



29/50

---

---

**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**  
**Facultad de Ciencias**

**TECNICAS ESTADISTICAS PARA EL ANALISIS  
DE BIOENSAYOS CUALITATIVOS**

**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 :  
**Renata Ma. Villalba Cohen**

---

---

México, D. F.

Enero 1989

**FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## PROLOGO

El propósito de este trabajo es presentar, en forma introductoria algunas técnicas estadísticas que se utilizan para analizar resultados provenientes de experimentos del tipo que se denomina con el nombre genérico de bioensayo.

Específicamente, se tratará el caso de bioensayos cuantales comparativos y se discutirán algunos procedimientos estadísticos que se han desarrollado desde el punto de vista frecuentista así como procedimientos desarrollados dentro del enfoque bayesiano.

Los elementos básicos de un bioensayo así como las características fundamentales de la inferencia bayesiana y frecuentista se presentan en el Capítulo 1. En el Capítulo 2 se presentan algunos procedimientos que utilizan el enfoque frecuentista y en el Capítulo 3, se describen algunas técnicas que se han propuesto en un contexto bayesiano.

Es importante mencionar que este trabajo no pretende una revisión exhaustiva del tema sino mas bien introducir al lector a una serie de métodos poco usados y que son útiles en el análisis de una gran variedad de experimentos.

## INDICE

CAPITULO 1. INTRODUCCION	2
CAPITULO 2. ENFOQUE FRECUENTISTA	20
CAPITULO 3. ENFOQUE BAYESIANO	31
CAPITULO 4. CONCLUSIONES	59
BIBLIOGRAFIA	63

# **CAPITULO I**

## **INTRODUCCION**

## INTRODUCCION

### I.1. CONCEPTOS BASICOS.

En el campo de la investigación científica la Estadística se ha desarrollado, en muchos casos, respondiendo a necesidades de diversas ciencias. Su carácter interdisciplinario ha generado una gran variedad de técnicas que son útiles para analizar y resolver diferentes tipos de problemas específicos. En particular, en el campo de la experimentación biológica existe una amplia gama de métodos; entre ellos, los mas relevantes son los agrupados bajo el nombre genérico de Diseño de Experimentos. Sin embargo, existen otros procedimientos no menos importantes como es el caso de las técnicas de Bioensayo.

Desde un punto de vista general, el Bioensayo está asociado a experimentos en unidades biológicas y puede ser descrito como el proceso de evaluar la potencia de un estímulo analizando las respuestas que produce en esas unidades. El estímulo podría ser por ejemplo, un veneno y el experimento consistiría en aplicar diferentes dosis del veneno a diferentes sujetos y observar si el sujeto "muere" o "sobrevive". En otra variante, se podría tratar de un experimento en el cual el estímulo fuera un cierto tipo de insulina aplicada en diferentes dosis a distintos sujetos y en el que, como respuesta, se mida la cantidad de azúcar en la sangre de los sujetos después de un determinado periodo de tiempo.

En el primer caso se podría estar interesado en estimar la dosis letal o bien la mínima dosis con la cual al menos la mitad de los sujetos mueren (dosis letal mediana). En el segundo ejemplo, podría ser de interés estimar la dosis de insulina que produzca una respuesta (nivel de azúcar) específica.

En general, el esquema de Bioensayo involucra un conjunto de estímulos que se aplica a sujetos experimentales los cuales generan un conjunto de respuestas (variable de respuesta).

De acuerdo a Finney (1978), un bioensayo posee dos características principales: la variable de respuesta (Y) y la variable "dosis" (X) que mide la intensidad o cantidad en que se administran los estímulos a los sujetos en el ensayo.

De acuerdo a Mendoza (1986), un bioensayo puede ser clasificado en dos tipos dependiendo de la naturaleza cualitativa o cuantitativa de la variable de respuesta. Usualmente, un bioensayo cualitativo es aquel en el que la variable de respuesta es cuantitativa (dicotómica) mientras que en un bioensayo cuantitativo la variable de respuesta es continua. También un bioensayo puede ser descriptivo si sólo involucra a un estímulo o comparativo si se ensayan dos o más estímulos bajo las mismas condiciones para fijar sus potencias relativas. Además, un bioensayo puede ser "directo" si se selecciona de antemano una respuesta específica y se ensayan diversas dosis en el experimento hasta alcanzar por primera vez esa respuesta seleccionada o bien puede ser "indirecto" si las dosis se fijan de antemano y se aplican a los sujetos para producir una respuesta aleatoria.

La Estadística aporta al Bioensayo los principios estadísticos generales para el método de ensayo, los diseños experimentales que proporcionen resultados útiles y confiables y el análisis de los datos experimentales optimizando el uso de la información proporcionada. Es importante mencionar que las técnicas estadísticas de Bioensayo pueden ser utilizadas en otro tipo de investigaciones, no biológicas, de características similares.

De acuerdo a la tipificación de bioensayos que se mencionó anteriormente, existen diversas técnicas estadísticas que se adecúan a cada situación. En forma muy general, en el caso de bioensayos cuantitativos si el estímulo es también cuantitativo, comúnmente se recurre a las técnicas usuales de Análisis de Regresión mientras que si el estímulo es cualitativo pueden utilizarse las técnicas conocidas como de Diseño de Experimentos. En el caso de bioensayos cualitativos por otra parte , las técnicas son variadas y existen diversos trabajos reportados en la literatura.

En este punto vale la pena recordar que en Estadística existen, principalmente, dos maneras de enfocar el planteamiento de cualquier problema de inferencia. En este tipo de problemas, la incertidumbre está presente y es natural entonces recurrir a la probabilidad como un medio de describirla. Precisamente en la interpretación y la forma de asignar probabilidades es que difieren los dos enfoques que se han mencionado y que se conocen con el nombre de frecuentista y bayesiano.

Desde un punto de vista matemático, la Probabilidad se define mediante una serie de axiomas establecidos por Kolmogorov (1933), a partir de los cuales se genera la Teoría Matemática de la Probabilidad. Esta estructura lógica es independiente de cualquier interpretación que se le dé a la Probabilidad.

Por otra parte, las diversas interpretaciones de la Probabilidad se distinguen particularmente en la forma en como ésta se asigna. La interpretación frecuentista considera que para poder asignar la probabilidad de ocurrencia de un evento, debe de ser posible realizar un número "suficientemente grande" de veces, en forma independiente e idéntica, un experimento, en el cual uno de sus posibles resultados sea el evento en cuestión. La



probabilidad básicamente se asigna como el cociente entre el número de veces que el evento de interés ocurre,  $n_A$ , y el número total de veces,  $n$ , que el experimento se repite. Es decir,  $p(A) = \frac{n_A}{n}$ , donde A es el evento de interés. Esta regla de asignación presupone entre otras cosas que:

- i) El límite cuando  $n$  tiende a infinito, del cociente  $\frac{n_A}{n}$  (la "verdadera probabilidad") existe y es único.
- ii) La probabilidad de un evento es una característica "objetiva" es decir, que depende únicamente del evento en cuestión.

Esta interpretación fué utilizada desde los siglos XVI y XVII en los problemas que surgen en los juegos de azar. En el siglo XVII Pierre de Fermat enmarcó el problema en la teoría general de combinatoria que fué desarrollada tiempo después, en los inicios del siglo XVIII, por el matemático J. Bernoulli. Este último puede ser considerado como el fundador de la teoría de la Probabilidad como una rama de las Matemáticas. Su libro póstumo "Ars Conjectandi" escrito en 1713, puede advertirse como la fusión de los métodos combinatorios probabilísticos y los inicios de los métodos estadísticos. Las investigaciones que se llevaron a cabo durante el siglo XVIII culminaron en un trabajo de Pierre Simon de Laplace, cuyas ideas dominaron a lo largo del siglo XIX.

