La Figura 2.3 muestra el mismo ejemplo anterior pero con datos reales, notar que algunos conjuntos de datos no siguen la distribución normal pero si la distribución t y en otros casos no siguen ninguna de las dos distribuciones.



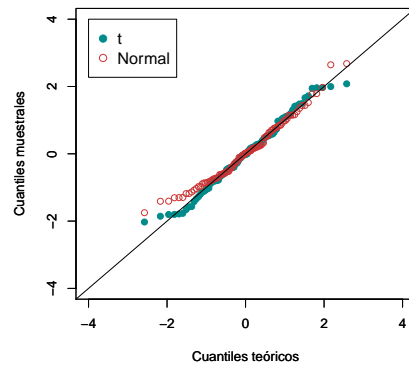
(a) CCC 0.35 y $n=25$



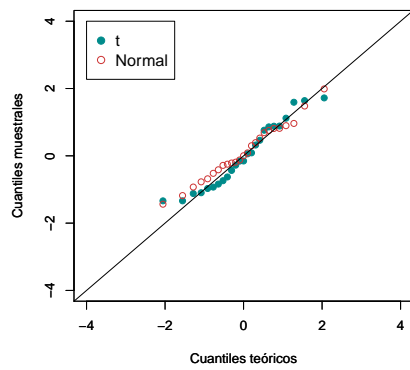
(b) CCC 0.35 y $n=100$



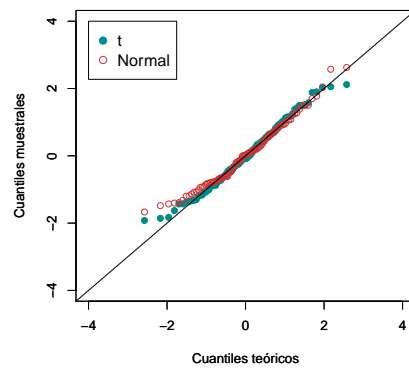
(c) CCC 0.75 y $n=25$



(d) CCC 0.75 y $n=100$

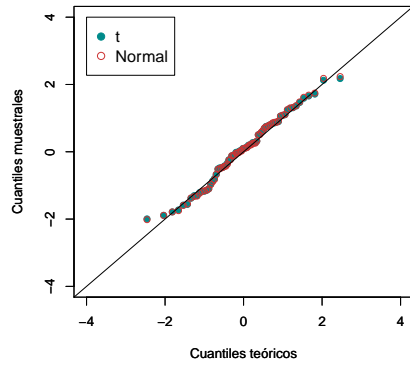


(e) CCC 0.95 y $n=25$

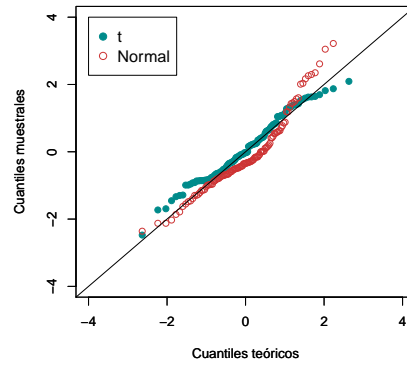


(f) CCC 0.95 y $n=100$

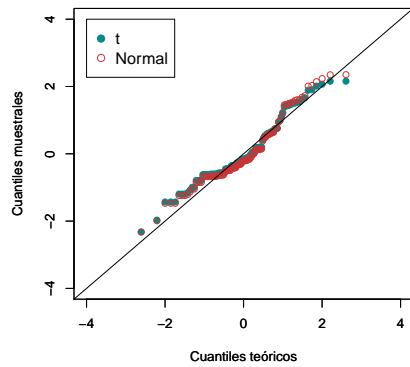
Figura 2.2: QQ plots para la transformación Wilson-Hilferty de la distancia de Mahalanobis considerando datos simulados bajo los escenarios de Lin (1989)



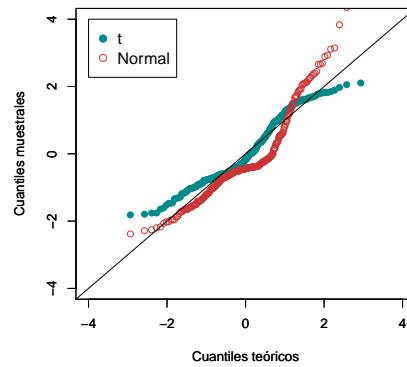
(a) Datos reales $n=72$



(b) Datos reales $n=118$



(c) Datos reales $n=110$



(d) Datos reales $n=299$

Figura 2.3: QQ plots para la transformación Wilson-Hilferty de la distancia de Mahalanobis considerando datos reales

2.6. Comentarios finales

Como se ha expuesto a lo largo del capítulo, la distribución t multivariada es una alternativa para modelar adecuadamente datos que presentan observaciones atípicas y/o extremas, además esta distribución queda especificada por el vector de medias y la matriz de varianza-covarianza, permitiendo una comparación con la distribución normal multivariada.

Coeficiente de correlación de concordancia

Este capítulo presenta el Coeficiente de Correlación de Concordancia (CCC), propuesto por Lin (1989), que mide el grado de acuerdo entre medidas generadas por dos instrumentos o métodos y una extensión, propuesta por Barnhart *et al.* (2002), que mide el grado de acuerdo entre medidas generadas por dos o más instrumentos o métodos.

Adicionalmente, se presenta la estimación y la derivación de la distribución asintótica para la construcción de intervalos de confianza y estadístico de prueba asintótico.

3.1. Concordancia entre dos medidas

Cuando $d = 2$ se tiene que $\mathbf{X}_i = (X_{i1}, X_{i2})^T$ para todo $i = 1, \dots, n$ es un vector aleatorio seleccionado independientemente desde una población con media $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ y matriz de varianza-covarianza dada por

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

Lin (1989) define que el grado de acuerdo entre X_{i1} y X_{i2} puede ser caracterizado por un índice absoluto llamado Desvío Cuadrado Medio (DCM) dado por

$$E[(X_{i1} - X_{i2})^2] = (\sigma_{11} + \sigma_{22} - 2\sigma_{12}) + (\mu_1 - \mu_2)^2 \quad (3.1.1)$$

para todo $i = 1, \dots, n$.

El DCM representa el valor esperado del desvío al cuadrado perpendicular desde la línea de concordancia. Si para cada par $(X_{i1}, X_{i2})^T$ existe un acuerdo perfecto en la población se espera que DCM sea cero.

Un índice escalado a partir de (3.1.1) es el CCC definido como

$$\begin{aligned} \rho_c &= 1 - \frac{E[(X_{i1} - X_{i2})^2]}{E[(X_{i1} - X_{i2})^2 | X_{i1}, X_{i2} \text{ no correlacionadas}]} \\ &= \frac{2\sigma_{12}}{\sigma_{11} + \sigma_{22} + (\mu_1 - \mu_2)^2} \\ &= \rho_{12}C_{12} \end{aligned} \quad (3.1.2)$$

donde ρ_{12} es el coeficiente de precisión (correlación de Pearson) entre las medidas generadas por dos métodos y C_{12} es el coeficiente de exactitud (sesgo) respecto a la línea de concordancia, tal que $0 < C_{12} \leq 1$, $-1 \leq \rho_{12} \leq 1$ y $-1 \leq \rho_c \leq 1$.

Además, cabe destacar que Lin (1989) descompone el coeficiente de exactitud en función de $a = (\mu_1 - \mu_2)/(\sigma_{11}\sigma_{22})^{1/4}$ (cambio de localización) y $b = \sqrt{\sigma_{11}/\sigma_{22}}$ (cambio de escala), de modo que el coeficiente de exactitud puede escribirse como $C_{12} = 2[b + b^{-1} + a^2]^{-1}$.

Si C_{12} toma valores cercanos a cero entonces las observaciones están más lejos de la línea de concordancia y si toma valores cercano a uno entonces las observaciones están más cerca de la línea de concordancia.

Algunas características del CCC son:

1. Si $\rho_{12} = \pm 1$ entonces $\rho_c = \pm 1$, lo que indica concordancia absoluta entre las medidas.
2. Si $\rho_{12} = 0$ entonces $\rho_c = 0$, lo que indica falta de concordancia entre las medidas
3. Si $C_{12} = 1$ (o equivalentemente $\mu_1 = \mu_2$ y $\sigma_{11} = \sigma_{22}$) entonces $\rho_c = \rho_{12}$ lo cual implica que no hay desviación de las observaciones desde la línea de concordancia.

La interpretación del coeficiente (3.1.2) es que mide el grado de acuerdo entre medidas cuantitativas, de modo que el coeficiente de precisión mide que tan cerca están las medidas de la mejor línea ajustada y el coeficiente de exactitud mide que tan cerca la mejor línea ajustada está de la línea de concordancia.

Un estimador natural del CCC es el obtenido a través del método de los momentos y se define por

$$\tilde{\rho}_c = \frac{2S_{12}}{S_{11} + S_{22} + (\bar{X}_1 - \bar{X}_2)^2}$$

donde

$$S_{12} = \frac{1}{n} \sum_{i=1}^n (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2), \quad \bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij} \quad \text{y} \quad S_{jj} = \frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2$$

para todo $j = 1, 2$.

Para encontrar la distribución asintótica de $\tilde{\rho}_c$ y construir un intervalo de confianza asintótico, Lin (1989) propone aplicar la Z -transformación de Fisher.

Sea $Z = (1/2) \ln \{(1 + \rho_c) / (1 - \rho_c)\} = \tanh^{-1}(\rho_c)$ la Z -transformación de Fisher para ρ_c , aplicando el Corolario 3.3 dado en Serfling (1980), Lin (1989) demuestra

que \tilde{Z} sigue una distribución asintótica con media Z y varianza

$$\sigma_{\tilde{Z}}^2 = \frac{1}{n} \left[\frac{(1 - \rho_{12}^2)\rho_c^2}{(1 - \rho_c^2)\rho_{12}^2} + \frac{4\rho_c^3(1 - \rho_c)a^2}{\rho_{12}(1 - \rho_c^2)^2} - \frac{\rho_c^4 a^4}{2\rho_{12}^2(1 - \rho_c^2)^2} \right]$$

donde ρ_{12} es el coeficiente de precisión (correlación de Pearson) y $a^2 = (\mu_1 - \mu_2)^2 / \sqrt{\sigma_{11}\sigma_{22}}$ el cambio de localización al cuadrado.

Sea $g(Z) = \tanh(Z)$, aplicando el método delta se tiene que $g(\hat{Z})$ sigue una distribución asintóticamente normal con media $g(Z)$ y varianza $\sigma_Z^2 [dg(Z)/dZ]$.

Como la derivada de la función tangente hiperbólica es $dg(Z)/dZ = 1 - \tanh(Z)^2$ y reemplazando $\tanh(Z) = \rho_c$ se tiene que

$$\frac{(\tilde{\rho}_c - \rho_c)}{\sqrt{\tilde{\sigma}_{\tilde{Z}}^2}} \xrightarrow{\mathfrak{D}} N(0, (1 - \rho_c^2)) \quad (3.1.3)$$

donde $\tilde{\sigma}_{\tilde{Z}}^2$ es consistente para $\sigma_{\tilde{Z}}^2$.

Por lo tanto, un intervalo de confianza asintótico del $(1 - \alpha)100\%$ de confianza, basado en la distribución asintótica de $\tilde{\rho}_c$ dada en (3.1.3), está dado por

$$IC(\rho_c) = \left\{ \tilde{\rho}_c \pm z_{1-\alpha/2} \sqrt{(1 - \tilde{\rho}_c)\tilde{\sigma}_{\tilde{Z}}^2} \right\},$$

y otro en base a la Z -transformación de Fisher, está dado por

$$\left\{ \frac{\exp(2z_{LI}) - 1}{\exp(2z_{LI}) + 1}, \frac{\exp(2z_{LS}) - 1}{\exp(2z_{LS}) + 1} \right\}$$

donde $z_{LI} = \tilde{Z} - z_{1-\alpha/2}\tilde{\sigma}_{\tilde{Z}}$ y $z_{LS} = \tilde{Z} + z_{1-\alpha/2}\tilde{\sigma}_{\tilde{Z}}$ con $z_{1-\alpha/2}$ el percentil de la distribución normal estándar.

Una función para calcular ambos intervalos de confianza, dada en Feng *et al.* (2014), está en la librería `agRee` de R.

3.2. Concordancia entre dos o más medidas

Sea $\mathbf{X}_i = (X_{i1}, \dots, X_{id})^T$ un vector aleatorio con vector de medias $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^T$ y matriz de varianza-covarianza $\boldsymbol{\Sigma}$ de dimensión $(d \times d)$, de modo que X_{ij} denota la medida para el i -ésima unidad de análisis obtenida con el j -ésimo instrumento o método, para todo $i = 1, \dots, n$ y $j = 1, \dots, d$.

Una extensión del CCC para dos o más medidas es el Coeficiente de Correlación de Concordancia Global (CCCG) propuesto por Barnhart *et al.* (2002).

Debido a que la diferencia al cuadrado $(X_{ij} - X_{ik})^2$ para todo $j \neq k$ es usada para describir desacuerdo entre medidas obtenidas por dos instrumentos, es posible considerar, para describir el desacuerdo de medidas obtenidas entre dos o más instrumentos, la variabilidad muestral inter-instrumento como

$$\mathbf{V} = \frac{\sum_{j=1}^d (X_{ij} - \bar{X}_i)^2}{(d-1)} \quad (3.2.1)$$

donde $\bar{X}_i = d^{-1} \sum_{j=1}^d X_{ij}$ es el promedio entre las medidas para el i -ésimo individuo y es usado en lugar de las diferencia al cuadrado para describir la variabilidad entre medidas generadas por múltiples instrumentos.

La expresión (3.2.1) puede ser escrita como una combinación lineal de las diferencias al cuadrado entre las medidas, es decir

$$\mathbf{V} = \frac{1}{d(d-1)} \sum_{j=1}^{d-1} \sum_{k=j+1}^d (X_{ij} - X_{ik})^2.$$

Luego Barnhart *et al.* (2002) propone el CCCG, sin considerar un instrumento de referencia, como

$$\rho_g = 1 - \frac{E[\mathbf{V}]}{E[\mathbf{V}|X_{i1}, \dots, X_{id} \text{ no correlacionadas}]}$$

donde la esperanza de \mathbf{V} , cuando las medidas están correlacionadas, es

$$\begin{aligned}
E[\mathbf{V}] &= \frac{1}{d(d-1)} \sum_{j=1}^{p-1} \sum_{k=j+1}^d \{V[(X_{ij} - X_{ik})] + E[(X_{ij} - X_{ik})]^2\} \\
&= \frac{1}{d(d-1)} \sum_{j=1}^{p-1} \{\sigma_{jj} + \sigma_{kk} - 2\sigma_{jk} + (\mu_j - \mu_k)^2\}
\end{aligned}$$

y cuando no lo están

$$E[\mathbf{V}|X_{i1}, \dots, X_{id} \text{ no correlacionadas}] = \frac{1}{d(d-1)} \sum_{j=1}^{d-1} \{\sigma_{jj} + \sigma_{kk} + (\mu_j - \mu_k)^2\}.$$

Si $\xi_{jk} = \sigma_{jj} + \sigma_{kk} + (\mu_j - \mu_k)^2$ el CCCG toma la forma dada por

$$\rho_g = \frac{\sum_{j=1}^{d-1} \sum_{k=j+1}^d 2\sigma_{jk}}{\sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk}} \quad (3.2.2)$$

y al amplificar por ξ_{jk} el numerador de (3.2.2) se obtiene el CCCG en función del CCC de a pares, definido como

$$\rho_g = \frac{\sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk} \rho_c^{jk}}{\sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk}} \quad (3.2.3)$$

donde ρ_c^{jk} es el CCC para todo $j \neq k$.

La expresión (3.2.3) puede ser interpretada como un promedio ponderado de todos los CCC con pesos ξ_{jk} , de modo que los pares de medidas son penalizados proporcionalmente por el grado de desacuerdo entre pares. Además, $\rho_c^{jk} = C_{jk} \rho_{jk}$ donde C_{jk} es el coeficiente de exactitud de a pares y ρ_{jk} es el coeficiente de precisión de a pares, entonces (3.2.3) queda

$$\rho_g = \frac{\sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk} C_{jk} \rho_{jk}}{\sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk}}. \quad (3.2.4)$$

Luego, al igual que el CCC bivariado, (3.2.4) está en función de la precisión y la exactitud, más aún, si $d = 2$ se tiene que ρ_g es igual a ρ_c definido en (3.1.2).

Finalmente, el CCCG definido en (3.2.2), puede ser reescrito en términos de sus medias, varianzas y covarianzas.

Sea

$$\begin{aligned} \sum_{j=1}^{d-1} \sum_{k=j+1}^d \xi_{jk} &= \sum_{j=1}^{d-1} \sum_{k=j+1}^d (\sigma_{jj} + \sigma_{kk}) + \sum_{j=1}^{d-1} \sum_{k=j+1}^d (\mu_j - \mu_k)^2 \\ &= (d-1) \sum_{j=1}^d \sigma_{jj} + \sum_{j=1}^{d-1} \sum_{k=j+1}^d (\mu_j - \mu_k)^2 \end{aligned}$$

entonces

$$\rho_g = \frac{\sum_{j=1}^{d-1} \sum_{k=j+1}^d 2\sigma_{jk}}{(d-1) \sum_{j=1}^d \sigma_{jj} + \sum_{j=1}^{d-1} \sum_{k=j+1}^d (\mu_j - \mu_k)^2} \quad (3.2.5)$$

donde (3.2.5) tiene estructura matricial como

$$\rho_g = \frac{2\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T)}{(d-1)\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_D) + (\mathbf{C}\boldsymbol{\mu})^T (\mathbf{C}\boldsymbol{\mu})} \quad (3.2.6)$$

con $\boldsymbol{\Sigma}_T$ la matriz triangular inferior sin la diagonal de $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_D = \text{Diag}(\sigma_{11}, \dots, \sigma_{pp})$ la matriz diagonal y \mathbf{C} matriz de dimensión $(r \times d)$ que genera las diferencias del tipo $(\mu_j - \mu_k)$ con $r = \binom{d}{2}$ comparaciones posibles, de modo que, $\boldsymbol{\mu}^T \mathbf{C}^T \mathbf{C} \boldsymbol{\mu}$ genera estas diferencias al cuadrado. Además, el numerador de (3.2.6) se puede reescribir como $\text{vech}(\boldsymbol{\Sigma}_T) = \text{vech}(\mathbf{E}_T)$ con $\mathbf{E}_T = \mathbf{M}_\xi \odot \mathbf{M}_C \odot \mathbf{M}_\rho$ donde \odot indica el producto Hadamard, dado en Magnus y Neudecker (1988), y

$$\begin{aligned} \mathbf{M}_\xi &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \xi_{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{1d} & \xi_{2d} & \cdots & 0 \end{pmatrix}, \quad \mathbf{M}_C = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ C_{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_{1d} & C_{2d} & \cdots & 0 \end{pmatrix}, \\ \mathbf{M}_\rho &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \rho^{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{1d} & \rho^{2d} & \cdots & 0 \end{pmatrix}. \end{aligned}$$

3.3. Inferencia

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios iid tal que $\mathbf{X}_i \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ para todo $i = 1, \dots, n$ y sea $\boldsymbol{\rho}(\boldsymbol{\theta})$ el coeficiente de correlación de concordancia entre medidas generadas por dos o más instrumentos, donde $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \eta)^T \in \Theta \subseteq \mathbb{R}^{d+q+1}$ con $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$.

Esta sección presenta el estimador MV, el intervalo de confianza asintótico y estadístico de prueba asintótico para el $\boldsymbol{\rho}(\boldsymbol{\theta})$.

Se propone someter a prueba las siguientes hipótesis estadísticas: (a) $H_0: \boldsymbol{\rho}(\boldsymbol{\theta}) = 0$ versus $H_1: \boldsymbol{\rho}(\boldsymbol{\theta}) \neq 0$ y (b) $H_0: \boldsymbol{\rho}(\boldsymbol{\theta}) \leq \boldsymbol{\rho}^*(\boldsymbol{\theta})$ versus $H_0: \boldsymbol{\rho}(\boldsymbol{\theta}) > \boldsymbol{\rho}^*(\boldsymbol{\theta})$.

La hipótesis (a) permite probar si el coeficiente es estadísticamente significativo y la hipótesis (b) permite probar si existe concordancia entre las medidas respecto a un valor preestablecido por el investigador.

3.3.1. Estimación máximo verosímil

El algoritmo EM, presentado en el capítulo 2, permite encontrar el estimador MV de $\boldsymbol{\theta}$ denotado por $\hat{\boldsymbol{\theta}}$ y por la propiedad de invarianza de los estimadores MV, dado en Lehmann y Casella (1998), el estimador MV del CCC y CCCG es $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}})$.

3.3.2. Distribución asintótica

Aplicando el Teorema 2.4.1 dado en Lehmann y Casella (1998), se tiene que si $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ una secuencia de estimadores MV que es consistente para $\boldsymbol{\theta}$ entonces $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n)$ es una secuencia de estimadores MV que es consistente para $\boldsymbol{\rho}(\boldsymbol{\theta})$.

Luego, bajo ciertas condiciones de regularidad dadas en Lehmann (1999) pag. 456, se sabe que $\hat{\boldsymbol{\theta}}$ sigue una distribución asintótica dada en el siguiente teorema.

Teorema 3.3.1. Sea $\hat{\boldsymbol{\theta}}_n$ una secuencia de estimadores MV consistentes para $\boldsymbol{\theta}$, bajo ciertas condiciones de regularidad, se tiene que

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{I}(\boldsymbol{\theta})^{-1}) \quad (3.3.1)$$

donde $n\mathbf{I}(\boldsymbol{\theta})$ es la matriz de información de Fisher.

Luego, la distribución asintótica del coeficiente de concordancia es dada en la siguiente proposición.

Proposición 3.3.1. Sea $\hat{\boldsymbol{\theta}}_n$ una secuencia de estimadores consistentes que sigue una distribución asintótica dada en (3.3.1) y sea $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n)$ una secuencia de estimadores consistentes para $\boldsymbol{\rho}(\boldsymbol{\theta})$. Entonces,

$$\sqrt{n}(\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n) - \boldsymbol{\rho}(\boldsymbol{\theta})) \xrightarrow{\mathcal{D}} N(0, \sigma_\rho^2)$$

donde $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^{d+q+1}$, $\sigma_\rho^2 = \mathbf{d}^T \mathbf{I}(\boldsymbol{\theta})^{-1} \mathbf{d}$ y $\mathbf{d} = [\partial \boldsymbol{\rho}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}]$ está dado en el Apéndice A.4.

Demostración: Aplicando el método delta definido en Serfling (1980), se obtiene la distribución asintótica para $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n)$.

Observación: La Proposición 3.3.1 permite encontrar la distribución asintótica del CCC entre medidas generadas por dos o más de dos instrumentos o métodos.

3.3.3. Intervalo de confianza asintótico

Sea $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n)$ una secuencia de estimadores máximo verosímiles para $\boldsymbol{\rho}(\boldsymbol{\theta})$ y sea $\sigma_\rho^2 = \mathbf{d}^T \mathbf{I}(\boldsymbol{\theta})^{-1} \mathbf{d}$, en base a la Proposición 3.3.1, se define una cantidad pivotal dada por

$$\mathbf{Q} = \frac{\sqrt{n}(\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_n) - \boldsymbol{\rho}(\boldsymbol{\theta}))}{\sqrt{\sigma_\rho^2}} \xrightarrow{\mathcal{D}} N(0, 1) \quad (3.3.2)$$

con media $\boldsymbol{\rho}(\boldsymbol{\theta})$ y varianza $n^{-1}\sigma_\rho^2$.

Por el Teorema 2.1.4, dado en Lehmann (1999), y considerando a $\mathbf{I}(\boldsymbol{\theta})$ como una función continua de $\boldsymbol{\theta}$, se sigue que $\mathbf{I}(\hat{\boldsymbol{\theta}}_n)$ converge en probabilidad a $\mathbf{I}(\boldsymbol{\theta})$, por lo tanto $\hat{\sigma}_\rho^2$ es un estimador consistente para σ_ρ^2 .

Luego, en base a (3.3.2), un intervalo de confianza asintótico para $\boldsymbol{\rho}(\boldsymbol{\theta})$ con nivel de confianza del $(1 - \alpha)100\%$ es

$$IC(\boldsymbol{\rho}(\boldsymbol{\theta})) = \left\{ \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) \pm z_{1-\alpha/2} \frac{\hat{\sigma}_\rho}{\sqrt{n}} \right\} \quad (3.3.3)$$

donde $z_{1-\alpha/2}$ es el percentil asociado a la distribución normal estándar.

Por otro lado, se sabe que la matriz de información observada

$$\mathbf{I}_o(\boldsymbol{\theta}) = - \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

calculada en el Apéndice A.3, al ser dividida por n es también un estimador consistente para $\mathbf{I}(\boldsymbol{\theta})$ (ver Lehmann (1999)). Como consecuencia se tiene otro estimador consistente para σ_ρ^2 , de modo que se puede construir otro intervalo de confianza como en (3.3.3).

3.3.4. Test de razón de verosimilitud

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios iid desde una distribución $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$, se considera $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \eta)^T$ el vector con contiene los parámetros de la distribución t , donde $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ es la matriz de varianza-covarianza vectorizada sin elementos repetidos, el cual, puede ser reescrito como $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}_1^T, \boldsymbol{\phi}_2^T, \eta)^T$ donde $\boldsymbol{\phi}_1$ contiene las covarianzas y $\boldsymbol{\phi}_2$ contiene las varianzas de $\boldsymbol{\Sigma}$.

La hipótesis (a) $H_0: \boldsymbol{\rho}(\boldsymbol{\theta}) = 0$ es equivalente a probar para $d = 2$ la hipótesis $H_0: \sigma_{12} = 0$ y para $d > 2$ la hipótesis $H_0: \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ donde $\boldsymbol{\Sigma}_0$ matriz diagonal, es decir, $\boldsymbol{\phi}_1 = 0$.

Sea $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}^T, \hat{\boldsymbol{\phi}}^T, \hat{\eta})^T$ el estimador MV bajo todo el espacio paramétrico Θ y sea

$\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\mu}}^T, \tilde{\boldsymbol{\phi}}_2^T, \tilde{\eta})^T$ el estimador MV bajo el espacio paramétrico restringido Θ_0 , es decir bajo H_0 .

El estadístico de prueba basado en la razón de verosimilitud derivado en el Apéndice A.5 queda expresado como

$$\Lambda = \left[\frac{K_d(\tilde{\eta})}{K_d(\hat{\eta})} \right]^n \left[\frac{|\tilde{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}|} \right]^{-n/2} \times \frac{\prod_{i=1}^n \left(1 + c(\tilde{\eta}) \tilde{\delta}_{i0}^2 \right)^{-(1/2\tilde{\eta}+d/2)}}{\prod_{i=1}^n \left(1 + c(\hat{\eta}) \hat{\delta}_i^2 \right)^{-(1/2\hat{\eta}+d/2)}} \quad (3.3.4)$$

con región crítica igual a $RC = \{\mathbf{x} : \Lambda(\mathbf{x}) \leq \lambda_0\}$ de modo que λ_0 es una constante tal que el estadístico sea de tamaño α y $\mathbf{x} \in \mathbb{R}^{n \times d}$.

Luego, por el Teorema 5.2.1 dado en Mardia *et al.* (1997), la distribución asintótica en base a (3.3.4) es

$$-2 \log(\Lambda) = -2 \left[\ell(\tilde{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}) \right] \xrightarrow{\mathfrak{D}} \chi^2(m)$$

con región crítica $RC = \{\mathbf{x} : -2 \log \Lambda(\mathbf{x}) > \chi_\alpha^2(m)\}$ donde $m = \dim(\Theta) - \dim(\Theta_0) = d + q + 1 - (2d + 1) = q - d$ son los grados de libertad y $\mathbf{x} \in \mathbb{R}^{n \times d}$.

Observación: Si $\eta \rightarrow 0$ la muestra de vectores aleatorios $\mathbf{X}_1, \dots, \mathbf{X}_n$ sigue una distribución $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y el TRV queda

$$\Lambda = \left[\frac{|\hat{\boldsymbol{\Sigma}}|}{|\tilde{\boldsymbol{\Sigma}}|} \right]^{n/2} = |\mathbf{R}|^{n/2}$$

donde \mathbf{R} es la matriz de correlación (ver Apéndice A.5) y por el Teorema 5.2.1 dado en Mardia *et al.* (1997) se tiene que el TRV asintótico es

$$-2 \log(\Lambda) = -n \log(|\mathbf{R}|) \xrightarrow{\mathfrak{D}} \chi^2(m)$$

y si $d = 2$ se tiene que $-2 \log(\Lambda) = -n \log(1 - \rho_{12}^2) \xrightarrow{\mathfrak{D}} \chi^2(1)$.

3.3.5. Prueba de Wald

La prueba de Wald se basa en la distribución aproximada de $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}})$ bajo la hipótesis nula $H_0 : \boldsymbol{\rho}(\boldsymbol{\theta}) = \boldsymbol{\rho}^*(\boldsymbol{\theta})$, es decir

$$\frac{\sqrt{n} \left(\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\rho}^*(\boldsymbol{\theta}) \right)}{\sqrt{\hat{\sigma}_\rho^2}} \xrightarrow{\mathcal{D}} N(0, 1).$$

Luego, estadístico de prueba es

$$W = \frac{\sqrt{n} \left(\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\rho}^*(\boldsymbol{\theta}) \right)}{\sqrt{\hat{\sigma}_\rho^2}} \quad (3.3.5)$$

donde $\hat{\sigma}_\rho^2$ es un estimador consistente para σ_ρ^2 y su región crítica está dada por $RC = \{\boldsymbol{x} : |W(\boldsymbol{x})| \geq z_{\alpha/2}\}$ o equivalentemente $RC = \{\boldsymbol{x} : W^2(\boldsymbol{x}) > \chi_\alpha^2(1)\}$ con $\boldsymbol{x} \in \mathbb{R}^{n \times d}$.

Por otro lado, si se desea someter a prueba (b) $H_0 : \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) \leq \boldsymbol{\rho}^*(\boldsymbol{\theta})$ el estadístico de prueba es el definido en (3.3.5) pero con región crítica dada por $RC = \{\boldsymbol{x} : W(\boldsymbol{x}) \geq z_\alpha\}$ con $\boldsymbol{x} \in \mathbb{R}^{n \times d}$.

3.4. Comentarios finales

Se ha estudiado que el CCC permite evaluar el grado de acuerdo entre medidas generadas por distintos métodos o instrumentos y que además se puede reescribir en función el coeficiente de precisión y exactitud. La inferencia sobre el CCC se basa en la distribución t multivariada, además utilizando propiedades asintóticas de los estimadores MV se construye un intervalo de confianza asintótico y estadísticos de prueba asintótico. El desempeño de los intervalos y estadístico de prueba son estudiados en el Apéndice B.

En el siguiente capítulo se estudia y desarrolla un método de diagnóstico para los estimadores MV y para el estimador del CCC.

Capítulo 4

Diagnóstico de influencia sobre el CCC

Este capítulo presenta los conceptos geométricos usados para el desarrollo y aplicación del método de diagnóstico de influencia local. Las Definiciones, Teoremas y Proposiciones están dadas en Do Carmo (1976) y Poon y Poon (2010).

4.1. Conceptos geométricos

Definición 4.1.1. Sea $f(\boldsymbol{\omega})$ una función real valorada sobre un conjunto abierto Ω en \mathbb{R}^n . La diferencial de f sobre $\boldsymbol{\omega}$ como una aplicación lineal es

$$df_{\boldsymbol{\omega}} : T_{\boldsymbol{\omega}}\mathbb{R}^n \rightarrow T_{f(\boldsymbol{\omega})}\mathbb{R}, \quad (4.1.1)$$

donde $T_{\boldsymbol{\omega}}\mathbb{R}^n$ es el espacio tangente de \mathbb{R}^n sobre el punto $\boldsymbol{\omega}$ y $T_{f(\boldsymbol{\omega})}\mathbb{R}$ es el espacio tangente de \mathbb{R} sobre el punto $f(\boldsymbol{\omega})$.

Definición 4.1.2. Suponga que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida sobre un conjunto abierto Ω en \mathbb{R}^n en un punto $\boldsymbol{\omega}_0$. Dado un vector unitario \mathbf{h} llamamos, provisto de que el límite exista,

$$d_{\mathbf{h}}f(\boldsymbol{\omega}_0) = \lim_{a \rightarrow 0} \frac{f(\boldsymbol{\omega}_0 + a\mathbf{h}) - f(\boldsymbol{\omega}_0)}{a} \quad (4.1.2)$$

la derivada direccional de f en la dirección \mathbf{h} sobre $\boldsymbol{\omega}_0$.

Observación: La derivada direccional (4.1.2) también se puede definir por

$$\left. \frac{d}{da} f(\boldsymbol{\omega}_0 + a\mathbf{h}) \right|_{a=0}.$$

Definición 4.1.3. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida en un subconjunto abierto Ω en \mathbb{R}^n y diferenciable, el gradiente de f en $\boldsymbol{\omega}$ es el vector en \mathbb{R}^n dado por

$$\nabla_f(\boldsymbol{\omega}) = \left(\frac{\partial f}{\partial \omega_1}, \frac{\partial f}{\partial \omega_2}, \dots, \frac{\partial f}{\partial \omega_n} \right)^T. \quad (4.1.3)$$

Teorema 4.1.1. Suponga que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida en un subconjunto abierto Ω en \mathbb{R}^n que contiene el punto $\boldsymbol{\omega}_0$. Entonces, para cualquier vector unitario \mathbf{h} , $d_{\mathbf{h}}f(\boldsymbol{\omega}_0)$ existe y es

$$d_{\mathbf{h}}f(\boldsymbol{\omega}_0) = \langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle. \quad (4.1.4)$$

donde $\langle \cdot, \cdot \rangle$ corresponde al producto punto.

Demostración:

Sea $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{h}$ con $a \in \mathbb{R}$, $\mathbf{h} \in \mathbb{R}^n$ y $\boldsymbol{\omega} \in \Omega \subset \mathbb{R}^n$ tal que $f(\boldsymbol{\omega}_0 + a\mathbf{h}) = f(\boldsymbol{\omega}(a))$, luego aplicando la regla de la cadena se tiene

$$\frac{d}{da} f(\boldsymbol{\omega}(a)) = \sum_{i=1}^n \frac{\partial f}{\partial \omega_i} \frac{d\omega_i}{da} = \left(\frac{\partial f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \right)^T \frac{d\boldsymbol{\omega}(a)}{da} = \nabla_f(\boldsymbol{\omega})^T \mathbf{h} = \langle \nabla_f(\boldsymbol{\omega}), \mathbf{h} \rangle,$$

luego (4.1.4) queda demostrado.

Observación: Si $a = 0$ entonces $\boldsymbol{\omega}(0) = \boldsymbol{\omega}_0$ luego

$$\left. \frac{d}{da} f(\boldsymbol{\omega}(a)) \right|_{a=0} = \langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle$$

donde $\nabla_f(\boldsymbol{\omega}_0)$ está dado en (4.1.3)

Proposición 4.1.1. Suponga que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en un subconjunto abierto Ω en \mathbb{R}^n que contiene el punto $\boldsymbol{\omega}_0$, sea \mathbf{h} un vector unitario y sea $\|\cdot\|$ la norma de un vector. Luego $d_{\mathbf{h}}f(\boldsymbol{\omega}_0)$ tiene un valor máximo $\|\nabla_f(\boldsymbol{\omega}_0)\|$ cuando \mathbf{h} es la dirección de $\nabla_f(\boldsymbol{\omega}_0)$ y un valor mínimo $-\|\nabla_f(\boldsymbol{\omega}_0)\|$ cuando \mathbf{h} es la dirección de $-\nabla_f(\boldsymbol{\omega}_0)$.

Demostración:

Por el Teorema 4.1.1 se sabe que $d_{\mathbf{h}}f(\boldsymbol{\omega}_0) = \langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle$, luego por definición de producto escalar entre dos vectores $\langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle = \|\nabla_f(\boldsymbol{\omega}_0)\| \|\mathbf{h}\| \cos(\gamma)$. Como $\|\mathbf{h}\| = 1$ se tiene que

$$d_{\mathbf{h}}f(\boldsymbol{\omega}_0) = \begin{cases} \|\nabla_f(\boldsymbol{\omega}_0)\| & \text{si } \gamma = 0 \\ -\|\nabla_f(\boldsymbol{\omega}_0)\| & \text{si } \gamma = 1 \end{cases}$$

Definición 4.1.4. Sea $f(\boldsymbol{\omega})$ una función real valorada sobre un subconjunto abierto Ω en \mathbb{R}^n , entonces el gráfico de influencia, es el conjunto de puntos en \mathbb{R}^{n+1} dado por

$$\mathcal{S} = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y}_i = \omega_i \text{ para } 1 \leq i \leq n, \mathbf{y}_{n+1} = f(\boldsymbol{\omega}), \boldsymbol{\omega} \in \Omega\}.$$

Analíticamente, el conjunto \mathcal{S} es la imagen de una función $\varphi : \Omega \rightarrow \mathbb{R}^{n+1}$. Explícitamente, esto está dado por la función

$$\varphi(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ f(\boldsymbol{\omega}) \end{pmatrix} \quad \text{con } \boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T \in \Omega.$$

Observación: La función φ representa el gráfico analíticamente, donde el conjunto $\mathcal{S} = \varphi(\Omega)$ en \mathbb{R}^{n+1} representa el gráfico geoméricamente.

Definición 4.1.5. La curvatura normal del gráfico \mathcal{S} sobre el punto $\varphi(\boldsymbol{\omega})$ en la dirección \mathbf{h} se define como la razón

$$C_{\mathbf{h}} = \frac{\Pi(\mathbf{h}, \mathbf{h})}{\mathbb{I}(\mathbf{h}, \mathbf{h})} \quad (4.1.5)$$

donde $\Pi(\mathbf{v}, \mathbf{v}) = -\langle d_{\mathbf{h}}\mathbf{u}(a), \mathbf{h} \rangle$ es la segunda forma fundamental, $\mathbf{u}(a)$ un vector normal tangente respecto a \mathbf{h} tal que $\|\mathbf{u}\| = 1$ y $\mathbb{I}(\mathbf{h}, \mathbf{h}) = \|d_{\mathbf{h}}\varphi(\boldsymbol{\omega})\|^2$ la primera forma fundamental.

Proposición 4.1.2. La curvatura normal $C_{\mathbf{h}}$ del gráfico \mathcal{S} sobre un punto $\boldsymbol{\omega}_0$ en dirección \mathbf{h} , está dado por

$$C_{\mathbf{h}} = \frac{\mathbf{h}^T \ddot{\mathbf{F}} \mathbf{h}}{(1 + \nabla_f(\boldsymbol{\omega}_0)^T \nabla_f(\boldsymbol{\omega}_0))^{1/2} \mathbf{h}^T (\mathbf{I} + \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T) \mathbf{h}} \quad (4.1.6)$$

con $\nabla_f(\boldsymbol{\omega}_0) = \partial f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} |_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ y $\ddot{\mathbf{F}} = \partial^2 f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T |_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$.

Demostración:

Sea $\varphi(\boldsymbol{\omega}(a)) = (\boldsymbol{\omega}(a)^T, f(\boldsymbol{\omega}(a)))^T$ el gráfico de influencia de modo que $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{h}$ con $a \in \mathbb{R}$, $\boldsymbol{\omega}_0 \in \Omega \subset \mathbb{R}^n$ y $\mathbf{h} \in \mathbb{R}^n$ y sea $\varphi(\boldsymbol{\omega}_0) = (\boldsymbol{\omega}_0, f(\boldsymbol{\omega}_0))^T$ un punto en \mathcal{S} .

