

UNIVERSIDAD DE PUERTO RICO
RECINTO DE RÍO PIEDRAS
FACULTAD DE CIENCIAS NATURALES
DEPARTAMENTO DE MATEMÁTICAS

TESIS DE MAESTRÍA EN MATEMÁTICA APLICADA

Análisis Bayesiano Comprensivo de Múltiples Pruebas Clínicas

Autor:

Jovadell CRUZ HERNÁNDEZ

Supervisor:

Dr. Luis Raúl PERICCHI GUERRA

15 de diciembre de 2014

APROBADO POR EL COMITÉ ASESOR EN CUMPLIMIENTO PARCIAL DE
LOS REQUISITOS PARA EL GRADO DE MAESTRÍA EN CIENCIA EN
MATEMÁTICAS APLICADAS EN LA UNIVERSIDAD DE PUERTO RICO

ASESOR: _____

Luis R. Pericchi Guerra, Ph.D.

LECTORES: _____

María E. Pérez Hernández, Ph.D.

Pedro J. Rodríguez Esquerdo, Ph.D., J.D.

UNIVERSIDAD DE PUERTO RICO

Fecha: **15 de diciembre de 2014**

Autor: **Jovadell Cruz Hernández**

Título: **Análisis Bayesiano Comprensivo de
Múltiples Pruebas Clínicas**

Departamento: **Matemáticas**

Grado: **M.S.** Convocatoria: **Diciembre** Año: **2014**

Se concede permiso adjunto a la Universidad de Puerto Rico a circular y hacer copia para fines no comerciales, a su discreción, del título anterior, a solicitud de personas o instituciones.

Firma del Autor

EL AUTOR SE RESERVA OTROS DERECHOS DE PUBLICACIÓN, Y NI LA TESIS NI EXTENSOS EXTRACTOS DE ELLA SE PUEDEN IMPRIMIR O DE OTRA MANERA REPRODUCIR SIN EL PERMISO ESCRITO DEL AUTOR.

EL AUTOR CERTIFICA QUE HA OBTENIDO EL PERMISO PARA EL USO DE CUALQUIER MATERIAL CON DERECHOS DE AUTOR QUE APARECEN EN ESTA TESIS Y QUE TODO USO ESTÁ CLARAMENTE RECONOCIDO.

A mi madre.

Tabla de Contenido

Tabla de Contenido	ii
Resumen	iii
Agradecimientos	iv
Introducción	v
1 Datos	1
2 Notación	2
3 Función de Verosimilitud	7
4 Factor de Bayes	8
5 Método Objetivo Bayesiano	10
6 Conclusión	17
7 Referencias	18
Apéndice	19

Resumen

Las pruebas diagnósticas que son utilizadas en la práctica clínica ayudan a determinar el estado de una enfermedad en el paciente. Conocer si estas pruebas son condicionalmente independientes o no, ayudará a determinar dicho estado con mayor precisión. Se encontró que las probabilidades de tener la enfermedad dado que ambas pruebas son positivas, negativas o combinadas se sobrestiman/subestiman con el modelo tradicional, que supone independencia condicional, y que existe un conflicto cuando las pruebas tienen resultados opuestos. Este conflicto puede generar evidencias dramáticamente discrepantes si se supone independencia condicional o no. En este trabajo construimos una prueba bayesiana objetiva basada en el factor de Bayes de los modelos con y sin independencia condicional. Además, se calculan las densidades posteriores de las probabilidades de diagnóstico e intervalos a posteriori por métodos de Monte Carlo y de Cadenas de Markov (MCMC). Este trabajo completa el trabajo de Pereira y Pericchi (1990) y generaliza el trabajo de Berger & Mossman (2001).

Agradecimientos

Quiero agradecer y dar reconocimiento a todos aquellos que me ayudaron durante el desarrollo de esta tesis. Principalmente al Dr. Luis Raúl Pericchi Guerra por su gran cooperación y paciencia en estos años como profesor y consejero. Su conocimiento me ayudó a ser mejor persona e investigador. También quiero agradecer a Raúl O. Sánchez por su apoyo y motivación en el momento de solicitar la admisión al programa de maestría. Por último, quiero agradecer a mi familia y amigos por su gran apoyo. Gracias a todos por estar ahí cuando más los necesite.

San Juan, Puerto Rico

15 de diciembre de 2014

Jovadell Cruz Hernández

Introducción

¿Estarás enfermo si tu resultado fue positivo en una prueba? Si no te convence, ¿una segunda prueba lo determinará? Las pruebas de diagnóstico clínico son la fuente de información más importante para ayudar a determinar si un paciente tiene o no la enfermedad. Aunque a veces no estamos dispuestos a aceptar el resultado de una prueba; consideramos hacernos una segunda prueba, igual o diferente a la primera, para reafirmar o invalidar el resultado de la primera prueba. Pero además de obtener los resultados de estas pruebas, lo más importante es cómo utilizamos sus resultados para dar el diagnóstico final.

Debemos conocer si las pruebas parten de la premisa de independencia condicional o no. Esta consideración es un factor importante a la hora de determinar la probabilidad de tener la enfermedad. Es decir, pruebas que están bajo la suposición de independencia condicional nos darán más información sobre el paciente que si no las consideráramos condicionalmente independientes. Por consiguiente, al suponer independencia condicional se estaría subestimando o sobrestimando el resultado, lo cual afectaría la decisión final del médico y a su vez al paciente.

Los datos que analizaremos fueron sacados del trabajo de Pereira & Pericchi (1990) en el Hospital das Clínicas, São Paulo. Usaremos la función de verosimilitud que no supone independencia condicional presentada en este trabajo y construiremos una función de verosimilitud que supone independencia condicional para determinar bajo cuál suposición es mejor considerar las pruebas.

Por otro lado, Mossman & Berger (2001) presentan cinco métodos para obtener intervalos de credibilidad de la probabilidad de tener la enfermedad dado el resultado de una prueba. Entre los considerados allí, el método Objetivo Bayesiano resultó ser el mejor de los métodos presentados. Este nos facilitará observar las diferencias de

suponer las pruebas condicionalmente independientes o no. Los autores implementaron un algoritmo para una sola prueba y en este trabajo lo ajustamos para incluir una segunda prueba. Hyndman (1996) implementó un código para calcular intervalos de alta densidad posterior que lo estaremos utilizando para comparar el resultado obtenido del método Objetivo Bayesiano. Ambos métodos se alejan del método frecuentista “Plug-In”, el cual se utilizan las probabilidades como valores fijos.

1 Datos

Los datos se obtuvieron del trabajo de Pereira & Pericchi (1990). El conjunto de datos se compone de observaciones en 100 niños con obstrucción biliar, clasificados como; intrahepático (D) y extrahepático (D^c). Se hicieron dos pruebas clínicas, cada una con resultados positivo o negativo, para diferenciar entre ambos estados. Nuestro estudio utiliza probabilidades condicionales, para lo cual necesitamos datos de pacientes que incluyan los resultados de ambas pruebas. Por esta razón, no podemos considerar los siete niños que no participaron en la segunda prueba. En la Tabla 1 se muestran los resultados de los 93 niños restantes.

