



**Facultad de Matemáticas**  
**Máster en Técnicas Estadísticas**

Proyecto de Fin de Máster

Metodología *ROC* en la Evaluación de Medidas  
Antropométricas como Marcadores de la Hipertensión  
Arterial.

Aplicación a Población Gallega Adulta.

Directora: **Carmen María Cadarso Suárez**

Por Teresa Luisa Noguera Moreno

Santiago de Compostela - 2010



# Índice general

<b>Resumen</b>	<b>v</b>
<b>Agradecimientos</b>	<b>VII</b>
<b>1. Presentación</b>	<b>1</b>
1.1. Objetivos	2
1.2. <i>Software</i> Estadístico	3
<b>2. Las Curvas <i>ROC</i></b>	<b>5</b>
2.1. Capacidad de Discriminación de un Clasificador o Marcador	5
2.1.1. Probabilidades de Clasificación de los Marcadores	6
2.1.2. Valores Predictivos	8
2.2. Espacio <i>ROC</i>	9
2.3. Curvas en el Espacio <i>ROC</i>	11
2.4. Propiedades de la Curva <i>ROC</i>	14
<b>3. Índices Resumen de la Curva <i>ROC</i></b>	<b>17</b>
3.1. Área Bajo la Curva <i>ROC</i>	17
3.2. Puntos y Áreas Parciales	18
3.3. Índice de Youden	18
3.3.1. Características del <i>YI</i>	19
<b>4. Métodos de Estimación del Índice de Youden</b>	<b>23</b>
4.1. Métodos Paramétricos	24
4.2. Métodos No Paramétricos	25

<b>5. Estudios de Simulación</b>	<b>27</b>
5.1. Resultados de los Estudios de Simulación . . . . .	28
5.1.1. Índice de Youden . . . . .	28
5.1.2. Puntos de Corte . . . . .	39
<b>6. Aplicación a Datos Reales</b>	<b>49</b>
6.1. Resultados . . . . .	51
6.1.1. Estudio del Índice de Masa Corporal como Marcador Diagnóstico de Hipertensión Arterial . . . . .	52
6.1.2. Estudio de la Circunferencia de la Cintura como Marcador Diagnóstico de Hipertensión Arterial . . . . .	62
6.1.3. Estudio del Marcador Índice Cintura/Cadera . . . . .	75
<b>7. Conclusiones</b>	<b>89</b>
<b>Bibliografía</b>	<b>91</b>

# Resumen

La curva ROC es una técnica para visualizar, organizar y seleccionar clasificadores o marcadores basándose en la capacidad de discriminación de los mismos en todo el rango de puntos de corte. Para evaluar esta habilidad discriminatoria de un marcador, es común resumir la información de la curva *ROC* en un único valor global o índice. Varios índices son utilizados con este propósito en la evaluación de pruebas diagnósticas, entre los que pueden citarse el Área Bajo la Curva ROC (en sus siglas inglesas *AUC*) o el Índice de Youden. En este proyecto se examinaron diferentes métodos de estimación paramétricos y no paramétricos para la estimación del Índice de Youden y sus puntos de corte asociados basados en: 1) Suposiciones de Normalidad, 2) Transformaciones para la Normalidad, 3) La Función Empírica de Distribución y 4) Suavización Tipo Núcleo. Fueron comparados en términos de sesgo y raíz del Error Cuadrático Medio mediante varios estudios de simulación, encontrándose que el método empírico (3) tuvo el peor comportamiento. Los métodos fueron utilizados para estimar el Índice de Youden y sus puntos de corte asociados en las medidas antropométricas Índice de Masa Corporal, Circunferencia de la Cintura y el Índice Cintura/Cadera, referidas por la Organización Mundial de la Salud como indicadores de obesidad y predictores de hipertensión arterial, en población gallega adulta. Fueron calculadas las curvas *ROC* de los tres marcadores diagnósticos y comparadas sus áreas bajo la curva, atendiendo al sexo y a la edad. No se encontraron diferencias estadísticamente significativas.



# Agradecimientos

Quiero dar gracias a la vida por haberme ofrecido la posibilidad, materializada en este máster, de sobrevivir profesionalmente, y expresar mi agradecimiento:

- A César Andrés Sánchez Sello, por la gentileza, afabilidad y buena voluntad con la que me recibió el primer día que entré a la Facultad de Matemáticas de la USC, ...completamente desorientada..
- A Carmen María Suárez Cadarso por haber aceptado ser la directora de mi proyecto, lo que me ha permitido volver a estar en contacto con la Bioestadística, que tanto me motiva.
- A Mónica López Rato por su paciencia, ayuda y buena disposición.
- A mis compañeras del máster Anaderli Torres Ortiz e Isabel María Martínez Silva por su apoyo.
- A María Xosé Rodríguez Álvarez porque siempre se puede contar con su ayuda.
- A Manuel Febrero Bande por su disponibilidad como coordinador.
- A mi familia, que siempre me ha apoyado.

A mi madre



# Capítulo 1

## Presentación

El gráfico *Receiver Operating Characteristics (ROC)* es una técnica para visualizar, organizar y seleccionar pruebas diagnósticas o marcadores, basándose en su capacidad de discriminación para clasificar correctamente un sujeto o un caso. Su uso se ha extendido para concebir y analizar el comportamiento de sistemas de diagnóstico, destacándose su aplicación en el campo de la medicina.

El diagnóstico es una parte familiar de la vida diaria. En medicina, una prueba diagnóstica a la que las personas están familiarizadas incluye, por ejemplo, cultivo de bacterias, imágenes radiológicas o exámenes bioquímicos. Éstas son pruebas diagnósticas en el sentido clásico. Sin embargo, se puede considerar el diagnóstico médico en un contexto mucho más amplio, donde la metodología estadística pertinente para diagnosticar una enfermedad también es aplicada a otros problemas importantes en medicina, en particular la prognosis.

La prognosis puede ser considerada como un tipo especial de diagnóstico, donde la condición para ser detectada no es una enfermedad, *per se*, pero es un resultado clínico de interés. Por ejemplo, se podrían usar los datos clínicos o laboratoriales de un paciente para predecir o “diagnosticar” hospitalización dentro de un período de tiempo determinado. Un resultado clínico, hospitalización en este caso, es la condición para ser diagnosticada. La exploración (*screening*) para encontrar la presencia o ausencia de un atributo o característica particular, es un caso especial importante del problema de diagnóstico general médico.

La exploración de poblaciones saludables para buscar enfermedades ocultas, que no han debutado, es atractiva para los investigadores porque las enfermedades son, a menudo, exitosamente tratadas si son detectadas precozmente. Desde el punto de vista estadístico las pruebas de exploración son pruebas diagnósticas, es decir, son procedimientos orientados

hacia la detección de una condición y su evaluación es básicamente la misma que la que se hace para las pruebas diagnósticas clásicas. El contexto, en el cual son aplicadas en la práctica, es lo que las separa de esas pruebas clásicas:

- Son aplicadas a individuos “saludables” y a gran escala, por lo que deben ser no-invasivas y económicas.
- Una prueba de exploración positiva es usualmente seguida, no directamente con tratamiento, sino con procedimientos de diagnóstico adicionales y más definitivos.
- Las enfermedades exploradas son a menudo progresivas (enfermedad cardiovascular) y su evolución en el tiempo introduce complicaciones interesantes en su evaluación.

La obesidad es uno de los mayores factores de riesgo para el desarrollo de enfermedades crónicas. El riesgo de eventos cardiovasculares se eleva con el aumento de masa corporal. La acumulación de grasa abdominal ha sido considerada como un factor de riesgo independiente para la diabetes tipo 2 y para condiciones de riesgo cardiovascular, como la hipertensión arterial. Para la exploración de la obesidad, la Organización Mundial de la Salud ha recomendado la utilización del **Índice de Masa Corporal** como un indicador universal del sobrepeso y la obesidad y la **Circunferencia de la Cintura** y el **Índice Cintura/Cadera** como medidas de la distribución de la grasa abdominal.

En este proyecto se pretende evaluar la capacidad discriminadora de estas medidas, mediante la selección de puntos de corte óptimos, como predictores de hipertensión arterial utilizando el análisis de las curvas *ROC* que supone el cálculo de valores globales o índices, como el Área Bajo la Curva (en sus siglas inglesas *AUC*) o el Índice de Youden, para resumir la información proporcionada por las curvas.

## 1.1. Objetivos

- Evaluar la capacidad de discriminación de las medidas antropométricas **Índice de Masa Corporal**, **Circunferencia de la Cintura** y el **Índice Cintura/Cadera** mediante el cálculo de las curvas *ROC*.
- Comparar la efectividad de los diferentes métodos de estimación paramétricos y no paramétricos en la estimación del Índice de Youden y sus puntos de corte asociados mediante varios estudios de simulación.
- Estimar el Índice de Youden y sus puntos de corte asociados para el **Índice de Masa Corporal**, la **Circunferencia de la Cintura** y el **Índice Cintura/Cadera**

utilizando los cuatro métodos en una base de datos de población gallega adulta.

## 1.2. *Software* Estadístico

Para los análisis realizados fue utilizado el **software** [R Development Core Team \(2009\)](#), la librería [Sing y otros \(2009\)](#), la librería [Brasil \(2010\)](#) y rutinas elaboradas con el *software* R por Mónica López Ratón.



## Capítulo 2

# Las Curvas *ROC*

### 2.1. Capacidad de Discriminación de un Clasificador o Marcador

Un clasificador o marcador es considerado una variable explicativa que clasifica en cada caso o sujeto, la presencia o ausencia de una condición, que se considera la variable respuesta. La información que suministra un clasificador, depende de su nivel de medición: *nominal* si sus "valores" son categorías, *binaria* si es nominal con sólo dos categorías, *ordinal* si las categorías pueden ser ordenadas de alguna manera, *discreta* si puede tomar un número finito de posibles valores distintos y *continua* si puede tomar cualquier valor en un rango finito o infinito de valores.

En situaciones de clasificación en las que se usan sólo dos clases, cada objeto, sujeto o caso es etiquetado con uno de los elementos del conjunto {"positivo (P)", "negativo (N)"}, augurándose la clase a la que pertenece dicho caso. Algunos modelos de clasificación producen una salida continua, como la estimación de la probabilidad de un caso de pertenecer a una clase, situación en la que diferentes umbrales de decisión o puntos de corte pueden ser aplicados para predecir la clase a la que pertenece dicho caso, y otros producen apenas la etiqueta discreta de una clase, indicándose con esto la clase predicha de ese caso.

La aplicación de un marcador, en situaciones como la descrita, produce cuatro posibles resultados:

- Si el caso es positivo y es clasificado como positivo se cuenta como un *verdadero positivo*
- Si el caso es positivo y es clasificado como negativo se cuenta como un *falso negativo*.

Clase Hipotetizada	Clase Verdadera	
	P	N
Si	Verdaderos Positivos	Falsos Positivos
No	Falsos Negativos	Verdaderos Negativos

Tabla 2.1: Matriz de Confusión

- Si el caso es negativo y es clasificado como negativo se cuenta como un *verdadero negativo*.
- Si el caso es negativo y es clasificado como positivo se cuenta como un *falso positivo*.

Dados un marcador y un conjunto de sujetos, una matriz de confusión o tabla de contingencia 2 x 2 puede ser construida para representar la disposición de dicho conjunto (ver Tabla 2.1).

### 2.1.1. Probabilidades de Clasificación de los Marcadores

Los números que se encuentran a lo largo de la diagonal principal de la matriz de confusión representada en la Tabla 2.1 representan las clasificaciones correctas y los que están a lo largo de la diagonal secundaria representan los errores (la confusión) entre las clases. Esta matriz es la base o soporte para varios indicadores comúnmente utilizados en sistemas de diagnóstico:

- Fracción de Verdaderos Positivos (*TPF*) denominada también Tasa de Aciertos o *Recall*

$$TPF = \frac{\text{Positivos Clasificados Correctamente (Verdaderos Positivos)}}{\text{Total de Positivos}} \quad (2.1)$$

- Fracción de Falsos Positivos (*FPF*) denominada también Tasa de Falsa Alarma

$$FPF = \frac{\text{Negativos Clasificados Incorrectamente (Falsos Positivos)}}{\text{Total de Negativos}} \quad (2.2)$$

- Precisión

$$\text{Precisión} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Positivos}} \quad (2.3)$$

- Exactitud

$$\text{Exactitud} = \frac{\text{Verdaderos Positivos} + \text{Verdaderos Negativos}}{\text{Positivos} + \text{Negativos}} \quad (2.4)$$

Fawcett (2006) comenta que algunos términos adicionales asociados con las curvas ROC son:

## 2.1 Capacidad de Discriminación de un Clasificador o Marcador 7

- *Sensibilidad* = Fracción de Verdaderos Positivos
- *Especificidad* (*TNF*) definida como la fracción siguiente:

$$\frac{\text{Verdaderos Negativos}}{\text{Falsos Positivos} + \text{Verdaderos Negativos}} = 1 - FPF \quad (2.5)$$

- *Valor Predictivo Positivo* = *Precisión*

Pepe (2003) adapta la matriz de confusión presentada en la Tabla 2.1 usando:

- Una variable binaria denominada *D* para denotar el estado de la enfermedad, en vez de Clase Verdadera:

$$D = \begin{cases} 1 & \text{para Enfermedad Presente} \\ 0 & \text{para Enfermedad No Presente} \end{cases}$$

- Una variable denominada *Y* que representa el resultado de la prueba, en vez de Clase Hipotetizada:

$$Y = \begin{cases} 1 & \text{positivo para Enfermedad} \\ 0 & \text{negativo para Enfermedad} \end{cases}$$

y explica que como  $1 - TPF$  es igual a la Fracción de Falsos Negativos, el par  $(FPF, TPF)$  define las probabilidades con las que los errores son cometidos al utilizar el marcador. Esta autora comenta que, aunque a menudo se utilicen los términos *tasa* de falsos positivos y *tasa* de verdaderos positivos, no son en realidad tasas. Son probabilidades o fracciones y como tal deben ser denominadas.

Sobre esta matriz adaptada al campo de la medicina, Pepe (2003) define:

- la **prevalencia de la población para la enfermedad** como

$$\rho = P[D = 1] \quad (2.6)$$

- la **probabilidad general del error de clasificación** en función del par  $(FPF, TPF)$  y de la prevalencia de la enfermedad:

$$P[Y \neq D] = \rho(1 - TPF) + (1 - \rho)FPF \quad (2.7)$$

En la investigación biomédica los términos Sensibilidad y Especificidad son con frecuencia utilizados para detallar el comportamiento del marcador. [Pepe \(2003\)](#) opina que, acompañando a esta información, deben ser reportados también los errores de clasificación  $(1 - TPF)$  y  $FPF$  para describir la exactitud diagnóstica del marcador, en vez de la probabilidad general. Toda esta información  $((FPF, TPF), (1 - TPF)$  y  $FPF)$  la considera necesaria por dos razones fundamentales:

- La primera razón es que los costes y consecuencias de los dos tipos de errores son, usualmente, verdaderamente diferentes. Los errores **Falsos Negativos**, es decir sujetos enfermos no diagnosticados como tal, pueden resultar en individuos que no reciban el tratamiento que necesitan para su enfermedad. La consecuencia puede ser tan grave como la muerte. Los errores **Falsos Positivos** tienden a ser menos graves: personas sometidas a procedimientos adicionales de diagnóstico o incluso a tratamiento. Los impactos negativos incluyen los inconvenientes personales y el estrés, pero las consecuencias a largo plazo son, usualmente, menores.
- La segunda razón para preferir el par  $(FPF, TPF)$  en vez de la probabilidad general del error de clasificación, es que esta última depende de la prevalencia de la enfermedad. Para ilustrar el problema con esto pone el ejemplo de una prueba totalmente no-informativa, que clasifica todos los sujetos como negativos (libres de la enfermedad). Si la prevalencia de la enfermedad es baja, esta prueba tendrá una probabilidad general de error baja, pero claramente es una prueba inútil.

### 2.1.2. Valores Predictivos

De acuerdo con [Pepe \(2003\)](#), como alternativa para considerar la frecuencia de los errores de clasificación para cada estatus de la enfermedad, la exactitud puede ser cuantificada por “cómo de bien” el resultado de la prueba diagnóstica puede predecir el verdadero estatus de la enfermedad, definiendo los valores predictivos como:

$$\text{Valor Predictivo Positivo} = PPV = P[D = 1/Y = 1] \quad (2.8)$$

$$\text{Valor Predictivo Negativo} = NPV = P[D = 0/Y = 0] \quad (2.9)$$

Una prueba perfecta predecirá la enfermedad perfectamente con valores  $PPV = 1$  y  $NPV = 1$ , pero por otro lado estos valores indicarían una prueba inútil que no tiene información sobre el verdadero estatus de la enfermedad, ya que  $P[D = 1/Y = 1] = P[D = 1]$  y  $P[D = 0/Y = 0] = P[D = 0]$ , es decir  $PPV = \rho$  (ver ecuación 2.6 de la página 7) y  $NPV = 1 - \rho$ .

Los valores predictivos dependen no sólo de la capacidad discriminatoria de la prueba en la clasificación de los sujetos, sino también de la prevalencia de la enfermedad. Un bajo *PPV* podría ser, simplemente, el resultado de una baja prevalencia de la enfermedad o podría ser debido a que a la prueba no refleja muy bien el verdadero estatus de la enfermedad, razón por la que los valores predictivos no son usados para cuantificar la exactitud inherente de la prueba, pero pueden ser utilizados para cuantificar el valor clínico de la prueba.

Hay una relación directa entre los valores predictivos y las probabilidades de clasificación, aunque el conocimiento de la prevalencia de la enfermedad también sea requerido para calcular un valor a partir del otro, por lo que en realidad la distribución conjunta de  $(D, Y)$  requiere tres parámetros. Parametrizaciones naturales serían:

1.  $(TPF, FPF, \rho)$
2.  $(PPV, NPV, \tau)$  donde  $\tau = P[Y = 1]$  es la probabilidad de una prueba positiva

## 2.2. Espacio *ROC*

Los gráficos *ROC* son gráficos bidimensionales en los que la *Fracción de Verdaderos Positivos* (*TPF*) es representada en el eje *Y* y la *Fracción de Falsos Positivos* (*FPF*) es representada en el eje *X*. Muestra las compensaciones relativas entre los beneficios (verdaderos positivos) y los costes (falsos positivos).

Un clasificador discreto es el que produce una salida que representa sólo la etiqueta de la clase. Cada clasificador discreto produce un par  $(FPF, TPF)$  correspondiente a un único punto en el espacio *ROC*. En el gráfico A de la Figura 2.1 se muestran cinco marcadores discretos identificados con las letras A hasta E.

Algunos puntos importantes del espacio *ROC* deben ser destacados, como:

- El punto  $(0,0)$  que representa la estrategia de “nunca emitir una clasificación positiva”. Un clasificador así no comete errores de falsos positivos, pero también no reconoce verdaderos positivos.
- El punto  $(1,1)$  que representa la estrategia opuesta de “emitir incondicionalmente clasificaciones positivas”.
- El punto  $(0,1)$  que representa la clasificación perfecta. En el gráfico A de la figura 2.1, el punto D es un marcador con una actuación perfecta.

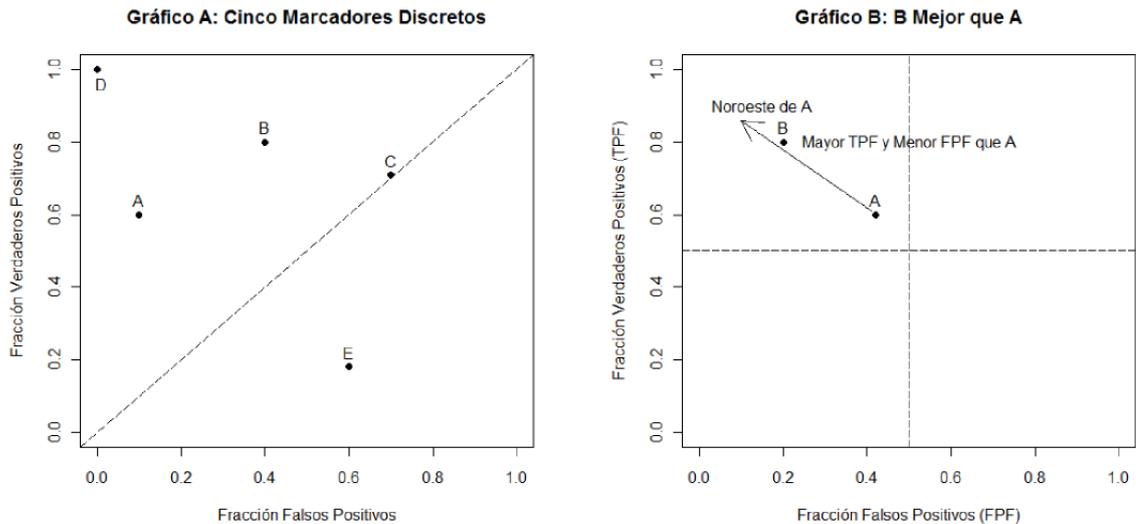


Figura 2.1: Gráficos *ROC* Básicos

Fawcett (2006) apunta que de manera informal, un punto en el espacio *ROC* es mejor que otro si está al noroeste (*alta fracción de verdaderos positivos, baja fracción de falsos positivos o ambas*) del primero (en el gráfico B de la Figura 2.1 el punto B es mejor que el A). Los clasificadores que figuren en el lado izquierdo de un gráfico *ROC*, cerca del eje *X*, podrían ser considerados “conservadores” (clasifican positivamente sólo con una fuerte evidencia, por lo que tienen una baja fracción de falsos positivos y también, a menudo, una baja fracción de verdaderos positivos), en cuanto que los clasificadores que figuren en la parte superior del lado derecho pueden ser considerados “liberales” clasifican positivamente con una débil evidencia, por lo que clasifican casi todos los positivos correctamente, pero a menudo, tienen altas fracciones de falsos positivos. En el gráfico A de la Figura 2.1 A es más conservador que B.

La diagonal  $y = x$  representa la estrategia de “adivinar al azar una clase”. Si un marcador clasifica aleatoriamente “positivo” la mitad de las veces, se puede esperar que obtenga mitad positivos y mitad negativos, produciendo el punto (0.5, 0.5) en el espacio *ROC*; similarmente si clasifica al azar “positivo” 90 % de las veces produciría el punto (0.9, 0.9), es decir un clasificador aleatorio produce un punto *ROC* que se desliza sobre la diagonal (hacia atrás y hacia adelante) basado en la frecuencia con la que adivina la clase positiva. Así, la actuación del marcador C (gráfico A de la Figura 2.1) es virtualmente aleatoria.

Cualquier clasificador que aparezca en el triángulo inferior derecho tiene una actuación o comportamiento peor que la clasificación al azar, por tanto, usualmente este triángulo está vacío en los gráficos *ROC*. Si se invierten los resultados de este clasificador en todos los casos, sus verdaderos positivos se convierten en falsos negativos y sus falsos positivos se convierten en verdaderos negativos. Esto significa que cualquier clasificador que produzca un punto en el triángulo inferior derecho puede invertirse para producir un punto en el triángulo superior izquierdo. En el gráfico A de la Figura 2.1, E actúa peor que un clasificador aleatorio y de hecho es la negación de B. Resumiendo, se puede decir que:

- Cualquier clasificador sobre la diagonal no tiene información sobre la clase.
- Un clasificador debajo de la diagonal tiene información útil, pero aplicada de forma incorrecta.

### 2.3. Curvas en el Espacio *ROC*

Como se mencionó en la sección 2.1 (página 5) hay clasificadores que son diseñados para producir sólo la decisión de la clase, es decir, **Sí** o **No** sobre cada sujeto. Cuando un marcador de este tipo es aplicado al conjunto de sujetos origina una única matriz de confusión que a su vez corresponde a un punto *ROC*. Un clasificador de este tipo produce solamente un único punto en el espacio *ROC*.

Existen otros clasificadores que, de forma natural, producen la probabilidad o puntuación (*score*) del sujeto, que es un valor numérico que representa el grado en que ese sujeto es un miembro de la clase. Estos valores pueden ser probabilidades estrictas, en cuyo caso se opera con teoremas de probabilidad, o pueden ser puntuaciones, en cuyo caso la única propiedad que tienen es que una alta puntuación indica una alta probabilidad.

Los marcadores que producen rangos o puntuaciones pueden ser usados con un punto de corte o umbral de decisión ( $c$ ), obteniendo así un clasificador binario: si el marcador está por encima o es igual al punto de corte produce un **Sí** y si no es así produce un **No**. Utilizando la notación de Pepe (2003) una prueba continua binaria se definiría como:

$$\begin{aligned} & \textit{Positivo si } Y \geq c \\ & \textit{Negativo si } Y < c \end{aligned}$$

Cada punto de corte origina un punto diferente en el espacio *ROC*, que representa un par ( $FPF, TPF$ ) correspondiente a un nivel de decisión determinado. Conceptualmente,

se puede imaginar el punto de corte variando desde  $(-\infty, +\infty)$  y trazando una curva en el espacio *ROC*. Las fracciones de verdaderos y falsos positivos generadas por  $c$  pueden representarse como:

$$TPF(c) = P[Y \geq c/D = 1] \quad (2.10)$$

$$FPF(c) = P[Y \geq c/D = 0] \quad (2.11)$$

La curva *ROC* es todo el conjunto de posibles fracciones de verdaderos positivos y falsos positivos obtenidas por la utilización de diferentes puntos de corte:

$$ROC(.) = \{(FPF(c), TPF(c)), c \in (-\infty, \infty)\} \quad (2.12)$$

Si el punto de corte  $c$  se incrementa ambas fracciones disminuyen y si  $c$  disminuye ambas fracciones aumentan. La curva *ROC* también se puede escribir como:

$$ROC(.) = \{(t, ROC(t)), t \in (0, 1)\} \quad (2.13)$$

donde la función *ROC* representa cada puntuación  $t$  con su correspondiente  $TPF(c)$ , siendo  $c$  el punto de corte correspondiente para  $FPF(c) = t$ .

La Figura 2.2 muestra un ejemplo de “curva” *ROC* de un conjunto de 20 casos, presentados por Fawcett (2006), que son mostrados en la Tabla 2.2. En el gráfico A de la figura se muestran los puntos en el espacio *ROC* generados por los diferentes puntos de corte y en el gráfico B se representa una curva *ROC* estimada a partir de esos puntos discretos.

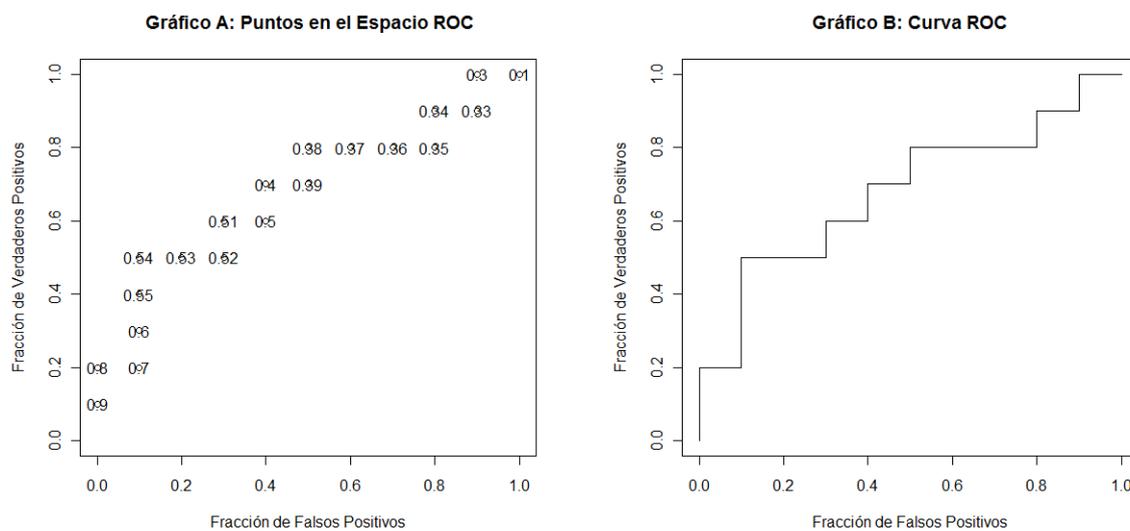
Cualquier curva generada a partir de un conjunto finito de sujetos es una función empírica que se acerca a una verdadera curva en la medida en que el número de individuos tiende a infinito. La función escalonada del gráfico B de la figura 2.2 es obtenida a partir de casos que han sido clasificados por sus puntuaciones y cada punto en el espacio *ROC* es etiquetado con la puntuación del punto de corte que lo origina. Al cambiar el nivel de decisión, la inclusión de un resultado verdadero positivo produce una línea vertical, a la vez que la inclusión de un falso positivo produce una horizontal. Un umbral de decisión de  $+\infty$  produce el punto  $(0, 0)$ . Si el punto de corte desciende a 0,9 el primer caso positivo es clasificado como positivo, produciendo el punto  $(0, 0,1)$  en el espacio *ROC*. En la Tabla 2.3 se muestra la matriz de confusión producida para este nivel de decisión, a partir de la cual se obtienen las dos fracciones utilizadas para obtener el punto de la curva *ROC*  $(0, 0,1)$ :

$$FPF = \frac{0}{10} = 0$$

$$TPF = \frac{1}{10} = 0,1$$

Caso	Clase	Puntuación	Caso	Clase	Puntuación
1	P	0.9	11	P	0.4
2	P	0.8	12	N	0.39
3	N	0.7	13	P	0.38
4	P	0.6	14	N	0.37
5	P	0.55	15	N	0.36
6	P	0.54	16	P	0.35
7	N	0.53	17	P	0.34
8	N	0.52	18	N	0.33
9	P	0.51	19	P	0.30
10	N	0.505	20	N	0.1

Tabla 2.2: Casos

Figura 2.2: Ejemplo Curva *ROC*

<i>Clase Hipotetizada</i>	<i>Clase Verdadera</i>	
	<b>P</b>	<b>N</b>
<b>Si</b>	1	0
<b>No</b>	9	10
<b>Totales Columna</b>	<b>10</b>	<b>10</b>

Tabla 2.3: Matriz de Confusión - Punto de Corte 0.9

Este proceso se repite para cada uno de los niveles de decisión. En la medida en que el punto de corte reduce su valor, la curva sube y se va desplazando a la derecha, terminando en el punto (1, 1) con un punto de corte asociado de 0,1. La disminución del nivel de decisión hace que la curva se mueva del área de los marcadores “conservadores” al área de los “liberales”.

