57
2ej

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

CONTRASTE DE HIPOTESIS EN LA FAMILIA ALFA

T E S I S

Que para obtener el Título de

A C T U A R I O

p r e s e n t a

EFRAIN EDUARDO SANTOS JAUREGUI

México, D. F. 1993

Dedicado a

mi mamá

mi papá

mis hermanos

Agradezco profundamente

al M. en C. Raúl Rueda Díaz del Campo, por haber aceptado ser mi director de tésis y por todo el tiempo que me dedicó para poder realizarla.

a los profesores de la comisión dictaminadora:

M. en C. Raúl Rueda Díaz del Campo
Dr. Federico O'Reilly Togno
M. en C. Beatríz Ospina Varón
Act. José Roberto Bautista Atenógenes
Act. Noé Moacir Vallejo González

por todas sus observaciones y sugerencias, mismas que enriquecieron la calidad de la tésis.

P r e f a c i o

La familia de distribuciones de probabilidad Alfa fue creada con la idea de contar con una familia de distribuciones alternativa a las usadas hasta entonces y que pudiera ser empleada en algunos de los métodos encaminados a resolver los problemas de bondad de ajuste, específicamente en aquellos métodos bayesianos en los que este problema se plantea como uno de contraste de hipótesis parámetricas dentro de la familia.

A manera de ejercicio académico surgió la inquietud de abordar no sólo el contraste asociado con la prueba de bondad de ajuste, sino de abordar un número mayor de contrastes de hipótesis sobre los parámetros de la familia Alfa.

Entre los problemas estadísticos que históricamente han recibido mayor atención se encuentran los contrastes de hipótesis y diversos han sido los procedimientos propuestos para resolver esta clase de problemas. Desde una perspectiva bayesiana, la teoría de la decisión proporciona un tratamiento general para resolver el problema de contrastar hipótesis.

Dentro del marco de la teoría de la decisión se encuentra el procedimiento bayesiano para el contraste de hipótesis paramétricas propuesto por Rueda(1992). Este procedimiento se ha aplicado con éxito a los contrastes en la familia exponencial y, bajo ciertas condiciones, las soluciones obtenidas coinciden con las obtenidas mediante procedimientos clásicos.

Este trabajo está encaminado a abordar algunos tipos de contrastes de hipótesis paramétricas en la familia de distribuciones Alfa y a obtener sus soluciones mediante el uso del procedimiento bayesiano mencionado en el párrafo anterior y a desarrollar un método bayesiano para probar Bondad de Ajuste dado un modelo probabilístico específico. En este procedimiento se utilizaron algunos elementos específicos como parte integral del mismo. Estos elementos son la divergencia logarítmica de kullback, como medida de la discrepancia etre modelos probabilísticos y el empleo de una distribución inicial de referencia sobre los parámetros de la familia Alfa.

En el capítulo 1 se presenta una breve introducción a las hipótesis paramétricas y a los conceptos básicos de la teoría bayesiana de toma de decisiones; se presenta el procedimiento de contraste de hipótesis utilizado; así como, las elecciones particulares para la función de utilidad y la función de distribución inicial de los parámetros.

En el capítulo 2 se introduce a la familia Alfa y se presentan algunos resultados sobre ella y que serán usados posteriormente; se presentan las soluciones a los contrastes seleccionados aplicando el procedimiento del capítulo 1; y se incluyen las regiones críticas para los mismos contrastes en el caso de resolverlos desde una perspectiva clásica del problema.

Finalmente, en el capítulo 3 se presentan los resultados de las simulaciones realizadas para algunos de los contrastes presentados en el capítulo 2.

Deseo profundamente que esta tésis, a diferencia de muchas otras, tenga la fortuna de ser consultada no sólo por quienes tuvieron una intervención directa o indirecta en su desarrollo, sino especialmente por personas totalmente ajenas a ella y que constituya una ayuda para el logro de sus metas particulares.

Efraín Eduardo Santos Jáuregui.
Noviembre de 1992.

C o n t e n i d o

Capítulo 1

Hipótesis Paramétricas y Enfoque Bayesiano

Capítulo 1
Hipótesis Paramétricas y Enfoque Bayesiano

1.1 Hipótesis Paramétricas

1.1.1 Parametrización

La necesidad por entender mejor a los fenómenos aleatorios, como por ejemplo la variación diaria del índice general de precios de la Bolsa Mexicana de Valores (B. M. V.), el nivel de contaminación dado por el índice metropolitano de la calidad del aire en la Ciudad de México del día siguiente, etc; ha tenido su orígen en muchas ocaciones por la manera en que pueden influir en algunas de las actividades de nuestra vida.

Considere que por alguna razón particular está interesado en el estudio de un fenómeno aleatorio específico. También considere que este fenómeno se representa por medio de una variable aleatoria X con distribución de probabilidad P(x). Al estar realizando el estudio es común suponer, en alguna de sus etapas, que el modelo probabilístico tiene una forma particular conocida, estableciéndola en frases tales como:

" .. Se sabe que el tiempo transcurrido entre dos emisiones consecutivas del Uranio 235, material radioactivo, tiene una distribución exponencial .. ".

" .. Si el diferencial en las utilidades generadas por dos acciones de emisoras distintas se comporta como una normal con varianza 2.4 y media comprendida entre 12 y 15, entonces .. ".

" .. Suponiendo que la proporción de personas afectadas por la última epidemia de cólera en Perú se puede describir mediante una distribución beta con media 0.1 y desviación estándar 0.05 ..".

En estos casos, lo que se está haciendo es tomar a P(x) como un elemento perteneciente a una determinada familia paramétrica de distribuciones $\mathcal{F}$.

Una familia paramétrica de distribuciones se define como un conjunto de distribuciones de probabilidad indexadas por medio de un parámetro θ, el cual toma valores en un conjunto Θ llamado espacio paramétrico; además, cualquier elemento dentro de la familia queda completamente identificado si se especifica un valor particular del parámetro y viceversa. Así pues, se está asumiendo que $P(x) \in \mathcal{F}$, con $\mathcal{F} = \{ P(x|\theta) : \theta \in \Theta \}$.

Una vez que se ha definido a la familia, se da una de dos situaciones: i) Se conoce completamente el valor del parámetro, o ii) No se conoce completamente el valor del parámetro. En el primer caso, como se conoce el valor del parámetro, se le asigna a X una distribución de probabilidad totalmente especificada, y en principio sólo se requiere calcular las probabilidades de que X tome valores en subconjuntos particulares, lo que representa un problema de probabilidad. En el segundo caso, como el valor del parámetro es desconocido, desde el punto de vista estadístico, para conocer más sobre él es necesario obtener información a través de una muestra aleatoria de X; en este caso se tiene un problema de inferencia estadística.

En este trabajo nos enfocaremos exclusivamente al caso en que no se conoce completamente el valor del parámetro

1.1.2 Contraste de Hipótesis

Un camino para hacer inferencias sobre el valor del parámetro es el contraste de hipótesis paramétricas, en el que una hipótesis paramétrica es una afirmación sobre el valor del parámetro.

De esta manera, el problema de contraste de hipótesis paramétricas consiste en determinar cuál de las hipótesis consideradas es la más adecuada, en base a los datos proporcionados por una muestra aleatoria de X. Estos contrastes, planteados en forma general, son del tipo

$$H_0: \theta \in \Theta_0 \quad \text{vs.} \quad H_1: \theta \in \Theta_1$$

con Θ_0 y Θ_1 subconjuntos no vacíos de Θ, tales que $\Theta_0 \cap \Theta_1 = \emptyset$. Si el subconjunto asociado con la hipótesis tiene un sólo elemento se dice que la hipótesis es simple y en caso contrario se dice que es compuesta. Históricamente se le ha llamado a la hipótesis H_0 la "Hipótesis nula" y a la H_1 la "Hipótesis alternativa".

Una vez resuelto el problema, se obtiene el valor particular del parámetro o en su caso el subconjunto (y por lo tanto la subfamilia paramétrica) que mejor explica el comportamiento de X, dependiendo de si el subconjunto asociado con la hipótesis aceptada tiene uno o más elementos.

Diversos procedimientos han sido desarrollados para contrastar hipótesis. Entre los más conocidos están el Lema de Neyman-Pearson cuando las hipótesis son simples y el método de cociente de verosimilitudes generalizado cuando al menos una de ellas es compuesta, con los que se encuentra una región del espacio muestral, la región crítica C^*, tal que la hipótesis nula se rechaza si la muestra cae dentro de C^* para un nivel de significancia predeterminado. Entre los procedimientos bayesianos desarrollados para resolver el problema de contraste de hipótesis está el propuesto por Rueda(1992), en el que se plantea este problema como uno de decisión, en donde se debe escoger una de las dos decisiones posibles:

$$d_0 : \text{Aceptar } H_0 \qquad \text{y} \qquad d_1 : \text{Rechazar } H_0$$

Este procedimiento sigue el método general dado por la Teoría de la Decisión para tomar una decisión y será el que se ocupe para resolver los contrastes planteados en este trabajo.

1.2 Perspectiva Bayesiana

1.2.1 Teoría de la decisión

Supóngase que el estudio está motivado porque durante el desarrollo de alguna actividad, y bajo ciertas condiciones, se presenta más de una forma de actuar y se tiene el problema de decidir cuál de los caminos posibles es el que se va a seguir. La dificultad presente es que las consecuencias de tomar cualquiera de estos caminos se desconocen, por depender de un conjunto de sucesos llamados sucesos inciertos por ignorase cuál de ellos es el que ocurre. Para ejemplificar este tipo de problemas considere las siguientes situaciones.

El inversionista en acciones que cotizan en la B. M. V. que busca obtener las mejores ganancias con el capital que ha invertido. Para lograr su propósito, cada cierto período de tiempo tiene que decidirse por una de tres posibles formas de actuar: mantener su posición accionaria, vender las acciones que posee para adquirir otras de una emisora distinta o comprar más acciones de la misma emisora. La utilidad o pérdida registrada al invertir en acciones, y en general en cualquier instrumento de inversión bursátil, está dada por la diferencia entre el precio al que se venden y el precio al que se compran estos valores. Las operaciones de compra y venta de valores se realizan en el piso de remates de la B. M. V.. Ahora bien, los precios de las acciones son fijados por las fuerzas de mercado de la oferta y la demanda, estas fuerzas dependen a su vez de la situación de la empresa emisora (utilidades generadas hasta ese momento, perspectivas de crecimiento, giro y penetración en el mercado, etc.), de la situación económica del país (índices de inflación, producto interno bruto, etc.), de la colocación de nuevas emisiones en bolsa, de acontecimientos políticos (nacionales e internacionales), del comportamiento de las bolsas de valores de Nueva York y Tokio, de temores de los propios inversionistas, etc.

Todo estos factores dan lugar a continuas fluctuaciones en los precios de las acciones lo que impide que el inversionista tenga la certeza sobre cuáles acciones subirán, cuáles bajarán y cuáles permanecerán sin cambio al final del período. Por lo que le resulta muy difícil decidir qué debe hacer; puesto que si se deshace de las actuales acciones para comprar otras, y las condiciones en el mercado en el corto plazo hacen que el precio de las acciones recién vendidas tenga un mayor crecimiento que el de las recién adquiridas, habrá perdido una buena oportunidad de incrementar sus utilidades. Más aún, si el precio de las acciones adquiridas baja, registrará pérdidas. Así pues, dependiendo de la decisión que tome el inversionista y del comportamiento del mercado accionario, es decir, de los precios de las acciones, las consecuencias pueden ir desde sufrir una gran pérdida hasta la de registrar una ganancia enorme en la operación.

El gerente de una sucursal bancaria, con el objetivo de ofrecer un servicio rápido y eficaz, se ve ante el problema de decidir cuántas cajas de atención al público debe aumentar en su sucursal para lograr que el tiempo promedio que invierte cada uno de sus clientes dentro del banco sea menor a 10 minutos, todo esto sin incurrir en gastos innecesarios. El, por desgracia, desconoce con exactitud el tiempo invertido por cada cliente puesto que éste depende, según consideraciones del gerente, del número de clientes que acuden a la sucursal durante el día y del número de operaciones que cada uno de ellos realiza; por lo que los posibles valores para el tiempo promedio por cliente consituyen el conjunto de sucesos inciertos. Además él está conciente de que el tomar alguna decisión puede tener consecuencias que van desde la de ofrecer el servicio deseado con perspectivas de incrementar su clientela y con un manejo eficiente tanto de recursos humanos como materiales (evitando la presencia de cajas ociosas), hasta consecuencias muy negativas como seguir con un servicio lento, la lógica molestia de la clientela (que bien pudiera cambiar de banco) y tener cajas frecuentemente ociosas ocacionando un inútil incremento de los gastos operativos de la sucursal.