Por definición basta encontrar la razón entre $\Pi(\mathbf{h}, \mathbf{h})$ y $\mathbb{I}(\mathbf{h}, \mathbf{h})$. Luego, la primera forma fundamental es

$$\begin{aligned} \mathbb{I}(\mathbf{h}, \mathbf{h}) &= \left\| d_{\mathbf{h}} \varphi(\boldsymbol{\omega}(a)) \Big|_{a=0} \right\|^2 = \left\| \frac{d}{da} \varphi(\boldsymbol{\omega}(a)) \Big|_{a=0} \right\|^2 \\ &= \left\| (\mathbf{h}^T, \langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle)^T \right\|^2 \\ &= \mathbf{h}^T (\mathbf{I} + \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T) \mathbf{h} \end{aligned}$$

con $\nabla_f(\boldsymbol{\omega}_0) = \partial f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} |_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$.

Sea $\mathbf{n}(a) \Big|_{a=0} = (-\nabla_f(\boldsymbol{\omega}_0), 1)^T \in \mathbb{R}^{n+1}$ el vector normal del gráfico \mathcal{S} sobre $\varphi(\boldsymbol{\omega})$ evaluado en $\boldsymbol{\omega}_0$, resultado obtenido al calcular el producto vectorial del gradiente de $\varphi(\boldsymbol{\omega})$ respecto $\boldsymbol{\omega}$ y sea $\mathbf{u}(a) \Big|_{a=0} = \mathbf{n} / \|\mathbf{n}\|$ vector unitario. Luego, la segunda forma fundamental está dada por

$$\Pi(\mathbf{h}, \mathbf{h}) = - \left\langle d_{\mathbf{h}} \mathbf{u}(a) \Big|_{a=0}, d_{\mathbf{h}} \varphi(\boldsymbol{\omega}(a)) \Big|_{a=0} \right\rangle = - \frac{1}{\|\mathbf{n}\|} \left\langle d_{\mathbf{h}} \mathbf{n}(a) \Big|_{a=0}, d_{\mathbf{h}} \varphi(\boldsymbol{\omega}(a)) \Big|_{a=0} \right\rangle$$

donde

$$\begin{aligned} d_{\mathbf{h}} \mathbf{n}(a) \Big|_{a=0} &= \begin{pmatrix} -\ddot{\mathbf{F}}^T \\ \mathbf{0} \end{pmatrix} \mathbf{h}, & d_{\mathbf{h}} \varphi(\boldsymbol{\omega}(a)) \Big|_{a=0} &= \begin{pmatrix} \mathbf{h} \\ \langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle \end{pmatrix} \quad y \\ \|\mathbf{n}\| &= (1 + \nabla_f(\boldsymbol{\omega}_0)^T \nabla_f(\boldsymbol{\omega}_0))^{1/2} \end{aligned}$$

con $\ddot{\mathbf{F}} = \partial^2 f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T |_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ y $\nabla_f(\boldsymbol{\omega}_0) = \partial f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} |_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$.

Finalmente, la curvatura normal, en base a la razón de la segunda y primera forma fundamental dada en (4.1.5) se demuestra (4.1.6).

4.2. Análisis de influencia local

Cook (1986) presenta el método de influencia local para evaluar la sensibilidad de los estimadores MV a pequeñas perturbaciones sobre los datos y/o los supuestos del modelo, este método permite evaluar el efecto simultáneo de las observaciones sobre los estimadores Máximos Verosímiles (MV), sin eliminarlos del conjunto de datos.

El modelo estadístico paramétrico $\mathfrak{F} = \{p(\mathbf{x}|\boldsymbol{\theta}); \boldsymbol{\theta} \in \Theta\}$ se define como la familia de densidades y el modelo estadístico paramétrico perturbado $\mathfrak{F}_\boldsymbol{\omega} = \{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega}) : \boldsymbol{\theta} \in \Theta, \boldsymbol{\omega} \in \Omega\}$ se define como la familia de densidades perturbadas donde $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T \in \Omega \subset \mathbb{R}^n$ es el vector de perturbación.

Sea $L(\boldsymbol{\theta})$ la función de verosimilitud correspondiente al modelo estadístico paramétrico y $\ell(\boldsymbol{\theta})$ su logaritmo, y sea $L(\boldsymbol{\theta}|\boldsymbol{\omega})$ la función de verosimilitud perturbada correspondiente al modelo estadístico paramétrico perturbado y $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ su logaritmo.

Sea $\boldsymbol{\omega}_0 \in \Omega \subset \mathbb{R}^n$ un vector de no perturbación de modo que $\mathfrak{F}_{\boldsymbol{\omega}_0} = \mathfrak{F}$ entonces $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$.

Para detectar si las observaciones influyen sobre los estimadores MV en el modelo estadístico propuesto, el método de influencia local permite comparar $\hat{\boldsymbol{\theta}}$ con $\hat{\boldsymbol{\theta}}_\boldsymbol{\omega}$, los cuales maximizan a $\ell(\boldsymbol{\theta})$ y $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ respectivamente, cuando $\boldsymbol{\omega}$ varía en Ω .

Si las estimaciones son similares implica que las perturbaciones han tenido poco efecto en la estimación de los parámetros y por el contrario si las diferencias son

grandes implica que las perturbaciones han tenido una influencia potencialmente importante en la estimación de los parámetros (ver Figura 4.1).

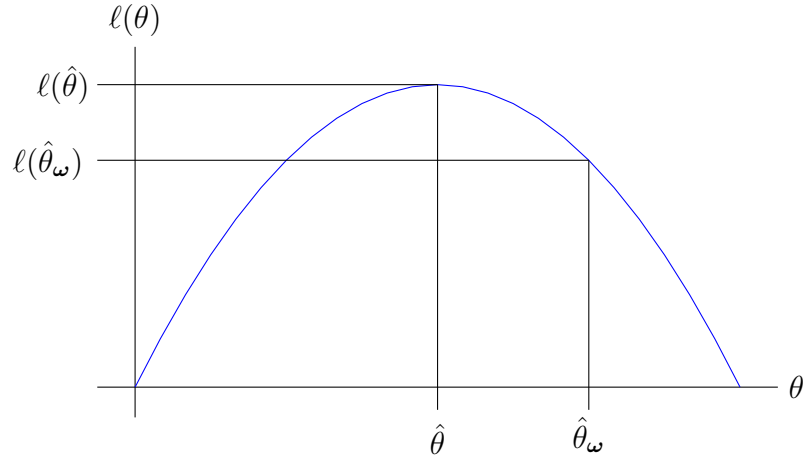


Figura 4.1: Comparación de estimadores

Para comparar $\hat{\theta}$ y $\hat{\theta}_\omega$ Cook (1986) propone el Desplazamiento de Verosimilitud (DV) como la función objetivo definida por

$$DV(\omega) = 2[\ell(\hat{\theta}) - \ell(\hat{\theta}_\omega)].$$

El gráfico $\varphi(\omega) = (\omega, DV(\omega))^T$ es llamado gráfico de influencia del desplazamiento de verosimilitud y estudiar su comportamiento en torno a ω_0 permite evaluar la influencia del esquema de perturbación sobre los estimadores MV.

En geometría diferencial, este tipo de gráfico es llamada Monge Patch y el estudio de influencia local se centra en analizar la curvatura de secciones normales del gráfico de influencia para medir que tan rápido la superficie se aleja de su plano tangente $T_{\varphi(\omega_0)}\mathbb{R}^n$ en una vecindad de $\omega_0 \in \Omega$.

El objetivo final es buscar la dirección \mathbf{h} en dónde se produce la mayor curvatura, ya que valores grandes de la curvatura normal indican sensibilidad a las perturbaciones consideradas en dicha dirección.

4.2.1. Perturbación

La perturbación de un modelo estadístico es cualquier modificación en los supuestos del modelo y/o los datos con el objetivo de constatar diferencias relevantes en los resultados del análisis estadístico. En general, en la literatura se presentan varios tipos de esquemas de perturbación, sin embargo en este trabajo se consideran dos.

1. **Ponderación de casos:** Su objetivo es evaluar si la contribución de las observaciones influyen en los estimadores MV. En este caso el logaritmo de la función de verosimilitud queda $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$ donde $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ pertenece a un subconjunto abierto $\Omega \subset \mathbb{R}^n$ y $\omega_i > 0$ para todo $i = 1, \dots, n$.
2. **Perturbación de datos:** El objetivo es perturbar la variable de estudio para evaluar si valores atípicos y/o extremos influyen sobre los estimadores MV. En este caso se considera el esquema de perturbación de la respuesta aditivo dado por $\mathbf{X}_{i\omega} = \mathbf{X}_i + \mathbf{1}_d \omega_i$ donde $\omega_i > 0$ para todo $i = 1, \dots, n$, $\mathbf{X}_i = (X_{i1}, \dots, X_{id})^T$ y $\mathbf{1}_d = (1, \dots, 1)^T$ de dimensión $(d \times 1)$.

En el caso del esquema de perturbación de casos el vector de no perturbación es $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ de dimensión $(n \times 1)$, mientras que en el caso del esquema de perturbación de la respuesta el vector de no perturbación es $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$ de dimensión $(n \times 1)$.

La dimensión del vector de perturbación puede cambiar, por ejemplo, Galea *et al.* (2002) consideran, para cada unidad experimental $i = 1, \dots, n$, un modelo de calibración comparativa con $j = 0, 1, \dots, p$ instrumentos, donde $j = 0$ indica el modelo de referencia. Asumen que el modelo es homocedástico, es decir, $V(\epsilon_{ij}) = \phi$. Luego proponen perturbar el supuesto de homogeneidad del modelo, de modo que $V(\epsilon_{ij}) = \phi/\omega_j$ para todo $j = 0, 1, \dots, p$ con $\omega_j > 0$, es decir, el vector de perturbación es $\boldsymbol{\omega} = (\omega_1, \dots, \omega_p)$ donde $p \neq n$.

4.2.2. Curvatura normal para el desplazamiento de verosimilitud

A continuación se deriva la curvatura normal para el desplazamiento de verosimilitud como función del vector de perturbación basado en Poon y Poon (1999).

Sea $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$ el vector de parámetros, $\hat{\boldsymbol{\theta}}_\omega$ el estimador MV bajo el modelo perturbado y $f(\boldsymbol{\omega}) = DV(\boldsymbol{\omega})$ la función objetivo. Desde la Proposición 4.1.2 basta encontrar $\partial f(\boldsymbol{\omega})/\partial \boldsymbol{\omega}$ y $\partial^2 f(\boldsymbol{\omega})/\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T$ y evaluarlas en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Se tiene que la primera derivada de la función objetivo respecto al vector de perturbación es

$$\frac{\partial f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} = \left(\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1}, \dots, \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_i}, \dots, \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_n} \right)^T,$$

y por la regla de la cadena, para todo $i = 1, \dots, n$, el i -ésimo elemento es

$$\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_i} = -2 \sum_{k=1}^K \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0)}{\partial \theta_k} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_\omega} \frac{\partial \hat{\theta}_{k,\omega}}{\partial \omega_i}$$

pero la derivada de la verosimilitud evaluada en su estimador MV, para todo $\boldsymbol{\omega}$, es cero, es decir

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_k} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_\omega} = 0 \tag{4.2.1}$$

y en particular para $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, se tiene

$$\frac{\partial f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \nabla_f(\boldsymbol{\omega}_0) = \mathbf{0}.$$

La segunda derivada de la función objetivo es

$$\frac{\partial^2 f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T} = \left[\frac{\partial^2 f(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} \right]_{(n \times n)}$$

donde el (i, j) -ésimo elemento es

$$\frac{\partial^2 f(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} = -2 \sum_{k=1}^K \sum_{l=1}^K \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}_0)}{\partial \theta_k \partial \theta_l} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}} \begin{pmatrix} \hat{\theta}_{l,\boldsymbol{\omega}} \\ \partial \omega_i \end{pmatrix} \begin{pmatrix} \hat{\theta}_{k,\boldsymbol{\omega}} \\ \partial \omega_j \end{pmatrix}. \quad (4.2.2)$$

Notar que, (4.2.2) es una forma cuadrática y al evaluar $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ se tiene

$$\frac{\partial^2 f(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -2 J_i^T \mathbf{L} J_j$$

para todo $i = 1, \dots, n$ y $j = 1, \dots, n$ con

$$\mathbf{L} = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}_0)}{\partial \theta_1 \partial \theta_1} & \cdots & \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}_0)}{\partial \theta_1 \partial \theta_K} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}_0)}{\partial \theta_K \partial \theta_1} & \cdots & \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}_0)}{\partial \theta_K \partial \theta_K} \end{pmatrix} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad \text{y} \quad J_i = \begin{pmatrix} \frac{\partial \hat{\theta}_{1,\boldsymbol{\omega}}}{\partial \omega_i} \\ \vdots \\ \frac{\partial \hat{\theta}_{K,\boldsymbol{\omega}}}{\partial \omega_i} \end{pmatrix} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$$

Por otro lado, por la regla de la cadena al derivar (4.2.1) respecto a ω_i , se tiene

$$\frac{\partial}{\partial \omega_i} \left(\frac{\partial \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}} \right) = \sum_{l=1}^K \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_k \partial \theta_l} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}} \frac{\partial \hat{\theta}_{l,\boldsymbol{\omega}}}{\partial \omega_i} \quad (4.2.3)$$

y al evaluar en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, la expresión (4.2.3) queda

$$\Delta_{ki} = \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_k \partial \omega_i} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} = L_k J_i$$

con L_k es el vector de la k -ésima fila de \mathbf{L} .

Luego, la matriz $\boldsymbol{\Delta}$ es de dimensión $(K \times n)$ y se puede escribir como $\boldsymbol{\Delta} = \mathbf{L} \mathbf{J}$.

Entonces, $\mathbf{J} = \mathbf{L}^{-1} \boldsymbol{\Delta}$ donde $\mathbf{J} = [J_1 \dots J_n]$ es una matriz de dimensión $(K \times n)$.

Por lo tanto, la segunda derivada de la función objetivo respecto al vector de perturbación, al evaluarlo en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ queda

$$\frac{\partial^2 f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -2 \mathbf{J}^T \mathbf{L} \mathbf{J} = 2 \boldsymbol{\Delta}^T [-\mathbf{L}^{-1}] \boldsymbol{\Delta}.$$

Finalmente,

$$\ddot{\mathbf{F}} = \frac{\partial^2 f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \text{ y } \nabla_f(\boldsymbol{\omega}_0) = \frac{\partial f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}.$$

A continuación se presenta la definición de la curvatura normal para el desplazamiento de verosimilitud como función del vector de perturbación basado en la Proposición 4.1.2.

Definición 4.2.1. La curvatura normal del gráfico \mathcal{S} generado por $DV(\boldsymbol{\omega})$ sobre un punto crítico $\boldsymbol{\omega}_0$ en la dirección \mathbf{h} es $C_{\mathbf{h}} = \mathbf{h}^T \ddot{\mathbf{F}} \mathbf{h}$ con $\ddot{\mathbf{F}} = 2\boldsymbol{\Delta}^T [-\mathbf{L}^{-1}] \boldsymbol{\Delta}$ donde $\mathbf{L} = (\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ y $\boldsymbol{\Delta} = (\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0, \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ es la matriz de perturbación.

En el gráfico \mathcal{S} la curvatura de ciertas curvas pasan a través de $\varphi(\boldsymbol{\omega}_0)$, estas curvas son llamadas secciones normales y están formadas por la intersección de la superficie con planos que contienen el vector normal (ortogonal) al plano tangente $T_{\varphi(\boldsymbol{\omega}_0)}$ de $\varphi(\boldsymbol{\omega})$ en $\boldsymbol{\omega}_0$. Las curvaturas de las secciones normales son llamadas curvaturas normales y son las que se utilizan para estudiar el comportamiento del gráfico de influencia alrededor de $\boldsymbol{\omega}_0$.

En cuanto a las secciones normales éstas se generan a partir de una línea recta en Ω que pasa por $\boldsymbol{\omega}_0$ y está representada por $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{h}$ con $a \in \mathbb{R}$, $\mathbf{h} \in \mathbb{R}^n$ tal que $\|\mathbf{h}\| = 1$, la cual genera una línea alzada (lifted line) en el gráfico \mathcal{S} y pasa a través de $\varphi(\boldsymbol{\omega}_0)$. Cada dirección \mathbf{h} específica una línea alzada a la cual le corresponde una sección normal.

La curvatura $C_{\mathbf{h}}$ de la sección normal $\varphi(\boldsymbol{\omega}(a))$ en la dirección \mathbf{h} tiene la estructura de una forma cuadrática y por la Proposición 1.2.3 y 1.2.4, dadas en Zhu *et al.* (2007), se puede encontrar la dirección donde se produce la mayor curvatura encontrando el mayor valor propio ($\lambda_{max} = C_{\mathbf{h}_{max}}$) asociada a la matriz simétrica $\ddot{\mathbf{F}}$. Luego, el vector propio asociado al mayor valor propio es la dirección donde se produce la mayor curvatura y lo denotamos como \mathbf{h}_{max} .

Finalmente, para visualizar si un conjunto de casos es potencialmente influyente

sobre los estimadores MV, dos tipos de gráficos son utilizados. El primero es el gráfico de \mathbf{h}_{max} versus los casos, la idea es detectar aquellas observaciones con valores grandes que indican una influencia local notable. El segundo es el gráfico de \mathbf{h}_{max} versus \mathbf{h}_{seg} donde *seg* indica el segundo vector propio asociado al segundo mayor valor propio, el cual fue propuesto por Kim (1996), éste permite detectar aquellos casos que se alejan en dirección de \mathbf{h}_{max} y/o \mathbf{h}_{seg} , indicando casos potencialmente influyentes sobre los estimadores MV.

4.2.3. Conformal curvatura normal

La curvatura normal no es invariante bajo transformación de escala uniforme sobre cualquier $\boldsymbol{\omega}$, por ejemplo, en Zhu *et al.* (2007) consideran el gráfico $\varphi^*(\boldsymbol{\omega}) = (\boldsymbol{\omega}, kf(\boldsymbol{\omega}))^T$, luego por la Proposición 4.1.2 su curvatura es

$$C_{\mathbf{h}}^* = \frac{k\mathbf{h}^T \ddot{\mathbf{F}}\mathbf{h}}{(1 + k^2 \nabla_f(\boldsymbol{\omega}_0)^T \nabla_f(\boldsymbol{\omega}_0))^{1/2} \mathbf{h}^T (\mathbf{I} + k^2 \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T) \mathbf{h}},$$

la cual es distinta a la curvatura normal de $\varphi(\boldsymbol{\omega}) = (\boldsymbol{\omega}, f(\boldsymbol{\omega}))^T$.

En particular, si $f(\boldsymbol{\omega}) = DV(\boldsymbol{\omega})$ entonces $\nabla_f(\boldsymbol{\omega}_0) = 0$, por ende, se sigue teniendo que $C_{\mathbf{h}}^* \neq C_{\mathbf{h}}$. En este escenario Poon y Poon (1999) proponen la curvatura normal invariante bajo transformación de escala uniforme.

Definición 4.2.2. La conformal curvatura normal del gráfico \mathcal{S} sobre un punto crítico $\boldsymbol{\omega}_0$ en dirección \mathbf{h} es

$$\mathbf{B}_{\mathbf{h}} = \frac{\mathbf{h}^T \ddot{\mathbf{F}}\mathbf{h}}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} \quad (4.2.4)$$

donde $\ddot{\mathbf{F}} = -2\boldsymbol{\Delta}^T \mathbf{L}^{-1} \boldsymbol{\Delta}$ evaluada en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ y $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

El Teorema 14.3.1 dado en Poon y Poon (2010) indica que esta medida invariante está acotada entre 0 y 1, es decir $0 \leq |\mathbf{B}_{\mathbf{h}}| \leq 1$, de modo que es posible comparar las curvaturas bajo diferentes modelos estadísticos.

Por otro lado, se tiene que el i -ésimo caso es potencialmente influyente si $B_i > \bar{B} + 2de(B)$ con $\bar{B} = n^{-1} \sum_{i=1}^n B_i$ y $de(B)$ es la desviación estándar de B_1, \dots, B_n , donde B_i es el i -ésimo elemento de $|\mathbf{B}_h|$.

4.2.4. Ponderación de casos

El logaritmo de la función de verosimilitud perturbada bajo el esquema de ponderación de casos es $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$ con $\ell_i(\boldsymbol{\theta}) = \log \left\{ |\boldsymbol{\Sigma}|^{-1/2} g(\delta_i^2) \right\}$ y $g : \mathbb{R} \rightarrow [0, \infty)$. La fdp de \mathbf{X} puede escribirse como $p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g(\delta_i^2)$ donde $\delta_i^2 = (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$ con $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\phi})^T \in \Theta \subseteq \mathbb{R}^{p+q}$ para el caso normal y $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\phi}, \eta)^T \in \Theta \subseteq \mathbb{R}^{p+q+1}$ para el caso t .

Caso Normal

Se tiene que $g(\delta_i^2) = (2\pi)^{-d/2} \exp\{-\delta_i^2/2\}$ con matriz de perturbación

$$\boldsymbol{\Delta} = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\sum_{i=1}^n \ell_i(\boldsymbol{\theta}) \frac{\partial \omega_i}{\partial \boldsymbol{\omega}^T} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} = \sum_{i=1}^n \mathbf{U}_i(\boldsymbol{\theta}) e_{in}^T$$

donde el vector score $\mathbf{U}_i(\boldsymbol{\theta}) = (U_i(\boldsymbol{\mu})^T, U_i(\boldsymbol{\phi})^T)^T$, $\boldsymbol{\omega}_0 = \mathbf{1}_n$ es el vector de no perturbación, e_{in} es un vector con un uno en el i -ésimo elemento de dimensión $(n \times 1)$ y $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\phi}})^T$ los estimadores MV de $\boldsymbol{\theta}$.

Luego, la matriz de influencia es

$$\boldsymbol{\Delta} = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1i} & \cdots & \Delta_{1n} \\ \Delta_{21} & \cdots & \Delta_{2i} & \cdots & \Delta_{2n} \end{pmatrix}$$

con elementos

$$\begin{aligned} \Delta_{1i} &= U_i(\boldsymbol{\mu})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}) \\ \Delta_{2i} &= U_i(\boldsymbol{\phi})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \frac{1}{2} \mathbf{D}_d^T \text{vec} \left(\hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}) (\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T \hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1} \right). \end{aligned}$$

Caso t

Se tiene que $g(\delta_i^2) = K_d(\eta) (1 + c(\eta)\delta_i^2)^{-(1/2\eta+d/2)}$ con

$$K_d(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} \text{ donde } c(\eta) = \frac{\eta}{(1-2\eta)}$$

y la matriz de perturbación

$$\Delta = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\sum_{i=1}^n \ell_i(\boldsymbol{\theta}) \frac{\partial \omega_i}{\partial \boldsymbol{\omega}^T} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} = \sum_{i=1}^n \mathbf{U}_i(\boldsymbol{\theta}) e_{in}^T$$

donde el vector score $\mathbf{U}_i(\boldsymbol{\theta}) = (U_i(\boldsymbol{\mu})^T, U_i(\boldsymbol{\phi})^T, U_i(\eta))^T$, $\boldsymbol{\omega}_0 = \mathbf{1}_n$ es el vector de no perturbación, e_{in} un vector con un uno en el i -ésimo elemento de dimensión $(n \times 1)$ y $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\phi}}, \hat{\eta})^T$ son los estimadores MV $\boldsymbol{\theta}$.

En este caso la matriz de perturbación está dada por

$$\Delta = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1i} & \cdots & \Delta_{1n} \\ \Delta_{21} & \cdots & \Delta_{2i} & \cdots & \Delta_{2n} \\ \Delta_{31} & \cdots & \Delta_{3i} & \cdots & \Delta_{3n} \end{pmatrix}$$

con elementos

$$\begin{aligned} \Delta_{1i} &= U_i(\boldsymbol{\mu})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \hat{u}_i \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}) \\ \Delta_{2i} &= U_i(\boldsymbol{\phi})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \frac{1}{2} \mathbf{D}_d^T \text{vec} \left(\hat{u}_i \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}) (\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T \hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1} \right) \\ \Delta_{3i} &= U_i(\eta)|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \frac{1}{2\hat{\eta}^2} \left\{ c(\hat{\eta})(d - \hat{u}_i \hat{\delta}_i^2) - \psi(1/2\hat{\eta} + d/2) + \psi(1/2\hat{\eta}) \right\} \\ &\quad + \frac{1}{2\hat{\eta}^2} \left\{ \log(1 + c(\hat{\eta})\hat{\delta}_i^2) \right\}. \end{aligned}$$

4.2.5. Perturbación de la respuesta

El esquema de perturbación de la respuesta es $\mathbf{X}_{i\omega} = \mathbf{X}_i + \mathbf{1}_d \omega_i$. El logaritmo de la función de verosimilitud perturbada es $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_{i\omega}(\boldsymbol{\theta})$ con $\ell_{i\omega}(\boldsymbol{\theta}) = \log \left\{ |\boldsymbol{\Sigma}|^{-1/2} g(\delta_{i\omega}^2) \right\}$ y $g: \mathbb{R} \rightarrow [0, \infty)$. La fdp de \mathbf{X} puede escribirse como $p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g(\delta_{i\omega}^2)$ donde $\delta_{i\omega}^2 = (\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{i\omega} - \boldsymbol{\mu})$ y $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\phi})^T \in \Theta \subseteq \mathbb{R}^{p+q}$

para el caso normal y $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\phi}, \eta)^T \in \Theta \subseteq \mathbb{R}^{p+q+1}$ para el caso t .

Caso Normal

Se tiene que $g(\delta_{i\omega}^2) = (2\pi)^{-d/2} \exp\{-\delta_{i\omega}^2/2\}$ y por conveniencia la matriz de influencia es calculada como

$$\Delta^T = \frac{\partial}{\partial \boldsymbol{\omega}} \left[\sum_{i=1}^n \frac{\partial \ell_{i\omega}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\partial}{\partial \boldsymbol{\omega}} \left[\sum_{i=1}^n \mathbf{U}_{i\omega}^T(\boldsymbol{\theta}) \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} .$$

Se define $\mathbf{U}_{i\omega}^T(\boldsymbol{\theta}) = (U_{i\omega}^T(\boldsymbol{\mu}), U_{i\omega}^T(\boldsymbol{\phi}))$ el vector score perturbado con elementos dados por

$$U_{i\omega}(\boldsymbol{\mu}) = -\frac{1}{2} \frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} = \boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})$$

$$U_{i\omega}(\boldsymbol{\phi}) = -\frac{1}{2} \frac{\partial \log(|\boldsymbol{\Sigma}|)}{\partial \boldsymbol{\phi}} - \frac{1}{2} \frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\phi}} = \frac{1}{2} \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1})$$

donde

$$\frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\phi}} = [-\mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1})] , \quad \frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\mu}} = 2\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(-1)$$

y $\frac{\partial \log(|\boldsymbol{\Sigma}|)}{\partial \boldsymbol{\phi}} = \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}) .$

Finalmente al derivar el vector score perturbado respecto a $\boldsymbol{\omega}$ y evaluarlo en $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ y $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ se obtiene

$$\Delta = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1i} & \cdots & \Delta_{1n} \\ \Delta_{21} & \cdots & \Delta_{2i} & \cdots & \Delta_{2n} \end{pmatrix}$$

con elementos

$$\Delta_{1i} = \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_d e_{in}^T$$

$$\Delta_{2i}^T = \mathbf{D}_d^T (\hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1}) \text{vec}((\mathbf{x}_i - \hat{\boldsymbol{\mu}}) \mathbf{1}_d^T) e_{in}^T$$

donde e_{in}^T es un vector con un uno en el i -ésimo elemento de dimensión $(n \times 1)$.

Caso t

Como $g(\delta_{i\omega}^2) = K_d(\eta) (1 + c(\eta)\delta_{i\omega}^2)^{-(1/\eta+d)/2}$ con

$$K_d(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} \text{ donde } c(\eta) = \frac{\eta}{(1-2\eta)},$$

sin pérdida de generalidad el cálculo de la matriz de perturbación es

$$\Delta^T = \frac{\partial}{\partial \boldsymbol{\omega}} \left[\sum_{i=1}^n \frac{\partial \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta}^T} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\partial}{\partial \boldsymbol{\omega}} \left[\sum_{i=1}^n \mathbf{U}_{i\omega}^T(\boldsymbol{\theta}) \right]$$

donde el vector score perturbado es $\mathbf{U}_{i\omega}^T(\boldsymbol{\theta}) = (U_{i\omega}^T(\boldsymbol{\mu}), U_{i\omega}^T(\boldsymbol{\phi}), U_{i\omega}(\eta))$ con elementos

$$\begin{aligned} U_{i\omega}(\boldsymbol{\mu}) &= -\frac{1}{2} \frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} = u_{i\omega} \boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu}) \\ U_{i\omega}(\boldsymbol{\phi}) &= -\frac{1}{2} \frac{\partial \log(|\boldsymbol{\Sigma}|)}{\partial \boldsymbol{\phi}} - \frac{1}{2} \frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\phi}} = \frac{1}{2} \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}) \\ U_{i\omega}(\eta) &= \frac{1}{2} \left(\frac{\partial \log K_d(\eta)}{\partial \eta} - \frac{\partial \delta_{i\omega}^2}{\partial \eta} \right) \\ &= \frac{1}{2\eta^2} (c(\eta)(d - u_{i\omega}\delta_{i\omega}^2) - \psi_0(1/2\eta + d/2) + \psi_0(1/2\eta) + \log(1 + c(\eta)\delta_{i\omega}^2)) \end{aligned}$$

donde $u_{i\omega} = (1/\eta + d)/(1/c(\eta) + \delta_{i\omega}^2)$ y

$$\begin{aligned} \frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\mu}} &= \frac{(1/\eta + d)}{(1/c(\eta) + \delta_{i\omega}^2)} 2\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(-1) \\ &= u_{i\omega} 2\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(-1) \\ \frac{\partial \delta_{i\omega}^2}{\partial \boldsymbol{\phi}} &= \frac{(1/\eta + d)}{(1/c(\eta) + \delta_{i\omega}^2)} [-\mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1})] \\ &= u_{i\omega} [-\mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_{i\omega} - \boldsymbol{\mu})(\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1})] \\ \frac{\partial \log(|\boldsymbol{\Sigma}|)}{\partial \boldsymbol{\phi}} &= \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}) \\ \frac{\partial \delta_{i\omega}^2}{\partial \eta} &= \frac{1}{\eta^2} [u_{i\omega} c(\eta) \delta_{i\omega}^2 - \log(1 + c(\eta) \delta_{i\omega}^2)] \\ \frac{\partial \log K_d(\eta)}{\partial \eta} &= \frac{1}{2\eta^2} [c(\eta)d - \psi(1/2\eta + d/2) + \psi(1/2\eta)]. \end{aligned}$$

Finalmente, al evaluar el vector score perturbado en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ y en $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, se obtiene la matriz de influencia

$$\boldsymbol{\Delta} = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1i} & \cdots & \Delta_{1n} \\ \Delta_{21} & \cdots & \Delta_{2i} & \cdots & \Delta_{2n} \\ \Delta_{31} & \cdots & \Delta_{3i} & \cdots & \Delta_{3n} \end{pmatrix}$$

con elementos

$$\begin{aligned} \Delta_{1i} &= \left[\hat{u}_i \left(\hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1} \left(2 \frac{(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} \right) \hat{\boldsymbol{\Sigma}}^{-1} \right) \mathbf{1}_d e_{in}^T \right] \\ \Delta_{2i}^T &= \left[\hat{u}_i \mathbf{D}_d^T (\hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1}) \text{vec} \left((\mathbf{x}_i - \hat{\boldsymbol{\mu}}) \mathbf{1}_d^T \right) e_{in}^T \right] \\ \Delta_{3i} &= 2(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_d \left[\frac{c(\hat{\eta})^2 \hat{u}_i \hat{\delta}_i^2}{(1 + c(\hat{\eta}) \hat{\delta}_i^2)} - c(\hat{\eta}) \hat{u}_i + \frac{c(\hat{\eta})}{(1 + c(\hat{\eta}) \hat{\delta}_i^2)} \right] \end{aligned}$$

y e_{in}^T es un vector con un uno en el i -ésimo elemento de dimensión $(n \times 1)$.

Observación: Si $\eta \rightarrow 0$ se tiene que $c(\eta) \rightarrow 0$ entonces $u_i \rightarrow 1$, luego el vector score bajo el supuesto t converge al vector score bajo el supuesto normal.

4.3. Influencia local generalizada

Casi siempre se está interesado en detectar casos potencialmente influyentes sobre el estimador MV. Sin embargo, a veces el interés está en detectar casos potencialmente influyente sobre una función del estimador MV como es el caso del CCC.

Cadigan y Farrel (2002) estudian la estabilidad de una función objetivo escalar, bajo algún esquema de perturbación, usando la derivada direccional para encontrar la dirección en dónde se produce la mayor pendiente local, siendo una generalización del método de influencia local propuesto por Cook (1986).

Sea $f(\boldsymbol{\omega}) = \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$ la función objetivo que depende del estimador MV perturbado.

Se llama enfoque de primer orden a la siguiente derivada direccional

$$S(\mathbf{h}) = \left. \frac{d\rho(\hat{\boldsymbol{\theta}}_\omega)}{da} \right|_{a=0}$$

y por el Teorema 4.1.1 se tiene

$$S(\mathbf{h}) = \mathbf{h}^T \left. \frac{\partial \rho(\hat{\boldsymbol{\theta}}_\omega)}{\partial \boldsymbol{\omega}} \right|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \quad (4.3.1)$$

con $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + a\mathbf{h}$, donde $\boldsymbol{\omega}_0 \in \Omega \subset \mathbb{R}^n$ es el vector de no perturbación, $\mathbf{h} \in \mathbb{R}^n$ vector de dirección y $a \in \mathbb{R}$ escalar que indica la magnitud de la perturbación.

Cadigan y Farrel (2002) muestran que

$$\left. \frac{\partial \rho(\hat{\boldsymbol{\theta}}_\omega)}{\partial \boldsymbol{\omega}} \right|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -\boldsymbol{\Delta}^T \mathbf{L}(\boldsymbol{\theta})^{-1} \left(\left. \frac{\partial \rho(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \quad (4.3.2)$$

Para encontrar la dirección donde se produce la mayor pendiente local, bajo el enfoque de primer orden, se tiene que por Cauchy- Schwartz la derivada direccional de $f(\boldsymbol{\omega})$ en cualquier dirección sobre $\boldsymbol{\omega}_0$ nunca excede el largo de $\nabla_f(\boldsymbol{\omega}_0)$, es decir, $\langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle \leq \|\nabla_f(\boldsymbol{\omega}_0)\| \|\mathbf{h}\|$ donde $\nabla_f(\boldsymbol{\omega}_0) = \left(\left. \frac{\partial \rho(\hat{\boldsymbol{\theta}}_\omega)}{\partial \boldsymbol{\omega}} \right) \right|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ y como $\|\mathbf{h}\| = 1$ entonces $\langle \nabla_f(\boldsymbol{\omega}_0), \mathbf{h} \rangle \leq \|\nabla_f(\boldsymbol{\omega}_0)\|$.

Por otro lado, por la Proposición 4.1.1, la desigualdad anterior se transforma en igualdad cuando \mathbf{h} es paralelo a $\nabla_f(\boldsymbol{\omega}_0)$, es decir, el $\cos(\gamma) = 1$ cuando $\gamma = 0$.

Luego, si $\nabla_f(\boldsymbol{\omega}_0) \neq 0$ la dirección donde se produce la mayor pendiente local es $\mathbf{h} = \nabla_f(\boldsymbol{\omega}_0) / \|\nabla_f(\boldsymbol{\omega}_0)\|$.

La influencia local generalizada a través del enfoque de primer orden, se basa en las propiedades de la pendiente, que no siempre son adecuadas. Wu y Luo (1993) señalan que en la dirección donde la superficie tiene mayor curvatura podrían haber casos potencialmente influyentes, sin embargo, en la dirección donde se produce la mayor pendiente no necesariamente ésta detecta una influencia importante.

Por lo anterior, Wu y Luo (1993) proponen estudiar la superficie de la función objetivo a través del enfoque de segundo orden definido por

$$\frac{dS(\mathbf{h})}{da} \Big|_{a=0} = \frac{d^2 \rho(\hat{\boldsymbol{\theta}}_\omega)}{da^2} \Big|_{a=0} = \mathbf{h}^T \left(\frac{\partial^2 \rho(\hat{\boldsymbol{\theta}}_\omega)}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \right) \mathbf{h} = \mathbf{h}^T \mathbf{H}_{f(\boldsymbol{\omega}_0)} \mathbf{h}$$

donde $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + a\mathbf{h}$.

Basado en la Proposición 4.1.2 la curvatura normal del gráfico \mathcal{S} sobre $\boldsymbol{\omega}_0$ en dirección \mathbf{h} toma la forma

$$C_{\mathbf{h}} = \frac{\mathbf{h}^T \mathbf{H}_{f(\boldsymbol{\omega}_0)} \mathbf{h}}{\mathbf{h}^T \nabla_{f(\boldsymbol{\omega}_0)} \mathbf{h}} \quad (4.3.3)$$

con $\nabla_{f(\boldsymbol{\omega}_0)} = (1 + \nabla_{f(\boldsymbol{\omega}_0)}^T \nabla_{f(\boldsymbol{\omega}_0)}) (\mathbf{I} + \nabla_{f(\boldsymbol{\omega}_0)} \nabla_{f(\boldsymbol{\omega}_0)}^T)$.

Luego, de todas las curvaturas (4.3.3), la máxima curvatura local y su correspondiente dirección se obtienen encontrando la solución de $|\mathbf{H}_{f(\boldsymbol{\omega}_0)} - \lambda \nabla_{f(\boldsymbol{\omega}_0)}| = 0$. En otras palabras, se debe encontrar el vector propio generalizado asociado al mayor valor propio generalizado de $\nabla_{f(\boldsymbol{\omega}_0)}^{-1} \mathbf{H}_{f(\boldsymbol{\omega}_0)}$.

Para realizar el análisis de influencia local generalizada, basado en el enfoque de primer y segundo orden, basta obtener la primera y segunda derivada de la función objetivo respecto a $\boldsymbol{\omega}$.

Bajo el supuesto normal, para realizar el diagnóstico sobre el CCC, es posible obtener el enfoque de primer y segundo orden, sin embargo, bajo el supuesto t solo es posible obtener el enfoque de primer orden, usando la expresión (4.3.2).

Por otro lado, para realizar el diagnóstico sobre el CCCG, obtener la primera y segunda derivada resulta complejo por lo que sólo es posible obtener la primera derivada a partir de (4.3.2).

4.3.1. Ponderación de casos

Caso Normal

Suponiendo que los datos siguen una distribución normal bivariada el vector de parámetros de interés es $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22})^T \in \Theta \subseteq \mathbb{R}^5$.