La teoría estadística basada en la interpretación frecuentista de la Probabilidad surge a finales del siglo XIX con los trabajos de Karl Pearson y recibe un impulso muy importante en las primeras décadas del siglo XX con la teoría desarrollada por R. Fisher, J. Neyman y E. S. Pearson. La mayoría de los conceptos estadísticos relacionados con problemas de estimación y pruebas de

hipótesis son fundamentalmente debidos a ellos.

El desarrollo de la teoría estadística frecuentista causó un fuerte impacto en esa época condicionando casi en su totalidad el desarrollo de otras escuelas. Aún más, puede decirse que en los tiempos actuales la teoría es básicamente la misma y tiene un uso muy generalizado.

Por otra parte, la interpretación subjetiva, que corresponde al enfoque bayesiano, considera a la probabilidad como el grado de creencia que cada persona tiene sobre una proposición dada y de ahí su carácter subjetivo o personal. A diferencia de la escuela frecuentista, en el caso subjetivo no es necesario suponer que un experimento puede ser repetido, aún más, no es necesario experimentar. En contrapartida, no es posible establecer una regla general que permita asignar probabilidades como en el caso frecuentista, pues la asignación, como se ha indicado, depende de cada persona.

Los primeros intentos por formalizar estas ideas se deben a F. P. Ramsey en 1926 y a B. de Finetti en 1937. Fué hasta 1954 que L. J. Savage retomó las ideas de Ramsey y De Finetti, generando un sistema consistente de principios que serían la base de las escuelas probabilística y estadística subjetivas. L. J. Savage desarrolló su trabajo desde la perspectiva de la Teoría de Decisiones enfatizando el carácter subjetivo que existe cuando la persona involucrada en la toma de decisiones cuantifica su incertidumbre haciendo uso de la Probabilidad. La teoría estadística que resulta de esta interpretación es conocida como estadística bayesiana debido al intensivo y fundamental uso que hace del Teorema de Bayes. En forma muy general este teorema permite actualizar el conocimiento, descrito mediante una probabilidad, que una persona posee sobre un fenómeno.

La estadística bayesiana no tuvo el desarrollo de la frecuentista pues como ya se mencionó fué hasta 1954 que se dieron las bases sólidas para su aplicación, mientras que el enfoque frecuentista, para esas fechas, gozaba de gran aceptación debido a su fuerte desarrollo teórico y la gran variedad de situaciones en las que podía ser aplicado.

El enfoque bayesiano ha sido criticado debido a su carácter subjetivo. La mayoría de los estadísticos frecuentistas opinan que los resultados obtenidos en un análisis estadístico no pueden depender de la persona que lo lleve a cabo sino que debe ser totalmente "objetivo", es decir, que los experimentos deben proporcionar toda la información relevante teniéndose así que si dos personas realizan el mismo experimento deben concluir lo mismo. Esta posición genera una controversia pues los estadísticos bayesianos argumentan que si se cuenta con información extra-experimental ésta debe ser incorporada a la estructura del problema dando como resultado que, aún cuando dos personas realicen el mismo experimento, al contar con información personal (extra-experimental) diferente, pueden llegar a conclusiones distintas.

En la actualidad la escuela bayesiana ha cobrado importancia debido fundamentalmente a que cuenta con una serie de principios normativos que se aplican a todo problema de inferencia y garantizan la coherencia y admisibilidad de los resultados que produce. Es un hecho que el enfoque frecuentista carece de una estructura lógica de este tipo. Un creciente número de estadísticos desarrollan su trabajo dentro del enfoque bayesiano, acentuando la controversia puesto que la estadística frecuentista también prosigue su desarrollo.

Dadas las dos interpretaciones, las correspondientes

escuelas estadísticas podrían considerarse antagónicas, sin embargo existen situaciones en las que los resultados numéricos coinciden y los procedimientos utilizados por ambas son similares. Más aún, bajo ciertas condiciones es posible probar que algunos resultados frecuentistas se pueden considerar casos particulares de las correspondientes soluciones bayesianas.

Por un lado la estadística frecuentista, debido a su gran desarrollo a lo largo de este siglo, cuenta con una variedad de procedimientos que resuelven una gran gama de problemas en forma relativamente sencilla; sin embargo no se cuenta con una teoría formal que proporcione los lineamientos a seguir en la solución de cualquier problema teniéndose así un conjunto de técnicas que en algunos casos son contradictorias.

Por otro lado, como ya se indicó, la estadística bayesiana tiene una estructura formal basada en un conjunto de principios que norma los pasos a seguir en la solución de problemas, sin embargo debido generalmente a las dificultades de orden técnico, aún no se cuenta con soluciones explícitas para algunos problemas que frecuentistamente están resueltos. Además, las dificultades para expresar el grado subjetivo de creencia acerca de un fenómeno mediante una probabilidad, hacen más complicada su aplicación.

Independientemente de los problemas que cada enfoque presenta, vale la pena tomar en cuenta lo que cada uno de ellos ofrece y tomar una posición crítica y consistente con respecto a las soluciones que aportan.

Desde luego las técnicas estadísticas para el análisis de Bioensayos no se encuentran al margen de esta controversia por lo que existen trabajos desarrollados bajo los dos enfoques.

Debido a la tendencia histórica, la literatura relativa al tema es más extensa en la parte frecuentista que en la bayesiana, sin embargo en años recientes se han incrementado los trabajos desde esta última perspectiva.

Este trabajo retoma algunas técnicas asociadas a bioensayos cuantales comparativos e indirectos y presenta con relativo detalle algunos de los trabajos más representativos de cada corriente.

## I.2. ELEMENTOS DE LOS MODELOS DE BIOENSAYO CUALITATIVO.

Como se mencionó en la sección I.1, existen diversos tipos de bioensayos. Específicamente, en este trabajo se tratará el caso particular de bioensayos cuantales comparativos. En este tipo específico de bioensayos la situación que se tiene es la siguiente:

Suponga que para una dosis fija  $X$ , se calcula la fracción de individuos de una población que presentaron una respuesta positiva. A esta cantidad se le llama "potencia" de la dosis  $X$  y se denota por  $P(X)$ . Al graficar diferentes dosis contra sus potencias se obtiene lo que se conoce como la curva de respuesta o función potencia. En un experimento diseñado para este tipo de bioensayo, se selecciona un conjunto de dosis  $X_1, \dots, X_m$  y cada una de ellas se administra a un conjunto de sujetos experimentales. Cada dosis  $X_i$  se aplica a  $n_i$  sujetos y los resultados se miden en respuestas cuantales (dicotómicas) de tal forma que las mediciones son justamente el número de sujetos que reaccionó positivamente en cada dosis, es decir, si  $Y$  es la variable de respuesta entonces

$$Y = \begin{cases} 1 & \text{reacción positiva} \\ 0 & \text{reacción negativa} \end{cases}$$

en donde  $Y$  es una variable aleatoria tal que

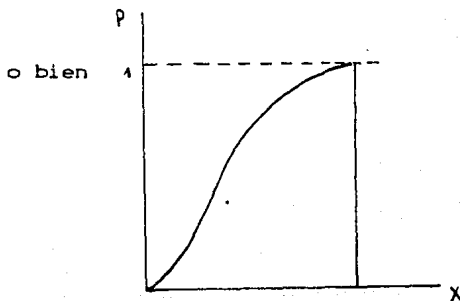
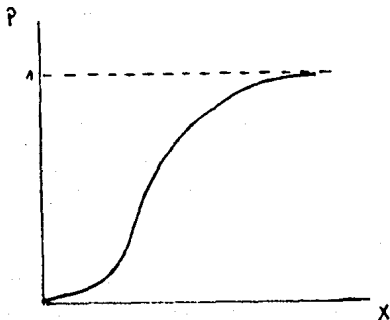
$$P \{ Y = 1 \} = P(X) \quad y$$

$$P \{ Y = 0 \} = 1 - P(X)$$

con  $P(X)$  la función potencia en la dosis  $X$ .