## 2.4. Propiedades de la Curva *ROC*

- La curva *ROC* es una función monótona creciente en (0,1), con  $\lim_{t \rightarrow 0} ROC(t) = 0$  y  $\lim_{t \rightarrow 1} ROC(t) = 1$
- Si se denotan por  $F_D$  y  $F_{\bar{D}}$  las funciones de distribución del resultado del marcador  $Y$  en las poblaciones enferma y no-enferma, la función *ROC* se puede escribir como:

$$ROC(t) = 1 - F_D \left( F_{\bar{D}}^{-1}(1 - t) \right), t \in (0, 1) \quad (2.14)$$

- La curva *ROC* es invariante si los resultados o puntuaciones del marcador son sometidas a transformaciones monótonas estrictamente crecientes.
- La curva *ROC* es la función en (0,1) con  $ROC(0) = 0$  y  $ROC(1) = 1$  con pendiente:

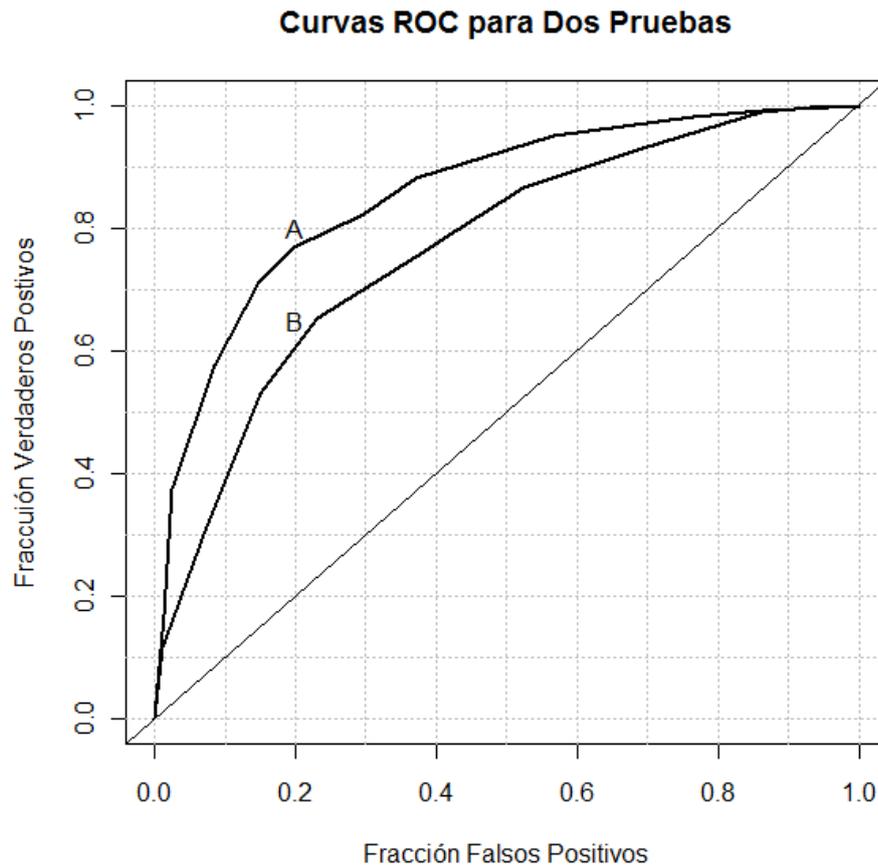
$$\frac{\partial ROC(t)}{\partial(t)} = \frac{f_D \left( F_D^{-1}(t) \right)}{f_{\bar{D}} \left( F_{\bar{D}}^{-1}(t) \right)} \quad (2.15)$$

donde  $f_D$  y  $f_{\bar{D}}$  son las funciones de densidad de los resultados del marcador  $Y$ .

Las demostraciones de estas propiedades se pueden revisar en [Pepe](#) (específicamente [2003](#), pág.68); [Krzanowski y Hand](#) (específicamente [2009](#), pág.20).

La mayoría de las pruebas tienen curvas que se encuentran entre una prueba perfecta y una prueba en que las distribuciones de  $Y$  son las mismas para las poblaciones enferma y

no-enferma y por lo tanto, para cualquier punto de corte  $c$  se tiene que  $TPF(c) = FPF(c)$ . Las mejores pruebas tienen curvas que están cerca de la esquina superior izquierda. En la Figura 2.4 se observa que la prueba A es la mejor de las dos: el correspondiente valor de la Fracción de Verdaderos Positivos para cualquier valor de la Fracción de Falsos Positivos es mayor que el de B.





## Capítulo 3

# Índices Resumen de la Curva *ROC*

La curva *ROC* es un resumen gráfico conveniente de toda la información que podría ser necesaria para una descripción comprensible de un marcador en todo el rango de sus puntos de corte. Sin embargo, es posible que en determinadas circunstancias un gráfico sea difícil de construir o que sea necesario comparar varios clasificadores diferentes, por eso la derivación de resúmenes más simples ha despertado un gran interés. Una especial atención se ha dedicado a la obtención de valores escalares únicos que puedan capturar las características y atributos esenciales de una curva *ROC*, emulando la forma en que algunas medidas de resumen, como la media y la varianza, capturan las características esenciales de los conjuntos de datos.

La esencia de estas medidas o índices es comparar las distribuciones de los valores o puntuaciones del marcador para las poblaciones enferma y no-enferma. Un buen marcador tiende a producir altos valores para la población enferma y bajos valores para la población no-enferma, y será mejor cuanto mayor sea el grado en que estas distribuciones difieren. La curva *ROC* es una manera de mostrar de forma conjunta estas dos distribuciones.

### 3.1. Área Bajo la Curva *ROC*

Probablemente el índice o valor global más ampliamente utilizado es el área bajo la curva *ROC*, comúnmente denotada por *AUC* (*Área Under Curve*). Ha sido estudiada por [Bamber \(1975\)](#); [Hanley y McNeil \(1982\)](#); [Faraggi y Reiser \(2002\)](#), entre otros autores. Como *AUC* es una porción del área de un cuadrado de lado 1 su valor siempre estará entre 0 y 1, sin embargo como una prueba no-informativa (clasificación aleatoria) produce la diagonal entre (0,0) y (1,1), la cual tiene un área de 0,5, ningún clasificador verosímil debe tener

un *AUC* menor que 0,5. La definición formal es:

$$AUC = \int_0^1 ROC(t)dt \quad (3.1)$$

Una interpretación inmediata del *AUC* es considerarla una media de *TPF*, tomada uniformemente sobre todo el rango de las fracciones de falsos positivos en  $(0, 1)$ . Otra interpretación, menos obvia, es considerarla igual a la probabilidad de que los resultados del marcador  $Y$  para un par de individuos, enfermo y sano, seleccionados al azar, estén ordenados correctamente. Si  $Y_D$  e  $Y_{\bar{D}}$  son los resultados del marcador para enfermos y sanos, respectivamente, se tiene que:

$$AUC = P[Y_D > Y_{\bar{D}}] \quad (3.2)$$

La demostración de esta interpretación se puede verificar en [Pepe](#) (específicamente [2003](#), pág.78); [Krzanowski y Hand](#) (específicamente [2009](#), pág.27).

Para explicar la interpretación de esta medida, [Zweig y Campbell \(1993\)](#) ponen el siguiente ejemplo: un área de 0.8 significa que un individuo seleccionado al azar del grupo de enfermos tiene un valor mayor del marcador que el de un individuo elegido al azar del grupo no-enfermo un 80 % de las veces. Esto no significa que un resultado positivo ocurre con una probabilidad de 0.80, ni que un resultado positivo se asocia con la enfermedad el 80 % de las veces.

## 3.2. Puntos y Áreas Parciales

Si una fracción específica de falsos positivos es de interés, por ejemplo  $t_0$  entonces  $ROC(t_0)$  proporciona el índice relevante. Comúnmente, el interés se centra sobre un rango de valores  $(0, t_0)$  de *FPF* y en este caso un índice adecuado es el área parcial bajo la curva ([Thompson y Zucchini, 1989](#)) :

$$pAUC(a, b) = \int_0^{t_0} ROC(t)dt \quad (3.3)$$

## 3.3. Índice de Youden

La *Sensibilidad* ( $Se$ ) y la *Especificidad* ( $Sp$ ) de un marcador son usualmente utilizadas de forma simultánea como una medida conjunta del comportamiento del marcador o prueba diagnóstica. Esto se debe a que son complementarias: en general, al disminuir  $c$  la fracción

de verdaderos positivos ( $TPF$ ) aumenta, dando lugar a que la fracción de verdaderos negativos ( $TNF$ ) disminuya. En esta situación, hay que alcanzar un compromiso aceptable. Una de las soluciones propuestas es seleccionar el punto de corte que maximice la diferencia entre las fracciones de verdaderos positivos y falsos positivos. El valor máximo de esta cantidad es el Índice de Youden ( $YI$ ) y el punto de corte  $c$  en el punto de la curva  $ROC$  correspondiente a este índice es, a menudo, seleccionado como el punto de corte óptimo del marcador. Por esta razón, aunque el  $AUC$  es el índice global más comúnmente usado el  $YI$  es también frecuentemente usado en la práctica médica (Aoki y otros, 1997; Schisterman y otros, 2005; Bao y otros, 2008; Matoba y otros, 2008).

Youden (1950) propuso un índice en el que se obtiene un único valor “mezclando” *Sensibilidad* y *Especificidad*:

$$YI = \text{Sensibilidad} + \text{Especificidad} - 1 \quad (3.4)$$

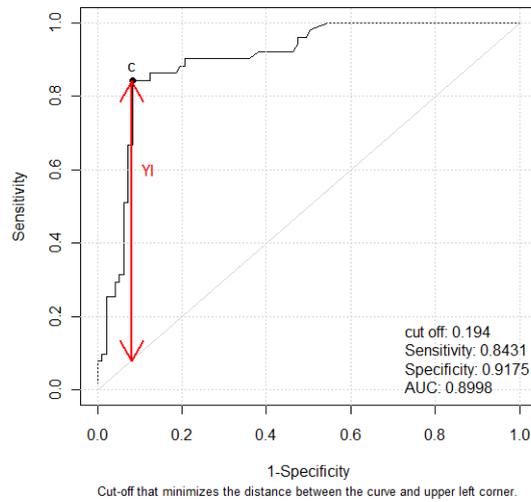
Como el  $YI$  maximiza las diferencias entre  $TPF$  y  $FPF$  (sobre todos los puntos de corte), puede escribirse de la siguiente manera:

$$YI = \max(\text{Sensibilidad} + \text{Especificidad} - 1) \quad (3.5)$$

En una curva  $ROC$ ,  $YI$  es la distancia vertical máxima desde la curva  $ROC$  a la diagonal principal (marcador no-informativo), siendo  $c_{YI}$  el punto de corte “óptimo”. La interpretación intuitiva del  $YI$  es que es el punto más alejado de esa diagonal (ver Figura 6.5).

### 3.3.1. Características del $YI$

1. El rango posible de valores es de 0 a 1. Se espera que el marcador muestre una proporción de resultados positivos mayor para el grupo de enfermos que para el grupo de no-enfermos.
2. Toma el valor cero si el marcador da la misma proporción de positivos para ambos grupos, indicando que un marcador como éste no es útil. Ver Tabla 3.1.
3. Alcanza valor 1 sólo cuando las dos fracciones  $FPF$  y  $FNF$  no están presentes. Si sólo unos de los dos tipos de errores está presente, el índice es controlado por ese error. Ver Tabla 3.2.
4. Es independiente de los tamaños relativos de los grupos no-enfermo y enfermo. Ver Tablas 3.3 y 3.4.

Figura 3.1: YI y Punto de Corte  $c$  Asociado

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	40	10	50
-	8	2	10
$YI = \frac{40}{50} + \frac{2}{10} - 1 = 0$			

Tabla 3.1: Misma Proporción Ambos Grupos

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	50	0	50
-	10	40	50
$YI = \frac{50}{50} + \frac{40}{50} - 1 = 0,8$			

Tabla 3.2: Índice Controlado por  $FPF$  y  $FNF$ 

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	90	10	100
-	10	40	50
$YI = \frac{90}{100} + \frac{40}{50} - 1 = 0,7$			

Tabla 3.3: Independencia Tamaños Relativos Grupos-Ejemplo A

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	9	1	10
-	10	40	50
$YI = \frac{9}{10} + \frac{40}{50} - 1 = 0,7$			

Tabla 3.4: Independencia Tamaños Relativos Grupos-Ejemplo B

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	23	2	25
-	1	9	10
$YI = \frac{23}{25} + \frac{9}{10} - 1 = 0,82$			

Tabla 3.5: Independencia Tamaños Absolutos Grupos-Ejemplo A

5. Es independiente de los tamaños absolutos de los grupos no-enfermo y enfermo. Ver Tablas 3.5 y 3.6.

<i>Diagnóstico</i>			
<i>Grupo</i>	+	-	Total
+	230	20	250
-	10	90	100
$YI = \frac{230}{250} + \frac{90}{100} - 1 = 0,82$			

Tabla 3.6: Independencia Tamaños Absolutos Grupos-Ejemplo B

En este trabajo, el interés principal está focalizado en el estudio del Índice de Youden como medida que proporciona un criterio para elegir el punto de corte óptimo, como el punto que maximiza dicho índice. En la actualidad existen varias aproximaciones para estimar este índice y su punto de corte asociado, las cuales serán comentadas en el siguiente capítulo.



## Capítulo 4

# Métodos de Estimación del Índice de Youden

Existen diversos enfoques para estimar la curva *ROC* y consecuentemente su *YI* asociado. Faraggi y Reiser (2002) examinaron varios enfoques paramétricos y no paramétricos para estimar el *AUC* y Fluss y otros (2005) revisaron estos procedimientos, estudiando su aplicación en la estimación del *YI* y su punto de corte asociado ( $c^*$ ). Estos procedimientos difieren debido a las varias alternativas utilizadas para estimar las funciones de distribución de los valores del marcador, en las muestras tomadas de los grupos de enfermos y no-enfermos. Usando estas funciones de distribución estimadas se pueden estimar *Se* y *Sp* para cualquier punto de corte y por tanto el Índice de Youden.

El primer enfoque paramétrico asume que los valores del marcador para ambas poblaciones (enferma y no-enferma) siguen una distribución Normal y estima el *YI* usando el método paramétrico estándar. El segundo enfoque paramétrico asume que existe una transformación monótona, por la que los valores transformados del marcador siguen una distribución Normal. Después de estimar la transformación y aplicarla a los datos, se utiliza la teoría Normal. El primer método no paramétrico, referido como método Empírico estima la función de distribución con la función de distribución empírica de la muestra. Es muy popular debido a su simplicidad. El segundo método no paramétrico usa una función de suavización tipo núcleo sobre la función de distribución de la muestra. Tiene la ventaja de proporcionar una estimación suave de la función de distribución. A partir de la estimación del *YI* para cada uno de estos procedimientos se obtiene la correspondiente estimación del punto de corte óptimo asociado  $c^*$ . Los distintos métodos de estimación proporcionarán estimaciones diferentes de *YI* y de  $c^*$  para los mismos datos, por lo que una comparación

de la efectividad de sus resultados debe ser llevada a cabo.

Supóngase que se tienen los resultados de un marcador pertenecientes a dos muestras aleatorias, provenientes de una población enferma y una no-enferma ( $x_1, x_2, \dots, x_m$  y  $y_1, y_2, \dots, y_n$ ), denotándose sus funciones de distribución por  $G_D$  y  $F_{\bar{D}}$  respectivamente. Para cualquier punto de corte dado  $c$ , las medidas de *Sensibilidad* y *Especificidad* pueden calcularse a partir de las distribuciones como  $Se(c) = 1 - G_D(c)$ ,  $Sp = F_{\bar{D}(c)}$  y mediante estas medidas obtener su  $YI$  asociado:

$$YI = \max_c \{Se(c) + Sp(c) - 1\} = \max_c \{F_{\bar{D}}(c) - G_D(c)\} \quad (4.1)$$

El valor  $c$  para el que se alcanza este máximo será considerado el punto de corte óptimo  $c^*$ . Para estimar el  $YI$  se sustituirán en la ecuación 4.1 las funciones de distribución por sus estimaciones ( $\hat{G}_D, \hat{F}_{\bar{D}}$ ), es decir:

$$\hat{YI} = \max_c \{\hat{F}_{\bar{D}}(c) - \hat{G}_D(c)\} \quad (4.2)$$

#### 4.1. Métodos Paramétricos

Sean  $X$  e  $Y$  los resultados de un marcador para sujetos enfermos y no-enfermos respectivamente. Un enfoque paramétrico simple para estimar  $YI$  es asumir que tanto  $X$  como  $Y$  tienen una distribución Normal con varianzas diferentes o iguales:  $X \sim (\mu_D, \sigma_D^2)$  y  $Y \sim (\mu_{\bar{D}}, \sigma_{\bar{D}}^2)$ . Sin perder generalidad se asume que  $\mu_D > \mu_{\bar{D}}$ . Consecuentemente:

$$Se(c) = 1 - \phi\left(\frac{c - \mu_D}{\sigma_D}\right),$$

$$Sp(c) = 1 - \phi\left(\frac{c - \mu_{\bar{D}}}{\sigma_{\bar{D}}}\right),$$

$$S(c) = \phi\left(\frac{c - \mu_{\bar{D}}}{\sigma_{\bar{D}}}\right) - \phi\left(\frac{c - \mu_D}{\sigma_D}\right) \text{ y}$$

$$YI \equiv \max_c \{S(c)\} \quad (4.3)$$

donde  $\phi$  es la función de distribución Normal estándar. Para llevar a cabo la maximización en la ecuación 4.3 se calcula la primera derivada de  $S(c)$ , se iguala a cero y se resuelve la ecuación cuadrática resultante. La raíz

$$c^* = \frac{(\mu_D \sigma_D^2 - \mu_{\bar{D}} \sigma_{\bar{D}}^2) - \sigma_{\bar{D}} \sigma_D \sqrt{(\mu_{\bar{D}} - \mu_D)^2 + (\sigma_D^2 - \sigma_{\bar{D}}^2) \log(\sigma_{\bar{D}}^2 / \sigma_D^2)}}{(\sigma_{\bar{D}}^2 - \sigma_D^2)} \quad (4.4)$$

proporcionará el máximo. Cuando se asume igualdad de varianzas,  $\sigma_{\bar{D}} = \sigma_D = \sigma^2$ ,  $c^* = (m\mu_{\bar{D}} + \mu_D)/2$ . Sustituyendo  $c^*$  en la ecuación 4.3 se obtiene la solución:

$$YI = \phi\left(\frac{c^* - \mu_{\bar{D}}}{\sigma_{\bar{D}}}\right) - \phi\left(\frac{c^* - \mu_D}{\sigma_D}\right) \quad (4.5)$$

Sustituyendo en las fórmulas 4.4 y 4.5 los valores desconocidos de los parámetros  $\mu_D$ ,  $\mu_{\bar{D}}$ ,  $\sigma_D$  y  $\sigma_{\bar{D}}$  por sus correspondientes medias y desviaciones típicas muestrales se obtienen las estimaciones de  $YI$  y  $c^*$ . A la estimación de  $YI$  y  $c^*$  obtenida de esta manera se le denominará Método Normal ( $N$ ).

Sin embargo, en muchas situaciones asumir que los valores del marcador siguen una distribución Normal es cuestionable. Un enfoque menos restrictivo es asumir que existe alguna transformación monótona  $t(\cdot)$ , tal que  $t(X)$  y  $t(Y)$  están normalmente distribuidos. Después de aplicar la transformación a los datos de la muestra se usa  $N$  para estimar  $(c_t^*, YI_t)$  para los datos transformados. Sigue la estimación de los datos de la muestra original con:  $\hat{Y}I = \hat{Y}I_t$  y  $\hat{c}^* = t^{-1}(\hat{c}_t^*)$ . Para usar este método se debe definir  $t(\cdot)$ . Algunos autores (Zou y Hall, 2000; Faraggi y Reiser, 2002; Schisterman y otros, 2004; Molodianovitch y otros, 2006) han usado la transformación de potencia de tipo Box-Cox, la cual consideran útil en una amplia variedad de situaciones:

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0; \\ \log y & \text{si } \lambda = 0 \end{cases} \quad (4.6)$$

Asumiendo que  $y^\lambda$  y  $x^\lambda$  siguen la distribución Normal el valor de  $\lambda$  puede ser estimado por el método de Máxima Verosimilitud. Usar el resultado de  $\hat{\lambda}$  para transformar los datos y aplicar después el método Normal es el denominado método de la Transformación Normal ( $TN$ ).

## 4.2. Métodos No Paramétricos

El enfoque no paramétrico más simple usa las funciones de distribución empíricas de  $X$  e  $Y$ . La función empírica es conocida por ser un estimador consistente (Vilar, 1991). En el caso de  $X$  e  $Y$  pueden ser escritas como:

$$\hat{G}_D(c) = \frac{1}{m} \sum_{i=1}^m I(x_i \leq c), \quad \hat{F}_{\bar{D}}(c) = \frac{1}{n} \sum_{i=1}^n I(y_i \leq c)$$

donde:

$$I(u \leq c) = \begin{cases} 1 & \text{si } u \leq c \\ 0 & \text{si } u > c \end{cases}$$

es la función indicadora.

Por tanto la estimación del Índice de Youden resulta en:

$$\hat{YI} = \max_c \left\{ \hat{F}_{\bar{D}}(c) - \hat{G}_H(c) \right\} \quad c \in \{x_1, \dots, x_m, y_1, \dots, y_n\} \quad (4.7)$$

Para estimar  $c^*$  se utilizó la observación donde se alcanzó el máximo. Este procedimiento fue denotado por *EMP*. El método Núcleo ( $K$ ) fue el otro método de estimación no paramétrico empleado, el cual proporciona una mayor suavización de la función de distribución empírica. En este trabajo fue utilizado el núcleo Gaussiano, pero se podrían haber utilizado otras funciones núcleo:

$$\hat{F}_{\bar{D}}(t) = \frac{1}{n} \sum_{i=1}^n \phi\left(\frac{t-y_i}{h_y}\right), \quad \hat{G}_D(t) = \frac{1}{m} \sum_{i=1}^m \phi\left(\frac{t-x_i}{h_x}\right)$$

Más importante que el núcleo empleado es la selección del parámetro ventana de suavizado. En este caso para controlar la suavización fueron utilizados los anchos de ventana siguientes:

$$h_y = (0,9) \min \{s_y, iqr_y/1,34\} (n^{-0,2}) \text{ y}$$

$$h_x = (0,9) \min \{s_x, iqr_x/1,34\} (m^{-0,2})$$

donde:

- $s_y = \left(\frac{1}{n-1} \sum_1^n (y_i - \bar{y})^2\right)^{0,5}$  y  $s_x = \left(\frac{1}{m-1} \sum_1^m (x_i - \bar{x})^2\right)^{0,5}$  son las desviaciones típicas de las muestras no-enferma y enferma respectivamente.
- $iqr_y$  e  $iqr_x$  son los rangos intercuantílicos de las dos muestras.

El ancho de ventana utilizado ha sido recomendado por [Silverman \(1986\)](#) como válido para un amplio rango de densidades. [Faraggi y Reiser \(2002\)](#) compararon diferentes procedimientos para la selección de los anchos de ventanas en el contexto de estimación del *AUC* y encontraron que procedimientos más complejos no se tradujeron en ninguna mejora. El *YI* estimado por el método Núcleo es:

$$\hat{YI} = \max_c \left\{ \hat{F}_{\bar{D}}(c) - G_{\hat{D}}(c) \right\} \quad (4.8)$$

## Capítulo 5

# Estudios de Simulación

Para evaluar y comparar las propiedades de los estimadores de los  $YI$  obtenidos por los métodos comentados en el capítulo 4 (página 23) y sus correspondientes puntos de corte, fueron realizados varios estudios de simulación. Las estimaciones obtenidas fueron comparadas en términos del sesgo y del error cuadrático medio, concretamente de la raíz de dicho error (“*Root Mean Square Error*”, RMSE). La suposición de Normalidad frecuentemente no es apropiada para los valores de un marcador diagnóstico. Pero se consideró aquí el método  $N$  para examinar su robustez y proporcionar una base de comparación para el procedimiento  $TN$ .

Las simulaciones cubrieron una gran variedad de formas distribucionales (simétricas, sesgadas, bimodales), que a menudo se observan en datos reales. Estas formas de las distribuciones son similares a las utilizadas por Faraggi y Reiser (2002) y Fluss y otros (2005). En las Tablas 5.1 y 5.2 se presentan los parámetros de los modelos usados en los estudios de simulación.

Distribución $YI$	$\mu_{\bar{D}}$	$\sigma_{\bar{D}}$	$\sigma_{\bar{D}}^2$	$\mu_D$ correspondientes a $YI$			
				0.2	0.4	0.6	0.8
Normal Varianzas Iguales	6.5	0.25	0.25	6.753	7.024	7.342	7.782
Normal Varianzas Diferentes	6.5	0.09	0.25	6.618	6.874	7.144	7.506
Normal <sup>-1/3</sup>	3.5	0.09	0.25	3.382	3.126	2.856	2.494
Lognormal	2.5	0.09	0.25	2.618	2.874	3.144	3.506
	$v_{\bar{D}}$	$r_{\bar{D}}$	$r_D$	$v_D$			
Gamma	0.5	2	2	0.344	0.23	0.143	0.072

Tabla 5.1: Parámetros de los Modelos Usados en Estudios de Simulación

Mixturas <i>YI</i>	$P_D$	$P_{\bar{D}}$	$\mu_{\bar{D}1}$	$\sigma_{\bar{D}1}^2$	$\mu_{\bar{D}2}$	$\sigma_{\bar{D}2}^2$	$\sigma_{D1}^2$	$\sigma_{D2}^2$	$\mu_{D1}, \mu_{D2} = \mu_{D1} + 4$			
									0.2	0.4	0.6	0.8
<i>Mix 1</i>	0.5	1	10	1	0	0	1	5	7.64	9.98	11.12	12.16
<i>Mix 2</i>	0.8	1	10	1	0	0	1	5	10.22	10.85	11.53	12.44
<i>Mix 3</i>	0.5	0.5	10	1	13	1	1	5	10.28	12.39	13.58	14.74
<i>Mix 4</i>	0.5	0.5	10	1	13	1.5	1	5	10.55	12.24	13.58	14.91
<i>Mix 5</i>	0.8	0.5	10	1	13	1	1	5	10.28	12.39	13.58	14.74
<i>Mix 6</i>	0.8	0.5	10	1	13	1.5	1	5	10.55	12.24	13.58	14.91

Tabla 5.2: Parámetros de las Mixturas Usadas en los Estudios de Simulación

Fueron utilizados diferentes valores del *YI* (0.2, 0.4, 0.6 y 0.8) con el objetivo de abarcar un rango de valores de este índice observados en la práctica, utilizando diferentes tamaños muestrales:  $n = m = 20, 50, 100$ ,  $(m, n) = (20, 60), (50, 100)$  y  $(100, 300)$ . Las medias de la población enferma fueron ajustadas a los valores del *YI* anteriormente mencionados fijadas las varianzas en ambas poblaciones y las medias en la población sana. Para estimar el sesgo y el *RMSE* se realizaron 1000 simulaciones de cada escenario para cada población (enferma y no-enferma). Para cada conjunto de datos fueron estimados el *YI* y su punto de corte asociado  $c^*$ , utilizando los cuatro métodos mencionados anteriormente. Estos valores fueron usados para calcular las estimaciones del sesgo y del *RMSE* de cada método. Las formas de las distribuciones se pueden observar en las Figuras 5.1, 5.2, 5.3 y 5.4<sup>1, 2</sup>,

## 5.1. Resultados de los Estudios de Simulación

En esta sección se presentan el sesgos y el *RMSE* de las estimaciones obtenidas en los escenarios elaborados, atendiendo a la selección realizada de los valores de *YI*, sus puntos de corte asociados y los tamaños de muestra utilizados.