La primera y segunda derivada del CCC bivariado $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)$ respecto a $\boldsymbol{\omega}$ evaluada en $\boldsymbol{\omega}_0$ son

$$\nabla_f(\boldsymbol{\omega}_0) = \left\{ \frac{\hat{\rho}_c}{n\hat{\sigma}_{12}} \left(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}\hat{\sigma}_{12} \right) - \frac{\hat{\rho}_c^2}{2n\hat{\sigma}_{12}} \mathbf{Z}^* \right\} \quad (4.3.4)$$

$$\mathbf{H}_{f(\boldsymbol{\omega}_0)} = \{ \boldsymbol{\Gamma}_3 - (\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2) \} \quad (4.3.5)$$

donde \odot indica el punto Hadamard y \mathbf{Z}^* , $\mathbf{Z}_{(1)}$, $\mathbf{Z}_{(2)}$, $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_2$ y $\boldsymbol{\Gamma}_3$ están dadas en el Apéndice A.6.1.

En este escenario es posible realizar el análisis de influencia local generalizada de primer y segundo orden al CCC bivariado. No obstante, para el CCCG es posible realizar la influencia local generalizada de primer orden con $\nabla_f(\boldsymbol{\omega}_0)$ dada en (4.3.2).

Caso t

Para el caso que los datos siguen una distribución t el vector de parámetros de interés es $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22}, \eta)^T \in \Theta \subseteq \mathbb{R}^6$.

Sin embargo, debido a que los estimadores MV se obtiene a través del algoritmo EM, no es posible calcular la primera derivada de $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)$ respecto a $\boldsymbol{\omega}$ directamente. Bajo este escenario se usa $\nabla_f(\boldsymbol{\omega}_0)$ definida en (4.3.2) para el análisis de influencia local generalizada de primer orden del CCC y CCCG.

4.3.2. Perturbación de la respuesta

Caso Normal

Suponiendo que los datos siguen una distribución normal bivariada, al igual que el esquema de perturbación anterior, se puede realizar el análisis de influencia local generalizada de primer y segundo orden al CCC bivariado.

La primera y segunda derivada de $\rho(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$ respecto a $\boldsymbol{\omega}$ evaluada en $\boldsymbol{\omega}_0$ y $\hat{\boldsymbol{\theta}}$ son

$$\nabla_f(\boldsymbol{\omega}_0) = \left(\frac{\hat{\rho}}{\hat{\sigma}_{12}} - \frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \quad (4.3.6)$$

$$\mathbf{H}_{f(\boldsymbol{\omega}_0)} = \left\{ 2 \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right) \mathbf{P}_n + 2 \left(\left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right)^2 - \frac{\hat{\rho}_c}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \mathbf{Z}^T \mathbf{P}_n^T \right\} \quad (4.3.7)$$

con \mathbf{Z} y \mathbf{P}_n dadas en el Apéndice A.6.1.

En el caso del CCCG sólo se realiza la influencia local generalizada de primer orden usando $\nabla_f(\boldsymbol{\omega}_0)$ dada en (4.3.2).

Caso t

Para el caso que los datos siguen una distribución t reparametrizada, al igual que el esquema de perturbación anterior, sólo es posible realizar la influencia local generalizada de primer orden para el CCC y el CCCG usando $\nabla_f(\boldsymbol{\omega}_0)$ dada en (4.3.2).

4.4. Medidas de influencia

Se tiene que la curvatura normal de la superficie $\varphi(\boldsymbol{\omega}) = (\boldsymbol{\omega}^T, f(\boldsymbol{\omega}))^T$ es $C_{\mathbf{h}}$ y la curvatura normal de otra superficie dada por $\varphi^*(\boldsymbol{\omega}) = (\boldsymbol{\omega}^T, kf(\boldsymbol{\omega}))^T$ es

$$C_{\mathbf{h}}^* = \frac{k \mathbf{h}^T \ddot{\mathbf{F}} \mathbf{h}}{\left(1 + k^2 \nabla_f^T(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0) \right)^{1/2} \mathbf{h}^T \left(\mathbf{I} + k^2 \nabla_f(\boldsymbol{\omega}_0) \nabla_f^T(\boldsymbol{\omega}_0) \right) \mathbf{h}}.$$

Luego, notar que si $\nabla_f(\boldsymbol{\omega}_0)$ es cero o distinto de cero siempre $C_{\mathbf{h}}^* \neq C_{\mathbf{h}}$.

Cuando $\nabla_f(\boldsymbol{\omega}_0) = 0$, Poon y Poon (1999), proponen la curvatura normal invariante bajo transformación de escala uniforme, llamada conformal curvatura normal, definida en (4.2.4). Sin embargo, si $\nabla_f(\boldsymbol{\omega}_0) \neq 0$ la conformal curvatura normal, para $\varphi(\boldsymbol{\omega})$, es

$$\mathbf{B}_{\mathbf{h}} = \frac{\mathbf{h}^T \ddot{\mathbf{F}} \mathbf{h}}{\mathbf{h}^T (\mathbf{I} + \nabla_f(\boldsymbol{\omega}_0) \nabla_f^T(\boldsymbol{\omega}_0)) \mathbf{h} \sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}}$$

y la conformal curvatura normal de $\varphi^*(\boldsymbol{\omega})$ es

$$\mathbf{B}_{\mathbf{h}}^* = \frac{\mathbf{h}^T \ddot{\mathbf{F}} \mathbf{h}}{\mathbf{h}^T (\mathbf{I} + k^2 \nabla_f(\boldsymbol{\omega}_0) \nabla_f^T(\boldsymbol{\omega}_0)) \mathbf{h} \sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}}$$

por lo que $\mathbf{B}_{\mathbf{h}}^* \neq \mathbf{B}_{\mathbf{h}}$ no es invariante bajo transformación uniforme de escala.

Finalmente, cuando $\nabla_f(\boldsymbol{\omega}_0)$ es distinta de cero, la curvatura normal y la conformal curvatura normal no son invariantes bajo transformación de escala uniforme.

Bajo el escenario anterior, Zhu *et al.* (2007), desarrollan medidas de influencia para cualquier función objetivo derivable, que siempre serán invariantes bajo transformación de escala uniforme y además exponen cómo verificar si el esquema de perturbación seleccionado es adecuado.

La metodología aplicada se basa en conceptos de geometría diferencial para una variedad n -dimensional (n -dimensional manifold en inglés). Basados en Amari (1985) describen el modelo estadístico perturbado $\mathfrak{F}_{\boldsymbol{\omega}}$ como una variedad n -dimensional donde el vector de perturbación $\boldsymbol{\omega}$ juega el rol de un sistema de coordenadas. Estas medidas de influencia son conocidas como medidas de primer orden y segundo orden.

En esta sección, por simplicidad, el análisis de influencia local basado en las

medidas de primer y segundo orden, sólo será aplicado al estimador del CCC entre medidas generadas por dos instrumentos o métodos, asumiendo que los datos siguen una distribución normal bivariada.

4.4.1. Selección del esquema de perturbación

Un mecanismo para verificar y por ende seleccionar un esquema de perturbación adecuado es a través de la estructura de una matriz métrica, específicamente la matriz de información de Fisher respecto al vector de perturbación $\boldsymbol{\omega}$.

Definición 4.4.1. La matriz de información de Fisher con respecto al vector de perturbación es $G(\boldsymbol{\omega})$, llamada matriz métrica y su (i, j) -ésimo elemento es

$$g_{ij}(\boldsymbol{\omega}) = E_{\boldsymbol{\omega}} \left[\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_i} \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} \right] \quad \forall \quad i, j = 1, \dots, n \quad (4.4.1)$$

donde $E_{\boldsymbol{\omega}}$ denota la esperanza con respecto a $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega})$.

Los elementos $g_{ii}(\boldsymbol{\omega})$ indican la cantidad de perturbación introducida por el i -ésimo componente del vector de perturbación $\boldsymbol{\omega}$, mientras que los elementos $g_{ij}(\boldsymbol{\omega})$ indican la asociación entre los diferentes componentes del vector de perturbación.

Zhu *et al.* (2007) indican que si el esquema de perturbación es adecuado para el modelo estadístico debe satisfacer:

- (i) $G(\boldsymbol{\omega})$ es definida positiva en torno de $\boldsymbol{\omega}_0$.
- (ii) Los elementos de fuera de la diagonal de $G(\boldsymbol{\omega}_0)$ deben ser lo más pequeño posible o la matriz debe ser $G(\boldsymbol{\omega}_0) = c\mathbf{I}_d$.

Si (ii) no se cumple, los autores proponen un nuevo vector de perturbación dado por $\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega}_0 + c^{-1/2}G(\boldsymbol{\omega})^{1/2}(\boldsymbol{\omega} - \boldsymbol{\omega}_0)$ de modo que sea apropiado, es decir $G(\tilde{\boldsymbol{\omega}})|_{\tilde{\boldsymbol{\omega}}=\boldsymbol{\omega}_0} = c\mathbf{I}_d$.

4.4.2. Medidas de influencia de primer y segundo orden

Para construir las medidas de influencia de primer y segundo orden Zhu *et al.* (2007) introducen la matriz métrica $G(\boldsymbol{\omega})$ y una matriz simétrica basada en el

símbolo de Christoffel para la conexión de Levi-Civita del tensor métrico llamada matriz Hessiana covariante.

Además Zhu *et al.* (2007) demuestran que las medidas de primer y segundo orden, presentadas a continuación, son invariante bajo transformación de escala.

Definición 4.4.2. La medida de influencia de primer orden en la dirección \mathbf{h} es

$$FI_{f,\mathbf{h}} = \frac{\mathbf{h}^T \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T \mathbf{h}}{\mathbf{h}^T G(\boldsymbol{\omega}_0) \mathbf{h}}, \quad (4.4.2)$$

donde $\nabla_f(\boldsymbol{\omega}_0) = \partial f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ y $G(\boldsymbol{\omega}_0)$ la matriz métrica evaluada en $\boldsymbol{\omega}_0$

Definición 4.4.3. La medida de influencia de segundo orden en la dirección \mathbf{h} es

$$SI_{f,\mathbf{h}} = \frac{\mathbf{h}^T \tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} \mathbf{h}}{\mathbf{h}^T G(\boldsymbol{\omega}_0) \mathbf{h}}, \quad (4.4.3)$$

con $G(\boldsymbol{\omega}_0)$ la matriz métrica evaluada en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ y $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)}$ la matriz hessiana covariante evaluada en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ donde el (i, j) -ésimo elemento es

$$\left[\tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} \right]_{ij} = \frac{\partial}{\partial \omega_i} \left(\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_j} \right) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} - \sum_{s,r} g^{r,s}(\boldsymbol{\omega}) \Gamma_{ijs}^0(\boldsymbol{\omega}) \left(\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_r} \right) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

donde $g^{r,s}(\boldsymbol{\omega})$ es el (r, s) -ésimo elemento de $G(\boldsymbol{\omega})^{-1}$ y

$$\Gamma_{ijs}^0(\boldsymbol{\omega}) = \frac{1}{2} \left[\frac{\partial}{\partial \omega_i} g(\boldsymbol{\omega})_{js} + \frac{\partial}{\partial \omega_j} g(\boldsymbol{\omega})_{is} - \frac{\partial}{\partial \omega_s} g(\boldsymbol{\omega})_{ij} \right]$$

es el símbolo de Christoffel para conexión Lévi-Civita del tensor métrico (ver Do Carmo (1976)).

Definición 4.4.4. La medida de influencia de segundo orden estandarizada en la dirección \mathbf{h} es definida como

$$SSI_{f,\mathbf{h}} = \frac{\mathbf{h}^T \tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} \mathbf{h}}{\sqrt{\text{tr} \left[G(\boldsymbol{\omega}_0)^{-1} \tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} \right]^2 \mathbf{h}^T G(\boldsymbol{\omega}_0) \mathbf{h}}}.$$

Luego, para la influencia de primer orden, basta encontrar la dirección donde se produce la mayor influencia sobre la función objetivo $f(\boldsymbol{\omega})$, la cual es el vector propio asociado al mayor valor propio de $\nabla_f(\boldsymbol{\omega}_0)\nabla_f(\boldsymbol{\omega}_0)^T$ y para la influencia de segundo orden es el vector propio asociado al mayor valor propio de $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)}$. La Tabla 4.1 resume las similitudes y diferencias entre las medidas de influencia propuesta por Zhu *et al.* (2007), la curvatura normal, la conformal curvatura normal y la máxima pendiente, estudiadas en las secciones anteriores.

Zhu *et al.* (2007) muestran que la medida de primer y segundo orden son invariantes bajo transformación de escala al considerar un sistema de coordenadas $\boldsymbol{\tau}$ que es un difeomorfismo de $\boldsymbol{\omega}$ para el modelo perturbado (variedad n -dimensional) basado en un tensor métrico. Por lo tanto, por el Teorema 1 parte (ii) y Teorema 2 parte (ii) dados en Zhu *et al.* (2007), se tiene que $FI_{k,f,\mathbf{h}} = k^2 FI_{f,\mathbf{h}}$, $SI_{k,f,\mathbf{h}} = k SI_{f,\mathbf{h}}$ y $SSI_{k,f,\mathbf{h}} = SSI_{f,\mathbf{h}}$ para $k \in \mathbb{R}$.

Tabla 4.1: Comparación entre las medidas de influencia según la derivada de la función objetivo sobre un punto crítico

	Medida de influencia	Ventajas-Desventajas
$\nabla_f(\boldsymbol{\omega}_0) = 0$	$SI_{f,\mathbf{h}} = C_{\mathbf{h}}$ $SSI_{f,\mathbf{h}} = \mathbf{B}_{\mathbf{h}}$	Coinciden si el esquema de perturbación es apropiado y son invariantes.
$\nabla_f(\boldsymbol{\omega}_0) \neq 0$	$SI_{f,\mathbf{h}} \neq C_{\mathbf{h}}$ $SSI_{f,\mathbf{h}} \neq \mathbf{B}_{\mathbf{h}}$ $FI_{f,\mathbf{h}} = \mathbf{S}(\mathbf{h})$	Son invariantes, pero no coinciden con $C_{\mathbf{h}}$ y $\mathbf{B}_{\mathbf{h}}$ las que no son invariantes. Coincide con la máxima pendiente local

A continuación se presenta la construcción de las medidas de influencia de primer y segundo orden para el coeficiente de correlación de concordancia como función objetivo $f(\boldsymbol{\omega}) = \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$, asumiendo que $\mathbf{X}_1, \dots, \mathbf{X}_n$ son iid de modo que \mathbf{X}_i sigue una distribución normal para todo $i = 1, \dots, n$, la fdp conjunta está dada por

$$p(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^n \left\{ (2\pi)^{-d/2} |\boldsymbol{\Sigma}|^{-1/2} e^{-\delta_i^2/2} \right\}$$

con $\mathbf{x} \in \mathbb{R}^{d \times n}$, $\delta_i^2 = (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$, $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T)^T$ y $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ es el vector columna que contiene los diferentes elementos de $\boldsymbol{\Sigma}$.

4.4.3. Ponderación de casos

Sea $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T \in \Omega \subset \mathbb{R}^n$ el vector de perturbación y $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$ el esquema de ponderación de casos. Sea $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ el vector de no perturbación de modo que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$ y $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega}_0) = p(\mathbf{x}|\boldsymbol{\theta})$.

Aplicando el Teorema 3 dado en Zhu *et al.* (2007), para el esquema de ponderación de casos, la densidad del modelo perturbado está dada por

$$p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n \{ \exp \{ \omega_i \ell_i(\mathbf{x}_i, \boldsymbol{\theta}) \} [c_i(\omega_i, \boldsymbol{\theta})]^{-1} \} \quad (4.4.4)$$

donde

$$c_i(\omega_i, \boldsymbol{\theta}) = \int_{\mathbb{R}^p} \exp \{ \omega_i \ell_i(\mathbf{x}_i, \boldsymbol{\theta}) \} d\mathbf{x}_i = e^{\omega_i b} (2\pi)^{d/2} |\omega_i^{-1} \boldsymbol{\Sigma}|^{1/2}$$

con $b = -(1/2) [d \log(2\pi) + \log |\boldsymbol{\Sigma}|]$.

Luego, al reemplazar $c_i(\omega_i, \boldsymbol{\theta})$ en (4.4.4) el esquema de ponderación de caso coincide con el esquema de perturbación de la matriz de varianza-covarianza, es decir

$$p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n \left\{ (2\pi)^{-d/2} |\omega_i^{-1} \boldsymbol{\Sigma}|^{-1/2} \exp \left(-\frac{\omega_i \delta_i^2}{2} \right) \right\}.$$

Bajo ciertas condiciones de regularidad para $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\omega})$ dadas en Amari (1985) y por el Teorema 3 dado en Zhu *et al.* (2007), los elementos de $G(\boldsymbol{\omega})$ están dados por $g_{ij}(\boldsymbol{\omega}) = V_{\boldsymbol{\omega}}[\ell(\mathbf{x}_i, \boldsymbol{\theta})] \delta_{ij}$ donde $V_{\boldsymbol{\omega}}[\ell(\mathbf{x}_i, \boldsymbol{\theta})] = \partial^2 \log(c_i(\omega_i, \boldsymbol{\theta})) / \partial \omega_i^2$.

Luego, se tiene que $g_{ij}(\boldsymbol{\omega}) = d\delta_{ij}/2\omega_i^2$ para todo $i = j$ y $g_{ij}(\boldsymbol{\omega}) = 0$ para todo $i \neq j$, con $\delta_{ij} = 1$ si $i = j$ o $\delta_{ij} = 0$ si $i \neq j$, donde δ_{ij} es conocido como delta de Kronecker.

Por lo tanto, la matriz métrica está dada por

$$G(\boldsymbol{\omega}) = \frac{d}{2} \begin{pmatrix} \frac{1}{\omega_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\omega_2^2} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\omega_n^2} \end{pmatrix}$$

y al ser evaluada en $\boldsymbol{\omega}_0$ queda $G(\boldsymbol{\omega}_0) = (d/2)\mathbf{I}_{n \times n}$ lo que indica que la perturbación de ponderación de casos es adecuada.

Por otro lado, desde la Definición 4.4.3, la matriz Hessiana covariante es $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega})} = \mathbf{H}_{f(\boldsymbol{\omega})} - \mathbf{T}_{f(\boldsymbol{\omega})}$ donde los elementos (i, j) -ésimo de $\mathbf{H}_{f(\boldsymbol{\omega})}$ es la segunda derivada de la función objetivo respecto a $\boldsymbol{\omega}$ y los elementos de $\mathbf{T}_{f(\boldsymbol{\omega})}$ son

$$\tau_{ij}(\boldsymbol{\omega}) = \sum_{s,r} g^{r,s}(\boldsymbol{\omega}) \Gamma_{ijs}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_r} \quad \forall i, j = 1, \dots, n \quad (4.4.5)$$

con $g^{s,r}(\boldsymbol{\omega}) = 2\omega_s^2/d$ para todo $s = r$ y $g^{s,r}(\boldsymbol{\omega}) = 0$ en cualquier otro caso, y $\Gamma_{ijs}^0(\boldsymbol{\omega}) = -(d/2)(1/\omega_i^3)\delta_{ij}\delta_{is}$ para $i = j = s$ donde $\delta_{ij} = 1$ cuando $i = j$ y $\delta_{ij} = 0$ en cualquier otro caso.

Para calcular (4.4.5) se considera $n = 2$ tal que

$$\begin{aligned} \tau_{ij}(\boldsymbol{\omega}) &= \sum_{s=1}^2 \left[\left(g^{1,s}(\boldsymbol{\omega}) \Gamma_{ijs}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} \right) + \left(g^{2,s}(\boldsymbol{\omega}) \Gamma_{ijs}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_2} \right) \right] \\ &= \left(g^{1,1}(\boldsymbol{\omega}) \Gamma_{ij1}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} \right) + \left(g^{1,2}(\boldsymbol{\omega}) \Gamma_{ij2}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} \right) \\ &\quad + \left(g^{2,1}(\boldsymbol{\omega}) \Gamma_{ij1}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_2} \right) + \left(g^{2,2}(\boldsymbol{\omega}) \Gamma_{ij2}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_2} \right) \end{aligned}$$

pero $g^{r,s}(\boldsymbol{\omega}) = 0 \quad \forall r \neq s$, entonces

$$\tau_{ij}(\boldsymbol{\omega}) = \sum_{s=1}^n g^{s,s}(\boldsymbol{\omega}) \Gamma_{ijs}^0(\boldsymbol{\omega}) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_s}$$

con $g^{s,s}(\boldsymbol{\omega}) = 2\omega_s^2/d$ y $\Gamma_{ijs}^0(\boldsymbol{\omega}) = -(d/2)(1/\omega_i^3)\delta_{ij}\delta_{is}$ para $i = j = s$ y 0 en otro

caso.

Luego, para $n = 2$ se tiene que los elementos quedan

$$\begin{aligned}\tau_{11}(\boldsymbol{\omega}) &= \frac{2\omega_1^2}{d} \left(-\frac{d}{2\omega_1^3} \delta_{11} \delta_{11} \right) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} + \frac{2\omega_2^2}{d} \left(-\frac{d}{2\omega_1^3} \delta_{11} \delta_{12} \right) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} \\ &= \frac{2\omega_1^2}{d} \left(-\frac{d}{2\omega_1^3} \right) \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} = -\frac{1}{\omega_1} \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} \\ \tau_{22}(\boldsymbol{\omega}) &= -\frac{1}{\omega_2} \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_2} \\ \tau_{12}(\boldsymbol{\omega}) &= \tau_{21}(\boldsymbol{\omega}) = 0\end{aligned}$$

de modo que la matriz queda

$$\mathbf{T}(\boldsymbol{\omega}) = \begin{pmatrix} -\frac{1}{\omega_1} \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_1} & 0 & \dots & 0 \\ 0 & -\frac{1}{\omega_2} \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_2} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & -\frac{1}{\omega_n} \frac{\partial f(\boldsymbol{\omega})}{\partial \omega_n} \end{pmatrix}$$

y evaluada en $\boldsymbol{\omega}_0$ queda $\mathbf{T}(\boldsymbol{\omega}_0) = -\text{Diag}(\nabla_f(\boldsymbol{\omega}_0))$.

Luego, para $d = 2$, la matriz Hessiana covariante evaluada en $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ es $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} = \mathbf{H}_{f(\boldsymbol{\omega}_0)} + \text{Diag}(\nabla_f(\boldsymbol{\omega}_0))$ y la matriz métrica queda $G(\boldsymbol{\omega}_0) = \mathbf{I}_n$.

Por lo tanto, como $\|\mathbf{h}\|^2 = 1$, las medidas de influencia de primer y segundo orden están dadas respectivamente por

$$\begin{aligned}FI_{\mathbf{h},f} &= \mathbf{h}^T \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T \mathbf{h} \\ SI_{\mathbf{h},f} &= \mathbf{h}^T \tilde{\mathbf{H}}_{f(\boldsymbol{\omega}_0)} \mathbf{h}.\end{aligned}$$

donde $\nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T$ y $\mathbf{H}_{f(\boldsymbol{\omega}_0)}$ están calculadas en el Apéndice A.6.1.

4.4.4. Perturbación de la respuesta

Sea $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T \in \Omega \subset \mathbb{R}^n$ el vector de perturbación y $\mathbf{X}_i(\boldsymbol{\omega}) = \mathbf{X}_i + \omega_i \mathbf{1}_d$ el esquema de perturbación de la respuesta y $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$ el vector de no perturbación de modo que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$ y $p(\mathbf{x}|\boldsymbol{\omega}_0) = p(\mathbf{x})$.

La función de verosimilitud perturbada está definida por

$$L(\boldsymbol{\theta}|\boldsymbol{\omega}) = \prod_{i=1}^n \{(2\pi)^{-d/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\{-\delta_{i\boldsymbol{\omega}}^2/2\}\}$$

con $\delta_{i\boldsymbol{\omega}}^2 = (\mathbf{x}_i - \boldsymbol{\mu} + \omega_i \mathbf{1}_d)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu} + \omega_i \mathbf{1}_d)$ y el logaritmo de la función de verosimilitud perturbada es

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \left\{ -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu} + \omega_i \mathbf{1}_d)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu} + \omega_i \mathbf{1}_d) \right\}.$$

Para calcular los elementos $g_{ij}(\boldsymbol{\omega})$ usamos la siguiente igualdad

$$E_{\boldsymbol{\omega}} \left[\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_i} \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} \right] = -E_{\boldsymbol{\omega}} \left[\frac{\partial}{\partial \omega_i} \left(\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} \right) \right]. \quad (4.4.6)$$

Luego, se tiene que

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} = -\mathbf{1}_d^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_d - \omega_j \mathbf{1}_d^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_d \text{ y } \frac{\partial}{\partial \omega_i} \left(\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} \right) = -\mathbf{1}_d^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_d$$

y aplicando la esperanza del lado izquierdo de (4.4.6) se obtienen los elementos $g_{ij}(\boldsymbol{\omega}) = \mathbf{1}_d^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_d \delta_{ij}$, con $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$, de modo que la matriz $G(\boldsymbol{\omega})$ al ser evaluada en $\boldsymbol{\omega}_0$ y $\widehat{\boldsymbol{\Sigma}}$ queda

$$G(\boldsymbol{\omega}_0) = \begin{pmatrix} \mathbf{1}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} & 0 & \dots & 0 \\ 0 & \mathbf{1}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{1}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} \end{pmatrix}. \quad (4.4.7)$$

La matriz Hessiana covariante es $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega})} = \mathbf{H}_{f(\boldsymbol{\omega})} - \mathbf{T}_{f(\boldsymbol{\omega})}$ donde los (i, j) -ésimo elementos de $\mathbf{H}_{f(\boldsymbol{\omega})}$ es la segunda derivada de la función objetivo respecto a $\boldsymbol{\omega}$ y

los elementos de $\mathbf{T}_{f(\boldsymbol{\omega})}$ están dados como en (4.4.5).

Se tiene que $g^{r,s}(\boldsymbol{\omega})$ no depende de ω_i para $r = s$ y $g^{r,s}(\boldsymbol{\omega}) = 0$ para $r \neq s$ entonces $\Gamma_{ijs}^0(\boldsymbol{\omega}) = 0$ para todo i, j, s , lo que implica que $\tau_{ij}(\boldsymbol{\omega}) = 0$.

Por lo tanto, la matriz $\mathbf{T}_{f(\boldsymbol{\omega})} = \mathbf{0}$ entonces la matriz Hessiana covariante es $\tilde{\mathbf{H}}_{f(\boldsymbol{\omega})} = \mathbf{H}_f(\boldsymbol{\omega})$.

Por otro lado, notar que la matriz métrica $G(\boldsymbol{\omega}_0) = (\mathbf{1}_d^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_d) \mathbf{I}_n = c \mathbf{I}_n$ con $c \in \mathbb{R}$, indicando que el esquema de perturbación es adecuado.

Finalmente, las medidas de influencia de primer y segundo orden bajo el esquema de perturbación de la respuesta están dadas por

$$\begin{aligned} FI_{\mathbf{h},f} &= \mathbf{h}^T \nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T \mathbf{h} \\ SI_{\mathbf{h},f} &= \mathbf{h}^T [\mathbf{H}_{f(\boldsymbol{\omega}_0)}] \mathbf{h}. \end{aligned}$$

donde $\nabla_f(\boldsymbol{\omega}_0) \nabla_f(\boldsymbol{\omega}_0)^T$ y $\mathbf{H}_{f(\boldsymbol{\omega}_0)}$ están calculadas en el Apéndice A.6.2.

4.5. Comentarios finales

El método de diagnóstico presentado, se basa principalmente, en estudiar la superficie de una función que depende del estimador MV perturbado. El estudio de la superficie generada por la función objetivo es a través de la curvatura y la idea es encontrar la dirección donde se produce la mayor curvatura, pues en esa dirección es posible detectar casos que pueden influir sobre los estimadores MV o sobre una función de ellos.

Cook (1986) estudia la superficie del desplazamiento de verosimilitudes y su método permite detectar casos potencialmente influyentes sobre el estimador MV, Cadigan y Farrel (2002) estudian la dirección de máxima pendiente de una función del estimador MV, Wu y Luo (1993) estudian la superficie, a través de la curvatu-

ra de una función del estimador MV, llamada influencia generalizada de segundo orden, y la influencia de primer orden es la derivada direccional de esta función, que coincide con la máxima pendiente local. Finalmente, debido a que las curvaturas no son siempre invariantes Zhu *et al.* (2007) proponen otra medida de influencia que siempre serán invariantes, denominadas de primer y segundo orden, que permiten estudiar la curvatura de la superficie generada por cualquier función de los estimadores MV.

Aplicación

La prueba de la Polisomnografía (PSG) es un procedimiento que registra las medidas de un conjunto de variables que permiten establecer un diagnóstico de trastornos relacionados con el sueño. Svetnik *et al.* (2007) presentan un experimento clínico donde consideran a 82 pacientes que sufren de insomnio transitorio. Los registros de las variables de la PGS fueron recolectados durante dos noches, una bajo un placebo y la otra bajo un tratamiento (se administra una droga llamada Zolpidem 10 mg).

Para la prueba de la PGS se aplicaron varios métodos de registro de las variables consideradas en el estudio del sueño, éstos son:

1. Manual (M): El método manual registra las variables de la PGS a través de un operador experto.
2. Automático (A): El método automático registra las variables de la PGS con un instrumento desarrollado por la compañía WideMed, llamado Morpheus.
3. Parcial (P): El método parcial registra las variables de la PGS con el instrumento automático (Morpheus) pero con revisión manual realizada por un operador experto.

En la práctica es común usar el tiempo que se demora un paciente en conciliar el sueño, esta variable es llamada latencia del sueño persistente (LPS: Latency to

persistent sleep, en inglés) y corresponde al tiempo (en minutos) de sueño consecutivo ininterrumpido.

Es de interés estudiar el grado de acuerdo entre las medidas de LPS generadas por los distintos métodos, ya que se requiere saber si el instrumento Morpheus genera medidas concordantes respecto al método manual.

El objetivo de la aplicación es ilustrar la eficacia del método de diagnóstico, el cual permite detectar casos potencialmente influyentes sobre el CCC y además ilustrar que bajo el supuesto distribucional t la estimación del CCC es menos sensible en presencia de datos atípicos y/o extremos, respecto a la estimación obtenida cuando se supone normalidad de datos.

5.1. Aplicación 1: Estudio del sueño

La aplicación 1 considera los registros del LPS en escala logaritmo que han sido generados con el método Manual (M) y el Automático (A).

5.1.1. Descripción de los datos

El CCC muestral para los registros en logaritmo del LPS es de 0.675 ($C_b = 0.943$ y $\rho_{12} = 0.715$) limitando con el punto de corte de 0.65 propuesto por McBride (2005), lo que sugiere un bajo acuerdo entre las medidas. Sin embargo, la Figura 5.1 no muestra una gran pérdida de exactitud (sesgo) pero sí de precisión (grado de dispersión). Se observa que los casos 1, 30 y 79 están alejados del conjunto de datos y posiblemente estén afectando el valor del CCC.

Por otro lado, la Figura 5.2 no presenta una relación clara entre las diferencias del $\log(\text{LPS})$ y su promedio. El promedio de las diferencias (0.245) es un estimador del sesgo (diferencia sistemática entre los métodos) y su desviación estándar (0.786) mide las fluctuaciones alrededor de su media. Existen algunos casos con un sesgo mayor respecto a aquellos que fluctúan alrededor del promedio de las diferencias, además los casos 1, 30, 35 y 79 están fuera de los límites de acuerdo

(-1.327 y 1.817). Se observa que no existe un aumento o disminución del sesgo (diferencias) con la magnitud (promedio de las diferencias).

La Tabla 5.1 resume las estadísticas descriptivas univariadas, se observa que la distribución del logaritmo de los registros generado por el método M presenta una leve asimetría a la izquierda, mientras que el método A presenta una leve asimetría a la derecha. En cuanto a la dispersión de los datos, el método A tiene mayor variabilidad respecto al método M ($CV_A = 48\%$, $CV_M = 34\%$). Notar que el método A genera un máximo registro de 6.17 y un mínimo registro de 0.00 ($Rango_A = 6.175$, $Rango_M = 4.067$).

El vector de medias y la matriz de varianza-covarianza muestral para el método M y A es

$$\bar{\mathbf{x}} = \begin{pmatrix} 2.554 \\ 2.308 \end{pmatrix}, \mathbf{S}_n = \begin{pmatrix} 0.762 & 0.694 \\ 0.694 & 1.237 \end{pmatrix} \text{ y } \mathbf{S}_{n-1} = \begin{pmatrix} 0.771 & 0.703 \\ 0.703 & 1.252 \end{pmatrix}$$

donde \mathbf{S}_n es el estimador muestral no insesgado y \mathbf{S}_{n-1} es el estimador muestral insesgado.

La varianza generalizada es $|\mathbf{S}_n| = 0.460$ ($|\mathbf{S}_{n-1}| = 0.472$) y la desviación típica generalizada es $|\mathbf{S}_n|^{1/2} = 0.678$ ($|\mathbf{S}_{n-1}|^{1/2} = 0.687$) indicando que el área que ocupan los datos es más amplio, más aún, ésto también es indicado por el coeficiente de correlación entre las medidas que es $\rho_{12} = 0.715$.

Finalmente, los registros de $\log(\text{LPS})$ generados con el método A y M siguen una distribución normal marginal (ver Tabla 5.2 prueba de Shapiro-Wilk calculado usando la librería stats de R) pero no siguen una distribución normal bivariada (ver Tabla 5.3 prueba Henze-Zirkler dado en Henze y Zirkler (1990) y calculado usando la librería MVN de R, y prueba Shapiro-Wilk (\mathcal{N})). Sin embargo, la distancia transformada de los registros $\log(\text{LPS})$ sigue una distribución t bivariada. La Figura 5.3 (b) y la Tabla 5.3 prueba Shapiro-Wilk (\mathcal{T}) apoyan ésta afirmación.

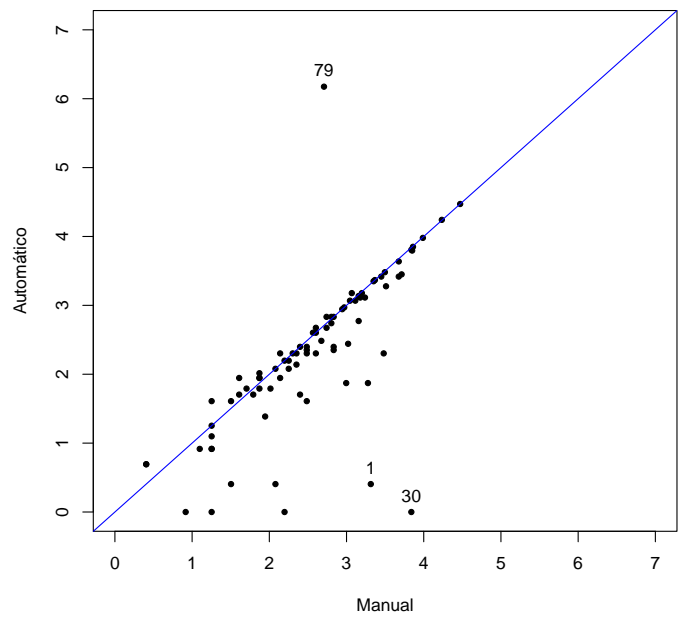


Figura 5.1: Dispersión entre los registros $\log(\text{LPS})$ obtenidos con el proceso Manual y Automático.

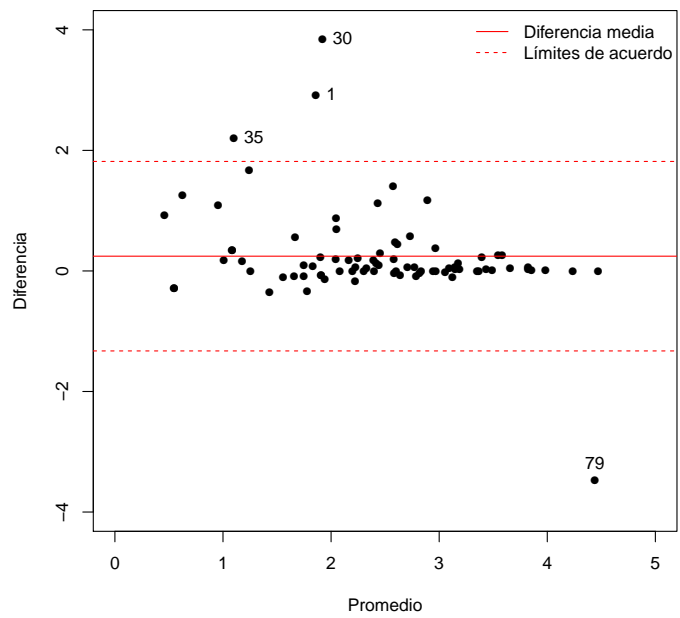


Figura 5.2: Límites de acuerdo de Altman para los registros $\log(\text{LPS})$.

Tabla 5.1: Estadística descriptiva univariada para los registros log(LPS) generados por el método manual y automático.

	n	Media	Est. Desv.	Mediana	Min	Max	Q(25%)	Q(75%)	Asimetría	Curtosis
Manual	82	2.55	0.88	2.60	0.40	4.47	1.96	3.17	-0.23	-0.43
Automático	82	2.31	1.12	2.30	0.00	6.17	1.73	3.07	0.12	0.76

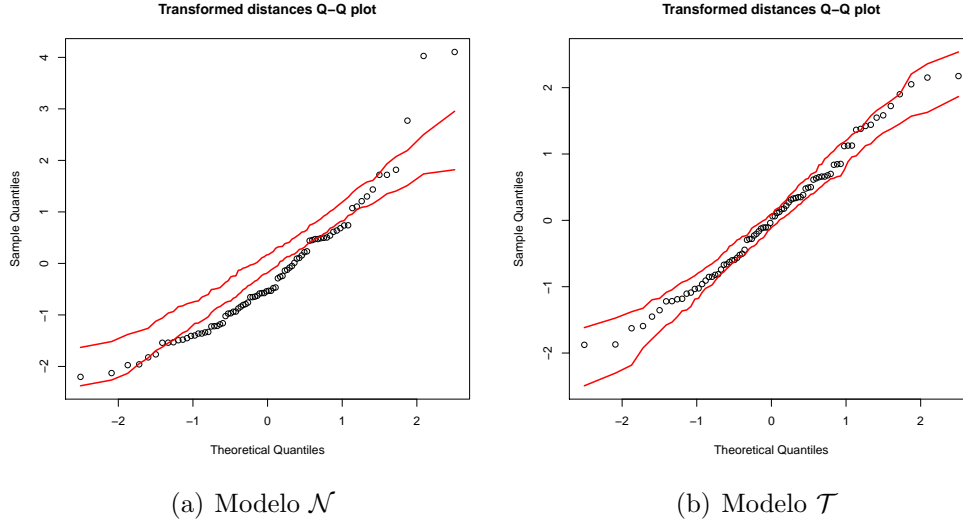


Figura 5.3: QQ-plots para las distancia transformada, asumiendo el modelo \mathcal{N} (a) y asumiendo el modelo \mathcal{T} (b), considerando el método M y A.

Tabla 5.2: Prueba de normalidad Shapiro-Wilk para los registros log(LPS).

	Estadístico	p-valor	Normalidad
Manual	0.9890	0.7132	Sí
Automático	0.9734	0.0864	Sí

Tabla 5.3: Prueba de normalidad multivariada para los registros log(LPS) y prueba de normalidad para la distancia transformada, asumiendo el modelo \mathcal{N} y el modelo \mathcal{T} .

	Estadístico	p-valor	Normalidad
Henze-Zirkler	4.786	0.000	No
Shapiro-Wilk (\mathcal{N})	0.920	0.000	No
Shapiro-Wilk (\mathcal{T})	0.983	0.338	Sí

5.1.2. Modelo estadístico y ajuste

Sea $\mathbf{X}_1, \dots, \mathbf{X}_{82}$ una muestra de vectores aleatorios con $\mathbf{X}_i = (X_{i1}, X_{i2})^T$. La variable aleatoria X_{i1} representa el registro en logaritmo generado por el método M y X_{i2} el registro en logaritmo generado por el método A. Se asumen dos modelos estadísticos para representar los registros en logaritmo \mathbf{X}_i del i -ésimo paciente.