A pesar de que en esta clase de problemas existe un conjunto de sucesos inciertos, se tiene un cierto grado de creencia sobre su ocurrencia, ya que la experiencia del inversionista y su conocimiento sobre la salud financiera de algunas empresas emisoras le hace creer que los precios de algunas de las acciones subirán más que otras y que habrán otras que bajen, habiendo precios más probables que otros para cada una de ellas. Además, si lo estima necesario puede recurrir a indicadores macroeconómicos, al análisis técnico bursátil, al análisis de estados financieros, etc. que bien pueden modificar la percepción sobre los precios de las acciones estimados inicialmente. Por otra parte, al gerente le parece poco probable que el número de personas que acuden a la sucursal en un día cualquiera sea menor a 20 o mayor a 100,000.

También es claro que para quien debe tomar una decisión hay consecuencias que son preferidas a otras. Así, el inversionista tiene una obvia y marcada preferencia por las grandes ganancias y aversión por las pérdidas; y para el gerente sería preferible aumentar en cinco el número de cajas y tener cajas frecuentemente ociosas pero captar más clientela que tener que incrementarlas en dos y perder clientela, lo que podría significar no ser considerado para ocupar un puesto superior en el futuro.

Los anteriores son ejemplos de problemas de decisión en ambiente de incertidumbre; Savage formalizó las ideas presentes en esta clase de problemas mediante la especificación del espacio de decisiones, conjunto formado por las posibles decisiones o formas de actuar; del espacio de sucesos inciertos, conjunto formado por los sucesos sobre los que existe incertidumbre sobre cuál de ellos ocurre, cada uno de estos conjuntos con la característica de ser conjuntos exclusivos y exhaustivos; y la especificación del espacio de consecuencias, conjunto formado por los resultados de tomar una decisión cuando ocurre alguno de los sucesos inciertos, ver por ejemplo Savage(1972).

También describió una forma de tomar decisiones, en la que se tiene que cuantificar la ocurrencia de los sucesos inciertos, y que de acuerdo con la interpretación subjetiva de la probabilidad consiste en una asignación de probabilidades que refleje el grado de creencia que se tiene sobre la ocurrencia de ellos (Lindley, 1971); así como cuantificar las preferencias que se tiene sobre las posibles consecuencias mediante una función de utilidad. Se tiene que calcular el valor esperado de la utilidad (utilidad esperada) para cada decisión, tomado con respecto al espacio de sucesos inciertos; por último, se debe escoger como solución al problema a aquella decisión que maximice el valor esperado de la utilidad, ésta decisión es denominada la decisión óptima.

1.2.2 Teorema de Bayes

Siguiendo con la variable aleatoria X con distribución de probabilidad P(x) desconocida y suponiendo que su distribución puede considerarse como un elemento de la familia paramétrica $\mathcal{F}$, es decir, $P(x) = P(x|\theta)$ para algún $\theta \in \Theta$ desconocido; con Θ el espacio paramétrico.

Enfocando el problema desde la perspectiva bayesiana se tiene que como hay incertidumbre sobre el valor del parámetro, debido al conocimiento o información que se tiene de él, se considera a θ una variable aleatoria con una distribución de probabilidad $P(\theta)$ sobre Θ. La reacción natural de alguien con un problema de decisión es la de tratar de reducir la incertidumbre existente en él. Bajo estas circunstancias, es común querer saber más sobre el parámetro, y también es común obtener más información a través de una muestra aleatoria z de X. La información contenida en z y la relación que guarda con el parámetro están dadas por la función de verosimilitud de la muestra $P(z|\theta)$, la cuál se incorpora al conocimiento inicial representado por $P(\theta)$ para obtener la distribución posterior (o distribución final) $P(\theta|z)$. Esta distribución posterior de θ representa el grado de creencia que se tiene, después de haber observado z, de la ocurrencia de los valores de θ modificado por la información adicional en z. Esta distribución posterior se obtiene a través del Teorema de Bayes que a continuación se enuncia.

Teorema de Bayes: Sea $P(\theta)$ la distribución de probabilidad inicial de θ, con $\theta \in \Theta$ y $P(z|\theta)$ la función de verosimilitud de la muestra aleatoria $z=\{x_1, x_2,.., x_n\}$ de X, entonces la distribución de probabilidad posterior de θ dada la información proporcionada por z y denotada por $P(\theta|z)$ está dada por

$$P(\theta|z) = \frac{P(z|\theta)P(\theta)}{\int_{\Theta} P(z|\theta)P(\theta)d\theta} \tag{1}$$

donde la integral es en el sentido de la medida de Lebesgue, es decir, se trata de una suma cuando Θ es discreto y de una integral de Riemman si es continuo; y como el denominador en (1) es una constante, esta relación suele manejarse en su forma proporcional

$$P(\theta|z) \propto P(z|\theta)P(\theta)$$

En algunas ocaciones se puede desear que únicamente la información en z determine a la distribución posterior $P(\theta|z)$, esto puede deberse por tener un conocimiento inicial vago o simplemente por no querer incorporar este conocimiento en la asignación de probabilidades. En este caso se debe tomar como distribución inicial a una distribución de referencia, $\pi(\theta)$, para poder aplicar el teorema de Bayes y obtener la distribución posterior de referencia $\pi(\theta|z)$.

De lo anterior, podemos decir que si el problema de decisión puede ser representado con un espacio de decisiones $D=\{d\}$, un espacio de sucesos inciertos Θ y un espacio de consecuencias $C=D\times\Theta$; y además se define una función de utilidad u(c) sobre el espacio C y una distribución de probabilidad inicial $P(\theta)$, o bien una distribución inicial de referencia, sobre Θ; entonces, la utilidad esperada de d está dada por

$$u^{*}(d) = \int_{\Theta} u(d,\theta) P(\theta|z)\, d\theta$$

con $P(\theta|z)$ dado por (1). La decisión óptima $d^{*} \epsilon D$, solución del problema después de haber obtenido información adicional sobre θ a través de z, será aquella que maximice la utilidad esperada $u^{*}(d)$, es decir, aquella que cumpla con

$$u(d^{*}) = \max_{d \epsilon D} u^{*}(d)$$

en donde " max " denota el valor máximo.

1.3 Un Procedimiento de Contraste de Hipótesis

1.3.1 Planteamiento

Una vez que se ha descrito la forma de tomar decisiones, según la teoría bayesiana de toma de decisiones, se presentará el procedimiento de contraste de hipótesis propuesto por Rueda(1992), en el que explícitamente se toma en cuenta la presencia de un parámetro de ruido.

Sea X una variable aleatoria con distribución de probabilidad denotada por $P(x|\theta,\omega)$, en la que $\theta \epsilon \Theta$ y $\omega \epsilon \Omega$ son el parámetro de interés y el parámetro de ruido respectivamente. Sea z una muestra aleatoria de X, entonces la familia paramétrica es $\mathcal{F} = \{ P(x|\theta,\omega) : \theta \epsilon \Theta, \omega \epsilon \Omega\}$. Supongamos que se quiere resolver el contraste de hipótesis (paramétricas):

$$H_0: \theta \epsilon \Theta_0 \quad \text{vs.} \quad H_1: \theta \epsilon \Theta_1 \ , \ \text{con} \quad \Theta_0 \cap \Theta_1 = \emptyset. \qquad (2)$$

Como se mencionó anteriormente, este problema se puede plantear como un problema de decisión en el que el espacio de decisiones, D, tiene únicamente dos elementos

$$d_0 : \text{ Aceptar } H_0 \qquad \text{y} \qquad d_1 : \text{ Rechazar } H_0$$

y el espacio de sucesos inciertos es $\Theta x \Omega$. Por lo que el problema es decidir en cuál de los dos subconjuntos de Θ está θ. Así pues, parece razonable estimar primeramente el valor de θ dentro de cada subconjunto (siguiendo algún criterio particular) y posteriormente decidir cuál de estas dos estimaciones es la que mejor lo aproxima. Obviamente, no se requiere estimar el valor del parámetro cuando el subconjunto tiene sólo un elemento. Se termina escogiendo la decisión d_1 si la mejor estimación está en el subconjunto sustentado por la hipótesis alternativa, en caso contrario se escoge la decisión d_0

Por trabajar dentro de la familia la hipótesis H_i: $\theta \in \Theta_i$ es equivalente a la hipótesis H_i: $P(x|\theta,\omega) \in \mathcal{F}_i$, donde $\mathcal{F}_i = \{P(x|\theta_i,\omega) : \theta_i \in \Theta_i, \omega \in \Omega\}$, con $i=1,2$; y el problema se puede interpretar como el decidir si la verdadera distribución de X la aproxima mejor algún elemento de $\mathcal{F}_0$ o alguno de $\mathcal{F}_1$. De esta forma resulta conveniente emplear una función de utilidad que incluya a alguna medida de la aproximación entre dos modelos probabilísticos.

Más que hablar de la aproximación entre los modelos se hablará de la discrepancia entre ellos, entendiendo por discrepancia el qué tanto difieren los modelos entre sí. Es claro que un modelo aproxima mejor a otro entre menor sea la discrepancia entre ellos. Denotaremos por $\delta(\theta,\omega:\theta_i)$ a la función de discrepancia entre el modelo verdadero y el especificado por θ_i.

Como todos los valores esperados utilizados en el procedimiento son tomados con respecto a la distribución posterior, se denota en lo que sigue a

$$E_{\theta,\omega|z}(\cdot) = \iint_{\Omega\Theta} (\cdot)\, P(\theta,\omega|z)\, d\theta\, d\omega \tag{3}$$

Tomando en cuenta lo anterior, la solución bayesiana al contraste de hipótesis planteado en (2) consistirá en:

a) Tomar como distribución de probabilidad a una distribución posterior $P(\theta,\omega|z)$, o si se prefiere a una distribución posterior de referencia, para (θ,ω).

b) Si Θ_i tiene un elemento hay que definir a θ_i^* como $\theta_i^* = \theta_i$.
Si Θ_i tiene más de un elemento y si $\tilde{\theta}_i$ denota a un estimador de θ en Θ_i, hay que definir a θ_i^* como aquel estimador que minimice el valor esperado de la discrepancia, es decir, aquel que satisfaga la expresión

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_i^*)] = \inf_{\tilde{\theta}_i \in \Theta_i} E_{\theta,\omega|z}[\delta(\theta,\omega:\tilde{\theta}_i)] \tag{4}$$

donde " inf " denota el ínfimo.

c) Tomar como función de utilidad a una función lineal de la discrepancia entre las distribuciones especificadas por θ y θ_i^*, de la forma

$$u(d_i,\theta,\omega) = -A\delta(\theta,\omega:\theta_i^*) + B_i \qquad (i=1,2)$$

donde $A > 0$ y $B_i \in \mathbb{R}$ son constantes arbitrarias elegidas por quien se enfrenta al problema de decisión.

d) Por último, se debe escoger d_1:Rechazar H_0, como la solución óptima al problema si y sólo si

$$E_{\theta,\omega|z}[u(d_1,\theta,\omega)] > E_{\theta,\omega|z}[u(d_0,\theta,\omega)]$$

o equivalentemente, en términos de la diferencia de discrepancias definida como $\delta(\theta,\ \omega:\theta_0^*,\ \theta_1^*) = \delta(\theta,\ \omega:\theta_1^*) - \delta(\theta,\ \omega:\theta_0^*)$, si y sólo si

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_0^*,\theta_1^*)] < \frac{B_1 - B_0}{A} \tag{5}$$

En caso contrario, se debe escoger d_0 : Aceptar H_0 .

Es claro que la solución al problema de decisión depende de las elecciones particulares de la función de utilidad y de la distribución inicial de los parámetros involucrados. Vale la pena notar que el haber tomado como forma de estimar el parámetro de interés la enunciada en el punto (b) y una utilidad con la forma mencionada en el punto (c) también constituyen elecciones particulares puesto que el decisor puede tomar, por ejemplo, los estimadores máximo-verosímiles dentro de cada subconjunto y/o una utilidad completamente diferente; sin embargo, las dadas en (b) y (c) son consideradas como parte integral del procedimiento. Las elecciones utilizadas en este trabajo se presentan en la siguiente sección.

1.3.2 Elecciones particulares

a) Función de utilidad. Siguiendo la propuesta de Bernardo(1982) se tomó a la Divergencia Logarítmica de Kullback como la medida de discrepancia entre los modelos probabilísticos especificados por los valores del parámetro θ y θ_i, definida por

$$\delta(\theta,\omega:\theta_i) = \int P(x|\theta,\omega)\ln\frac{P(x|\theta,\omega)}{P(x|\theta_i,\omega)}\,dx \tag{6}$$

y representa la información esperada proporcionada por la observación $X = x$ para discriminar a favor de la verdadera distribución de X, $P(x|\theta)$, contra el modelo propuesto $P(x|\theta_i)$, (Kullback, 1959). Dicho de otra forma, es la información esperada proporcionada por la obsevación $X = x$ necesaria para pasar del modelo propuesto al modelo verdadero.