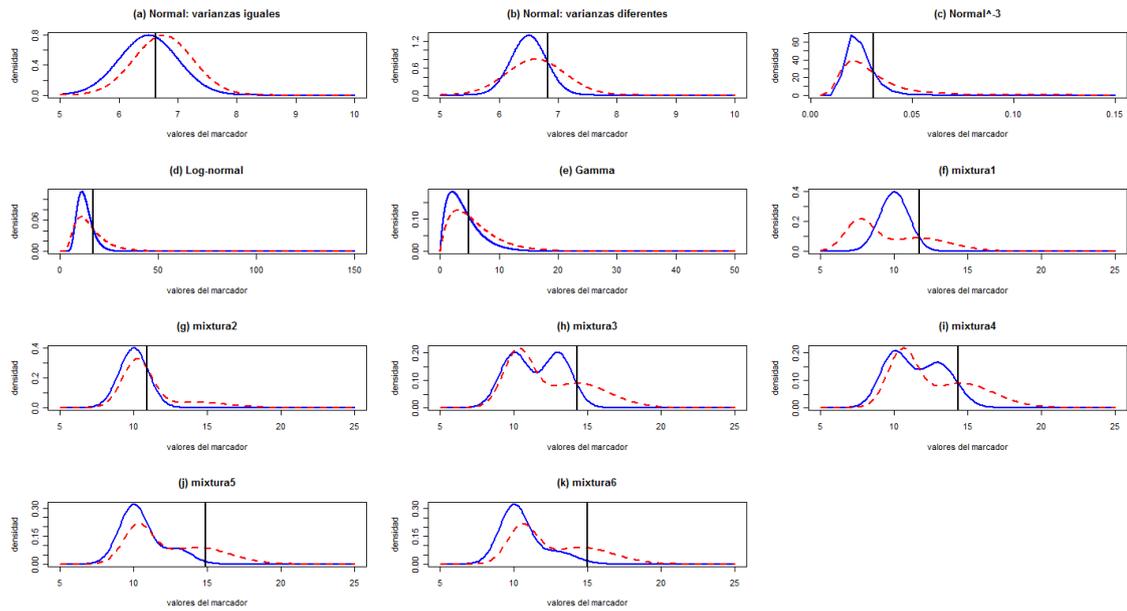
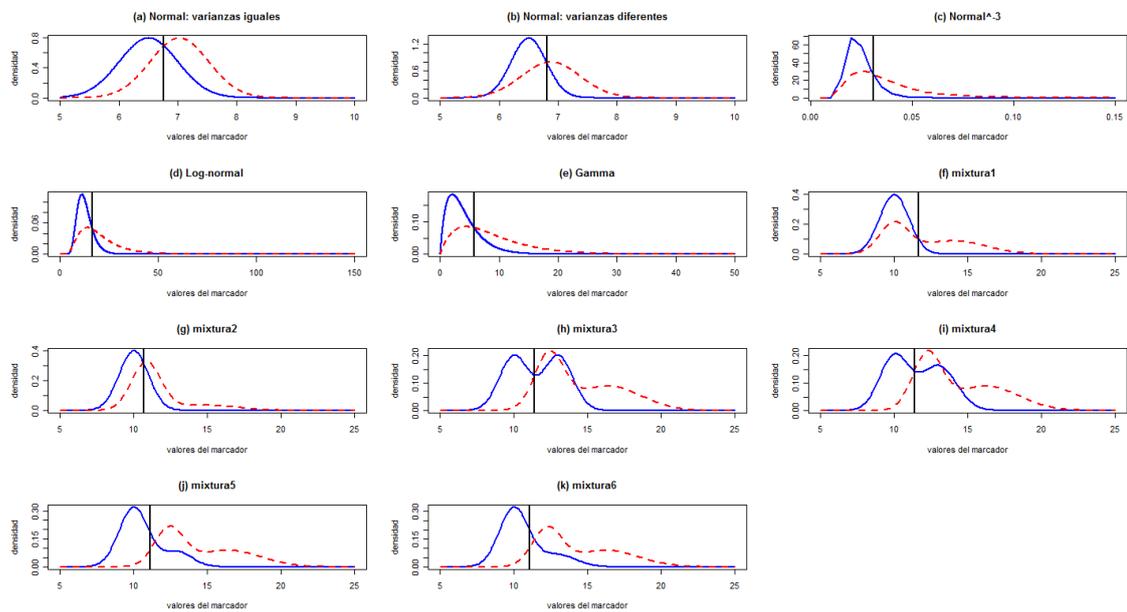
### 5.1.1. Índice de Youden

#### Simulaciones con Distribuciones Normales - Varianzas y Tamaños Muestrales Iguales

En la Tabla 5.3 se presentan los resultados obtenidos para muestras de poblaciones normales con varianzas y tamaños muestrales iguales. El valor esperado de la población enferma

<sup>1</sup>Línea Sólida: Sanos, Línea Discontinua: Enfermos, Línea Vertical: Punto de Corte Óptimo

<sup>2</sup>Ver definición de las Mixturas en la Tabla 5.2.

Figura 5.1: Distribuciones Usadas en el Estudio de Simulación con  $YI = 0,2$ Figura 5.2: Distribuciones Usadas en el Estudio de Simulación con  $YI = 0,4$

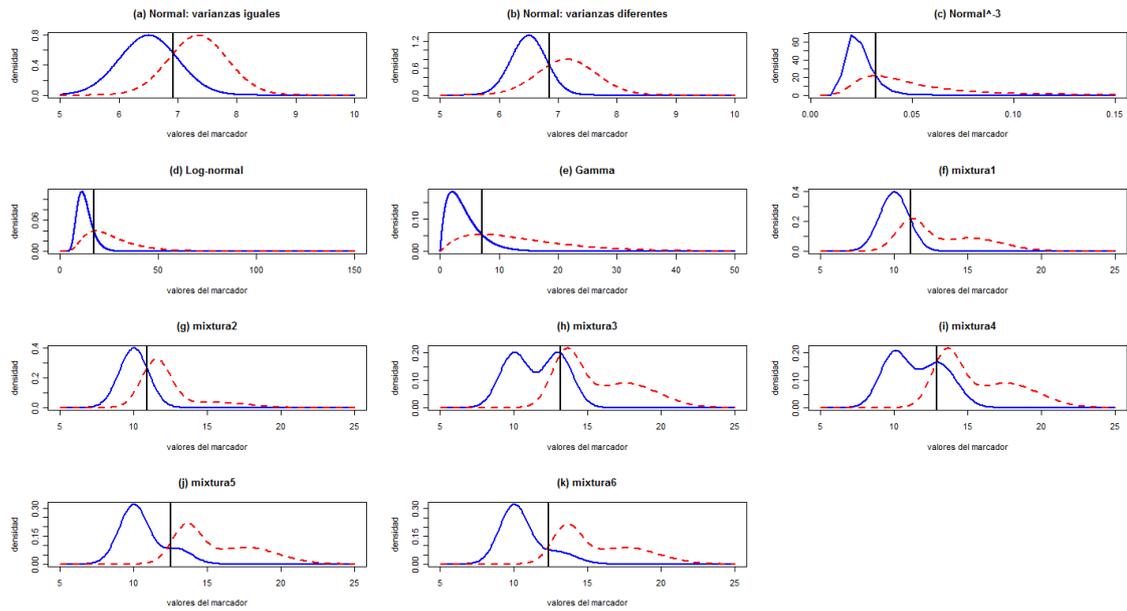


Figura 5.3: Distribuciones Usadas en el Estudio de Simulación con  $YI = 0,6$

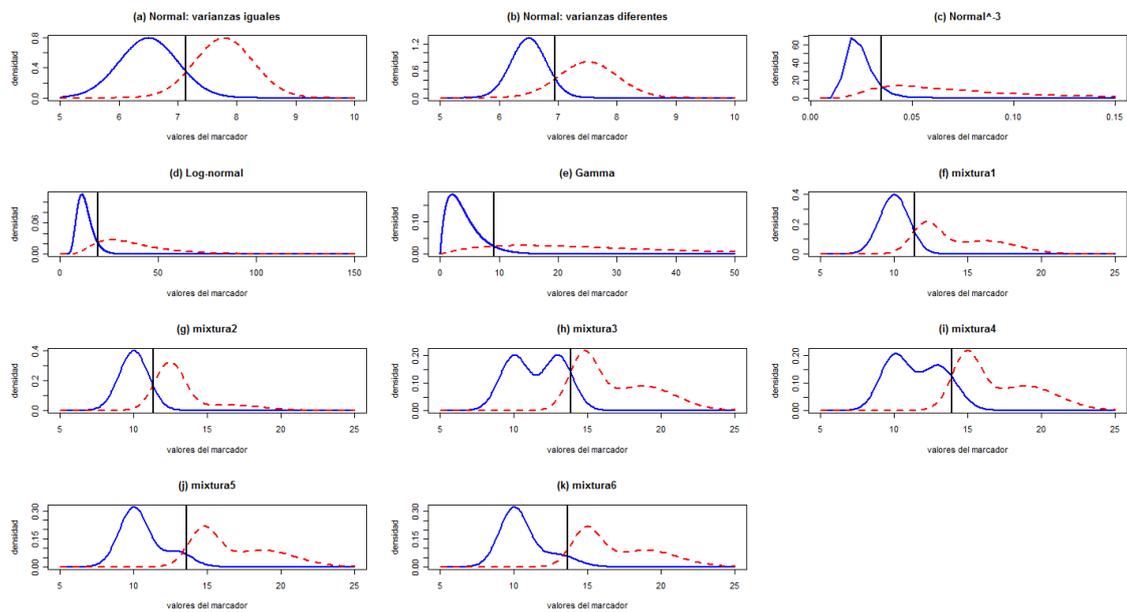


Figura 5.4: Distribuciones Usadas en el Estudio de Simulación con  $YI = 0,8$

( $\mu_D$ ) es elegido para corresponder al valor de  $YI$  (ver capítulo 5, tercer párrafo). Los cuatro métodos muestran una tendencia decreciente de los valores de sesgo y  $RMSE$ , en la medida que el tamaño de las muestras aumenta, siendo  $K$  el único método que presentó sesgos negativos. Al examinar los correspondientes sesgos mínimos de cada método, se nota que el del procedimiento  $K$  es el menor (**-0.028**) para  $YI = 0,8$  y tamaño muestral 50. Al comparar los procedimientos (para los valores de  $YI$  seleccionados en cada uno de los tamaños muestrales) se observa que:

- El procedimiento  $EMP$  fue, en términos de sesgo, el que peor comportamiento mostró, con valores sustancialmente altos, en tanto que  $K$  presentó los valores mínimos en todas las situaciones, con excepción de la combinación  $YI = 0,2$  y  $n = m = 20$ .
- Para  $YI = 0,2$ , independientemente del tamaño de la muestra,  $K$  aparece como el método que exhibe el menor  $RMSE$ , coincidiendo para  $n = m = 100$  con  $N$  y  $TN$ , mientras que para  $YI = 0,8$  fueron  $TN$  (para muestras de tamaño 20) y  $N$  para el resto de los tamaños de muestra empleados, los que presentaron menores valores de  $RMSE$ .
- Para todas las combinaciones de valores de  $YI$  y tamaños de muestra usados  $EMP$  mostró los valores más altos del  $RMSE$ .
- No parece haber grandes diferencias entre los sesgos y los  $RMSE$  obtenidos por  $N$  y  $TN$ . Esto es debido a que se está suponiendo una distribución Normal para el marcador diagnóstico.

Estos resultados coinciden con los obtenidos por [Fluss y otros \(2005\)](#).

### Simulaciones con Distribuciones Normales - Varianzas Iguales y Tamaños Muestrales Diferentes

Los resultados alcanzados en este escenario son semejantes a los presentados en la Tabla 5.3: los sesgos de los cuatro procedimientos tienden a decrecer cuando los tamaños de las muestras aumentan y  $K$  es el único que presenta sesgos negativos. En el examen de los respectivos sesgos mínimos de cada método se observa que  $N$  y  $TN$  muestran los valores menores para  $YI=0,4$  y  $(m, n)=(100, 300)$ , mientras que  $EMP$ , con  $YI=0,2$  y  $(m, n)=(20, 60)$  exhibe el máximo valor. Al cotejar los sesgos de los procedimientos, para todo el rango de valores de  $YI$  con  $n$  fijo, se percibe que  $K$  muestra el valor menor en todas las situaciones, mientras que  $EMP$  produce el máximo, alcanzando también los mayores  $RMSEs$  (ver Tabla 5.4).

Escenario Normal con Varianzas y Tamaños Muestrales Iguales									
	<b>YI</b>	<b>Sesgo</b>				<b>RMSE</b>			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.032	0.019	0.014	0.006	0.117	0.119	0.105	0.076
	<b>TN</b>	0.035	0.025	0.020	0.011	0.121	0.122	0.106	0.075
	<b>EMP</b>	0.140	0.112	0.090	0.061	0.187	0.165	0.140	0.099
	<b>K</b>	0.034	0.002	-0.016	-0.026	0.116	0.116	0.108	0.088
$(m = n = 50)$	<b>N</b>	0.010	0.006	0.004	0.002	0.078	0.076	0.067	0.049
	<b>TN</b>	0.011	0.009	0.008	0.004	0.079	0.077	0.068	0.049
	<b>EMP</b>	0.082	0.064	0.051	0.035	0.115	0.103	0.088	0.065
	<b>K</b>	0.008	-0.012	-0.024	-0.028	0.077	0.078	0.074	0.061
$(m = n = 100)$	<b>N</b>	0.004	0.003	0.004	0.002	0.055	0.052	0.046	0.033
	<b>TN</b>	0.005	0.005	0.005	0.004	0.055	0.053	0.046	0.033
	<b>EMP</b>	0.053	0.043	0.035	0.026	0.080	0.072	0.062	0.046
	<b>K</b>	0.001	-0.013	-0.020	-0.020	0.055	0.055	0.052	0.042

Tabla 5.3: Sesgo y RMSE para Estimadores de  $YI$ :  $Y \sim N(6,5;0,25)$ ,  $X \sim N(\mu_D;0,25)$

### Simulaciones con Distribuciones Normales - Varianzas Diferentes y Tamaños Muestrales Iguales

Se mantiene la similitud de los resultados conseguidos en este escenario con los observados en los anteriores. En el análisis de los correspondientes sesgos mínimos, se observa que el método  $K$  muestra el menor valor con  $YI = 0,8$  y  $n = 50$ , en cuanto que  $EMP$  exhibe el valor máximo para  $YI = 0,8$  y  $n = 100$ . Comparados los métodos, dentro de un tamaño muestral fijo para el rango de valores de  $YI$ , se nota que  $K$  continúa siendo el procedimiento que obtiene el menor sesgo y  $EMP$  el mayor, mostrando también, este último, los mayores  $RMSEs$  (ver Tabla 5.5).

### Simulaciones con Distribuciones Asimétricas

Para datos normales la similitud de los resultados proporcionados por  $N$  y  $TN$  no es sorprendente, pero si las distribuciones son asimétricas se deberían esperar mayores diferencias. Con el objetivo de comparar los resultados de este trabajo con los obtenidos por Fluss y otros (2005) fueron generadas muestras de poblaciones normales (con varianzas diferentes) elevadas a la potencia  $-3$ . Los sesgos y  $RMSEs$  correspondientes se presentan en la Tabla 5.6. Para este escenario altamente asimétrico el método  $N$ , el cual asume

Escenario Normal con Varianzas Iguales y Tamaños Muestrales Diferentes									
	<b>YI</b>	<b>Sesgo</b>				<b>RMSE</b>			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = 20, n = 60)$	<b>N</b>	0.021	0.013	0.010	0.004	0.098	0.097	0.085	0.062
	<b>TN</b>	0.023	0.016	0.013	0.007	0.100	0.098	0.086	0.061
	<b>EMP</b>	0.113	0.089	0.073	0.053	0.150	0.133	0.114	0.085
	<b>K</b>	0.020	-0.006	-0.020	-0.025	0.093	0.095	0.089	0.073
$(m = 50, n = 150)$	<b>N</b>	0.008	0.005	0.005	0.002	0.063	0.061	0.053	0.039
	<b>TN</b>	0.009	0.007	0.007	0.004	0.064	0.061	0.054	0.039
	<b>EMP</b>	0.066	0.053	0.042	0.030	0.094	0.085	0.072	0.054
	<b>K</b>	0.005	-0.011	-0.020	-0.022	0.062	0.064	0.060	0.049
$(m = 100, n = 300)$	<b>N</b>	0.000	-0.001	0.000	0.000	0.044	0.042	0.037	0.027
	<b>TN</b>	0.001	0.000	0.001	0.001	0.044	0.042	0.037	0.027
	<b>EMP</b>	0.041	0.032	0.025	0.019	0.063	0.055	0.047	0.037
	<b>K</b>	-0.003	-0.014	-0.020	-0.020	0.045	0.045	0.043	0.036

Tabla 5.4: Sesgo y RMSE para Estimadores de  $YI$ :  $Y \sim N(6,5;0,25)$ ,  $X \sim N(\mu_D;0,25)$ 

normalidad, es el que presenta el peor comportamiento en términos de sesgo y este sesgo no disminuye al aumentar el tamaño de la muestra. Los sesgos correspondientes a los procedimientos  $TN$  y  $K$  son los menores y no parece que haya grandes diferencias entre los mismos, disminuyendo en la medida que el tamaño de las muestras aumenta. Al comparar los métodos se nota que:

- Para los tamaños de muestra 20, el procedimiento  $TN$  es el que tiene menores sesgos para los valores 0.2 y 0.4 del  $YI$ .
- El método  $K$ , con la excepción del caso de  $YI = 0,8$ , presenta sesgos menores que el método  $N$ .
- Los valores del  $RMSE$  del procedimiento  $N$  para los valores 0.2 y 0.4 del  $YI$  son los más altos, independientemente del tamaño de la muestra, pero para  $YI = 0,6$  con  $n \leq 50$  es ligeramente superado por  $TN$ .

Estos resultados son similares a los obtenidos por [Fluss y otros \(2005\)](#). Siendo las conclusiones semejantes para los casos de datos obtenidos de una distribución log-normal y los obtenidos de una distribución gamma se decidió no presentar los resultados por razones de simplicidad.

Escenario Normal con Varianzas Diferentes y Tamaños Muestrales Iguales									
	<b>YI</b>	<b>Sesgo</b>				<b>RMSE</b>			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.029	0.015	0.012	0.006	0.096	0.114	0.103	0.076
	<b>TN</b>	0.027	0.014	0.014	0.009	0.097	0.116	0.105	0.076
	<b>EMP</b>	0.111	0.102	0.083	0.057	0.162	0.156	0.134	0.097
	<b>K</b>	0.022	0.003	-0.015	-0.025	0.105	0.111	0.105	0.086
$(m = n = 50)$	<b>N</b>	0.010	0.005	0.004	0.001	0.062	0.072	0.065	0.048
	<b>TN</b>	0.008	0.004	0.005	0.003	0.063	0.074	0.067	0.049
	<b>EMP</b>	0.063	0.057	0.046	0.033	0.098	0.098	0.086	0.063
	<b>K</b>	0.006	-0.009	-0.022	-0.026	0.066	0.074	0.072	0.060
$(m = n = 100)$	<b>N</b>	0.004	0.003	0.003	0.002	0.045	0.049	0.044	0.033
	<b>TN</b>	0.003	0.003	0.004	0.003	0.046	0.050	0.045	0.033
	<b>EMP</b>	0.040	0.038	0.033	0.025	0.066	0.067	0.059	0.045
	<b>K</b>	0.002	-0.010	-0.018	-0.020	0.047	0.051	0.050	0.041

Tabla 5.5: Sesgo y RMSE para Estimadores de  $YI$ :  $Y \sim N(6,5;0,09)$ ,  $X \sim N(\mu_D;0,25)$

### Simulaciones con Mixturas de Distribuciones Normales

Con el propósito de considerar escenarios adicionales que no hubieran sido obtenidos a partir de transformaciones de la familia Box-Cox, fueron examinadas mixturas de dos distribuciones normales, que dan lugar a diferentes grados de bimodalidad. Los resultados obtenidos para la Mixtura 1 están presentados en la Tabla 5.7. Al examinar los sesgos, dentro de cada método, se observa que todos los procedimientos obtuvieron su correspondiente sesgo mínimo en la combinación  $YI=0.2$  y  $n = 100$ . Al comparar los métodos se observa:

- El menor sesgo lo presenta  $K$ , independientemente del tamaño muestral, y el mayor sesgo  $EMP$ , con dos excepciones:  $YI=0.2$  y  $n = 20$ ;  $YI=0.2$  y  $n = 50$ .
- Los valores más altos de  $RMSE$  los presenta también el método  $EMP$ , excepto para  $YI = 0,2$  independientemente de  $n$ .
- Los métodos  $N$  y  $TN$  muestran un patrón de sesgos similar.

La Tabla 5.8 contiene los resultados para la Mixtura 2. En la revisión de los sesgos, dentro de cada método, se nota que todos los procedimientos obtienen su respectivo sesgo mínimo para  $YI=0.8$ , siendo  $K$  el que presentó el menor valor con  $n = 20$ . Al cotejar los 4 métodos

Escenario Normal Potencia (-1/3)									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.110	0.102	0.052	-0.045	0.167	0.147	0.088	0.075
	<b>TN</b>	0.026	0.011	0.013	0.008	0.097	0.118	0.102	0.071
	<b>EMP</b>	0.096	0.086	0.075	0.056	0.150	0.144	0.130	0.095
	<b>K</b>	0.032	0.017	-0.003	-0.039	0.113	0.113	0.094	0.078
$(m = n = 50)$	<b>N</b>	0.119	0.110	0.045	-0.068	0.149	0.127	0.064	0.084
	<b>TN</b>	0.010	0.007	0.007	0.003	0.066	0.075	0.066	0.047
	<b>EMP</b>	0.058	0.053	0.045	0.035	0.096	0.095	0.083	0.063
	<b>K</b>	0.019	0.011	-0.006	-0.035	0.076	0.076	0.062	0.056
$(m = n = 100)$	<b>N</b>	0.124	0.110	0.040	-0.080	0.140	0.119	0.052	0.093
	<b>TN</b>	0.006	0.007	0.007	0.004	0.049	0.052	0.046	0.033
	<b>EMP</b>	0.041	0.040	0.034	0.025	0.067	0.069	0.061	0.046
	<b>K</b>	0.015	0.011	-0.003	-0.029	0.054	0.053	0.044	0.044

Tabla 5.6: Sesgo y RMSE para Estimadores de  $YI$ :  $Y^{-1/3} \sim N(3,5,0,09)$ ,  $X^{-1/3} \sim N(\mu_D, 0,25)$

se nota que  $K$  tiene sesgos mínimos para todas las situaciones y  $EMP$  presenta los sesgos máximos, repitiéndose la misma situación para los errores. Los procedimientos  $N$  y  $TN$  muestran patrones similares de sesgos y errores.

Para los datos provenientes de la Mixtura 3 los resultados, mostrados en la Tabla 5.9, indican una situación parecida a la anterior:

- El examen de los sesgos de cada método indicó que todos obtienen su correspondiente sesgo mínimo para  $YI = 0,8$ , siendo  $K$  el que mostró el menor valor para  $n = 20$ .
- La comparación de los métodos, para los valores de  $YI$  considerados con  $n$  fijo, muestra que  $K$  tiene sesgos mínimos para todas las situaciones y  $EMP$  presenta los sesgos máximos. Esto se repite para los errores.
- $N$  y  $TN$  continúan mostrando sesgos y errores muy similares.

La Tabla 5.10 contiene los resultados obtenidos para la Mixtura 4. La revisión de los sesgos de cada método mostró que alcanzan su respectivo sesgo mínimo para  $YI=0.8$ , siendo  $K$  el que presenta el menor valor para  $n = 20$ . La comparación de los métodos, realizada como las anteriores, mostró que  $K$  tiene los menores sesgos y  $RMSE$  en todas las combinaciones

Escenario Mixtura 1									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	-0.016	0.237	0.242	0.137	0.100	0.256	0.251	0.142
	<b>TN</b>	-0.013	0.243	0.246	0.139	0.103	0.262	0.254	0.143
	<b>EMP</b>	-0.099	0.308	0.291	0.168	0.130	0.326	0.300	0.173
	<b>K</b>	-0.160	0.206	0.212	0.120	0.171	0.230	0.226	0.130
$(m = n = 50)$	<b>N</b>	-0.038	0.230	0.240	0.138	0.080	0.238	0.243	0.140
	<b>TN</b>	-0.037	0.233	0.242	0.139	0.080	0.242	0.245	0.140
	<b>EMP</b>	-0.147	0.275	0.271	0.158	0.153	0.283	0.275	0.161
	<b>K</b>	-0.179	0.202	0.213	0.122	0.181	0.213	0.219	0.126
$(m = n = 100)$	<b>N</b>	-0.047	0.232	0.242	0.139	0.069	0.236	0.243	0.140
	<b>TN</b>	-0.046	0.234	0.243	0.140	0.069	0.238	0.245	0.141
	<b>EMP</b>	-0.164	0.262	0.264	0.154	0.166	0.267	0.266	0.155
	<b>K</b>	-0.184	0.208	0.220	0.127	0.185	0.213	0.223	0.128

Tabla 5.7: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 1

Escenario Mixtura 2									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.221	0.220	0.180	0.108	0.248	0.242	0.196	0.118
	<b>TN</b>	0.227	0.227	0.185	0.110	0.256	0.249	0.201	0.120
	<b>EMP</b>	0.313	0.293	0.239	0.146	0.335	0.312	0.253	0.154
	<b>K</b>	0.204	0.190	0.147	0.086	0.233	0.216	0.171	0.105
$(m = n = 50)$	<b>N</b>	0.208	0.212	0.176	0.108	0.221	0.221	0.183	0.112
	<b>TN</b>	0.211	0.216	0.179	0.109	0.224	0.225	0.185	0.113
	<b>EMP</b>	0.264	0.257	0.212	0.133	0.276	0.266	0.218	0.137
	<b>K</b>	0.189	0.183	0.147	0.088	0.204	0.195	0.156	0.095
$(m = n = 100)$	<b>N</b>	0.210	0.214	0.178	0.110	0.216	0.219	0.181	0.112
	<b>TN</b>	0.212	0.216	0.180	0.110	0.218	0.221	0.183	0.112
	<b>EMP</b>	0.249	0.246	0.203	0.127	0.255	0.251	0.206	0.129
	<b>K</b>	0.193	0.190	0.154	0.093	0.201	0.197	0.159	0.097

Tabla 5.8: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 2

Escenario Mixtura 3									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.173	0.470	0.364	0.192	0.204	0.474	0.364	0.192
	<b>TN</b>	0.170	0.474	0.365	0.192	0.204	0.478	0.366	0.192
	<b>EMP</b>	0.255	0.515	0.387	0.199	0.281	0.519	0.388	0.199
	<b>K</b>	0.160	0.446	0.354	0.191	0.194	0.451	0.356	0.191
$(m = n = 50)$	<b>N</b>	0.170	0.473	0.366	0.193	0.184	0.474	0.367	0.193
	<b>TN</b>	0.169	0.475	0.367	0.193	0.184	0.476	0.367	0.193
	<b>EMP</b>	0.220	0.500	0.382	0.198	0.233	0.502	0.382	0.198
	<b>K</b>	0.157	0.451	0.358	0.192	0.172	0.453	0.358	0.192
$(m = n = 100)$	<b>N</b>	0.166	0.472	0.367	0.193	0.172	0.473	0.367	0.193
	<b>TN</b>	0.166	0.473	0.367	0.193	0.173	0.474	0.367	0.193
	<b>EMP</b>	0.201	0.491	0.377	0.197	0.208	0.492	0.378	0.197
	<b>K</b>	0.155	0.453	0.359	0.192	0.163	0.454	0.359	0.192

Tabla 5.9: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 3

de  $YI$  y  $n$  y  $EMP$  los mayores.  $N$  y  $TN$  presentan sesgos y  $RMSE$  similares.

Los datos provenientes de las distribuciones de la Mixtura 5 se exponen en la Tabla 5.11. Los 4 procedimientos presentan sus respectivos sesgos mínimos con  $YI = 0,2$  y  $n = 100$ , siendo el de  $K$  el menor de todos. Para el caso de los errores, todos los métodos mostraron su respectivo mínimo para  $YI = 0,2$ . En la comparación de los métodos, para el rango de valores de  $YI$  seleccionados con  $n$  fijo, se observó que:

- El procedimiento  $K$  presentó los sesgos mínimos para todos los tamaños muestrales.
- $TN$ , independientemente del tamaño muestral, mostró los sesgos máximos para  $YI = 0,2$ .
- $EMP$  presentó los sesgos máximos para  $YI \geq 0,4$ .
- $K$  tiene los errores mínimos para  $YI \geq 0,4$  con todos los tamaños muestrales.

En la Tabla 5.12 se presentan los resultados para la Mixtura 6. El examen de los sesgos dentro de cada método indica que todos los procedimientos obtuvieron su respectivo mínimo con  $YI = 0,2$  y  $n = 100$ , siendo  $K$  el que tiene el menor valor. Todos los procedimientos

Escenario Mixtura 4									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.233	0.431	0.355	0.191	0.258	0.436	0.356	0.191
	<b>TN</b>	0.233	0.436	0.357	0.191	0.260	0.441	0.358	0.191
	<b>EMP</b>	0.314	0.483	0.381	0.199	0.335	0.488	0.382	0.199
	<b>K</b>	0.215	0.403	0.343	0.190	0.242	0.411	0.346	0.191
$(m = n = 50)$	<b>N</b>	0.231	0.433	0.358	0.192	0.241	0.436	0.358	0.192
	<b>TN</b>	0.231	0.436	0.359	0.192	0.242	0.438	0.359	0.192
	<b>EMP</b>	0.280	0.465	0.376	0.198	0.290	0.467	0.376	0.198
	<b>K</b>	0.212	0.408	0.347	0.191	0.224	0.411	0.348	0.191
$(m = n = 100)$	<b>N</b>	0.226	0.432	0.358	0.193	0.231	0.433	0.358	0.193
	<b>TN</b>	0.227	0.434	0.359	0.193	0.233	0.435	0.359	0.193
	<b>EMP</b>	0.261	0.453	0.370	0.197	0.267	0.455	0.371	0.197
	<b>K</b>	0.211	0.410	0.349	0.191	0.217	0.412	0.349	0.191

Tabla 5.10: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 4

Escenario Mixtura 5									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	0.043	0.310	0.322	0.185	0.117	0.324	0.325	0.185
	<b>TN</b>	0.048	0.315	0.324	0.185	0.122	0.329	0.327	0.185
	<b>EMP</b>	-0.126	0.376	0.357	0.197	0.144	0.388	0.360	0.197
	<b>K</b>	-0.176	0.279	0.303	0.182	0.183	0.296	0.309	0.183
$(m = n = 50)$	<b>N</b>	0.036	0.310	0.325	0.187	0.082	0.315	0.326	0.187
	<b>TN</b>	0.038	0.313	0.327	0.187	0.085	0.318	0.328	0.187
	<b>EMP</b>	-0.161	0.349	0.349	0.195	0.165	0.355	0.350	0.196
	<b>K</b>	-0.185	0.281	0.309	0.183	0.187	0.288	0.310	0.184
$(m = n = 100)$	<b>N</b>	0.030	0.306	0.325	0.187	0.060	0.309	0.325	0.187
	<b>TN</b>	0.031	0.308	0.325	0.187	0.062	0.311	0.326	0.187
	<b>EMP</b>	-0.177	0.333	0.341	0.194	0.179	0.336	0.342	0.194
	<b>K</b>	-0.191	0.282	0.310	0.183	0.191	0.285	0.311	0.184

Tabla 5.11: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 5

Escenario Mixtura 6									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$(m = n = 20)$	<b>N</b>	-0.052	0.241	0.307	0.184	0.100	0.262	0.310	0.184
	<b>TN</b>	-0.050	0.247	0.309	0.184	0.101	0.267	0.313	0.184
	<b>EMP</b>	-0.066	0.313	0.346	0.196	0.114	0.330	0.350	0.197
	<b>K</b>	-0.139	0.211	0.285	0.180	0.156	0.236	0.292	0.182
$(m = n = 50)$	<b>N</b>	-0.085	0.239	0.310	0.186	0.103	0.248	0.311	0.186
	<b>TN</b>	-0.084	0.242	0.311	0.186	0.103	0.250	0.312	0.186
	<b>EMP</b>	-0.121	0.284	0.335	0.195	0.133	0.292	0.337	0.195
	<b>K</b>	-0.163	0.211	0.291	0.182	0.168	0.221	0.293	0.183
$(m = n = 100)$	<b>N</b>	-0.102	0.236	0.309	0.186	0.112	0.240	0.309	0.186
	<b>TN</b>	-0.101	0.237	0.309	0.186	0.111	0.241	0.310	0.186
	<b>EMP</b>	-0.149	0.265	0.326	0.193	0.153	0.270	0.327	0.193
	<b>K</b>	-0.176	0.211	0.292	0.182	0.178	0.216	0.293	0.183

Tabla 5.12: Sesgo y RMSE para Estimadores de  $YI$  - Mixtura 6

logran su correspondiente error mínimo con  $YI = 0,2$  y  $n = 20$ . Al comparar los métodos para los valores de  $YI$  con  $n$  fijo, se nota que  $K$  tiene los sesgos mínimos en todas las situaciones y es el procedimiento que obtiene los  $RMSE$  menores para  $YI \geq 0,4$ . El procedimiento  $EMP$  muestra los sesgos máximos para  $YI \geq 0,4$

### 5.1.2. Puntos de Corte

Aunque  $YI$  y  $c^*$  están fuertemente relacionados, un buen método para estimar uno de los dos no es necesariamente bueno para estimar el otro (la mayoría de las propiedades de los estimadores, como el sesgo, no se conserva bajo transformaciones no lineales). En esta sección se presentan los resultados obtenidos en la estimación del punto de corte  $c^*$  por los cuatro métodos. Las tablas siguientes tienen una nueva fila, cuando se comparan con las anteriores, en la que se muestran los valores de los puntos de corte calculados, usando los verdaderos valores de los parámetros, correspondientes a los valores de  $YI$ .