- **Modelo \mathcal{N} :** Asume que $\mathbf{X}_i \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ para todo $i = 1, \dots, 82$ pacientes.
- **Modelo \mathcal{T} :** Asume que $\mathbf{X}_i \sim \mathcal{T}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ para todo $i = 1, \dots, 82$ pacientes.

El ajuste del modelo \mathcal{N} , con $\log Lik(\mathcal{N}) = -200.890$ y convergencia del algoritmo EM en 1 iteración, está dado por

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 2.5539 \\ 2.3090 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.762 & 0.694 \\ 0.694 & 1.237 \end{pmatrix},$$

mientras que el ajuste del modelo \mathcal{T} , con $\log Lik(\mathcal{T}) = -165.5504$ y convergencia del algoritmo EM en 570 iteraciones, está dado por

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 2.6171 \\ 2.5345 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 11.42685 & 11.25016 \\ 11.25016 & 12.31703 \end{pmatrix} \quad \text{y} \quad \hat{\eta} = 0.480.$$

El vector de medias estimado por el modelo \mathcal{N} es similar al vector de medias estimado por el modelo \mathcal{T} , pero la estimación de las matrices de varianza-covarianza son diferentes. El modelo \mathcal{T} recoge mayor variabilidad de los datos, pues la varianza generalizada estimada es de 14.179 siendo mayor que la varianza generalizada estimada del modelo \mathcal{N} que es de 0.460. Además, el parámetro de forma estimado está lejos de cero indicando que los datos están alejados de la normalidad, más aún, el valor estimado está cercano a 0.5, estando en el límite de la existencia del parámetro de forma.

La región de confianza del 95 % para el vector de medias, representada por la elipse basada en los estimadores MV del modelo \mathcal{N} y el modelo \mathcal{T} , se muestra en la Figura 5.4. La elipse está inclinada a la derecha puesto que la covarianza es

positiva entre las dos variables, además el área que ocupan los datos es mayor en el modelo \mathcal{T} ($|\hat{\Sigma}| = 14.179$) que en el modelo \mathcal{N} ($|\hat{\Sigma}| = 0.460$).

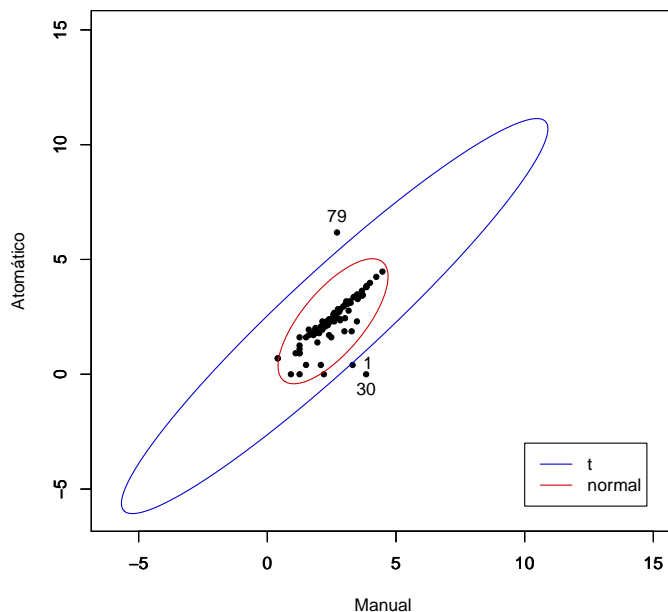


Figura 5.4: Elipse del 95 % basada en el vector de medias estimado y matriz de varianza-covarianza estimada, asumiendo el modelo \mathcal{N} (línea roja) y el modelo \mathcal{T} (línea azul).

La Tabla 5.4 resume el intervalo de confianza asintótico (ICA) para los dos modelos propuestos y el intervalo de confianza basado en la Z-transformación de Fisher propuesta por Lin (1989).

Se observa una gran diferencia en las estimas del CCC y en sus respectivos intervalos al comparar el modelo \mathcal{T} respecto al modelo \mathcal{N} , de igual manera ocurre con el coeficiente de precisión. Sin embargo, no hay una gran diferencia en el coeficiente del sesgo.

5.1.3. Diagnóstico

Desde el análisis descriptivo el gráfico de Altman detectó el siguiente conjunto de casos potencialmente influyente $I_A = \{1, 30, 79\}$. La Figura 5.2 muestra que los

Tabla 5.4: Estimador del CCC, error estándar de estimación (eee), intervalo de confianza asintótico (ICA), precisión y sesgo.

	$\hat{\rho}_c(\text{eee})$	ICA	Precisión	Sesgo
Modelo \mathcal{N}	0.674 (0.056)	(0.566, 0.783)	0.715	0.943
Modelo \mathcal{T}	0.947 (0.014)	(0.921, 0.974)	0.948	0.999
Fisher	0.675 (0.058)	(0.561, 0.788)	0.715	0.943

casos están fuera de los límites de acuerdo indicando un pobre acuerdo entre las medidas.

Por otro lado, la Figura 5.5 muestra las observaciones que se detectaron a través de la distribución de los pesos \hat{u} versus la distancia de Mahalanobis $\hat{\delta}$ al asumir el modelo \mathcal{T} . Luego, el conjunto $I_u = \{1, 3, 30, 35, 48, 79\}$ es potencialmente influyente, ya que están por debajo del percentil 5% de la distribución de los pesos dada en Osorio y Galea (2016).

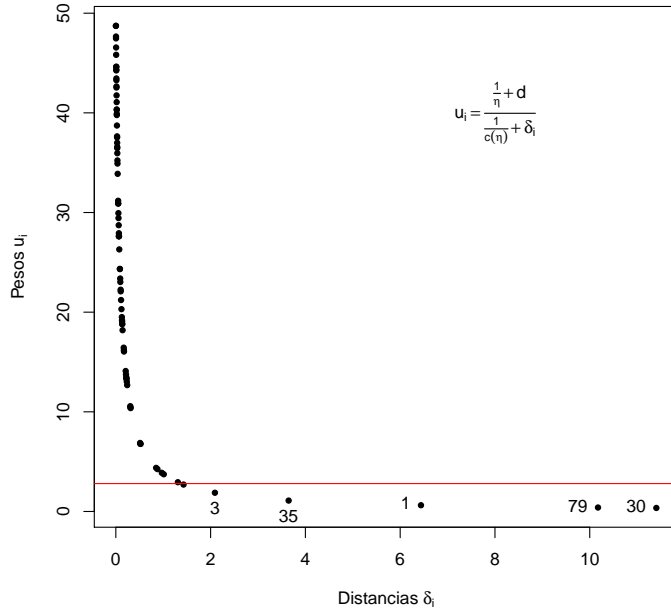


Figura 5.5: Observaciones detectadas como potencialmente influyentes a través de la distribución de los pesos versus la distancia de Mahalanobis, asumiendo el modelo \mathcal{T} .

En el estudio de influencia, bajo el esquema de perturbación de ponderación de casos y asumiendo el modelo \mathcal{N} , se detectaron los casos 1, 30 y 79 como potencialmente influyentes sobre el estimador MV. Al asumir el modelo \mathcal{T} , si bien hay casos que limitan con el punto de corte, ningún caso se considera como potencialmente influyente sobre el estimador MV (Ver Figura 5.7). El mismo resultado ocurre cuando se considera la dirección donde se produce la mayor conformal curvatura normal (ver Figura 5.8).

Bajo el modelo \mathcal{N} , la Figura 5.9 gráfico (a), muestra que el caso 79 es detectado por la máxima dirección \mathbf{h}_{max} y la segunda máxima dirección \mathbf{h}_{seg} , mientras que el caso 1 y 30 son detectados sólo por \mathbf{h}_{max} . En el caso del modelo \mathcal{T} , gráfico (b), no existen casos con influencia importante.

En cuanto a la influencia local generalizada, el método detectó que los casos 1, 30 y 79 pueden ser influyentes directamente sobre el estimador MV del CCC al asumir el modelo \mathcal{N} (ver Figura 5.10, gráfico (a)). No obstante, si se asume el modelo \mathcal{T} , el método no detecta casos con influencia importante (ver Figura 5.10, gráfico (b)).

Por lo tanto, bajo el esquema de ponderación de casos, el método detectó el caso 1, 30 y 79 como potencialmente influyentes sobre el estimador MV y el estimador del CCC, bajo el modelo \mathcal{N} .

Bajo el esquema de perturbación de la respuesta, el cual permite detectar datos extremos, para el modelo \mathcal{N} se detectan tres casos potencialmente influyentes que son el 34, 53 y 59 (ver Figura 5.11, gráfico (a)), y en el modelo \mathcal{T} no se detectan casos potencialmente influyentes. Sin embargo, para el modelo \mathcal{N} , basado en la conformal curvatura normal, el método detecta los casos 1, 30 y 79 como potencialmente influyentes (ver Figura 5.12, gráfico (a)), y asumiendo el modelo \mathcal{T} el método no detecta casos potencialmente influyentes (gráfico (b)).

La Figura 5.13 muestra que los casos 1, 30, 35 y 79 son potencialmente influ-

yentes a lo largo de la segunda dirección máxima \mathbf{h}_{seg} , mientras que el caso 34 y 53 son potencialmente influyentes a lo largo de la primera dirección máxima \mathbf{h}_{max} .

Por otro lado, bajo el modelo \mathcal{N} , la influencia local generalizada para la máxima pendiente (que coincide con la medida de primer orden) detecta los casos 8, 34, 53, 59 y 79 como posiblemente influyentes sobre el estimador del CCC, donde el caso 34, 53 y 59 son clasificados como extremos. En contraste, asumiendo el modelo \mathcal{T} , no se detectan casos potencialmente influyentes sobre el estimador del CCC (ver Figura 5.14).

Finalmente, en el caso de usar las medidas de segundo orden, asumiendo el modelo \mathcal{N} , se detectan los casos 8, 30, 34 y 79 como potencialmente influyentes, bajo el esquema de ponderación de casos y bajo el esquema de perturbación de la respuesta se detectan los casos 4, 6, 7 y 10 (ver Figura 5.15).

La Tabla 5.5 resume los casos detectados como potencialmente influyentes sobre los estimadores MV y el estimador del CCC, usando las medidas de influencia local, las medidas de influencia local generalizada y las medidas de primer y segundo orden para ambos modelos bajo el esquema de ponderación de casos y de perturbación de la respuesta.

Se presenta, en la Tabla 5.6, que los casos 1, 30 y 79 tienen alta variabilidad y las medidas son diferentes, mientras que el resto de los casos están en acuerdo pero su magnitud es muy alta o muy baja, y los considerados extremos son casos detectados bajo el esquema de perturbación de la respuesta.

Para evaluar la sensibilidad de los estimadores MV y del estimador del CCC, en ambos modelos, se consideran los siguientes conjuntos de casos potencialmente influyentes $I_1 = \{1, 30, 79\}$, $I_2 = \{4, 6, 7, 8, 10, 34, 53, 59\}$.

La Tabla 5.7 muestra la Razón de Cambio (RC) para el estimador del CCC al eliminar los casos considerados en desacuerdo I_1 , bajo el modelo \mathcal{N} el estimador

Tabla 5.5: Resumen de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCC respecto al esquema de perturbación, medida de influencia y modelo estadístico.

Influencia en	EMV		EMV del CCC	
	Modelo \mathcal{N}	Modelo \mathcal{T}	Modelo \mathcal{N}	Modelo \mathcal{T}
Ponderación				
\mathbf{h}_{max} C	1-30-79	No detecta	1-30-79	No es
\mathbf{h}_{max} B	30-79	No detecta	influyentes	influyente
$\mathbf{h}_{max}/\mathbf{h}_{seg}$	1-30-79	No detecta	indirectamente	indirectamente
\mathbf{h}_{FI}	*	*	1-30-79	No detecta
\mathbf{h}_{SI}	*	*	8-30-34-79	*
Respuesta				
\mathbf{h}_{max} C	No detecta	No detecta	1-30-79	No es
\mathbf{h}_{max} B	1-30-79	No detecta	influyentes	influyente
$\mathbf{h}_{max}/\mathbf{h}_{seg}$	1-30-79	No detecta	indirectamente	indirectamente
\mathbf{h}_{max} FI	*	*	8-34-53-59-79	No detecta
\mathbf{h}_{max} SI	*	*	4-6-7-10	*

Tabla 5.6: Valores, diferencia y estado de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCC.

Casos	Método M	Método A	Diferencia	Desv. Est.	Estado
1	3.314	0.405	2.909	2.057	desacuerdo
4	2.833	2.398	0.435	0.308	extremo
6	3.045	3.068	-0.024	0.017	acuerdo
7	3.350	3.350	0.000	0.000	acuerdo
8	0.916	0.000	0.916	0.648	extremo
10	4.234	4.241	-0.007	0.005	acuerdo
30	3.839	0.000	3.839	2.715	desacuerdo
34	4.472	4.472	0.000	0.000	acuerdo
53	0.405	0.693	-0.288	0.203	extremo
59	0.405	0.693	-0.288	0.203	extremo
79	2.708	6.175	-3.467	2.451	desacuerdo

del CCC presenta una mayor RC respecto al modelo \mathcal{T} , indicando que la estimación del CCC basada en este modelo es más estable.

Sin embargo, al realizar el análisis de sensibilidad para los casos I_2 , la RC es menor al 7% asumiendo el modelo \mathcal{N} y alrededor del 1% asumiendo el modelo \mathcal{T} , por lo que se considera que en ambos casos el estimador del CCC varía marginalmente.

Para los estimadores de MV la Tabla 5.8 muestra que, bajo el modelo \mathcal{N} y el modelo \mathcal{T} , la RC no es importante sobre el estimador del vector de medias pero sí lo es sobre el estimador de la varianza generalizada, la cual es considerablemente mayor en el modelo \mathcal{T} .

Por otro lado, la RC sobre el estimador del parámetro de forma es marginal y está bastante lejos de cero, indicando que no es adecuado modelar los datos bajo la distribución normal.

El estimador del CCC y su respectivo intervalo de confianza asintótico (ICA) bajo el modelo \mathcal{N} , el modelo \mathcal{T} , la Z-transformación de Fisher y Bootstrap no paramétrico (Boot) con todos los casos y eliminando los casos 1, 30 y 79, se muestran en la Tabla 5.9.

El ICA bajo el modelo \mathcal{N} es similar al de la Z-transformación de Fisher y más precisos respecto al intervalo Boot no paramétrico. Al eliminar los casos potencialmente influyentes los intervalos cambian significativamente. En el modelo \mathcal{T} , el ICA es muy diferente al ser comparado con los otros, notar que el proceso de estimación provee un estimador del CCC más alto y al eliminar los casos potencialmente influyentes el intervalo no se ve afectado.

Feng *et al.* (015a) llegan a una conclusión similar, la Tabla 5.10 muestra los resultados del estimador puntual del CCC e intervalo de credibilidad/confianza obtenidos por diferentes métodos; el método de Bayes basado en la distribución t multivariada (Bayes), Jackknife (J), Jackknife con Z-transformación de Fisher (JZ) y Bootstrap (Boot). Todas las estimas con J, JZ y Boot son similares y al compararlos con Bayes éste provee un estimador del CCC robusto y un intervalo de credibilidad estable (ver Figura 5.6).

Finalmente, el proceso de estimación sin el conjunto I_1 para el modelo \mathcal{N} , con convergencia del algoritmo EM en 1 es

$$\hat{\mu}_{I_1} = \begin{pmatrix} 2.526 \\ 2.313 \end{pmatrix}, \quad \hat{\Sigma}_{I_1} = \begin{pmatrix} 0.761 & 0.769 \\ 0.769 & 0.981 \end{pmatrix}$$

y para el modelo \mathcal{T} , con convergencia del algoritmo EM en 449, es

$$\hat{\mu}_{I_1} = \begin{pmatrix} 2.623 \\ 2.547 \end{pmatrix}, \quad \hat{\Sigma}_{I_1} = \begin{pmatrix} 3.839 & 3.792 \\ 3.792 & 4.058 \end{pmatrix} \text{ y } \hat{\eta}_{I_1} = 0.439.$$

Tabla 5.7: Razón de de cambio para el estimador del CCC y medidas de calidad del ajuste al eliminar casos potencialmente influyentes.

Casos eliminados	CCC				Comparación del ajuste			
	Modelo \mathcal{N}	RC	Modelo \mathcal{T}	RC	$\log Lik(\mathcal{N})$	$\log Lik(\mathcal{T})$	$AIC(\mathcal{N})$	$AIC(\mathcal{T})$
Ninguno	0.674	*	0.947	*	-200.890	-165.550	411.780	343.101
1	0.715	6 %	0.953	1 %	-192.990	-156.628	395.979	325.257
30	0.749	11 %	0.952	1 %	-186.448	-155.482	382.897	322.964
79	0.724	7 %	0.952	1 %	-185.907	-155.723	381.815	323.445
1-30	0.795	18 %	0.957	1 %	-175.741	-146.370	361.482	304.739
1-79	0.770	14 %	0.957	1 %	-175.982	-146.620	361.963	305.239
30-79	0.808	20 %	0.956	1 %	-166.688	-145.465	343.376	302.931
1-30-79	0.860	28 %	0.959	1 %	-150.728	-136.129	311.456	284.258

Tabla 5.8: Razón de cambio para los estimadores MV del modelo \mathcal{N} y \mathcal{T} al eliminar casos potencialmente influyentes.

Estimador	Modelo \mathcal{N}			Modelo \mathcal{T}		
	Ninguno	1-30-79	RC	Ninguno	1-30-79	RC
$\hat{\mu}_1$	2.554	2.526	1 %	2.617	2.623	0 %
$\hat{\mu}_2$	2.309	2.313	0 %	2.535	2.547	0 %
$ \hat{\Sigma} $	0.460	0.156	66 %	14.179	1.204	92 %
$\hat{\eta}$	*	*	*	0.480	0.439	9 %

Tabla 5.9: Estimador puntual, error estándar de estimación e intervalo de confianza asintótico/confianza para el CCC basado en distintos métodos.

Métodos	Todos los casos		Sin los casos 1-30-79	
	$\hat{\rho}_c$ (eee)	ICA/IC	$\hat{\rho}_c$ (eee)	ICA/IC
Modelo \mathcal{N}	0.674 (0.056)	(0.566, 0.783)	0.860 (0.055)	(0.753, 0.967)
Modelo \mathcal{T}	0.947 (0.014)	(0.921, 0.974)	0.959 (0.011)	(0.939, 0.980)
Fisher	0.675 (0.058)	(0.561, 0.788)	0.861 (0.030)	(0.803, 0.919)
Boot	0.675 (0.104)	(0.472, 0.878)	0.861 (0.041)	(0.782, 0.944)

Tabla 5.10: Estimador puntual e intervalo de credibilidad/confianza para del CCC basado en distintos métodos.

Métodos	Todos los casos		Sin los casos 1-30-79	
	$\hat{\rho}_c$	IC	$\hat{\rho}_c$	IC
Bayes	0.870	(0.799, 0.927)	0.894	(0.835, 0.939)
J	0.672	(0.457, 0.888)	0.862	(0.777, 0.948)
JZ	0.656	(0.346, 0.837)	0.856	(0.729, 0.926)
Boot	0.675	(0.475, 0.861)	0.861	(0.771, 0.931)

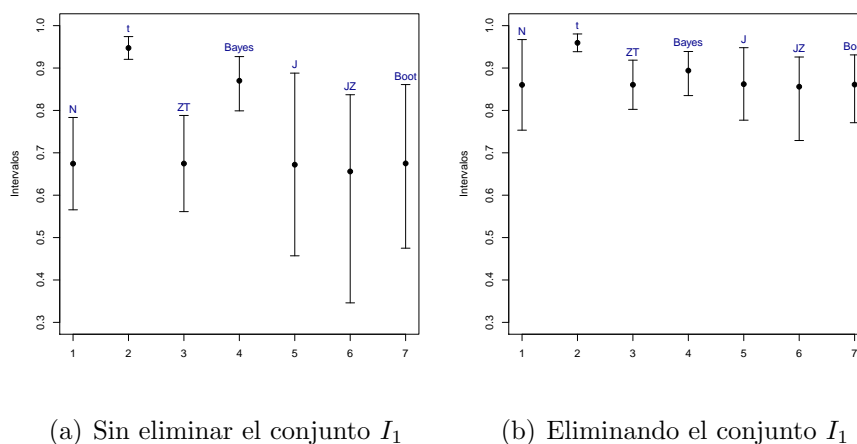


Figura 5.6: Comparación de los intervalos de confianza y credibilidad basado en diferentes métodos.

5.1.4. Conclusión

En la práctica, es importante obtener estimaciones robustas del CCC, ya que este coeficiente permite cuantificar el acuerdo entre los registros del $\log(\text{LPS})$. Los registros del generados por el método A presentan una mayor variabilidad respecto a los registros generados por el método M.

El método de diagnóstico de influencia local y la influencia local generalizada de primer orden determinaron que los casos 1, 30 y 79 son potencialmente influyentes sobre los estimadores MV y sobre el estimador del CCC. El análisis de sensibilidad llevado a cabo muestra que efectivamente estos casos influyen sobre el estimador del CCC bajo el esquema de ponderación de casos al asumir el modelo \mathcal{N} . Sin embargo, cuando se asume el modelo \mathcal{T} el método no detecta casos potencialmente influyentes, más aún, al eliminar los casos 1, 30 y 79 se obtienen estimaciones del CCC que no son sensibles.

Por lo tanto, en base al criterio del logaritmo de la verosimilitud y el criterio AIC, el modelo \mathcal{T} modela mejor los datos que el modelo \mathcal{N} , entregando estimaciones del CCC e ICA menos sensibles en presencia del conjunto I_1 .

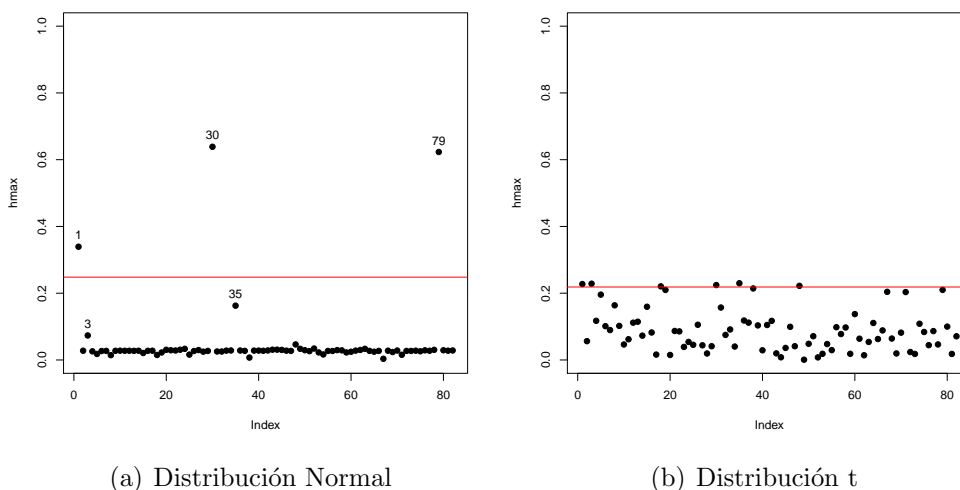
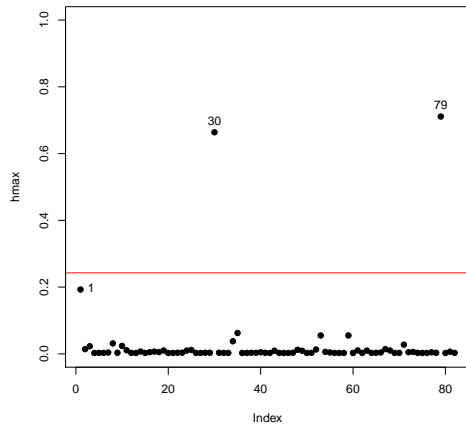
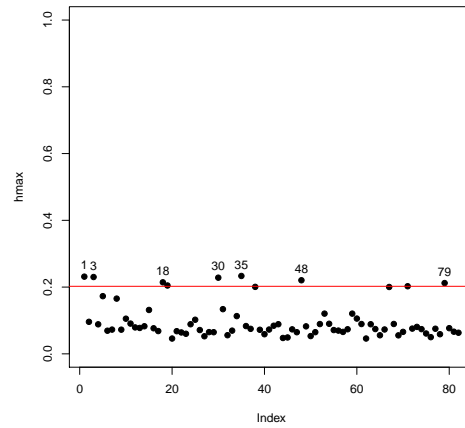


Figura 5.7: Gráfico de influencia de h_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.

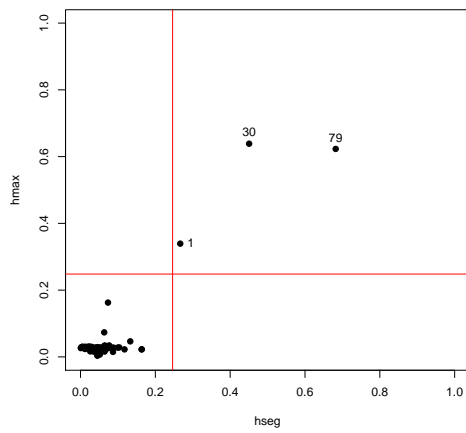


(a) Distribución Normal

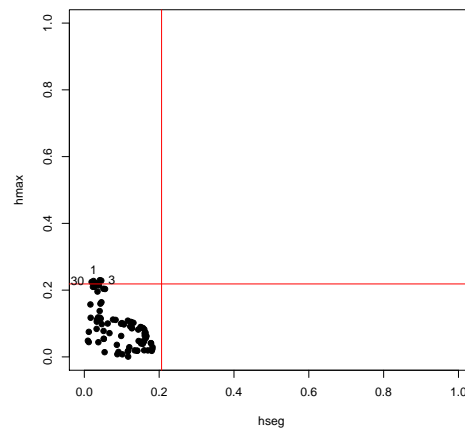


(b) Distribución t

Figura 5.8: Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.

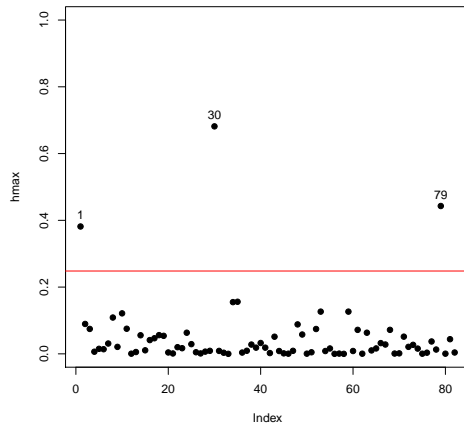


(a) Distribución Normal

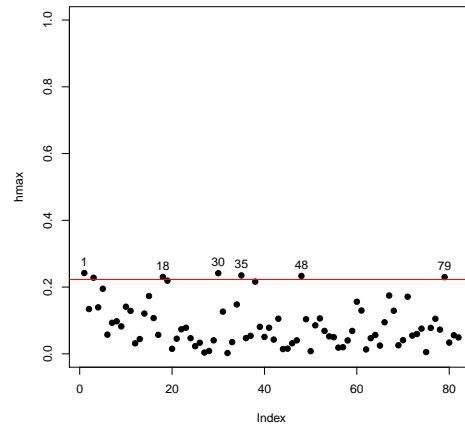


(b) Distribución t

Figura 5.9: Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.

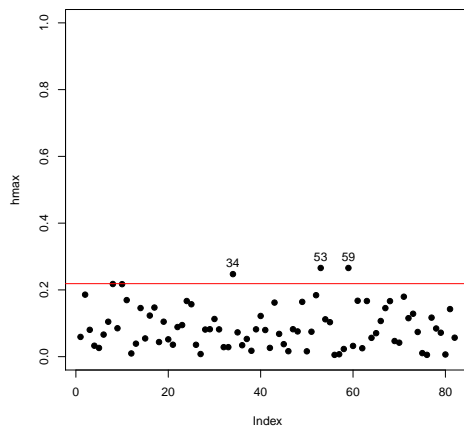


(a) Distribución Normal

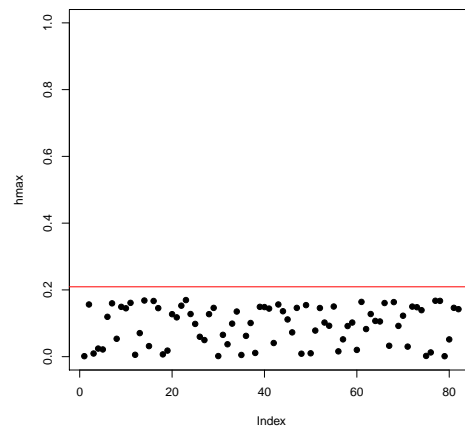


(b) Distribución t

Figura 5.10: Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local (o medida de primer orden) asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de ponderación de casos.



(a) Distribución Normal



(b) Distribución t

Figura 5.11: Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.

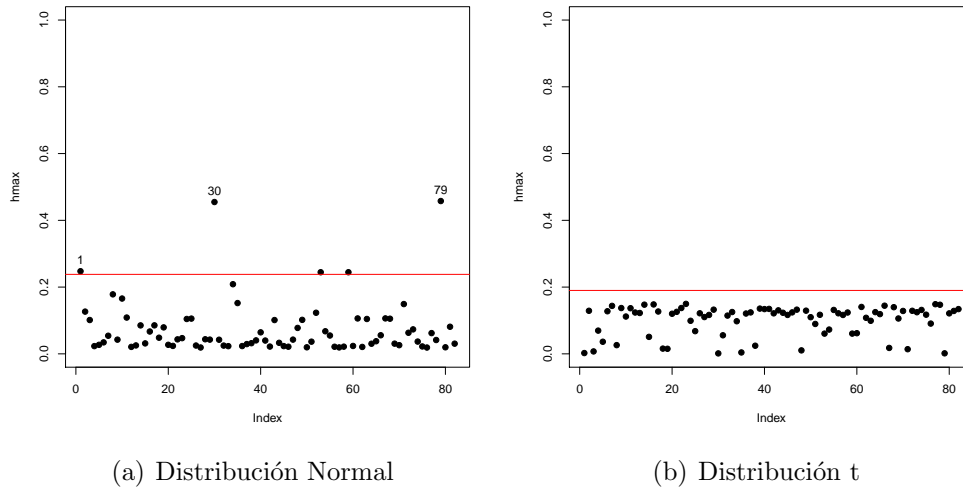


Figura 5.12: Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.

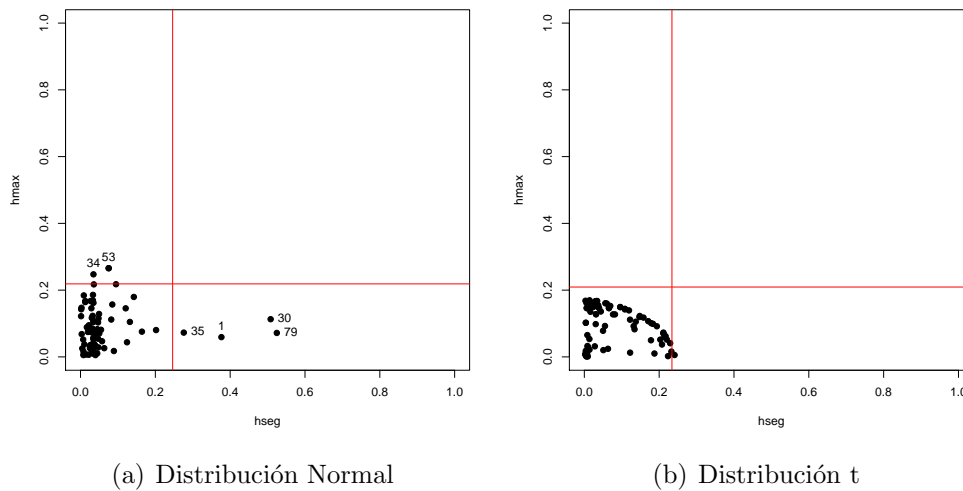
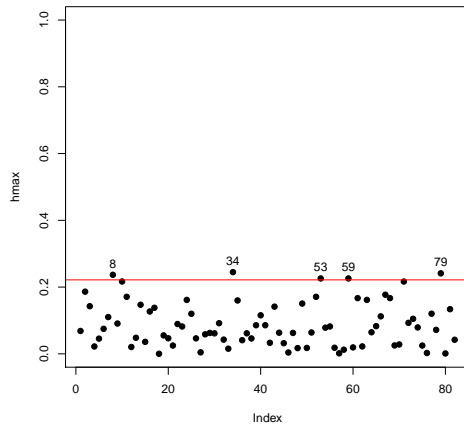
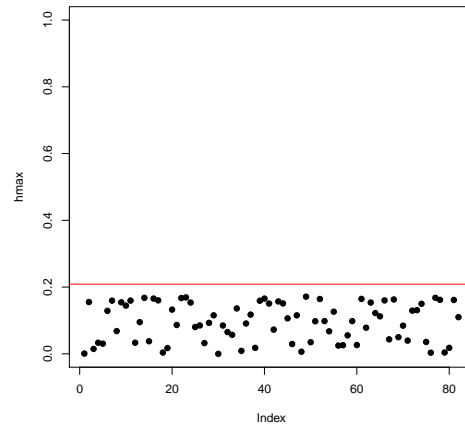


Figura 5.13: Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.

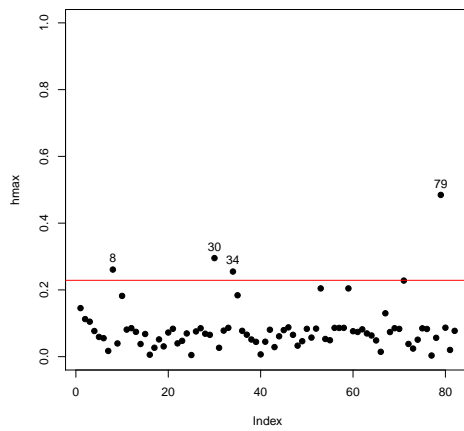


(a) Distribución Normal

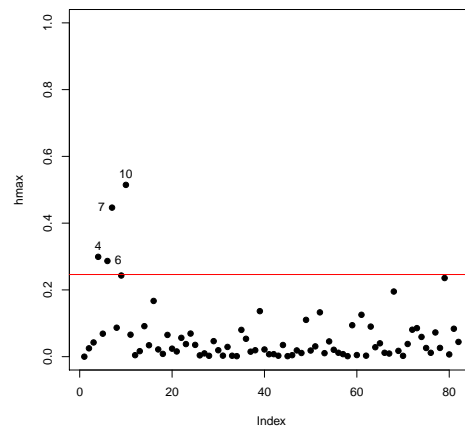


(b) Distribución t

Figura 5.14: Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local (o medida de primer orden) asumiendo que los datos siguen una distribución normal (a) y una distribución t (b), bajo el esquema de perturbación de la respuesta.



(a) Ponderación de casos



(b) Perturbación de la respuesta

Figura 5.15: Gráfico de influencia de \mathbf{h}_{max} para la medida de influencia de segundo orden asumiendo que los datos siguen una distribución normal bajo ambos esquemas de perturbación.

5.2. Aplicación 2: Estudio del sueño con tres métodos

Además de los métodos manual y automático usados en el estudio del sueño (aplicación 1) para analizar el efecto del zolpidem 10 mg en un grupo de 82 pacientes que sufren de insomnio transitorio, se aplicó un tercer método para obtener los registros del log(LPS) en la PSG. El tercer método consiste en usar de forma semiautomática el instrumento Morpheus y es denominado como el método Parcial (P).

5.2.1. Descripción de los datos

El estimador muestral del CCCG para los registros de log(LPS) es de 0.752, con una precisión global de 0.780 y exactitud global de 0.964. La Tabla 5.11 muestra el resumen del acuerdo entre las medidas de a pares y global de los registros de log(LPS).

Los registros generados por los métodos M y P son concordantes, pues el CCC es de 0.919, esto se debe a que el coeficiente de exactitud y precisión son altos, en el caso de los registros generados por los métodos M y A presentan un pobre acuerdo, pues el CCC es de 0.675 y en el caso del acuerdo de los registros generados por los métodos P y A el CCC es de 0.704. En estos dos últimos casos el bajo acuerdo es afectado por la baja precisión indicando más dispersión de los datos.

La Figura 5.16 muestra la dispersión entre los registros generados por el método M y P, notar que se observa una alta concentración de puntos alrededor de la línea de acuerdo, donde los casos 3, 17, 18, 19 y 59 están poco alejados de dicha línea. En el caso de los registros generados por el método P y A, los casos 1, 30, 17 y 79 están más alejados de la línea de acuerdo, lo que hace disminuir el CCC a través del coeficiente de precisión, la misma situación anterior se repite al analizar la dispersión entre los registros generados por el método M y A ya estudiado en la aplicación 1.

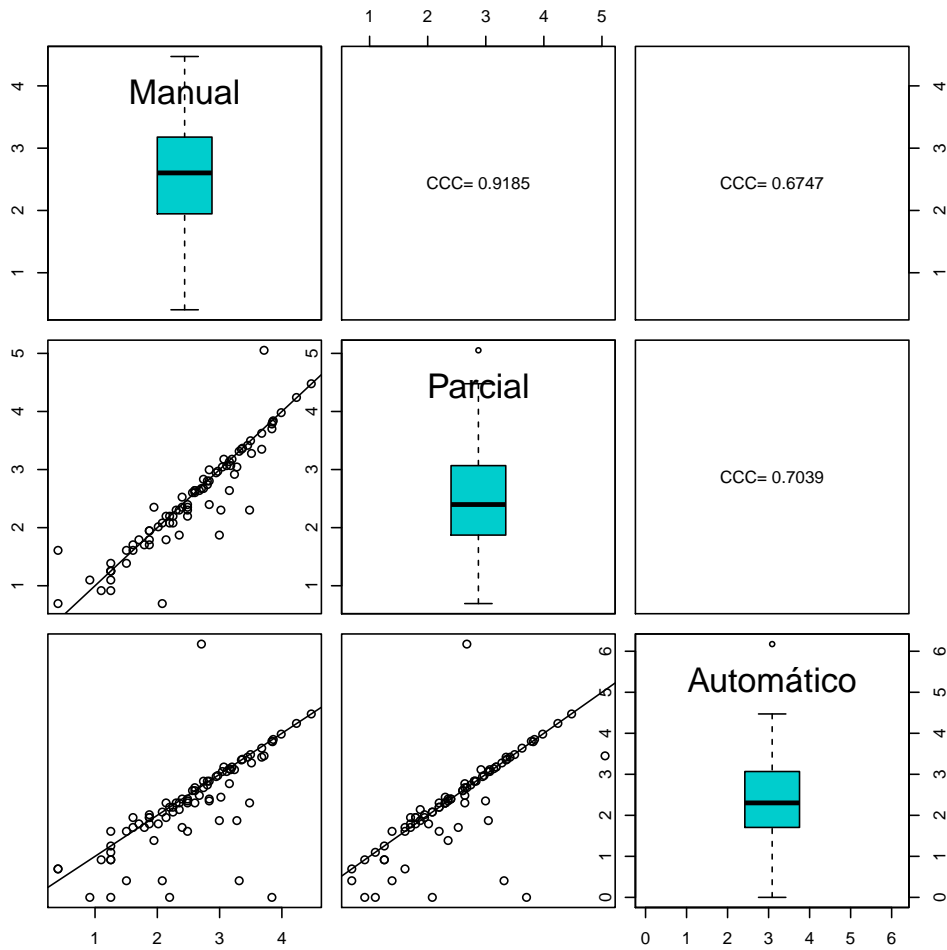


Figura 5.16: Dispersión entre los registros $\log(\text{LPS})$ obtenidos con el proceso Manual (M), Parcial (P) y Automático (A).