Estado	E^{++}	E^{+-}	E^{-+}	E^{--}	Tamaño de la muestra
D	28	12	1	4	45
D^c	2	9	2	35	48

Tabla 1: Frecuencia de niños con obstrucción biliar según los resultados de las pruebas

E^{++} , E^{+-} , E^{-+} y E^{--} es el número total de personas que, según el superíndice, respondieron positivo o negativo a cada prueba.

2 Notación

La probabilidad de tener la enfermedad D, conocida como la prevalencia $P(D)$, se denota por d . Se supone que las pruebas de diagnóstico solo tienen dos posibles resultados, positivo y negativo. La probabilidad de obtener un resultado positivo en la primera prueba dado que el paciente tiene la enfermedad D, se conoce como la sensibilidad, $P(T_1^+|D)$, de la prueba y es representada por p . Y la probabilidad de obtener un resultado negativo en la primera prueba dado que el paciente no tiene la enfermedad D, que se conoce como la especificidad, $P(T_1^-|D^c)$, y es representada por q .

La probabilidad de obtener un resultado positivo en la segunda prueba dado que el paciente tiene la enfermedad D y el resultado de la primera prueba es positivo, $P(T_2^+|D, T_1^+)$, es representada por p_+ . La probabilidad de obtener un resultado positivo en la segunda prueba dado que el paciente tiene la enfermedad D y el resultado de la primera prueba es negativo, $P(T_2^+|D, T_1^-)$, y se representada por p_- . La probabilidad de obtener un resultado negativo en la segunda prueba dado que el paciente no tiene la enfermedad D y el resultado de la primera prueba es positivo, $P(T_2^-|D^c, T_1^+)$, se representa con q_+ . La probabilidad de obtener un resultado negativo en la segunda prueba dado que el paciente no tiene la enfermedad D y el resultado de la primera prueba es negativo, $P(T_2^-|D^c, T_1^-)$, se representa con q_- .

Sabemos que dos eventos A y B son condicionalmente independientes dado C si y solo si $P(B|D, A) = P(B|D)$. Por lo que, suponiendo independencia condicional entre ambas pruebas dado D, definimos a p como la probabilidad de un resultado positivo en la segunda prueba dado que el paciente está enfermo $P(T_2^+|D)$. Y definimos a q como la probabilidad de un resultado negativo en la segunda prueba dado que el paciente no está enfermo, $P(T_2^-|D^c)$.

Definimos a ϕ_+^j , para $j = 1, 2$, como la probabilidad de tener la enfermedad D dado el resultado positivo de la primera o segunda prueba $P(D|T_j^+)$. Y a ϕ_-^j como la probabilidad de tener la enfermedad D dado el resultado negativo de la primera o segunda prueba $P(D|T_j^-)$. Y finalmente, $\phi_{++} = P(D|T_1^+, T_2^+)$, como la probabilidad de tener la enfermedad D dado que los resultados de ambas pruebas son positivas. De forma análoga, definimos para ϕ_{+-} , ϕ_{-+} y ϕ_{--} . Utilizaremos i y ni para identificar cuando las pruebas son condicionalmente independientes o no, respectivamente.

Parámetro	Probabilidad
d	$P(D)$
p	$P(T_1^+ D)$
p_+	$P(T_2^+ D, T_1^+)$
p_-	$P(T_2^+ D, T_1^-)$
p.	$P(T_2^+ D)$
q	$P(T_1^- D^c)$
q_+	$P(T_2^- D^c, T_1^+)$
q_-	$P(T_2^- D^c, T_1^-)$
q.	$P(T_2^- D^c)$
ϕ_+^1	$P(D T_1^+)$
ϕ_-^1	$P(D T_1^-)$
ϕ_+^2	$P(D T_2^+)$
ϕ_-^2	$P(D T_2^-)$
ϕ_{++}	$P(D T_1^+, T_2^+)$
ϕ_{+-}	$P(D T_1^+, T_2^-)$
ϕ_{-+}	$P(D T_1^-, T_2^+)$
ϕ_{--}	$P(D T_1^-, T_2^-)$

Tabla 2: Notación para cada parámetro.

Se utiliza el Teoremas de Bayes para obtener las siguiente ecuaciones:

$$P(D|T_j^+) = \frac{P(D)P(T_j^+|D)}{P(D)P(T_j^+|D) + P(D^c)P(T_j^+|D^c)}$$

$$P(D|T_j^-) = \frac{P(D)P(T_j^-|D)}{P(D)P(T_j^-|D) + P(D^c)P(T_j^-|D^c)}$$

Bajo la suposición de que las pruebas son condicionalmente independientes:

$$P(D|T_1^+, T_2^+) = \frac{P(T_2^+|D)P(D|T_1^+)}{P(T_2^+|D)P(D|T_1^+) + P(T_2^+|D^c)P(D^c|T_1^+)}$$

$$P(D|T_1^-, T_2^-) = \frac{P(T_2^-|D)P(D|T_1^-)}{P(T_2^-|D)P(D|T_1^-) + P(T_2^-|D^c)P(D^c|T_1^-)}$$

$$P(D|T_1^-, T_2^+) = \frac{P(T_2^+|D)P(D|T_1^-)}{P(T_2^+|D)P(D|T_1^-) + P(T_2^+|D^c)P(D^c|T_1^-)}$$

$$P(D|T_1^+, T_2^-) = \frac{P(T_2^-|D)P(D|T_1^+)}{P(T_2^-|D)P(D|T_1^+) + P(T_2^-|D^c)P(D^c|T_1^+)}$$

Sin la suposición de que las pruebas son condicionalmente independientes:

$$P(D|T_1^+, T_2^+) = \frac{P(T_2^+|D, T_1^+)P(D|T_1^+)}{P(T_2^+|D, T_1^+)P(D|T_1^+) + P(T_2^+|D^c, T_1^+)P(D^c|T_1^+)}$$

$$P(D|T_1^-, T_2^-) = \frac{P(T_2^-|D, T_1^-)P(D|T_1^-)}{P(T_2^-|D, T_1^-)P(D|T_1^-) + P(T_2^-|D^c, T_1^-)P(D^c|T_1^-)}$$

$$P(D|T_1^-, T_2^+) = \frac{P(T_2^+|D, T_1^-)P(D|T_1^-)}{P(T_2^+|D, T_1^-)P(D|T_1^-) + P(T_2^+|D^c, T_1^-)P(D^c|T_1^-)}$$

$$P(D|T_1^+, T_2^-) = \frac{P(T_2^-|D, T_1^+)P(D|T_1^+)}{P(T_2^-|D, T_1^+)P(D|T_1^+) + P(T_2^-|D^c, T_1^+)P(D^c|T_1^+)}$$