Un objetivo usual en este tipo de experimentos es determinar la dosis  $X$  a la cual se obtenga una potencia específica. También hay que notar que la forma de la función potencia es desconocida, sólo se estima en los puntos en los cuales se tomaron observaciones, por lo que resulta interesante preguntarse por el valor de la función potencia en otra dosis que no sea alguna de las que ya se aplicaron. En consecuencia, el problema a resolver es encontrar la dosis tal que su potencia sea un valor predeterminado y/o bien dada una dosis, conocer el valor de su potencia.

Dado que  $P(X)$  representa la fracción de sujetos en la población que presentaron una respuesta positiva, una posible gráfica de  $P(X)$  podría ser



La función potencia es típicamente desconocida, pero en general tiene las siguientes características: es monótona no decreciente, su mínimo valor es cero y su máximo es uno.

Si la curva fuese estrictamente creciente, el problema podría ser visto como uno de calibración, sin embargo ese no es el caso. Aún más, en términos generales, este problema es similar al caso típico de Regresión. De hecho,  $P(X)$  podría considerarse lineal por pedazos. Sin embargo si sólo se aplica cada dosis a un solo sujeto, la distribución de  $Y$  no es razonablemente aproximada por una distribución Normal (para incorporar la suposición usual de normalidad en Regresión). En cambio si cada dosis se aplica a un lote de  $m$  unidades experimentales y se registra  $Y'$ , el número de éxitos, la distribución de  $Y'$  (por dosis) es Binomial y el problema parece simplificarse. Sin embargo, en este caso, se tiene que  $E(Y')=mP(X)$  y  $Var(Y')=mP(X)[1-P(X)]$  (por dosis), es decir no hay igualdad de varianzas y habría que recurrir al empleo de alguna transformación para estabilizar varianzas (e.g. Draper & Smith, 1981. Cap. 5).

De la forma de la curva y de sus propiedades habría posibilidades de aproximarla mediante una función de distribución o bien buscar otro tipo de funciones que cumplan con las características de  $P(X)$ . El ajustar  $P(X)$  mediante una función de distribución puede ser difícil ya que es posible que su gráfica no sea fácilmente identificada con las curvas de distribuciones conocidas.

Un camino que varios autores han explorado es el de ajustar  $P(X)$  mediante la función de distribución logística, que está definida por

$$P(X) = \frac{1}{1 + e^{-(\alpha + \beta x)}} \quad \beta > 0$$

Como se mencionó anteriormente, no sólo se desea ajustar la curva  $P(X)$  sino encontrar la relación entre la dosis  $X$  y la respuesta  $Y$ . De hecho, es de interés encontrar  $X_{p^*}$  tal que

$$X_{p^*} = \min \{ X : P(X) \geq p^* \} \quad \text{con } p^* \text{ un valor predeterminado.}$$

En particular, uno de los casos más estudiados y reportados en la literatura es el de  $p^* = 1/2$  en cuyo caso  $X_{p^*}$  recibe el nombre de dosis mediana efectiva.

Estadísticamente, el planteamiento general del problema se describe a continuación:

Si se emplean  $M$  dosis  $(X_1, X_2, \dots, X_M)$  y cada una de ellas se aplica a cada uno de  $n_i$  sujetos (ó unidades experimentales) en forma independiente, se tiene que la información queda modelada como

$$p(Y_{ij} | X_i) = [P(X_i)]^{y_{ij}} [1 - P(X_i)]^{1-y_{ij}} \quad \begin{matrix} j=1,2,\dots,n_i \\ i=1,2,\dots,M \end{matrix}$$

entonces, la función de verosimilitud es

$$\begin{aligned} p(\underline{Y} | \underline{X}) &= \prod_{i=1}^M \prod_{j=1}^{n_i} [P(X_i)]^{y_{ij}} [1 - P(X_i)]^{1-y_{ij}} \\ &= \prod_{i=1}^M [P(X_i)]^{s_i} [1 - P(X_i)]^{n_i - s_i} \end{aligned}$$

con  $s_i = \sum_{j=1}^{n_i} Y_{ij}$



Particularmente, si se adopta una estructura específica para  $P(X)$ , el problema se reduce a la determinación de un conjunto de parámetros. Por ejemplo, si se elige

$$P(X) = \frac{1}{1 + e^{-(\alpha + \beta X)}} \quad (\beta > 0)$$

se tiene que

$$P(X) = \frac{1}{1 + e^{-\beta(\gamma - X)}} \quad \text{con } \gamma = \frac{\alpha}{\beta}$$

Es interesante notar que en este modelo  $X_{1/2} = \gamma$ , de modo que la estimación de la dosis mediana efectiva se reduce a estimar  $\gamma$ . Mas aún se puede verificar que

$$X_{p^*} = \gamma + \frac{1}{\beta} \log\left(\frac{1 - p^*}{p^*}\right)$$

de modo que si  $\beta$  es conocida, de nueva cuenta el problema se limita a estimar  $\gamma$ .

Es importante mencionar que el problema de ajustar  $P(X)$  por una determinada función puede evitarse y como alternativa se puede desarrollar una técnica apropiada sin imponer una estructura específica para la función potencia. Parecería razonable esperar una mayor complicación si se renuncia a una estructura de  $P(X)$ , sin embargo, existen algunos resultados generales en esta línea, que resultan muy útiles (cf. Ramsey, 1972).

Antes de presentar con relativo detalle algunas técnicas particulares, es conveniente recordar los conceptos básicos de la Inferencia Estadística desde los puntos de vista bayesiano y frecuentista.

### I.3. CONCEPTOS BASICOS DE LOS ENFOQUES BAYESIANO Y FRECUENTISTA

El enfoque bayesiano en la parte inferencial, consiste esencialmente en lo siguiente:

Sea  $\theta$  un parámetro de interés. Esta cantidad es una constante desconocida. El conocimiento que el investigador posea sobre  $\theta$  debe describirse mediante una distribución de probabilidad que se denota por  $p(\theta)$ . Esta distribución representa entonces, el conocimiento inicial que el investigador tiene sobre el fenómeno bajo estudio y comúnmente se le llama distribución apriori o inicial.

Por otra parte es usual que el investigador desee contar con información adicional a la que posee sobre el parámetro y que de esta manera recurrirá a la realización de un experimento con el fin de recolectar nueva información que al combinarse con su conocimiento inicial, le proporcione un nuevo estado de conocimiento sobre  $\theta$ . Si  $z$  representa el resultado obtenido en el experimento, el conocimiento sobre  $\theta$  después de haber obtenido  $z$  quedará descrito en la distribución  $p(\theta|z)$  a la que se llama distribución posterior o final. Los términos inicial y final son relativos a la obtención de  $z$ .

La función de verosimilitud  $p(z|\theta)$  como función de  $\theta$ , mide que tan verosímiles son los posibles valores de  $\theta$  para un conjunto de datos  $z$  dado.

La forma natural de incorporar y combinar esta información con la inicial  $p(\theta)$  es a través de la expresión

$$p(\theta|z) = p(\theta)p(z|\theta)/p(z)$$

que es equivalente a

$$p(\theta|z) = \frac{p(\theta)p(z|\theta)}{\int p(\theta)p(z|\theta)d\theta}$$

donde la integral es sobre  $\Omega$ , el espacio parametral.

A esta expresión se le conoce con el nombre de Teorema de Bayes debido a que fué T. Bayes quien la introdujo en 1763. Puede notarse que la expresión del denominador  $\int p(\theta)p(z|\theta)d\theta$  no es más que una constante de normalización, ésto es, lo único que asegura es que  $\int p(\theta|z)d\theta=1$ . Por esta razón, el teorema de Bayes se enuncia frecuentemente como

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

de aquí que  $p(\theta|z)$  se interprete como el conocimiento que se adquiere sobre  $\theta$  a partir de los datos  $z$ .