Por otro lado, la Figura 5.17, gráfico (a) muestra la dispersión entre la media y la desviación estándar (d_e) y el gráfico (b) muestra la dispersión entre la mediana y la desviación media (d_m) de los registros generados por los tres métodos. Los casos 1, 30, 35 y 79 tienen una mayor desviación estándar y desviación media respecto al resto de casos. Las medidas del caso 79 en promedio son de magnitud mayor respecto a las medidas del caso 1, 30 y 35, desde el punto de vista descriptivo, estos casos aumentan la variabilidad de los datos. Este hecho se refleja en el coeficiente de precisión global y por ende en el poco acuerdo global.

Tabla 5.11: Acuerdo bivariado y global para los registros del logaritmo de LPS generados por los métodos manual, parcial y automático.

	Coefficiente	Exactitud	Precisión
CCCG	0.752	0.964	0.780
CCC M-P	0.919	0.997	0.921
CCC M-A	0.675	0.943	0.715
CCC P-A	0.704	0.954	0.734

La Tabla 5.12 muestra los registros de los casos 1, 30, 35 y 79, los cuales presentan alta variabilidad al ser comparados, por ejemplo, con los casos 34 y 84 que presentan registros entre los métodos bastantes similares. Notar que el método automático es el que genera más diferencia entre los casos 1, 30, 35 y 79. Por lo tanto, el estimador muestral del CCCG es bajo debido a la alta variabilidad que presentan los registros generados por el método A.

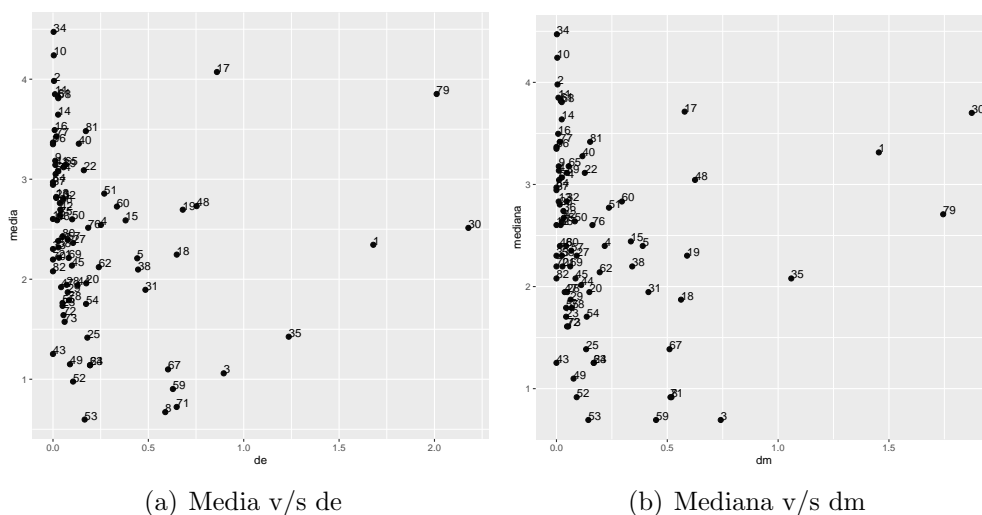


Figura 5.17: Media versus desviación estándar (de) (a) y mediana versus desviación media (dm) (b) de los registros obtenidos por los tres métodos.

La Tabla 5.13 resume las estadísticas univariadas de los registros generados por los tres métodos. El método A presenta una mayor variabilidad con respecto al método M y P ($CV_M = 34\%$, $CV_P = 36\%$ y $C_A = 48\%$), esto se debe a que la magnitud de los registros obtenidos por A es mayor respecto a los regis-

Tabla 5.12: Registros obtenidos por tres métodos, media, mediana y medidas de variabilidad.

Caso	Manual	Parcial	Automático	Media	Desv. Est.	Mediana	Desv. Media
34	4.472	4.477	4.472	4.474	0.003	4.472	0.002
82	2.079	2.079	2.079	2.079	0.000	2.079	0.000
1	3.314	3.314	0.405	2.345	1.679	3.314	1.454
30	3.839	3.701	0.000	2.514	2.178	3.701	1.874
35	2.197	2.079	0.000	1.426	1.236	2.079	1.059
79	2.708	2.674	6.175	3.852	2.011	2.708	1.745

tros obtenidos por los otros dos métodos ($\text{Rango}_M = 4.067$, $\text{Rango}_P = 4.360$, $\text{Rango}_A = 6.175$).

La distribución de los registros muestra una leve asimetría y curtosis siendo parecidas a la distribución normal univariada. La prueba de normalidad Shapiro-Wilk apoya la hipótesis de normalidad de los datos para los métodos M (p-valor es de 0.7132), A (p-valor es de 0.086) y P (p-valor es de 0.8688). Sin embargo, esto no necesariamente sugiere que los datos puedan ser modelados por una distribución normal multivariada, como se verá más adelante. Por otro lado, el vector de media

Tabla 5.13: Estadística descriptiva univariada para los puntajes log(LPS) obtenidos por los tres métodos.

	n	Media	Desv. Est.	Mediana	Min	Max	Q(25%)	Q(75%)	Asimetría	Curtosis
Manual	82	2.55	0.88	2.60	0.40	4.47	1.96	3.17	-0.23	-0.43
Parcial	82	2.48	0.89	2.40	0.69	5.05	1.87	3.06	0.22	-0.13
Automático	82	2.31	1.12	2.30	0.00	6.17	1.73	3.07	0.12	0.76

y la matriz de varianzas-covarianzas muestral para los métodos M, P y A es

$$\bar{\mathbf{x}} = \begin{pmatrix} 2.554 \\ 2.485 \\ 2.308 \end{pmatrix}, \mathbf{S}_n = \begin{pmatrix} 0.762 & 0.710 & 0.694 \\ 0.710 & 0.780 & 0.721 \\ 0.694 & 0.721 & 1.237 \end{pmatrix} \text{ y } \mathbf{S}_{n-1} = \begin{pmatrix} 0.771 & 0.719 & 0.703 \\ 0.719 & 0.790 & 0.729 \\ 0.703 & 0.729 & 1.252 \end{pmatrix}$$

donde \mathbf{S}_n es el estimador muestral no insesgado y \mathbf{S}_{n-1} es el estimador muestral insesgado.

La varianza generalizada es de $|\mathbf{S}_n| = 0.050$ ($|\mathbf{S}_{n-1}| = 0.052$) y la desviación típica generalizada es de $|\mathbf{S}_n|^{1/2} = 0.224$ ($|\mathbf{S}_{n-1}|^{1/2} = 0.228$), indicado una mayor dispersión de los datos, y por ende una menor precisión global, que es de 0.780. Este aumento de variabilidad posiblemente es debido a los casos 1, 30, 35 y 79 que se han estudiado anteriormente.

Respecto a la normalidad de los registros, la Figura 5.18 muestra los QQ-plots de la distancia transformada, asumiendo el modelo \mathcal{N} , gráfico (a) y asumiendo el modelo \mathcal{T} , gráfico (b). Se observa que los datos podrían seguir una distribución t . La Tabla 5.14 muestra los resultados de la prueba de normalidad multivariada del $\log(\text{LPS})$, la prueba de normalidad Shapiro-Wilk (\mathcal{N}) para la distancia transformada y la prueba Shapiro-Wilk (\mathcal{T}) para la distancia transformada. En el caso de la normalidad multivariada para el $\log(\text{LPS})$, éstos no siguen una distribución normal multivariada (p-valor es de 0.000) y en el caso de la normalidad de la distancia transformada, se concluye que los datos podrían ser modelados por una distribución t , aunque se debe destacar que el p-valor está al límite para rechazar la normalidad de la distancia transformada.

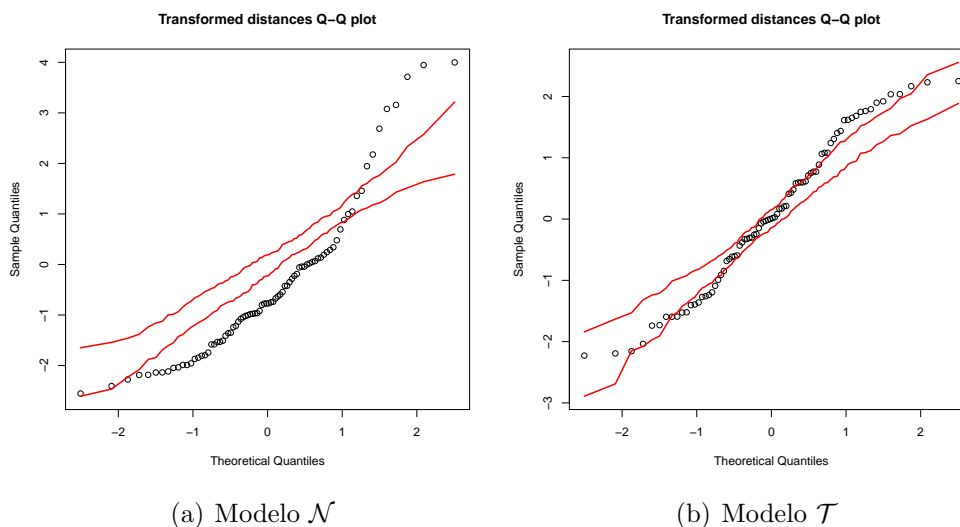


Figura 5.18: QQ-plots para las distancia transformada, asumiendo el modelo \mathcal{N} (a) y asumiendo el modelo \mathcal{T} (b), considerando el método M,P y A.

Tabla 5.14: Prueba de normalidad multivariada para los registros log(LPS) y prueba de normalidad para la distancia transformada, asumiendo el modelo \mathcal{N} y el modelo \mathcal{T} .

	Estadístico	p-valor	Normalidad
Henze-Zirkler (Normal)	6.926	0.000	No
Shapiro-Wilk (\mathcal{N})	0.894	0.000	No
Shapiro-Wilk (\mathcal{T})	0.969	0.046	Sí

5.2.2. Modelo estadístico y ajuste

Con $\mathbf{X}_1, \dots, \mathbf{X}_{82}$ una muestra de vectores aleatorios se tiene que $\mathbf{X}_i = (X_{i1}, X_{i2}, X_{i3})^T$. La variable aleatoria X_{i1} representa el registro del log(LPS) generado por el método M, X_{i2} registro del log(LPS) generado por el método P y X_{i3} registro del log(LPS) generado por el método A. Se asumen dos modelos estadísticos para representar los registros \mathbf{X}_i del i -ésimo paciente.

- **Modelo \mathcal{N} :** Asume que $\mathbf{X}_i \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ para todo $i = 1, \dots, 82$ pacientes.
- **Modelo \mathcal{T} :** Asume que $\mathbf{X}_i \sim \mathcal{T}_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ para todo $i = 1, \dots, 82$ pacientes.

El ajuste del modelo \mathcal{N} , con $\log \text{Lik}(\mathcal{N}) = -226.3502$ y convergencia del algoritmo EM en 1 iteración, está dado por

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 2.554 \\ 2.485 \\ 2.309 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.762 & 0.710 & 0.694 \\ 0.710 & 0.780 & 0.721 \\ 0.694 & 0.721 & 1.237 \end{pmatrix},$$

mientras que el ajuste del modelo \mathcal{T} , con $\log \text{Lik}(\mathcal{N}) = -141.3678$ y convergencia del algoritmo EM en 493 iteraciones, está dado por

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 2.617 \\ 2.625 \\ 2.535 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 29.356 & 28.794 & 29.047 \\ 28.794 & 29.066 & 29.162 \\ 29.047 & 29.162 & 30.374 \end{pmatrix} \quad \text{y} \quad \hat{\eta} = 0.493.$$

Basado en el criterio del logaritmo de verosimilitud, el modelo \mathcal{T} genera menos pérdida relativa de la información estadística que el modelo \mathcal{N} . El criterio de

Akaike indica la misma situación ($\text{AIK}_{\mathcal{N}} = 470.700$ y $\text{AIK}_{\mathcal{T}} = 302.736$).

Los vectores de medias estimados son similares en ambos modelos pero las matrices de varianzas-covarianzas son muy diferentes, siendo las varianzas y covarianzas del modelo \mathcal{T} mayor que las del modelo \mathcal{N} debido a que el estimador del parámetro de forma es cercano 0.5.

La varianza generalizada estimada y desviación típica generalizada, para el modelo \mathcal{N} , es de 0.050 y 0.224 respectivamente, mientras que para el modelo \mathcal{T} es 26.252 y 5.124 respectivamente. Por lo tanto, el modelo \mathcal{T} recoge mayor variabilidad de los registros $\log(\text{LPS})$.

Lo anterior, es confirmado por la estimación muestral de los coeficientes de asimetría y curtosis multivariada y la prueba de normalidad multivariada dado en Mardia *et al.* (1997), ver Tabla 5.15.

Tabla 5.15: Prueba de normalidad multivariada de Mardia basado en los coeficientes de asimetría y curtosis multivariada.

	Coeficiente	Estadístico	p-valor
Asimetría	2.743	37.493	0.000
Curtosis	35.707	17.117	0.000

El ICA, considerando el modelo \mathcal{N} , tiene un mayor error estándar de estimación (eee) y por ende una mayor amplitud al compararlo con el intervalo del modelo \mathcal{T} . Por otro lado, los intervalos son marcadamente diferentes, notar que no están superpuestos (ver Tabla 5.16).

Tabla 5.16: Estimador del CCCG, error estándar de estimación (eee) e intervalo de confianza asintótico (ICA).

	$\hat{\rho}_g$ (eee)	ICA
Modelo \mathcal{N}	0.7518 (0.0415)	(0.6705, 0.8331)
Modelo \mathcal{T}	0.9798 (0.0046)	(0.9708, 0.9888)

5.2.3. Diagnóstico

El conjunto $I_D = \{1, 30, 35, 79\}$ contiene los casos atípicos y extremos detectados en el análisis descriptivo y que posiblemente están influyendo sobre el estimador MV y el estimador del CCCG.

Por otro lado, la Figura 5.19 muestra la dispersión de los pesos \hat{u} versus la distancia de Mahalanobis $\hat{\delta}$. Los casos bajo el percentil del 5% de la distribución de los pesos son detectados como potencialmente influyentes, a este conjunto lo denotaremos por $I_u = \{1, 3, 30, 35, 59, 79\}$. La Tabla 5.17 resume los casos de-

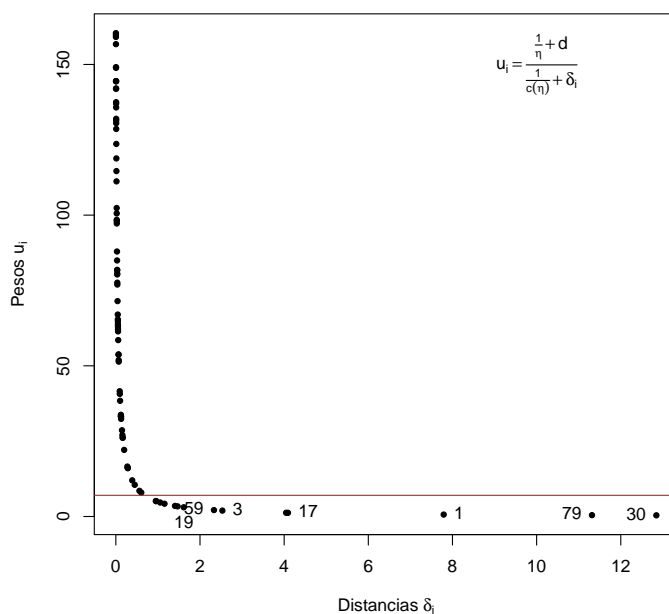


Figura 5.19: Observaciones detectadas como potencialmente influyentes a través de la distribución de los pesos versus la distancia de Mahalanobis.

tectados como potencialmente influyentes sobre los estimadores MV y sobre el estimador del CCCG, según la medida de influencia, esquema de perturbación y modelo estadístico propuesto.

El análisis de influencia local, bajo el esquema de ponderación de casos, detectó para el modelo \mathcal{N} al conjunto $I_1 = \{1, 30, 79\}$ como potencialmente influyente,

mientras que en el caso del modelo \mathcal{T} no se detectan casos potencialmente influyentes, pues los casos que fluctúan alrededor del punto de corte no presentan una influencia importante (ver Figura 5.20).

El método basado en la conformal curvatura normal, detecta como potencialmente influyente al conjunto $I_2 = \{3, 17, 30, 59, 79\}$ al asumir el modelo \mathcal{N} y no detecta casos potencialmente influyentes al asumir el modelo \mathcal{T} (ver Figura 5.21).

La Figura 5.22 muestra la dirección máxima \mathbf{h}_{max} versus la segunda dirección máxima \mathbf{h}_{seg} , para el modelo \mathcal{N} , gráfico (a) y el modelo \mathcal{T} , gráfico (b). El método de influencia detecta, para el modelo \mathcal{N} , los casos 1, 30 y 79 como potencialmente influyentes a través de \mathbf{h}_{max} y los casos 3, 17, 18, 19 y 59 a través de \mathbf{h}_{seg} , mientras que al asumir el modelo \mathcal{T} no se detectan casos con influencia importante.

El análisis de influencia local generalizada de primer orden (o máxima pendiente local) detectó, para el modelo \mathcal{N} , los casos 1, 17, 30 y 59, mientras que para el modelo \mathcal{T} no detectó casos con influencia importante (ver Figura 5.23).

Bajo el esquema de perturbación de la respuesta, el método de influencia local detecta, en la dirección en dónde se produce la mayor curvatura \mathbf{h}_{max} , los casos 34 y 53 como potencialmente influyentes al asumir el modelo \mathcal{N} , sin embargo la influencia de estos casos no es importante. En el modelo \mathcal{T} el método de influencia no detecta ningún caso potencialmente influyente (ver Figura 5.24).

Al considerar la dirección en dónde se produce la mayor conformal curvatura normal \mathbf{h}_{max} , el método de influencia detecta los casos 3, 17, 30, 59 y 79 como potencialmente influyentes, asumiendo el modelo \mathcal{N} y en el caso del modelo \mathcal{T} no se detecta ningún caso potencialmente influyente (ver Figura 5.25).

La Figura 5.26 muestra la dirección máxima \mathbf{h}_{max} versus la segunda dirección máxima \mathbf{h}_{seg} . Considerando el modelo \mathcal{N} se detectan los casos 3, 17, 18, 19 y 59

como potencialmente influyentes presentando una mayor influencia a través de \mathbf{h}_{seg} , mientras que los casos 34 y 53 se detectan como potencialmente influyentes presentando una mayor influencia a través de \mathbf{h}_{max} , asumiendo el modelo \mathcal{T} no se detectó ningún caso con alta influencia.

La Figura 5.27 muestra la dirección máxima \mathbf{h}_{max} donde se produce la mayor pendiente local. El método detecta los casos 1, 17, 30 y 79 como potencialmente influyentes directamente sobre el estimador del CCCG cuando se asume el modelo \mathcal{N} , mientras que al asumir el modelo \mathcal{T} no detecta casos con alta influencia.

Tabla 5.17: Resumen de casos potencialmente influyentes sobre los estimadores MV y el estimador del CCCG respecto al esquema de perturbación, medida de influencia y modelo estadístico.

Esquema	EMV		EMV para el CCC	
	Modelo \mathcal{N}	Modelo \mathcal{T}	Modelo \mathcal{N}	Modelo \mathcal{T}
Ponderación	Modelo \mathcal{N}	Modelo \mathcal{T}	Modelo \mathcal{N}	Modelo \mathcal{T}
$\mathbf{h}_{max} C$	1-30-79	No detecta	Sólo casos	Los casos
$\mathbf{h}_{max} \mathbf{B}$	3-17-30-59-79	No detecta	1-30-79	no son
$\mathbf{h}_{max}/\mathbf{h}_{seg}$	1-30-79 3-17-18-19-59	No detecta	son influyentes indirectamente	influyentes indirectamente
$\mathbf{h}_{max} \text{FI}$	**	**	1-17-30-59	No detecta
Respuesta	Modelo \mathcal{N}	Modelo \mathcal{T}	Modelo \mathcal{N}	Modelo \mathcal{T}
$\mathbf{h}_{max} C$	No detecta	No detecta	Sólo casos	Los casos
$\mathbf{h}_{max} \mathbf{B}$	3-17-30-59-79	No detecta	1-30-79	no son
$\mathbf{h}_{max}/\mathbf{h}_{seg}$	No detecta 3-17-18-19-59	No detecta	son influyentes indirectamente	influyentes indirectamente
$\mathbf{h}_{max} \text{FI}$	**	**	1-17-30-59	No detecta

Para evaluar la sensibilidad de los estimadores MV y el estimador del CCCG se consideran los conjuntos $I_1 = \{1, 30, 70\}$ y $I_2 = \{1, 17, 30, 59, 79\}$ como potencialmente influyentes.

La Tabla 5.18 muestra la razón de cambio (RC) para el estimador del CCCG y la comparación del ajuste al eliminar el conjunto I_1 . Notar que bajo el modelo \mathcal{N} la RC es mayor que la RC bajo el modelo \mathcal{T} , bajo este modelo se puede decir que el estimador del CCCG no se ve influenciado por el conjunto I_1 . Además la

$\log Lik(\mathcal{T})$ es siempre menor que la $\log Lik(\mathcal{N})$ al igual que el criterio de Akaike (AIC). Respecto al conjunto I_2 , éste no supera el 2% de RC, de modo que no se considera como potencialmente influyente sobre el estimador del CCCG, ya que los valores varían marginalmente.

Respecto a los estimadores MV, la Tabla 5.19 muestra la RC del modelo \mathcal{N} y el modelo \mathcal{T} . El estimador de las medias no presenta una RC importante pero la varianza generalizada estimada sí, tanto en el modelo \mathcal{N} como en el modelo \mathcal{T} . Notar que al eliminar el conjunto I_1 la variabilidad de los datos disminuye en ambos modelos, pero la RC para el estimador del parámetro de forma es nula indicando que el estimador es poco sensible al eliminar los casos 1, 30 y 79.

El ICA del 95% de confianza se presenta en la Tabla 5.20. El intervalo en base al modelo \mathcal{N} y en base al Bootstrap no paramétrico (Boot) cambian considerablemente al eliminar el conjunto I_1 y el intervalo en base al modelo \mathcal{T} no se ve afectado.

Feng *et al.* (015a) muestran que al eliminar el conjunto I_1 , considerando los tres métodos de registros, los intervalos de confianza Jackknife (J), Jackknife con Z-transformación de Fisher (JZ) y Bootstrap (Boot) cambian drásticamente no así el intervalo de credibilidad basado en la distribución t (ver Tabla 5.21).

Finalmente, el ajuste para el modelo \mathcal{N} sin el conjunto I_1 , con convergencia del algoritmo EM en 1, es

$$\hat{\mu}_{I_1} = \begin{pmatrix} 2.526 \\ 2.457 \\ 2.313 \end{pmatrix}, \quad \hat{\Sigma}_{I_1} = \begin{pmatrix} 0.761 & 0.708 & 0.769 \\ 0.708 & 0.781 & 0.794 \\ 0.769 & 0.794 & 0.981 \end{pmatrix}$$

y para el modelo \mathcal{T} , con convergencia del algoritmo EM en 549, es

$$\widehat{\mu}_{I_1} = \begin{pmatrix} 2.684 \\ 2.649 \\ 2.653 \end{pmatrix}, \quad \widehat{\Sigma}_{I_1} = \begin{pmatrix} 28.153 & 27.677 & 27.780 \\ 27.677 & 27.962 & 27.939 \\ 27.780 & 27.939 & 28.521 \end{pmatrix} \text{ y } \widehat{\eta}_{I_1} = 0.493.$$

Tabla 5.18: Razón de cambio para el estimador del CCCG y medidas de calidad del ajuste al eliminar casos potencialmente influyentes.

Casos eliminados	$\widehat{\rho}_g$				Comparación del ajuste			
	Modelo \mathcal{N}	RC	Modelo \mathcal{T}	RC	$\log Lik(\mathcal{N})$	$\log Lik(\mathcal{T})$	$AIC(\mathcal{N})$	$AIC(\mathcal{T})$
Ninguno	0.752		0.980		-226.350	-141.368	470.700	302.736
1	0.783	4%	0.982	0%	-217.601	-129.384	453.202	278.767
30	0.805	7%	0.982	0%	-210.849	-128.185	439.697	276.371
79	0.792	5%	0.982	0%	-210.480	-128.459	438.959	276.917
1-30	0.839	12%	0.984	0%	-198.404	-115.799	414.807	251.597
1-79	0.826	10%	0.984	0%	-199.027	-116.086	416.054	252.171
30-79	0.850	13%	0.984	0%	-188.905	-114.917	395.810	249.834
1-30-79	0.888	18%	0.985	1%	-168.410	-102.146	354.819	224.293

Tabla 5.19: Razón de cambio para los estimadores MV del modelo \mathcal{N} y \mathcal{T} al eliminar casos potencialmente influyentes.

Estimador	\mathcal{N}			\mathcal{T}		
	Ninguno	1-30-79	RC	Ninguno	1-30-79	RC
$\widehat{\mu}_1$	2.554	2.526	1%	2.660	2.684	1%
$\widehat{\mu}_2$	2.485	2.457	1%	2.625	2.649	1%
$\widehat{\mu}_3$	2.309	2.313	0%	2.622	2.653	1%
$ \widehat{\Sigma} $	0.050	0.014	72%	26.252	12.377	53%
$\widehat{\eta}$	*	*	*	0.493	0.492	0%

Tabla 5.20: Estimador puntual, error estándar de estimación e intervalo de confianza asintótico/confianza para el CCCG basado en distintos métodos.

Métodos	Todos los casos		Sin los casos 1-30-79	
	$\hat{\rho}_g$ (eee)	ICA/IC	$\hat{\rho}_g$ (eee)	ICA/IC
Modelo \mathcal{N}	0.752 (0.042)	(0.671, 0.833)	0.888 (0.020)	(0.849, 0.927)
Modelo \mathcal{T}	0.980 (0.005)	(0.971, 0.989)	0.985 (0.003)	(0.979, 0.992)
Boot	0.752 (0.076)	(0.606, 0.905)	0.888 (0.030)	(0.830, 0.948)

Tabla 5.21: Estimador puntual e intervalo de credibilidad/confianza para el CCCG basado en distintos métodos.

Métodos	Todos los casos		Sin los casos 1-30-79	
	$\hat{\rho}_g$	IC	$\hat{\rho}_g$	IC
Bayes	0.914	(0.870, 0.946)	0.931	(0.898, 0.958)
J	0.752	(0.591, 0.914)	0.890	(0.831, 0.949)
JZ	0.740	(0.502, 0.873)	0.886	(0.804, 0.935)
Boot	0.752	(0.602, 0.900)	0.888	(0.816, 0.936)

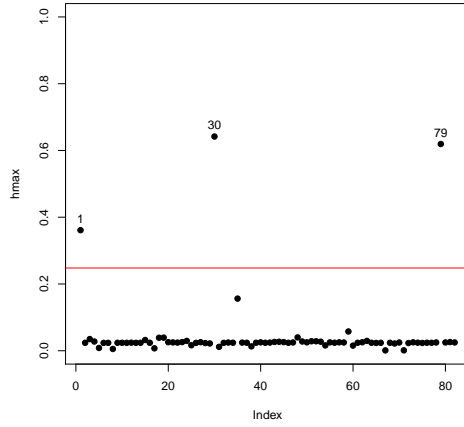
Fuente: Feng *et al.* (015a)

5.2.4. Conclusión

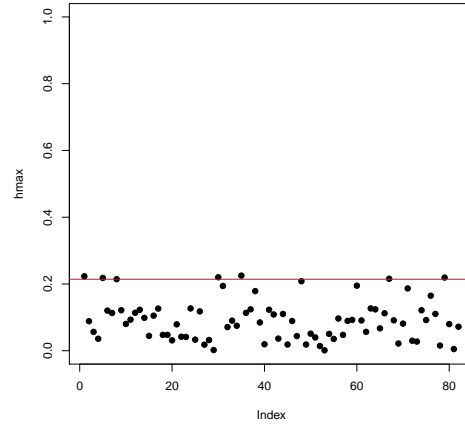
En la aplicación 1 se estudió el acuerdo entre las medidas generadas por el método M y A. Al incorporar un tercer método P, no sólo es de interés evaluar el acuerdo entre pares de medidas sino que también evaluar el acuerdo global entre las medidas.

El método de diagnóstico de influencia local detectó varios casos potencialmente influyentes sobre los estimadores MV y sobre el estimador del CCCG. Sin embargo, al realizar el análisis de sensibilidad sólo los casos 1, 30 y 79 ejercen una influencia importante sobre la matriz de varianza-covarianza estimada en ambos modelos, y sólo bajo el modelo \mathcal{N} estos casos influyen sobre el estimador del CCCG y del ICA.

Por lo tanto, la distribución t genera un procedimiento de estimación más robusto respecto que la distribución normal.

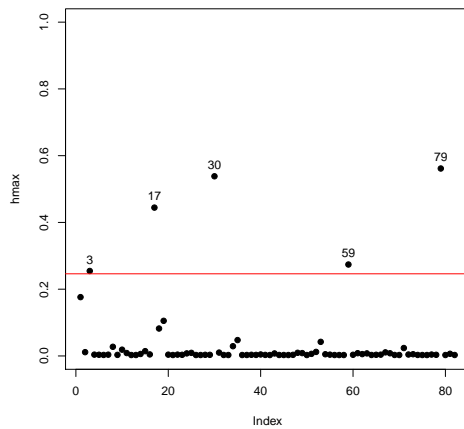


(a) Distribución Normal

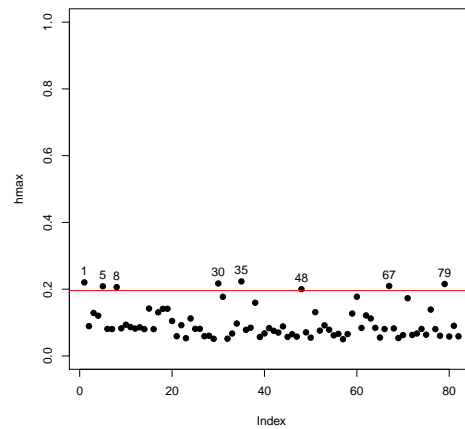


(b) Distribución t

Figura 5.20: Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.



(a) Distribución Normal



(b) Distribución t

Figura 5.21: Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.

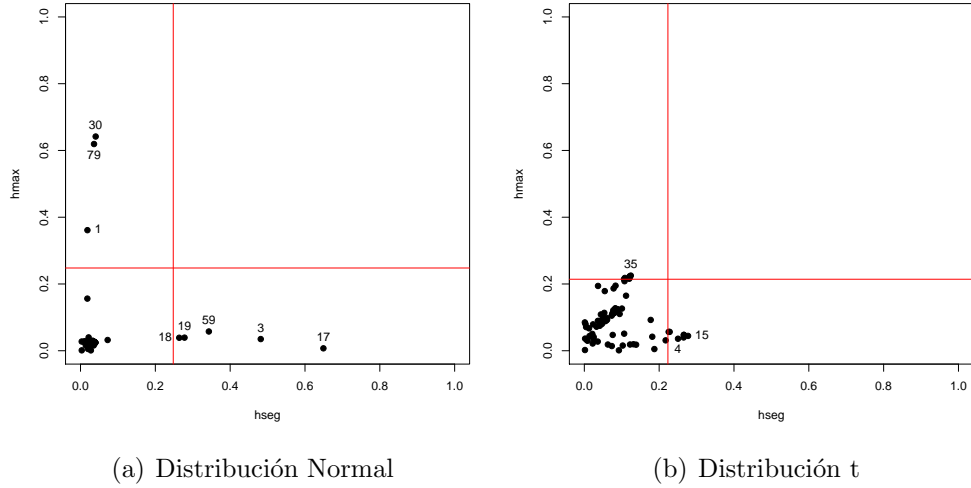


Figura 5.22: Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.

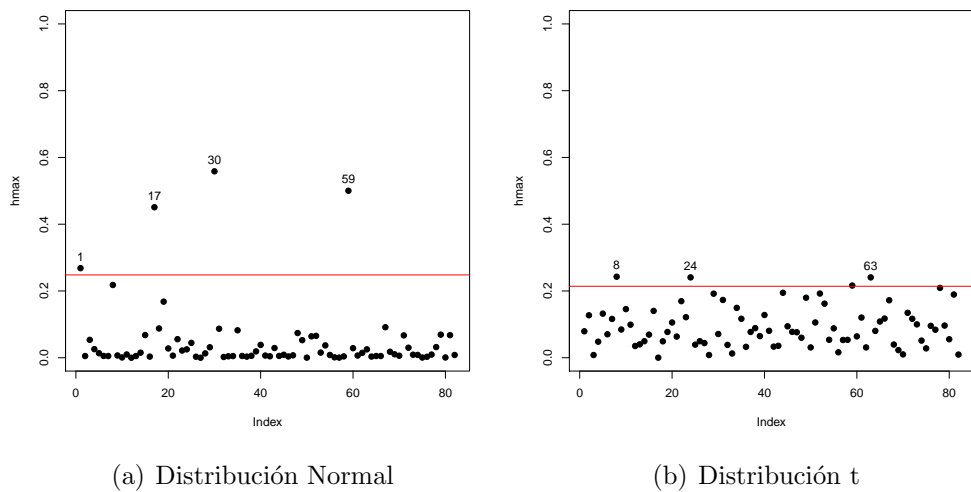
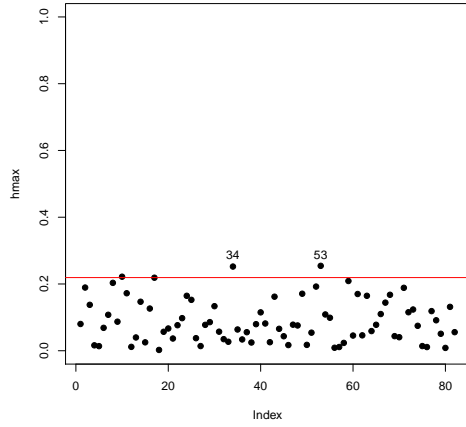
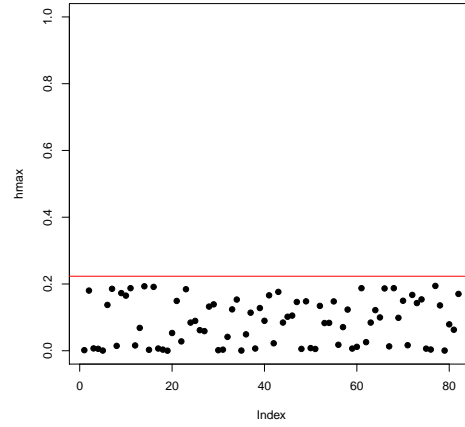


Figura 5.23: Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de ponderación de casos.

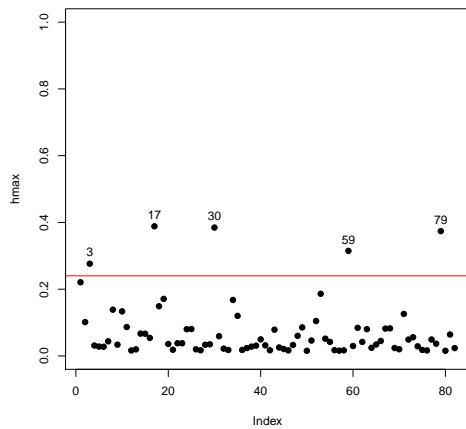


(a) Distribución Normal

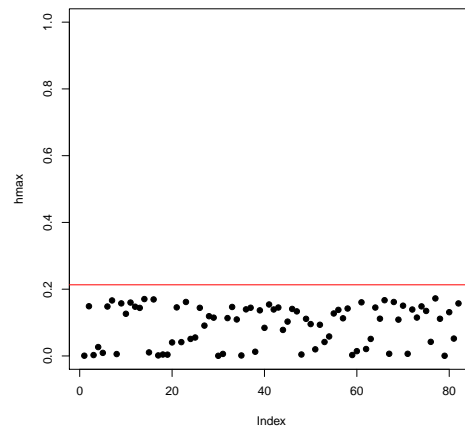


(b) Distribución t

Figura 5.24: Gráfico de influencia de \mathbf{h}_{max} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.

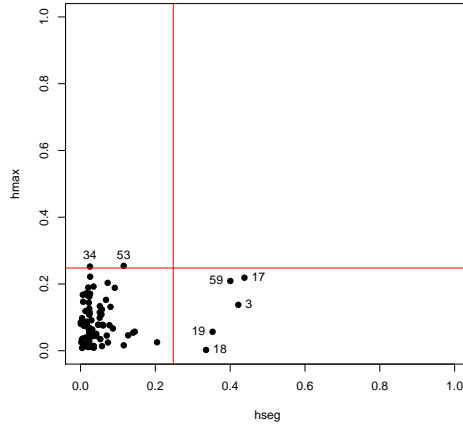


(a) Distribución Normal

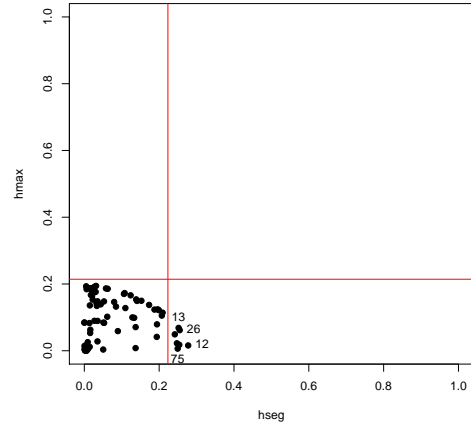


(b) Distribución t

Figura 5.25: Gráfico de influencia de \mathbf{h}_{max} para la conformal curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.

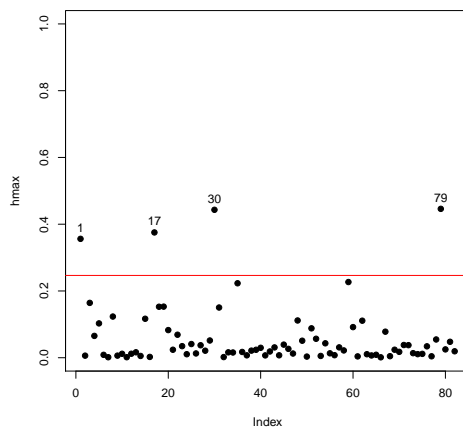


(a) Distribución Normal

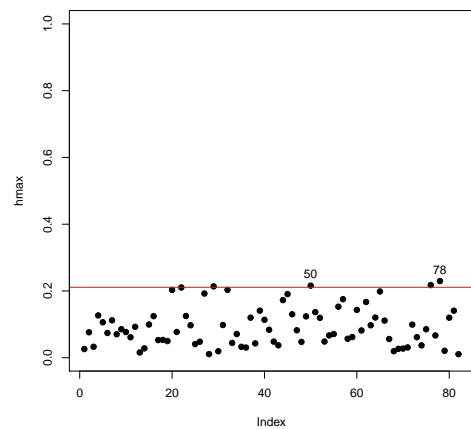


(b) Distribución t

Figura 5.26: Gráfico de influencia de \mathbf{h}_{max} versus \mathbf{h}_{seg} para la curvatura normal asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.



(a) Distribución Normal



(b) Distribución t

Figura 5.27: Gráfico de influencia de \mathbf{h}_{max} para la máxima pendiente local asumiendo que los datos siguen una distribución normal (a) y una distribución t (b) multivariada, bajo el esquema de perturbación de la respuesta.