Utilizando la notación presentada anteriormente podemos reescribir estas ecuaciones

de la siguiente forma:

$$\phi_+^1 = \frac{d \cdot p}{d \cdot p + (1 - d) \cdot (1 - q)} \quad (2.1)$$

$$\phi_-^1 = \frac{d \cdot (1 - p)}{d \cdot (1 - p) + (1 - d) \cdot q} \quad (2.2)$$

$$\phi_+^2 = \frac{d \cdot p.}{d \cdot p. + (1 - d) \cdot (1 - q.)} \quad (2.3)$$

$$\phi_-^2 = \frac{d \cdot (1 - p.)}{d \cdot (1 - p.) + (1 - d) \cdot q.} \quad (2.4)$$

Bajo la suposición de que las pruebas son condicionalmente independientes:

$$\phi_{++}^i = \frac{p. \cdot \phi_+}{p. \cdot \phi_+ + (1 - q.) \cdot (1 - \phi_+)} \quad (2.5)$$

$$\phi_{--}^i = \frac{(1 - p.) \cdot \phi_-}{(1 - p.) \cdot \phi_- + q. \cdot (1 - \phi_-)} \quad (2.6)$$

$$\phi_{+-}^i = \frac{(1 - p.) \cdot \phi_+}{(1 - p.) \cdot \phi_+ + q. \cdot (1 - \phi_+)} \quad (2.7)$$

$$\phi_{-+}^i = \frac{p. \cdot \phi_-}{p. \cdot \phi_- + (1 - q.) \cdot (1 - \phi_-)} \quad (2.8)$$

Sin la suposición de que las pruebas son condicionalmente independientes:

$$\phi_{++}^{ni} = \frac{p_+ \cdot \phi_+}{p_+ \cdot \phi_+ + (1 - q_+) \cdot (1 - \phi_+)} \quad (2.9)$$

$$\phi_{--}^{ni} = \frac{(1 - p_-) \cdot \phi_-}{(1 - p_-) \cdot \phi_- + q_- \cdot (1 - \phi_-)} \quad (2.10)$$

$$\phi_{+-}^{ni} = \frac{(1 - p_+) \cdot \phi_+}{(1 - p_+) \cdot \phi_+ + q_+ \cdot (1 - \phi_+)} \quad (2.11)$$

$$\phi_{-+}^{ni} = \frac{p_- \cdot \phi_-}{p_- \cdot \phi_- + (1 - q_-) \cdot (1 - \phi_-)} \quad (2.12)$$

3 Función de Verosimilitud

Pereira & Pericchi (1990) presentan la función de verosimilitud para los datos:

$$f(x, y, x^{++}, x^{-+}, x^{+-}, x^{--}, y^{++}, y^{-+}, y^{+-}, y^{--}, m, n | p, q, p_+, p_-, q_+, q_-) =$$

$$c p^x (1-p)^{m-x} q^y (1-q)^{n-y} p_+^{x^{++}} (1-p_+)^{x^{+-}} q_+^{y^{++}} (1-q_+)^{y^{+-}}$$

$$\cdot p_-^{x^{-+}} (1-p_-)^{x^{--}} q_-^{y^{-+}} (1-q_-)^{y^{--}} \quad (3.1)$$

donde m es el número total de pacientes con D que respondieron la primera prueba y x es el número de pacientes con D que su resultado fue positivo en la primera prueba. Y n es el número total de pacientes con D^c que respondieron la primera prueba y y es el número de pacientes con D^c que su resultado fue negativo en la primera prueba. Cuando las dos pruebas son condicionalmente independientes dado D entonces $p_+ = P(T_2^+ | D, T_1^+) = P(T_2^+ | D) = p$. y $p_- = P(T_2^+ | D, T_1^-) = P(T_2^+ | D) = p$. Y $q_+ = P(T_2^- | D^c, T_1^+) = P(T_2^- | D^c) = q$. y $q_- = P(T_2^- | D^c, T_1^-) = P(T_2^- | D^c) = q$. La función de verosimilitud para este caso es:

$$f(x, y, x^{++}, x^{-+}, x^{+-}, x^{--}, y^{++}, y^{-+}, y^{+-}, y^{--}, m, n | p, q, p_+, q_+) =$$

$$c p^x (1-p)^{m-x} q^y (1-q)^{n-y} p_+^{x^{++}+x^{+-}} (1-p_+)^{x^{+-}+x^{--}} q_+^{y^{++}+y^{+-}} (1-q_+)^{y^{+-}+y^{--}} \quad (3.2)$$

donde $c = \binom{m}{x} \binom{n}{y} \binom{x^{++}+x^{+-}}{x^{++}} \binom{y^{++}+y^{+-}}{y^{++}} \binom{x^{+-}+x^{--}}{x^{+-}} \binom{y^{+-}+y^{--}}{y^{+-}}$. Esta constante proviene de distribuciones binomiales.

4 Factor de Bayes

El Factor de Bayes se utiliza para comparar modelos, decidiendo cuál de los modelos se ajusta mejor a los datos. Consideramos nuestro primer modelo, M_1 , donde ambas pruebas son condicionalmente independientes como la hipótesis nula; mientras el segundo modelo, M_2 , donde no hay suposición de independencia condicional, que después llamaremos el método bayesiano general, corresponde a la hipótesis alternativa.

Según planteado por Pereira & Pericchi (1990), se le asigna a p , q , p_+ y q_+ una distribución previa $Beta(1, 1)$ y para p_+ , p_- , q_+ y q_- una distribución $Beta$ con parámetros $\frac{1}{2}$ y $\frac{1}{2}$. Usando (3.2) y (3.1) como la función de verosimilitud para M_1 y M_2 , respectivamente, obtenemos el Factor de Bayes:

$$B_{01} = \frac{B(x^{++} + x^{-+} + 1, x^{+-} + x^{--} + 1) \cdot B(y^{+-} + y^{--} + 1, y^{++} + y^{-+} + 1) \cdot \pi^4}{B(x^{++} + \frac{1}{2}, x^{+-} + \frac{1}{2}) \cdot B(y^{+-} + \frac{1}{2}, y^{++} + \frac{1}{2}) \cdot B(x^{-+} + \frac{1}{2}, x^{--} + \frac{1}{2}) \cdot B(y^{--} + \frac{1}{2}, y^{-+} + \frac{1}{2})}. \quad (4.1)$$

Al aplicar este factor al conjunto de datos bajo estudio, obtuvimos un resultado de 0.6224, lo que implica que M_2 se ajusta mejor a los datos que M_1 . Según la escala de Jeffreys es de grado 1, evidencia débil en contra la hipótesis nula.