Escenario Normal con Varianzas y Tamaños Muestrales Iguales-Puntos de Corte									
	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		6.627	6.762	6.921	7.141	6.627	6.762	6.921	7.141
$(m = n = 20)$	<b>N</b>	0.006	-0.003	-0.002	-0.001	0.210	0.118	0.086	0.087
	<b>TN</b>	-0.005	-0.013	-0.011	-0.006	0.212	0.126	0.097	0.092
	<b>EMP</b>	-0.012	-0.015	-0.028	-0.049	0.280	0.207	0.174	0.168
	<b>K</b>	0.003	0.008	0.002	-0.002	0.333	0.186	0.137	0.121
$(m = n = 50)$	<b>N</b>	0.005	0.002	0.000	-0.001	0.140	0.068	0.051	0.054
	<b>TN</b>	-0.001	-0.005	-0.005	-0.004	0.145	0.080	0.064	0.059
	<b>EMP</b>	-0.013	-0.010	-0.013	-0.013	0.226	0.161	0.140	0.121
	<b>K</b>	-0.004	0.002	0.001	0.001	0.231	0.125	0.092	0.079
$(m = n = 100)$	<b>N</b>	0.001	0.001	0.000	-0.001	0.099	0.049	0.037	0.039
	<b>TN</b>	-0.001	0.000	-0.002	-0.003	0.104	0.058	0.046	0.043
	<b>EMP</b>	-0.007	0.000	-0.005	-0.009	0.201	0.133	0.113	0.094
	<b>K</b>	-0.004	0.003	0.002	0.000	0.181	0.096	0.068	0.058

Tabla 5.13: Sesgo y RMSE para Estimadores de  $c^*$ :  $Y \sim N(6,5, 0,25)$ ,  $X \sim N(\mu_D, 0,25)$

### Simulaciones con Distribuciones Normales - Varianzas y Tamaños Muestrales Iguales

En la Tabla 5.13 se presentan los resultados obtenidos para los puntos de corte óptimos en el escenario normal con varianzas iguales. Los métodos  $N$  y  $TN$  logran sus correspondientes sesgos mínimos con  $YI = 0,4$  y  $n = 20$ . En cuanto a los errores se observa que todos los procedimientos logran sus respectivos mínimos con  $n = 100$ . Al comparar los métodos, para los valores de  $YI$  con  $n$  fijo, se nota que para  $n \leq 50$  el procedimiento  $EMP$  logra los sesgos mínimos y el método  $N$  es el que presenta los errores mínimos para todas las situaciones.

### Simulaciones con Distribuciones Normales - Varianzas y Tamaños Muestrales Diferentes

Los resultados para el escenario normal con tamaños muestrales diferentes son mostrados en la Tabla 5.14. Se observa que los métodos  $N$  y  $TN$  obtienen sus correspondientes sesgos mínimos en  $YI = 0,2$  con  $n = 20$ . Al cotejar los métodos, para los diferentes valores de  $YI$

Escenario Normal con Varianzas Iguales y $(m, n)$ Diferentes-Puntos de Corte									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
	$c^*$	6.627	6.762	6.921	7.141	6.627	6.762	6.921	7.141
$(m = 20, n = 60)$	<b>N</b>	-0.003	0.003	0.004	0.005	0.186	0.091	0.066	0.069
	<b>TN</b>	-0.014	-0.007	-0.005	0.000	0.189	0.101	0.077	0.073
	<b>EMP</b>	0.011	0.009	0.004	0.003	0.272	0.195	0.158	0.137
	<b>K</b>	0.022	0.014	0.007	0.001	0.300	0.161	0.110	0.094
$(m = 50, n = 150)$	<b>N</b>	0.001	0.003	0.003	0.003	0.114	0.056	0.041	0.042
	<b>TN</b>	-0.001	0.001	0.002	0.003	0.116	0.064	0.049	0.044
	<b>EMP</b>	0.007	0.017	0.018	0.006	0.212	0.152	0.121	0.105
	<b>K</b>	0.029	0.016	0.011	0.002	0.201	0.105	0.073	0.063
$(m = 100, n = 300)$	<b>N</b>	0.001	0.001	0.001	0.001	0.080	0.040	0.029	0.031
	<b>TN</b>	0.001	0.001	0.001	0.001	0.084	0.048	0.037	0.034
	<b>EMP</b>	0.007	0.002	0.010	0.006	0.176	0.121	0.099	0.092
	<b>K</b>	0.019	0.007	0.004	0.000	0.153	0.079	0.055	0.049

Tabla 5.14: Sesgo y RMSE para Estimadores de  $c^*$ :  $Y \sim N(6,5, 0,25)$ ,  $X \sim N(\mu_D, 0,25)$ 

con  $n$  fijo, se nota que:

- El procedimiento  $TN$ , para el primer conjunto de tamaños muestrales, muestra los sesgos mínimos para todos los valores de  $YI$ .
- Para  $(m = 50, n = 150)$  y  $YI \leq 0,6$  es también  $TN$  el que logra los sesgos mínimos.
- Los métodos  $N$  y  $TN$  para  $YI \leq 0,6$  presentan los valores menores para el sesgo.
- Estos dos procedimientos tiene patrones similares de errores, presentando  $TN$  valores ligeramente inferiores a  $N$ .

### Simulaciones con Distribuciones Normales - Varianzas Diferentes y Tamaños Muestrales Iguales

En la Tabla 5.15 se presentan los resultados para el escenario normal con varianzas diferentes. Al examinar los sesgos de los métodos se nota que, con excepción del procedimiento  $K$ , todos logran su sesgo mínimo para  $YI = 0,2$  con  $n = 20$ . La comparación de los métodos, realizada como ha sido descrita anteriormente, muestra que el procedimiento  $EMP$  tiene

Escenario Normal con Varianzas Diferentes y $(m = n)$ -Puntos de Corte									
	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
	$c^*$	6.828	6.806	6.851	6.951	6.828	6.806	6.851	6.951
$(m = n = 20)$	<b>N</b>	-0.142	-0.011	-0.005	-0.002	0.277	0.082	0.065	0.067
	<b>TN</b>	-0.150	-0.022	-0.015	-0.007	0.281	0.092	0.072	0.069
	<b>EMP</b>	-0.078	-0.036	-0.039	-0.050	0.180	0.144	0.131	0.128
	<b>K</b>	0.011	0.013	0.013	0.011	0.183	0.127	0.100	0.092
$(m = n = 50)$	<b>N</b>	-0.064	-0.002	-0.001	-0.001	0.196	0.048	0.039	0.041
	<b>TN</b>	-0.073	-0.008	-0.007	-0.004	0.205	0.053	0.046	0.044
	<b>EMP</b>	-0.035	-0.024	-0.012	-0.017	0.140	0.114	0.098	0.088
	<b>K</b>	0.022	0.018	0.017	0.012	0.121	0.084	0.066	0.061
$(m = n = 100)$	<b>N</b>	-0.018	-0.002	-0.002	-0.002	0.105	0.034	0.029	0.030
	<b>TN</b>	-0.019	-0.005	-0.004	-0.003	0.107	0.037	0.033	0.032
	<b>EMP</b>	-0.021	-0.016	-0.012	-0.012	0.113	0.091	0.079	0.073
	<b>K</b>	0.019	0.013	0.010	0.007	0.087	0.063	0.049	0.044

Tabla 5.15: Sesgo y RMSE para Estimadores de  $c^*$ :  $Y \sim N(6,5, 0,09)$ ,  $X \sim N(\mu_D, 0,25)$

los sesgos mínimos en todas las situaciones, con dos excepciones:  $YI = 0,2$  con  $n = 20$  y  $n = 50$  y  $N$  logra los errores mínimos para  $YI \geq 0,4$ , independientemente del tamaño muestral.

### Simulaciones con Distribuciones Asimétricas

Los sesgos y  $RMSE$  obtenidos para los puntos de corte en este escenario, comentado en la sub-sección 5.1.1 (página 32), se muestran en la Tabla 5.16. Se observa que todos los procedimientos logran sus sesgos mínimos para  $YI = 0,2$ . Al comparar los métodos se nota que  $TN$  presenta el sesgo mínimo para  $YI = 0,2$ , con todos los tamaños muestrales utilizados, en cuanto que  $EMP$  alcanza el mínimo para  $YI \geq 0,4$ . Para los errores se observa que  $N$  muestra  $RMSE$  mínimo para  $YI = 0,2$  y  $TN$  lo obtiene para  $YI \geq 0,4$ , independientemente del tamaño muestral.

Escenario Normal Potencia (-1/3)-Puntos de Corte									
	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		3.135	3.069	3.201	3.526	3.135	3.069	3.201	3.526
$(m = n = 20)$	<b>N</b>	0.024	0.362	0.442	0.392	0.458	0.467	0.551	0.552
	<b>TN</b>	-0.285	-0.014	-0.001	0.020	0.594	0.243	0.224	0.252
	<b>EMP</b>	-0.131	-0.089	-0.081	-0.118	0.505	0.423	0.372	0.409
	<b>K</b>	0.195	0.144	0.193	0.241	0.694	0.391	0.377	0.451
$(m = n = 50)$	<b>N</b>	0.155	0.404	0.478	0.421	0.300	0.450	0.527	0.496
	<b>TN</b>	-0.098	0.000	0.004	0.009	0.377	0.154	0.146	0.153
	<b>EMP</b>	-0.044	-0.030	-0.034	-0.043	0.385	0.321	0.290	0.306
	<b>K</b>	0.129	0.111	0.163	0.188	0.413	0.260	0.269	0.317
$(m = n = 100)$	<b>N</b>	0.189	0.422	0.496	0.440	0.233	0.447	0.522	0.481
	<b>TN</b>	-0.038	-0.001	0.004	0.006	0.239	0.108	0.102	0.108
	<b>EMP</b>	-0.026	-0.019	-0.005	-0.025	0.335	0.256	0.241	0.251
	<b>K</b>	0.056	0.075	0.128	0.164	0.295	0.189	0.202	0.260

Tabla 5.16: Sesgo y RMSE para Estimadores de  $c^*$ :  $Y^{-1/3} \sim N(3,5,0,09)$ ,  $X^{-1/3} \sim N(\mu_D, 0,25)$

### Simulaciones con Mezclas de Distribuciones Normales

La Tabla 5.17 muestra los sesgos y *RMSE* obtenidos para el escenario denominado Mezcla 1. El examen de los sesgos de cada método señala que *N*, *EMP* y *K* consiguen sus mínimos pertinentes para  $YI = 0,2$  y se observa que todos los procedimientos muestran sesgos negativos para  $YI \leq 0,4$ , independientemente de cuál sea el tamaño muestral. En términos de *RMSE*, el comportamiento de los procedimientos es muy variado. Al confrontar los métodos se percibe que, para  $YI = 0,2$ , el procedimiento *N* es el que tiene los sesgos mínimos, pero cuando  $YI \geq 0,4$  es *EMP* el que consigue esos valores mínimos, ambos casos independientemente de  $n$ .

Los resultados obtenidos para la Mezcla 2 se presentan en la Tabla 5.18. El examen de los sesgos de los 4 procedimientos muestra que todos logran sus mínimos para  $YI = 0,2$ , siendo la totalidad de los sesgos para este valor de  $YI$  negativos. En la comparación de los métodos se observa que el procedimiento *EMP* tiene los sesgos menores para  $n \leq 50$  en todo el rango de valores de  $YI$  usados. En cuanto a los errores, el comportamiento es muy variable, tal y como se observó en el escenario de la Mezcla 1.

## Escenario Mixtura 1-Puntos de Corte

	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		11.730	11.660	11.100	11.370	11.730	11.660	11.100	11.370
$(m = n = 20)$	<b>N</b>	-2.122	-0.651	0.381	0.561	2.202	0.677	0.430	0.612
	<b>TN</b>	-2.067	-0.664	0.374	0.569	2.173	0.699	0.427	0.624
	<b>EMP</b>	-1.107	-0.722	0.260	0.282	1.555	0.810	0.433	0.481
	<b>K</b>	-1.214	-0.617	0.408	0.579	2.245	0.681	0.487	0.653
$(m = n = 50)$	<b>N</b>	-2.301	-0.648	0.384	0.565	2.339	0.658	0.403	0.584
	<b>TN</b>	-2.211	-0.649	0.380	0.566	2.298	0.669	0.401	0.586
	<b>EMP</b>	-0.627	-0.672	0.334	0.455	1.254	0.733	0.427	0.538
	<b>K</b>	-0.665	-0.626	0.397	0.574	2.030	0.656	0.432	0.606
$(m = n = 100)$	<b>N</b>	-2.383	-0.648	0.382	0.561	2.400	0.653	0.392	0.571
	<b>TN</b>	-2.318	-0.647	0.381	0.559	2.373	0.656	0.392	0.569
	<b>EMP</b>	-0.310	-0.657	0.355	0.511	1.030	0.698	0.413	0.557
	<b>K</b>	-0.173	-0.633	0.392	0.564	1.665	0.648	0.414	0.582

Tabla 5.17: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 1

## Escenario Mixtura 2-Puntos de Corte

	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		10.910	10.660	10.900	11.310	10.910	10.660	10.900	11.310
$(m = n = 20)$	<b>N</b>	-0.444	0.159	0.285	0.358	0.498	0.228	0.329	0.410
	<b>TN</b>	-0.453	0.147	0.277	0.357	0.514	0.235	0.328	0.412
	<b>EMP</b>	-0.474	0.124	0.239	0.194	0.621	0.356	0.394	0.395
	<b>K</b>	-0.449	0.166	0.295	0.359	0.579	0.301	0.373	0.437
$(m = n = 50)$	<b>N</b>	-0.458	0.153	0.283	0.359	0.477	0.182	0.300	0.378
	<b>TN</b>	-0.457	0.149	0.280	0.358	0.486	0.194	0.301	0.377
	<b>EMP</b>	-0.473	0.147	0.252	0.305	0.570	0.295	0.338	0.389
	<b>K</b>	-0.465	0.148	0.277	0.355	0.524	0.228	0.315	0.390
$(m = n = 100)$	<b>N</b>	-0.452	0.155	0.283	0.357	0.462	0.171	0.292	0.368
	<b>TN</b>	-0.448	0.155	0.283	0.356	0.464	0.180	0.294	0.367
	<b>EMP</b>	-0.451	0.157	0.272	0.325	0.519	0.263	0.335	0.379
	<b>K</b>	-0.453	0.154	0.282	0.354	0.490	0.203	0.305	0.373

Tabla 5.18: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 2

Escenario Mixtura 3-Puntos de Corte									
	YI	Sesgo				RMSE			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
	$c^*$	14.270	11.410	13.180	13.840	14.270	11.410	13.180	13.840
$(m = n = 20)$	<b>N</b>	-2.062	1.307	-0.068	-0.324	2.075	1.318	0.238	0.433
	<b>TN</b>	-2.074	1.291	-0.039	-0.251	2.101	1.311	0.250	0.438
	<b>EMP</b>	-2.161	1.164	-0.417	-1.032	2.189	1.195	0.537	1.096
	<b>K</b>	-1.994	1.337	-0.057	-0.346	2.019	1.356	0.277	0.495
$(m = n = 50)$	<b>N</b>	-2.035	1.299	-0.084	-0.347	2.038	1.303	0.162	0.388
	<b>TN</b>	-2.035	1.277	-0.070	-0.310	2.043	1.298	0.167	0.389
	<b>EMP</b>	-2.086	1.241	-0.255	-0.786	2.103	1.260	0.350	0.845
	<b>K</b>	-1.988	1.323	-0.074	-0.344	1.998	1.331	0.190	0.422
$(m = n = 100)$	<b>N</b>	-2.033	1.310	-0.066	-0.324	2.035	1.313	0.118	0.347
	<b>TN</b>	-2.033	1.298	-0.062	-0.311	2.036	1.310	0.122	0.352
	<b>EMP</b>	-2.060	1.280	-0.153	-0.610	2.071	1.291	0.248	0.664
	<b>K</b>	-1.990	1.327	-0.057	-0.320	1.995	1.332	0.142	0.373

Tabla 5.19: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 3

Los datos obtenidos de la Mixtura 3 produjeron los resultados que se exponen en la Tabla 5.19. La revisión de los sesgos en cada uno de los métodos indica que todos alcanzan sus pertinentes mínimos para  $YI = 0,2$  con  $n = 20$ , observándose que sólo para  $YI = 0,4$  los sesgos son positivos. Al cotejar los procedimientos, para el rango de valores de  $YI$  seleccionado en cada uno de los tamaños muestrales utilizados, se nota que el método *EMP* logra los sesgos mínimos en todas las combinaciones. En términos de los errores la pauta de los procedimientos es muy variable.

Los sesgos y *RMSE* obtenidos para los datos que provienen de la Mixtura 4 son presentados en la Tabla 5.20. El examen de los sesgos en cada método mostró la misma pauta que se observó en el escenario denominado Mixtura 3: todos los procedimientos lograron sus respectivos mínimos para  $YI = 0,2$  con  $n = 20$ , notándose que para  $YI = 0,2$  e  $YI = 0,8$  los sesgos son negativos. Al comparar los métodos se observa que *EMP* exhibe los sesgos mínimos para todos el rango de valores de  $YI$  y tamaños muestrales utilizados. El comportamiento, en términos de errores, es muy variable.

En la Tabla 5.21 se encuentran los resultados conseguidos con los datos de la Mixtura 5. El

Escenario Mixtura 4-Puntos de Corte									
	<b>YI</b>	<b>Sesgo</b>				<b>RMSE</b>			
		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		14.350	11.370	12.870	13.890	14.350	11.370	12.870	13.890
$(m = n = 20)$	<b>N</b>	-2.105	1.354	0.337	-0.182	2.115	1.365	0.411	0.355
	<b>TN</b>	-2.134	1.333	0.359	-0.112	2.147	1.352	0.437	0.393
	<b>EMP</b>	-2.226	1.234	0.011	-0.925	2.256	1.270	0.360	1.013
	<b>K</b>	-2.074	1.383	0.348	-0.198	2.099	1.403	0.446	0.425
$(m = n = 50)$	<b>N</b>	-2.092	1.348	0.322	-0.205	2.096	1.352	0.352	0.276
	<b>TN</b>	-2.104	1.332	0.332	-0.172	2.111	1.345	0.365	0.298
	<b>EMP</b>	-2.143	1.300	0.171	-0.650	2.161	1.319	0.304	0.736
	<b>K</b>	-2.059	1.374	0.332	-0.197	2.070	1.382	0.377	0.322
$(m = n = 100)$	<b>N</b>	-2.092	1.358	0.339	-0.181	2.094	1.360	0.354	0.224
	<b>TN</b>	-2.082	1.345	0.341	-0.172	2.089	1.356	0.357	0.243
	<b>EMP</b>	-2.114	1.329	0.270	-0.466	2.127	1.342	0.342	0.546
	<b>K</b>	-2.052	1.375	0.348	-0.177	2.058	1.380	0.373	0.265

Tabla 5.20: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 4

análisis de los sesgos de cada método indica que todos los procedimientos obtuvieron sus correspondientes mínimos para  $YI = 0,2$  y  $n \leq 50$ . Al confrontar los métodos, se observa que con  $n \leq 50$  los sesgos mínimos para  $YI = 0,2$  fueron obtenidos por el procedimiento  $N$  y para  $YI \geq 0,4$  por el procedimiento  $EMP$ . Con  $n = 100$  la diferencia que se revela es que el procedimiento  $TN$  es quien produce el sesgo mínimo para  $YI = 0,4$ , en vez de  $EMP$ . En términos de errores se mantiene la pauta observada en los datos provenientes de las mixturas anteriores.

Los resultados de los datos de la Mixtura 6 se exponen en la Tabla 5.22. La revisión de los sesgos apunta que todos los procedimientos logran sus mínimos para  $YI = 0,2$ . La comparación de los métodos muestra que para  $n \geq 50$ , los procedimientos  $N$  y  $TN$  obtuvieron los sesgos mínimos para  $YI = 0,2$  e  $YI = 0,4$  respectivamente, mientras que para  $YI \geq 0,6$  es el método  $EMP$  el que logra los valores menores. Para  $n = 20$ ,  $N$  muestra el menor valor para  $YI = 0,2$  y con  $YI \geq 0,4$  los sesgos menores son alcanzados por el procedimiento  $EMP$ . En términos de  $RMSE$  el patrón de comportamiento es muy variable.

## Escenario Mixtura 5-Puntos de Corte

	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		14.870	11.100	12.530	13.590	14.870	11.100	12.530	13.590
$(m = n = 20)$	<b>N</b>	-3.801	1.234	0.286	-0.282	3.814	1.241	0.336	0.369
	<b>TN</b>	-3.709	1.215	0.291	-0.243	3.777	1.223	0.342	0.363
	<b>EMP</b>	-2.727	1.174	0.110	-0.796	2.840	1.203	0.293	0.868
	<b>K</b>	-2.948	1.249	0.296	-0.285	3.317	1.249	0.365	0.403
$(m = n = 50)$	<b>N</b>	-3.873	1.230	0.274	-0.301	3.877	1.233	0.294	0.333
	<b>TN</b>	-3.764	1.204	0.276	-0.283	3.835	1.223	0.296	0.331
	<b>EMP</b>	-2.307	1.204	0.199	-0.575	2.381	1.220	0.279	0.637
	<b>K</b>	-2.459	1.249	0.280	-0.298	2.856	1.256	0.311	0.351
$(m = n = 100)$	<b>N</b>	-3.869	1.235	0.286	-0.283	3.871	1.236	0.296	0.302
	<b>TN</b>	-3.735	1.214	0.286	-0.275	3.813	1.229	0.296	0.302
	<b>EMP</b>	-2.061	1.220	0.253	-0.439	2.120	1.230	0.302	0.489
	<b>K</b>	-2.229	1.242	0.293	-0.278	2.662	1.246	0.311	0.313

Tabla 5.21: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 5

## Escenario Mixtura 6-Puntos de Corte

	Sesgo					RMSE			
	YI	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
$c^*$		15.000	11.050	12.360	13.630	15.000	11.050	12.360	13.630
$(m = n = 20)$	<b>N</b>	-3.659	1.232	0.515	-0.148	3.693	1.240	0.545	0.291
	<b>TN</b>	-3.601	1.209	0.515	-0.113	3.678	1.220	0.546	0.301
	<b>EMP</b>	-3.208	1.183	0.357	-0.685	3.310	1.218	0.460	0.785
	<b>K</b>	-3.194	1.239	0.522	-0.149	3.450	1.259	0.568	0.334
$(m = n = 50)$	<b>N</b>	-3.760	1.230	0.503	-0.167	3.791	1.233	0.515	0.223
	<b>TN</b>	-3.654	1.207	0.503	-0.153	3.745	1.223	0.514	0.234
	<b>EMP</b>	-2.928	1.211	0.443	-0.442	3.038	1.231	0.489	0.531
	<b>K</b>	-2.738	1.244	0.508	-0.163	3.014	1.252	0.526	0.253
$(m = n = 100)$	<b>N</b>	-3.875	1.232	0.514	-0.149	3.895	1.234	0.520	0.186
	<b>TN</b>	-3.720	1.213	0.513	-0.143	3.830	1.228	0.519	0.194
	<b>EMP</b>	-2.732	1.215	0.489	-0.304	2.852	1.228	0.520	0.380
	<b>K</b>	-2.585	1.235	0.520	-0.145	2.905	1.240	0.531	0.209

Tabla 5.22: Sesgo y RMSE para Estimadores de  $c^*$  - Mixtura 6



## Capítulo 6

# Aplicación a Datos Reales

En el año 2004 fue llevado a cabo un estudio en Galicia para evaluar la presencia de los principales factores de riesgo de las enfermedades cardiovasculares, entre los que se encuentra el aumento de la tensión arterial ([Perez-Fernandez y otros, 2007](#)). Los individuos del estudio fueron seleccionados a partir de la base de datos del Servicio Gallego de Salud (SERGAS), que cubre más del 95 % de la población. Fueron tomadas medidas antropométricas directas incluyendo el peso corporal (en kilogramos), altura (en metros), circunferencia de la cintura (en centímetros) y circunferencia de la cadera (en centímetros). Fueron calculados el Índice de Masa Corporal (el peso en kilogramos dividido por el cuadrado de la altura en metros ( $kg/m^2$ ) y el Índice Cintura/Cadera (perímetro de la cintura dividido por el perímetro máximo de la cadera).

En la [Tabla 6.1](#) se muestran los límites del IMC, iguales en hombres y en mujeres, para clasificar la obesidad y el sobrepeso, según la Conferencia de Consenso 2000 de la Sociedad Española para el Estudio de la Obesidad ([SEEDO, 2000](#)):

	<b>Valores Límites del IMC (<math>kg/m^2</math>)</b>
Peso Insuficiente	$< 18,5$
Normopeso	$18,5 - 24,9$
Sobrepeso	$25 - 29,9$
Obesidad Grado I	$30 - 34,9$
Obesidad Grado II	$35 - 39,9$
Obesidad Grado III	$\geq 40$

Tabla 6.1: Clasificación del Sobrepeso y la Obesidad según el IMC

De acuerdo con SEEDO (2000) el Índice Cintura/Cadera ( $C/C$ ) es aceptado como un buen indicador de la obesidad central. En este documento también se explica que algunos autores sugieren que la circunferencia de la cintura tiene una buena correlación con la acumulación de grasa perivisceral (envuelve y protege órganos vitales como el corazón y riñones). Aunque no están claramente definidos los valores a partir de los cuales se observa un aumento del riesgo cardiovascular, estudios epidemiológicos transversales de diferentes comunidades autónomas españolas, como el realizado por Botana y otros (2007), sitúan los valores de riesgo para la circunferencia de la cintura de la población española en  $\geq 102$  cm en los hombres y  $\geq 88$  cm en las mujeres. En este trabajo se utilizará el criterio de la Federación Internacional de Diabetes para la Circunferencia de la Cintura y el de la Organización Mundial de la Salud para el Índice Cintura/Cadera (Ascaso y otros, 2006):

Criterio	Valores Límites	
	Varones	Mujeres
Índice Cintura/Cadera	$> 0,9$	$> 0,85$
Circunferencia de la Cintura	$\geq 94$	$\geq 80$

Tabla 6.2: Valores de Riesgo según la Distribución de la Grasa Corporal

La presión arterial o tensión arterial es la presión que ejerce la sangre contra la pared de las arterias y la **Hipertensión Arterial** (HTA) es el término médico empleado para la presión arterial alta. Las lecturas de la presión arterial se miden en milímetros de mercurio (mmHg) y generalmente se dan como dos números. Por ejemplo, 120 sobre 80 (escrito como 120/80):

- El número superior corresponde a la presión sistólica, que es la presión creada cuando el corazón se contrae. De acuerdo con el criterio formulado por Human Health Services~Department (2003) se considera alta si constantemente es o está por encima de 140.
- El número inferior corresponde a la presión diastólica que es la presión entre latidos cardíacos. Se considera alta si es o está por encima de 90 (Human Health Services~Department, 2003).

Para este trabajo se cuenta con una muestra de 2860 individuos con edades de 18 a 104 años, de los cuales el 53.7% son mujeres (1218 (79.3%) no hipertensas y 318 (20.7%) hipertensas) y el 46.3% son hombres (913 (79.96%) no hipertensos y 411 (31.04%) hipertensos). Perez-Fernandez y otros (2007) especifican que el conocimiento de la HTA de los individuos

clasificados como hipertensos en la muestra es el resultado de un diagnóstico médico previo.