Conclusión general

Como se ha expuesto a lo largo de este trabajo es de interés obtener estimaciones más confiables del coeficiente de correlación de concordancia. La distribución t ha permitido obtener estimas menos sensibles a observaciones atípicas y/o extremas respecto a las estimas obtenidas bajo normalidad, más aún, el intervalo de confianza asintótico es más preciso al ser comparado con el intervalo de confianza asintótico basado en el supuesto de normalidad.

Por otro lado, se ha ilustrado que, a través del estudio de simulación, la tasa de cobertura de los intervalos de confianza asintóticos propuestos son similares al valor nominal. Además, el estadístico de prueba asintótico basado en la de razón de verosimilitud tiene un desempeño adecuado, ya que la capacidad de la prueba para detectar diferencias significativas, converge rápidamente a una potencia empírica alta cuando la hipótesis alternativa se aleja poco de la hipótesis nula. En cuanto al desempeño de la prueba de Wald éste es relativo, pues la capacidad de la prueba para detectar diferencia significativa depende del valor que se desea docimar, el nivel de significación y el tamaño de la muestra.

El método de diagnóstico propuesto por Cook (1986) es ampliamente usado para detectar observaciones, que tras pequeñas perturbaciones en el modelo y/o los datos, son potencialmente influyentes sobre los estimadores MV. Se ha ilustrado que el método es eficiente en detectar estas observaciones cuando se asume

que los datos siguen una distribución normal, sin embargo, el método no detecta casos potencialmente influyentes, asumiendo una distribución t , lo cual permite confirmar que dicha distribución es más adecuada para obtener estimaciones más confiables del CCC.

Ahora bien, como se ha discutido, el interés del diagnóstico no es sólo sobre los estimadores MV, sino que también sobre el estimador del CCC. Se ha presentado que la generalización del método propuesto por Cook (1986), basado en medidas de influencia local generalizada y las medidas de primer y segundo orden, también han sido eficaces para detectar observaciones potencialmente influyentes sobre el estimador del CCC al asumir normalidad y al comparar los resultados del diagnóstico, al asumir que los datos siguen una distribución t , el método no detecta casos potencialmente influyentes sobre el estimador del CCC, lo cual confirma que la distribución t genera inferencia robusta para el CCC.

Finalmente, a modo de conclusión evaluar el acuerdo entre las medidas generadas por varios instrumentos (o métodos) no es suficiente para sustituir un método por otro sólo asumiendo normalidad de los datos. Si los datos, como se ha visto, presentan observaciones atípicas o extremas es adecuado considerar para el proceso de inferencia del CCC la distribución t pues su estimador es poco sensible a pequeñas perturbaciones sobre el modelo y/o los datos.

Bibliografía

- Amari, S. (1985). Differential-Geometrical Methods in Statistics. Lecture notes in statist. *Springer*, 28.
- Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley, Hoboken, New Jersey.
- Arellano, R. (1994). *Distribuições Elípticas: Propriedades, inferência e aplicações a modelos de regressão*. Tesis doctoral, Instituto de matemáticas e estatística. Universidade de São Paulo.
- Arellano, R. y Bolfarine, H. (1995). On some characterizations of the t-distribution. *Statistics and Probability Letters*, 25:79–85.
- Banhart, H. y Williamson, J. (2001). Modelling concordance correlation via gee to evaluate reproducibility. *Biometrics*, 57(3):931–940.
- Barnhart, H., Haber, M., y Lin, L. (2007). An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics*, 17:529–569.
- Barnhart, H., Haber, M., y Song, J. (2002). Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics*, 58:1020–1027.
- Bastiani, F. D., Aquino, A. D., y Uribe, M. (2014). Influence diagnostics in elliptical spatial linear models. *TEST*, 23:322–340.

- Bland, J. y Altman, D. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8:135–160.
- Bolfarine, H. y Galea, M. (1996). On structural comparative calibration under a t -model. *Computational Statistics*, 11:63–85.
- Cadigan, N. y Farrel, P. (2002). Generalized local influence with applications to fish stock cohort analysis. *Appl. Statist.*, 51:469–483.
- Carrasco, J. y Jover, L. (2003). Estimating the generalized concordance correlation coefficient through variance components. *Biometrics*, 59(4):849–858.
- Carrasco, J., Jover, L., King, T., y Chinchilli, V. (2007). Comparison of concordance correlation coefficient estimating approaches with skewed data. *Journal of Biopharmaceutical Statistics*, 17:673–684.
- Choudhary, P. y Nagaraja, H. (2015). *Measuring Agreement: Methodology and Applications*. Wiley Series in Probability and Statistics, New York.
- Cook, R. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society B*, 48:133–169.
- Cornish, E. (1954). The multivariate t distribution associated with a set of normal standard deviates. *Australian Journal of Physics*, 7:531–542.
- Dempster, A., Laird, N., y Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm. *J. R. Statist. Soc. B*, 39:1–37.
- Diaz, J., Galea, M., y Leiva, V. (2003). Influence diagnostics for elliptical multivariate linear regression models. *Communications in Statistics - Theory and Methods*, 32(3):625–641.
- Do Carmo, P. (1976). *Differential Geometry of Curves and Surfaces*. Prentice Hall, New Jersey.
- Dunnett, C. y Sobel, M. (1954). A bivariate generalization of student's t -distribution with tables for certain special cases. *Biometrika*, 41:153–169.

- Feng, D., Baumgartner, R., y Svetnik, V. (2015a). A robust bayesian estimate of the concordance correlation coefficient. *Journal of Biopharmaceutical Statistics*, 25:490–507.
- Feng, D., Baumgartner, R., y Svetnik, V. (2015b). A bayesian estimate of the concordance correlation coefficient with skewed data. *Pharmaceutical Statistics*, 14:350–358.
- Feng, D., Svetnik, V., Coimbra, A., y Baumgartner, R. (2014). A comparison of confidence interval methods for the concordance correlation coefficient and intraclass correlation coefficient with small number raters. *Journal of Biopharmaceutical Statistics*, 24:272–293.
- Fung, W. K. y Kwan, C. (1997). A note on local influence based on normal curvature. *J. R. Statist. Soc. B*, 59:839–843.
- Galea, G., Riquelme, M., y Paula, G. (2000). Diagnostic methods in elliptical linear regression models. *Brazilian Journal of Probability and Statistics*, 14:167–184.
- Galea, M. (1995). *Calibração comparativa estrutural e funcional*. Tesis doctoral, Instituto de matemáticas e estatística. Universidade de São Paulo.
- Galea, M., Bolfarine, H., y de Castro, M. (2002). Local influence in comparative calibration models. *Biometrical Journal*, 44:59–81.
- Galea, M., Bolfarine, H., y Vilca, F. (2005). Local influence in comparative calibration models under elliptical-distributions. *Biometrical Journal*, 47(5):691–706.
- Galea, M., Paula, G., y Bolfarine, H. (1997). Local influence in elliptical linear regression model. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 46:71–79.
- Graybill, F. (1983). *Matrices with applications in statistics*. Wadsworth International Group, Belmont.

- Henze, N. y Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Commun. Statist. Theor. Meth.*, 19:3595–3617.
- Hogg, R. y Craig, A. (1978). *Introduction to Mathematical of Statistics*. Macmillan, New York.
- Johnson, N., Kotz, S., y Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Wiley, New York.
- Kim, M. (1996). Local influence in multivariate normal data. *Journal of Applied Statistics*, 23:535–541.
- King, T. y Chinchilli, V. (2001). Robust estimators of the concordance correlation coefficient. *Journal of Biopharmaceutical Statistics*, 11:83–105.
- Lange, K., Little, R., y Taylor, J. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, 84:881–896.
- Lawrance, A. (1988). Regression transformation diagnostics using local influence. *Journal of the American Statistical Association*, 83:1067–1072.
- Lehmann, E. (1999). *Elements of Large Sample Theory*. Springer, New York.
- Lehmann, E. y Casella, G. (1998). *Theory of Point Estimation*. Springer, New York.
- Lesaffre, E. y Verbeke, G. (1998). Local influence in linear mixed models. *International Biometric Society*, 54:570–582.
- Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45:225–268.
- Lin, L., A.S., H., y Wu, W. (2004). *Statistical Tools for Measuring Agreement*. Springer, New York.
- Little, R. y Rubin, D. (1980). *Statistical Analysis with Missing Data*. Wiley, New York.

- Magnus, J. y Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. John Wiley and Sons, New York.
- Mardia, K., Kent, J., y Bibby, J. (1997). *Multivariate Analysis*. ACADEMIC PRESS, San Diego.
- McBride, G. B. (2005). A proposal for strength-of-agreement criteria for lins concordance correlation coefficient. *NIWA Client Report: HAM 2005-062. Report to Ministry of Health*.
- McLachlan, G. y Peel, D. (1999). The emmix algorithm for the fitting of normal and t-components. *Journal of Statistical Software*, 4(2):1–14.
- McLachlan, G. y Peel, D. (2000). *Finite Mixture Models*. Wiley, New York.
- Nadarajah, S. y Kotz, S. (2004). *Multivariate t Distributions and Their Applications*. Cambridge, New York.
- Osorio, F. y Galea, M. (2016). Statistical inference in multivariate analysis using the t distribution. *Unpublished manuscript*.
- Osorio, F., Galea, M., y Paula, G. (2007). Assessment of local influence in elliptical linear models with longitudinal structure. *Computational Statistics and Data Analysis*, 51:4354–4368.
- Paula, G. (1993). Assessing local influence in restricted regression models. *Computational Statistics and Data Analysis*, 16:63–79.
- Poon, W. y Poon, Y. (1999). Conformal normal curvature and assessment of local influence. *J. R. Statist. Soc.B*, 61:51–61.
- Poon, Y. y Poon, W. (2010). *Application of Elementary Differential Geometry to Influence Analysis*. International Press, Beijing.
- Resnick, S. (1999). *A Probability Path*. Birkhäuser, Boston.
- Rubin, D. (1983). Iteratively reweighted least squares. *In Encyclopedia of Statistical Sciences*, 4:272–275.

- Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New Jersey.
- Sutradhar, B. (1993). Score test for the covariance matrix of elliptical t - distribution. *Journal of Multivariate Analysis*, 46:1–12.
- Svetnik, V., Ma, J., Soper, K., Doran, S., Renger, J., Deacon, S., y Koblan, K. (2007). Evaluation of automated and semi-automated scoring of polysomnographic recordings from a clinical trial using zolpidem in the treatment of insomnia. *SLEEP*, 30(11):1562–1574.
- Thomas, W. y Cook, R. (1990). Assessing influence on predictions from generalized linear model. *Technometrics*, 32:59–65.
- Tong, Y. (1990). *The Multivariate Normal Distribution*. Springer, New York.
- Wu, X. y Luo, Z. (1993). Second order approach to local influence. *J. R. Statist. Soc. B*, 55:929–936.
- Zhu, H., Ibrahim, J., Lee, S., y Zhang, H. (2007). Perturbation selection and influence measures in local influence analysis. *The Annals of Statistics*, 35:2565–2588.

Distribuciones y derivación de algunos resultados

A.1. Distribuciones de probabilidad utilizadas

En este apéndice se presentan las funciones de distribución de probabilidad (fdp) utilizadas para derivar la fdp de la distribución t multivariada reparametrizada. Las definiciones se adoptan de Mardia *et al.* (1997), Anderson (2003), Nadarajah y Kotz (2004) y Osorio y Galea (2016).

A.1.1. Distribución Normal

Una variable aleatoria d -dimensional \mathbf{Y} con fdp normal puede ser representada por su media $\boldsymbol{\mu}$ y matriz de varianza-covarianza $\boldsymbol{\Sigma}$ definida positiva. Su densidad está dada por

$$p(\mathbf{y}) = (2\pi)^{-d/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\} \text{ con } \mathbf{y} = (y_1, \dots, y_d)^T \in \mathbb{R}^d.$$

Notación: $\mathbf{Y} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

A.1.2. Distribución Gamma

Una variable aleatoria X admite una función de distribución de probabilidad Gamma con parámetro de forma $r > 0$ y parámetro de tasa $\lambda > 0$, su densidad está dada por

$$p(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-x\lambda} \text{ con } x > 0$$

donde $\Gamma(r) = \int_0^{+\infty} t^{r-1} e^{-t} dt$ es la función gamma.

Notación: $X \sim \Gamma(r, \lambda)$.

Observación: Usando el teorema de transformación de variables aleatorias continuas, se tiene que si $X \sim \Gamma(r, \lambda)$ entonces $aX \sim \Gamma(r, \lambda/a)$ con a constante.

A.1.3. Distribución Chi-cuadrado

La distribución chi-cuadrado puede ser vista como un caso particular de la distribución Gamma. Si $r = \nu/2$ y $\lambda = 1/2$ se tiene que su densidad está dada por

$$p(u) = \frac{(1/2)^{\nu/2}}{\Gamma(\nu/2)} u^{\nu/2-1} e^{-u/2} \text{ con } u > 0.$$

Notación: $U \sim \chi^2(\nu)$.

Observación: Como $X \sim \Gamma(\nu/2, 1/2)$ es equivalente a decir $X \sim \chi^2(\nu)$ y como $X/\nu \sim \Gamma(\nu/2, \nu/2)$ se tiene que $X/\nu \sim \chi^2(\nu)$.

A.1.4. Distribución t multivariada

Dada $V \sim \Gamma(1/2\eta, 1/2c(\eta))$ independiente de $\mathbf{Y} \sim N_d(\mathbf{0}, \Sigma)$ de modo que la fdp conjunta es

$$\begin{aligned} p(\mathbf{y}, v) &= p(\mathbf{y})p(v) \\ &= (2\pi)^{-d/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{y}^T \Sigma^{-1} \mathbf{y} \right\} \times \frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} v^{(1/2\eta-1)} \exp \left\{ -\frac{v}{2c(\eta)} \right\}. \end{aligned}$$

donde $c(\eta) = \eta/(1 - 2\eta)$ con $0 < \eta < 1/2$.

El objetivo es encontrar la función de distribución de probabilidad de \mathbf{X} aplicando el teorema de transformación de variables aleatorias continuas.

Sea $U = V$ y $\mathbf{Y} = (\mathbf{X} - \boldsymbol{\mu}) \sqrt{U}$ luego se tiene que la fdp conjunta entre \mathbf{X} y U es $p(\mathbf{x}, u) = p_{\mathbf{Y}}((\mathbf{x} - \boldsymbol{\mu})\sqrt{u})p_V(u)|J|$ donde

$$|J| = \begin{vmatrix} \partial \mathbf{y} / \partial \mathbf{x}^T & \partial \mathbf{y} / \partial u \\ \partial w / \partial \mathbf{x}^T & \partial w / \partial u^T \end{vmatrix} = u^{p/2}.$$

Por lo tanto, la fdp conjunta para \mathbf{X} y U , está dada por

$$\begin{aligned} p(\mathbf{x}, u) &= (2\pi)^{-d/2} |\Sigma/u|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T [\Sigma/u]^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\ &\times \frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} u^{(1/2\eta)-1} \exp \left\{ -\frac{u}{2c(\eta)} \right\}. \end{aligned}$$

Para obtener la fdp de \mathbf{X} basta marginalizar $p(\mathbf{x}, u)$ con respecto a u , es decir

$$\begin{aligned}
p(\mathbf{x}) &= \int_0^{+\infty} p(\mathbf{x}, u) du \\
&= \left(\frac{c(\eta)}{\pi} \right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)} \\
&\quad \times \int_0^{+\infty} \frac{(\delta^2/2 + 1/2c(\eta))^{1/2\eta+d/2}}{\Gamma(1/2\eta + d/2)} u^{1/2\eta+d/2-1} \exp \{ -u(\delta^2/2 + 1/2c(\eta)) \} du \\
&= \left(\frac{c(\eta)}{\pi} \right)^{d/2} \frac{\Gamma(1/2\eta + d/2)}{\Gamma(1/2\eta)} |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)} \\
&= K_d(\eta) |\boldsymbol{\Sigma}|^{-1/2} (1 + c(\eta)\delta^2)^{-(1/2\eta+d/2)}
\end{aligned}$$

con $\delta^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$ la distancia de Mahalanobis.

Por lo tanto, el vector aleatorio \mathbf{X} sigue una fdp $p(\mathbf{x})$ llamada distribución t multivariada reparametrizada.

Notación: $\mathbf{X} \sim \mathcal{T}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$.

A.2. Algoritmo EM

Se define $\mathbf{X}_c = (\mathbf{X}_{obs}^T, U_1, \dots, U_n)^T$ el vector aleatorio completo, donde $\mathbf{X}_{obs} = (\mathbf{X}_1^T, \dots, \mathbf{X}_n^T)$ es el vector observado y U_i para todo $i = 1, \dots, n$ variables aleatorias no observadas o perdidas.

Para cada $i = 1, \dots, n$ se tiene que $\mathbf{X}_i | U_i = u_i \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}/u_i)$ y $U_i \sim \Gamma(1/2\eta, 1/2c(\eta))$ de modo que la fdp conjunta es $p(\mathbf{x}_i, u_i) = p(\mathbf{x}_i | u_i) p(u_i)$. Su función de verosimilitud completa se define por

$$\begin{aligned}
L_c(\boldsymbol{\theta}) &= (2\pi)^{-nd/2} |\boldsymbol{\Sigma}|^{-n/2} \left(\prod_{i=1}^n u_i^{d/2} \right) \exp \left\{ -\frac{1}{2} \sum_{i=1}^n u_i \delta_i^2 \right\} \\
&\quad \times \left[\frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} \right]^n \left(\prod_{i=1}^n u_i^{(1/2\eta)-1} \right) \exp \left\{ -\sum_{i=1}^n \left(\frac{1}{2c(\eta)} \right) u_i \right\}
\end{aligned}$$

con $\delta_i^2 = (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$ y el logaritmo está dado por

$$\begin{aligned} \ell_c(\boldsymbol{\theta}) &= -\frac{nd}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n u_i \delta_i^2 + n \log \left[\frac{((1/2c(\eta))^{1/2\eta})}{\Gamma(1/2\eta)} \right] \\ &\quad + \left(\frac{1}{2\eta} - 1 + \frac{d}{2} \right) \sum_{i=1}^n \log(u_i) - \frac{1}{2c(\eta)} \sum_{i=1}^n u_i \\ &= -\frac{nd}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n u_i \delta_i^2 + \frac{n}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) - n \log (\Gamma(1/2\eta)) \\ &\quad + \left(\frac{1}{2\eta} - 1 + \frac{d}{2} \right) \sum_{i=1}^n \log(u_i) - \frac{1}{2c(\eta)} \sum_{i=1}^n u_i \end{aligned}$$

Paso E

Este paso requiere el calculo de $E(\ell_c(\boldsymbol{\theta}) | \mathbf{x}_{obs}, \boldsymbol{\theta}^{(k)})$, entonces se tiene

$$\begin{aligned} E(\ell_c(\boldsymbol{\theta}) | \mathbf{x}_{obs}, \boldsymbol{\theta}^{(k)}) &= -\frac{nd}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n E(U_i | \mathbf{x}_i, \boldsymbol{\theta}^{(k)}) \delta_i^2 \\ &\quad + \frac{n}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) - n \log (\Gamma(1/2\eta)) + \left(\frac{1}{2\eta} + \frac{d}{2} - 1 \right) \sum_{i=1}^n E(\log(U_i) | \mathbf{x}_i, \boldsymbol{\theta}^{(k)}) \\ &\quad - \frac{1}{2c(\eta)} \sum_{i=1}^n E(U_i | \mathbf{x}_i, \boldsymbol{\theta}^{(k)}). \end{aligned}$$

Para obtener $E(U_i | \mathbf{x}_i, \boldsymbol{\theta}^{(k)})$ y $E(\log(U_i) | \mathbf{x}_i, \boldsymbol{\theta}^{(k)})$ es necesario encontrar $p(u_i | \mathbf{x}_i)$.

Por el teorema de Bayes se tiene,

$$\begin{aligned} p(u_i | \mathbf{x}_i) &= \frac{p(u_i) p(\mathbf{x}_i | u_i)}{\int_0^{+\infty} p(u_i) p(\mathbf{x}_i | u_i) du_i} = \frac{p(u_i) p(\mathbf{x}_i | u_i)}{p(\mathbf{x}_i)} \\ &= \frac{\frac{(1+2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} u_i^{(1/2\eta)-1} \exp \{-u_i(1/2\eta)\} \times (2\pi)^{-d/2} |\boldsymbol{\Sigma}/u_i|^{-1/2} \exp \{(-1/2)u_i \delta_i^2\}}{K_d(\eta) |\boldsymbol{\Sigma}|^{-1/2} (1+c(\eta)\delta_i^2)^{(-1/2)(1/\eta+d)}} \\ &= \left(\frac{1+c(\eta)\delta_i^2}{2c(\eta)} \right)^{(1/2\eta+d/2)} \frac{u_i^{(1/2\eta+d/2-1)}}{\Gamma(1/2\eta+d/2)} \exp \left\{ -u_i \left(\frac{1+c(\eta)\delta_i^2}{2c(\eta)} \right) \right\}. \end{aligned}$$

Por lo tanto, $U_i | \mathbf{X}_i = \mathbf{x}_i \sim \Gamma(r, \lambda)$ con $r = (1/2\eta + p/2)$ y $\lambda = (1+c(\eta)\delta_i^2)/2c(\eta)$, con esperanza dada por

$$E(U_i | \mathbf{x}_i, \boldsymbol{\theta}^{(k)}) = r/\lambda = \frac{(1/\eta^{(k)} + d)}{1/c(\eta^{(k)}) + \delta_i^2(k)} = u_i^{(k)}.$$

Por otro lado, si $R \sim \Gamma(r, \lambda)$ entonces $E(\log(R)) = \psi(r) - \log(\lambda)$ con $\psi(r) = (\partial\Gamma(r)/\partial r)\Gamma(r)$, lo cual permite calcular

$$\begin{aligned}
E(\log(U_i)|\mathbf{x}_i, \boldsymbol{\theta}^{(k)}) &= \psi(r) - \log(\lambda) \\
&= \psi(1/2\eta^{(k)} + d/2) - \log((1 + c(\eta^{(k)})\delta_i^2(k))/2c(\eta^{(k)})) \\
&= \psi(1/2\eta^{(k)} + d/2) - \log((1/2)(1/c(\eta^{(k)}) + \delta_i^2(k))) \\
&\quad - \log((1\eta^{(k)} + d)/(1\eta^{(k)} + d)) \\
&= \psi(1/2\eta^{(k)} + d/2) - \log(1/2\eta^{(k)} + d/2) + \log(u_i^{(k)}).
\end{aligned}$$

Finalmente $E(\ell_c(\boldsymbol{\theta})|\mathbf{x}, \boldsymbol{\theta}^{(k)}) = Q((\boldsymbol{\mu}, \boldsymbol{\phi}), \boldsymbol{\theta}^{(k)}) + Q(\eta, \boldsymbol{\theta}^{(k)})$ donde

$$\begin{aligned}
Q((\boldsymbol{\mu}, \boldsymbol{\phi}), \boldsymbol{\theta}^{(k)}) &= -\frac{nd}{2} \log(2\pi) - \frac{n}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n u_i^{(k)} \delta_i^2 \\
Q(\eta, \boldsymbol{\theta}^{(k)}) &= n \log \left[\frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} \right] - \frac{1}{2c(\eta)} \sum_{i=1}^n u_i^{(k)} \\
&\quad + \left(\frac{1}{2\eta} - 1 + \frac{d}{2} \right) \sum_{i=1}^n \left\{ \psi(1/2\eta^{(k)} + d/2) - \log(1/2\eta^{(k)} + d/2) + \log(u_i^{(k)}) \right\}.
\end{aligned}$$

Paso M

Este paso consiste en actualizar $\boldsymbol{\theta}^{(k+1)}$ maximizando $Q((\boldsymbol{\mu}, \boldsymbol{\phi}), \boldsymbol{\theta}^{(k)})$ y $Q(\eta, \boldsymbol{\theta}^{(k)})$ para obtener las actualizaciones de $(\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\Sigma}^{(k+1)})$ y $\eta^{(k+1)}$ por separado.

Para estimar $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ se considera la maximización de $Q((\boldsymbol{\mu}, \boldsymbol{\phi}), \boldsymbol{\theta}^{(k)})$ conociendo $u_i^{(k)}$ desde el paso anterior para todo $i = 1, \dots, n$.

McLachlan y Peel (1999) indican que $Q((\boldsymbol{\mu}, \boldsymbol{\phi}), \boldsymbol{\theta}^{(k)})$ es el logaritmo de la función de verosimilitud de n observaciones independientes con media común $\boldsymbol{\mu}$ y matriz de varianza-covarianza $\boldsymbol{\Sigma}$ y es equivalente a calcular la media muestral y matriz de varianza-covarianza muestral ponderada por los $u_1^{(k)}, \dots, u_n^{(k)}$.

Entonces los estimadores MV iterativos de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ están dados por

$$\begin{aligned}\boldsymbol{\mu}^{(k+1)} &= \frac{\sum_{i=1}^n u_i^{(k)} \mathbf{x}_i}{\sum_{i=1}^n u_i^{(k)}} \\ \boldsymbol{\Sigma}^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n u_i^{(k)} (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})(\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})^T.\end{aligned}$$

Luego, como en Osorio y Galea (2016), se actualiza $\eta^{(k+1)}$ como la solución de la ecuación $\partial Q(\eta, \boldsymbol{\theta}^{(k)})/\partial \eta = 0$, donde

$$\begin{aligned}\frac{\partial Q(\eta, \boldsymbol{\theta}^{(k)})}{\partial \eta} &= \frac{n}{2\eta^2} \left\{ \psi\left(\frac{1}{2\eta}\right) - \log\left(\frac{1}{2c(\eta)}\right) + \frac{c(\eta)}{\eta} - \psi\left(\frac{1}{2\eta^{(k)}} + \frac{d}{2}\right) \right. \\ &\quad \left. + \log(1/2\eta^{(k)} + d/2) - \frac{1}{n} \sum_{i=1}^n \left(\log(u_i^{(k)}) - u_i^{(k)} \right) \right\}.\end{aligned}$$

A.3. Vector score y matriz de información de Fisher

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios tal que $\mathbf{X}_i \sim \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$ para todo $i = 1, \dots, n$.

Sea $g : \mathbb{R} \rightarrow [0, \infty)$ una función de modo que la fdp de \mathbf{X} puede escribirse como $p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g(\delta_i^2)$ donde $g(\delta_i^2) = K_d(\eta)(1 + c(\eta)\delta_i^2)^{-(1/2\eta+d/2)}$ y $\delta_i^2 = (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$.

Sea $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \eta)^T \in \mathbb{R}^{d+q+1}$ el vector de parámetros con $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ donde $d = \dim(\boldsymbol{\mu})$, $q = \dim(\boldsymbol{\phi})$ y $1 = \dim(\eta)$ y $\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta})$ con $\ell_i(\boldsymbol{\theta}) = \log \left\{ |\boldsymbol{\Sigma}|^{-1/2} g(\delta_i^2) \right\}$ es el logaritmo de la función de verosimilitud.

A.3.1. Derivación del vector score

La primera derivada de $\ell(\boldsymbol{\theta})$ es el vector score definido como $\mathbf{U}(\boldsymbol{\theta}) = \sum_{i=1}^n \mathbf{U}_i(\boldsymbol{\theta})$ donde $\mathbf{U}_i(\boldsymbol{\theta}) = (\mathbf{U}_i(\boldsymbol{\mu})^T, \mathbf{U}_i(\boldsymbol{\phi})^T, U_i(\eta))^T$ con elementos dados por

$$\begin{aligned}
U_i(\boldsymbol{\mu}) &= \frac{\partial \log(g(\delta_i^2))}{\partial \boldsymbol{\mu}} \\
&= -\frac{1}{2} \frac{(1/\eta + d)}{(1 + c(\eta)\delta_i^2)} c(\eta) \left\{ \frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} \right\} \\
U_i(\boldsymbol{\phi}) &= -\frac{n}{2} \frac{\partial \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi}} + \frac{\partial \log(g(\delta_i^2))}{\partial \boldsymbol{\phi}} \\
&= -\frac{n}{2} \frac{\partial \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi}} - \frac{1}{2} \frac{(1/\eta + d)}{(1 + c(\eta)\delta_i^2)} c(\eta) \left\{ \frac{\partial \delta_i^2}{\partial \boldsymbol{\phi}} \right\} \\
U_i(\eta) &= \frac{\partial \log(g(\delta_i^2))}{\partial \eta} \\
&= \frac{\partial \log(K_d(\eta))}{\partial \eta} - \frac{\partial}{\partial \eta} \left\{ \frac{1}{2} (1/\eta + d) \log(1 + c(\eta)\delta_i^2) \right\}.
\end{aligned}$$

Las siguientes derivadas son de interés para obtener el vector score;

$$\begin{aligned}
\frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} &= -2\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}), \quad \frac{\partial \delta_i^2}{\partial \boldsymbol{\phi}} = \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}) \\
\frac{\partial \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi}} &= \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}), \quad \frac{\partial \log K_d(\eta)}{\partial \eta} = \frac{1}{2\eta^2} \{c(\eta)d - \psi(1/2\eta + d/2) + \psi(1/2\eta)\} \\
\text{y } \frac{\partial}{\partial \eta} \left\{ \frac{1}{2} (1/\eta + d) \log(1 + c(\eta)\delta_i^2) \right\} &= \frac{1}{2\eta^2} \{u_i c(\eta)\delta_i^2 - \log(1 + c(\eta)\delta_i^2)\}.
\end{aligned}$$

Por lo tanto, los elementos del vector score para el i -ésimo caso son

$$\begin{aligned}
U_i(\boldsymbol{\mu}) &= u_i \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}) \\
U_i(\boldsymbol{\phi}) &= \frac{1}{2} \mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}) \\
U_i(\eta) &= \frac{1}{2\eta^2} \{c(\eta)(d - u_i \delta_i^2) - \psi(1/2\eta + d/2) + \psi(1/2\eta) + \log(1 + c(\eta)\delta_i^2)\}
\end{aligned}$$

donde $\psi(z) = d \log(\Gamma(z))/dz$, \mathbf{D}_d la matriz de duplicación de dimensión $d^2 \times (d/2)(d+1)$ definida en Magnus y Neudecker (1988) y el operador $\text{vec}(\cdot)$ permite vectorizar una matriz apilando sus columnas. Notar que $q = (d/2)(d+1)$.

A.3.2. Derivación de la matriz de información observada

La segunda derivada de $\ell(\boldsymbol{\theta})$ evaluada en $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ es

$$\mathbf{L} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \eta} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\mu}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \eta} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta \partial \boldsymbol{\mu}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta \partial \boldsymbol{\phi}^T} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta^2} \end{pmatrix} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

luego la matriz de información observada queda $\mathbf{I}_o(\boldsymbol{\theta}) = -\mathbf{L}$.

Considerando las siguientes derivadas

$$\begin{aligned} \frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} &= -2(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}, \quad \frac{\partial u_i}{\partial \boldsymbol{\mu}} = -\frac{(1/\eta + d)}{1/c(\eta) + \delta_i^2} \frac{\partial \delta_i^2}{\partial \boldsymbol{\mu}} \\ \frac{\partial \delta_i^2}{\partial \boldsymbol{\phi}} &= -\mathbf{D}_d^T \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}) \\ \frac{\partial u_i}{\partial \boldsymbol{\phi}} &= -\frac{u_i}{(1/c(\eta) + \delta_i^2)} \frac{\partial \delta_i^2}{\partial \boldsymbol{\phi}} \end{aligned}$$

y evaluando en los estimadores MV, se obtienen los elementos de \mathbf{L} , dados por

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= -\frac{n}{1 - 2\hat{\eta}} \hat{\boldsymbol{\Sigma}}^{-1} + 2\hat{\boldsymbol{\Sigma}}^{-1} \left[\sum_{i=1}^n \frac{\hat{u}_i (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} \right] \hat{\boldsymbol{\Sigma}}^{-1} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= n \mathbf{D}_d^T (\hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1}) \mathbf{D}_d + \frac{1}{2} \mathbf{D}_d^T \left[\sum_{i=1}^n \frac{\hat{u}_i}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} (\mathbf{M} \mathbf{M}^T \otimes \mathbf{M} \mathbf{M}^T) \right] \mathbf{D}_d \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= \sum_{i=1}^n \frac{2(1/\hat{\eta} + d)}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} (\mathbf{M}^T \otimes \mathbf{M} \mathbf{M}^T) \mathbf{D}_d \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\mu} \partial \eta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= \sum_{i=1}^n \frac{1}{\hat{\eta}^2} \left(\frac{(\hat{u}_i - 1)}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} \right) \mathbf{M} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \eta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= \sum_{i=1}^n \frac{1}{2\hat{\eta}^2} \frac{(\hat{u}_i)}{(1/c(\hat{\eta}) + \hat{\delta}_i^2)} \mathbf{D}_d^T \text{vec}(\mathbf{M} \mathbf{M}^T) \end{aligned}$$

con $\mathbf{M} = \hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})$.

Por otro lado, considerando

$$U_i(\eta) = a(\eta) - \frac{\tau_i}{2\eta^2} \text{ donde } a(\eta) = \frac{1}{2\eta^2}(dc(\eta) - \beta(\eta)), \beta(\eta) = \psi(1/2\eta + d/2) - \psi(1/2\eta)$$

$$\text{ y } \tau_i = c(\eta)u_i\delta_i^2 - \log(1 + c(\eta)\delta_i^2)$$

se obtiene

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta^2} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &= \sum_{i=1}^n \frac{\partial U_i(\eta)}{\partial \eta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \sum_{i=1}^n \frac{\partial}{\partial \eta} \left[a(\eta) - \frac{\tau_i}{2\eta^2} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &= \sum_{i=1}^n \left[\frac{da(\eta)}{d\eta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} + \frac{\hat{\tau}_i}{\hat{\eta}^2} - \frac{1}{2\hat{\eta}^2} \frac{d\tau_i}{d\eta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \right] \end{aligned}$$

donde

$$\begin{aligned} \frac{da(\eta)}{d\eta} &= \frac{1}{2\eta^2} \left(d \frac{dc(\eta)}{d\eta} - \frac{d\beta(\eta)}{d\eta} \right) - \frac{1}{\eta^3} (c(\eta) - \beta(\eta)), \quad \frac{dc(\eta)}{d\eta} = \frac{1}{(1 - 2\eta^2)} \text{ y} \\ \frac{d\beta(\eta)}{d\eta} &= -\frac{1}{2\eta^2} \left[\psi' \left(\frac{1}{2\eta} + \frac{d}{2} \right) - \psi' \left(\frac{1}{2\eta} \right) \right] \text{ con } \psi'(z) = \frac{d^2 \log(\Gamma(z))}{dz^2}. \end{aligned}$$

A.3.3. Derivación de la matriz de información de Fisher

El primer momento poblacional del vector score es $E[\mathbf{U}(\boldsymbol{\theta})] = \mathbf{0}$ y el segundo momento del vector score es la matriz de información de Fisher y es denotada por $\mathcal{I}(\boldsymbol{\theta}) = \sum_{i=1}^n E[\mathbf{U}_i(\boldsymbol{\theta})\mathbf{U}_i^T(\boldsymbol{\theta})]$.

La representación matricial es

$$\mathcal{I}(\boldsymbol{\theta}) = \sum_{i=1}^n \begin{pmatrix} E[\mathbf{U}_i(\boldsymbol{\mu})\mathbf{U}_i^T(\boldsymbol{\mu})] & E[\mathbf{U}_i(\boldsymbol{\mu})\mathbf{U}_i^T(\boldsymbol{\phi})] & E[\mathbf{U}_i(\boldsymbol{\mu})\mathbf{U}_i^T(\eta)] \\ E[\mathbf{U}_i(\boldsymbol{\phi})\mathbf{U}_i^T(\boldsymbol{\mu})] & E[\mathbf{U}_i(\boldsymbol{\phi})\mathbf{U}_i^T(\boldsymbol{\phi})] & E[\mathbf{U}_i(\boldsymbol{\phi})\mathbf{U}_i^T(\eta)] \\ E[\mathbf{U}_i(\eta)\mathbf{U}_i^T(\boldsymbol{\mu})] & E[\mathbf{U}_i(\eta)\mathbf{U}_i^T(\boldsymbol{\phi})] & E[\mathbf{U}_i(\eta)\mathbf{U}_i^T(\eta)] \end{pmatrix}$$

Para calcular los elementos de $\mathbf{I}(\boldsymbol{\theta})$ se utilizan los Lemas dados en Sutradhar (1993), Lange *et al.* (1989), Bolfarine y Galea (1996) y Osorio y Galea (2016), y resultados matriciales dados en Graybill (1983) y Magnus y Neudecker (1988).

Lema A.3.1. Si $\mathbf{Z} \sim \mathcal{T}_d(\mathbf{0}, \boldsymbol{\Sigma}, \eta)$ y $W = 1 + c(\eta)\mathbf{Z}^T\boldsymbol{\Sigma}^{-1}\mathbf{Z}$, entonces

- (i) $E(W^{-1}) = (1 + \eta d)^{-1}$
- (ii) $E(W^{-1}\mathbf{Z}) = \mathbf{0}$
- (iii) $E(W^{-1}\mathbf{Z}\mathbf{Z}^T) = [(1 - 2\eta)/(1 + \eta d)]\boldsymbol{\Sigma}$
- (iv) $E(W^{-2}) = (1 - 2\eta)/[(1 + \eta d)(1 + \eta(d + 2))]$
- (v) $E(W^{-2}\mathbf{Z}) = \mathbf{0}$
- (vi) $E(W^{-2}\mathbf{Z}\mathbf{Z}^T) = \{(1 - 2\eta)/[(1 + \eta d)(1 + \eta(d + 2))]\}\boldsymbol{\Sigma}$
- (vii) $E(\log W) = \psi(1/2\eta + d/2) - \psi(1/2\eta)$

donde $\psi(\cdot)$ es la derivada de la función $\Gamma(\cdot)$.

Lema A.3.2. Si $\mathbf{Z} \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$, entonces

$$E(\mathbf{Z}\mathbf{Z}^T \otimes \mathbf{Z}\mathbf{Z}^T) = 2(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\mathbf{N}_d + \text{vec}(\boldsymbol{\Sigma})\text{vec}^T(\boldsymbol{\Sigma})$$

donde $\mathbf{N}_d = (\mathbf{I}_{d^2} + \mathbf{K}_d)/2$ con \mathbf{K}_d la matriz de conmutación de orden d .

Lema A.3.3. Si $\mathbf{Z} \sim \mathcal{T}_d(\mathbf{0}, \boldsymbol{\Sigma}, \eta)$ y $W = 1 + c(\eta)\mathbf{Z}^T\boldsymbol{\Sigma}^{-1}\mathbf{Z}$, entonces

$$E(W^{-2}\mathbf{Z}\mathbf{Z}^T \otimes \mathbf{Z}\mathbf{Z}^T) = \frac{(1 - 2\eta)^2}{(1 + \eta d)(1 + \eta(d + 2))} \{2(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\mathbf{N}_d + \text{vec}(\boldsymbol{\Sigma})\text{vec}^T(\boldsymbol{\Sigma})\}$$

pues $1 + c(\eta)\mathbf{Z}^T\boldsymbol{\Sigma}^{-1}\mathbf{Z} \stackrel{\mathcal{D}}{=} 1 + d\eta F(d, 1/\eta)$, donde $\stackrel{\mathcal{D}}{=}$ indica igual en distribución.