Ahora, sólo hay que especificar las constantes A, B_0 y B_1 para especificar completamente a la función de utilidad. Estas constantes son arbitrarias y las debe fijar quien tiene el problema de decisión, sin embargo puede no ser una tarea sencilla. Se puede observar que si la contribución de la discrepancia a la utilidad esperada es la misma para las dos decisiones, la diferencia de las utilidades esperadas estaría dada por $B_1 - B_0$. Si esta diferencia es mayor que cero, representa la preferencia por escoger a d_1 sobre d_0 y si es menor que cero representa la preferencia por escoger a d_0 sobre d_1; por lo que esta diferencia es la predisposición por escoger alguna de las hipótesis. Por otra parte, debe tomarse $B_1 = B_0$ si no hay una predisposición por alguna decisión mas que la reflejada por la discrepancia. Por lo que toca a la constante A, se trata únicamente de un factor de ponderación o factor de escala sobre la discrepancia.

Las soluciones obtenidas en el siguiente capítulo se analizarán para algunas combinaciones de estas constantes.

b) Distribución inicial. Con la intención de contar con un procedimiento que también contemple los casos en que no se dispone o no se desea incorporar la información inicial sobre los parámetros, se va a utilizar como distribución inicial de los parámetros involucrados a una distribución inicial de referencia $\pi(.)$, concretamente la obtenida a través del método de maximización de la información desconocida y reportada por Bernardo(1979).

En el siguiente capítulo se introduce una familia paramétrica particular y se presentan algunos contrastes sobre sus parámetros a los cuales se les va a aplicar el procedimiento recién descrito para resolverlos.

Capítulo 2

Contraste de Hipótesis en la Familia Alfa

Capítulo 2
Contraste de Hipótesis en la Familia Alfa

2.1 Familia Alfa

La familia paramétrica con la que se va trabajar es la familia Alfa creada por Bernardo(1982) y los contrastes de hipótesis que se van a analizar serán sobre los parámetros de la distribución Alfa.

La principal razón que motivó la creación de esta familia tuvo su orígen en el estudio del problema de Bondad de Ajuste. Para poder parametrizar el problema de Bondad de Ajuste y reducirlo a un problema de contraste de hipótesis paramétricas se necesita contar con una familia de distribuciones que tenga entre sus elementos a la función de densidad uniforme en el intervalo [0,1] y que sea densa en el entorno de la densidad uniforme. Dentro de las familias que han sido empleadas para ese propósito se encuentra la familia de distribuciones Beta; sin embargo, ésta tiene el inconveniente de complicar los desarrollos matemáticos para obtener una solución. Es por esto que se se pensó en alguna otra familia de distribuciones que pudiera ser empleada para tal propósito y que resultara matemáticamente más tratable que las anteriores, es por eso que se crea la familia de distribuciones Alfa.

A continuación se define a la distribución y a la familia Alfa y se presentan algunos resultados que se serán usados en secciones posteriores.

Definición. Sea X una variable aleatoria contínua en el intervalo [0,1], se dice que X se distribuye Alfa con parámetros (θ,ω), con $\theta > 0$ y $0 < \omega < 1$, si su función de densidad de probabilidad es de la forma

$$Al(x|\theta,\omega) = \begin{cases} \theta\left(\dfrac{x}{\omega}\right)^{\theta-1} & , 0 \le x \le \omega \\ \theta\left(\dfrac{1-x}{1-\omega}\right)^{\theta-1} & , \omega < x \le 1 \end{cases} \tag{7}$$

Es fácil probar que la función de distribución está dada por

$$F(x|\theta,\omega) = \begin{cases} \omega\left(\dfrac{x}{\omega}\right)^{\theta} & , \ 0 \le x \le \omega \\ 1-(1-\omega)\left(\dfrac{1-x}{1-\omega}\right)^{\theta} & , \ \omega < x \le 1 \end{cases} \tag{8}$$

Se define a la familia Alfa como el conjunto de funciones de densidad de probabilidad dado por

$$\mathcal{F} = \{ \, Al(x|\theta,\omega) : \theta > 0, \ 0 < \omega < 1 \}$$

Se han propuesto algunas familias más densas en las que se mantienen las características esenciales de la familia Alfa, por ejemplo las hechas por Bayarri(1984); sin embargo, únicamente se considerará en este trabajo la que se acaba de definir.

De (7) se observa que si $\theta = 1$, se cumple que $Al\,(x|\theta,\omega) = U\,(x|0,1)$ para toda ω; con $U\,(x|0,1)$ la función de densidad uniforme en el intervalo $[0,1]$. Si $\theta < 1$, la distribución es convexa y tiene un mínimo. Si $\theta > 1$, la distribución es cóncava y tiene un máximo. Además, la distribución alcanza el valor extremo θ cuando $x = \omega$.

La verosimilitud de la muestra aleatoria $z = \{x_1, x_2, \ldots, x_n\}$ de X, es

$$P(z|\theta,\omega) = \theta^{n} \exp\left[\, - n\,(\theta - 1)\, t\,(\omega)\,\right]$$

en donde

$$t(\omega) = \frac{1}{n}\left[\, \sum_{i \in I(\omega)} \ln \frac{\omega}{x_i} + \sum_{j \in J(\omega)} \ln \frac{1-\omega}{1-x_j} \,\right]$$

con $I\,(\omega) = \{\, i : x_i \le \omega \,\}$ y $J\,(\omega) = \{\, j : x_j > \omega \,\}$

La función de densidad de probabilidad inicial de referencia para los parámetros de la distribución Alfa (θ,ω) que se va a usar, como se mencionó en la sección 1.3.2, es la obtenida por Bernardo(1979); debido a que no es la intención determinar como fue deducida, únicamente se hará uso de dicho resultado.

$$\pi(\theta, \omega) \propto \theta^{-1}$$

Combinando esta distribución inicial con la verosimilitud de z, y por el teorema de Bayes, se obtiene que la función de densidad posterior de referencia es

$$\pi(\theta, \omega | z) \propto \theta^{n-1} \exp[-n(\theta - 1) t(\omega)]$$

que puede escribirse como

$$\pi(\theta, \omega | z) = \frac{t^{-n}(\omega)\, e^{n t(\omega)}}{\int_{\Omega} t^{-n}(\omega) e^{n t(\omega)} d\omega} Ga[\theta | n, n t(\omega)]$$

donde $Ga[\theta \mid n, n t(\omega)]$ denota a la función de densidad Gamma con parámetros n y $n t(\omega)$.

2.2 Contrastes de hipótesis

Ahora, se van a enunciar algunos contrastes de hipótesis sobre los parámetros de la distribución Alfa, mismos que serán resueltos usando el procedimiento de la sección 1.3. Para simplificar un poco la notación se tomará

$$E_{\omega|z}(\cdot) = \frac{\int_{\Omega} (\cdot)\, t^{-n}(\omega)\, e^{n t(\omega)}\, d\omega}{\int_{\Omega} t^{-n}(\omega)\, e^{n t(\omega)}\, d\omega},$$

2.2.1 $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$

En este caso ω es tomado como el parámetro de ruido y debido a que los subconjuntos para el parámetro de interés para cada hipótesis tienen un sólo elemento no hay que llevar a cabo el proceso de estimación. Ahora bien, para resolver el contraste se debe calcular primeramente la discrepancia entre el modelo verdadero

(θ,ω) y uno propuesto (θ_l,ω), usando (6) se tiene

$$\delta(\theta,\omega:\theta_l) = \int P(x|\theta,\omega)\ln\frac{P(x|\theta,\omega)}{P(x|\theta_l,\omega)}\,dx$$

Por la forma de la distribución Alfa, para llevar a cabo la integración con respecto a x, se debe partir la integral de (0,1] en la suma de una integral sobre $[0,\omega)$ y otra en $(\omega,1]$. Sustituyendo las distribuciones Alfa correspondientes se tiene que

$$\delta(\theta,\omega:\theta_l) = \int_0^{\omega}\theta\left(\frac{x}{\omega}\right)^{\theta-1}\ln\left[\frac{\theta(x/\omega)^{\theta-1}}{\theta_l(x/\omega)^{\theta_l-1}}\right]dx + \int_{\omega}^{1}\theta\left(\frac{1-x}{1-\omega}\right)^{\theta-1}\ln\left[\frac{\theta\left(\frac{1-x}{1-\omega}\right)^{\theta-1}}{\theta_l\left(\frac{1-x}{1-\omega}\right)^{\theta_l-1}}\right]dx$$

y resolviendo las integrales por cambio de variable se obtiene que la discrepancia entre ellos es

$$\delta(\theta,\omega:\theta_l) = \ln\frac{\theta}{\theta_l} + \frac{\theta_l-\theta}{\theta}$$

por lo que la diferencia de discrepancias está dada por

$$\delta(\theta,\omega:\theta_0,\theta_1) = \ln\frac{\theta_0}{\theta_1} + \frac{\theta_1-\theta_0}{\theta} \tag{9}$$

como puede observarse, ninguna de estas dos últimas expresiones depende de ω.

Para poder aplicar el criterio dado en (5) para decidir si se acepta o se rechaza la hipótesis nula se tiene que calcular el valor esperado de la diferencia de discrepancias. Sustituyendo la diferencia de discrepancias dada por (3) se tiene que

$$E_{\theta,\omega|x}[\delta(\theta,\omega:\theta_0,\theta_1)] = E_{\theta,\omega|x}\left[\ln\frac{\theta_0}{\theta_1} + \frac{\theta_1-\theta_0}{\theta}\right]$$

y como el término que contiene al logarítmo natural es una constante, al desarrollar se obtiene

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_0,\theta_1)] = \ln\frac{\theta_0}{\theta_1} + E_{\omega|z}\{(\theta_1-\theta_0)\int_{\Theta}\frac{1}{\theta}\,Ga[\theta|n,nt(\omega)]\,d\theta\}$$

por otra parte se sabe que si Y se distribuye Ga(y|a,b), entonces $E_Y(1/\theta) = b/(a-1)$. Al usar este resultado se obtiene que el valor esperado de la diferencia de discrepancias es de la forma

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_0,\theta_1)] = \ln\frac{\theta_0}{\theta_1} + \frac{n(\theta_1-\theta_0)}{n-1}\,E_{\omega|z}[t(\omega)]$$

y por (5) se concluye que se escogerá a d_1 : Rechazar H_0 como la decisión óptima, si y sólo si

$$\ln\frac{\theta_0}{\theta_1} + \frac{n(\theta_1-\theta_0)}{n-1}\,E_{\omega|z}[t(\omega)] < \frac{B_1-B_0}{A} \tag{10}$$

y en caso contrario se deberá escoger d_0 : Aceptar H_0.

2.2.2 H_0: $\theta = \theta_0$ vs. H_1: $\theta \neq \theta_0$

También en este caso ω es el parámetro de ruido, pero a diferencia del contraste anterior el subconjunto especificado en la hipótesis alternativa tiene más de un elemento. De acuerdo con nuestro procedimiento, como Θ_1 tiene más de un elemento, para encontrar la solución al contraste se tiene que realizar el proceso de estimación para poder encontrar el valor $\theta_1^* \in \Theta_1$ que satisfaga a (4). Para ello, se debe encontrar primeramente el valor esperado de la discrepancia entre θ y θ_1 dada en (9). De modo que se tiene que

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_1)] = E_{\theta,\omega|z}\left[\ln\frac{\theta}{\theta_1} + \frac{\theta_1-\theta}{\theta}\right]$$

distribuyendo el valor esperado y simplificando

$$E_{\theta,\omega|z}[\,\delta(\theta,\omega:\theta_1)\,] = E_{\theta,\omega|z}[\,\ln\theta + \frac{\theta_1}{\theta}\,] - E_{\theta,\omega|z}[\,\ln\theta_1 + 1\,]$$

hay que notar que el segundo término es una constante, por lo que

$$E_{\theta,\omega|z}[\,\delta(\theta,\omega:\theta_1)\,] = E_{\omega|z}\left\{\int_{\Theta} (\ln\theta + \frac{\theta_1}{\theta})\, Ga[\,\theta|n,nt(\omega)]\, d\theta\right\} - \ln\theta_1 - 1$$

y como se cumple que si Y se distribuye Ga(y|a,b), entonces $E_Y(\ln Y) = \Psi(a) - \ln b$ y $E_Y(1/\theta) = b/(a-1)$; en donde $\Psi(\,.\,)$ es la función digamma (ver por ejemplo Abramowitz & Stegun, 1968); al aplicar esto a la ecuación anterior y desarrollar se obtiene que

$$E_{\theta,\omega|z}[\,\delta(\theta,\omega:\theta_1)\,] = \frac{n\theta_1}{n-1}E_{\omega|z}[\,t(\omega)\,] - \ln\theta_1 + E_{\omega|z}\{\,\Psi(n) - \ln[\,nt(\omega)\,]\,\} - 1 \tag{11}$$

Una vez que se cuenta con el valor esperado de la discrepancia, para encontrar el valor θ_1^* en el que esta expresión alcanza el mínimo hay que calcular la derivada de esta función con respecto a θ_1 y resolver la ecuación correspondiente. Como los dos últimos términos no dependen de θ_1, al final de estos pasos se obtiene que

$$\theta_1^* = \frac{n-1}{n\,E_{\omega|z}[\,t(\omega)\,]}$$

Por el criterio de la segunda derivada se puede comprobar que este valor corresponde a un mínimo.