El diseño de estudios de evaluación de las pruebas o marcadores diagnósticos puede ser complejo puesto que existen covariables que pueden influir en la *accuracy* de la prueba diagnóstica. Dichas covariables pueden referirse a características propias del paciente, como la edad y el sexo del mismo, o a propiedades más específicas de la enfermedad en estudio. Por eso es necesario evaluar o controlar los posibles efectos de las covariables en la eficacia, en términos de discriminación, de un marcador, es decir, cuantificar dichos efectos en alguna medida resumen de la curva *ROC*, como el *AUC* o el *YI* (Thompson y Zucchini, 1989; Obuchowski, 1995). En este caso dada la influencia de la edad y el sexo en los valores de las medidas antropométricas, parece adecuado realizar el estudio separadamente por grupos de edad: <30, 30-39, 40-49, 50-59, 60-69,  $\geq 70$  años y sexo.

La prevalencia de la HTA en la muestra estudiada se presenta en la Tabla 6.3. Se observa que la misma se incrementa con la edad en ambos sexos. La HTA fue más prevalente en hombres, excepto en el grupo de edad  $\geq 70$  años donde las mujeres tuvieron una prevalencia de la enfermedad más elevada que los hombres.

Grupos de Edad	Hombres		Mujeres	
	n	%	n	%
< 30 años	47	13.66 %	16	4.00 %
30-39 años	70	20.71 %	23	5.67 %
40-49 años	96	32.32 %	51	16.24 %
50-59 años	85	50.00 %	72	39.34 %
60-69 años	65	59.09 %	78	56.52 %
$\geq 70$ años	48	73.85 %	78	82.11 %
<b>Total</b>	<b>411</b>	<b>31.04 %</b>	<b>318</b>	<b>20.70 %</b>

Tabla 6.3: Prevalencia de HTA por Grupos de Edad y Sexo

## 6.1. Resultados

En esta sección se presentan los resultados obtenidos para cada una de las medidas antropométricas que se estudiaron en este trabajo: el Índice de Masa Corporal (IMC), la Circunferencia de la Cintura y el Índice Cintura/Cadera (*C/C*), como predictores de la presencia de HTA. El análisis efectuado, para cada uno de los marcadores, fue organizado de la siguiente manera:

- Fue realizado un histograma atendiendo a la clasificación por Sexo y Grupo de Edad.
- Para evaluar la capacidad de discriminación de las medidas antropométricas se trazaron las curvas *ROC* para hombres y para mujeres por grupos de edad.
- Se compararon las curvas *ROC* a través del contraste de sus correspondientes *AUCs* con la prueba de homogeneidad de áreas mediante el estadístico Ji-Cuadrado (DeLong y otros, 1988). El contraste fue realizado por grupos de edad para hombres y para mujeres.
- Se calcularon las *AUCs*, sus varianzas e Intervalos de Confianza (IC) al 95 % para cada grupo de edad para hombres y para mujeres.
- Se trazaron las curvas *ROC* en cada uno de los grupos de edad en hombres y mujeres y se compararon sus correspondientes *AUCs*.
- Se representaron las *AUCs* junto con sus correspondientes IC al 95 % y el *YI* (obtenido por los métodos de estimación examinados en este trabajo) por grupos de edad en hombres y en mujeres.
- Se determinaron los puntos de corte óptimos mediante la maximización del *YI* utilizando los métodos de estimación mencionados y se representaron sobre las curvas *ROC* por grupo de edad y sexo.

### 6.1.1. Estudio del Índice de Masa Corporal como Marcador Diagnóstico de Hipertensión Arterial

Los valores medios del IMC se presentan en la Figura 6.1. En los tres primeros grupos de edad, es decir, hasta los 49 años los hombres tuvieron un IMC más elevado. Sin embargo, a partir de los 50 años el IMC es superior en las mujeres, aunque en los dos últimos grupos de edad ambos valores son similares.

Esta medida antropométrica se incrementó con la edad en ambos sexos, excepto en el último grupo, donde en la muestra estudiada hubo una disminución en el IMC con respecto al grupo de edad anterior, siendo algo más marcada esta diferencia en mujeres.

Las curvas *ROC* calculadas, en hombres y en mujeres, para los diferentes grupos de edad considerados están representadas en la Figura 6.2). Se puede decir que este marcador tiene mejor capacidad de discriminación en las mujeres, excepto en la población de  $\geq 70$  años, donde la discriminación disminuye de manera importante. En esta población el marcador

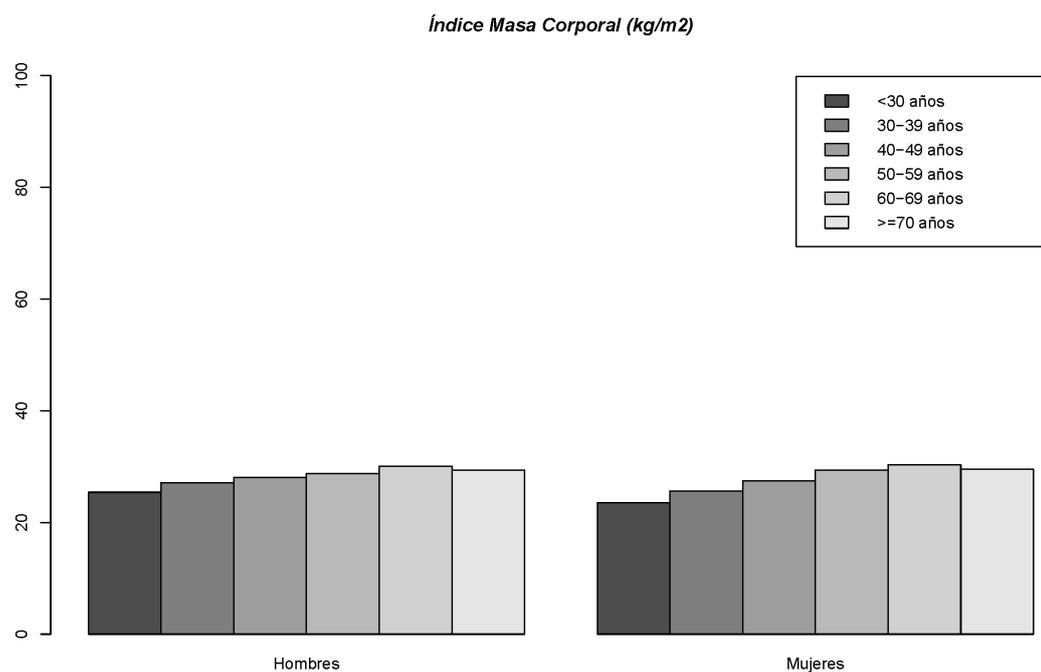


Figura 6.1: Valor Medio del IMC por Sexo y Grupos de Edad

discrimina mejor en hombres.

En la Tabla 6.4 se presentan las *AUCs*, sus varianzas e Intervalos de Confianza (IC) al 95 % para el IMC, por grupos de edad, para los hombres. Se observa que las áreas bajo las curvas *ROC* en los grupos de edad estudiados variaron entre 0.721 (30-39 años) y 0.574 ( $\geq 70$  años). En general la capacidad de discriminación del IMC disminuye a medida que aumenta la edad, lo que indicaría que el marcador discrimina mejor en las poblaciones jóvenes.

Los resultados de la comparación de las diferentes curvas *ROC*, para cada grupo de edad en hombres, se presentan en la Tabla 6.5. Se encontraron diferencias significativas entre la curva *ROC* del grupo de edad **30-39 años** y las curvas de los grupos de edad de la población más adulta (**> 60 años**) y la curva *ROC* para la población de **40-49 años** y la de la población del grupo  $\geq 70$  años.

En la Tabla 6.6 se muestran las *AUCs* y sus ICs al 95 %, por grupos de edad para las mujeres. Las áreas bajo las curvas *ROC* para el IMC en los grupos de edad estudiados

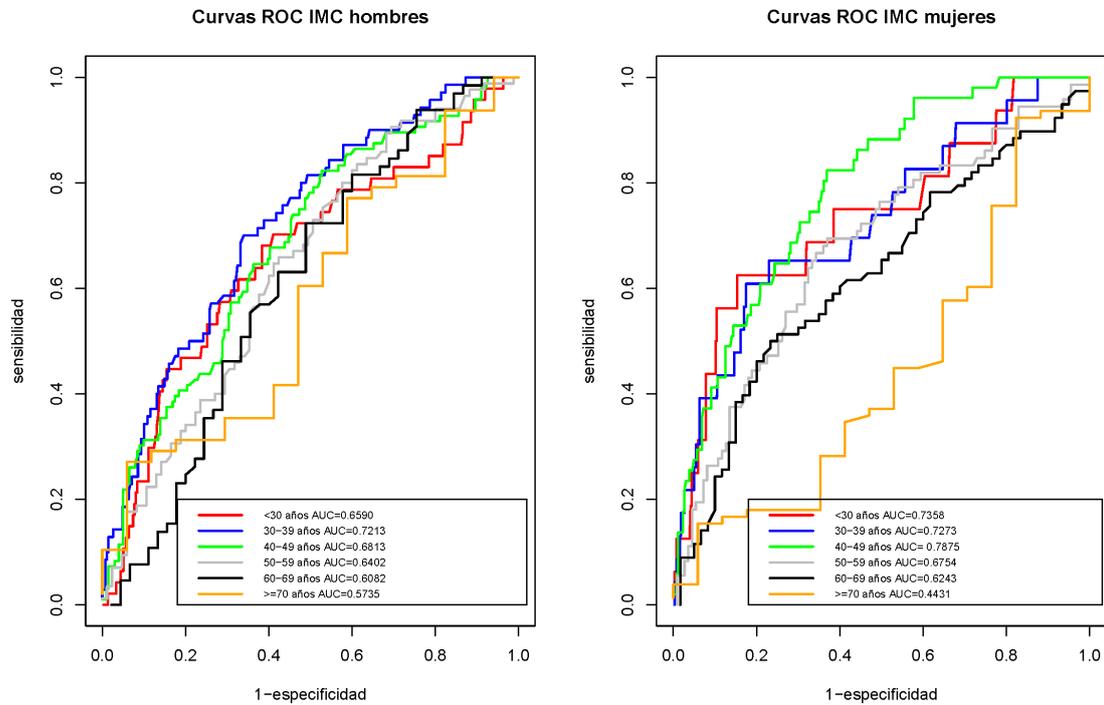


Figura 6.2: Curvas ROC del IMC por Sexo y Grupos de Edad

Grupo Edad	<i>AUC</i>	Varianza	Intervalo de Confianza (95 %)
H < 30 años	0.6589655	0.002098408	0.5691828 - 0.7487483
H 30-39 años	0.7213486	0.001109919	0.6560515 - 0.7866457
H 40-49 años	0.6812552	0.001081961	0.6167857 - 0.7457246
H 50-59 años	0.6402076	0.001797279	0.5571163 - 0.7232990
H 60-69 años	0.6082051	0.003266718	0.4961830 - 0.7202273
H $\geq$ 70 años	0.5735294	0.006676025	0.4133867 - 0.7336721

Tabla 6.4: *AUC* para el IMC e IC (95 %) en Hombres por Grupo de Edad

Curvas a Comparar	$\chi^2$	gl	p-valor
H <30 años H 30-39 años	1.2130	1	0.2707
H <30 años H 40-49 años	0.1549	1	0.6939
H <30 años H 50-59 años	0.1097	1	0.7405
H <30 años H 60-69 años	0.8031	1	0.3702
H <30 años H $\geq 70$ años	2.2751	1	0.1315
H 30-39 años H 40-49 años	0.5010	1	0.4790
H 30-39 años H 50-59 años	2.0521	1	0.1520
<b>H 30-39 años H 60-69 años</b>	3.9901	1	<b>0.0458</b>
<b>H 30-39 años H <math>\geq 70</math> años</b>	6.8106	1	<b>0.0091</b>
H 40-49 años H 50-59 años	0.5252	1	0.4686
H 40-49 años H 60-69 años	1.6633	1	0.1972
<b>H 40-49 años H <math>\geq 70</math> años</b>	3.6171	1	<b>0.0572</b>
H 50-59 años H 60-69 años	0.3192	1	0.5721
H 50-59 años H $\geq 70$ años	1.3858	1	0.2391
H 60-69 años H $\geq 70$ años	0.3748	1	0.5404

Tabla 6.5: Resultados Prueba Ji-Cuadrado IMC para cada Grupo de Edad - Hombres

variaron entre 0.788 (40-49 años) y 0.443 ( $\geq 70$  años).

Los resultados de la comparación de las curvas *ROC* en mujeres para cada uno de los grupos de edad, mediante la prueba de homogeneidad de áreas mencionada anteriormente, evidencian claramente las diferencias entre el poder de discriminación en las poblaciones jóvenes y en las poblaciones más adultas (ver Tabla 6.7).

También resulta de interés la comparación de la capacidad discriminativa del marcador IMC en cada uno de los grupos de edad en hombres y en mujeres (ver Figura 6.3), para ver si el comportamiento en mujeres y hombres de la misma edad es similar. Los resultados obtenidos se presentan en la Tabla 6.8.

Destaca la importante diferencia en la discriminación entre hombres y mujeres mayores ( $\geq 70$  años), siendo más elevada en hombres, pudiéndose considerar que hay también diferencias significativas en las curvas *ROC* en hombres y mujeres con edades comprendidas entre los 40 y los 49 años. En el resto de los grupos de edad el comportamiento parece que

Grupo Edad	AUC	Varianza	Intervalo de Confianza (95 %)
M < 30 años	0.7358398	0.005372287	0.5921825 - 0.8794971
M 30-39 años	0.7273243	0.003596026	0.6097914 - 0.8448572
M 40-49 años	0.7875196	0.001033667	0.7245054 - 0.8505338
M 50-59 años	0.6754254	0.001710256	0.5943706 - 0.7564802
M 60-69 años	0.6242521	0.002296694	0.5303231 - 0.7181811
M $\geq$ 70 años	0.4430618	0.006356449	0.2867991 - 0.5993246

Tabla 6.6: AUC para el IMC e IC (95 %) en Mujeres por Grupo de Edad

Curvas a Comparar	$\chi^2$	gl	p-valor
M <30 años M 30-39 años	0.0226	1	0.8805
M <30 años M 40-49 años	0.8325	1	0.3616
M <30 años M 50-59 años	1.1376	1	0.2862
<b>M &lt;30 años M 60-69 años</b>	3.8811	1	<b>0.0488</b>
<b>M &lt;30 años M <math>\geq</math>70 años</b>	26.7177	1	<b>2.3546e-07</b>
M 30-39 años M 40-49 años	1.1294	1	0.2879
M 30-39 años M 50-59 años	0.8395	1	0.3595
<b>M 30-39 años M 60-69 años</b>	3.3113	1	<b>0.0688</b>
<b>M 30-39 años M <math>\geq</math>70 años</b>	25.1861	1	<b>5.2057e-07</b>
<b>M 40-49 años M 50-59 años</b>	3.9164	1	<b>0.0478</b>
<b>M 40-49 años M 60-69 años</b>	8.3085	1	<b>0.0039</b>
<b>M 40-49 años M <math>\geq</math>70 años</b>	36.9823	1	<b>1.1921e-09</b>
M 50-59 años M 60-69 años	0.8162	1	0.3663
<b>M 50-59 años M <math>\geq</math>70 años</b>	16.8290	1	<b>4.0904e-05</b>
<b>M 60-69 años M <math>\geq</math>70 años</b>	10.2327	1	<b>0.0014</b>

Tabla 6.7: Resultados Prueba Ji-Cuadrado IMC para cada Grupo de Edad - Mujeres

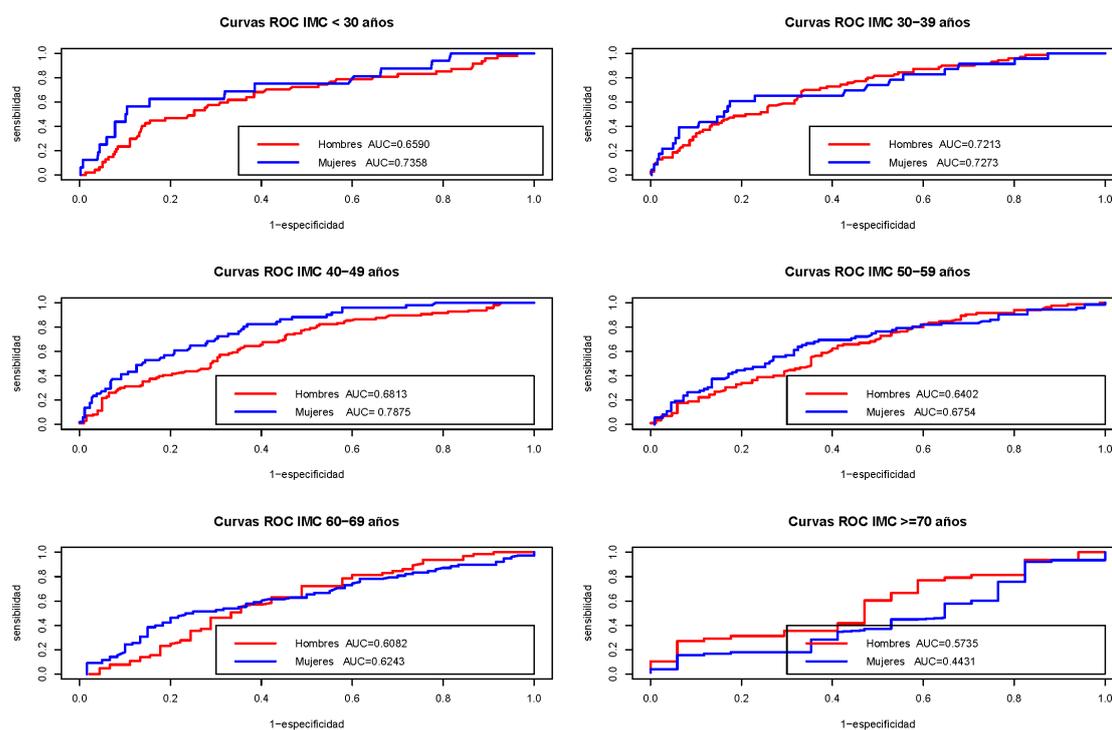


Figura 6.3: Curvas *ROC* del IMC en Cada Grupo de Edad por Sexo

es similar en ambos sexos, aunque el *AUC* es ligeramente superior en mujeres.

A modo de resumen se representó como varían las medidas globales de la curva *ROC*: el *AUC* junto con sus correspondientes intervalos de confianza al 95% y el Índice de Youden (obtenido mediante los diferentes métodos de estimación revisados en este trabajo) por grupos de edad en hombres y en mujeres (ver Figuras 6.4 y 6.5)<sup>1</sup>. Se observa que en las mujeres el *AUC* presenta un comportamiento más variable que en los hombres.

Los procedimientos Normal y Transformación Normal proporcionaron Índices de Youden similares para la población masculina con edades comprendidas entre los 18 y los 59 años. Pero a partir de los 60 años dichas estimaciones fueron diferentes. En la población de 60-69 años, se obtuvieron valores del *YI* similares mediante el método de la Transformación Normal y el método Empírico.

En los hombres con edades de 30 a 59 años, las estimaciones del *YI* obtenidas mediante el

<sup>1</sup>grupo 1 : <30, grupo 2: 30-39, grupo 3: 40-49, grupo 4: 50-59, grupo 5: 60-69 y grupo 6: ≥70 años.

Curvas a Comparar	$\chi^2$	gl	p-valor
H <30 años M <30 años	1.8420	1	0.1747
H 30-39 años M 30-39 años	0.0111	1	0.9160
<b>H 40-49 años M 40-49 años</b>	3.5196	1	<b>0.0606</b>
H 50-59 años M 50-59 años	0.3866	1	0.5341
H 60-69 años M 60-69 años	0.0803	1	0.7769
<b>H <math>\geq 70</math> años M <math>\geq 70</math> años</b>	5.3055	1	<b>0.0213</b>

Tabla 6.8: Resultados Prueba Ji-Cuadrado IMC por Grupos de Edad - Hombres y Mujeres

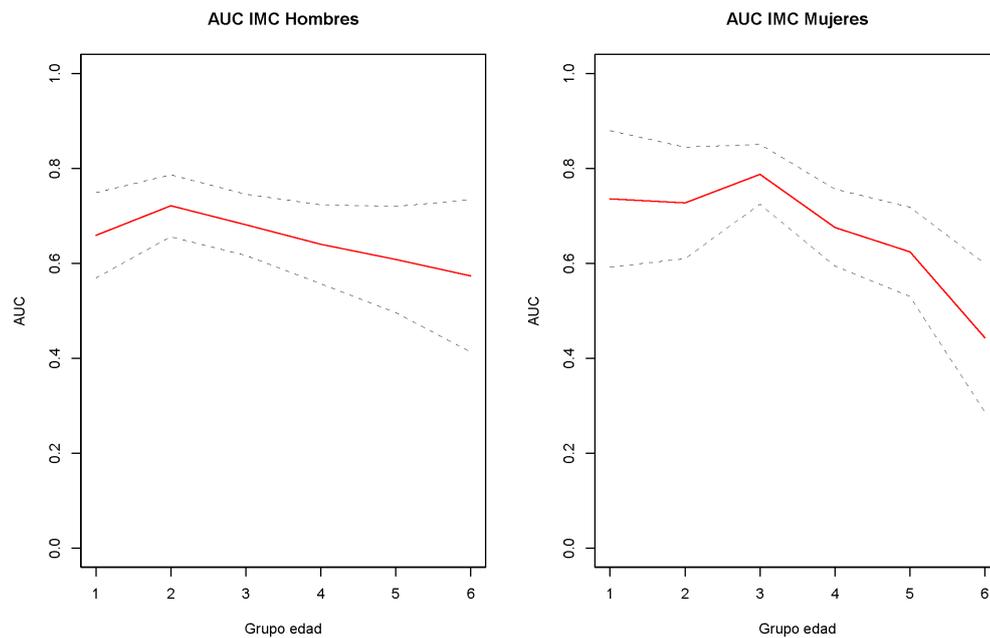


Figura 6.4: AUC del IMC e IC para Cada Grupo de Edad - Hombres y Mujeres

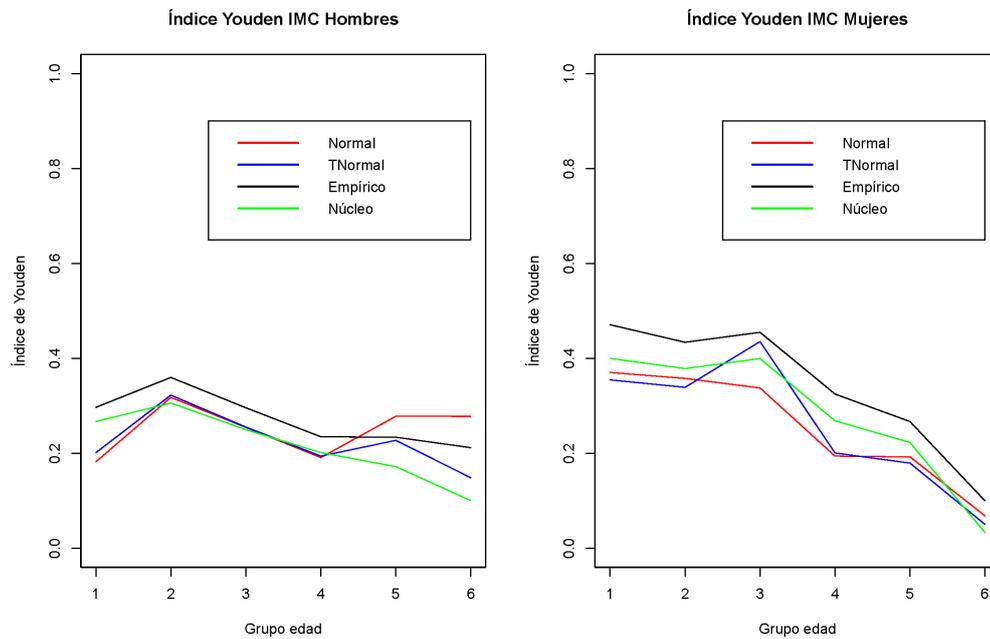


Figura 6.5:  $YI$  IMC para Cada Grupo de Edad Hombres y Mujeres-Métodos Examinados

método Núcleo fueron similares a las de los métodos  $N$  y  $TN$ .

En realidad se puede decir que hasta los 59 años todos los métodos de estimación siguieron el mismo comportamiento (en el grupo de edad de 30-39 años el Índice de Youden fue superior al de la población de 18 a 30 años, y luego a partir de los 39 años hasta los 59 años disminuyó progresivamente). En esta población masculina (18 a 59 años) el método  $EMP$  fue el que detectó los valores del  $YI$  más elevados, una discriminación en torno al 30% hasta los 49 años y en torno al 24% para la población con edades entre los 50 y los 59 años. En los hombres de 60 a 69 años, los métodos  $N$  y  $TN$  detectaron un incremento del  $YI$  con respecto al grupo de edad anterior, siendo más pronunciado dicho incremento con el método  $N$ , el método  $EMP$  proporcionó un valor  $YI$  similar al de la población de 50 a 59 años y el método Núcleo seleccionó un  $YI$  inferior.

En la población femenina, los cuatro métodos de estimación proporcionaron  $YIs$  más diferentes entre sí que en el caso de los varones. Los métodos  $TN$ ,  $EMP$  y  $K$  parece que siguieron el mismo comportamiento (el  $YI$  presentó la misma forma), mientras que el método  $N$ , en general, siguió un patrón más diferente.

En todos los grupos de edad el método *EMP* fue el que proporcionó los valores del *YI* más elevados, un valor en torno al 40 % para las mujeres con edades entre los 18 y los 49 años, al 30 % para las mujeres de 50 a 69 años y del 10 % en la población femenina mayor de 70 años.

En las mujeres de 18 a 39 años y de 60 a 69 años el método *TN* seleccionó el menor *YI*, en las poblaciones femeninas de 40 a 59 años fue el método *N* y en las mujeres mayores de 70 años el método *K*.

En las mujeres menores de 30 años los métodos *N* y *TN* proporcionaron *YIs* similares y en las mujeres con edades comprendidas entre los 30 y los 39 años y en las mayores de 70 años fueron los métodos *N*, *TN* y *K* los que dieron una discriminación similar. En la población femenina de 40 a 49 años, los procedimientos *TN* y *EMP* originaron un *YI* similar, en la de 50 a 69 años los métodos *N* y *TN*.

Los puntos de corte óptimos para el IMC para hombres y mujeres, obtenidos mediante los métodos de estimación estudiados, por grupos de edad, se muestran en las Tablas 6.9 y 6.10. En el caso del método de Transformación Normal se agregó el valor de la potencia de Box-Cox obtenida para cada grupo de edad y en el método Núcleo se añadieron los valores de las ventanas de suavización utilizadas en enfermos y sanos. Se observa que los valores varían entre  $22.857 \text{ kg/m}^2$  y  $35.422 \text{ kg/m}^2$  para los hombres, según el grupo de edad y el método considerado, y entre  $23.390 \text{ kg/m}^2$  y  $34.730 \text{ kg/m}^2$  para las mujeres.

Teniendo en cuenta los valores límite del IMC (ver Tabla 6.1 en la página 49) los métodos *N* y *TN* seleccionaron el punto de corte óptimo para detectar la HTA dentro de los límites de peso normal. Sin embargo para la población de 60 a 69 años el método *N* seleccionó el punto de corte óptimo para el IMC dentro de la categoría de Obesidad Grado II y en los hombres mayores de 70 años los métodos *N* y *EMP* proporcionaron el punto de corte óptimo dentro de los valores definidos para la Obesidad de Grado I.

En las restantes situaciones (grupo de edad y método de estimación) los puntos de corte óptimos se incluyeron en la categoría de sobrepeso.

En la Tablas 6.11 y 6.12 se muestran los valores de las medidas de Sensibilidad y Especificidad en los puntos de corte óptimos obtenidos por maximización del Índice de Youden para cada grupo de edad, en hombres y en mujeres. Dichas medidas permiten visualizar los puntos de corte óptimos sobre la curva *ROC* (ver Figuras 6.6 y 6.7).

Hombres							
Método		Grupos de Edad					
		1	2	3	4	5	6
$N$	$c^*$	22.857	28.457	29.122	28.466	35.422	32.793
	$YI$	0.182	0.318	0.255	0.191	0.278	0.278
$TN$	$c^*$	24.703	27.513	28.088	28.192	26.329	26.461
	$YI$	0.201	0.323	0.255	0.194	0.227	0.149
	Potencia	-1.852	-0.200	-0.300	0.500	-1.260	-1.700
$EMP$	$c^*$	25.060	27.970	26.600	28.375	28.160	31.630
	$YI$	0.297	0.360	0.296	0.235	0.234	0.212
$K$	$c^*$	26.350	27.333	27.768	27.284	28.130	26.111
	$YI$	0.267	0.306	0.250	0.202	0.172	0.101
	$h_{enf}$	1.734	1.185	1.479	1.141	1.227	1.568
	$h_{sanos}$	0.936	1.025	1.081	1.336	1.594	1.893

Tabla 6.9:  $c^*$  e  $YI$  para IMC Hombres

Mujeres							
Método		Grupos de Edad					
		1	2	3	4	5	6
$N$	$c^*$	27.961	30.602	32.574	25.647	31.804	25.239
	$YI$	0.370	0.358	0.338	0.194	0.193	0.068
$TN$	$c^*$	24.434	26.605	27.025	29.258	30.896	25.979
	$YI$	0.355	0.339	0.436	0.201	0.179	0.050
	Potencia	-1.970	-1.752	-1.900	-1.055	-1.000	-0.500
$EMP$	$c^*$	26.930	28.380	26.620	28.430	31.040	23.390
	$YI$	0.471	0.434	0.455	0.325	0.267	0.100
$K$	$c^*$	26.170	27.948	27.773	29.246	31.292	34.730
	$YI$	0.400	0.379	0.400	0.269	0.223	0.033
	$h_{enf}$	3.921	3.887	2.399	2.074	1.879	1.317
	$h_{sanos}$	0.960	1.029	1.222	1.265	1.385	1.444

Tabla 6.10:  $c^*$  e  $YI$  para IMC Mujeres

Hombres							
Método	Grupos de Edad						
		1	2	3	4	5	6
$N$	$Se$	0.809	0.571	0.521	0.635	0.046	0.125
	$Sp$	0.333	0.724	0.701	0.588	0.956	0.941
$TN$	$Se$	0.702	0.729	0.646	0.659	0.846	0.750
	$Sp$	0.572	0.608	0.632	0.553	0.289	0.412
$EMP$	$Se$	0.681	0.700	0.823	0.659	0.723	0.271
	$Sp$	0.603	0.653	0.468	0.576	0.489	0.882
$K$	$Se$	0.532	0.729	0.656	0.788	0.723	0.771
	$Sp$	0.741	0.590	0.597	0.424	0.489	0.412

Tabla 6.11:  $Se$  y  $Sp$  para IMC Hombres

En relación a la estimación del punto de corte óptimo, se observó que en los hombres menores de 30 años los métodos de la Transformación Normal y el Empírico proporcionaron óptimos similares, mientras que en las mujeres del mismo grupo de edad, los procedimientos  $N$ ,  $EMP$  y  $K$  originaron puntos de corte óptimos análogos.