Teorema A.3.1. Sean $\mathbf{A}(n \times q)$, $\mathbf{B}(q \times s)$ y $\mathbf{C}(n \times s)$ matrices. Entonces

$$\text{vec}(\mathbf{A}\mathbf{B}\mathbf{C}^T) = (\mathbf{A} \otimes \mathbf{C})\text{vec}(\mathbf{B}).$$

Teorema A.3.2. Sean $\mathbf{A}(n \times n)$ y $\mathbf{B}(n \times n)$ matrices, a, b escalar y sean $\mathbf{x}(n \times 1)$ y $\mathbf{y}(n \times 1)$ vectores, entonces

1. $\text{vec}(a\mathbf{A} + b\mathbf{B}) = a\text{vec}(\mathbf{A}) + b\text{vec}(\mathbf{B})$
2. $\text{vec}(\mathbf{x}\mathbf{y}^T) = (\mathbf{x} \otimes \mathbf{y})$

Teorema A.3.3. Si $\mathbf{A}(m_1 \times n_1)$, $\mathbf{B}(m_2 \times n_2)$, $\mathbf{F}(n_1 \times k_1)$ y $\mathbf{G}(n_2 \times k_2)$ matrices, entonces

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{F} \otimes \mathbf{G}) = \mathbf{AF} \otimes \mathbf{BG}.$$

Elementos de la matriz de Fisher

Sea el vector aleatorio $\mathbf{Z}_i = (\mathbf{X}_i - \boldsymbol{\mu})$ para todo $i = 1, \dots, n$ con dimensión $(d \times 1)$ y sea $\mathbf{I}_{\boldsymbol{\mu}\boldsymbol{\mu}}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} [U_i(\boldsymbol{\mu})U_i^T(\boldsymbol{\mu})]$ luego aplicando el Lema (A.3.1) parte (vi) se obtiene

$$\begin{aligned} E [U_i(\boldsymbol{\mu})U_i^T(\boldsymbol{\mu})] &= (1/\eta + d)^2 c(\eta)^2 \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{\theta}} [W_i^{-2} \mathbf{Z}_i \mathbf{Z}_i^T] \boldsymbol{\Sigma}^{-1} \\ &= (1/\eta + d)^2 c(\eta)^2 \left\{ \frac{(1 - 2\eta)}{[(1 + \eta d)(1 + \eta(d + 2))]} \right\} \boldsymbol{\Sigma}^{-1} \\ &= c_{\boldsymbol{\mu}}(\eta) \boldsymbol{\Sigma}^{-1} \end{aligned}$$

con $c_{\boldsymbol{\mu}}(\eta) = c_{\phi}(\eta)/(1 - 2\eta)$ y $c_{\phi}(\eta) = (1 + \eta d)/(1 + \eta(d + 2))$.

Sea $\mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) = E [U_i(\phi)U_i^T(\phi)]$, aplicando los Teoremas A.3.1, A.3.2, A.3.3 y los Lemas A.3.3, A.3.1 parte (iii) se obtiene

$$\begin{aligned} E [U_i(\phi)U_i^T(\phi)] &= \frac{1}{4} \mathbf{D}_d^T \{ E[\text{vec}(u_i \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1}) \text{vec}^T(u_i \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1})] \\ &\quad - E[\text{vec}(u_i \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1}) \text{vec}^T(\boldsymbol{\Sigma}^{-1})] - E[\text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^T(u_i \boldsymbol{\Sigma}^{-1} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1})] \\ &\quad + E[\text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^T(\boldsymbol{\Sigma}^{-1})] \} \mathbf{D}_d \\ &= \frac{1}{4} \mathbf{D}_d^T \{ 2c_{\phi}(\eta)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) + (c_{\phi}(\eta) - 1) \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^T(\boldsymbol{\Sigma}^{-1}) \mathbf{N}_d \} \mathbf{D}_d \end{aligned}$$

con $c_{\phi}(\eta) = (1 + \eta d)/(1 + \eta(d + 2))$.

Sea $\mathbf{I}_{\eta\eta}(\boldsymbol{\theta}) = E(U_i^2(\eta))$ con $U_i(\eta) = a(\eta) - \tau_i/2\eta^2$ donde

$$\begin{aligned} a(\eta) &= \frac{1}{2\eta^2} \{ c(\eta)d - \beta(\eta) \}, \quad \beta(\eta) = \left\{ \psi \left(\frac{1}{2\eta} + \frac{d}{2} \right) - \psi \left(\frac{1}{2\eta} \right) \right\} \\ \text{y } \tau_i &= \frac{1}{2\eta^2} \{ c(\eta)u_i \delta_i^2 - \log(1 + c(\eta)\delta_i^2) \}. \end{aligned}$$

Como $E(U_i^2(\eta)) = -E(\partial U_i(\eta)/\partial \eta)$ se tiene

$$\begin{aligned} E \left[\frac{\partial U_i(\eta)}{\partial \eta} \right] &= \frac{da(\eta)}{d\eta} - \frac{1}{2\eta^2} \left[E \left(\frac{\partial \tau_i}{\partial \eta} \right) - \frac{2}{\eta} E(\tau_i) \right] \\ &= \frac{1}{2\eta^2} \left\{ \frac{d}{(1-2\eta)^2} \left(\frac{1+\eta d(1-4\eta) - 8\eta^2}{(1+\eta d)(1+(d+2)\eta)} \right) - \frac{d\beta(\eta)}{d\eta} \right\} \end{aligned}$$

donde

$$\begin{aligned} \frac{da(\eta)}{d\eta} &= \frac{1}{2\eta^2} \left\{ \frac{d}{(1-2\eta)^2} - \frac{d\beta(\eta)}{d\eta} \right\} - \frac{1}{\eta^3} (dc(\eta) - \beta(\eta)) \\ \frac{d\beta(\eta)}{d\eta} &= -\frac{1}{2\eta^2} \left\{ \psi' \left(\frac{1}{2\eta} + \frac{d}{2} \right) - \psi' \left(\frac{1}{2\eta} \right) \right\} \quad \text{y} \quad \psi'(\cdot) = \frac{\partial^2 \Gamma(\cdot)}{\partial \eta^2}. \end{aligned}$$

Sea $\mathbf{I}_{\mu\phi}(\boldsymbol{\theta}) = E[U_i(\boldsymbol{\mu})U_i^T(\boldsymbol{\phi})] = -E[\partial U_i^T(\boldsymbol{\phi})/\partial \boldsymbol{\mu}]$, aplicando el Lema A.3.1 parte (ii) se tiene

$$\begin{aligned} E \left[\frac{\partial U_i^T(\boldsymbol{\phi})}{\partial \boldsymbol{\mu}} \right] &= E \left[\frac{\partial}{\partial \boldsymbol{\mu}} \left\{ \frac{1}{2} \mathbf{D}_d^T \text{vec} (u_i \boldsymbol{\Sigma}^{-1} \mathbf{Z}_i \mathbf{Z}_i^T \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}) \right\}^T \right] \\ &= \frac{1}{2} E \left[u_i \frac{\partial}{\partial \boldsymbol{\mu}} \left\{ \text{vec}^T (\mathbf{Z}_i \mathbf{Z}_i^T) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_d \right\} \right] \\ &+ \frac{1}{2} E \left[\left\{ \frac{\partial u_i}{\partial \boldsymbol{\mu}} \right\} \text{vec}^T (\mathbf{Z}_i \mathbf{Z}_i^T) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_d \right] \\ &= \frac{1}{2} (1/\eta + d) c(\eta) E \left[W_i^{-1} \frac{\partial}{\partial \boldsymbol{\mu}} \left\{ \text{vec}^T (\mathbf{Z}_i \mathbf{Z}_i^T) \right\} \right] (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_d \\ &+ (1/\eta + d) c(\eta)^2 \boldsymbol{\Sigma}^{-1} E [W_i^{-2} \mathbf{Z}_i \text{vec}^T (\mathbf{Z}_i \mathbf{Z}_i^T)] (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_d \\ &= \mathbf{0}_{d \times q} \end{aligned}$$

pues la primera esperanza de la expresión anterior es una matriz nula de dimensión $(d \times d^2)$ y al ser multiplicada por $(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_d$ se tiene $\mathbf{0}_{d \times q}$ y la esperanza del lado izquierdo de la expresión anterior también es una matriz nula de dimensión $(d \times q)$.

Sea $\mathbf{I}_{\mu\eta}(\boldsymbol{\theta}) = E[U_i(\boldsymbol{\mu})U_i(\eta)] = -E[\partial U_i(\boldsymbol{\mu})/\partial \eta]$. Por el Lema A.3.2 parte (ii) y (v) se tiene

$$E \left[\frac{\partial U_i(\boldsymbol{\mu})}{\partial \eta} \right] = \frac{1}{\eta(1-2\eta)^2} \left\{ (1+\eta d) \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{\theta}} [W_i^{-2} \mathbf{Z}_i] - (1-2\eta) \boldsymbol{\Sigma}^{-1} E [W_i^{-1} \mathbf{Z}_i] \right\} = \mathbf{0}_{d \times 1}$$

Sea $\mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) = E[\mathbf{U}_i(\phi)U_i(\eta)] = -E[\partial\mathbf{U}_i(\phi)/\partial\eta]$, aplicando el Lema A.3.2 parte (iii) y (vi) se tiene

$$\begin{aligned} E\left[\frac{\partial\mathbf{U}_i(\phi)}{\partial\eta}\right] &= \frac{1}{2}\mathbf{D}_d^T(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \\ &\quad \times \text{vec}\left(\frac{(1+\eta d)}{\eta(1-2\eta)^2}E[W_i^{-2}\mathbf{Z}_i\mathbf{Z}_i^T] - \frac{(1-2\eta)}{\eta(1-2\eta)^2}E[W_i^{-2}\mathbf{Z}_i\mathbf{Z}_i^T]\right) \\ &= \frac{c(\eta)(d+2)}{(1+\eta(d+2))(1+\eta d)}\mathbf{D}_d^T\text{vec}(\boldsymbol{\Sigma}^{-1}). \end{aligned}$$

Finalmente la matriz de información de Fisher esperada es

$$\mathcal{I}(\boldsymbol{\theta}) = n\mathbf{I}(\boldsymbol{\theta}) = n \begin{pmatrix} \mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) & \mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) & \mathbf{I}_{\eta\eta}(\boldsymbol{\theta}) \end{pmatrix}$$

y sus elementos son sub-matrices dadas por

$$\begin{aligned} \mathbf{I}_{\mu\mu}(\boldsymbol{\theta}) &= c_\mu(\eta)\boldsymbol{\Sigma}^{-1} \\ \mathbf{I}_{\phi\phi}(\boldsymbol{\theta}) &= \frac{1}{4}\mathbf{D}_d^T\{2c_\phi(\eta)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\mathbf{N}_d + (c_\phi(\eta) - 1)\text{vec}(\boldsymbol{\Sigma}^{-1})\text{vec}^T(\boldsymbol{\Sigma}^{-1})\}\mathbf{D}_d \\ \mathbf{I}_{\phi\eta}(\boldsymbol{\theta}) &= -\frac{c(\eta)(d+2)}{(1+\eta(d+2))(1+\eta d)}\mathbf{D}_d^T\text{vec}(\boldsymbol{\Sigma}^{-1}) \\ \mathbf{I}_{\eta\eta}(\boldsymbol{\theta}) &= -\frac{1}{2\eta^2}\left\{\frac{d}{(1-2\eta)^2}\left[\frac{1+\eta d(1-4\eta)-8\eta^2}{(1+\eta d)(1+(d+2)\eta)}\right] - \frac{d\beta(\eta)}{d\eta}\right\} \end{aligned}$$

donde $u_i = (1/\eta + d)/(1/c(\eta) + \delta_i^2)$, $c_\mu(\eta) = c_\phi(\eta)/(1-2\eta)$ y $c_\phi(\eta) = (1+2\eta)/(1+(d+2)\eta)$.

A.4. Derivada del CCC

Sea $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^{d+q+1}$ y sea $\boldsymbol{\rho}(\boldsymbol{\theta}) : \mathbb{R}^{d+q+1} \rightarrow \mathbb{R}$ una función de $\boldsymbol{\theta}$ que representa el coeficiente de correlación de concordancia bivariado y global.

A.4.1. Derivada para CCC bivariado

Sea $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22}, \eta)^T \in \Theta \subseteq \mathbb{R}^6$ y sea

$$\rho(\boldsymbol{\theta}) = \frac{2\sigma_{12}}{\sigma_{11} + \sigma_{22} + (\mu_1 - \mu_2)^2}$$

el coeficiente de correlación bivariado. La derivada de $\rho(\boldsymbol{\theta})$ respecto a $\boldsymbol{\theta}$ es

$$\mathbf{d} = \left(\frac{\partial \rho(\boldsymbol{\theta})}{\partial \mu_1}, \frac{\partial \rho(\boldsymbol{\theta})}{\partial \mu_2}, \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{11}}, \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{12}}, \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{22}} \right)^T$$

donde

$$\begin{aligned} \frac{\partial \rho(\boldsymbol{\theta})}{\partial \mu_1} &= \frac{(\mu_1 - \mu_2)}{\sigma_{12}} \rho_c^2, \quad \frac{\partial \rho(\boldsymbol{\theta})}{\partial \mu_2} = -\frac{(\mu_1 - \mu_2)}{\sigma_{12}} \rho_c^2, \quad \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{11}} = -\frac{\rho_c^2}{2\sigma_{11}}, \quad \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{22}} = -\frac{\rho_c^2}{2\sigma_{22}} \\ \text{y } \frac{\partial \rho(\boldsymbol{\theta})}{\partial \sigma_{12}} &= \frac{\rho_c}{\sigma_{12}}. \end{aligned}$$

A.4.2. Derivada para el CCC global

Sea $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \eta)^T \in \Theta \subseteq \mathbb{R}^{d+q+1}$ con $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ que se puede re ordenar como $(\boldsymbol{\phi}_1^T, \boldsymbol{\phi}_2^T)^T$ de modo que $\boldsymbol{\phi}_1$ es un vector con $m = P_2^d/2$ elementos no nulos de $\boldsymbol{\Sigma}_T$ y $\boldsymbol{\phi}_1$ contiene d elementos no nulos de $\boldsymbol{\Sigma}_D$ tal que $q = m + d$.

El coeficiente de correlación de concordancia global es

$$\rho(\boldsymbol{\theta}) = \frac{2\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T)}{(d-1)\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_D) + (\mathbf{C}\boldsymbol{\mu})^T (\mathbf{C}\boldsymbol{\mu})}$$

donde $\boldsymbol{\Sigma}_T$ la matriz triangular inferior sin la diagonal de la matriz de varianzas-covarianzas y $\boldsymbol{\Sigma}_D = \text{Diag}(\sigma_{11}, \dots, \sigma_{dd})$ la matriz diagonal, es decir

$$\boldsymbol{\Sigma}_T = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \sigma_{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1d} & \sigma_{2d} & \cdots & 0 \end{pmatrix} \quad \text{y} \quad \boldsymbol{\Sigma}_D = \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & \vdots \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_{dd} \end{pmatrix}$$

y \mathbf{C} es una matriz de dimensión $(m \times d)$ que genera las diferencias del tipo

$(\mu_j - \mu_k)$ con $\binom{d}{2}$ comparaciones posibles, dada por

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & -1 \end{pmatrix}$$

de modo que $\boldsymbol{\mu}^T \mathbf{C}^T \mathbf{C} \boldsymbol{\mu}$ genera las diferencias al cuadrado.

Por otro lado, notar que $\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T) = \mathbf{1}_m^T \boldsymbol{\phi}_1$ y $\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_D) = \mathbf{1}_d^T \boldsymbol{\phi}_2$.

Luego la derivada de $\boldsymbol{\rho}(\boldsymbol{\theta})$ respecto a $\boldsymbol{\theta}$ es

$$\mathbf{d} = \left(\frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\partial \boldsymbol{\mu}^T}, \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\partial \boldsymbol{\phi}^T}, \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\partial \eta} \right)^T$$

con

$$\begin{aligned} \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\partial \boldsymbol{\mu}} &= -\rho_g^2 \frac{\mathbf{C} \boldsymbol{\mu}}{\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T)} \quad \text{de dimensión } (d \times 1) \\ \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\boldsymbol{\phi}_1} &= \rho_g \frac{\mathbf{1}_m}{\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T)} \quad \text{de dimensión } (m \times 1) \\ \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\boldsymbol{\phi}_2} &= -\rho_g^2 \frac{(d-1)}{2} \frac{\mathbf{1}_d}{\mathbf{1}_q^T \text{vech}(\boldsymbol{\Sigma}_T)} \quad \text{de dimensión } (d \times 1) \\ \frac{\partial \boldsymbol{\rho}(\boldsymbol{\theta})}{\eta} &= 0. \end{aligned}$$

donde $\rho_g = \boldsymbol{\rho}(\boldsymbol{\theta})$.

Notar que $\partial \boldsymbol{\rho}(\boldsymbol{\theta}) / \partial \boldsymbol{\phi}^T$ es de dimensión $(m+d) \times 1$ donde $m+d=q$ por lo que \mathbf{d} tiene dimensión $(d+q+1) \times 1$.

A.5. Derivación del Test de Razón de Verosimilitud

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra de vectores aleatorios iid con fdp $\mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. El Test de Razón de Verosimilitud (TRV) para probar la hipótesis estadística $H_0 : \boldsymbol{\rho}(\boldsymbol{\theta}) = 0$ o equivalentemente $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ matriz diagonal, está dado por

$$\Lambda = \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta})} = \frac{L(\tilde{\boldsymbol{\theta}})}{L(\hat{\boldsymbol{\theta}})}$$

donde $\hat{\boldsymbol{\theta}}$ es el estimador MV bajo todo el espacio paramétrico $\Theta \subseteq \mathbb{R}^{d+q+1}$, $\tilde{\boldsymbol{\theta}}$ es el estimador MV bajo el espacio paramétrico restringido $\Theta_0 \subseteq \mathbb{R}^{2d+1}$, es decir bajo la hipótesis nula.

La región crítica está dada por $RC = \{\underline{\mathbf{x}} : \Lambda(\underline{\mathbf{x}}) \leq \lambda_0\}$ de modo que λ_0 es una constante tal que el estadístico sea de tamaño α y $\mathbf{x} \in \mathbb{R}^{n \times d}$.

A.5.1. Estimador MV bajo todo el espacio paramétrico

Los estimadores MV bajo todo el espacio paramétrico $\Theta \subseteq \mathbb{R}^{d+q+1}$ son entregados por el algoritmo EM implementado por Osorio y Galea (2016) en R-project llamado librería MVT y descrito en el Apéndice A.2.

A.5.2. Estimador MV bajo la hipótesis nula

Sin pérdida de generalidad es posible encontrar los estimadores MV bajo H_0 considerando la función de verosimilitud completa bajo el espacio paramétrico restringido Θ_0 , la cuál está dada por

$$\begin{aligned} L_c(\boldsymbol{\theta}_0) &= (2\pi)^{-nd/2} |\boldsymbol{\Sigma}_0|^{-n/2} \left(\prod_{i=1}^n u_i^{d/2} \right) \exp \left\{ -\frac{1}{2} \sum_{i=1}^n u_i \delta_{i0}^2 \right\} \\ &\times \left[\frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} \right]^n \left(\prod_{i=1}^n u_i^{(1/2\eta)-1} \right) \exp \left\{ -\sum_{i=1}^n u_i \left(\frac{1}{2c(\eta)} \right) \right\} \end{aligned}$$

con $\boldsymbol{\theta}_0 = (\boldsymbol{\mu}^T, \boldsymbol{\phi}_0^T, \eta)^T$ donde $\boldsymbol{\phi}_0 = \text{vech}(\boldsymbol{\Sigma}_0)$ y $\delta_{i0}^2 = (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$.

Como $\ell_c(\boldsymbol{\theta}_0) = \log L_c(\boldsymbol{\theta}_0)$ al aplicar el paso E del algoritmo EM al logaritmo de la función de verosimilitud se obtiene

$$\begin{aligned} E(\ell_c(\boldsymbol{\theta}_0) | \mathbf{x}_i, \boldsymbol{\theta}_0^{(k)}) &= -\frac{nd}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}_0| + n \log \left[\frac{(1/2c(\eta))^{1/2\eta}}{\Gamma(1/2\eta)} \right] \\ &+ \left(\frac{1}{2\eta} + \frac{d}{2} - 1 \right) \sum_{i=1}^n \left\{ \psi \left(\frac{1}{2\eta^{(k)}} + \frac{d}{2} \right) - \log \left(\frac{1}{2\eta^{(k)}} + \frac{d}{2} \right) + \log(u_i^{(k)}) \right\} \\ &- \frac{1}{2} \sum_{i=1}^n u_i^{(k)} \delta_{i0}^2 - \frac{1}{2c(\eta^{(k)})} \end{aligned}$$

donde

$$u_i^{(k)} = E(U_i | \mathbf{x}_i, \boldsymbol{\theta}_0^{(k)}) = \frac{1/\eta^{(k)} + d}{1/c(\eta^{(k)}) + \delta_{i0}^2(k)}$$

con $\delta_{i0}^2(k) = (\mathbf{x}_i - \boldsymbol{\mu}^{(k)})^T [\boldsymbol{\Sigma}_0^{(k)}]^{-1} (\mathbf{x}_i - \boldsymbol{\mu}^{(k)})$. Luego al aplicar el paso M del algoritmo EM se obtienen los estimadores MV bajo la hipótesis nula, dados por

$$\begin{aligned} \boldsymbol{\mu}^{(k+1)} &= \frac{\sum_{i=1}^n u_i^{(k)} \mathbf{x}_i}{\sum_{i=1}^n u_i^{(k)}} \quad \text{y} \quad \boldsymbol{\Sigma}_0^{(k+1)} = \text{Diag}(\mathbf{S}^{(k+1)}) \quad \text{donde} \\ \mathbf{S}^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n u_i^{(k)} (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)}) (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})^T \end{aligned}$$

y $\eta^{(k+1)}$ es actualizada por $\partial Q(\eta, \boldsymbol{\theta}_0^{(k)}) / \partial \eta$.

Los estimadores MV bajo la hipótesis nula son entregados por el paquete implementado en R-project por McLachlan y Peel (1999) llamado librería EMMIX.

Notación: Si el algoritmo EM encuentra convergencia en la iteración $(k+1)$, entonces se denotarán los estimadores MV como $\boldsymbol{\mu}^{(k+1)} = \tilde{\boldsymbol{\mu}}$, $\boldsymbol{\Sigma}_0^{(k+1)} = \tilde{\boldsymbol{\Sigma}}$ y $\eta^{(k+1)} = \tilde{\eta}$.

A.5.3. Estadístico de prueba

El estadístico de prueba basado en el TRV queda como

$$\Lambda = \left[\frac{K_d(\tilde{\eta})}{K_d(\hat{\eta})} \right]^n \left[\frac{|\tilde{\Sigma}|}{|\hat{\Sigma}|} \right]^{-n/2} \times \frac{\prod_{i=1}^n \left(1 + c(\tilde{\eta})\tilde{\delta}_{i0}^2 \right)^{-(1/2\tilde{\eta}+d/2)}}{\prod_{i=1}^n \left(1 + c(\hat{\eta})\hat{\delta}_i^2 \right)^{-(1/2\hat{\eta}+d/2)}}.$$

Si η tiende a cero, se sabe que $\lim_{1/\eta \rightarrow \infty} \mathcal{T}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta) = N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ entonces el TRV queda

$$\Lambda = \left[\frac{|\tilde{\Sigma}|}{|\hat{\Sigma}|} \right]^{-n/2}$$

y como $\tilde{\Sigma}$ es diagonal se tiene que $|\tilde{\Sigma}|^{-1}|\hat{\Sigma}| = |\mathbf{R}|$ donde \mathbf{R} es la matriz de correlación, el TRV asumiendo normalidad se puede escribir como $\Lambda = |\mathbf{R}|^{n/2}$.

A.6. Derivadas de influencia

Sea el coeficiente de correlación de concordancia bivariado una función de los estimadores MV perturbados definido por

$$\rho(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}) = \frac{2\hat{\sigma}_{12}(\boldsymbol{\omega})}{\hat{\sigma}_{11}(\boldsymbol{\omega}) + \hat{\sigma}_{22}(\boldsymbol{\omega}) + (\hat{\mu}_1(\boldsymbol{\omega}) - \hat{\mu}_2(\boldsymbol{\omega}))^2},$$

donde $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + a\mathbf{h}$ y $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}} = (\hat{\mu}_1(\boldsymbol{\omega}), \hat{\mu}_2(\boldsymbol{\omega}), \hat{\sigma}_{11}(\boldsymbol{\omega}), \hat{\sigma}_{12}(\boldsymbol{\omega}), \hat{\sigma}_{22}(\boldsymbol{\omega}))^T$ es el estimador MV perturbado.

La derivada direccional de $\rho(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$ respecto a a se define como las derivadas de influencia del coeficiente de correlación de concordancia.

A.6.1. Ponderación de casos

Para el modelo estadístico con densidad normal multivariada el logaritmo de la función de verosimilitud perturbada está dada por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \left\{ -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right\} \quad (\text{A.6.1})$$

derivando (A.6.1) respecto a $\boldsymbol{\mu}$ y $\boldsymbol{\phi}$ con $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ se tiene que los estimadores MV en función de $\boldsymbol{\omega}$ están dados por

$$\hat{\mu}_j(\boldsymbol{\omega}) = \frac{\sum_{i=1}^n \omega_i X_{ij}}{\sum_{i=1}^n \omega_i} \quad \text{y} \quad \hat{\sigma}_{jl}(\boldsymbol{\omega}) = \frac{\sum_{i=1}^n \omega_i (X_{ij} - \hat{\mu}_j(\boldsymbol{\omega})) (X_{il} - \hat{\mu}_l(\boldsymbol{\omega}))}{\sum_{i=1}^n \omega_i}$$

con $\omega_i = \omega_{i0} + a h_i$ para todo $i = 1, \dots, n$.

La primera derivada de los estimadores MV perturbados es

$$\beta_j(\boldsymbol{\omega}) = \frac{d\hat{\mu}_j(\boldsymbol{\omega})}{da} \quad \text{y} \quad \alpha_{jl}(\boldsymbol{\omega}) = \frac{d\hat{\sigma}_{jl}(\boldsymbol{\omega})}{da} \quad \text{para todo } j, l = 1, \dots, d.$$

Para $d = 2$ la primera derivada de los estimadores MV perturbados son

$$\begin{aligned} \alpha_{11}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(1)} - \mathbf{1}_n \hat{\sigma}_{11})}{n} \\ \alpha_{12}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{12})}{n} \\ \alpha_{22}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{(\mathbf{Z}_{(2)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{22})}{n} \\ \beta_1(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{\mathbf{Z}_{(1)}}{n} \\ \beta_2(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{\mathbf{Z}_{(2)}}{n} \end{aligned}$$

donde \odot es el punto Hadamard, $\mathbf{Z}_{(1)} = (Z_{11}, \dots, Z_{n1})^T$, $\mathbf{Z}_{(2)} = (Z_{12}, \dots, Z_{n2})^T$ y $Z_{ij} = (X_{ij} - \hat{\mu}_j) \quad \forall i = 1, \dots, n \quad \forall j = 1, 2$.

Luego la primera derivada de influencia de $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})$ es

$$\begin{aligned}
\left. \frac{d\rho(\hat{\boldsymbol{\theta}}_\omega)}{da} \right|_{a=0} &= \left. \frac{2(d\hat{\sigma}_{12}(\omega)/da)}{(\hat{\sigma}_{11}(\omega) + \hat{\sigma}_{22}(\omega) + (\hat{\mu}_1(\omega) - \hat{\mu}_2(\omega))^2)} \right|_{a=0} \\
&- \left. \frac{2\hat{\sigma}_{12}(\omega)d[(\hat{\sigma}_{11}(\omega) + \hat{\sigma}_{22}(\omega) + (\hat{\mu}_1(\omega) - \hat{\mu}_2(\omega))^2)/da]}{[\hat{\sigma}_{11}(\omega) + \hat{\sigma}_{22}(\omega) + (\hat{\mu}_1(\omega) - \hat{\mu}_2(\omega))^2]^2} \right|_{a=0} \\
&= \left(\frac{\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)}{\hat{\sigma}_{12}(\omega)} \alpha_{12}(\omega) \right) \Big|_{a=0} \\
&- \left(\frac{\boldsymbol{\rho}^2(\hat{\boldsymbol{\theta}}_\omega)}{2\hat{\sigma}_{12}(\omega)} [\alpha_{11}(\omega) + \alpha_{22}(\omega) - 2(\beta_1(\omega) - \beta_2(\omega))] \right) \Big|_{a=0} \\
&= \mathbf{h}^T \left\{ \frac{\hat{\rho}_c}{n\hat{\sigma}_{12}} (\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{12}) - \frac{\hat{\rho}_c^2}{2n\hat{\sigma}_{12}} \mathbf{Z}^* \right\}
\end{aligned}$$

donde $\mathbf{Z}^* = [(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(1)} - \mathbf{1}_n \hat{\sigma}_{11}) + (\mathbf{Z}_{(2)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{22}) + 2(\hat{\mu}_1 - \hat{\mu}_2)(\mathbf{Z}_{(1)} - \mathbf{Z}_{(2)})]$.

Para calcular la segunda derivada se define

$$A(\boldsymbol{\omega}) = \frac{\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)}{\hat{\sigma}_{12}(\omega)} \alpha_{12}(\omega) \quad \text{y} \quad B(\boldsymbol{\omega}) = \frac{\boldsymbol{\rho}^2(\hat{\boldsymbol{\theta}}_\omega)}{2\hat{\sigma}_{12}(\omega)} [\alpha_{11}(\omega) + \alpha_{22}(\omega) - 2(\beta_1(\omega) - \beta_2(\omega))]$$

de modo que

$$\left. \frac{d^2 \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)}{da^2} \right|_{a=0} = \left. \frac{dA(\boldsymbol{\omega})}{da} \right|_{a=0} - \left. \frac{dB(\boldsymbol{\omega})}{da} \right|_{a=0} = \mathbf{h}^T (\boldsymbol{\Gamma}_3 - (\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2)) \mathbf{h}$$

con

$$\begin{aligned}
\boldsymbol{\Gamma}_1 &= \left(\frac{\nabla_f(\boldsymbol{\omega}_0)}{\hat{\sigma}_{12}} - \left(\frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}} \right)^2 \frac{(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{12})}{n} \right) \\
&\times \left(\frac{(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(1)} - \mathbf{1}_n \hat{\sigma}_{11})^T}{n} + \frac{(\mathbf{Z}_{(2)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \hat{\sigma}_{22})^T}{n} \right) \\
&+ \frac{1}{2} \frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}} \frac{2}{n^2} (\mathbf{1}_n \mathbf{1}_n^T \hat{\sigma}_{11}^2 - \mathbf{1}_n (\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(1)})^T - \mathbf{Z}_{(1)} \mathbf{Z}_{(1)}^T) \\
&+ \frac{1}{2} \frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}} \frac{2}{n^2} (\mathbf{1}_n \mathbf{1}_n^T \hat{\sigma}_{22}^2 - \mathbf{1}_n (\mathbf{Z}_{(2)} \odot \mathbf{Z}_{(2)})^T - \mathbf{Z}_{(2)} \mathbf{Z}_{(2)}^T)
\end{aligned}$$

$$\begin{aligned}
\Gamma_2 &= \frac{\widehat{\rho}_c^2}{n^2 \widehat{\sigma}_{12}} \left((\mathbf{Z}_{(1)} - \mathbf{Z}_{(2)})(\mathbf{Z}_{(1)} - \mathbf{Z}_{(2)})^T - (n+1)(\widehat{\mu}_1 - \widehat{\mu}_2)(\mathbf{Z}_{(1)} - \mathbf{Z}_{(2)}) \mathbf{1}_n^T \right) \\
\Gamma_3 &= \frac{2}{n^2} \frac{\widehat{\rho}_2}{\widehat{\sigma}_{12}} \left(\mathbf{1}_n \mathbf{1}_n^T \widehat{\sigma}_{12} - \mathbf{1}_n (\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)})^T - \mathbf{Z}_{(1)} \mathbf{Z}_{(2)}^T \right) - \frac{1}{2n^2} \left(\frac{\widehat{\rho}_c}{\widehat{\sigma}_{12}} \right)^2 \\
&\quad \times \left((\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(1)} - \mathbf{1}_n \widehat{\sigma}_{11}) + (\mathbf{Z}_{(2)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \widehat{\sigma}_{22}) + 2(\mu_1 - \mu_2)(\mathbf{Z}_{(1)} - \mathbf{Z}_{(2)}) \right) \\
&\quad \left((\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \widehat{\sigma}_{12}) \right)^T.
\end{aligned}$$

donde $\widehat{\rho}_c = \boldsymbol{\rho}(\widehat{\boldsymbol{\theta}}_\omega)|_{\omega=\omega_0}$ y los estimadores MV $\widehat{\mu}_1 = \widehat{\mu}_1(\boldsymbol{\omega})|_{\omega=\omega_0}$, $\widehat{\mu}_2 = \widehat{\mu}_2(\boldsymbol{\omega})|_{\omega=\omega_0}$, $\widehat{\sigma}_{11} = \widehat{\sigma}_{11}(\boldsymbol{\omega})|_{\omega=\omega_0}$, $\widehat{\sigma}_{12} = \widehat{\sigma}_{12}(\boldsymbol{\omega})|_{\omega=\omega_0}$, $\widehat{\sigma}_{22} = \widehat{\sigma}_{22}(\boldsymbol{\omega})|_{\omega=\omega_0}$.

Finalmente el vector y matriz para construir las medidas de influencia de primer y segundo orden quedan

$$\begin{aligned}
\nabla_f(\boldsymbol{\omega}_0) &= \left\{ \frac{\widehat{\rho}_c}{n \widehat{\sigma}_{12}} \left(\mathbf{Z}_{(1)} \odot \mathbf{Z}_{(2)} - \mathbf{1}_n \widehat{\sigma}_{12} \right) - \frac{\widehat{\rho}_c^2}{2n \widehat{\sigma}_{12}} \mathbf{Z}^* \right\} \\
\mathbf{H}_{f(\boldsymbol{\omega}_0)} &= \Gamma_3 - (\Gamma_1 + \Gamma_2).
\end{aligned}$$

A.6.2. Perturbación de la respuesta

Para el modelo estadístico con densidad normal multivariada el logaritmo de la función de verosimilitud perturbada está dada por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \left\{ -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{x}_{i\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{i\omega} - \boldsymbol{\mu}) \right\} \quad (\text{A.6.2})$$

donde $\mathbf{x}_{i\omega} = \mathbf{x}_i + \omega_i \mathbf{1}_d$. Derivando (A.6.2) respecto a $\boldsymbol{\mu}$ y $\boldsymbol{\phi}$ con $\boldsymbol{\phi} = \text{vech}(\boldsymbol{\Sigma})$ se tiene que los estimadores MV perturbados son

$$\widehat{\mu}_j(\boldsymbol{\omega}) = n^{-1} \sum_{i=1}^n (X_{ij} + \omega_i) \quad \text{y} \quad \widehat{\sigma}_{jl}(\boldsymbol{\omega}) = n^{-1} \sum_{i=1}^n (X_{ij} - \widehat{\mu}_j(\boldsymbol{\omega}) + \omega_i)(X_{il} - \widehat{\mu}_l(\boldsymbol{\omega}) + \omega_i)$$

donde $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + a\mathbf{h}$ y sus derivadas por

$$\beta_j(\boldsymbol{\omega}) = \frac{d\widehat{\mu}_j(\boldsymbol{\omega})}{da} \quad \text{y} \quad \alpha_{jl}(\boldsymbol{\omega}) = \frac{d\widehat{\sigma}_{jl}(\boldsymbol{\omega})}{da} \quad \text{para todo } j, l = 1, \dots, d.$$

Para $d = 2$ la primera derivada de los estimadores MV perturbados evaluados en $a = 0$ quedan

$$\begin{aligned}\alpha_{11}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T 2\mathbf{P}_n \mathbf{Z}_{(1)} \\ \alpha_{12}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T 2\mathbf{P}_n \mathbf{Z}_{(2)} \\ \alpha_{22}(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T 2\mathbf{P}_n \mathbf{Z} \\ \beta_1(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{\mathbf{1}_n}{n} \\ \beta_2(\boldsymbol{\omega})|_{a=0} &= \mathbf{h}^T \frac{\mathbf{1}_n}{n}\end{aligned}$$

donde $\mathbf{P}_n = [n^{-1}\mathbf{I}_n - n^2\mathbf{1}_n\mathbf{1}_n^T]$ y $\mathbf{Z} = [\mathbf{Z}_{(1)} + \mathbf{Z}_{(2)}]$, $\mathbf{Z}_{(1)} = (Z_{11}, \dots, Z_{n1})^T$, $\mathbf{Z}_{(2)} = (Z_{12}, \dots, Z_{n2})^T$ y $Z_{ij} = (X_{ij} - \hat{\mu}_j) \quad \forall i = 1, \dots, n \quad \forall j = 1, 2$.

La primera derivada de influencia de $\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)$ es

$$\frac{d\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)}{da}\Big|_{a=0} = \mathbf{h}^T \left\{ \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} - \frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \right\}$$

y la segunda derivada es

$$\begin{aligned}\frac{d^2\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_\omega)}{da^2}\Big|_{a=0} &= \mathbf{h}^T \left\{ 2\mathbf{P}_n \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right) - 2\mathbf{P}_n \mathbf{Z} \mathbf{Z}^T \mathbf{P}_n \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}^2} \right) + 2\mathbf{P}_n \mathbf{Z} \mathbf{Z}^T \mathbf{P}_n \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right)^2 \right\} \mathbf{h} \\ &= \mathbf{h}^T \left\{ 2 \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right) \mathbf{P}_n + 2 \left(\left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right)^2 - \frac{\hat{\rho}_c}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \mathbf{Z}^T \mathbf{P}_n \right\} \mathbf{h}.\end{aligned}$$

Finalmente el vector y matriz para construir las medidas de influencia de primer y segundo orden quedan

$$\begin{aligned}\nabla_f(\boldsymbol{\omega}_0) &= \left\{ \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} - \frac{\hat{\rho}_c^2}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \right\} \\ \mathbf{H}_{f(\boldsymbol{\omega}_0)} &= \left\{ 2 \left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right) \mathbf{P}_n + 2 \left(\left(\frac{\hat{\rho}_c}{\hat{\sigma}_{12}} \right)^2 - \frac{\hat{\rho}_c}{\hat{\sigma}_{12}^2} \right) \mathbf{P}_n \mathbf{Z} \mathbf{Z}^T \mathbf{P}_n \right\}.\end{aligned}$$

Estudio de simulación

Esta sección tiene por objetivo evaluar y comparar el desempeño de la inferencia estadística sobre el coeficiente de correlación de concordancia bivariado y global asumiendo que los datos son generados a partir de una distribución t y una distribución normal.

Es de interés estudiar características tales como: tasas de cobertura de los intervalos de confianza, tasas superior e inferior de los intervalos de confianza y potencia empírica de la prueba estadística.

Por otro lado, se desea evaluar el porcentaje de observaciones atípicas y/o extremas que son detectados según la medida de influencia considerada, simulando datos desde una distribución normal.

B.1. Escenarios de simulación

Coefficiente de correlación de concordancia bivariado

Se considera el vector $\boldsymbol{\mu} = (0, 0)^T$ y parámetro de forma $\eta = 0.25$ con tres matrices de varianzas-covarianzas basada en Lin (1989).