Una vez que el proceso de estimación ha sido concluido se puede hacer uso de la expresión (10) encontrada en la sección anterior, de donde se concluye que la decisión óptima será escoger d_1 : Rechazar H_0, si y sólo si

$$\ln\frac{\theta_0}{\theta_1^*} - \frac{\theta_0}{\theta_1^*} + 1 < \frac{B_1 - B_0}{A}$$

en caso contrario se debe escoger a d_0 : Aceptar H_0 .

Se puede observar que si el miembro izquierdo de esta expresión se ve como una función de θ_1^*, esta función es siempre menor o igual a cero. Si θ_1^* es menor que θ_0 y se acerca a cero, la función se hace más negativa; lo mismo sucede cuando el valor de θ_1^* crece si es mayor que θ_0 y es cero cuando es igual a θ_0.

De lo anterior se deduce que si no existiera una predisposición por alguna de las hipótesis, $B_1 - B_0 = 0$, invariablemente sería rechazada la hipótesis nula sin importar la evidencia en z.

Para evitar tener "resuelto" el problema de antemano, es necesario darle un peso mayor a la hipótesis nula, es decir, es necesario que nuestra función de utilidad refleje una predisposición por la hipótesis nula. Esto se logra pidiendo que B_1 sea menor que B_0. Esta situación debe ser tomada muy en cuenta en el momento de hacer la asignación de utilidades.

Resulta razonable pedir que $B_1 > B_0$ ya que un conjunto de valores más específico en la hipótesis nula indica un mayor interés por un modelo más sencillo y por lo tanto es de esperarse que este modelo reporte una mayor utilidad que la reportada por el modelo más general de la hipótesis alternativa.

2.2.3 H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta = \theta_1$, $\omega = \omega_1$

A diferencia de los dos casos anteriores ω ya no es parámetro de ruido sino que es parte del parámetro de interés (θ,ω), este es el único contraste de hipótesis simples posible. De la expresión en (6) se tiene que la discrepancia entre los modelos especificados por (θ,ω) y (θ_i,ω_i) es

$$\delta(\theta,\omega:\theta_i,\omega_i) - \int P(x|\theta,\omega) \ln \frac{P(x|\theta,\omega)}{P(x|\theta_i,\omega_i)}\, dx$$

tomando nuevamente esta integral como la suma de una sobre $(0,\omega)$ y otra sobre $(\omega,1]$ y si se define a $f(\omega)$ como

$$f(\omega) = \omega \ln \omega + (1-\omega) \ln (1-\omega)$$

después de resolver las integrales empleando cambios de variable se obtiene que

$$\delta(\theta,\omega:\theta_i,\omega_i) = \ln\frac{\theta}{\theta_i} + \frac{\theta_i-\theta}{\theta} - (\theta_i-1)\left[f(\omega) - \omega\ln\frac{\omega_i}{1-\omega_i}\right] + (\theta_i-1)\ln(1-\omega_i) \tag{12}$$

Esta discrepancia obviamente depende de ω, pero si se compara con la obtenida en los casos anteriores se observa que los dos primeros términos corresponden a la discrepancia cuando ω es parámetro de ruido, el tercer término es la parte de la discrepancia que depende únicamente de ω y el término restante no dependen de θ ni de ω.

Si se calcula la diferencia de discrepancias, se agrupan términos y se definen las constantes G y H como

$$G = \ln\left[\left(\frac{\omega_1}{1-\omega_1}\right)^{\theta_1-1}\left(\frac{1-\omega_0}{\omega_0}\right)^{\theta_0-1}\right] \quad , \quad H = \ln\left[\frac{(1-\omega_1)^{\theta_1-1}}{(1-\omega_0)^{\theta_0-1}}\right] \tag{13}$$

es inmediato que

$$\delta(\theta,\omega:\theta_0,\omega_0,\theta_1,\omega_1) = \ln\frac{\theta_0}{\theta_1} + \frac{\theta_1-\theta_0}{\theta} - (\theta_1-\theta_0)f(\omega) + \omega G + H$$

Sustituyendo esta expresión en (3) para encontrar el valor esperado y distribuyendo se obtiene directamente que

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_0,\omega_0,\theta_1,\omega_1)] = \ln\frac{\theta_0}{\theta_1} + E_{\theta,\omega|z}\left(\frac{\theta_1-\theta_0}{\theta}\right) + E_{\theta,\omega|z}[\,\omega G - (\theta_1-\theta_0)f(\omega)\,] + H$$

realizando la integración con respecto a θ teniendo en cuenta que el último término no depende de θ y reagrupando términos se obtiene

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_0,\omega_0,\theta_1,\omega_1)] = \ln\frac{\theta_0}{\theta_1} + E_{\omega|z}\left\{(\theta_1-\theta_0)\left[\frac{nt(\omega)}{n-1} - f(\omega)\right] + \omega G\right\} + H$$

de donde se concluye que se escogerá d_1 : Rechazar H_0 , si y sólo si

$$\ln\frac{\theta_0}{\theta_1} + E_{\omega|z}\left\{(\theta_1-\theta_0)\left[\frac{nt(\omega)}{n-1} - f(\omega)\right] + \omega G\right\} + H < \frac{B_1 - B_0}{A} \tag{14}$$

con G y H dados por (13). En caso contrario se debe escoger a d_0 : Aceptar H_0.

2.2.4 H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta \neq \theta_0$, $\omega \neq \omega_0$

En este caso no hay parámetros de ruido. Para resolver este contraste primeramente se tiene que llevar a cabo la estimación de ambos parámetros dentro del subconjunto determinado por la hipótesis alternativa, en otras palabras, hay que encontrar los valores θ_1^* y ω_1^*.

De la sección anterior se encontró que la discrepancia entre θ,ω y θ_1,ω_1 está dada por (12), ahora se va a encontrar su valor esperado. Al realizar la integración con respecto a θ para la parte que depende exclusivamente de este parámetro se tiene que

$$E_{\theta,\omega|z}\left[\ln\theta + \frac{\theta_1}{\theta}\right] = E_{\omega|z}\left\{\Psi(n) - \ln[nt(\omega)] + \frac{nt(\omega)}{n-1}\theta_1\right\}$$

y finalmente se obtiene que el valor esperado de la discrepancia es

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_1,\omega_1)] = E_{\omega|z}\left\{\frac{nt(\omega)}{n-1}\theta_1 - \ln t(\omega) + \psi(\theta_1-1)\left[\omega\ln\frac{\omega_1}{1-\omega_1} + \ln(1-\omega_1) - f(\omega)\right]\right\} - \ln\theta_1 + R \tag{15}$$

con $R = \Psi(n) - \ln n - 1$. Para determinar los valores θ_1^* y ω_1^* que hacen mínima a la expresión anterior se siguió el método de derivación análogo al empleado en la sección 2.2.2 para encontrar los puntos críticos de (11), con la particularidad de que ahora se trata de una función real de dos variables θ_1 y ω_1. Despúes de usar dicho método se obtiene que estos valores son

$$\omega_1^* = E_{\omega|z}(\omega) \qquad \theta_1^* = \left\{ E_{\omega|z}\left[\frac{nt(\omega)}{n-1} - f(\omega)\right] + f(\omega_1^*) \right\}^{-1} \tag{16}$$

Para determinar si estos valores corresponden a un mínimo de (15), al aplicar el criterio de las segundas derivadas y definiendo

$$D^* = \frac{1 - \theta_1^*}{\theta_1^{*2}\omega_1^*(1-\omega_1^*)}$$

se llega a que los valores en (16) coorresponden a un mínimo de (15) siempre y cuando D^* sea mayor que cero, lo cuál se cumple cuando $\theta_1^* < 1$ ya que el denominador es mayor que cero; y corresponderán a un punto silla si D^* es menor que cero, que se cumple cuando $\theta_1^* > 1$, en este caso se necesita de algún otro método para encontrar el mínimo, si es que éste existe. Sólo faltaría comentar que cuando D^* es igual a cero, que se cumple cuando $\theta_1^* = 1$, se trata de los puntos llamados puntos degenerados en cuyo caso no hay criterio por lo que se debe recurrir a algún otro método, sin embargo hay que tener en mente que la distribución Alfa es independiente de ω para este valor y no tendría sentido un contraste como el tratado en esta sección.

Una vez que se ha logrado encontrar los valores θ_1^* y ω_1^*, y empleando la expresión (14) podemos concluir que se escogerá la decisión d_1 : Rechazar H_0 , si y sólo si se cumple que

$$\ln\frac{\theta_0}{\theta_1^*} + E_{\omega|z}\left\{(\theta_1^* - \theta_0)\left[\frac{nt(\omega)}{n-1} - f(\omega)\right] + \omega G\right\} + H < \frac{B_1 - B_0}{A}$$

en donde las constantes G y H están dadas por

$$G = \ln\left[\left(\frac{\omega_1^*}{1-\omega_1^*}\right)^{\theta_1^*-1}\left(\frac{1-\omega_0}{\omega_0}\right)^{\theta_0-1}\right] \qquad H = \ln\left[\frac{(1-\omega_1^*)^{\theta_1^*-1}}{(1-\omega_0)^{\theta_0-1}}\right]$$

en caso contrario se escogerá d_0 : Aceptar H_0 .

2.2.5 H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta = \theta_1$, $\omega \neq \omega_0$

Para resolver este contraste se necesita realizar únicamente la estimación de ω sobre el subespacio especificado por la hipótesis alternativa, es decir, se va a buscar el valor ω_1^* sobre este espacio. En la sección anterior se obtuvo la expresión para el valor esperado de la discrepancia (15). Viendo esta expresión como una función de ω_1, para encontrar sus puntos críticos hay que derivar la función con respecto a ω_1 y posteriormente resolver la ecuación resultante para obtener a ω_1^*.

Después de realizar los pasos anteriores se llega a que ω_1^* está dado por

$$\omega_1^* = E_{\omega|z}(\omega) \tag{17}$$

Ahora, para identificar si en este punto corresponde a un mínimo de (15) se va a utilizar el criterio de la segunda derivada. Al tomar la segunda derivada y evaluarla en ω_1^* se obtiene un valor que se va a denotar por D^* y que es de la forma

$$D^* = \frac{1-\theta_1}{\omega_1^*(1-\omega_1^*)}$$

de donde se concluye que ω_1^* corresponde a un mínimo si D^* es mayor que cero, que se cumple cuando $\theta_1 < 1$. Por otro lado cuando $\theta_1 > 1$, D^* es menor que cero y por lo tanto ω_1^* corresponde a un máximo; para encontrar el mínimo dentro del subconjunto especificado en H_1, si es que este existe, se requiere de algún otro método. Por último está el caso D^* igual a cero, $\theta_1 = 1$, pero en tal caso no tiene sentido plantearse un contraste como el abordado en esta sección.

Ahora vamos a analizar un poco más a fondo lo que pasa cuando $\theta_1 > 1$ (D^* menor que cero). El valor esperado de la discrepancia, (15), visto como una función de ω_1 se puede expresar como

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\theta_1,\omega_1)] = (\theta_1 - 1)\left\{\ln\frac{\omega_1}{1-\omega_1}\, E_{\omega|z}(\omega) + \ln(1-\omega_1)\right\} + r$$

con r una constante.

Claramente el comportamiento de esta función depende exclusivamente del comportamiento de los términos entre las llaves. Si se denota a la suma de estos términos con $g(\omega_1)$ se puede comprobar que conforme ω_1 se aproxima a ω_1^*, el valor de $g(\omega_1)$ se acerca al valor de $f(\omega_1^*)$ y que cuando ω_1 se aproxima a 0 o a 1 la función $g(\omega_1)$ decrece indefinidamente, lo que hace que el valor esperado de la discrepancia también decrezca indefinidamente, esto trae como consecuencia que no exista el mínimo; sin embargo, esta situación se presenta únicamente alrededor de los valores 0 y 1.

Por lo anterior se puede afirmar que si el contraste es tal que el subconjunto especificado por la hipótesis alternativa (sobre el cuál se minimiza a (15)) no tiene ni a cero ni a 1 como límites y $\theta_1 > 1$, entonces el valor esperado permanecerá acotado y por lo tanto existirá el mínimo. Además, el valor ω_1^* en el que alcanza el mínimo será alguno de los límites del subconjunto de valores de ω especificado en H_1.