Teorema 1. *El orden de las pruebas no afecta el resultado del Factor de Bayes.*

Demostración. Utilizaremos la siguiente notación $x^+ = x^{++} + x^{-+}$, $x^- = x^{+-} + x^{--}$, $y^+ = y^{++} + y^{-+}$ y $y^- = y^{+-} + y^{--}$ en la ecuación (4.1):

$$B_{01} = \frac{\beta(x^+ + 1, x^- + 1)}{\beta(x^{++} + \frac{1}{2}, x^{+-} + \frac{1}{2})\beta(x^{-+} + \frac{1}{2}, x^{--} + \frac{1}{2})} \cdot \frac{\beta(y^- + 1, y^+ + 1)}{\beta(y^{+-} + \frac{1}{2}, y^{++} + \frac{1}{2})\beta(y^{--} + \frac{1}{2}, y^{-+} + \frac{1}{2})} \cdot \pi^4.$$

Sabemos que

$$\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (4.2)$$

entonces

$$B_{01} = \frac{\frac{\Gamma(x^{\cdot+}+1)\Gamma(x^{\cdot-}+1)}{\Gamma(x^{\cdot+}+x^{\cdot-}+2)}}{\frac{\Gamma(x^{++}+\frac{1}{2})\Gamma(x^{+-}+\frac{1}{2})}{\Gamma(x^{\cdot+}+1)} \frac{\Gamma(x^{-+}+\frac{1}{2})\Gamma(x^{--}+\frac{1}{2})}{\Gamma(x^{\cdot-}+1)}} \cdot \frac{\frac{\Gamma(y^{\cdot-}+1)\Gamma(y^{\cdot+}+1)}{\Gamma(y^{\cdot-}+y^{\cdot+}+2)}}{\frac{\Gamma(y^{+-}+\frac{1}{2})\Gamma(y^{++}+\frac{1}{2})}{\Gamma(y^{\cdot+}+1)} \frac{\Gamma(y^{-+}+\frac{1}{2})\Gamma(y^{--}+\frac{1}{2})}{\Gamma(y^{\cdot-}+1)}} \cdot \pi^4.$$

Al reorganizar los numeradores y denominadores se obtiene:

$$B_{01} = \frac{\frac{\Gamma(x^{\cdot+}+1)\Gamma(x^{\cdot-}+1)}{\Gamma(x^{\cdot+}+x^{\cdot-}+2)}}{\frac{\Gamma(x^{++}+\frac{1}{2})\Gamma(x^{+-}+\frac{1}{2})}{\Gamma(x^{\cdot+}+1)} \frac{\Gamma(x^{+-}+\frac{1}{2})\Gamma(x^{--}+\frac{1}{2})}{\Gamma(x^{\cdot-}+1)}} \cdot \frac{\frac{\Gamma(y^{\cdot+}+1)\Gamma(y^{\cdot-}+1)}{\Gamma(y^{\cdot+}+y^{\cdot-}+2)}}{\frac{\Gamma(y^{++}+\frac{1}{2})\Gamma(y^{-+}+\frac{1}{2})}{\Gamma(y^{\cdot+}+1)} \frac{\Gamma(y^{+-}+\frac{1}{2})\Gamma(y^{--}+\frac{1}{2})}{\Gamma(y^{\cdot-}+1)}} \cdot \pi^4.$$

Reagrupando y usando (4.2) se obtiene:

$$B_{01} = \frac{B(x^{\cdot+}+1, x^{\cdot-}+1)}{B(x^{++}+\frac{1}{2}, x^{+-}+\frac{1}{2})B(x^{+-}+\frac{1}{2}, x^{--}+\frac{1}{2})} \cdot \frac{B(y^{\cdot+}+1, y^{\cdot-}+1)}{B(y^{++}+\frac{1}{2}, y^{-+}+\frac{1}{2})B(y^{+-}+\frac{1}{2}, y^{--}+\frac{1}{2})} \cdot \pi^4. \quad (4.3)$$

Demostramos que (4.1) y (4.3) coinciden, por lo tanto, el orden de las pruebas no afecta el resultado del Factor de Bayes. □

5 Método Objetivo Bayesiano

El enfoque bayesiano asigna distribuciones previas a cada parámetro. En nuestro caso los parámetros son: d , p , q , p_+ , q_+ , p_- , q_- , p y q . A partir de esas distribuciones previas podemos determinar la distribución posterior e intervalo de credibilidad para ϕ , ϕ^i y ϕ^{ni} . La selección de estas distribuciones previas depende del conocimiento que se tiene de la situación general y de la procedencia de los datos, este análisis se denomina análisis bayesiano subjetivo. Sin embargo, los intervalos para ϕ , ϕ^i y ϕ^{ni} se derivan usando únicamente los datos, en cuyo caso el enfoque será el análisis bayesiano objetivo. Este último escoge distribuciones a priori no-informativas o por defecto, para los parámetros, distribuciones previas que dependen sólo en el modelo de datos en cuestión y no en creencias subjetivas.

Los intervalos para ϕ , ϕ^i y ϕ^{ni} que resultarán del método objetivo bayesiano son generalmente llamados "intervalos de credibilidad". Estos intervalos son análogos a intervalos de confianza frecuentistas a ser considerados por sus propiedades de cobertura frecuentistas. Esto quiere decir que para un nivel de significancia α , los intervalos de confianza, en repetidas ocasiones, contienen a ϕ , ϕ^i y ϕ^{ni} el $(1 - \alpha)$ de las veces.

Usamos una distribución beta para describir la distribución a priori de los parámetros, y las distribuciones previas sugeridas por Pereira y Pericchi (1990). Esta selección de previas resultará en una distribución posterior $Beta(x_i + 1, n_i - x_i + 1)$ para d , p , q , p y q y $Beta(x_i + \frac{1}{2}, n_i - x_i + \frac{1}{2})$ para p_+ , q_+ , p_- y q_- .

El siguiente proceso de Monte Carlo simple producirá un intervalo de credibilidad de colas iguales con un nivel de significancia α para ϕ_+ y ϕ_{++} . Utilizando el programa estadístico R:

1. se simulan valores para cada uno de los parámetros con sus respectivas distribu-

ciones posteriores.

2. se usa la ecuación (2.1)-(2.4) para obtener el valor de ϕ .
3. repite este proceso, para un número N grande de veces para generar valores de ϕ .
4. se utilizan las ecuaciones (2.5)-(2.12) para hallar ϕ^i y ϕ^{ni} .
5. se ordena los valores y se encuentra el entero más cercano a $N \cdot \alpha/2$ y $N \cdot (1 - \alpha/2)$ de ϕ , ϕ^i y ϕ^{ni} , estos valores son el límite inferior y superior, respectivamente, de los intervalos.