Escenario 1

$$\Sigma = \begin{pmatrix} (4/2)^2 & 0.5 \times (4/3) \times (2/3) \\ 0.5 \times (4/3) \times (2/3) & (2/3) \end{pmatrix}$$

Escenario 2

$$\Sigma = \begin{pmatrix} 0.9^2 & 0.8 \times 0.9 \times 1.1 \\ 0.8 \times 0.9 \times 1.1 & 1.1^2 \end{pmatrix}$$

Escenario 3

$$\Sigma = \begin{pmatrix} 1 & 0.95 \\ 0.95 & 1 \end{pmatrix}$$

Coefficiente de correlación de concordancia global

Se considera el vector $\boldsymbol{\mu} = (0, 0, 0)^T$ y parámetro de forma $\eta = 0.25$, $\eta = 0.10$ y $\eta = 0.05$ con tres matrices de varianzas-covarianzas basada en Lin (1989).

Escenario 1

$$\Sigma = \begin{pmatrix} (4/2)^2 & 0.5 \times (4/3) \times (2/3) & 0.5 \times (4/3) \times (2/3) \\ 0.5 \times (4/3) \times (2/3) & (2/3) & 0.5 \times (4/3) \times (2/3) \\ 0.5 \times (4/3) \times (2/3) & 0.5 \times (4/3) \times (2/3) & (4/3)^2 \end{pmatrix}$$

Escenario 2

$$\Sigma = \begin{pmatrix} 0.9^2 & 0.8 \times 0.9 \times 1.1 & 0.8 \times 0.9 \times 1.1 \\ 0.8 \times 0.9 \times 1.1 & 1.1^2 & 0.8 \times 0.9 \times 1.1 \\ 0.8 \times 0.9 \times 1.1 & 0.8 \times 0.9 \times 1.1 & 1.1^2 \end{pmatrix}$$

Escenario 3

$$\Sigma = \begin{pmatrix} 1 & 0.95 & 0.95 \\ 0.95 & 1 & 0.95 \\ 0.95 & 0.95 & 1 \end{pmatrix}$$

B.2. Cobertura y consistencia

Para la cobertura del intervalo de confianza se consideran $R = 2000$ réplicas de datos simulados de tamaño 100, 150 y 200 desde una distribución normal y t . El proceso es el siguiente:

1. Generar R veces datos d variados desde una distribución normal y t .
2. Obtener los R estimadores MV y obtener el estimador del coeficiente de correlación de concordancia promedio.
3. Obtener las R matrices de Fisher y obtener el estimador promedio de la varianza estimada del coeficiente de concordancia.
4. Obtener la tasa inferior y superior de los intervalos de confianza asintóticos.
5. Obtener la tasa de cobertura de los intervalos de confianza asintóticos.

Para estudiar la consistencia de la varianza estimada del estimador del coeficiente de correlación de concordancia bivariado y global se realizó el siguiente procedimiento

1. Se consideran diferentes tamaños de muestra 10, 20 ,..., 500.
2. Se generan datos d variantes desde una distribución normal y t para cada tamaño de muestra.
3. Se obtienen los estimadores MV para cada tamaño de muestra.
4. Se obtiene la matriz de información de Fisher estimada para cada tamaño de muestra.

5. Se obtiene la varianza estimada para el estimador del CCC y CCCG para cada tamaño de muestra.

Comentarios

La Tabla B.1 y B.2 muestran el desempeño del estimador promedio del CCC, la varianza estimada promedio del estimador del CCC, la tasa de cobertura y la tasa de los límites de los intervalos de confianza asintóticos. A medida que crece el tamaño de muestra el estimador promedio del CCC se aproxima al verdadero valor del CCC, la varianza estimada promedio disminuye, la tasa de cobertura converge a la verdadera cobertura (95 %) y la tasa de los límites superior e inferior son más precisos.

En el caso del CCCG, la Tabla B.3 y B.4 muestran el desempeño del estimador promedio del CCCG, la varianza estimada promedio del estimador del CCCG, la tasa de cobertura y la tasa de los límites de los intervalos de confianza asintóticos. Nuevamente, a medida que crece el tamaño de muestra ocurre la misma situación anterior.

Por otro lado, bajo los escenarios 1, 2 y 3 la varianza estimada converge a la varianza verdadera del estimador del CCC y CCCG cuando se simulan datos desde una distribución t (ver Figura B.1, B.2 y B.3). Lo mismo ocurre al simular datos desde una distribución normal (ver Figura B.4, B.5 y B.6).

Se concluye que el desempeño de éstas características es adecuado, ya que presentarán tasas aproximadas a los valores reales.

B.3. Potencia empírica para el TRV y Wald

Se consideran $R = 2000$ réplicas de datos simulados generados desde una distribución t y normal de parámetros definidos en los escenarios 1, 2 y 3 para tres tamaños de muestra $n = 100$, $n = 150$ y $n = 200$.

TRV

Para cada réplica $r = 1, \dots, 2000$ se obtienen los estimadores MV bajo la hipótesis nula y bajo todo el espacio paramétrico usando la librería EMMIX de R-project. Luego, para cada réplica $r = 1, \dots, 2000$ se obtiene el estadístico de prueba asintótico y se obtiene el número de veces que se ha rechazado la hipótesis nula, es decir la potencia empírica.

Wald

Para cada réplica $r = 1, \dots, 2000$ se calcula el los estimadores MV y se construye el estadístico de prueba bajo la hipótesis nula. Luego, se obtiene el número de veces que se ha rechazado la hipótesis nula a medida que se aleja el valor del CCC bajo la hipótesis alternativa, es decir se obtiene la potencia empírica.

Comentarios

El estadístico de prueba asintótico, basado en la razón de verosimilitud para docimar si el coeficiente de correlación de concordancia es significativo, tiene la capacidad de detectar diferencias significativas cuando la hipótesis alternativa se aleja poco de la hipótesis nula.

Asumiendo datos simulados normal y t la Tabla B.5 y B.6 muestra la potencia empírica según el tamaño de muestra y el nivel de significancia para el CCC, mientras que la Tabla B.7 y B.8 muestra la potencia empírica según el tamaño de muestra y el nivel de significancia para el CCCG.

Se observa que al considerar niveles de significancia más flexibles la potencia empírica es mayor y por otro lado al disminuir el tamaño de la muestra la potencia empírica disminuye. Además, el tamaño de la prueba es similar al tamaño nominal.

Por lo tanto, el desempeño de la potencia en relación con el tamaño de la muestra y el nivel de significancia fue el esperado, es decir aumenta a medida que el tamaño de la muestra aumenta y el valor del coeficiente se aleja de cero.

En cuanto a la prueba de Wald, para docimar si el coeficiente es significativo respecto a un valor preestablecido, la potencia empírica aumenta a medida que el nivel de significancia aumenta y también el tamaño de muestra. Sin embargo, el desempeño del test es más bajo al simular datos t comparado con el desempeño al simular datos normales (ver Tabla B.9, B.10, B.11 y B.12). Además se observa que la potencia empírica en los diferentes escenarios no es estable por lo que el desempeño de la prueba de Wald es relativo ya que depende del valor del coeficiente a docimar.

B.4. Simulación para las medidas de influencia

Se consideran $R = 500$ réplicas de datos simulados de tamaño $n = 100$ y $d = 2$ asumiendo una distribución normal con media $\boldsymbol{\mu} = (0, 0)^T$ y $\boldsymbol{\phi} = (1, 0.95, 1)^T$ $\forall i = 1, \dots, n \quad \forall j = 1, 2$.

Se perturba el caso 1 como $x_{1j}^* = x_{1j} + \epsilon$ donde ϵ toma los valores 0.5, 1.5, 2.0, 2.5 y 3.0, posteriormente se realiza el análisis de influencia local, influencia local generalizada e influencia local con medidas de primer y segundo orden.

La Tabla B.13 y B.14 muestran el porcentaje de veces que es detectado el dato perturbado para diferentes valores de ϵ según las medidas de influencia local considerada.

Comentarios

En general, el porcentaje de detección del dato perturbado, según la medida de influencia, es superior al 75 % bajo el esquema de ponderación de casos. Sin embargo, la medida de segundo orden es más conservadora en detectar el dato perturbado ya que la magnitud de la influencia es menor a la magnitud de la influencia de las otras medidas (ver Figura B.7).

En el caso del esquema de perturbación de la respuesta, ocurre que la dirección

máxima de la conformal curvatura normal y la segunda dirección máxima de la curvatura detectan sobre un 75% el dato perturbado. La dirección máxima de las medidas de primer y/o segundo orden siguen siendo más conservadoras en el sentido de que la magnitud de la influencia es menor respecto a la magnitud de las otras medidas (ver Figura B.8). Por otro lado, la dirección máxima de la curvatura no detecta el dato perturbado, en Kim (1996) sugiere graficar la primera dirección versus la segunda dirección para observar en que dirección los datos son influyentes.

Tabla B.1: Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCC simulando datos desde una distribución t

Escenario	ρ_c	n	$\hat{\rho}_c$	$\hat{\sigma}_{\hat{\rho}_c}^2$	TC	LI	LS
1	0.40	150	0.3987	0.0035	0.9470	0.2827	0.5147
		100	0.3943	0.0052	0.9420	0.2532	0.5354
		50	0.3935	0.0105	0.9380	0.1952	0.5918
		25	0.3736	0.0200	0.9070	0.1021	0.6450
2	0.78	150	0.7819	0.0012	0.9440	0.7146	0.8491
		100	0.7794	0.0018	0.9390	0.6965	0.8623
		50	0.7745	0.0037	0.9370	0.6583	0.8906
		25	0.7560	0.0083	0.9260	0.5875	0.9245
3	0.95	150	0.9492	0.0001	0.9460	0.9311	0.9673
		100	0.9483	0.0001	0.9410	0.9258	0.9708
		50	0.9470	0.0003	0.9350	0.9151	0.9790
		25	0.9409	0.0007	0.9220	0.8926	0.9893

Tabla B.2: Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCC simulando datos desde una distribución normal

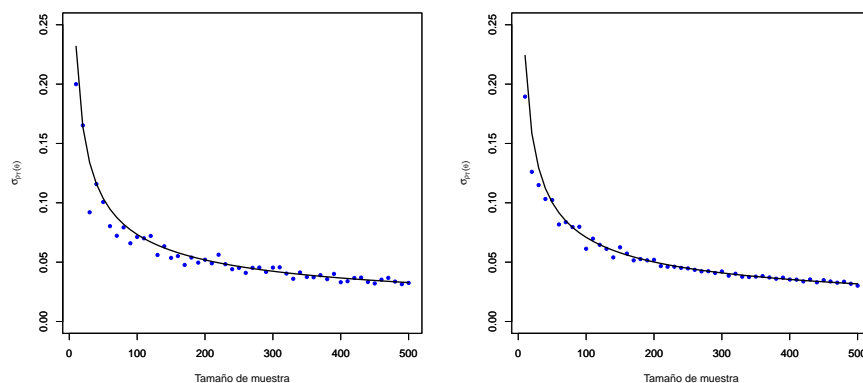
Escenario	ρ_c	n	$\hat{\rho}_c$	$\hat{\sigma}_{\hat{\rho}_c}^2$	TC	LI	LS
1	0.40	150	0.3979	0.0027	0.9460	0.2969	0.4989
		100	0.3956	0.0040	0.9390	0.2722	0.5190
		50	0.3922	0.0079	0.9405	0.2191	0.5653
		25	0.3816	0.0155	0.9260	0.1412	0.6221
2	0.78	150	0.7817	0.0009	0.9415	0.7230	0.8403
		100	0.7796	0.0014	0.9350	0.7074	0.8517
		50	0.7750	0.0029	0.9365	0.6722	0.8777
		25	0.7641	0.0062	0.9255	0.6162	0.9121
3	0.95	150	0.9492	0.0001	0.9485	0.9334	0.9650
		100	0.9485	0.0001	0.9450	0.9290	0.9681
		50	0.9471	0.0002	0.9455	0.9189	0.9752
		25	0.9435	0.0005	0.9330	0.9017	0.9853

Tabla B.3: Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCCG simulando datos desde una distribución t

Escenario	ρ_g	n	$\hat{\rho}_g$	$\hat{\sigma}_{\hat{\rho}_g}^2$	TC	LI	LS
1	0.33	150	0.3289	0.0033	0.9340	0.2171	0.4407
		100	0.3252	0.0048	0.9330	0.1894	0.4610
		50	0.3229	0.0092	0.9160	0.1353	0.5105
		25	0.3142	0.0166	0.9010	0.0631	0.5653
2	0.74	150	0.7303	0.0013	0.9300	0.6602	0.8004
		100	0.7264	0.0020	0.9310	0.6400	0.8128
		50	0.7235	0.0039	0.9380	0.6029	0.8441
		25	0.7091	0.0080	0.9210	0.5396	0.8787
3	0.95	150	0.9486	0.0001	0.9390	0.9331	0.9641
		100	0.9476	0.0001	0.9440	0.9284	0.9669
		50	0.9464	0.0002	0.9380	0.9189	0.9738
		25	0.9421	0.0005	0.9340	0.9017	0.9826

Tabla B.4: Promedio de los parámetros estimados, tasa de cobertura y tasa de los límites superior e inferior de los intervalos de confianza asintóticos para el CCCG simulando datos desde una distribución normal

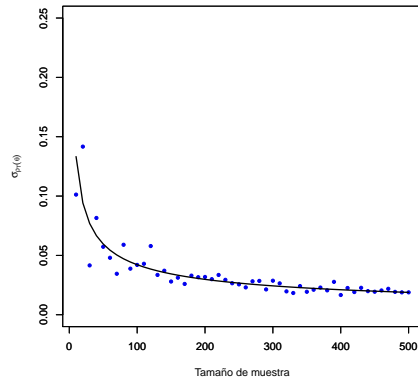
Escenario	ρ_g	n	$\hat{\rho}_g$	$\hat{\sigma}_{\hat{\rho}_g}^2$	TC	LI	LS
1	0.33	150	0.3303	0.0026	0.9520	0.2312	0.4295
		100	0.3310	0.0038	0.9455	0.2104	0.4517
		50	0.3266	0.0074	0.9320	0.1588	0.4944
		25	0.3204	0.0138	0.9195	0.0912	0.5496
2	0.78	150	0.7323	0.0010	0.9545	0.6705	0.7941
		100	0.7317	0.0015	0.9465	0.6562	0.8072
		50	0.7262	0.0031	0.9380	0.6189	0.8334
		25	0.7172	0.0063	0.9295	0.5656	0.8688
3	0.95	150	0.9491	0.0000	0.9515	0.9355	0.9627
		100	0.9488	0.0001	0.9505	0.9321	0.9655
		50	0.9472	0.0002	0.9490	0.9230	0.9714
		25	0.9442	0.0004	0.9405	0.9086	0.9799



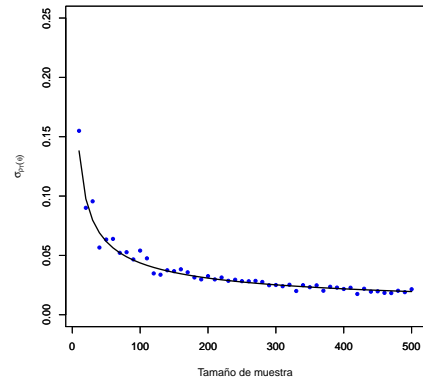
(a) Escenario 1 para CCC

(b) Escenario 1 para CCCG

Figura B.1: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 1

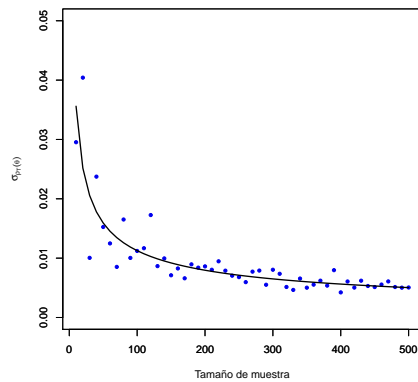


(a) Escenario 2 para CCC

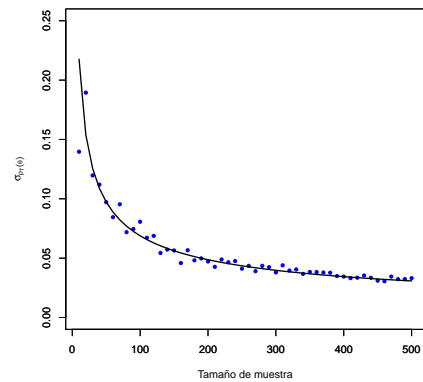


(b) Escenario 2 para CCCG

Figura B.2: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 2

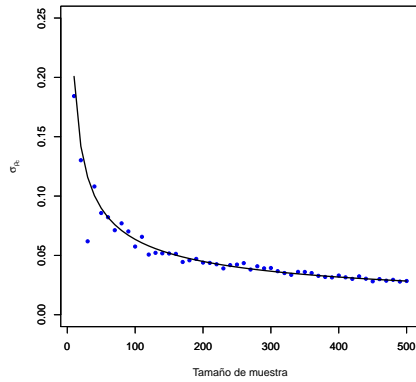


(a) Escenario 3 para CCC

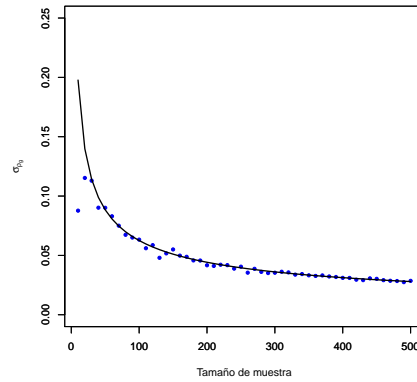


(b) Escenario 3 para CCCG

Figura B.3: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución t bajo el escenario 3

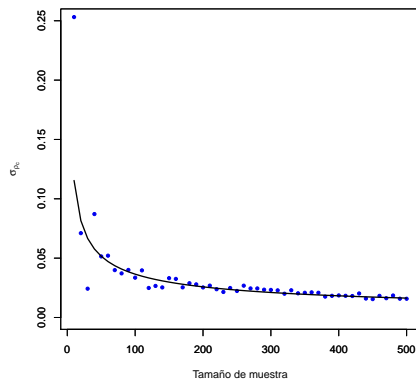


(a) Escenario 1 para CCC

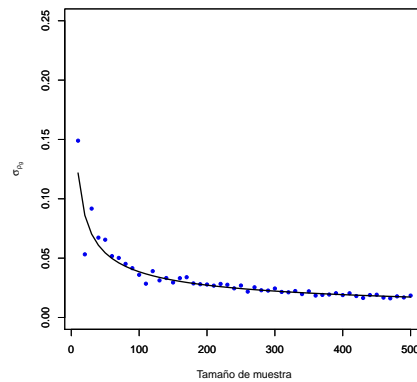


(b) Escenario 1 para CCCG

Figura B.4: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 1



(a) Escenario 2 para CCC



(b) Escenario 2 para CCCG

Figura B.5: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 2

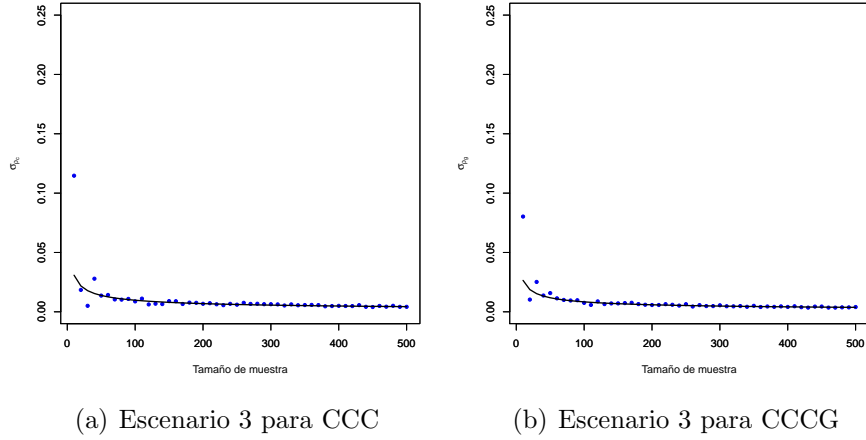


Figura B.6: Consistencia de la varianza estimada para el estimador del coeficiente de correlación de concordancia bivariado y global basado en datos simulados desde una distribución normal bajo el escenario 3

Tabla B.5: Poder empírico para el TRV y tamaño del test para el CCC simulando datos desde una distribución t

n	Hipótesis	ρ_c	$\eta = 0.25$			$\eta = 0.10$			$\eta = 0.05$		
			1 %	5 %	10 %	1 %	5 %	10 %	1 %	5 %	10 %
200	H ₀	0.00	0.008	0.044	0.094	0.013	0.060	0.117	0.010	0.054	0.104
		0.05	0.024	0.097	0.167	0.037	0.111	0.181	0.032	0.106	0.183
	0.10	0.094	0.237	0.343	0.115	0.268	0.378	0.113	0.285	0.407	
	0.15	0.242	0.462	0.584	0.283	0.509	0.633	0.306	0.557	0.673	
	0.20	0.471	0.688	0.809	0.533	0.767	0.852	0.591	0.800	0.879	
	0.25	0.704	0.891	0.943	0.794	0.917	0.955	0.831	0.938	0.966	
	0.30	0.903	0.971	0.987	0.930	0.979	0.991	0.950	0.987	0.995	
150	H ₀	0.00	0.008	0.054	0.096	0.010	0.048	0.098	0.011	0.052	0.100
		0.05	0.023	0.083	0.149	0.025	0.088	0.147	0.028	0.090	0.157
	0.10	0.064	0.185	0.288	0.078	0.198	0.304	0.080	0.217	0.330	
	0.15	0.168	0.375	0.488	0.187	0.410	0.534	0.210	0.431	0.551	
	0.20	0.352	0.581	0.705	0.397	0.639	0.751	0.426	0.660	0.768	
	0.25	0.551	0.791	0.871	0.632	0.831	0.902	0.658	0.852	0.915	
	0.30	0.776	0.923	0.953	0.834	0.944	0.975	0.859	0.962	0.982	
100	H ₀	0.00	0.013	0.058	0.111	0.010	0.043	0.087	0.011	0.056	0.110
		0.05	0.017	0.074	0.135	0.016	0.066	0.129	0.020	0.076	0.138
	0.10	0.044	0.144	0.232	0.040	0.146	0.235	0.051	0.157	0.249	
	0.15	0.105	0.266	0.377	0.117	0.283	0.406	0.125	0.296	0.429	
	0.20	0.215	0.433	0.562	0.236	0.465	0.593	0.260	0.495	0.612	
	0.25	0.373	0.621	0.728	0.421	0.654	0.762	0.450	0.673	0.777	
	0.30	0.562	0.767	0.849	0.621	0.825	0.895	0.647	0.832	0.897	

Tabla B.6: Poder empírico para el TRV y tamaño del test para el CCC simulando datos desde una distribución normal

n	Hipótesis	ρ_c	1 %	5 %	10 %
200	H ₀	0.00	0.007	0.055	0.109
		0.05	0.033	0.112	0.183
	H ₁	0.10	0.128	0.306	0.424
		0.15	0.345	0.586	0.707
		0.20	0.636	0.839	0.899
		0.25	0.861	0.952	0.980
		0.30	0.969	0.994	0.998
150	H ₀	0.00	0.008	0.048	0.100
		0.05	0.021	0.100	0.174
	H ₁	0.10	0.096	0.244	0.355
		0.15	0.251	0.479	0.597
		0.20	0.484	0.713	0.809
		0.25	0.720	0.884	0.938
		0.30	0.893	0.967	0.985
100	H ₀	0.00	0.011	0.058	0.111
		0.05	0.020	0.087	0.152
	H ₁	0.10	0.064	0.187	0.288
		0.15	0.161	0.349	0.459
		0.20	0.313	0.532	0.651
		0.25	0.505	0.723	0.806
		0.30	0.699	0.864	0.921

Tabla B.7: Poder empírico del TRV y tamaño del test para el CCCG simulando datos desde una distribución t

n	Hipótesis	ρ_g	$\eta = 0.25$			$\eta = 0.10$			$\eta = 0.05$		
			1 %	5 %	10 %	1 %	5 %	10 %	1 %	5 %	10 %
200	H ₀	0.00	0.011	0.051	0.105	0.014	0.055	0.105	0.015	0.050	0.104
		0.05	0.038	0.132	0.226	0.048	0.136	0.222	0.050	0.160	0.247
	H ₁	0.10	0.210	0.403	0.529	0.232	0.460	0.585	0.258	0.482	0.610
		0.15	0.550	0.762	0.859	0.620	0.818	0.882	0.664	0.849	0.907
		0.20	0.871	0.955	0.975	0.903	0.978	0.988	0.927	0.975	0.985
150	H ₀	0.00	0.014	0.060	0.106	0.012	0.057	0.106	0.013	0.058	0.117
		0.05	0.030	0.102	0.197	0.031	0.127	0.202	0.037	0.124	0.216
	H ₁	0.10	0.140	0.331	0.447	0.166	0.365	0.494	0.182	0.376	0.495
		0.15	0.410	0.629	0.727	0.475	0.694	0.789	0.486	0.701	0.796
		0.20	0.706	0.863	0.919	0.781	0.913	0.955	0.800	0.920	0.956
100	H ₀	0.00	0.022	0.069	0.125	0.012	0.063	0.117	0.019	0.062	0.115
		0.05	0.034	0.105	0.176	0.030	0.097	0.173	0.031	0.111	0.179
	H ₁	0.10	0.092	0.242	0.341	0.101	0.246	0.353	0.110	0.264	0.377
		0.15	0.249	0.462	0.586	0.274	0.499	0.620	0.300	0.527	0.652
		0.20	0.493	0.710	0.809	0.541	0.755	0.840	0.583	0.780	0.868

Tabla B.8: Poder empírico para el TRV y tamaño del test para el CCCG simulando datos desde una distribución normal

n	Hipótesis	ρ_g	1 %	5 %	10 %
200	H ₀	0.00	0.010	0.049	0.107
		0.05	0.044	0.135	0.222
	H ₁	0.10	0.233	0.453	0.571
		0.15	0.607	0.812	0.882
		0.20	0.901	0.971	0.985
150	H ₀	0.00	0.009	0.054	0.107
		0.05	0.036	0.127	0.211
	H ₁	0.10	0.172	0.359	0.477
		0.15	0.451	0.674	0.781
		0.20	0.771	0.906	0.944
100	H ₀	0.00	0.013	0.049	0.103
		0.05	0.025	0.093	0.176
	H ₁	0.10	0.102	0.263	0.368
		0.15	0.283	0.508	0.631
		0.20	0.552	0.762	0.840

Tabla B.9: Poder empírico de la prueba de Wald y tamaño del test para el CCC simulando datos desde una distribución t

Wald	n	Hipótesis	ρ_c	1 %	5 %	10 %
Escenario 1	200	H ₁	0.60	0.909	0.980	0.992
		H ₀	0.40	0.010	0.045	0.094
		H ₁	0.20	0.469	0.802	0.914
	150	H ₁	0.60	0.782	0.929	0.960
		H ₀	0.40	0.012	0.054	0.103
		H ₁	0.20	0.282	0.638	0.796
	100	H ₁	0.60	0.558	0.795	0.879
		H ₀	0.40	0.015	0.060	0.110
		H ₁	0.20	0.099	0.647	0.592
Escenario 2	200	H ₁	0.88	0.774	0.991	0.995
		H ₀	0.78	0.018	0.053	0.102
		H ₁	0.58	0.557	0.997	1.000
	150	H ₁	0.88	0.910	0.964	0.984
		H ₀	0.78	0.014	0.057	0.101
		H ₁	0.58	0.877	0.977	0.993
	100	H ₁	0.88	0.774	0.884	0.929
		H ₀	0.78	0.018	0.062	0.112
		H ₁	0.58	0.557	0.868	0.933
Escenario 3	200	H ₁	0.97	0.854	0.985	0.960
		H ₀	0.95	0.013	0.052	0.093
		H ₁	0.90	0.875	0.983	0.996
	150	H ₁	0.97	0.749	0.869	0.812
		H ₀	0.95	0.019	0.054	0.105
		H ₁	0.90	0.650	0.921	0.850
	100	H ₁	0.97	0.596	0.748	0.812
		H ₀	0.95	0.022	0.060	0.105
		H ₁	0.90	0.255	0.712	0.850

Tabla B.10: Poder empírico para la prueba de Wald y tamaño del test para el CCC simulando datos desde una distribución normal

Wald	n	Hipótesis	ρ_c	1 %	5 %	10 %
Escenario 1	200	H ₁	0.60	0.973	0.997	0.999
		H ₀	0.40	0.009	0.056	0.110
		H ₁	0.20	0.710	0.920	0.967
	150	H ₁	0.60	0.907	0.980	0.994
		H ₀	0.40	0.009	0.054	0.107
		H ₁	0.20	0.491	0.800	0.904
	100	H ₁	0.60	0.706	0.897	0.943
		H ₀	0.40	0.011	0.061	0.115
		H ₁	0.20	0.230	0.559	0.713
Escenario 2	200	H ₁	0.88	0.993	0.999	1.000
		H ₀	0.78	0.011	0.054	0.116
		H ₁	0.58	0.999	1.000	1.000
	150	H ₁	0.88	0.971	0.991	0.996
		H ₀	0.78	0.011	0.049	0.098
		H ₁	0.58	0.969	0.999	0.999
	100	H ₁	0.88	0.872	0.946	0.968
		H ₀	0.78	0.016	0.056	0.109
		H ₁	0.58	0.772	0.955	0.985
Escenario 3	200	H ₁	0.97	0.929	0.973	0.985
		H ₀	0.95	0.015	0.056	0.107
		H ₁	0.90	0.971	0.999	1.000
	150	H ₁	0.97	0.851	0.934	0.963
		H ₀	0.95	0.013	0.052	0.103
		H ₁	0.90	0.865	0.976	0.994
	100	H ₁	0.97	0.530	0.827	0.884
		H ₀	0.95	0.018	0.055	0.109
		H ₁	0.90	0.695	0.856	0.937

Tabla B.11: Poder empírico para la prueba de Wald y tamaño del test para el CCCG simulando datos desde una distribución t

Wald	n	Hipótesis	ρ_g	1 %	5 %	10 %
Escenario 1	200	H ₁	0.45	0.470	0.684	0.790
		H ₀	0.33	0.010	0.042	0.098
		H ₁	0.20	0.477	0.724	0.830
	150	H ₁	0.45	0.341	0.558	0.673
		H ₀	0.33	0.015	0.062	0.118
		H ₁	0.20	0.352	0.599	0.708
	100	H ₁	0.45	0.225	0.417	0.532
		H ₀	0.33	0.024	0.070	0.125
		H ₁	0.20	0.198	0.431	0.560
Escenario 2	200	H ₁	0.80	0.553	0.735	0.816
		H ₀	0.74	0.015	0.047	0.092
		H ₁	0.62	0.650	0.877	0.935
	150	H ₁	0.80	0.449	0.625	0.715
		H ₀	0.74	0.013	0.064	0.123
		H ₁	0.62	0.456	0.734	0.830
	100	H ₁	0.80	0.331	0.505	0.592
		H ₀	0.74	0.019	0.075	0.132
		H ₁	0.62	0.233	0.532	0.676
Escenario 3	200	H ₁	0.97	0.961	0.984	0.994
		H ₀	0.95	0.011	0.042	0.085
		H ₁	0.92	0.632	0.898	0.948
	150	H ₁	0.97	0.866	0.948	0.971
		H ₀	0.95	0.013	0.059	0.112
		H ₁	0.92	0.410	0.739	0.851
	100	H ₁	0.97	0.732	0.853	0.902
		H ₀	0.95	0.021	0.060	0.114
		H ₁	0.92	0.160	0.510	0.675

Tabla B.12: Poder empírico para la prueba de Wald y tamaño del test para el CCCG simulando datos desde una distribución normal

Wald	n	Hipótesis	ρ_g	1%	5%	10%
Escenario 1	200	H ₁	0.45	0.579	0.785	0.869
		H ₀	0.33	0.013	0.052	0.099
		H ₁	0.20	0.614	0.827	0.898
	150	H ₁	0.45	0.452	0.666	0.767
		H ₀	0.33	0.011	0.049	0.096
		H ₁	0.20	0.466	0.703	0.809
	100	H ₁	0.45	0.267	0.495	0.620
		H ₀	0.33	0.007	0.055	0.103
		H ₁	0.20	0.280	0.538	0.663
Escenario 2	200	H ₁	0.80	0.657	0.828	0.891
		H ₀	0.74	0.015	0.053	0.100
		H ₁	0.62	0.807	0.939	0.972
	150	H ₁	0.80	0.542	0.716	0.801
		H ₀	0.74	0.009	0.046	0.098
		H ₁	0.62	0.631	0.860	0.921
	100	H ₁	0.80	0.385	0.576	0.668
		H ₀	0.74	0.009	0.056	0.110
		H ₁	0.62	0.370	0.679	0.804
Escenario 3	200	H ₁	0.97	0.983	0.995	0.998
		H ₀	0.95	0.014	0.054	0.100
		H ₁	0.92	0.821	0.957	0.980
	150	H ₁	0.97	0.947	0.981	0.992
		H ₀	0.95	0.011	0.049	0.093
		H ₁	0.92	0.605	0.876	0.941
	100	H ₁	0.97	0.815	0.922	0.953
		H ₀	0.95	0.010	0.050	0.102
		H ₁	0.92	0.304	0.673	0.814

Tabla B.13: Porcentaje de detección del dato perturbado según medida de influencia bajo el esquema de ponderación de casos simulando datos desde una distribución normal bivariada

Medidas de Influencia	ϵ					
	0.5	1.0	1.5	2.0	2.5	3.0
\mathbf{h}_{max} para C	78.4	78.8	98.8	100.0	100.0	100.0
\mathbf{h}_{max} para B	88.4	87.2	99.8	100.0	100.0	100.0
\mathbf{h}_{max} para FI	76.6	77.4	97.8	99.8	100.0	100.0
\mathbf{h}_{max} para SI	17.6	18.4	31.8	47.8	80.2	99.8
\mathbf{h}_{max} para FI y SI	84.4	84.2	99.4	100.0	100.0	100.0

Tabla B.14: Porcentaje de detección del dato perturbado según medida de influencia bajo el esquema de perturbación de la respuesta simulando datos desde una distribución normal bivariada

Medidas de Influencia	ϵ					
	1	2	3	4	5	6
\mathbf{h}_{max} para C	6	6.8	2.8	0.2	0.0	0.0
\mathbf{h}_{seg} para C	84	87.6	84.8	86.0	84.8	83.4
\mathbf{h}_{max} para B	77	100.0	100.0	100.0	100.0	100.0
\mathbf{h}_{max} para FI	11	31.4	57.2	94.0	100.0	100.0
\mathbf{h}_{max} para SI	11	31.4	57.2	94.0	100.0	100.0
\mathbf{h}_{max} para FI y SI	11	31.4	57.2	94.0	100.0	100.0

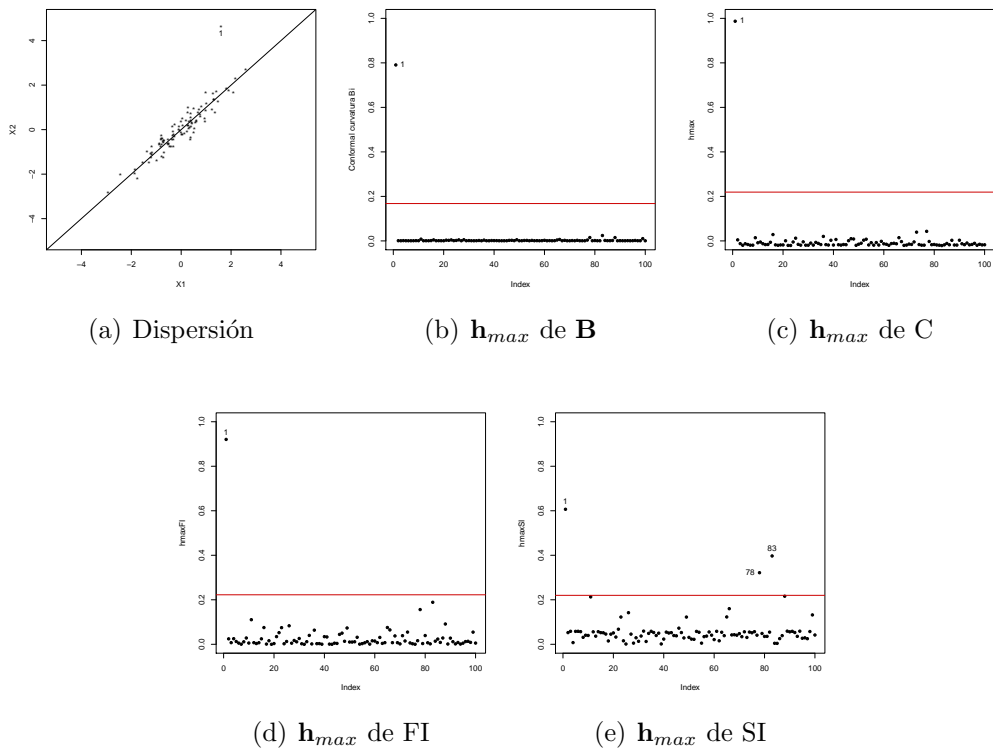


Figura B.7: Dispersión de datos simulados (a), h_{max} de la conformal curvatura normal (b), h_{max} de la curvatura (c), h_{max} de la medida de primer orden (d) y h_{max} de la medida de segundo orden (e), bajo el esquema de ponderación de casos

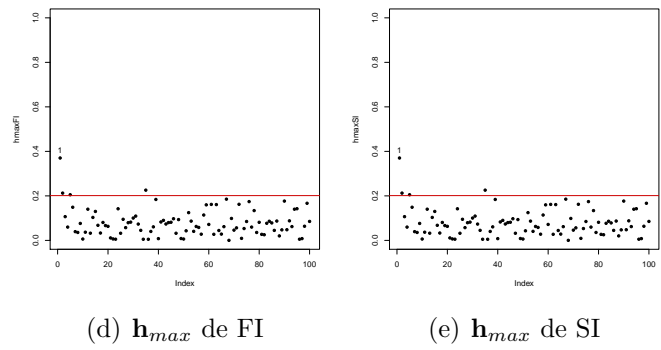
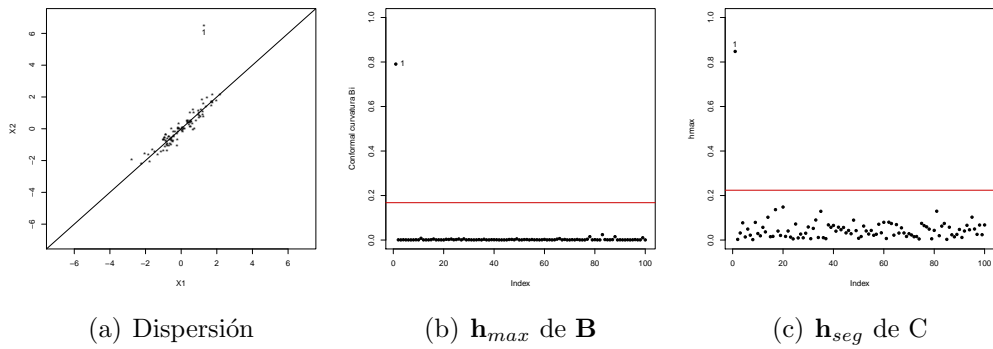


Figura B.8: Dispersión de datos simulados (a), \mathbf{h}_{max} de la conformal curvatura normal (b), \mathbf{h}_{seg} de la curvatura (c), \mathbf{h}_{max} de la medida de primer orden (d) y \mathbf{h}_{max} de la medida de segundo orden (e), bajo el esquema de perturbación de la respuesta