Cabe mencionar que la restricción en el contraste de la sección anterior para que exista el mínimo cuando θ_1^* es mayor que 1 tiene el mismo orígen que el encontrado en esta sección, el comportamiento de $g(\omega_1)$ cuando ω_1 se aproxima a 0 o a 1.

Concluimos que la solución al contraste cuando $\theta_1 < 1$, consiste en escoger d_1 : Rechazar H_0 como la solución óptima si y sólo si

$$\ln\frac{\theta_0}{\theta_1} + E_{\omega|z}\left\{(\theta_1-\theta_0)\left[\frac{nt(\omega)}{n-1} - f(\omega)\right] + \omega G\right\} + H < \frac{B_1 - B_0}{A}$$

con

$$G = \ln\left[\left(\frac{\omega_1^*}{1-\omega_1^*}\right)^{\theta_1-1}\left(\frac{1-\omega_0}{\omega_0}\right)^{\theta_0-1}\right] \quad , \quad H = \ln\left[\frac{(1-\omega_1^*)^{\theta_1-1}}{(1-\omega_0)^{\theta_0-1}}\right]$$

en donde ω_1^* está dado por (17). En caso contrario, se debe escoger d_0 : Aceptar H_0 como solución del problema.

2.2.6 $H_0: \omega = \omega_0$ vs. $H_1: \omega = \omega_1$

En todos los contrastes analizados en las secciones anteriores tuvieron a θ como el parámetro de interés (o uno de los de interés), mientras que ω se tomó como el parámetro de ruido (o en el mejor de los casos como uno de los parámetros de interés). En el contraste que ahora se va a analizar se va a tomar a ω como el parámetro de interés y a θ como el parámetro de ruido.

De la definición (6) y teniendo cuidado de la forma en que deben tomarse los parámetros, la discrepancia entre los modelos ω y ω_i está dada por

$$\delta(\theta,\omega:\omega_i) = \int P(x|\theta,\omega) \ln \frac{P(x|\theta,\omega)}{P(x|\theta,\omega_i)}\, dx$$

resolviendo esta integral, en la misma forma como se ha hecho en las secciones anteriores, se obtiene que la discrepancia es

$$\delta(\theta,\omega:\omega_i) = (\theta-1)\left[\omega \ln \frac{\omega_i}{\omega} + (1-\omega)\ln \frac{1-\omega_i}{1-\omega}\right] \tag{18}$$

aquí se puede notar una diferencia importante con los contrastes que consideran a ω como el parámetro de ruido ya que para éstos la discrepancia entre el modelo verdadero y el propuesto no depende del parámetro de ruido, mientras que en el contraste que estamos analizando la discrepancia sí depende del parámetro de ruido.

La diferencia de discrepancias de ω con ω_0 y con ω_1 es

$$\delta(\theta,\omega:\omega_0,\omega_1) = (\theta-1)\left[\omega \ln \frac{\omega_1}{\omega_0} + (1-\omega)\ln \frac{1-\omega_1}{1-\omega_0}\right]$$

y como era de esperarse la diferencia de discrepancias también depende de θ. Si ahora se calcula el valor esperado de esta diferencia de discrepancias, al tomar la integral con respecto a θ se tiene que

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\omega_0,\omega_1)] = E_{\omega|z}\left\{\left[\omega \ln\frac{\omega_1}{\omega_0} + (1-\omega)\ln\frac{1-\omega_1}{1-\omega_0}\right]\int_\theta (\theta-1)\, Ga[\theta | n, nt(\omega)]\right\}$$

la integral que aparece en esta ecuación corresponde al valor esperado de $(\theta - 1)$. También se cumple que si Y se distribuye $Ga(y \mid a,b)$, entonces $E_Y(y) = a/b$. Al aplicar los resultados anteriores se obtiene que el valor esperado de la diferencia de discrepancias es

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\omega_0,\omega_1)] = E_{\omega|z}\left\{\left[\omega \ln\frac{\omega_1}{\omega_0} + (1-\omega)\ln\frac{1-\omega_1}{1-\omega_0}\right](t^{-1}(\omega)-1)\right\}$$

por lo que se concluye que se escogerá d_1 : Rechazar H_0 como la solución óptima al contraste si y sólo si

$$E_{\omega|z}\left\{\left[\omega \ln\frac{\omega_1}{\omega_0} + (1-\omega)\ln\frac{1-\omega_1}{1-\omega_0}\right](t^{-1}(\omega)-1)\right\} < \frac{B_1 - B_0}{A} \qquad (19)$$

En caso contrario, se escogerá d_0 : Aceptar H_0 , como solución al contraste.

Al comparar la regla de decisión para el contraste de la sección 2.2.1 con la regla de decisión que se acaba de encontrar se aprecia que ésta última tiene una forma mucho más complicada. Esto se debe a que para este contraste la discrepancia depende del parámetro de ruido, este hecho se refleja en la regla de decisión a través del término en que aparece t (ω). Esto no es sorprendente si se recuerda que al crear a la distribución Alfa se pensó en que fuera matemáticamente manejable en uno de los parámetros, el parámetro θ, el otro parámetro no tuvo tanta atención pues siempre se tomó como el parámetro de ruido.

2.2.7 $\omega = \omega_0$ vs $\omega \neq \omega_0$

En este contraste nuevamente se toma a ω como el parámetro de interés y a θ como el parámetro de ruido. Por la forma que tiene la hipótesis alternativa, se debe realizar primero la estimación de ω en el espacio especificado por esta hipótesis para determinar el valor ω_1^*.

En el contraste anterior se encontró que la discrepancia entre los modelos especificados por ω y ω_1 está dada por (18), por lo que el valor esperado de la discrepancia es

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\omega_1)] = E_{\theta,\omega|z}\left\{(\theta-1)\left[\omega\ln\frac{\omega_1}{\omega}+(1-\omega)\ln\frac{1-\omega_1}{1-\omega}\right]\right\}$$

como solamente el término en paréntesis depende de θ, se observa la misma integral con respecto a θ que la encontrada en la sección anterior. Al efectuar la integración se obtiene

$$E_{\theta,\omega|z}[\delta(\theta,\omega:\omega_1)] = E_{\theta,\omega|z}\left\{\left[\omega\ln\frac{\omega_1}{\omega}+(1-\omega)\ln\frac{1-\omega_1}{1-\omega}\right](t^{-1}(\omega)-1)\right\} \tag{20}$$

Ahora, hay que encontrar ω_1^* que minimice al valor esperado de la discrepancia. Al tomar a (20) como función solamente de ω_1, calcular su derivada, igualar ésta a cero y resolver para ω_1 se llega a que el punto crítico es

$$\omega_1^* = 1 - \frac{E_{\omega|z}[\omega(t^{-1}(\omega)-1)]}{E_{\omega|z}[(t^{-1}(\omega)-1)]} \tag{21}$$

Para determinar bajo qué condiciones este valor corresponde a un mínimo de (20) se va a denotar con D^* a la segunda derivada con respecto a ω_1, evaluada en $\omega_1 = \omega_1^*$. Esto es

$$D^* = \frac{E_{\omega|z}[t^{-1}(\omega)-1]}{(1-\omega_1^*)^2}$$

Por el criterio de la segunda derivada se tiene que ω_1^* corresponde a un mínimo si D^* es mayor que cero, como el denominador en D^* es siempre mayor que cero, esta condición se reduce a pedir que $E_{\omega|z}[\,t^{-1}(\omega) - 1] > 0$. Por otro lado, si D^* es menor que cero, ω_1^* corresponde a un máximo, en tal situación se requiere de algún otro método para encontrar el valor que minimice a (20), siempre y cuando exista el mínimo.

Usando la regla de decisión dada en (19) se concluye que si D^* es menor que cero, la decisión óptima es d_1 : Rechazar H_0 , si y sólo si

$$E_{\omega|z}\left\{\left[\omega \ln\frac{\omega_1^*}{\omega_0} + (1-\omega)\ln\frac{1-\omega_1^*}{1-\omega_0}\right](t^{-1}(\omega) - 1)\right\} < \frac{B_1 - B_0}{A}$$

con ω_1^* dado en (21). En caso contrario se debe escoger d_0 : Aceptar H_0 como la solución óptima.

2.2.8 H_0: $\theta = 1$ vs. H_1: $\theta \neq 1$

El contraste que ahora se va a analizar es un caso particular del contraste estudiado en la sección 2.2.2. Como se apuntó en el capítulo 1, la distribución uniforme corresponde al valor particular $\theta = 1$ dentro de la familia Alfa lo que permite que este contraste se pueda interpretar como el probar uniformidad contra no uniformidad. Esto hace que dicho contraste tenga una importancia singular ya que presenta una forma de abordar los problemas de bondad de ajuste, a continuación se explica en que consiste esta clase de problemas.

De lo visto en el capítulo anterior se puede observar que el haber supuesto que una determinada familia paramétrica es lo suficientemente buena para considerar que alguno de sus elementos es una buena aproximación del modelo verdadero es muy importante en el proceso de toma de decisiones (debido a que la decisión óptima depende de $P(\theta|z)$ y ésta, a su vez, depende de la familia paramétrica empleada a través de $P(x|\theta)$). Pero, ¿ qué sucede si la familia no es lo suficientemente buena ?. Lo que puede suceder es que la decisión óptima encontrada puede resultar inadecuada en virtud del comportamiento verdadero de X.

En este momento puede volverse atractivo para quien está realizando el estudio, si lo juzga conveniente, contar con un procedimiento que le permita determinar si la familia paramétrica (o la distribución de probabilidad) propuesta es compatible con el comportamiento de X, dada una muestra aleatoria z. Esta clase de problemas son conocidos como problemas de bondad de ajuste (el libro de D'agostino & Stephens, 1986; es una excelente referencia). En el caso contínuo este problema se puede plantear como sigue:

Problema de Bondad de Ajuste.

Sea X una variable aleatoria contínua con distribución de probabilidad P(x). Sea $\mathcal{F}$ = { Distribuciones contínuas } y $\mathcal{F}_1 = \mathcal{F} - \mathcal{F}_0$, con $\mathcal{F}_0$ una familia de distribuciones con al menos un elemento, tal que $\mathcal{F}_0 \subset \mathcal{F}$. El problema de bondad de ajuste consiste en probar las hipótesis

$$H_0 : P(x) \in \mathcal{F}_0 \qquad \text{vs.} \qquad H_1 : P(x) \in \mathcal{F}_1$$

dada la muestra aleatoria $z = \{x_1, x_2, .., x_n\}$ de X.

Es bien sabido que si la variable aleatoria contínua X tiene una función de distribución F(x) y se le aplica la "Transformación Integral de Probabilidad" (Probability Integral Transformation, PIT) a las observaciones de la muestra aleatoria $z = \{x_1, x_2, .., x_n\}$, esto es $y_i = F(x_i)$ $(i=1,..n)$; entonces $z' = \{y_1, y_2, .., y_n\}$ es una muestra aleatoria de la distribución uniforme $U(Y|0,1)$, ver por ejemplo Harris(1966).

Una práctica comúnmente empleada por los métodos clásicos para probar bondad de ajuste consiste en transformar los datos en z, usando función de distribución del modelo que se propone como explicación de X, para obtener a z' y posteriormente aplicar algún método dedicado a probar uniformidad (por ejemplo la prueba de Anderson y Darling) a z' .

Si nos restringimos únicamente al caso en que $\mathcal{F}_0$ tiene sólo un elemento (esto sucede cuando se propone una distribución totalmente especificada) y se sigue el esquema anterior, el probar bondad de ajuste se reduce a probar uniformidad. Más aún, con la introducción de una familia paramétrica en la que la función de densidad uniforme se obtiene con el valor particular del parámetro $\theta = \theta_0$, probar bondad de ajuste se reduce a contrastar las hipótesis paramétricas

$$H_0 : \theta = \theta_0 \qquad \text{vs.} \qquad H_1 : \theta \neq \theta_0$$

De esta manera, si se trabaja con la familia Alfa y si se denota por $F_0(x)$ a la función de distribución correspondiente al modelo $P_0(x)$ que se propone como explicación, el problema de decidir si $P_0(x)$ es una buena explicación del comportamiento de X dada la muestra z consiste en contrastar las hipótesis paramétricas

$$H_0: \theta = 1 \quad \text{vs.} \quad H_1: \theta \neq 1 \tag{22}$$

en base a la muestra aleatoria $z' = \{y_1, y_2, \ldots, y_n\}$ obtenida al efectuar la transformación de los datos $y_i = F_0(x_i)$; con $i = 1 \ldots n$.