El orden que decidamos al escoger las pruebas no afecta el resultado, es decir, $P(D|T_1^+, T_2^+) = P(D|T_2^+, T_1^+)$. Además, el Teorema de Bayes es secuencial, esto quiere decir que la probabilidad posterior luego de observar un resultado se puede utilizar como la probabilidad a priori para calcular la probabilidad posterior futura, sin importar la suposición de independencia condicional. Es decir,

$$P(D|T_1^+, T_2^+) = \frac{P(D)P(T_1^+, T_2^+|D)}{P(D)P(T_1^+, T_2^+|D) + P(D^c)P(T_1^+, T_2^+|D^c)}.$$

De la definición de probabilidad condicional $P(B)P(A|B) = P(A \cap B)$ sabemos que

$$\begin{aligned} P(D)P(T_1^+, T_2^+|D) &= P(T_1^+ \cap T_2^+ \cap D) \\ &= P(T_2^+|D, T_1^+)P(D \cap T_1^+) \\ &= P(T_2^+|D, T_1^+)P(D|T_1)P(T_1^+). \end{aligned}$$

De igual forma, ya que

$$P(D^c)P(T_1^+, T_2^+|D^c) = P(T_2^+|D^c, T_1^+)P(D^c|T_1^+)P(T_1^+),$$

entonces

$$\begin{aligned}
 P(D|T_1^+, T_2^+) &= \frac{P(T_2^+|D, T_1^+)P(D|T_1^+)P(T_1^+)}{[P(T_2^+|D, T_1^+)P(D|T_1^+) + P(T_2^+|D^c, T_1^+)P(D^c|T_1^+)]P(T_1^+)} \\
 &= \frac{P(T_2^+|D, T_1^+)P(D|T_1^+)}{P(T_2^+|D, T_1^+)P(D|T_1^+) + P(T_2^+|D^c, T_1^+)P(D^c|T_1^+)}.
 \end{aligned}$$

Los intervalos de credibilidad de Alta Densidad Posterior (HPD) se obtienen de una región con un nivel de confianza de tal forma que la probabilidad de cada punto contenido en la región es al menos igual a la probabilidad de cada punto fuera de la región. Por lo tanto, este método contiene los valores que son más probables. Hyndman (1996) implanta un código para obtener estos intervalos de credibilidad.

Consideramos el método “Plug-In” que utiliza cada uno de los parámetros mencionados anteriormente como probabilidades fijas. Por ejemplo, para obtener el valor de d , la probabilidad de tener la enfermedad, simplemente dividimos n por $n + m$. El valor n representa el número de niños con el estado D y $n + m$ representa el número total de niños. De forma análoga se procede para p , q , p_+ , q_+ , p_- , q_- , p , y q .

Aplicando estos valores fijos a las ecuaciones (2.1)-(2.12) obtenemos los valores de ϕ , ϕ^{ni} y ϕ^i . Esto no toma en cuenta la variabilidad de las estimaciones que pueden tener estas probabilidades según menciona Berger & Mossman (2001).

En la Tabla 4 se muestran las probabilidades de tener la enfermedad dado que la primera o segunda prueba es positiva o negativa y que las dos pruebas son positivas, negativas o combinadas, cuando se considera la independencia condicional entre las pruebas o no. Estas se calcularon con el método de “Plug-In” (PI), Objetivo Bayesiano (MOB) y Alta Densidad Posterior (HDP). Además, se calculó la media y desviación estándar (DE) de los intervalos.

Parámetro	Plug-In
d	$\frac{45}{93}$
p	$\frac{40}{45}$
p_+	$\frac{28}{40}$
p_-	$\frac{1}{5}$
$p.$	$\frac{29}{45}$
q	$\frac{37}{48}$
q_+	$\frac{9}{11}$
q_-	$\frac{35}{37}$
$q.$	$\frac{44}{48}$

Tabla 3: Resultados del método “Plug-in” para cada uno de los parámetros.

Podemos observar que el PI está en el intervalo de credibilidad del MOB. Además, el PI de ϕ_{++}^i es mayor que ϕ_{++}^{ni} excepto en el caso de ϕ_{--} que se esperaría que fuese menor. Esto implica que considerar las pruebas condicionalmente independientes sobreestima la probabilidad de que la persona tenga la enfermedad. El caso de ϕ_{--} , el no suponer independencia condicional es aproximadamente dos veces más probable que suponer independencia condicional. El único caso que mostró una leve diferencia es ϕ_{+-} . Pero el caso que produce mayor conflicto es el de ϕ_{-+} . Este muestra un PI de 0.5110 que diría que el paciente esta enfermo contra un 0.3333 que diría lo contrario. Por lo tanto, el método bayesiano general recoge mejor el conflicto de información. Además, podemos notar que la media es aproximadamente igual al valor de PI. Esto nos dice que es buen estimador para los valores de PI.

Probabilidades	PIM	OMB	HPD	Media	Mediana
ϕ_+^1	0.7843	(0.6540 , 0.8742)	(0.6629 , 0.8811)	0.7737	0.0569
ϕ_-^1	0.1190	(0.0529 , 0.2481)	(0.0443 , 0.2347)	0.1362	0.0508
ϕ_+^2	0.8788	(0.7262 , 0.9506)	(0.7444 , 0.9628)	0.8574	0.0579
ϕ_-^2	0.2667	(0.1721 , 0.3898)	(0.1677 , 0.3844)	0.2738	0.0561
ϕ_{++}^{ni}	0.9333	(0.8035 , 0.9858)	(0.8253 , 0.9949)	0.9190	0.0482
ϕ_{++}^i	0.9657	(0.8965 , 0.9879)	(0.9087 , 0.9930)	0.9553	0.0237
ϕ_{--}^{ni}	0.1026	(0.0356 , 0.2245)	(0.0261 , 0.2089)	0.1122	0.0491
ϕ_{--}^i	0.0498	(0.0200 , 0.1255)	(0.0139 , 0.1138)	0.0603	0.0275
ϕ_{-+}^{ni}	0.3333	(0.0389 , 0.8248)	(0.0089 , 0.7706)	0.3767	0.2162
ϕ_{-+}^i	0.5110	(0.2254 , 0.7892)	(0.2232 , 0.7872)	0.5051	0.1470
ϕ_{+-}^{ni}	0.5714	(0.3623 , 0.7609)	(0.3682 , 0.7667)	0.5677	0.1033
ϕ_{+-}^i	0.5851	(0.3936 , 0.7494)	(0.3998 , 0.7552)	0.5783	0.0916

Tabla 4: Resultados del método "Plug-In", Objetivo Bayesiano y HDP (para $\alpha = 0.05$).