El papel que la distribución inicial juega en los métodos bayesianos es de vital importancia y su elección debe realizarse muy cuidadosamente. El problema de asignar una distribución inicial no es sencillo y a menudo se recurre a métodos sugeridos en la literatura (ver Winkler 1967, 1972, De Finetti 1962, Savage 1971, Lindley 1965, 1971, 1977). Sin embargo en diversas situaciones no es posible asignar una distribución inicial que exprese las opiniones del investigador ya sea por problemas técnicos o porque se considere conveniente introducir una distribución inicial que no involucre un conocimiento prejuiciado del investigador.

De hecho en algunas ocasiones es conveniente realizar el análisis con una distribución que no necesariamente represente el

conocimiento del investigador o bien que influya "poco" en los datos de forma tal que no los apoye ni tampoco los desmienta. Además, una distribución de este tipo puede ser utilizada por cualquier persona como un punto de referencia para establecer el efecto de su información inicial personal en los resultados del análisis. Las distribuciones de este tipo son llamadas de referencia y se emplean muy a menudo para comparar los enfoques clásico y bayesiano de la Estadística.

En el caso de combinar un conjunto de datos  $z$  con una distribución inicial de referencia vía el teorema de Bayes, a la distribución final que se obtiene se le llama distribución final de referencia.

Para construir distribuciones de referencia existen varios métodos reportados en la literatura, entre los cuales se encuentran los de Jeffreys (1946), DeGroot (1970) y Bernardo (1979).

El problema específico de estimación desde un punto de vista bayesiano puede plantearse en dos niveles: el primero de ellos dentro del contexto de toma de decisiones y el segundo a un nivel inferencial.

Desde el punto de vista de la Teoría de Decisiones, el problema de estimación consiste en elegir un valor de  $\hat{\theta}$  para aproximar el valor desconocido de  $\theta$ . De acuerdo a los fundamentos de la teoría para tomar una decisión, además de asignar una distribución de probabilidad sobre  $\theta$ , debe especificarse una función (llamada de utilidad) que cuantifique las consecuencias de tomar una determinada decisión. Aquella decisión que maximice la utilidad esperada será la decisión óptima (Savage 1954 y Bernardo 1981). En forma equivalente puede definirse una función de

pérdida para la cual la decisión óptima será la que minimice la pérdida esperada.

En el caso de estimación, una función de pérdida que se utiliza con frecuencia es la cuadrática, que mide la diferencia o discrepancia cuadrática entre  $\theta$  y  $\hat{\theta}$ , que se define como

$$L(\theta, \hat{\theta}) = a(\theta - \hat{\theta})^2 \quad \text{con } a > 0$$

y para la cual es fácil demostrar que el mínimo de la pérdida esperada, con respecto a la distribución en  $\theta$ , se alcanza en el correspondiente valor esperado de la distribución (apriori o posterior, según sea el caso). Es decir,  $\hat{\theta} = E(\theta)$  es el mejor estimador para  $\theta$  en el sentido de que minimiza la pérdida cuadrática esperada. Por supuesto, existen otro tipo de funciones de pérdida pero, con frecuencia involucran alguna función de la diferencia  $(\theta - \hat{\theta})$ . Una discusión amplia sobre este tema se puede consultar en DeGroot (1970).

Ahora bien, cuando el problema de estimación se plantea sólo desde el punto de vista inferencial, esto es, cuando no se pretende utilizar el valor estimado  $\hat{\theta}$  y por tanto no son claras las consecuencias, medidas a través de una función de utilidad o de pérdida, de una selección particular, es común elegir como estimador el valor de  $\theta$  mas probable o en su caso, el valor que maximiza la densidad de  $\theta$ . Así, cuando se utiliza la distribución posterior  $p(\theta|z)$  el estimador es la moda posterior de  $\theta$ .

Por otra parte, desde el punto de vista frecuentista,  $\theta$  es también el parámetro de interés cuyo valor es desconocido. Sin embargo en este enfoque no se involucra en ninguna forma el conocimiento previo que se tenga sobre  $\theta$  y se considera que la única manera de obtener información sobre el parámetro es en base a un conjunto de datos  $z$ . Existen diversos métodos para obtener,

como función de los datos  $z$ , estimadores  $\hat{\theta}=\hat{\theta}(z)$  que aproximen en algún sentido el valor de  $\theta$ . Entre ellos destaca el de máxima verosimilitud que produce estimadores generalmente razonables, dentro del contexto clásico. Para una discusión amplia sobre estimadores máximo verosímiles y sus propiedades puede consultarse Mood et al. (1974) y Kendall & Stuart (1973).

## **CAPITULO 2**

### **ENFOQUE FRECUENTISTA**

## ENFOQUE FRECUENTISTA

De acuerdo a lo que se mencionó en el capítulo anterior, es de interés estudiar la relación dosis-respuesta, cuando la respuesta es de tipo cuantal. Una descripción general para modelar esta situación es la siguiente:

Dada una dosis  $X$  aplicada a un grupo de sujetos se obtiene, para cada sujeto, una respuesta cuantal  $u$  ( $u=1$  en caso de "éxito" y  $u=0$  en caso de "fracaso"), de manera que el valor esperado de  $u$  dado  $X$  medirá la proporción de sujetos que presentaron respuesta positiva (éxito). Si se denota con  $U$  a este valor esperado, se tiene entonces que

$$E(u|X) = U \quad (2.1)$$

Más aún, la relación (2.1) puede plantearse como

$$U = f(X) \quad (2.2)$$

en donde  $f$  es típicamente, una función monótona creciente en el intervalo  $[0,1]$ , desconocida. El problema que se plantea para describir la relación dosis-respuesta es elegir una función dentro de la familia

$$\mathcal{F} = \{ f: f \text{ monótona creciente en } [0,1] \}$$

Seleccionar  $f \in \mathcal{F}$  es, desde el punto de vista estadístico, un problema no paramétrico de difícil solución dada la cardinalidad de  $\mathcal{F}$ . Sin embargo, una forma de simplificar esta situación es considerar una familia paramétrica



$$\mathcal{F}_\theta = \{ f_\theta, \theta \in \Omega \mid f_\theta \text{ monótona creciente en } [0,1] \}$$

de tal forma que (2.2) puede reescribirse como

$$U = f_\theta(X) \quad (2.3)$$

De hecho, en la situación de bioensayo planteada en el capítulo anterior, se tiene que  $f_\theta(X)$  es la función potencia  $P(X)$ , es decir

$$P(X) = U = f_\theta(X)$$

Un modelo paramétrico utilizado con frecuencia es

$$P_\theta(X) = P_0(X - \theta)$$

con  $P_0$  una función conocida y monótona creciente en el intervalo  $[0,1]$  en donde  $\theta \in \mathbb{R}$  es un parámetro de localización. Más aún,  $P_0(X)$  puede construirse como

$$P_\theta(X) = P_0(\theta_1(X - \theta_2)) \quad (2.4)$$

con  $P_0$  una función conocida monótona creciente en el intervalo  $[0,1]$  y  $\underline{\theta} = (\theta_1, \theta_2) \in \mathbb{R}^2$  con  $\theta_1$  un parámetro de escala y  $\theta_2$  un parámetro de localización.

Utilizando (2.3) y (2.4) se tiene que una función que modela la relación dosis-respuesta está dada por

$$P(X) = U = P_0(\theta_1(X - \theta_2))$$

Como  $P_0$  es monótona creciente se sigue que

$$P_0^{-1}(U) = \theta_1(X - \theta_2) \quad (2.5)$$

con  $P_0^{-1}$  la función inversa de  $P_0$  que existe siempre que  $P_0$  sea continua.

La expresión (2.5) representa una relación dosis-respuesta lineal, que puede reescribirse como

$$Y = \alpha + \beta X \quad (2.6)$$

con  $Y = P_0^{-1}(U)$ ,  $\alpha = -\theta_1\theta_2$  y  $\beta = \theta_1$ .

Finney (1978) argumenta que una transformación adecuada de la dosis permite construir una relación de la forma (2.6). El autor propone utilizar una transformación de la familia generada por Box & Cox (1964) sobre la dosis, de tal forma que, en general se considere

$$U = P_0(\alpha + \beta Z) \quad (2.7)$$

donde  $Z$  es la transformación de la dosis  $X$ , llamada metadosis o dosis metamétrica.