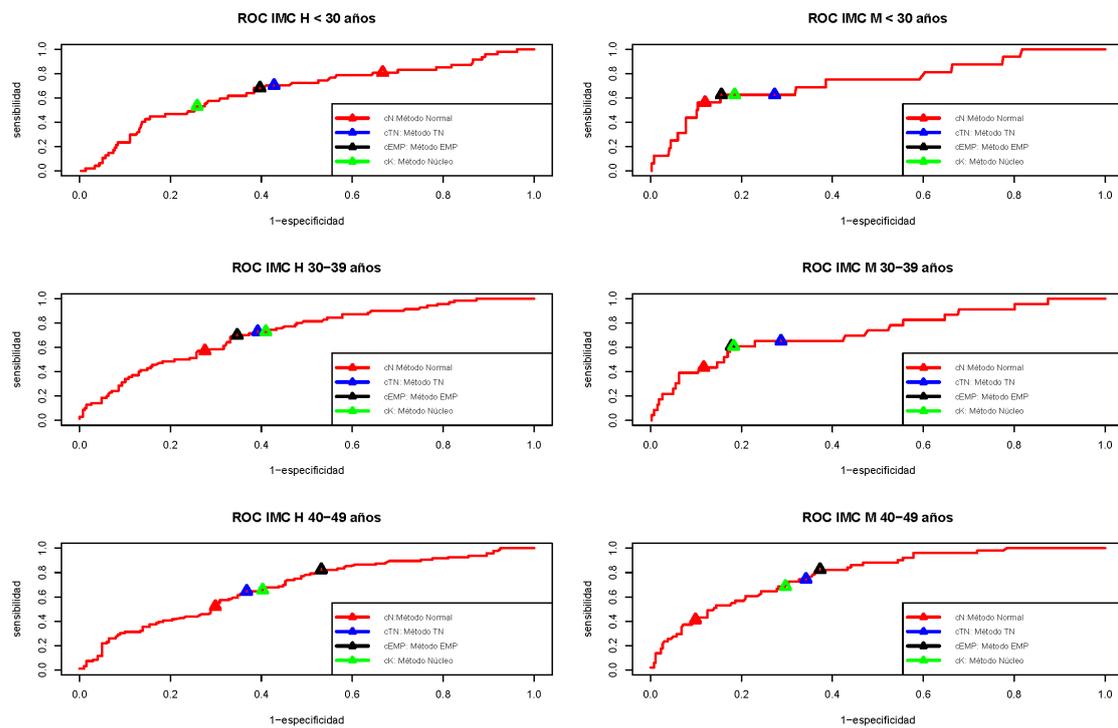
En la población masculina de 30 a 39 años los métodos  $TN$ ,  $EMP$  y  $K$  seleccionaron óptimos similares, mientras que el método Normal proporcionó un óptimo más elevado que los métodos anteriores. En la población femenina del mismo grupo los métodos  $EMP$  y  $K$  detectaron puntos de corte similares en los que la medida de Sensibilidad es exactamente la misma y la medida de Especificidad muy similar, por eso la ubicación de los dos puntos de corte óptimos prácticamente coincide (ver Figura 6.6).

Nótese que cuando la suposición de normalidad es rechazada no es sorprendente que las estimaciones del Índice de Youden y del punto de corte óptimo mediante el método Normal sean bastante diferentes de las estimaciones proporcionadas por los otros métodos, como se ve en la población de 40 a 49 años (ver Figura 6.6).

### 6.1.2. Estudio de la Circunferencia de la Cintura como Marcador Diagnóstico de Hipertensión Arterial

Los valores medios del valor de la Circunferencia de la Cintura (CC) se presentan en la Figura 6.8. En todos los grupos de edad los hombres tienen un valor de cintura más

Mujeres							
Método	Grupos de Edad						
	1	2	3	4	5	6	
$N$	$Se$	0.562	0.435	0.412	0.833	0.449	0.833
	$Sp$	0.880	0.883	0.901	0.288	0.800	0.176
$TN$	$Se$	0.625	0.652	0.745	0.569	0.500	0.795
	$Sp$	0.727	0.713	0.658	0.694	0.750	0.176
$EMP$	$Se$	0.625	0.609	0.824	0.694	0.500	0.923
	$Sp$	0.844	0.822	0.627	0.622	0.750	0.118
$K$	$Se$	0.625	0.6097	0.686	0.569	0.474	0.103
	$Sp$	0.815	0.817	0.703	0.685	0.783	0.941

Tabla 6.12:  $Se$  y  $Sp$  para IMC MujeresFigura 6.6:  $c^*$  sobre Curva  $ROC$  IMC-Hombres y Mujeres (1)

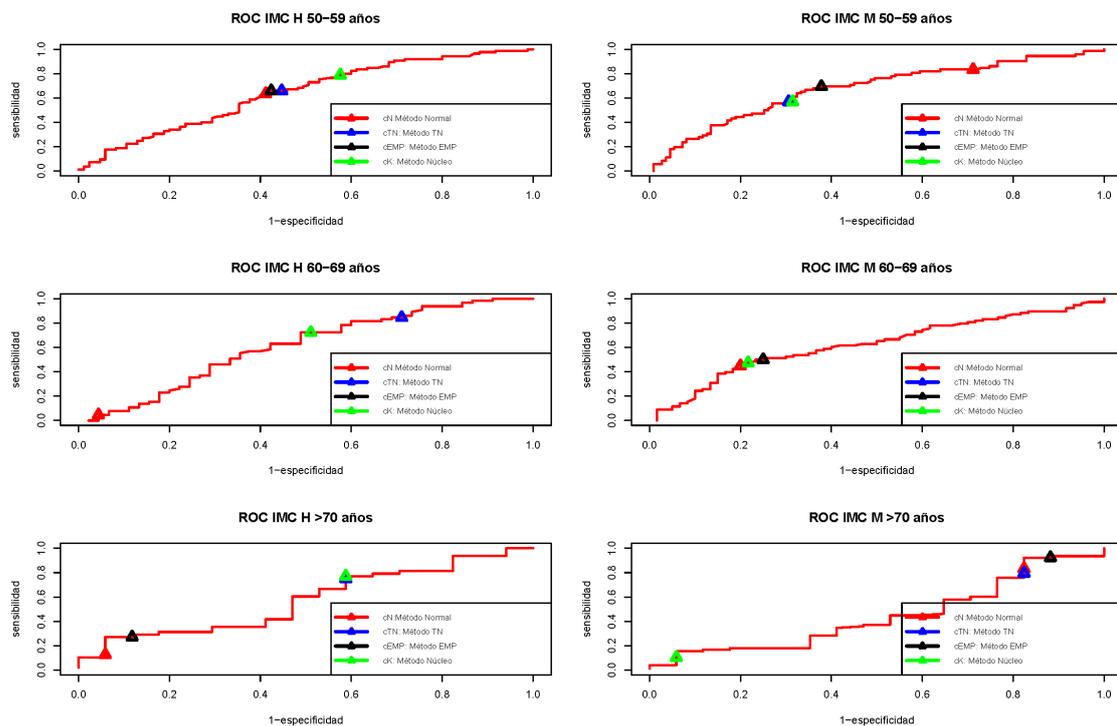


Figura 6.7:  $c^*$  sobre Curva *ROC* IMC-Hombres y Mujeres (2)

elevado que las mujeres. Hasta los 39 años los valores de cintura son menores y a partir de dicha edad van aumentando excepto en el último grupo de edad, donde se apreció una disminución de la cintura con respecto al grupo anterior (60-69 años), siendo superior dicha disminución en hombres.

Las curvas *ROC* calculadas, en hombres y en mujeres, para los diferentes grupos de edad considerados se muestran en la Figura 6.9. En general se puede decir que la capacidad de discriminación del valor de la medida de la CC, como marcador de HTA, fue mejor en las mujeres.

En la Tabla 6.13 se muestran las áreas bajo las curvas *ROC* del valor de esta medida en hombres. Se observó que la discriminación disminuye con la edad, aunque en este estudio para los hombres jóvenes menores de 30 años se detectó una discriminación menor que para la de los hombres con edades comprendidas entre los 30 y los 49 años, pero dichas diferencias no fueron significativas (ver Tabla 6.14).

Las *AUCs* de las curvas *ROC* de los diferentes grupos de edad en hombres variaron entre

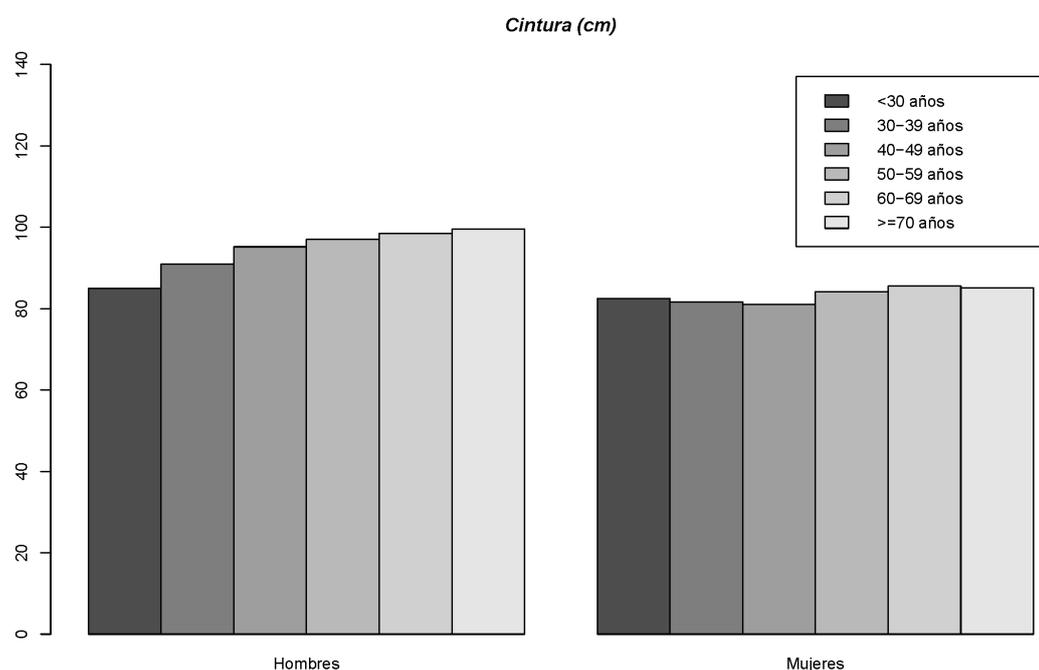


Figura 6.8: Valor Medio de Circunferencia de la Cintura

0.671 (30-39 años) y 0.553 ( $\geq 70$  años), resultando dicha diferencia estadísticamente significativa ( $p = 0,0437$ , ver Tabla 6.14). Se observó así, un comportamiento del valor de CC similar al del IMC, en el sentido de que la discriminación más elevada y la menos elevada se alcanzaron en los mismos grupos de edad, aunque la capacidad de discriminación de la CC fue ligeramente inferior a la del IMC.

En el caso de las mujeres también se puede decir que la capacidad de clasificación de la CC para detectar HTA disminuyó con la edad, puesto que aunque en la muestra aquí analizada la discriminación resultó superior en las mujeres de 40 a 49 años que en las mujeres más jóvenes, dicha diferencia no resultó significativa, al igual que en el caso de las mujeres mayores de 70 años donde el  $AUC$  fue ligeramente superior al de las mujeres con edades comprendidas entre los 60 y los 69 años (ver Tabla 6.15).

Cabe destacar que se detectaron diferencias significativas entre el comportamiento del marcador en la población menor de 60 años y en la población mayor (o igual) de 60 años (ver Tabla 6.16), por lo que se diferencian claramente ambas poblaciones.

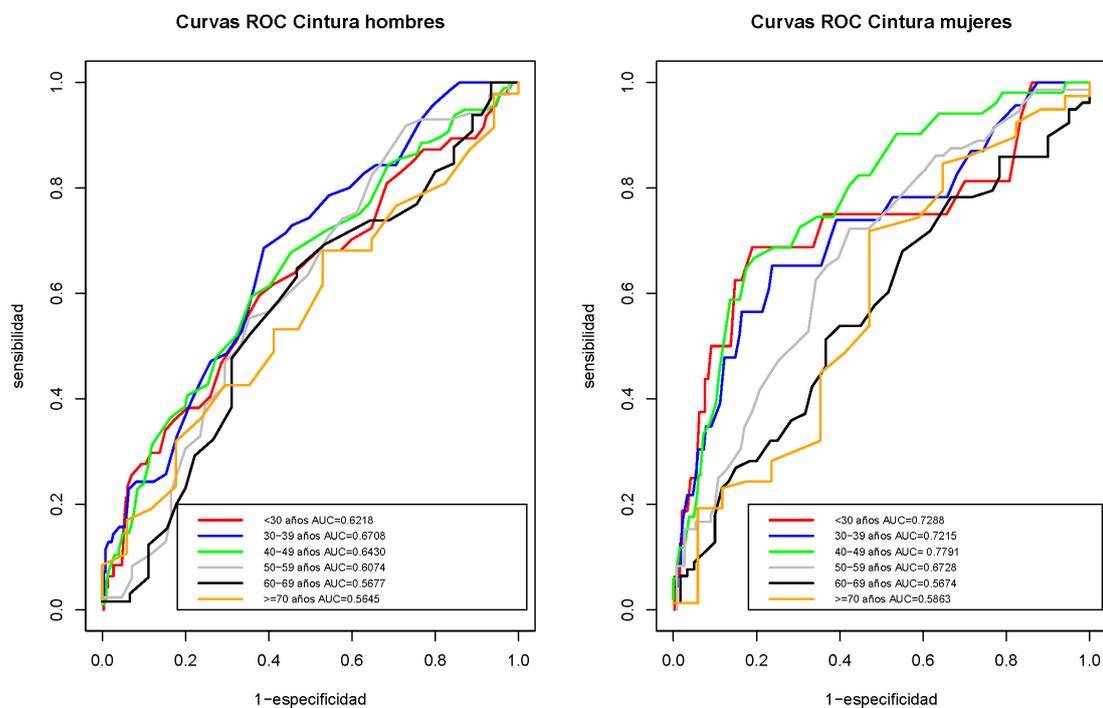


Figura 6.9: Curvas *ROC* Cintura por Sexo y Grupos de Edad

Grupo Edad	<i>AUC</i>	Varianza	Intervalo de Confianza (95 %)
H < 30 años	0.6218210	0.002242605	0.5290047 - 0.7146374
H 30-39 años	0.6707889	0.001218296	0.6023782 - 0.7391997
H 40-49 años	0.6429830	0.001207891	0.5748650 - 0.7111010
H 50-59 años	0.6074048	0.001901586	0.5219364 - 0.6928733
H 60-69 años	0.5676923	0.003237702	0.4561688 - 0.6792158
H $\geq$ 70 años	0.5644556	0.006243678	0.4095852 - 0.7193259

Tabla 6.13: *AUC* para CC e IC (95 %) en Hombres por Grupo de Edad

Curvas a Comparar	$\chi^2$	gl	p-valor
H <30 años H 30-39 años	0.69284	1	0.4052
H <30 años H 40-49 años	0.1294	1	0.7191
H <30 años H 50-59 años	0.0600	1	0.8064
H <30 años H 60-69 años	0.8466	1	0.3575
H <30 años H $\geq 70$ años	0.9509	1	0.3295
H 30-39 años H 40-49 años	0.2234	1	0.6365
H 30-39 años H 50-59 años	1.1608	1	0.2813
<b>H 30-39 años H 60-69 años</b>	3.0711	1	<b>0.0797</b>
<b>H 30-39 años H <math>\geq 70</math> años</b>	3.2670	1	<b>0.0707</b>
H 40-49 años H 50-59 años	0.3657	1	0.5453
H 40-49 años H 60-69 años	1.6379	1	0.2006
H 40-49 años H $\geq 70$ años	1.7818	1	0.1819
H 50-59 años H 60-69 años	0.4557	1	0.4996
H 50-59 años H $\geq 70$ años	0.5330	1	0.4654
H 60-69 años H $\geq 70$ años	0.0030	1	0.9561

Tabla 6.14: Resultados Prueba Ji-Cuadrado CC para cada Grupo de Edad - Hombres

Estudiando la discriminación en hombres y en mujeres de un mismo grupo de edad se detectaron diferencias por sexo en las poblaciones jóvenes, menores de 30 años, y en las poblaciones con edades comprendidas entre los 40 y los 49 años (ver Tabla 6.17).

En la Figura 6.11 se representó cómo varió el *AUC* en hombres y en mujeres por grupos de edad. A la vista de los resultados obtenidos se detectó una clara diferencia, en términos de discriminación, entre hombres y mujeres, siendo más variable en estas últimas (al igual que sucedía con el IMC).

En la Figura 6.12 puede verse el comportamiento del Índice de Youden por grupos de edad y sexo mediante los diferentes métodos de estimación estudiados. En la población masculina los métodos *N* y *TN* proporcionaron Índices de Youden prácticamente iguales en todos los grupos de edad. En los hombres jóvenes menores de 30 años, el método *K* proporcionó casi el mismo valor del *YI* que el método *N*.

El método *EMP* fue el que proporcionó los *YIs* más elevados en todos los grupos. Al igual

Grupo Edad	AUC	Varianza	Intervalo de Confianza (95 %)
M < 30 años	0.7287598	0.006643993	0.5690017 - 0.8885178
M 30-39 años	0.7214780	0.003760800	0.6012825 - 0.8416735
M 40-49 años	0.7791322	0.001255311	0.7096899 - 0.8485744
M 50-59 años	0.6727978	0.001623154	0.5938340 - 0.7517616
M 60-69 años	0.5674145	0.002422785	0.4709416 - 0.6638875
M $\geq$ 70 años	0.5863499	0.006862088	0.4239910 - 0.7487089

Tabla 6.15: AUC para CC e IC (95 %) en Mujeres por Grupo de Edad

Curvas a Comparar	$\chi^2$	gl	p-valor
M <30 años M 30-39 años	0.0153	1	0.9015
M <30 años M 40-49 años	0.7332	1	0.3919
M <30 años M 50-59 años	0.9049	1	0.3415
<b>M &lt;30 años M 60-69 años</b>	7.5218	1	<b>0.0061</b>
<b>M &lt;30 años M <math>\geq</math>70 años</b>	5.8599	1	0.0155
M 30-39 años M 40-49 años	0.9604	1	0.3271
M 30-39 años M 50-59 años	0.6847	1	0.4080
<b>M 30-39 años M 60-69 años</b>	6.8582	1	<b>0.0088</b>
<b>M 30-39 años M <math>\geq</math>70 años</b>	5.2760	1	0.0216
<b>M 40-49 años M 50-59 años</b>	3.2671	1	<b>0.0707</b>
M 40-49 años M 60-69 años	1.2952e+01	1	3.1964e-04
<b>M 40-49 años M <math>\geq</math>70 años</b>	10.7385	1	<b>0.0010</b>
<b>M 50-59 años M 60-69 años</b>	3.2089	1	<b>0.0732</b>
M 50-59 años M $\geq$ 70 años	2.1593	1	0.1417
M 60-69 años M $\geq$ 70 años	0.1036	1	0.7476

Tabla 6.16: Resultados Prueba Ji-Cuadrado CC para cada Grupo de Edad - Mujeres

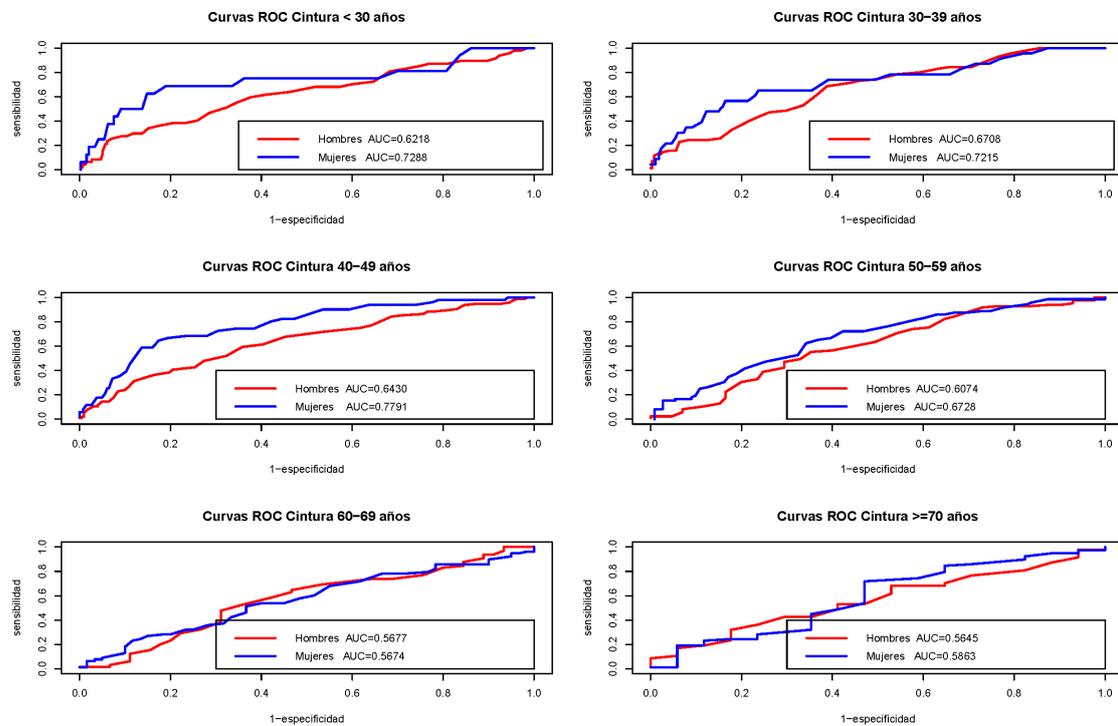


Figura 6.10: Curvas ROC CC en Cada Grupo de Edad por Sexo

Curvas a Comparar	$\chi^2$	gl	p-valor
H <30 años M <30 años	3.3043	1	<b>0.0691</b>
H 30-39 años M 30-39 años	0.7424	1	0.3889
<b>H 40-49 años M 40-49 años</b>	5.3560	1	<b>0.0207</b>
H 50-59 años M 50-59 años	1.2356	1	0.2663
H 60-69 años M 60-69 años	2.2295e-05	1	9.9623e-01
H $\geq$ 70 años M $\geq$ 70 años	0.1385	1	0.7098

Tabla 6.17: Resultados Prueba Ji-Cuadrado CC por Grupos de Edad - Hombres y Mujeres

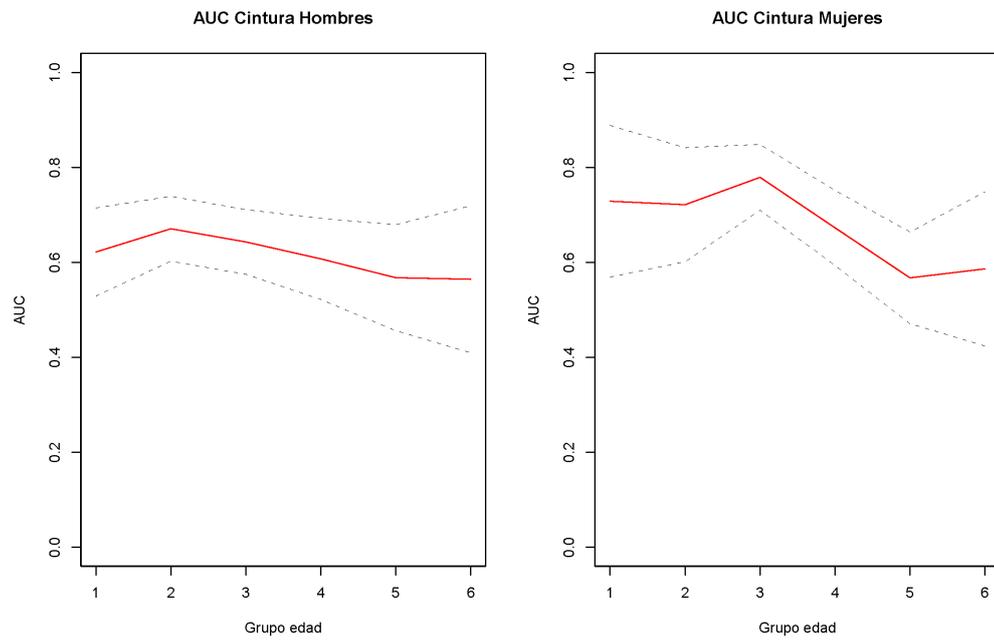


Figura 6.11: *AUC* e IC CC para Cada Grupo de Edad - Hombres y Mujeres

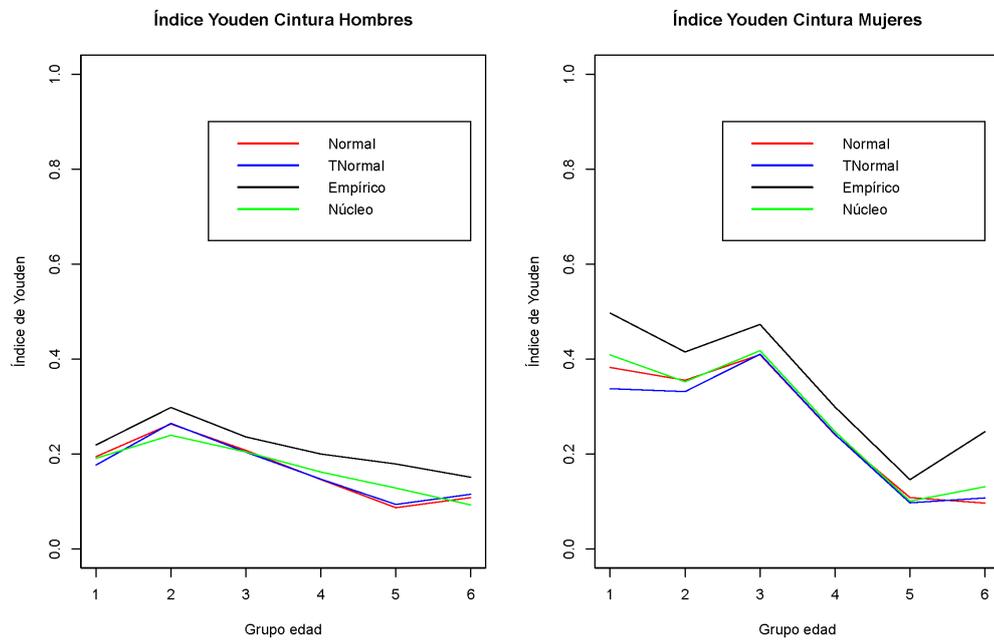


Figura 6.12: *YI* CC para Cada Grupo de Edad Hombres y Mujeres-Métodos Examinados

que los métodos  $N$  y  $TN$  siguieron el mismo patrón, los métodos  $EMP$  y  $K$  tuvieron un comportamiento similar aunque este último siempre registró unos valores inferiores a los del procedimiento  $EMP$ .

Los métodos paramétricos detectaron un decremento más pronunciado de la discriminación, desde los 39 años hasta los 69 años, que los métodos no paramétricos. En la población masculina de más (o igual) de 70 años el comportamiento de ambos grupos de métodos también fue diferente, las técnicas de estimación paramétricas detectaron un aumento de la discriminación con respecto al grupo de edad anterior, mientras que las no paramétricas registraron un  $YI$  inferior al de la población masculina de los 60 a los 69 años.

En las mujeres de 18 a 69 años, los métodos  $N$  y  $K$  proporcionaron Índices de Youden similares y en la población de 39 a 69 años, el método  $TN$  seleccionó estimaciones del  $YI$  similares a las de los métodos anteriores  $N$  y  $K$ . El método  $EMP$  fue el método que detectó los valores del  $YI$  más elevados y en las mujeres de más (o igual) de 70 años, los métodos  $N$  y  $TN$  dieron  $YIs$  próximos.

En realidad se puede decir que en la población femenina desde los 18 hasta los 69 años los cuatro métodos de estimación siguieron un patrón parecido, salvo en el último grupo de edad donde se detectaron quizás las mayores diferencias: el método  $EMP$  registró un pronunciado incremento del  $YI$  con respecto al grupo anterior, los métodos  $TN$  y  $K$  detectaron un incremento menos marcado, sin embargo el método  $N$  proporcionó un  $YI$  ligeramente inferior.

Los puntos de corte óptimos para la CC para hombres y mujeres, obtenidos mediante los métodos de estimación estudiados, por grupos de edad, se muestran en las Tablas 6.18 y 6.19, respectivamente, que tienen la misma configuración que para el marcador  $IMC^2$ . Los valores óptimos variaron entre 86 cm y 106.715 cm en hombres y entre 79.915 cm y 100.272 cm en mujeres, según el grupo de edad y el procedimiento de estimación considerado. Como resulta lógico, siempre se detectaron valores óptimos de la CC inferiores en mujeres que en hombres.