Al comparar el PI con el MOB vemos la amplia diferencia que existen en las probabilidades posteriores de ambos métodos. Por ejemplo, para ϕ_+^1 obtuvimos un PI de 0.7843 comparado con el MOB que varia entre 0.6540 y 0.8742. Por otro lado, la diferencia observada en el PI también está presente en el MOB. Los límites inferiores de los intervalos de credibilidad de las pruebas que no suponen independencia condicional tienden a ser menores que considerar el supuesto de independencia condicional excepto para ϕ_{--} . En este, el límite superior es mayor sin suponer independencia condicional. Nuevamente, en el caso de ϕ_{+-} no hay una diferencia notable pero para ϕ_{-+} hay una enorme discrepancia en los intervalos. Esto se puede apreciar en la desviación estándar para ϕ_{-+} . Sin suponer independencia condicional se obtuvo 0.2162 y bajo la suposición de independencia condicional es 0.1470. Además, son las desviaciones estándares mayores en comparación con los demás intervalos. Esto es consistente con lo que consiguieron Pereira & Pericchi (1990) sobre la diferencia que existe en suponer independencia condicional o no para este caso. En la Fig. 1, se observa las distribuciones posteriores de los intervalos de credibilidad. Estas muestran claramente que existe una diferencia amplia cuando se considera el supuesto de independencia o no, en cuyo caso estaríamos subestimando o sobrestimando las prob-

abilidades de tener la enfermedad D dado que ambas pruebas son positivas, negativas o combinadas.

Probabilidades	MOB	HPD
ϕ_+^1	0.2202	0.2182
ϕ_-^1	0.1952	0.1904
ϕ_+^2	0.2244	0.2184
ϕ_-^2	0.2178	0.2167
ϕ_{++}^{ni}	0.1824	0.1696
ϕ_{++}^i	0.0914	0.0844
ϕ_{--}^{ni}	0.1889	0.1829
ϕ_{--}^i	0.1056	0.0999
ϕ_{-+}^{ni}	0.7859	0.7617
ϕ_{-+}^i	0.5638	0.5640
ϕ_{+-}^{ni}	0.3985	0.3985
ϕ_{+-}^i	0.3557	0.3555

Tabla 5: Longitud de los intervalos de credibilidad para los métodos MOB y HPD

Los intervalos de credibilidad para HPD son aproximadamente igual a los intervalos del MOB. Por lo que, comparando con el método PI llegaremos a las mismas conclusiones. Ahora bien, si observamos la Tabla 5 la longitud de los intervalos de credibilidad para el método HPD son más corto que MOB, esto se debe a las características del método HPD.

Teorema 2. *Si dos pruebas son condicionalmente independientes dado D, no es necesario tener datos conjuntos.*

Demostración. Sean T_1 y T_2 pruebas condicionalmente independientes, entonces $P(T_2|D, T_1) = P(T_2|D)$. $P(T_2|D)$ no depende de T_1 , por lo tanto, no necesitamos datos conjuntos de T_1 y T_2 . □

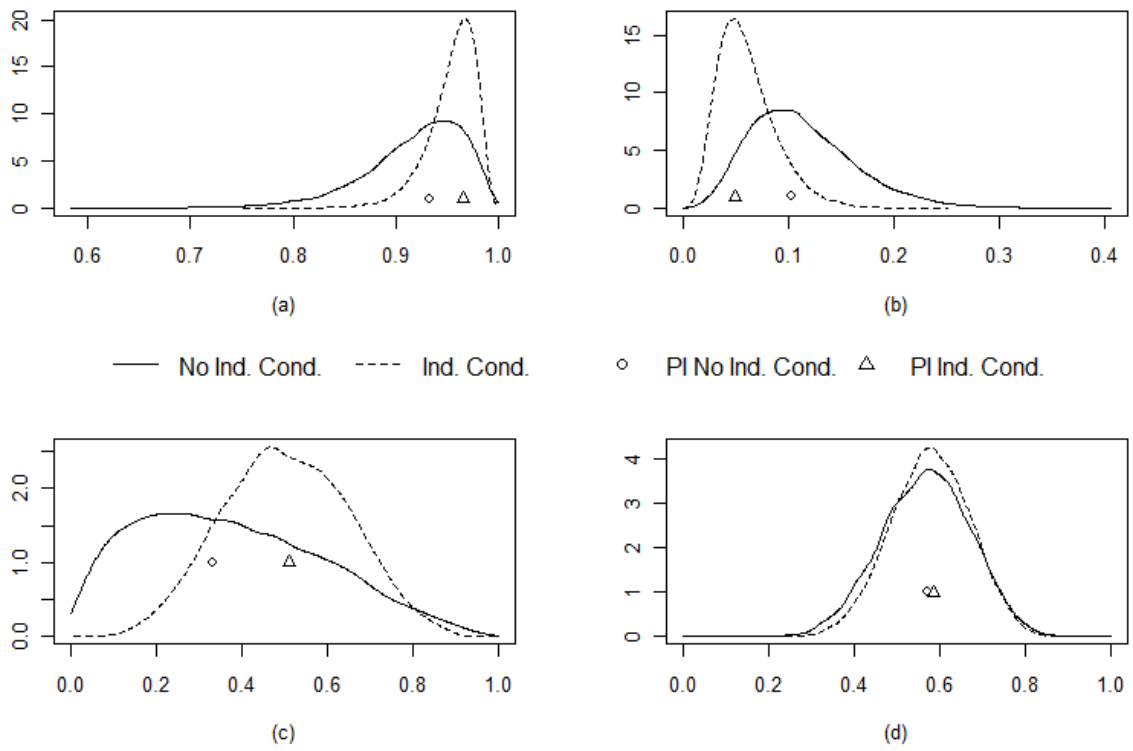


Figura 1: Gráficas de la simulación de Monte Carlo simple para: (a) ϕ_{++} , (b) ϕ_{--} , (c) ϕ_{-+} y (d) ϕ_{+-}

6 Conclusión

A menos que exista evidencia fuerte dada por el Factor de Bayes desarrollado, de que el modelo que supone independencia condicional es el adecuado, se deberán analizar las pruebas con el método bayesiano general, que no supone independencia condicional. Suponer independencia condicional indebidamente sobrestima la evidencia cuando se tienen dos positivos. Mientras que suponiendo independencia condicional indebidamente subestima la evidencia cuando se tienen dos negativos. Pero quizás lo más importante es cuando las pruebas, que suponen independencia condicional, con resultados combinados crean un conflicto donde la evidencia puede ser analizada de forma totalmente errónea. Por último, se comprobó que el método de alta densidad posterior crea unos intervalos más cortos que el método objetivo bayesiano.

Se desea realizar este análisis para un conjunto de datos donde una prueba haya sido realizada en dos veces. Además hacer una comparación entre los intervalos de máxima densidad posterior y los intervalos obtenidos en este trabajo.