Suponiendo  $P_0$  monótona y continua, se obtiene de (2.7) que

$$Y = P_0^{-1}(U) = \alpha + \beta Z$$

con  $Y$  la respuesta metamétrica o metarespuesta.

Ahora bien, desde el punto de vista frecuentista, dadas  $P_0$  y la transformación  $Z$ , el problema se reduce a estimar los

parámetros  $\alpha$  y  $\beta$  en la relación dosis-respuesta metamétrica.

En este capítulo se describirá el análisis de bioensayos cuantales bajo diferentes estructuras para modelar  $P_0$ , suponiendo que la transformación  $Z$  para lograr la linealidad ya está dada (debido a lo anterior, se denotará indistintamente por  $X$  tanto a la dosis original como a la dosis transformada). Se dará primero una breve descripción de cada uno de los modelos propuestos y posteriormente se comentarán globalmente.

El primer modelo que se dará como aproximación a  $P$  es la distribución Normal. Bliss (1944) presenta un ejemplo en el cual suponer que el logaritmo de la dosis sigue una distribución Normal, describe en forma muy precisa la relación dosis-respuesta. Este ejemplo sugiere que el uso de esta distribución para modelar a  $P(X)$  parece razonable. Por otro lado, el Teorema Central del Límite puede justificar el uso de la distribución Normal cuando la variable de respuesta es un promedio.

En esta situación y de acuerdo a la ecuación (2.7) se tendrá que

$$P(X) = \int_{-\infty}^{\alpha + \beta X} (2\pi)^{-1/2} \exp\left\{-\frac{1}{2} t^2\right\} dt \quad (2.8)$$

Denotando por  $\Phi$  a la función de distribución Normal, la respuesta original  $U$  debe transformarse como  $Y = \Phi^{-1}(U)$ , para obtener la relación (2.6).

Bliss (1934a,b) introduce la conocida función Probit que se define como

$$P(X) = \int_{-\infty}^{y-\alpha} (2\pi)^{-1/2} \exp\left\{-\frac{1}{2} t^2\right\} dt \quad (2.9)$$

con  $Y = \alpha + \beta X$

La función Probit obliga a que sea poco probable la obtención de valores negativos.

Finney (1978) menciona que en muchos experimentos el uso de la ecuación (2.9) es preferible a la ecuación (2.8). Una discusión más amplia sobre la transformación Probit puede verse en Finney (1971).

Berkson (1944) propone la distribución Logística como una alternativa a la Normal argumentando que las dos distribuciones son similares entre sí. El autor menciona que la Logística puede aplicarse a una gran variedad de experimentos físico-químicos.

En términos de la distribución Logística,  $P(X)$  queda descrita como

$$P(X) = \frac{1}{1 + e^{-(\alpha + \beta X)}} \quad \alpha, \beta > 0 \quad (2.10)$$

y la correspondiente transformación metamétrica es

$$Y = \ln \frac{P(X)}{1 - P(X)} \quad (2.11)$$

Berkson llamó a esta transformación el Logit de  $P(X)$ .

Wilson & Worcester (1943a,b,d) proponen otra transformación para generar la distribución Logística que difiere de la anterior por un factor de 2. La transformación es la siguiente

$$P(X) = \frac{1}{2} (1 + \tanh(\alpha + \beta X))$$

que es equivalente a

$$P(X) = \frac{e^{2Y}}{1 + e^{2Y}} \quad (2.12)$$

obteniéndose que

$$Y = \frac{1}{2} \ln \frac{P(X)}{1 - P(X)} \quad (2.13)$$

Finney (1978) menciona que una ventaja de la transformación (2.12) es que resulta muy similar en magnitud a la ecuación (2.8) y que entonces podrían hacerse comparaciones con la transformación Probit.

Por su parte, Wilson & Worcester (1943c) introducen una distribución para la cual

$$P(X) = \int_{-\infty}^{\alpha + \beta X} \frac{1}{2} (1 + t^2)^{-3/2} dt$$

$$= \frac{1}{2} \left[ 1 + \frac{Y}{(1 + Y^2)^{1/2}} \right]$$

con  $Y = \alpha + \beta X$ . De donde

$$Y = \frac{2P(X) - 1}{2[P(X)(1 - P(X))]^{1/2}}$$

Los autores mencionan que esta función es matemáticamente tratable y fácilmente comparable con las transformaciones Normal y Logística.

Como una alternativa a la distribución Normal, Urban (1909, 1910), basado en un estudio en Psicometría con respuestas cuantales, propone la siguiente distribución

$$\begin{aligned} P(X) &= \int_{-\infty}^{\alpha + \beta X} \frac{1}{\pi(1 + t^2)} dt \\ &= \frac{1}{2} + \pi^{-1} \tan^{-1}(\alpha + \beta X) \end{aligned}$$

Esta función es conocida como distribución Cauchy, para la cual se sabe que no existe la media, la varianza y en general, ningún momento.

La respuesta metamétrica resulta ser

$$Y = \cot[\pi(1 - P(X))]$$

Finney (1978) comenta que esta distribución es poco realista ya que asigna probabilidades grandes a valores muy extremos.

Por otra parte, Knudsen & Curtis (1947) proponen una distribución cuya función de densidad está dada por

$$p(t) = \begin{cases} \text{sen } 2t & 0 \leq t \leq \frac{\pi}{2} \\ 0 & \text{en otro caso} \end{cases}$$

y, entonces

$$P(X) = \text{sen}^2 Y$$

con  $Y = \alpha + \beta X$ , de donde

$$Y = \text{sen}^{-1} \sqrt{P(X)}$$

Finney (1978) menciona que, especialmente en métodos de análisis de varianza cuya variable de respuesta  $r$ , está expresada como porcentaje, una transformación muy usual de esta variable es  $\text{sen}^{-1} \sqrt{r}$ . Esta transformación tiene la ventaja de que su varianza es constante. De acuerdo a lo anterior, la respuesta metamétrica  $Y$  tendrá varianza constante. Finney (1978) comenta que esta distribución sería una representación pobre para cualquier relación dosis-respuesta en el contexto de bioensayo, pero podría ser una aproximación razonable siempre y cuando los datos evitaran los valores extremos.

Por último, la aproximación más sencilla que se puede hacer es suponer que la relación con la dosis es lineal, es decir

$$P(X) = \alpha + \beta X = Y$$

o equivalentemente, tomar la función de densidad Rectangular

$$p(t) = \begin{cases} 1 & \text{si } 0 \leq t \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

El hecho de que  $0 \leq P(X) \leq 1$  implica la existencia de una dosis  $-\alpha/\beta$  por debajo de la cual ningún sujeto puede responder y una dosis  $(1 - \alpha)/\beta$  arriba de la cual todos los sujetos deben responder.

Una comparación somera entre las seis aproximaciones descritas anteriormente está dada en Finney (1978) y, en general, sus comentarios son los siguientes:

El autor hace notar que, salvo las distribuciones Normal y Logística, las demás tienen poco interés teórico. A pesar de eso, al graficar las transformaciones Logística, Normal, Rectangular y la propuesta por Knudsen & Curtis (1947), estandarizándolas de manera que cada una de ellas tenga media cero y varianza uno (razón por la cual las dos restantes no entraron en este análisis), puede verse que la relación entre la dosis  $X$  y la función potencia  $P(X)$  es muy parecida en los cuatro casos. Aún más, para valores de  $P(X)$  entre 0.05 y 0.95, las cuatro curvas son prácticamente indistinguibles. Sin embargo, menciona que la distribución Rectangular no es recomendable pues en el caso de bioensayos cuantales, esta aproximación es poco sostenible.

Si se supone que  $P(X)$  es Normal y se utiliza la transformación Probit, la relación entre  $X$  y  $P(X)$  es una línea recta. Si la transformación Probit se aplicara sobre las otras transformaciones, las gráficas entre  $X$  y  $P(X)$  serían difícilmente distinguibles de la línea recta y por tanto no sería posible detectar la no normalidad de la distribución.