En hombres, los puntos de corte óptimos obtenidos fueron 94 cm en las poblaciones con edades a partir de los 40 años, por lo que se engloban en la categoría en la que parece haber un aumento del riesgo cardiovascular (ver Tabla 6.2 en la página 50). Sin embargo,

---

<sup>2</sup>grupo 1 : <30, grupo 2: 30-39, grupo 3: 40-49, grupo 4: 50-59, grupo 5: 60-69 y grupo 6:  $\geq 70$  años.

		<b>Hombres</b>					
<b>Método</b>		<b>Grupos de Edad</b>					
		1	2	3	4	5	6
<i>N</i>	$c^*$	91.552	93.939	98.542	94.679	98.678	106.715
	$YI$	0.194	0.263	0.208	0.146	0.087	0.108
<i>TN</i>	$c^*$	89.169	92.630	97.602	94.560	96.980	106.146
	$YI$	0.177	0.264	0.204	0.147	0.094	0.115
	Potencia	-0.700	0.305	0.600	0.900	-0.600	2.000
<i>EMP</i>	$c^*$	86	92	95	97	97	97
	$YI$	0.219	0.298	0.236	0.200	0.179	0.151
<i>K</i>	$c^*$	91.688	91.384	96.827	94.390	99.405	104.777
	$YI$	0.191	0.240	0.204	0.162	0.128	0.093
	$h_{enf}$	5.356	3.302	3.572	3.038	3.436	4.509
	$h_{sanos}$	3.010	2.897	2.790	3.315	3.451	4.776

Tabla 6.18:  $c^*$  e  $YI$  para CC Hombres

		<b>Mujeres</b>					
<b>Método</b>		<b>Grupos de Edad</b>					
		1	2	3	4	5	6
<i>N</i>	$c^*$	85.550	89.967	88.553	89.791	100.272	89.607
	$YI$	0.382	0.355	0.409	0.241	0.108	0.097
<i>TN</i>	$c^*$	79.915	85.422	85.930	88.700	98.867	88.907
	$YI$	0.338	0.331	0.410	0.241	0.097	0.107
	Potencia	-1.569	-0.678	-0.400	0.301	-0.752	-0.800
<i>EMP</i>	$c^*$	82	86	90	87	95	88
	$YI$	0.497	0.415	0.473	0.299	0.146	0.247
<i>K</i>	$c^*$	83.469	87.700	88.269	87.808	94.152	86.063
	$YI$	0.409	0.352	0.418	0.247	0.101	0.131
	$h_{enf}$	8.197	8.072	4.895	3.783	4.777	3.864
	$h_{sanos}$	2.656	3.066	3.195	3.837	4.220	6.098

Tabla 6.19:  $c^*$  e  $YI$  para CC Mujeres

		Hombres					
Método	Grupos de Edad						
	1	2	3	4	5	6	
<i>N</i>	<i>Se</i>	0.383	0.614	0.479	0.706	0.631	0.319
	<i>Sp</i>	0.768	0.634	0.726	0.459	0.533	0.824
<i>TN</i>	<i>Se</i>	0.468	0.686	0.521	0.706	0.692	0.319
	<i>Sp</i>	0.714	0.612	0.677	0.459	0.467	0.824
<i>EMP</i>	<i>Se</i>	0.617	0.714	0.615	0.565	0.692	0.681
	<i>Sp</i>	0.589	0.56	0.597	0.60	0.467	0.353
<i>K</i>	<i>Se</i>	0.383	0.714	0.542	0.706	0.585	0.426
	<i>Sp</i>	0.768	0.560	0.667	0.459	0.578	0.706

Tabla 6.20: *Se* y *Sp* para CC Hombres

en la población menor de 40 años, los valores óptimos fueron menores (<94 cm), lo que puede indicar que en la población joven sea necesario establecer un punto de corte menor, a partir del cual se detecte un aumento del riesgo de hipertensión arterial.

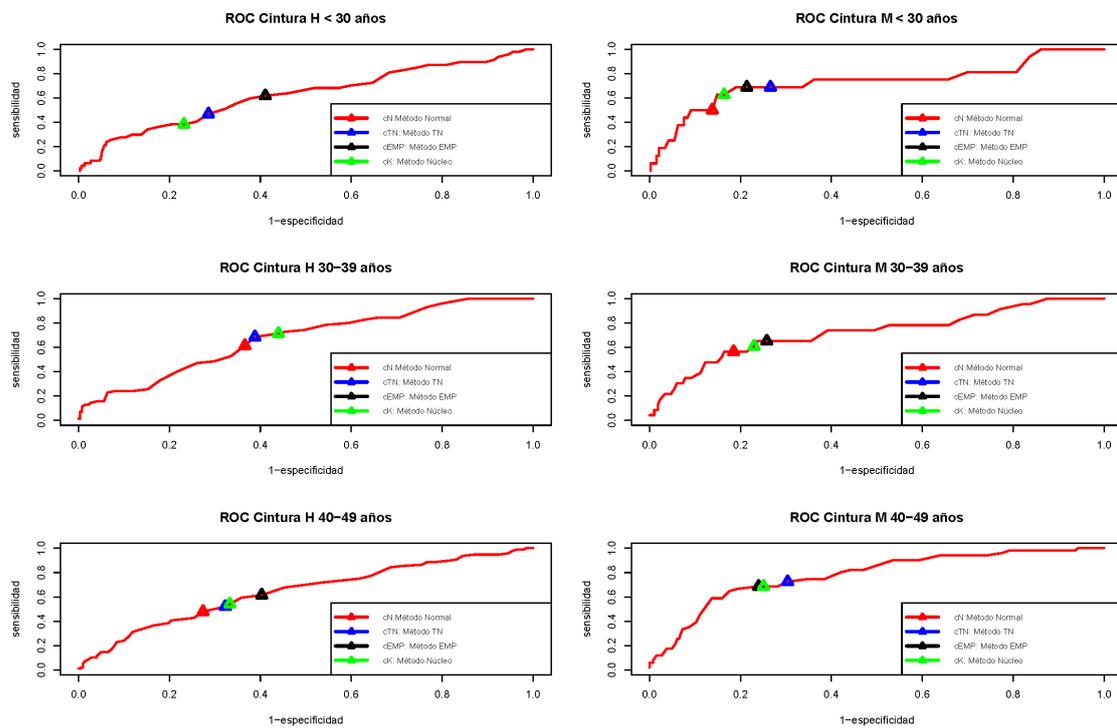
En las mujeres, sin embargo, los puntos óptimos que resultaron mediante los diferentes métodos de estimación fueron 80 cm, por lo que son valores que marcan un aumento del riesgo cardiovascular (Tabla ver Tabla 6.2).

Las medidas de Sensibilidad y Especificidad en los puntos de corte óptimos obtenidos por maximización del Índice de Youden para cada grupo de edad, en hombres y en mujeres se presentan en las Tablas 6.20 y 6.21. Estas medidas permiten visualizar los puntos de corte óptimos sobre la curva ROC (ver Figuras 6.13 y 6.14).

En la Figura 6.13 se nota que para la población masculina menor de 30 años sólo se ven tres puntos de corte óptimos, no siendo visible el punto producido por el método *N*. En este caso los métodos *N* y *K* proporcionaron puntos muy próximos (ver Tabla 6.18). Algo semejante ocurre con la población de hombres también, de 30 a 39 años en que los puntos provenientes de los procedimientos *EMP* Y *K* están posicionados muy cerca uno del otro.

En la población de mujeres de 40 a 49 años prácticamente coinciden los puntos proporcionados por los métodos *N* y *K*.

Mujeres							
Método	Grupos de Edad						
	1	2	3	4	5	6	
$N$	$Se$	0.500	0.565	0.686	0.653	0.321	0.679
	$Sp$	0.862	0.815	0.749	0.631	0.750	0.529
$TN$	$Se$	0.688	0.652	0.725	0.667	0.372	0.718
	$Sp$	0.734	0.742	0.696	0.604	0.683	0.529
$EMP$	$Se$	0.688	0.652	0.686	0.722	0.538	0.744
	$Sp$	0.786	0.742	0.76	0.532	0.60	0.412
$K$	$Se$	0.625	0.609	0.686	0.722	0.538	0.795
	$Sp$	0.836	0.770	0.749	0.577	0.600	0.353

Tabla 6.21:  $Se$  y  $Sp$  para CC MujeresFigura 6.13:  $c^*$  sobre Curva  $ROC$  CC-Hombres y Mujeres (1)

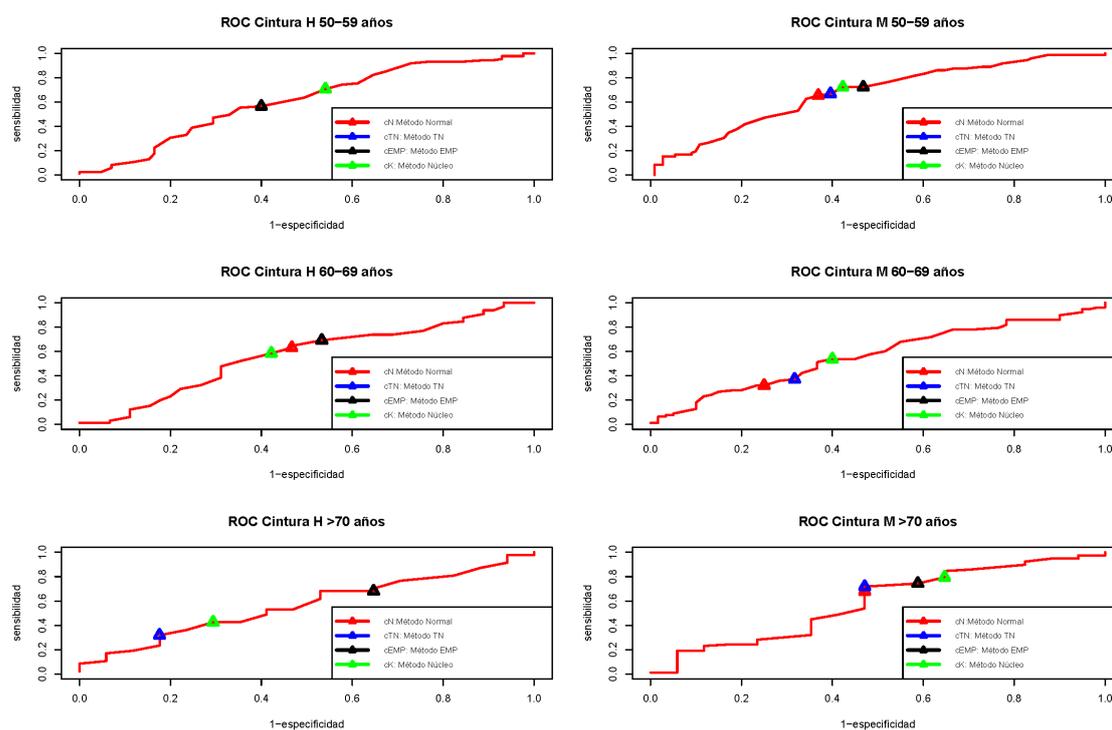


Figura 6.14:  $c^*$  sobre Curva ROC CC-Hombres y Mujeres (2)

### 6.1.3. Estudio del Marcador Índice Cintura/Cadera

Los valores medios del Índice de Cintura/Cadera ( $C/C$ ) se presentan en la Figura 6.15. En todos los grupos de edad los hombres tienen un valor del  $C/C$  más elevado que las mujeres. En hombres aumenta progresivamente con la edad, mientras que en mujeres los valores son más similares con el paso del tiempo y hay una ligera disminución del índice en los grupos de edad 30-39 años y 40-49 años con respecto a las mujeres de <30 años.

En general el Índice de Cintura/Cadera discriminó mejor en la población femenina (las áreas bajo las curvas ROC en mujeres fueron superiores a las áreas bajo las curvas ROC en hombres) (ver Figura 6.16).

En la Tabla 6.22 se presentan las áreas bajo las curvas ROC del  $C/C$  y sus IC al 95 % en hombres. Dichas  $AUC$ s variaron entre 0.633 (60-69 años) y 0.550 (50-59 años), pero dicha diferencia no resultó significativa ( $p = 0,1715$ , ver Tabla 6.23). De hecho no se detectaron diferencias significativas en la discriminación del Índice de Cintura/Cadera por grupos de edad en hombres, por lo que parece que la edad no influyó en la capacidad de discriminación

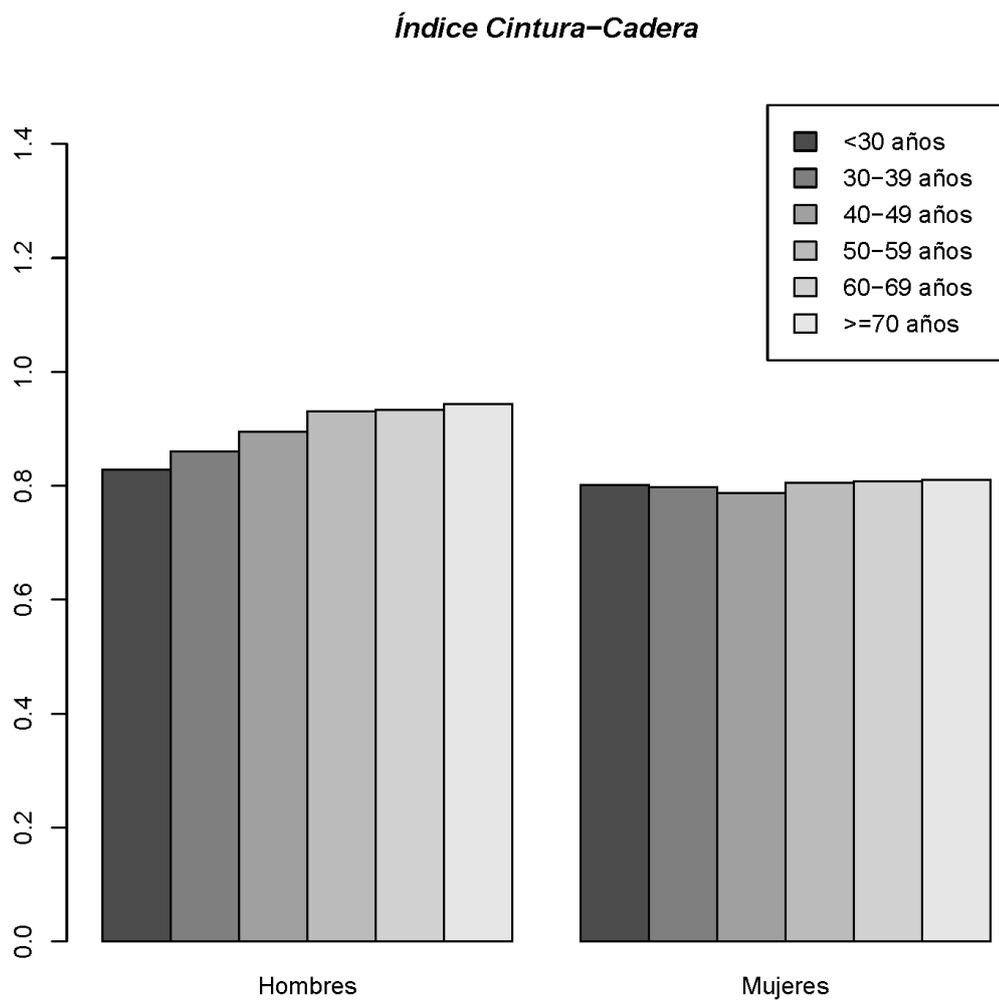
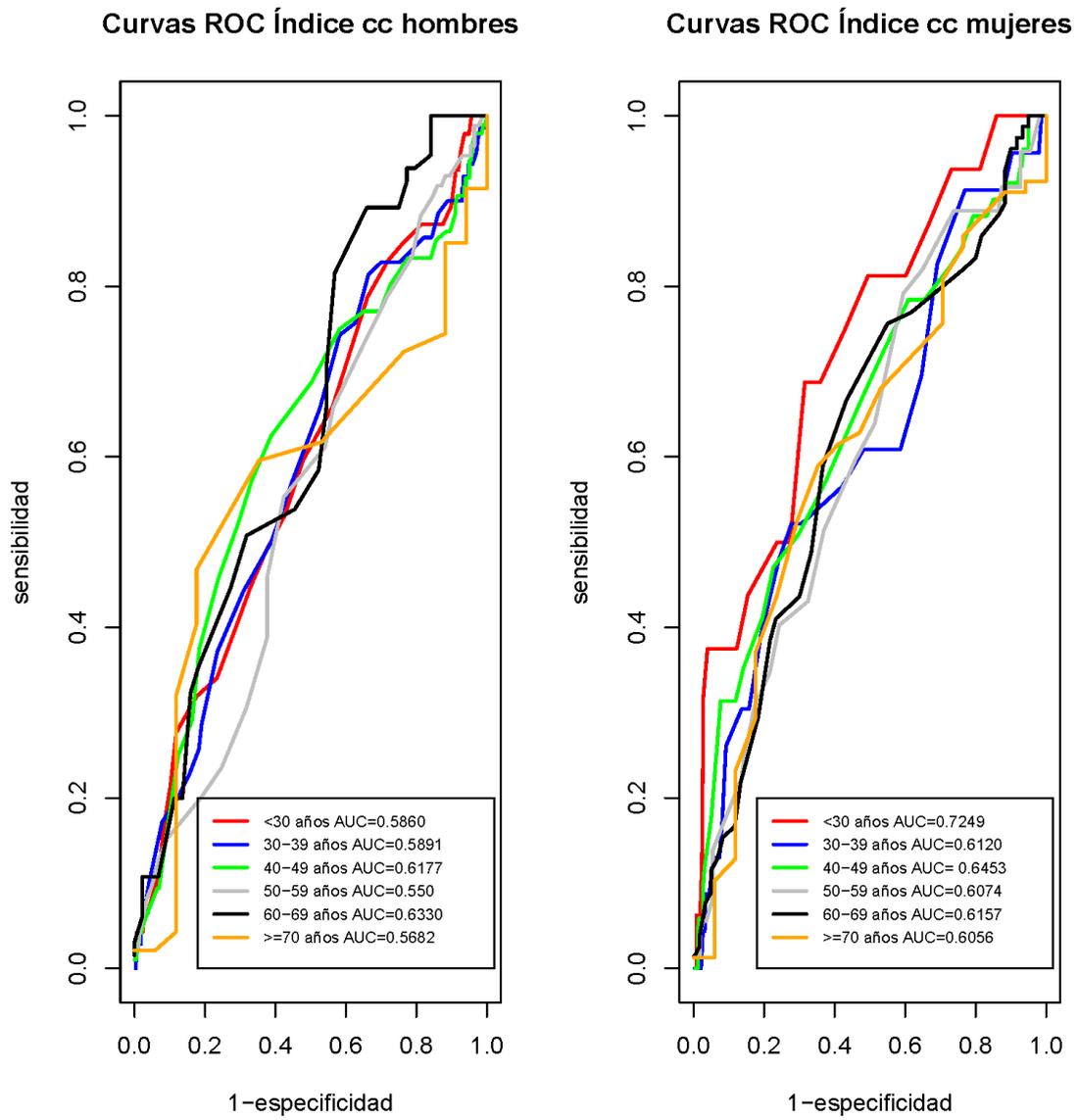


Figura 6.15: Valor Medio del Índice Cintura/Cadera

Figura 6.16: Curvas *ROC C/C* por Sexo y Grupos de Edad

Grupo Edad	<i>AUC</i>	Varianza	Intervalo de Confianza (95 %)
H < 30 años	0.5859719	0.002154336	0.4950005 - 0.6769433
H 30-39 años	0.5891258	0.001536198	0.5123063 - 0.6659453
H 40-49 años	0.6176928	0.001325813	0.5463271 - 0.6890584
H 50-59 años	0.5499654	0.001966001	0.4630614 - 0.6368694
H 60-69 años	0.6330420	0.003058670	0.5246457 - 0.7414383
H ≥ 70 años	0.5682103	0.006110271	0.4150034 - 0.7214172

Tabla 6.22: *AUC* para *C/C* e IC (95 %) en Hombres por Grupo de Edad

de este marcador. Sería necesario confirmar si esta suposición tiene sentido desde el punto de vista médico, es decir, ver si existe una justificación para tal hecho.

En las mujeres, las *AUCs* del *C/C* variaron entre 0.725 (<30 años) y 0.606 (≥ 70 años) (ver Tabla 6.24), y hay diferencia estadísticamente significativa ( $p = 0,0496$ , en la Tabla 6.25), por lo que se detecta la importancia de la edad en la capacidad de discriminación del Índice de Cintura/Cadera en mujeres, siendo más adecuado y útil en las poblaciones jóvenes.

En la Figura 6.17 se representaron las curvas *ROC* en hombres y en mujeres de cada uno de los grupos de edad considerados en el estudio. Solamente se detectaron diferencias significativas entre los hombres y las mujeres menores de 30 años, siendo bastante superior la efectividad del marcador en mujeres ( $p = 0,0222$ , ver Tabla 6.26). En los restantes grupos de edad parece que el comportamiento del Índice de Cintura/Cadera fue similar, en términos de discriminación, en hombres y en mujeres.

En la población masculina, los métodos *N* y *TN* proporcionaron *YIs* similares en todos los grupos de edad. En los hombres menores de 30 años y en los de edades comprendidas entre los 50 y los 59 años, el método *K* proporcionó valores similares del *YI* a los obtenidos mediante los procedimientos *N* y *TN*. En la población de 60 a 69 años el método *EMP* seleccionó un valor del *YI* similar al obtenido mediante el método *TN* y en las poblaciones de 30 a 39 años y más de 70 años el método *EMP* y el método *K* dieron Índices de Youden similares (ver Figura 6.19).

En realidad se puede decir que por un lado los métodos *N* y *TN*, en términos del *YI*, en general siguen un comportamiento similar y por otro lado los métodos *EMP* y *K* también

Curvas a Comparar	$\chi^2$	gl	p-valor
H <30 años H 30-39 años	0.0027	1	0.9586
H <30 años H 40-49 años	0.2726	1	0.6016
H <30 años H 50-59 años	0.3513	1	0.5534
H <30 años H 60-69 años	0.6003	1	0.4384
H <30 años H $\geq 70$ años	0.0855	1	0.7700
H 30-39 años H 40-49 años	0.2211	1	0.6382
H 30-39 años H 50-59 años	0.4155	1	0.5192
H 30-39 años H 60-69 años	0.5226	1	0.4697
H 30-39 años H $\geq 70$ años	0.1185	1	0.7306
H 40-49 años H 50-59 años	1.2429	1	0.2649
H 40-49 años H 60-69 años	0.0638	1	0.8005
H 40-49 años H $\geq 70$ años	0.6635	1	0.4153
H 50-59 años H 60-69 años	1.8701	1	0.1715
H 50-59 años H $\geq 70$ años	0.0902	1	0.7639
H 60-69 años H $\geq 70$ años	1.1389	1	0.2859

Tabla 6.23: Resultados Prueba Ji-Cuadrado  $C/C$  para cada Grupo de Edad - Hombres

Grupo Edad	<i>AUC</i>	Varianza	Intervalo de Confianza (95 %)
M < 30 años	0.7248535	0.004649333	0.5912114 - 0.8584957
M 30-39 años	0.6119964	0.004293826	0.4835652 - 0.7404275
M 40-49 años	0.6452934	0.002068049	0.5561624 - 0.7344243
M 50-59 años	0.6074199	0.001823927	0.5237148 - 0.6911250
M 60-69 años	0.6157051	0.002354565	0.5206001 - 0.7108101
M $\geq 70$ años	0.6055807	0.005626992	0.4585574 - 0.7526040

Tabla 6.24: *AUC* para  $C/C$  e IC (95 %) en Mujeres por Grupo de Edad

Curvas a Comparar	$\chi^2$	gl	p-valor
<b>M &lt;30 años M 30-39 años</b>	3.4512	1	<b>0.0632</b>
M <30 años M 40-49 años	1.7151	1	0.1903
<b>M &lt;30 años M 50-59 años</b>	3.7368	1	<b>0.0532</b>
<b>M &lt;30 años M 60-69 años</b>	3.2281	1	<b>0.0724</b>
<b>M &lt;30 años M <math>\geq 70</math> años</b>	3.8547	1	<b>0.0496</b>
M 30-39 años M 40-49 años	0.3004	1	0.5836
M 30-39 años M 50-59 años	0.0057	1	0.9400
M 30-39 años M 60-69 años	0.0037	1	0.9513
M 30-39 años M $\geq 70$ años	0.0112	1	0.9159
M 40-49 años M 50-59 años	0.3887	1	0.5330
M 40-49 años M 60-69 años	0.2372	1	0.6262
M 40-49 años M $\geq 70$ años	0.4273	1	0.5133
M 50-59 años M 60-69 años	0.0186	1	0.8915
M 50-59 años M $\geq 70$ años	0.0009	1	0.9758
M 60-69 años M $\geq 70$ años	0.0278	1	0.8676

Tabla 6.25: Resultados Prueba Ji-Cuadrado  $C/C$  para cada Grupo de Edad - Mujeres

Curvas a Comparar	$\chi^2$	gl	p-valor
<b>H &lt;30 años M &lt;30 años</b>	5.2264	1	<b>0.0222</b>
H 30-39 años M 30-39 años	0.1417	1	0.7066
H 40-49 años M 40-49 años	0.2064	1	0.6496
H 50-59 años M 50-59 años	0.8945	1	0.3443
H 60-69 años M 60-69 años	0.0814	1	0.7754
H $\geq 70$ años M $\geq 70$ años	0.3784	1	0.5385

Tabla 6.26: Resultados Prueba Ji-Cuadrado  $C/C$  por Grupos de Edad - Hombres y Mujeres

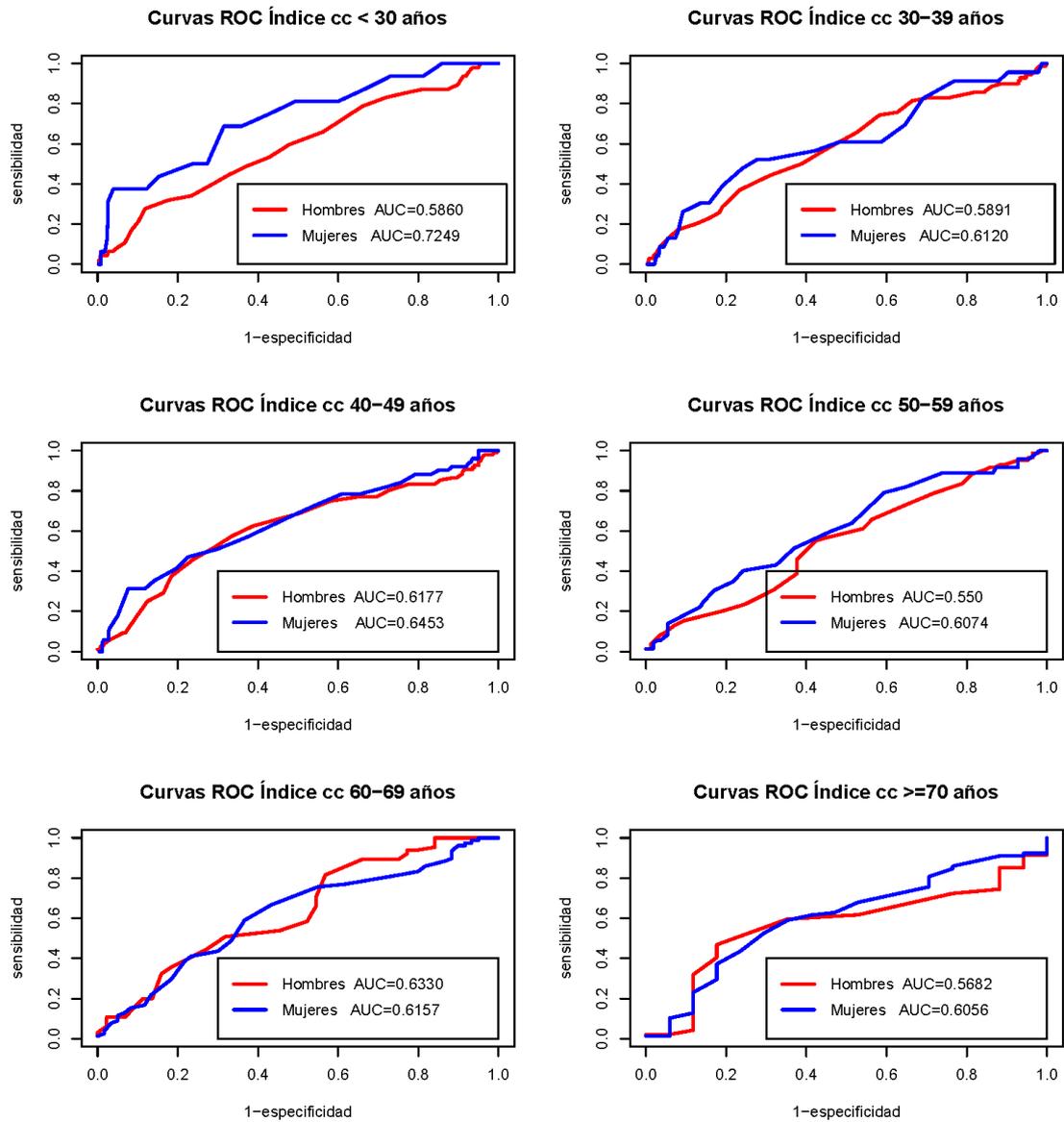


Figura 6.17: Curvas ROC C/C en Cada Grupo de Edad por Sexo

se comportan de forma análoga entre ellos.

En todos los grupos de edad el método *EMP* proporcionó los valores del *YI* más elevados, excepto en el grupo de 60 a 69 años, donde el método *TN* dió lugar a un *YI* superior. En las poblaciones de 30 a 59 años y 70 años hubo una clara diferencia entre la discriminación ofrecida por los métodos *K* y *EMP* por un lado y la proporcionada por los métodos *N* y *TN* por otro. Los primeros ofrecieron unas estimaciones más elevadas que los segundos.