Por lo anterior, se puede proponer el siguiente procedimiento para probar bondad de ajuste:

a) Obtener una muestra aleatoria $z = \{x_1, x_2, \ldots, x_n\}$ de X.

b) Transformar los datos de la muestra aleatoria z, con la función de distribución $F_0(x)$ correspondiente al modelo probabilístico que se propone, $P_0(x)$, para obtener la muestra aleatoria $z' = \{y_1, y_2, \ldots, y_n\}$; con $y_i = F_0(x_i)$.

c) Estimar a θ, dentro del subconjunto $\Theta_1 = \Theta - \{1\}$, por medio de

$$\theta_1^* = \frac{n-1}{n E_{\omega|z'}[t(\omega)]}$$

d) Decidir que $P_0(x)$ no representa una buena explicación del comportamiento de X, es decir, escoger como decisión óptima al problema de bondad de ajuste (22) a d_1 : Rechazar H_0 , si y sólo si

$$1 - \ln\theta_1^* - \frac{1}{\theta_1^*} < \frac{B_1 - B_0}{A}$$

con θ_1^* dado en el inciso (b) y $t(\omega)$ dado por

$$t(\omega) = \frac{1}{n}\left[\sum_{i\in I(\omega)} \ln\frac{\omega}{y_i} + \sum_{j\in J(\omega)} \ln\frac{1-\omega}{1-y_j}\right]$$

en donde $I(\omega) = \{ i : y_i \leq \omega \}$ y $J(\omega) = \{ j : y_j > \omega \}$.

e) En caso de que no se satisfaga la condición del inciso anterior se debe tomar a $P_0(x)$ como explicación del comportamiento de X.•

Cabe mencionar que el miembro izquierdo de esta expresión es siempre menor o igual que cero, hecho que debe ser tomado en cuenta cuando se realiza la asignación de valores a las constantes que intervienen en el miembro derecho de esta regla de decisión.

2.3 Solución Clásica de 2.2

Una vez que se han presentado y analizado las soluciones bayesianas a los contrastes abordados en la sección 2.2 puede existir la inquietud de conocer, aunque sea de una manera muy superficial, cómo son las soluciones clásicas de algunos de estos contrastes. A continuación se presentan las regiones críticas obtenidas por el Lema de Neymann-Pearson o por el método de razón de verosimilitudes generalizado, según lo amerite el contraste.

Si bien es cierto que no hay una argumentación formal para llevar a cabo una comparación entre soluciones clásicas y bayesianas, y que dicho sea de paso no es la intención, sí se harán algunos comentarios breves sobre ambas.

Para resolver un contraste clásicamente hay que encontrar la región crítica C^*, posteriormente encontrar a la distribución de la función dependiente de z que aparece en C^* bajo la hipótesis nula, tarea que puede resultar nada sencilla y que de hecho no se emprenderá aquí. Por último, con esta distribución se debe encontrar el valor de la constante k_0 especificada en C^* para un nivel de significancia preestablecido.

✔ En el contraste de hipótesis de la sección 2.2.1, H_0: $\theta = \theta_0$ vs. H_1: $\theta = \theta_1$, como ω es desconocido, hay que encontrar el estimador máximo-verosímil, $\hat{\omega}$. Como puede comprobarse, el estimador se obtiene resolviendo la ecuación $t'(\hat{\omega}) = 0$ siempre y cuando cumpla con la condición $t''(\hat{\omega})(\theta_i - 1) > 0$ $(i=0,1)$. Suponiendo $\theta_0 > \theta_1$, la región crítica está dada por

$$C^* = \{ z : t(\hat{\omega}) > k_0 \}$$

y para $\theta_0 < \theta_1$ se obteniene una C^* como la anterior pero con la desigualdad invertida.

Cabe mencionar que no se puede garantizar la existencia del estimador máximo-verosímil. Si no existe dicho estimador, no existe la solución al contraste utilizando este método. Por otra lado, la solución bayesiana siempre existe para este contraste y además no es necesario estimar ω.

✔ En el contraste de la sección 2.2.2, H_0: $\theta = \theta_0$ vs. H_1: $\theta \neq \theta_0$, como ambos parámetros son desconocidos se debe encontrar los estimadores máximo-verosímiles $\hat{\theta}$ y $\hat{\omega}$. Los estimadores se obtienen al resolver simultáneamente las ecuaciones

$$t'(\hat{\omega})(\hat{\theta} - 1) = 0 \qquad y \qquad \hat{\theta} = t^{-1}(\hat{\omega}) \tag{23}$$

siempre y cuando se cumpla alguna de estas dos condiciones: i) que $t(\hat{\omega}) < 1$ y que $\hat{\omega}$ minimice a $t(\omega)$, o ii) que $t(\hat{\omega}) > 1$ y que $\hat{\omega}$ maximice a $t(\omega)$. En caso de que los estimadores máximo-verosímiles existan, la región crítica está dada por

$$C^* = \{ z : \ln \hat{\theta} + \frac{\theta_0 - \hat{\theta}}{\hat{\theta}} > K_0 \}$$

Al igual que en el contraste anterior no se puede asegurar la existencia de los estimadores máximo-verosímiles y por tanto que exista la solución. Además, se tiene el problema de encontrar la distribución de la función de z que parece ser bastante compleja. Esto último no es tan preocupante ya que se puede recurrir a alguna aproximación de dicha distribución. Se puede notar que la regla de decisión bayesiana para este contraste tiene la misma forma que la región crítica que se acaba de encontrar, con la única diferencia en la forma de estimar θ.

✔ En el contraste en 2.2.3, H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta = \theta_1$, $\omega = \omega_1$, tenemos que las dos hipótesis son simples y después de aplicar el Lema de Neymann-Pearson se obtiene que la región crítica es

$$C^* = \{ z : (\theta_0 - 1)\, t(\omega_0) - (\theta_1 - 1)\, t(\omega_1) > k_0 \}$$

Aquí hay una dificultad adicional a las existentes en los contrastes anteriores para obtener la distribución de la función de la muestra (miembro izquierdo de la desigualdad en la región crítica) bajo H_0.

Se puede probar que si θ es conocido, la función $t(\omega)$ tiene una distribución $Ga(t(\omega) \mid n, n\theta)$, pero no sabemos como se distribuye $t(\omega_1)$; por si esto fuera poco, no se puede usar la aproximación $-2 \ln \Lambda \sim \chi^2_{(m)}$ para n grande (donde Λ es el cociente de las verosimilitudes correspondientes a las dos hipótesis y $\chi^2_{(m)}$ es la distribución χ-cuadrada con m grados de libertad) ya que ésta se aplica cuando se usan los estimadores máximo-verosímiles, este no es el caso por tratarse de un contraste de hipótesis simples.

✔ Para el contraste de la sección 2.2.4, H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta \neq \theta_0$, $\omega \neq \omega_0$, hay que usar los estimadores máximo-verosímiles de ambos parámetros. Estos son los mismos que los encontrados en el segundo contraste y se obtienen de (23) bajo las mismas condiciones que antes. La región crítica es

$$C^* = \{ z : t(\hat{\omega}) + (\theta_0 - 1)\, t(\omega_0) + \ln\hat{\theta} > K_0 \}$$

✔ Para el contraste de la sección 2.2.5, H_0: $\theta = \theta_0$, $\omega = \omega_0$ vs. H_1: $\theta = \theta_1$, $\omega \neq \omega_0$, hay que estimar ω únicamente. El estimador maximo-verosímil para ω está dado por la ecuación

$$t'(\hat{\omega}) = 0$$

con la condición t'' ($\hat{\omega}$) > 0, es decir, tal que $\hat{\omega}$ maximice a la función t (ω). Por lo tanto la región crítica es

$$C^{*} = \{ z : (\theta_1 - 1)\,t(\hat{\omega}) - (\theta_0 - 1)\,t(\omega_0) < k_0 \}$$

✔ Para el contraste en 2.2.6, H_0: $\omega = \omega_0$ vs. H_1: $\omega = \omega_1$, el parámetro θ es desconocido por lo que hay que trabajar con su estimador. El estimador máximo-verosímil para θ está dado por

$$\hat{\theta} = t^{-1}(\omega_1)$$

A diferencia de los casos anteriores, aquí no se tiene ninguna condición sobre el estimador para que efectivamente corresponda al punto donde se maximiza a la verosimilitud. La región crítica es

$$C^{*} = \{ z : \ln \frac{t(\omega_1)}{t(\omega_0)} + t(\omega_0) - t(\omega_1) < k_0 \}$$

✔ Para el contraste en 2.2.7, $\omega = \omega_0$ vs $\omega \neq \omega_0$, los estimadores nuevamente se encuentran a partir de (23) y por lo tanto la región crítica es

$$C^{*} = \{ z : \ln \frac{t(\hat{\omega})}{t(\omega_0)} + t(\omega_0) - t(\hat{\omega}) < k_0 \}$$

Ahora que se cuenta con las regiones críticas el siguiente paso es encontrar la distribución de la función de la muestra para fijar a k_0.

Recordando que la función t (ω) se distribuye Ga(t (ω) | n , n θ) si θ es conocido, y como en los contrastes 2.2.6 y 2.2.7 no se conoce el valor de θ, no se puede conocer completamente la distribución de t(ω) y esto, a su vez, hace que no se pueda determinar la distribución de la función de la muestra bajo la hipótesis nula y por lo tanto que no se pueda encontrar k_0 al nivel de significancia preestablecido.

Por otra parte, el procedimiento bayesiano sí permitió dar una solución a todos los contrastes analizados, al menos bajo ciertas condiciones, aún en aquellos contrastes en los que los métodos clásicos usados en esta sección no lo permitieron.

Capítulo 3

Simulaciones

Capítulo 3

Simulaciones

3.1 Procedimiento

Con el propósito de tener una idea más clara del comportamiento de las soluciones bayesianas de algunos de los contrastes analizados en el capítulo anterior, se llevó a cabo un trabajo de simulación que consistió básicamente en la determinación de los valores $(B_1 - B_0) / A$ para los cuáles el porcentaje de muestras aleatorias que llevaron a rechazar la hipótesis nula fue igual a un conjunto de porcentajes preestablecidos. Las muestras aleatorias Alfa generadas eran compatibles con la hipótesis nula correspondiente.

Por comodidad, se usará el término "cuantil empírico", $\delta^*_{(\alpha)}$, para denotar al valor particular de $(B_1 - B_0) / A$ tal que el α% de las muestras aleatorias generadas condujeron a rechazar la hipótesis nula. Hablamos de los cuantiles empíricos porque éstos fueron obtenidos a través de un proceso de simulación y no a través del modelo explícito de la distribución de la función de la muestra que aparece en la región crítica.

Los valores $\delta^*_{(\alpha)}$ que se acaban de definir pueden ser interpretados desde el punto de vista clásico como los cuantiles empíricos para los cuáles la probabilidad de rechazar la hipótesis nula, cuando es verdadera, es igual al nivel de significancia $\alpha' = \alpha / 100$.

Vale la pena enfatizar que desde nuestra perspectiva, no estamos relacionando a α con el nivel de significancia clásico. Para nosotros, el porcentaje α% representa simplemente el porcentaje de muestras que llevaron a rechazar la hipótesis nula.

A continuación se describe con mayor detalle el procedimiento de simulación realizado para cada uno de los tres contrastes abordados en este capítulo, para diferentes tamaños de muestra.

a) Se generaron 10,000 muestras aleatorias Alfa de tamaño n, con parámetros θ y ω compatibles con la hipótesis nula.

La generación de las muestras aleatorias Alfa se hizo usando el método de Montecarlo. Se empleó el algoritmo AS 183 propuesto en Applied Statistics, vol. 31, como el generador de muestras aleatorias uniformes y se transformaron cada una de las observaciones uniformes de las muestras con la función inversa de la función de distribución que aparece en (8).

b) Se calculó el valor esperado de la diferencia de discrepancias correspondiente a cada una de las muestras.

Debido a la dificultad que hay para calcular analíticamente las integrales involucradas, para su evaluación se empleó el método de integración de Romberg (Burden, 1978).

c) Se ordenaron de menor a mayor los 10,000 valores esperados obtenidos en (b).

d) Se encontraron los cuantiles empíricos, $\delta^*_{(\alpha)}$, para los valores α = 1, 5, 10, 20, 50, 80, 90, 95, 99.

De los valores esperados ordenados en (c), se tomó aquel valor que cumplió con estar en la posición $\alpha * 10{,}000$ como el cuantil $\delta^*_{(\alpha)}$; por ejemplo, el cuantil $\delta^*_{(90)}$ correspondió al valor esperado que estaba en la posición 9,000.

Ahora se van a presentar los resultados de las simulaciones realizadas para algunos de los contrastes vistos en el capítulo anterior.

3.2 $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$

Si bien es cierto que en este contraste ω es el parámetro de ruido, es necesario escoger un valor particular para él para poder realizar la generación de las muestras aleatorias. El valor que se tomó para este propósito fue $\omega = 0.5$.

A continuación se presentan en la Tabla 1 los cuantiles empíricos obtenidos en las simulaciones realizadas para diferentes combinaciones de θ_0 y θ_1 y para dos tamaños de muestra, n = 20 y n = 50. Cabe mencionar que también se hicieron simulaciones para los valores ω = 0.3 y 0.9, en las que se obtuvieron cuantiles empíricos similares a los que aquí se reportan.