7 Referencias

- [1] Mossman, D., & Berger, J. O. (2001). Intervals for posttest probabilities: a comparison of 5 methods. *Medical Decision Making*, 21(6), 498-507.
- [2] de Braganca Pereira, C. A., & Pericchi, L. R. (1990). Analysis of diagnosability. *Applied statistics*, 189-204.
- [3] DeGroot, M. H., & Schervish, M. J. *Probability and statistics*. 2002. pages: 72-73.
- [4] Hyndman, R. J. (1996). Computing and graphing highest density regions. *The American Statistician*, 50(2), 120-126.
- [5] Jeffreys, H. (1961). *The theory of probability*. Oxford University Press. page: 432
- [6] R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Apéndice

Factor de Bayes

Función de verosimilitud marginal para p . De forma análoga para q , p y q .

$$\begin{aligned}\int \binom{m}{x} p^x (1-p)^{m-x} \cdot \frac{p^0 (1-p)^0}{B(1,1)} dp &= \int \binom{m}{x} \frac{p^x (1-p)^{m-x}}{1} dp \\ &= \binom{m}{x} B(x+1, m-x+1) \int \frac{p^{(x+1)-1} (1-p)^{(m-x+1)-1}}{B(x+1, m-x+1)} dp \\ &= \binom{m}{x} B(x+1, m-x+1).\end{aligned}$$

Función de verosimilitud marginal para p_+ . De forma análoga para p_- , q_+ y q_- .

$$\begin{aligned}\int \binom{x^{++} + x^{+-}}{x^{++}} p_+^{x^{++}} (1-p_+)^{x^{+-}} \cdot \frac{p_+^{-\frac{1}{2}} (1-p_+)^{-\frac{1}{2}}}{B(\frac{1}{2}, \frac{1}{2})} dp_+ &= \int \binom{x^{++} + x^{+-}}{x^{++}} \frac{p_+^{x^{++}-\frac{1}{2}} (1-p_+)^{x^{+-}-\frac{1}{2}}}{B(\frac{1}{2}, \frac{1}{2})} dp_+ \\ &= \frac{\binom{x^{++} + x^{+-}}{x^{++}}}{B(\frac{1}{2}, \frac{1}{2})} \int p_+^{x^{++} + \frac{1}{2} - 1} (1-p_+)^{x^{+-} + \frac{1}{2} - 1} dp_+ \\ &= \frac{\binom{x^{++} + x^{+-}}{x^{++}} B(x^{++} + \frac{1}{2}, x^{+-} + \frac{1}{2})}{B(\frac{1}{2}, \frac{1}{2})} \int \frac{p_+^{x^{++} + \frac{1}{2} - 1} (1-p_+)^{x^{+-} + \frac{1}{2} - 1}}{B(x^{++} + \frac{1}{2}, x^{+-} + \frac{1}{2})} dp_+ \\ &= \frac{\binom{x^{++} + x^{+-}}{x^{++}} B(x^{++} + \frac{1}{2}, x^{+-} + \frac{1}{2})}{\pi}.\end{aligned}$$

hdr {hdrcde}

Regiones de Alta Densidad

Cálcula y grafica las regiones de alta densidad en una dimensión incluyendo el boxplot de HDR.

hdr(x, prob = c(50, 95, 99), den, h=hdrbw(BoxCox(x,lambda),mean(prob))),

```
lambda=1, nn=5000, all.modes=FALSE)
```

Código para calcular el Método Objetivo Bayesiano

El archivo de los datos debe contener la información de manera similar a la Tabla 1 pero sin el estado y tamaño de la muestra.

La siguiente función se creó para facilitar el cómputo del Teorema de Bayes.

```
TeoBayes <- function(a, b, c)
{
  d = b * a / ( b * a + c * (1-a) )
  return (d)
}
```

A continuación, con el siguiente código se calcularon los intervalos de credibilidad, plug-in, media y mediana.

```
ProbTests <- function(Data, sim, alpha=0.05)
{
  Data = Data
  sim = sim

  #Nivel de significancia
  low = sim*alpha/2
  upp = sim*(1-alpha/2)

  #Lectura de Datos
  xd = sum(Data[1,])
  nd = sum(Data)
```

```

x1.d = sum(Data[1,1:2])
x1.dc = sum(Data[2,1:2])
n1.dc = sum(Data[2,])
x2.d1 = Data[1,1]
x2.dc1 = Data[2,1]
x2.d = Data[1,1] + Data[1,3]
x2.dc = Data[2,1] + Data[2,3]
x2.d1.2 = Data[1,4]
n1.d1.2 = sum(Data[1,3:4])
x2.dc1.2 = Data[2,4]
n2.dc1.2 = sum(Data[2,3:4])
x2.d1.3 = Data[1,3]
x2.dc1.3 = Data[2,3]
x2.d1.4 = Data[1,2]
x2.dc1.4 = Data[2,2]

```

#Simulación de los parámetros

```

prev = rbeta(sim, xd + 1, nd - xd + 1)
p1.d = rbeta(sim, x1.d + 1, xd - x1.d + 1)
p1.dc = rbeta(sim, x1.dc + 1, n1.dc - x1.dc + 1)
p2.d1 = rbeta(sim, x2.d1 + 0.5, x1.d - x2.d1 + 0.5)
p2.dc1 = rbeta(sim, x2.dc1 + 0.5, x1.dc - x2.dc1 + 0.5)
p2.d = rbeta(sim, x2.d + 1, xd - x2.d + 1)
p2.dc = rbeta(sim, x2.dc + 1, n1.dc - x2.dc + 1)
p2.d1.2 = rbeta(sim, x2.d1.2 + 0.5 , n1.d1.2 - x2.d1.2 +
0.5 )
p2.dc1.2 = rbeta(sim, x2.dc1.2 + 0.5 , n2.dc1.2 - x2.dc1.2

```

```

+ 0.5)
p2.d1.3 = rbeta(sim , x2.d1.3 + 0.5 , n1.d1.2 - x2.d1.3 +
0.5)
p2.dc1.3 = rbeta(sim , x2.dc1.3 + 0.5 , n2.dc1.2 - x2.dc1.3
+ 0.5)
p2.d1.4 = rbeta(sim , x2.d1.4 + 0.5 , x1.d - x2.d1.4 + 0.5)
p2.dc1.4 = rbeta(sim , x2.dc1.4 + 0.5 , x1.dc - x2.dc1.4 +
0.5)

```

#Probabilidades Posteriores

```

pd.1 = TeoBayes(prev , p1.d , p1.dc)
pd.12 = TeoBayes(pd.1 , p2.d1 , p2.dc1)
pd.12.ind = TeoBayes(pd.1 , p2.d , p2.dc)
pd.1.2 = TeoBayes(prev , 1-p1.d , 1-p1.dc)
pd.12.2 = TeoBayes(pd.1.2 , p2.d1.2 , p2.dc1.2)
pd.12.ind2 = TeoBayes(pd.1.2 , 1-p2.d , 1-p2.dc)
pd.12.3 = TeoBayes(pd.1.2 , p2.d1.3 , p2.dc1.3)
pd.12.ind3 = TeoBayes(pd.1.2 , p2.d , p2.dc)
pd.12.4 = TeoBayes(pd.1 , p2.d1.4 , p2.dc1.4)
pd.12.ind4 = TeoBayes(pd.1 , 1-p2.d , 1-p2.dc)
pd.2 = TeoBayes(prev , p2.d , p2.dc)
pd.2.2 = TeoBayes(prev , 1-p2.d , 1-p2.dc)