Aún más, podrían sugerirse otras transformaciones que fueran igualmente difíciles de distinguir. Esta dificultad de decidir qué transformación utilizar en un conjunto de datos es aparente, pues si los datos no pueden discriminar entre alternativas, las conclusiones no se verán afectadas por la elección del modelo que aproxime a  $P(X)$ .

Como un último comentario y como se mencionó al inicio de este capítulo, al proponer un modelo paramétrico para aproximar a  $P(X)$ , el problema se reduce a estimar los parámetros involucrados en la relación  $Y = \alpha + \beta X$ . En general, las transformaciones que se proponen no cumplen con el supuesto de homocedasticidad y por tanto el proceso de estimación que se utiliza es el de Mínimos Cuadrados Ponderados.

## **CAPITULO 3**

### **ENFOQUE BAYESIANO**

## ENFOQUE BAYESIANO

Partiendo de la situación general presentada en el Capítulo 1, en la que se emplean  $M$  dosis  $\{X_1, X_2, \dots, X_M\}$  cada una de las cuales se aplica a cada uno de  $n_i$  sujetos, se tiene que la función de verosimilitud es

$$p(Y|X) = \prod_{i=1}^M [P(X_i)]^{s_i} [1 - P(X_i)]^{n_i - s_i} \quad (3.1)$$

$$\text{con } s_i = \sum_{j=1}^{n_i} y_{ij}$$

Desde la perspectiva bayesiana es necesario determinar la manera en que se ha de describir el conocimiento inicial para que al combinarse con la información experimental mediante el teorema de Bayes, se obtenga la distribución final que describe el nuevo estado de conocimiento. Dentro de esta línea se presentarán tres trabajos, dos de los cuales suponen una estructura logística para  $P(X)$  -Freeman (1970) y Tsutakawa (1980)- y el de Ramsey (1972), que no hace uso de esta suposición.

En primer término, Freeman (1970), considera la verosimilitud (3.1) y el caso en que  $P(X)$  puede describirse mediante una función logística, es decir

$$P(X) = \frac{1}{1 + e^{-\beta(\gamma - x)}} \quad \text{con } \beta > 0$$

Como una limitación del tratamiento propuesto por Freeman debe mencionarse que supone conocido al parámetro  $\beta$ . En estas condiciones, el fenómeno queda totalmente descrito al determinar

$\gamma$ . Esto es,  $\gamma$  es el único parámetro desconocido en el problema. Así, basta con asignar una distribución inicial para  $\gamma$  y obtener la correspondiente distribución final.

Fremman considera inicialmente  $M=1$  (es decir sólo una dosis  $X_1$ ) en donde

$$p(Y|X_1, \beta, \gamma) = [P(X_1)]^{s_1} [1 - P(X_1)]^{n_1 - s_1}$$

con  $s_1 = \sum_{j=1}^n y_{1j}$

pero

$$P(X_1) = [1 + e^{-\beta(\gamma - x_1)}]^{-1}$$

entonces

$$p(Y|X_1, \beta, \gamma) = \left[ \frac{1}{1 + e^{-\beta(\gamma - x_1)}} \right]^{s_1} \left[ \frac{e^{-\beta(\gamma - x_1)}}{1 + e^{-\beta(\gamma - x_1)}} \right]^{n_1 - s_1}$$

así

$$p(Y|X_1, \beta, \gamma) = \frac{[e^{-\beta(\gamma - x_1)}]^{n_1 - s_1}}{[1 + e^{-\beta(\gamma - x_1)}]^{n_1}} = \frac{e^{\beta(\gamma - x_1)s_1}}{[1 + e^{\beta(\gamma - x_1)}]^{n_1}}$$

Esta función de verosimilitud sugiere en forma natural el empleo de una distribución inicial conjugada (Raiffa & Schlaifer, 1961) para  $\gamma$  de la forma

$$p(\gamma|n_0, s_0) = \frac{e^{\beta(\gamma - x_1)s_0}}{[1 + e^{\beta(\gamma - x_1)}]^{n_0}}$$

que equivale a utilizar una Beta con parámetros  $s_0 - 1$ ,  $n_0 - s_0 - 1$  como distribución inicial para el valor de  $P(X_1)$ . (recuérdese que  $\beta$  es conocida). Freeman propone el empleo de una función de pérdida cuadrática para estimar  $\gamma$  lo cual implica que  $\hat{\gamma}$  coincide con el valor esperado posterior de  $\gamma$  que en general debe ser calculado por métodos numéricos. El autor incluye en el problema la decisión sobre el tamaño de muestra óptima suponiendo que cada observación tiene un costo fijo  $c > 0$  que debe de ser incluido en la función de pérdida. Entonces, si  $L$  es la función de pérdida se tiene que

$$L(\hat{\gamma}, \gamma) = K(\hat{\gamma} - \gamma)^2$$

se transforma en

$$L_1(\hat{\gamma} - \gamma) = K(\hat{\gamma} - \gamma) + nc$$

Freeman demuestra que el problema tiene solución para  $c > 0$  y que ésta se puede obtener resolviendo una ecuación de programación dinámica. El autor generaliza el procedimiento propuesto a los casos de dos y tres dosis los cuales se complican ya que adicionalmente hay que decidir en qué dosis se debe realizar cada observación. Freeman reconoce que si hay más de tres dosis el problema se vuelve intratable y finaliza comparando sus resultados, mediante una simulación con los del método de Dixon y Mood (1948), concluyendo que ambos métodos son igualmente efectivos y que el método frecuentista puede resultar superior si de antemano se eligen dosis cercanas a  $\gamma$ .

Por su parte Tsutakawa (1972), desarrolla un método con el objetivo de encontrar diseños experimentales óptimos. Considerando que el propósito del experimento sea la estimación de un parámetro real, el autor propone como criterio la selección del experimento que produzca una varianza posterior mínima. En particular, Tsutakawa también elige una estructura logística para  $P(X)$  que le permite discutir una variedad de situaciones.

Como se recordará, cuando se tiene una función de pérdida cuadrática y se desea estimar el valor de un parámetro  $\theta$ , el estimador de Bayes,  $\hat{\theta}$ , es la esperanza de  $\theta$  con respecto a la distribución a priori del parámetro (o posterior, dependiendo de si se toma información adicional). En particular, si  $\theta$  tiene asociado una distribución a priori  $\lambda(\theta) = \mathcal{M}(\mu, \tau^2)$  y se cuenta con una muestra de observaciones independientes con distribución condicional  $p(\underline{x}|\theta) = \mathcal{M}(\theta, r^2)$  con  $r^2$  conocida, entonces la distribución posterior para  $\theta$  será  $p(\theta|\underline{x}) = \mathcal{M}(\mu', \tau')$  donde

$$\mu' = \frac{r^2 \mu + n \bar{x} \tau^2}{r^2 + n \tau^2} \quad \text{y} \quad \tau' = (1/\tau^2 + n/r^2)^{-1}$$

En particular, la varianza posterior de  $\theta$  está dada por  $(1/\tau^2 + n/r^2)^{-1}$  y, en caso de utilizar una función de pérdida cuadrática para estimar el parámetro, coincide con el riesgo de Bayes.