T a b l a 1
Cuantiles Empíricos, $\delta^*_{(\alpha)}$,
para $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$

n = 20

	α %: 1	5	10	20	50	80	90	95	99
θ_0 = 0.5									
θ_1 = 0.6	-.066	-.044	-.030	-.014	.024	.058	.093	.114	.152
0.8	-.137	-.055	-.006	.041	.151	.278	.358	.423	.540
3.0	.889	1.639	2.012	2.491	3.419	4.464	5.013	5.430	6.544
7.0	4.884	6.225	7.231	8.512	11.017	13.702	15.387	16.781	19.670
25.0	21.462	29.265	32.682	37.442	46.547	57.303	63.365	68.426	78.252
θ_0 = 0.7									
θ_1 = 0.6	-.075	-.058	-.044	-.025	.005	.032	.045	.056	.073
0.8	-.051	-.036	-.027	-.012	.015	.045	.062	.081	.106
3.0	.418	.759	1.024	1.288	1.958	2.615	3.077	3.397	4.055
7.0	2.700	3.822	4.509	5.274	7.054	8.927	9.952	11.063	12.962
25.0	15.972	19.014	22.428	25.569	32.139	39.510	43.450	47.085	53.952
θ_0 = 1.0									
θ_1 = 0.6	-.199	-.079	-.028	.013	.106	.181	.217	.249	.263
0.8	-.116	-.074	-.052	-.026	.018	.059	.075	.092	.117
3.0	-.034	.245	.391	.558	.821	1.366	1.836	1.833	2.293
7.0	1.458	2.125	2.540	3.136	4.226	5.473	6.217	6.813	8.178
25.0	10.283	13.184	14.658	16.514	21.089	25.516	29.423	31.359	37.271
θ_0 = 5.0									
θ_1 = 0.6	.826	.833	.912	1.013	1.187	1.372	1.443	1.508	1.601
0.8	.423	.562	.664	.782	.964	1.120	1.199	1.258	1.344
3.0	-.168	-.089	-.049	.003	.095	.173	.207	.233	.291
7.0	-.109	-.060	-.032	.003	.080	.168	.219	.268	.363
25.0	.764	1.113	1.416	1.755	2.536	3.442	3.851	4.207	5.335
θ_0 = 10.0									
θ_1 = 0.6	1.239	1.428	1.519	1.637	1.840	2.006	2.082	2.146	2.279
0.8	1.007	1.193	1.292	1.382	1.576	1.756	1.829	1.890	1.994
3.0	.001	.180	.253	.342	.488	.602	.663	.709	.798
7.0	-.167	-.088	-.055	-.018	.050	.104	.126	.144	.171
25.0	-.031	.130	.235	.337	.641	.875	1.132	1.314	1.671
θ_0 = 50.0									
θ_1 = 0.6	2.692	2.863	3.070	3.192	3.410	3.579	3.666	3.733	3.821
0.8	2.499	2.676	2.775	2.892	3.111	3.304	3.389	3.457	3.564
3.0	1.198	1.417	1.521	1.644	1.829	2.004	2.082	2.165	2.295
7.0	.508	.714	.806	.907	1.090	1.250	1.326	1.371	1.490
25.0	-.148	-.046	.015	.078	.174	.271	.307	.342	.410

n = 50

	1	5	10	20	50	80	90	95	99
θ_0 = 0.5									
θ_1 = 0.6	-.044	-.027	-.018	-.005	.019	.044	.056	.069	.088
0.8	-.039	.010	.036	.060	.134	.211	.252	.286	.344
3.0	1.762	2.160	2.384	2.852	3.245	3.886	4.228	4.551	5.090
7.0	6.790	7.724	8.261	9.011	10.500	12.186	13.019	13.698	15.092
25.0	31.816	35.113	37.238	40.108	45.704	52.311	55.757	58.686	63.408
θ_0 = 0.7									
θ_1 = 0.6	-.040	-.029	-.020	-.010	.010	.028	.034	.041	.056
0.8	-.032	-.021	-.014	-.005	.012	.029	.040	.048	.064
3.0	.857	1.160	1.315	1.472	1.902	2.311	2.526	2.738	3.202
7.0	3.847	4.702	5.168	5.675	6.862	8.005	8.602	9.147	10.043
25.0	20.578	23.714	25.596	27.965	31.867	36.563	38.880	41.097	44.433
θ_0 = 1.0									
θ_1 = 0.6	-.050	-.008	.012	.053	.104	.153	.174	.199	.230
0.8	-.060	-.032	-.020	-.008	.021	.044	.056	.065	.078
3.0	.319	.479	.582	.876	.912	1.173	1.314	1.458	1.816
7.0	1.920	2.681	3.031	3.388	4.072	4.799	5.243	5.846	6.406
25.0	12.641	15.398	16.847	18.230	21.029	24.032	26.057	27.704	30.750
θ_0 = 5.0									
θ_1 = 0.6	.875	1.002	1.057	1.115	1.244	1.337	1.387	1.428	1.519
0.8	.675	.765	.809	.804	.972	1.080	1.127	1.158	1.208
3.0	-.050	-.001	.025	.057	.108	.155	.179	.198	.228
7.0	-.046	-.017	-.002	.024	.071	.125	.156	.161	.223
25.0	1.187	1.552	1.756	1.988	2.400	2.923	3.205	3.390	3.972
θ_0 = 10.0									
θ_1 = 0.6	1.494	1.616	1.681	1.738	1.854	1.971	2.024	2.074	2.152
0.8	1.261	1.362	1.413	1.481	1.598	1.694	1.750	1.794	1.842
3.0	.242	.305	.361	.414	.496	.581	.617	.650	.700
7.0	-.058	-.019	-.000	.013	.051	.091	.107	.120	.146
25.0	.177	.291	.350	.435	.620	.816	.908	.989	1.157
θ_0 = 50.0									
θ_1 = 0.6	3.057	3.153	3.223	3.280	3.412	3.528	3.584	3.639	3.704
0.8	2.758	2.886	2.949	3.012	3.138	3.258	3.322	3.382	3.420
3.0	1.497	1.625	1.682	1.743	1.859	1.969	2.013	2.058	2.124
7.0	.780	.871	.932	.998	1.103	1.194	1.243	1.279	1.332
25.0	.003	.060	.085	.121	.189	.250	.280	.300	.342

El cuantil empírico $\delta^*_{(\alpha)}$ denota al valor $(B_1 - B_0) / A$ para el cuál el porcentaje de muestras aleatorias que llevaron a rechazar la hipótesis nula fue α %, dado que ésta fue verdadera.

De la tabla anterior se puede observar que si se tomara, por ejemplo, el contraste H_0: θ_0 = .5 vs. H_1: θ_1 = .6, con n = 20 y se hubieran escogido las constantes A, B_0 y B_1 de forma tal que $(B_1 - B_0)/A = \delta^*_{(95)} = .114$, entonces la muestra aleatoria habría llevado erróneamente a rechazar la hipótesis nula, cuando ésta fue verdadera, en el 95% de los casos. Por otro lado, si $(B_1 - B_0) / A = \delta^*_{(1)} = -.066$, la hipótesis nula se habría rechazado en el 1% de los casos; o equivalentemente, se habría aceptado correctamente la hipótesis nula en el 99% de los casos.

También se calculó el promedio aritmético, Prom, de los 10,000 valores esperados obtenidos en el inciso (b) del proceso de simulación. Estos promedios se presentan en la Tabla 2 para un conjunto de contrastes particulares.

Tabla 2
Promedio, Prom,
para $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$

		Prom	
θ_0	θ_1	n = 20	n = 50
0.5	0.6	0.028	0.019
	0.8	0.163	0.140
	3.0	3.481	3.277
	7.0	11.164	6.790
	25.0	47.533	31.816
0.7	0.6	0.003	0.009
	0.8	0.017	0.012
	3.0	1.992	1.911
	7.0	7.170	6.857
	25.0	32.670	32.134
1.0	0.6	0.097	0.102
	0.8	0.016	0.019
	3.0	0.972	0.930
	7.0	4.335	4.104
	25.0	21.417	21.212
5.0	0.6	1.189	1.227
	0.8	0.946	0.970
	3.0	0.087	0.104
	7.0	0.088	0.075
	25.0	2.615	2.449
10.0	0.6	1.820	1.852
	0.8	1.565	1.589
	3.0	0.468	0.491
	7.0	0.041	0.051
	25.0	0.671	0.628
50.0	0.6	3.382	3.405
	0.8	3.096	3.132
	3.0	1.819	1.853
	7.0	1.071	1.092
	25.0	0.169	0.185

La cantidad Prom representa el promedio aritmético de los 10,000 valores esperados de la diferencia de discrepancias, obtenidos en la simulación, para valores particulares de θ_0 y θ_1.

De la tabla anterior se puede observar en todos los contrastes considerados que si no hay predisposición por alguna de las hipótesis contrastadas ($(B_1 - B_0) / A = 0$), es más probable que se hubiera aceptado la hipótesis nula como la solución del contraste.

3.3 $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

Como se recordará de la sección 2.2.2, en esta clase de contrastes primero hay que estimar θ con θ_1^* usando el método descrito en 2.2.2 y posteriormente usar la regla de decisión correspondiente. De esta forma, se obtuvo una estimación θ_1^* para cada una de las 10,000 muestras generadas en el proceso de simulación. Las simulaciones fueron hechas para tamaños de muestra n = 20 y n = 40. Los valores para θ_0 fueron escogidos de modo tal que se tomara en cuenta un rango amplio de posibilidades.

Con la intención de tener una idea de qué tan buenas fueron las estimaciones obtenidas se calcularon la estimación media, $\bar{\theta}_1^*$, y la varianza de las estimaciones, Var (θ_1^*), para cada uno de los contrastes considerados en esta sección. Estos resultados se muestran en la Tabla 3.

T a b l a 3
Estimación media, $\bar{\theta}_1^*$, y Varianza, Var(θ_1^*),
para $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$

	n = 20		n = 40	
θ_0	$\bar{\theta}_1^*$	Var(θ_1^*)	$\bar{\theta}_1^*$	Var(θ_1^*)
0.5	0.492	0.002	0.505	0.006
0.7	0.707	0.030	0.697	0.017
1.0	1.010	0.067	1.010	0.023
5.0	5.013	1.562	4.978	0.504
10.0	10.009	5.780	9.942	2.120
50.0	50.341	142.682	49.997	51.699

La Estimación media, $\bar{\theta}_1^*$, representa el promedio aritmético y Var(θ_1^*) representa la varianza de las 10,000 estimaciones para cada θ_0.

De la Tabla 3 se puede observar que el método de estimación proporcionó valores de θ_1^* muy cercanos al verdadero valor del parámetro θ en todo el rango considerado para θ_0. Sin embargo, se puede notar que la varianza de las estimaciones crece rápidamente conforme θ_0 crece. Por otra parte, y como era de esperarse, la varianza de las estimaciones disminuye al aumentar el tamaño de la muestra.

Después de cada estimación se calculó el valor esperado de la diferencia de discrepancias para cada una de las muestras y con ellos se determinaron los cuantiles empíricos para cada contraste. A continuación se presentan en la Tabla 4 los cuantiles empíricos obtenidos en las simulaciones.

Tabla 4

Cuantiles Empíricos, $\delta^*_{(\alpha)}$,

para $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$

	n = 20									n = 40								
θ_0 α%:	1	5	10	20	50	80	90	95	99	1	5	10	20	50	80	90	95	99
0.5	-.1684	-.1089	-.0743	-.0466	-.0129	-.0016	-.0005	-.0001	-7E-6	-.0706	-.0412	-.0266	-.0174	-.0050	-.0007	-.0001	-3E-5	-1E-6
0.7	-.1765	-.1004	-.0726	-.0453	-.0124	-.0017	-.0005	-.0002	-1E-5	-.0687	-.0426	-.0292	-.0171	-.0052	-.0008	-.0002	-8E-5	-4E-6
1.0	-.1681	-.1150	-.0860	-.0515	-.0136	-.0024	-.0007	-.0003	-9E-6	-.0771	-.0446	-.0303	-.0183	-.0046	-.0007	-.0002	-5E-5	-2E-6
5.0	-.2580	-.1341	-.0822	-.0451	-.0117	-.0014	-.0003	-.0001	-4E-6	-.0804	-.0378	-.0262	-.0160	-.0048	-.0007	-.0001	-4E-5	-2E-6
10.0	-.1928	-.1041	-.0717	-.0465	-.0135	-.0016	-.0003	-.0001	-3E-6	-.0729	-.0424	-.0292	-.0165	-.0048	-.0008	-.0001	-5E-5	-2E-6
50.0	-.2045	-.1120	-.0761	-.0466	-.0126	-.0020	-.0005	-.0001	-7E-6	-.0808	-.0398	-.0276	-.0160	-.0045	-.0007	-.0002	-6E-5	-8E-6

El cuantil empírico $\delta^*_{(\alpha)}$ denota al valor $(B_1 - B_0) / A$ para el cuál el porcentaje de muestras aleatorias que llevaron a rechazar la hipótesis nula fue α %, dado que ésta fue verdadera.