En la población femenina, cabe destacar que los métodos *N* y *TN* proporcionaron los mismos valores del *YI* para todos los grupos de edad, lo que parece indicar que el marcador Índice de Cintura/Cadera se distribuye como una Normal en mujeres. Además se observó el mismo orden en los diferentes métodos de estimación para todos los grupos de edad, es decir, los valores más elevados se obtuvieron siempre mediante el método *EMP* y las estimaciones más bajas fueron las proporcionadas por los métodos *N* y *TN*.

En las poblaciones menores de 30 años y de 50 a 69 años el método *K* proporcionó valores más próximos a los obtenidos mediante los métodos *N* y *TN*, mientras que en la población de 40 a 49 años el método *K* seleccionó un óptimo más próximo al proporcionado por el método *EMP*. En realidad los cuatro métodos de estimación siguieron un patrón bastante similar a lo largo de todos los grupos de edad en mujeres.

En las Tablas 6.27 y 6.28<sup>3</sup> aparecen los valores de corte óptimos del *C/C* obtenidos mediante los diferentes métodos de estimación del *YI* para cada grupo de edad y sexo. Dichos valores variaron entre 0.799 y 1.039 en hombres y entre 0.758 y 0.910 en mujeres, dependiendo del procedimiento de estimación seleccionado y del grupo de edad del paciente.

En los varones, los métodos *N*, *TN* y *EMP* para edades mayores que 30 años detectaron puntos de corte mayores de 0.9, valor que se había establecido como límite del Índice de Cintura/Cadera (Tabla 6.2), a partir del cual parece que se detecta un aumento de riesgo cardiovascular.

En las mujeres, el método *TN*, en el grupo de edad 50 a 59 años detectó un punto de corte mayor que 0.85 y los métodos *K* y *EMP* lo hacen para la población con edades de 60 o mayores.

---

<sup>3</sup>grupo 1 : <30, grupo 2: 30-39, grupo 3: 40-49, grupo 4: 50-59, grupo 5: 60-69 y grupo 6: ≥70 años.

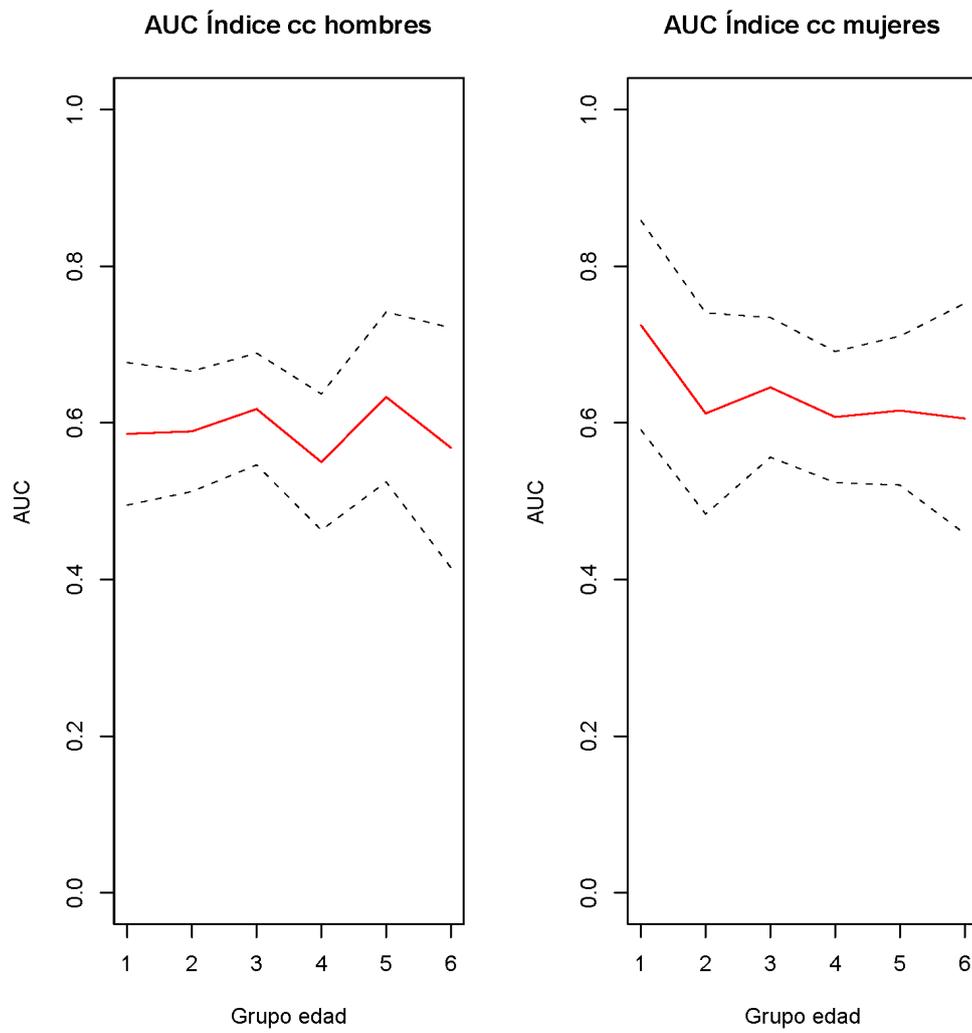


Figura 6.18: *AUC* e *IC C/C* para Cada Grupo de Edad - Hombres y Mujeres

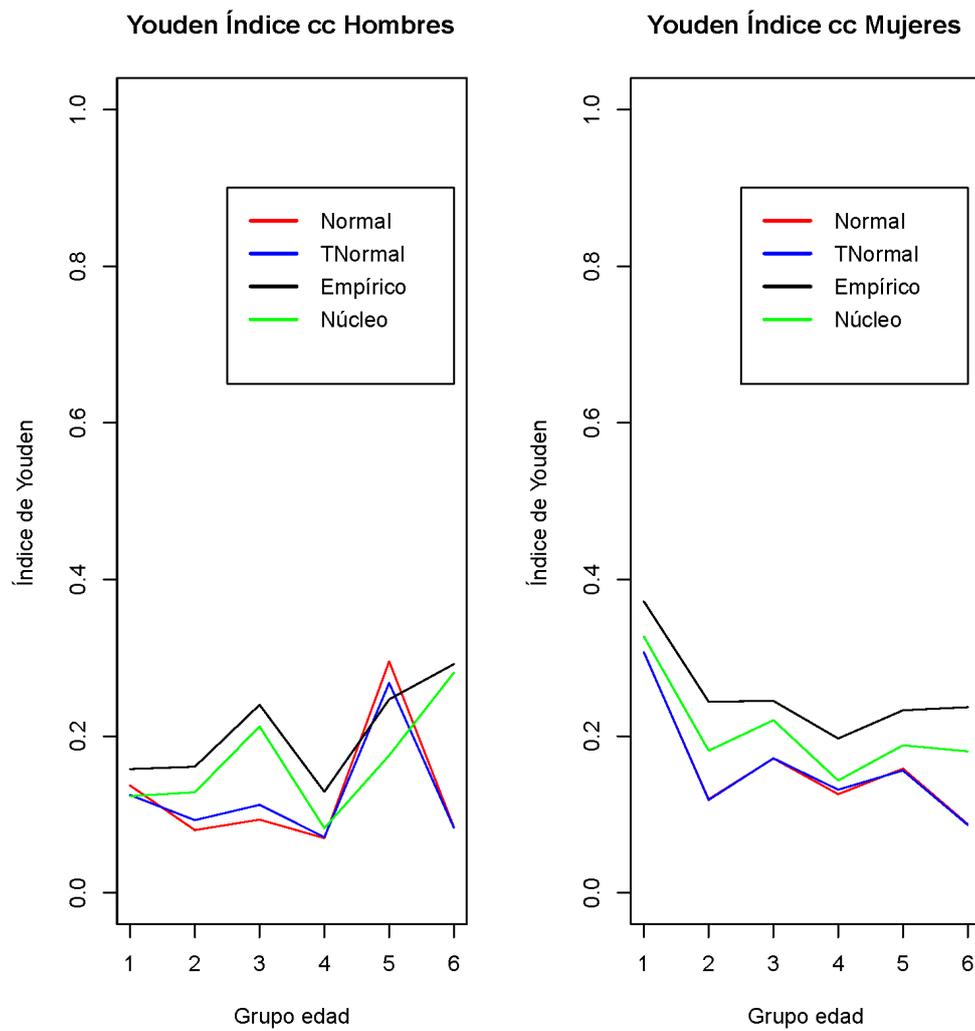


Figura 6.19:  $YI C/C$  para Cada Grupo de Edad Hombres y Mujeres-Métodos Examinados

Hombres							
Método		Grupos de Edad					
		1	2	3	4	5	6
<i>N</i>	$c^*$	0.799	0.932	0.963	0.897	0.874	0.835
	<i>YI</i>	0.137	0.080	0.093	0.070	0.295	0.083
<i>TN</i>	$c^*$	0.820	0.926	0.962	0.918	0.888	1.039
	<i>YI</i>	0.125	0.093	0.112	0.071	0.268	0.083
	Potencia	2.000	2.000	2.000	2.000	2.000	2.000
<i>EMP</i>	$c^*$	0.90	0.86	0.92	0.94	0.91	0.97
	<i>YI</i>	0.158	0.161	0.240	0.129	0.247	0.292
<i>K</i>	$c^*$	0.891	0.869	0.928	0.926	0.901	0.970
	<i>YI</i>	0.123	0.128	0.212	0.083	0.175	0.281
	$h_{enf}$	0.031	0.022	0.020	0.019	0.023	0.036
	$h_{sanos}$	0.019	0.018	0.019	0.022	0.026	0.008

Tabla 6.27:  $c^*$  e *YI* para *C/C* Hombres

Mujeres							
Método		Grupos de Edad					
		1	2	3	4	5	6
<i>N</i>	$c^*$	0.823	0.758	0.810	0.842	0.841	0.910
	<i>YI</i>	0.307	0.119	0.172	0.126	0.158	0.088
<i>TN</i>	$c^*$	0.811	0.759	0.815	0.863	0.847	0.899
	<i>YI</i>	0.307	0.119	0.172	0.132	0.156	0.087
	Potencia	0.700	0.900	1.316	1.800	1.241	2.000
<i>EMP</i>	$c^*$	0.78	0.83	0.84	0.80	0.85	0.88
	<i>YI</i>	0.372	0.244	0.245	0.197	0.233	0.237
<i>K</i>	$c^*$	0.826	0.843	0.851	0.821	0.857	0.888
	<i>YI</i>	0.327	0.181	0.220	0.143	0.188	0.181
	$h_{enf}$	0.057	0.045	0.031	0.026	0.022	0.025
	$h_{sanos}$	0.018	0.020	0.018	0.024	0.024	0.030

Tabla 6.28:  $c^*$  e *YI* para *C/C* Mujeres

Hombres							
Método	Grupos de Edad						
	1	2	3	4	5	6	
<i>N</i>	<i>Se</i>	0.830	0.257	0.292	0.812	0.923	0.894
	<i>Sp</i>	0.281	0.817	0.836	0.247	0.227	0.059
<i>TN</i>	<i>Se</i>	0.660	0.286	0.292	0.729	0.923	0.213
	<i>Sp</i>	0.383	0.810	0.836	0.353	0.227	0.882
<i>EMP</i>	<i>Se</i>	0.824	0.373	0.612	0.459	0.341	0.765
	<i>Sp</i>	0.158	0.161	0.240	0.129	0.247	0.292
<i>K</i>	<i>Se</i>	0.319	0.743	0.573	0.659	0.892	0.468
	<i>Sp</i>	0.824	0.418	0.667	0.435	0.341	0.824

Tabla 6.29: *Se* y *Sp* para *C/C* Hombres

En las Tablas 6.29 y 6.30 se presentan las medidas de Sensibilidad y Especificidad en los puntos de corte óptimos obtenidos para cada grupo de edad, en hombres y en mujeres. Estas medidas permiten visualizar los puntos de corte óptimos sobre la curva ROC (ver Figuras 6.20 y 6.21).

Mujeres							
Método	Grupos de Edad						
	1	2	3	4	5	6	
<i>N</i>	<i>Se</i>	0.438	0.826	0.627	0.556	0.718	0.372
	<i>Sp</i>	0.846	0.309	0.582	0.586	0.500	0.824
<i>TN</i>	<i>Se</i>	0.438	0.826	0.627	0.431	0.718	0.526
	<i>Sp</i>	0.805	0.309	0.582	0.676	0.500	0.706
<i>EMP</i>	<i>Se</i>	0.641	0.694	0.701	0.351	0.500	0.588
	<i>Sp</i>	0.372	0.244	0.245	0.197	0.233	0.237
<i>K</i>	<i>Se</i>	0.438	0.478	0.412	0.639	0.667	0.590
	<i>Sp</i>	0.846	0.759	0.805	0.486	0.567	0.647

Tabla 6.30: *Se* y *Sp* para *C/C* Mujeres

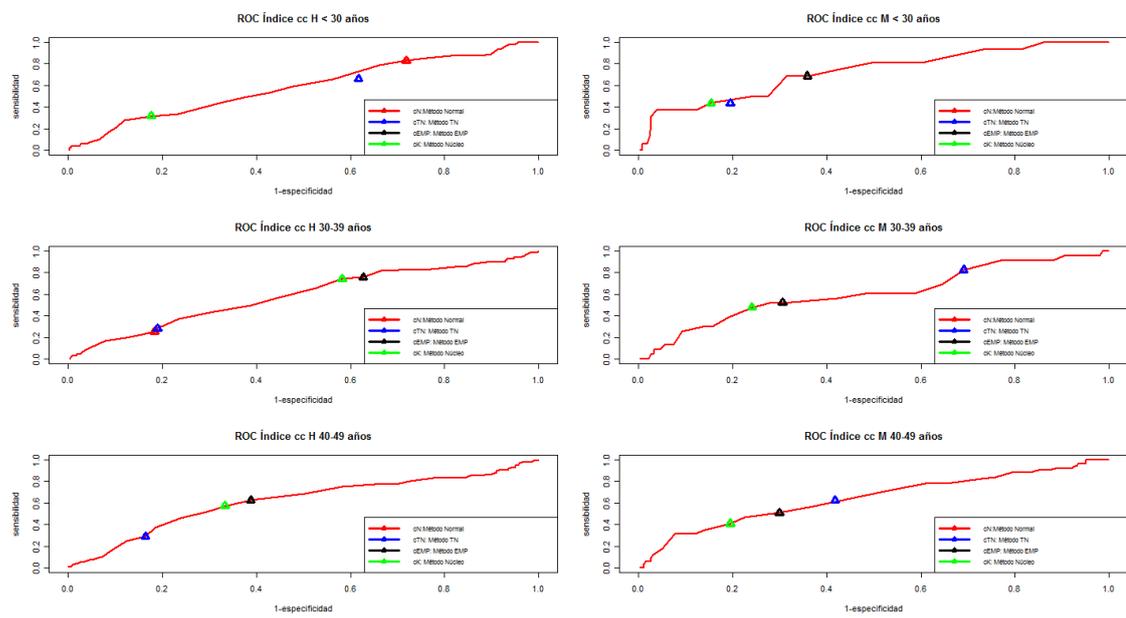


Figura 6.20:  $c^*$  sobre Curva ROC C/C-Hombres y Mujeres (1)

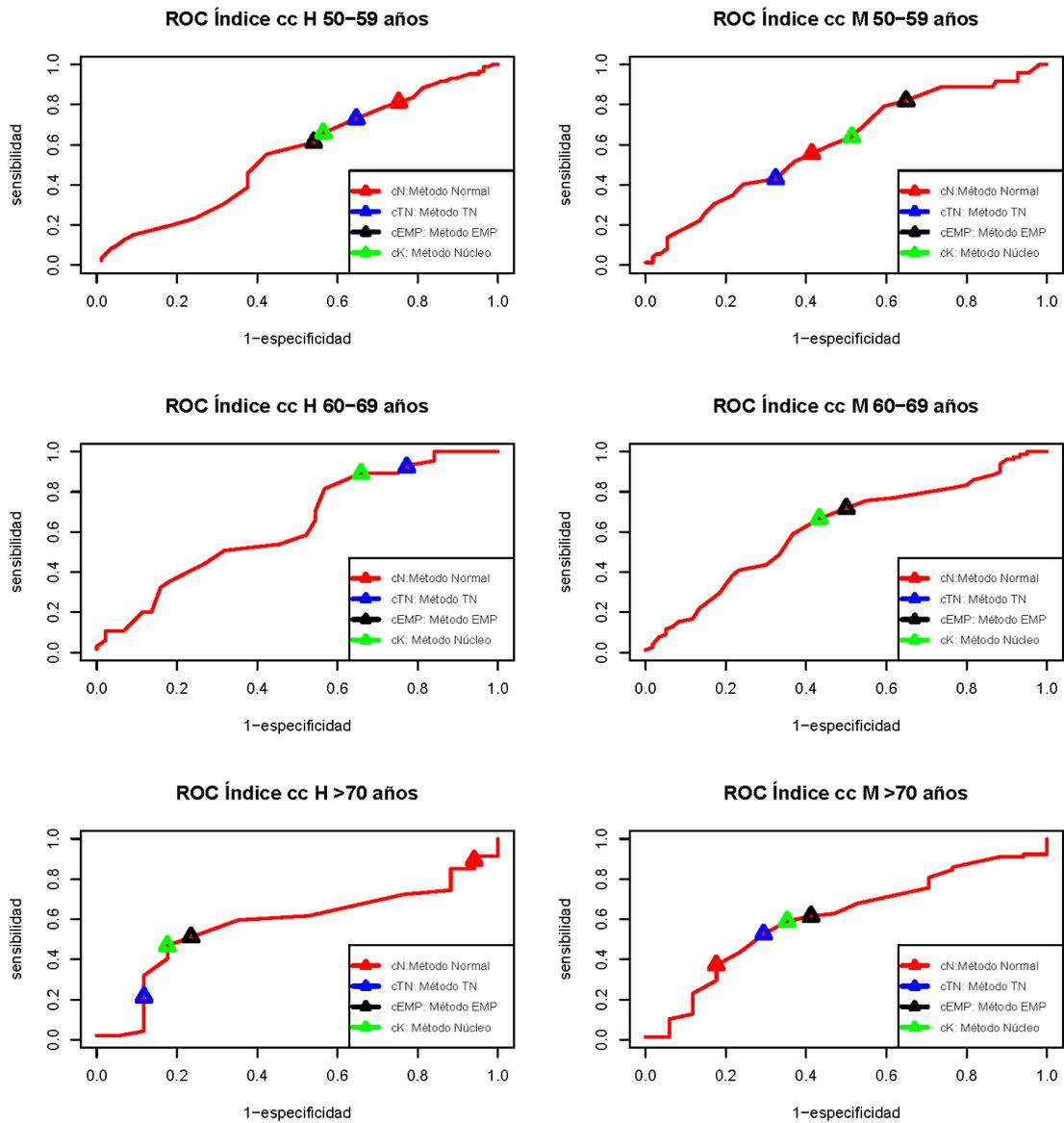


Figura 6.21:  $c^*$  sobre Curva  $ROC$  C/C-Hombres y Mujeres (2)

## Capítulo 7

# Conclusiones

De los resultados de los estudios de simulación realizados es evidente que:

1. Ninguno de los cuatro métodos examinados es superior en términos de sesgos y  $RMSE$  para todos los escenarios considerados.
2. Una buena estimación de  $YI$  no indica una buena estimación de  $c^*$ .
3. En la estimación del  $YI$  se observó que:
  - El método  $EMP$  tiene el peor comportamiento y no es recomendado, a menos que los tamaños muestrales sean grandes.
  - Para datos provenientes de distribuciones normales con varianzas iguales, aunque  $K$  fue el método que mostró los sesgos mínimos, el método  $N$  tiene valores de sesgo muy semejantes y se observó que cuando las varianzas son diferentes los métodos  $K$  y  $TN$  tienen sesgos menores que  $N$ .
  - Para datos provenientes de distribuciones asimétricas el método  $N$  presenta sesgos y  $RMSE$  mayores que los de  $TN$  y  $K$ . Se observó que para  $YI \leq 0,6$   $K$  sería una buena alternativa, mientras que cuando el nivel de separación es alto ( $YI = 0,8$ ) es  $TN$  el que tiene mejor comportamiento. Cuando los tamaños muestrales son grandes ( $n \geq 50$ ) y el nivel de separación es moderado ( $YI \leq 0,4$ ) el comportamiento de  $TN$  se asemeja a  $K$ .
  - En situaciones de bimodalidad  $K$  es generalmente el mejor método.
4. En la estimación del  $c^*$  se observó que:
  - Los métodos  $N$  y  $EMP$  muestran, de manera general, una pauta peor que los procedimientos  $TN$  y  $K$ .

- Para datos de distribuciones normales se observó que el método  $TN$  es, generalmente, el que tiene mejor comportamiento.
- Para datos de distribuciones bimodales los resultados son ambivalentes.

Teniendo en cuenta los resultados obtenidos en la aplicación a datos reales se pueden extraer las siguientes conclusiones en cuanto a la comparación de la capacidad de discriminación de las medidas antropométricas  $IMC$ ,  $CC$  y  $C/C$ :

- Para los hombres con edades comprendidas entre 18 y 59 años el  $IMC$  fue la medida antropométrica que registró valores del  $AUC$  ligeramente superiores a los de  $CC$ , pero dichas diferencias no fueron significativas.
- Sin embargo en la población masculina con edades a partir de los 60 años el  $C/C$  presentó un  $AUC$  ligeramente más elevado que los de las otras dos medidas, aunque tampoco se detectaron diferencias estadísticamente significativas. Entonces, sin tener en cuenta otros factores que puedan favorecer la utilización de una de las medidas, en la práctica médica inicialmente se obtendría una discriminación similar con cualquiera de ellas.
- En las mujeres con edades comprendidas entre los 18 y los 59 años el  $IMC$  tuvo valores de  $AUC$  algo más elevados, pero no estadísticamente significativos, en relación a la  $CC$ .
- En la población femenina de 60 a 69 años el  $IMC$  y el  $CC$  fueron los que presentaron los  $AUC$ s más elevados y en la población con edades superiores o iguales a 70 años fueron la  $CC$  y el  $C/C$ .
- Dado que no se detectaron diferencias significativas en ninguno de los casos no se puede afirmar que uno de los marcadores tenga mayor capacidad de discriminación en relación a los demás, así los resultados de este estudio apoyarían la controversia que existe en el ámbito médico acerca de la selección de la mejor medida antropométrica (Ashwell y otros, 1996) como marcador diagnóstico de HTA. De hecho, algunos autores (Pereira y otros, 1999; Silva y otros, 2006; Gomes y otros, 2006) recomiendan que se utilice la combinación de los tres marcadores para potencializar la capacidad predictiva de complicaciones cardiovasculares.
- Como conclusión general importante se confirma que el poder de discriminación de los marcadores diagnósticos de HTA estudiados se ve influenciado por las covariables sexo y edad del paciente y no tenerlo en cuenta podría llevar a conclusiones erróneas.

# Bibliografía

- Aoki, K.; Misumi, J.; Kimura, T.; Zhao, W. y Xie, T. (1997). «Evaluation of cutoff levels for screening of gastric cancer using serum pepsinogens and distributions of levels of serum pepsinogen I, II and of PG I/PG II ratios in a gastric cancer case-control study.» *Journal of Epidemiology*, **7(3)**, pp. 143–51.
- Ascaso, J.F.; González-Santos, P.; Hernández Mijaresa, A.; Mangas Rojas, A.; Masana Marín, L.; Millán Núñez-Cortés, J.; Pallaroda, L.F.; Pedro-Botet, J.; Pérez-Jiménez, F.; Pintó, X.; Plaza, I.; Rubiés-Prat, J. y Zúñiga, M. (2006). «Diagnóstico de Síndrome Metabólico. Adecuación de los Criterios Diagnósticos en nuestro Medio. Recomendaciones del foro HDL. Resumen Ejecutivo». *Clin Esp*, **206(11)**, pp. 576–82.
- Ashwell, M.; Lejeune, S. y McPherson, K. (1996). «Ratio of waist circumference to height may be better indicator of need for weight management». *BMJ*, **312**, p. 377.
- Bamber, D. (1975). «The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph». *Journal of Mathematical Psychology*, **12(4)**, pp. 387–415.
- Bao, Y.; Lu, J.; Wang, C.; Yang, M.; Li, H.; Zhang, X.; Zhu, J.; Lu, H.; Jia, W. y Xiang, K. (2008). «Optimal waist circumference cutoffs for abdominal obesity in Chinese». *Atherosclerosis*, **201**, p. 378–384.
- Botana, M.; Mato, J.; Cadarso-Suárez, C.; Tomé, M. A.; Pérez-Fernández, R.; Fernández-Mariño, A.; Rego-Iraeta, A. y Solache, I. (2007). «Overweight, obesity and central obesity prevalences in the region of Galicia in Northwest Spain». *Obesity and Metabolism*, **3(3)**, pp. 106–115.
- Brasil, Pedro (2010). *DiagnosisMed: Diagnostic test accuracy evaluation for medical pro-*

- professionals*. R package version 0.2.3.  
<http://CRAN.R-project.org/package=DiagnosisMed>
- DeLong, E. R.; DeLong, D. y Clarke-Pearson, D. L. (1988). «Comparing the area under two or more correlated receiver operating characteristic curves: a nonparametric approach». *Biometrics*, pp. 837–845.
- Faraggi, D. y Reiser, B. (2002). «Estimation of the Area under the ROC Curve». *Statistics in Medicine*, **21**, pp. 3093 – 3106.
- Fawcett, T. (2006). «An introduction to ROC analysis». *Pattern Recognition Letters*, **27(8)**, pp. 861–874.
- Fluss, R.; Faraggi, D. y Reiser, B. (2005). «Estimation of the Youden Index and its Associated Cutoff Point». *Biometrical Journal*, **47(4)**, pp. 458–472.
- Gomes, M. A.; Rech, C. R.; Gomes, M. B. A. y Santos, D. L. (2006). «Correlação entre índices antropométricos e distribuição de gordura corporal em mulheres idosas.» *Rev bras cineantropom desempenho hum.*, **8(3)**, pp. 16–22.
- Hanley, J. y McNeil, B. (1982). «The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve». *Diagnostic Radiology*, **143(1)**, pp. 29 – 36.
- Human Health Services Department, U.S. (2003). «The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure». *Informe técnico*, National Institutes of Health and National Heart, Lung, and Blood Institute.
- Krzanowski, W. y Hand, D. (2009). *ROC Curves for Continuous Data*. Taylor and Francis Group.
- Matoba, Y.; Inoguchi, T.; Nasu, S.; Suzuki, S.; Yanase, T.; Nawata, H. y Takayanagi, R. (2008). «Optimal Cut Points of Waist Circumference for the Clinical Diagnosis of Metabolic Syndrome in the Japanese Population». *Diabetes Care*, **31(3)**, pp. 590–592.
- Molodianovitch, K.; Faraggi, D. y Reiser, B. (2006). «Comparing the Areas Under Two Correlated ROC Curves: Parametric and Non-Parametric Approaches». *Biometrical Journal*, **48(5)**, pp. 745 – 757.
- Obuchowski, N. (1995). «Multireader, multimodality receiver operating characteristic curve studies: Hypothesis testing and sample size estimation using an analysis of variance approach with dependent observations». *Academic Radiology*, **2**, pp. 522 – 529..

- Pepe, Margaret Sullivan (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press Inc, New York.
- Pereira, R. A.; Sichieri, R. y Marins, V.M.R. (1999). «Razão cintura/quadril como preditor de hipertensão arterial». *Cad Saúde Pública*, **15(2)**, pp. 333–44.
- Perez-Fernandez, R.; Mariño, A. F.; Cadarso-Suarez, C.; Botana, M. A.; Tome, M. A.; Solache, I.; Rego-Iraeta, A. y Mato, A. J. (2007). «Prevalence, awareness, treatment and control of hypertension in Galicia (Spain) and association with related diseases». *Journal of Human Hypertension*, **21**, pp. 366–373.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.  
<http://www.R-project.org>
- Schisterman, E. F.; Faraggi, D. y Reiser, B. (2004). «Adjusting the Generalized ROC Curve for Covariates». *Stat Med*, **23**, p. Stat Med.
- Schisterman, E. F.; Perkins, N. J.; Liu, A. y Bondell, H. (2005). «Optimal Cut-point and Its Corresponding Youden Index to Discriminate Individuals Using Pooled Blood Samples». *Epidemiology*, **16(1)**, pp. 73–81.
- SEEDO (2000). «Consenso SEEDO'2000 para la evaluación del sobrepeso y la obesidad y el establecimiento de criterios de intervención terapéutica». En: *Conferencia de Consenso*, .
- Silva, J.L.T.; Barbosa, D. S.; Oliveira, J. A. y Guedes, D. P. (2006). «Distribuição centrípeta da gordura corporal, sobrepeso e aptidão cardiorrespiratória: associação com sensibilidade insulínica e alterações metabólicas». *Arq Bras Endocrinol Metab*, **50(6)**, pp. 1034–40..
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC.
- Sing, Tobias; Sander, Oliver; Beerenwinkel, Niko y Lengauer, Thomas (2009). *ROCR: Visualizing the performance of scoring classifiers..* R package version 1.0-4.  
<http://CRAN.R-project.org/package=ROCR>
- Thompson, M. L. y Zucchini, W. (1989). «On the statistical analysis of ROC curves.» *Statistics in Medicine*, **8(10)**, pp. 1277–90.

- Vilar, J.M. (1991). «Estimación no Paramétrica de la Función de Distribución». *Qüestió*, **15(1)**, pp. 3–20.
- Youden, W. J. (1950). «Index for Rating Diagnostic Tests». *Cancer*, **3(1)**, pp. 32–5.
- Zou, K. H. y Hall, W. J. (2000). «Two transformation models for estimating an ROC curve derived from continuous data». *Journal of Applied Statistics*, **27(5)**, pp. 621 – 631.
- Zweig, M.H. y Campbell, G. (1993). «Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine». *Clinical Chemistry*, **39(8)**, pp. 561–577.