Por otra parte, la información de Fisher que contiene la muestra respecto a  $\theta$  está dada por (Kendall & Stuart, 1973)

$$\begin{aligned} I_n(\theta) &= nI(\theta) \\ &= -nE_x \left[ \left\{ \frac{\partial}{\partial \theta} \ln p(x|\theta, r) \right\}^2 \right]^{-1} \\ &= -n \left[ E_x \{ -r^2 \} \right]^{-1} \\ &= \frac{n}{r^2} \end{aligned}$$

de manera que la varianza posterior de  $\theta$  se puede escribir como

$$\frac{1}{1/\tau^2 + nI(\theta)} \quad (3.2)$$

En general, encontrar el Riesgo de Bayes es mas complicado de lo que resulta en el caso particular de Normalidad y Tsutakawa propone aproximar la varianza posterior de  $\theta$  tomando en cuenta las siguientes propiedades asintóticas, válidas bajo condiciones de regularidad:

a) Si  $\theta_0$  es el verdadero valor del parámetro, el estimador máximo verosímil de  $\theta$ ,  $\hat{\theta}$ , es asintóticamente Normal con media  $\theta_0$  y varianza  $[I_n(X, \theta_0)]^{-1}$  con  $I_n(X, \theta_0)$  la información contenida en una muestra cuando se aplica  $X = (X_1, X_2, \dots, X_n)$ , el conjunto de dosis seleccionadas.

b) La distribución posterior de  $\theta$  puede ser aproximada por la distribución Normal con media  $\theta$  y varianza  $[L_2(\hat{\theta})]^{-1}$ , donde  $L_2(\hat{\theta})$  es el negativo de la segunda derivada (con respecto a  $\theta$ ) del logaritmo de la función de verosimilitud valuada en  $\hat{\theta}$ .

c) Si  $\theta_0$  es el verdadero valor de  $\theta$ , con probabilidad uno

$$\frac{L_2(\hat{\theta})}{I_n(X, \theta)} \rightarrow 1$$

Estas propiedades sugieren que, al menos para muestras grandes, la distribución posterior de  $\theta$  estará concentrada alrededor de  $\theta_0$  y que  $[I(X, \theta)]^{-1}$  será una aproximación razonable a la varianza posterior de  $\theta$ .

Sin embargo, cuando el tamaño de muestra es finito (como ocurre en la práctica), el estimador máximo verosímil  $\hat{\theta}$  puede distar mucho de  $\theta_0$  (el verdadero valor de  $\theta$ ) e  $[I_n(X, \theta)]^{-1}$  puede resultar ser una pobre aproximación de la varianza posterior de  $\theta$ . Este problema surge cuando el tamaño de muestra es pequeño como para que la aproximación asintótica sea apropiada o bien cuando la distribución apriori domina a la verosimilitud. Otro detalle que conviene notar es que  $[I_n(X, \theta_0)]^{-1}$  no depende de la distribución

apriori para  $\theta$ ,  $\lambda(\theta)$ , y que para cualquier tamaño de muestra fijo  $I_n(X, \theta_0)$  podría ser cercana a cero de manera que  $[I_n(X, \theta_0)]^{-1}$  resultaría no acotada, mientras que si  $\lambda(\theta)$  tiene varianza finita, la varianza posterior de  $\theta$ , para cualquier conjunto de observaciones, sería acotada.

Para evitar estos posibles problemas al usar  $[I_n(X, \theta_0)]^{-1}$  como estimador de la varianza, Tsutakawa considera añadir una constante mayor que cero (denotada por  $B(\lambda)$ ) a  $I_n(X, \theta_0)$  de tal forma que la estimación de la varianza posterior de  $\theta$  quede en términos de  $[I_n(X, \theta_0) + B(\lambda)]^{-1}$ .

Sea de esta forma

$$H(X, \theta_0, \lambda) = \frac{1}{I(X, \theta_0) + B(\lambda)}, \quad B(\lambda) \geq 0$$

la aproximación a la varianza posterior de  $\theta$ . Nótese que  $H(X, \theta_0, \lambda)$  está acotada por el inverso de  $B(\lambda)$ .

En particular, si  $\lambda(\theta)$  es una distribución Normal con varianza  $\tau^2$  y se cuenta con una función de verosimilitud Normal con varianza  $[I_n(X, \theta_0)]^{-1}$ , la estimación de la varianza posterior de  $\theta$  tomando  $B(\lambda) = 1/\tau^2$  sería exactamente la misma que la obtenida en (3.2), es decir

$$H(X, \theta_0, \lambda) = \frac{1}{nI(X, \theta_0) + \frac{1}{\tau^2}}$$

Ahora bien, con respecto al problema de diseño óptimo, Tsutakawa define el mejor experimento en términos de los valores de  $X$  (el conjunto de niveles o dosis a aplicar) que minimice la varianza posterior esperada, es decir, el problema se reduce a encontrar  $X$  tal que



$$E\{H(X, \theta, \lambda)\} = \int H(X, \theta, \lambda) \lambda(\theta) d\theta = G(X | \lambda)$$

sea mínima.

Como ya se mencionó, el autor elige la estructura logística para la distribución de tolerancia  $P(X)$ . La distribución es

$$P(X) = \frac{1}{1 + \exp(\theta - X)} \quad -\infty < X < \infty$$

en particular  $E(X) = \theta$  y  $V(X) = \frac{\pi^2}{3}$ .

Con fines de ilustración, utiliza para modelar  $\lambda(\theta)$ , la apriori de  $\theta$ , una distribución  $M(0, \tau^2)$ . Tomando una observación en cada nivel, la distribución posterior de  $\theta$  dado  $Y = (y_1, y_2, \dots, y_n)$  es

$$p(\theta | X, Y) \propto \lambda(\theta) \prod_{i=1}^n P_i^{y_i} [1 - P_i]^{1 - y_i} \quad y_i = 0, 1 \quad i = 1, \dots, n$$

Para esta situación el autor sugiere considerar a  $B(\lambda)$  de la forma

$$B(\lambda) = \left( \frac{1}{\tau \delta} \right)^2$$

con  $\delta$  una constante positiva elegida de tal forma que refleje la importancia de la distribución apriori sobre la distribución posterior. De acuerdo a los resultados asintóticos mencionados anteriormente, se usaría  $\delta = 1$ . El autor hace notar que si  $\delta < 1$ ,  $(\delta\tau)^2$  sería menor que la varianza apriori,  $\tau^2$  y entonces se estaría dando un énfasis, en términos de una mayor precisión, a la apriori de  $\theta$  que en realidad no tiene. Sin embargo, el autor indica también que valores de  $\delta$  mayores que uno podrían ser de interés para distinguir entre lo que llama, la distribución

apriori para el diseño del experimento de la distribución apriori para inferencia. Con respecto a estas dos distribuciones apriori, puede comentarse que si se cuenta con una distribución muy informativa, es decir, con varianza pequeña, es de esperarse que pocos experimentos proporcionen información útil. Dicho de otro modo, si el conocimiento apriori sobre  $\theta$  es muy preciso, se requeriría de un experimento muy refinado para que pudiera aportar nueva información. Por otra parte, si el conocimiento apriori sobre  $\theta$  es vago, casi cualquier experimento proporcionaría información adicional apreciable y además permitiría que inclusive resultados extremos fueran contemplados e incorporados en la distribución posterior.

Así, cuando el diseño ha sido seleccionado en base a un conocimiento apriori vago sobre  $\theta$ , el autor sugiere utilizar una distribución apriori de inferencia cuya varianza esté "inflada".

Con lo anterior, Tsutakawa espera obtener una distribución posterior con suficiente sensibilidad de manera que si la apriori del diseño contiene información que se contrapone con la información muestral, ésto se refleje en la posterior.

Ahora bien, si  $\delta = \infty$ , es decir,  $B(\lambda) = 0$ , la situación se traduce en una distribución apriori difusa. El autor indica que en este caso, se deben tomar precauciones pues  $G(X|\lambda)$  puede ser infinito cuando  $\lambda(\theta)$  no es Normal. (En particular, para el caso en el que  $\lambda(\theta)$  es una distribución Logística cuya varianza no es estrictamente menor que la varianza de  $P(X)$ , el autor demuestra que  $G(X|\lambda)$  diverge).

Es importante notar que no es posible justificar la dicotomía de las distribuciones apriori propuestas por Tsutakawa pues no hay un proceso de aprendizaje entre ambas. Aparentemente

se trata del mismo estado de conocimiento sobre  $\theta$  modificando la distribución de acuerdo a los intereses que se tengan.