Si en esta ocasión se hubieran tomado las constantes de forma tal que $(B_1 - B_0) /A$ fuera mayor o igual que cero, la regla de decisión siempre hubiera llevado a rechazar la hipótesis nula sin importar de qué contraste (de los considerados en esta sección) se estuviera estudiando. Por otra parte, si se considera, por ejemplo, el contraste $H_0: \theta = 1$ vs. $H_1: \theta \neq 1$, contraste de uniformidad contra no uniformidad, con $n = 40$ y si se escogen las constantes de manera que $(B_1 - B_0) / A$ sea igual a $\delta^*_{(1)} = -.0771$; entonces, en el 1% de los casos se hubiera rechazado la hipótesis nula, o en otras palabras, en el 99% de los casos se hubiera aceptado la hipótesis de normalidad sabiendo que es verdadera.

Si se calcula nuevamente la media del valor esperado de la diferencia de discrepancias, Prom, para cada uno de los contraste se obtiene la Tabla 5.

Tabla 5
Promedio, Prom,
para $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$

	Prom	
θ_0	n = 20	n = 40
0.5	-.0273	-.0106
0.7	-.0273	-.0107
1.0	-.0302	-.0110
5.0	-.0297	-.0100
10.0	-.0282	-.0106
50.0	-.0285	-.0100

La cantidad Prom representa el promedio aritmético de los 10,000 valores esperados de la diferencia de discrepancias, para valores particulares de θ_0.

Se puede observar de la Tabla 5 que si se hubieran escogido las constantes de manera que $(B_1 - B_0) / A \geq -.0285$, para n = 20, habría sido más probable que la muestra aleatoria hubiera llevado a aceptar la hipótesis nula en los contrastes con θ_0 = .5, .7, 10, 50. Por otro lado, para ese mismo valor, -.0285, la hipótesis nula probablemente hubría sido rechazada para aquellos contrastes con θ_0 = 1, 5.

3.4 $H_0 : \theta = \theta_0 , \omega = \omega_0$ vs. $H_1 : \theta = \theta_1 , \omega = \omega_1$

Ahora toca el turno a la única clase de contrastes en los que ambas hipótesis son simples.

Las hipótesis nula y alternativa escogidas en estos contrastes corresponden a distribuciones que son parecidas entre sí, ya sea porque el punto en el que las distribuciónes alcanzan su valor extremo es el mismo (primero y segundo contrastes) o porque los máximos de las distribuciones son iguales o muy cercanos entre sí.

A continuación se presenta las cuantiles empíricos obtenidos en estas simulaciones para n = 20 y n = 50.

T a b l a 6
Cuantiles Empíricos, $\delta^*_{(\alpha)}$,
para $H_0: \theta=\theta_0, \omega=\omega_0$ vs $H_1: \theta=\theta_1, \omega=\omega_1$

					n = 20									n = 50							
θ_0	ω_0	θ_1	ω_1	α%:	1	5	10	20	50	80	90	95	99	1	5	10	20	50	80	90	95	99
1.5	0.5	1.5	0.5		-.10	-.07	-.06	-.04	-.01	.01	.03	.05	.06	-.07	-.05	-.04	-.02	0	.01	.02	.03	.04
3.0	0.5	0.5	0.5		.57	.66	.73	.79	.96	1.12	1.19	1.24	1.32	.65	.754	.81	.85	.96	1.06	1.11	1.15	1.24
20.0	0.4	20.0	0.5		-.48	-.46	-.45	-.45	-.40	-.34	-.31	-.28	-.23	-.46	-.46	-.45	-.45	-.38	-.34	-.34	-.33	-.33
0.7	0.3	0.7	0.8		-.13	-.06	-.06	-.04	-.01	.03	.06	.08	.10	-.10	-.06	-.04	-.03	.02	.06	.09	.10	.12
5.0	0.5	0.2	0.1		2.95	3.22	3.31	3.43	3.64	3.82	3.91	3.99	4.10	3.31	3.41	3.46	3.53	3.66	3.77	3.82	3.86	3.96
1.2	0.4	0.8	0.7		-.06	0	.03	.06	.10	.13	.15	.17	.22	.01	.05	.06	.08	.10	.12	.14	.16	.20
10.0	0.2	0.2	0.9		3.36	3.57	3.66	3.76	4.00	4.19	4.27	4.34	4.45	3.64	3.76	3.81	3.86	4.00	4.12	4.18	4.22	4.30

El cuantil empírico $\delta^*_{(\alpha)}$ denota al valor $(B_1 - B_0) / A$ para el cuál el porcentaje de muestras aleatorias que llevaron a rechazar la hipótesis nula fue α %, dado que ésta fue verdadera.

Se pueden identificar en la Tabla 6 tanto valores menores como mayores que cero, confirmando que para esta clase de contrastes no es necesaria la predisposición por alguna de las hipótesis.

En esta ocasión también se calcularon los promedios de los valores esperados de la diferencia de discrepancias, Prom, de las 10,000 muestras para cada uno de los contrastes.

Los promedios para cada contraste aparecen en la Tabla 7 de la siguiente página.

Tabla 7
Promedio, Prom,
para $H_0: \theta = \theta_0, \omega = \omega_0$ vs $H_1: \theta = \theta_1, \omega = \omega_1$

				Prom	
θ_0	ω_0	θ_1	ω_1	n = 20	n = 50
1.5	0.5	1.8	0.5	-.014	-.007
3.0	0.5	0.5	0.5	.958	.958
20.0	0.4	20.0	0.5	-.389	-.390
0.7	0.3	0.7	0.6	-.005	.018
5.0	0.9	0.2	0.1	3.623	3.649
1.2	0.4	0.9	0.7	.092	.099
10.0	0.2	0.2	0.9	3.984	3.997

La cantidad Prom representa el promedio aritmético de los 10,000 valores esperados de la diferencia de discrepancias, para valores particulares de θ_0, ω_0, θ_1 y ω_1.

Se puede ver de la Tabla 7 que para el tamaño de muestra n = 20 y en el caso en que no hay predisposición por escoger alguna de las hipótesis, $(B_1 - B_0) / A = 0$, habría sido más probable que la muestra aleatoria genrada hubiera conducido a rechazar erróneamente la hipótesis nula en los primero, tercero y cuarto contrastes, mientras que habría conducido a aceptarla correctamente en el resto de los contrastes.

Para el tamaño n = 40, es más probable que la muestra hubiera conducido a la aceptación de la hipótesis nula en los mismos contrastes que para el tamaño de muestra anterior y adicionalmente en el cuarto contraste. En otras palabras, la regla de decisión discriminó bien en el cuarto contraste para el tamaño n = 40, aunque no lo hizo para n = 20.

Por último, si se escogieran las constantes de manera que se cumpla que $(B_1 - B_0) / A \leq -1$, habría sido más probable que la muestra aleatoria generada hubiera llevado a aceptar la hipótesis nula, en los siete contrastes analizados en esta sección.

Conclusiones

El procedimiento bayesiano que se usó en este trabajo para el contraste de hipótesis paramétricas en la familia Alfa, el cuál emplea como función de utilidad una función lineal de la discrepancia entre dos modelos probabilísticos, proporcionó una solución general a todos aquellos contrastes analizados en los que θ es el parámetro de interés y ω es el parámetro de ruido, en aquellos en los que ω es el parámetro de interés y θ es el parámetro de ruido y en aquellos en los que ambas hipótesis son simples.

El procedimiento dió una solución a los contrastes en los que ambos parámetros son de interés y en la hipótesis alternativa no se especifica un valor particular para ω siempre y cuando θ_1, o bien θ_1^* (dependiendo de si se da un sólo valor para θ en la hipótesis alternativa o no), sea menor que 1. La razón por la que el procedimiento no proporcionó una solución a los contrastes que no cumplen con la condición anterior tuvo su orígen en que el valor esperado de la discrepancia se hace más negativo conforme ω se acerca a 0 o a 1, provocando que no exista el valor ω_1^* de ω en el que el valor esperado de la discrepancia sea mínimo. Por otra parte, es de esperarse que sí exista la solución para estos contrastes siempre y cuando el subconjunto de valores para ω bajo la hipótesis alternativa no tenga como puntos frontera al 0 ni al 1.

Se hizo evidente que la elección de las constantes A, B_0 y B_1 que aparecen en la función de utilidad son importantes no sólo por proporcionar una familia de funciones de utilidad más amplia, sino por ser imprescindibles para evitar que en algunos contrastes se rechace o se acepte de antemano la hipótesis nula, esto es, sin tomar en cuenta la información contenida en la muestra. La cantidad $B_1 - B_0$ puede interpretarse como la predisposición por escoger la hipótesis alternativa sobre la hipótesis nula, cuando $A = 1$.

En los contrastes de la forma H_0: $\theta = \theta_0$ vs. H_1: $\theta \neq \theta_0$, es requisito indispensable pedir que la predisposición por escoger la hipótesis alternativa sobre la hipótesis nula sea menor que cero, es decir, que haya una predisposición neta por escoger la hipótesis nula.

También hay que mencionar que el procedimiento de estimación, mediante la minimización del valor esperado de la discrepancia, proporcionó muy buenas aproximaciones al verdadero valor de θ en todas las simulaciones realizadas para los contrastes en los que ω es el parámetro de ruido y no se especifica un sólo valor para el parámetro de interés θ en la hipótesis alternativa.

Sería interesante que en trabajos posteriores se pudiera estudiar las propiedades de los estimadores obtenidos mediante dicho método y se hicieran simulaciones en las que se estimen ambos parámetro para verificar qué tan buenas son las estimaciones en estos casos.

Así mismo, se propuso un método bayesiano para probar si un modelo probabilístico completamente especificado es compatible con el comportamiento de una variable aleatoria, basados en la información proporcionada por una muestra aleatoria. Puede resultar interesante llevar a cabo en un trabajo futuro algún proceso de simulación con la intención de ver qué tan buenas son las soluciones obtenidas con este método ante muestras generadas a partir de diversos modelos probabilísticos; así como, realizar un estudio comparativo de este método con respecto a otros métodos comúnmente empleados para probar bondad de ajuste.

Hay que señalar que en el contraste de hipótesis simples se encontró que la función de la muestra que aparece en la región crítica obtenida con el Lema de Neymann-Pearson tiene una distribución, cuando la hipótesis nula es verdadera, que resulta muy difícil de determinar; lo que trae serios problemas en la determinación de la constante k_0, cantidad esencial para tener resuelto el contraste desde el punto de vista clásico. El método de cociente de verosimilitudes generalizado fue incapaz de dar una solución a los contrastes en los que θ es el parámetro de ruido y en la hipótesis alternativa no se especifica un único valor para el parámetro de interés ω. Para los contrastes restantes no se puede asegurar que la solución exista, sino que está condicionada a la existencia de los estimadores máximo-verosímiles correspondientes.

Lo anterior constituye una desventaja del procedimiento clásico frente al bayesiano, ya que el procedimiento bayesiano permitió abordar y resolver un rango más amplio de contrastes dentro de la familia Alfa.

Bibliografía

Abramowitz, M. & Stegun, I.(1968). Handbook of mathematical functions.Dover publications Inc. New York.

Bayarri, M.(1984). Contraste bayesiano de modelos probabilísticos.(Tésis doctoral). Dpto. Bioestadística, Dpto. de Estadística e Inv. Operativa, U. de Valencia, España.

Bernardo, J.(1979). Reference posterior distribution for bayesian inference. J. R. Statist. Soc. B(1979),vol 41, No. 2,pp. 113-147.

Bernardo, J.(1982). Contraste de modelos probabilísticos desde una perspectiva bayesiana. Trabajos de estadística y de investigación operativa; vol 33, No. 2, 16-30.

Burden, R. L.(1978). Numerical Analysis. Prindle, Weber & Schmidt Incorporated.

D'agostino, R. & Stephens, M.(1986). Goodness-of-fit Techniques.Marcel Dekker inc.

Harris, B.(1966). Theory of probability, Addison-Wesley.

Applied Statistics. vol. 31, No. 2, 1982, pp. 189.

Kennedy, W. & Gentle, J.(1980). Statistical computing. Marcel Dekker Inc.

Kullback, S.(1959). Information theory and statistics. Wiley, New York.

Lindley, D.(1971). Making decisions. John Wiley & Sons Ltd.

Rueda, R.(1992). A Bayesian Alternative to Parametric Hypothesis Testing. Test. vol. 1, pp. 43-47.

Savage, L.(1972). The fundations of statistics. 2o. revised edition. Dover Publications Inc. New York.