```

#Plug-In

```

plug = TeoBayes(xd/nd , x1.d/xd , x1.dc/n1.dc)
plug.d = TeoBayes(plug , x2.d1/x1.d , x2.dc1/x1.dc)
plug.d.ind = TeoBayes(plug , x2.d/xd , x2.dc/n1.dc)

```

```

plug.2 = TeoBayes(xd/nd,1-x1.d/xd, 1-x1.dc/n1.dc )
plug.d.2 = TeoBayes(plug.2, x2.d1.2/n1.d1.2,
x2.dc1.2/n2.dc1.2)
plug.d.ind2 = TeoBayes(plug.2, 1-x2.d/xd, 1-x2.dc/n1.dc)
plug.d.3 = TeoBayes(plug.2, x2.d1.3/n1.d1.2,
x2.dc1.3/n2.dc1.2)
plug.d.ind3 = TeoBayes(plug.2, x2.d/xd, x2.dc/n1.dc)
plug.d.4 = TeoBayes(plug, x2.d1.4/x1.d, x2.dc1.4/x1.dc)
plug.d.ind4 = TeoBayes(plug, 1-x2.d/xd, 1-x2.dc/n1.dc)
plug2 = TeoBayes(xd/nd,x2.d/xd,x2.dc/n1.dc)
plug2.1 = TeoBayes(xd/nd,1-x2.d/xd,1-x2.dc/n1.dc)

```

#Tabla

```

probs <- matrix( c(plug,plug.2,plug2, plug2.1, plug.d,
plug.d.ind,plug.d.2,plug.d.ind2,plug.d.3,plug.d.ind3,
plug.d.4,plug.d.ind4, sort(pd.1)[low],sort(pd.1.2)[low],
sort(pd.2)[low], sort(pd.2.2)[low], sort(pd.12)[low],
sort(pd.12.ind)[low], sort(pd.12.2)[low],
sort(pd.12.ind2)[low], sort(pd.12.3)[low],
sort(pd.12.ind3)[low], sort(pd.12.4)[low],
sort(pd.12.ind4)[low], sort(pd.1)[upp],sort(pd.1.2)[upp],
sort(pd.2)[upp], sort(pd.2.2)[upp], sort(pd.12)[upp],
sort(pd.12.ind)[upp], sort(pd.12.2)[upp],
sort(pd.12.ind2)[upp], sort(pd.12.3)[upp],
sort(pd.12.ind3)[upp], sort(pd.12.4)[upp],
sort(pd.12.ind4)[upp], hdr(pd.1)$hdr[2,1],
hdr(pd.1.2)$hdr[2,1],hdr(pd.2)$hdr[2,1],

```



```

hdr(pd.2.2)$hdr[2,1],hdr(pd.12)$hdr[2,1],
hdr(pd.12.ind)$hdr[2,1],hdr(pd.12.2)$hdr[2,1],
hdr(pd.12.ind2)$hdr[2,1],hdr(pd.12.3)$hdr[2,1],
hdr(pd.12.ind3)$hdr[2,1],hdr(pd.12.4)$hdr[2,1],
hdr(pd.12.ind4)$hdr[2,1],hdr(pd.1)$hdr[2,2],
hdr(pd.1.2)$hdr[2,2],hdr(pd.2)$hdr[2,2],
hdr(pd.2.2)$hdr[2,2],hdr(pd.12)$hdr[2,2],
hdr(pd.12.ind)$hdr[2,2],hdr(pd.12.2)$hdr[2,2],
hdr(pd.12.ind2)$hdr[2,2],hdr(pd.12.3)$hdr[2,2],
hdr(pd.12.ind3)$hdr[2,2],hdr(pd.12.4)$hdr[2,2],
hdr(pd.12.ind4)$hdr[2,2],mean(pd.1),
mean(pd.1.2),mean(pd.2),mean(pd.2.2),mean(pd.12),
mean(pd.12.ind),mean(pd.12.2),mean(pd.12.ind2),
mean(pd.12.3),mean(pd.12.ind3),mean(pd.12.4),
mean(pd.12.ind4),sd(pd.1),sd(pd.1.2),sd(pd.2),sd(pd.2.2),
sd(pd.12),sd(pd.12.ind),sd(pd.12.2),sd(pd.12.ind2),
sd(pd.12.3),sd(pd.12.ind3),sd(pd.12.4),sd(pd.12.ind4)),
nrow = 12, dimnames=list(c("P(D|T1+)", "P(D|T1-)", "P(D|T2+)",
"P(D|T2-)", "P(D|T1+,T2+)-ni", "P(D|T1+,T2+)-i",
"P(D|T1-,T2-)-ni", "P(D|T1-,T2-)-i", "P(D|T1-,T2+)-ni",
"P(D|T1-,T2+)-i", "P(D|T1+,T2-)-ni", "P(D|T1+,T2-)-i" ),
c("PIM", "Lower_OBM", "Upper_OBM", "Lower_HDP", "Upper_HDP",
"Mean", "DE")))

```

#Gráfica

```
op <- par(mfrow=c(2,2))
```

```
Graf1 = plot(density(pd.12, to=1),main="",zero.line=0,ylim=
```

```

    c(0,max(density(pd.12.ind)$y)), xlab="(a)",ylab="")
lines(density(pd.12.ind, to=1), lty=2)
points(plug.d,y=1)
points(plug.d.ind, y=1,pch=2)
Graf2 = plot(density(pd.12.2,from=0),main="",zero.line=0,
ylim=c(0,max(density(pd.12.ind2)$y)), xlab="(b)",
ylab="")
lines(density(pd.12.ind2,from=0),lty=2)
points(plug.d.2,y=1)
points(plug.d.ind2, y=1,pch=2)
Graf3 = plot(density(pd.12.3,from=0,to=1),main="",
zero.line=0, ylim=c(0,max(density(pd.12.3)$y,
density(pd.12.ind3)$y)), xlab="(c)",ylab="")
lines(density(pd.12.ind3, to=1),lty=2)
points(plug.d.3,y=1)
points(plug.d.ind3, y=1,pch=2)
Graf4 = plot(density(pd.12.4,from=0,to=1),main="",
zero.line=0, ylim=c(0,max(density(pd.12.4)$y,
density(pd.12.ind4)$y)), xlab="(d)",ylab="")
lines(density(pd.12.ind4),lty=2)
points(plug.d.4,y=1)
points(plug.d.ind4, y=1,pch=2)

return(probs)
}

```