Retomando el problema de minimizar  $G(X|\lambda)$ , el autor menciona que el problema de minimización puede ser numéricamente intratable. Esto es, hay que encontrar  $X=(X_1, X_2, \dots, X_n)$  tal que minimice  $G(X|\lambda)$  y basta con que  $n$  (fijo) sea grande para que el problema se complique pues, por ejemplo, si el número de niveles es del orden de 20, habría que minimizar una función de 20 variables.

Tsutakawa propone una simplificación que consiste en tomar  $k$  (fija) observaciones en cada uno de  $m$  niveles, igualmente espaciados de manera que  $km=n$  (el total de observaciones). Así, dados  $m, k, \tau$  y  $\delta$  el problema se reduce a minimizar una función de una sola variable,  $w$ , que representa el ancho del intervalo entre los niveles. Específicamente, el autor propone la minimización de

$$\begin{aligned} G(w, \tau) &= mk \int \lambda(\theta) H(X, \theta, \lambda) d\theta \\ &= mk \int \lambda(\theta) / \left\{ \sum_{i=1}^{n=mk} I_1(X_i, \theta) + (\delta^2 \tau^2)^{-1} \right\} \\ &= mk \int \lambda(\theta) / \left\{ k \sum_{i=1}^m I_1(X_i, \theta) + (\delta^2 \tau^2)^{-1} \right\} \\ &= m \int \lambda(\theta) / \left\{ \sum_{i=1}^m I_1(X_i, \theta) + (k \delta^2 \tau^2)^{-1} \right\} \end{aligned}$$

Tsutakawa señala que  $G_1$  no es otra cosa que  $G$  multiplicada por  $mk$ , con el fin de poder comparar diseños con tamaños de muestra diferentes. En particular, cuando  $\delta = \omega$ ,  $G_1$  no depende de  $k$ .

Ahora bien, debido al hecho de que los  $m$  niveles son

igualmente espaciados se tiene que

$$X_i = \left\{ i - \frac{(m+1)}{2} \right\} w$$

y dado que  $m, k, \tau$  y  $\lambda$  son fijos, la única variable con respecto a la cual es necesario minimizar es  $w$ , lo cual simplifica enormemente el problema.

El valor  $w^0$  que minimiza  $G_1$  cuando las otras variables están fijas, es denominado por el autor el intervalo óptimo dados  $m, k, \tau$  y  $\lambda$  y denota por  $G_1^0(\tau)$  al valor  $G_1(w^0, \tau)$ .

El autor especifica que, salvo en el caso en que  $m=1$  y  $w^0=0$ , encontrar  $w^0$  requiere de evaluar la integral  $G_1$  por métodos numéricos con precisión suficiente de tal suerte que  $w^0$  pueda ser determinada por un método iterativo.

Tsutakawa propone un análisis exploratorio para estudiar la influencia que los diversos parámetros del problema tienen en la solución. Específicamente plantea las siguientes interrogantes:

- a) Dados  $m, k$  y  $\delta$  ¿cuál es el efecto sobre  $w^0$  y  $G_1^0$  al variar  $\tau$ ?
- b) Dado  $n=mk$ , el total de observaciones, ¿cuál es el efecto de  $m$  sobre  $G_1^0$ ?
- c) ¿Cuál es el efecto de  $\delta$  sobre  $w^0$  y  $G_1^0$ ? En particular, ¿cuándo se debe considerar  $\delta=\infty$ ?
- d) Dados  $\tau, m$  y  $\delta$  ¿cuál es el efecto sobre  $G_1^0$  si se varía  $k$ ?

El autor reporta valores de  $w^0$  y  $G_1^0$  considerando algunos valores específicos para  $\tau, n, \delta$  y  $m$ . Los resultados que reporta son, en términos generales, los siguientes:

Cuando  $\delta=1$  y  $n=20$  ó  $40$ ,  $w^0$  es igual a cero para toda  $\tau$  menor o igual que  $1.2$ . Para valores de  $\tau$  mayores a  $1.2$ ,  $w^0$  se incrementa a medida que  $m$  decrece, observándose que el cambio en  $w^0$  es brusco para  $\tau$  menor que  $2.5$  y a partir de este valor  $w^0$  se incrementa prácticamente en forma lineal. Por otra parte, los valores de  $G_1^0$  coinciden para diferentes valores de  $m$ , variando  $\tau$  hasta  $2.5$ . Los cambios abruptos en  $G_1^0$  para diferentes valores de  $m$ , se presentan cuando  $\tau$  es grande. En todos los casos  $G_1^0$  va disminuyendo a medida que  $m$  crece. El autor sugiere, para los tamaños de muestra que consideró en su análisis, que  $m$  debe ser al menos  $4$  para valores de  $\tau$  entre  $2.5$  y  $4$  y debe de ser  $8$  para valores de  $\tau$  mayores que  $4$ .

En el caso de  $\delta=\infty$ , el comportamiento de  $w^0$  y  $G_1^0$  es similar excepto que ambas variables aumentan su valor mas rápidamente que en los otros casos.

El autor indica que que  $G_1^0$  es monótona creciente con respecto a  $k$  y que este hecho sugiere que la ganancia por observación es mayor cuando las muestras son pequeñas.

Por otra parte, Tsutakawa también analiza la aproximación a la varianza posterior de  $\theta$  mediante la función  $H$  y encuentra que el error en la aproximación se incrementa si la varianza a priori,  $\tau^2$ , se incrementa. Además, para cada pareja de valores  $(m,k)$  considerados en el análisis, el error que se comete al tomar  $\delta=\infty$  es mayor que cuando  $\delta=1$  y éste aumenta a medida que el tamaño muestral disminuye. El autor comenta que la aproximación es razonable cuando  $\tau=1$  y resulta menos satisfactoria a medida que  $\tau$  crece.

El autor discute el diseño de un experimento cuya realización se lleve a cabo, secuencialmente, en dos etapas.

Comenta que en ese caso el diseño de la segunda etapa, no es más que un problema de una sola etapa, usando la distribución posterior obtenida en la primera etapa como distribución apriori para la segunda. Tsutakawa considera que en esta situación se debe elegir el diseño de la primera etapa de tal forma que optimice, en algún sentido, el experimento en forma global.

Aún cuando el autor no lo menciona, es importante recordar que en un procedimiento bayesiano en el cual se incorpora información secuencialmente, por ejemplo, en dos etapas, es equivalente tomar la distribución posterior de la primera etapa como apriori de la segunda, incorporar la información y construir la posterior final (después de involucrar las dos etapas), a trabajar la distribución apriori (de la primera etapa), incorporar (vía teorema de Bayes) toda la información y obtener la distribución final correspondiente. Sin embargo, si el objetivo es minimizar la varianza posterior esperada, ya no necesariamente es equivalente minimizar en forma secuencial la varianza posterior que incorporar toda la información (de las dos etapas) y con la distribución resultante minimizar la varianza esperada posterior.

El procedimiento de Tsutakawa no es estrictamente bayesiano, ya que sin recurrir al Teorema de Bayes para encontrar las varianzas posteriores las aproxima, en cada etapa, mediante la correspondiente función  $H$ . En cualquier caso, el problema de la minimización secuencial aparece. De hecho el autor ejemplifica su procedimiento con algunos casos particulares llegando a la conclusión de que el diseño en dos etapas es mejor que el diseño equivalente en una etapa.

El autor considera el caso en el cual se tienen dos diseños igualmente espaciados, cada uno con  $m$  niveles (o dosis) y  $k$  observaciones por nivel, suponiendo que la distribución

posterior obtenida en la primera etapa es aproximadamente Normal con media  $\theta'$  y varianza  $v$ . Más que encontrar el diseño óptimo (igualmente espaciado) que minimice  $G_1^0(v^{1/2})$ , Tsutakawa propone, en primer término, aproximar  $v^{1/2}$  por

$$t(w, \theta, \tau) = \{k \sum I_1(X_i, \theta) + \tau^2\}^{-1/2}$$

denotando a  $v$  con  $\tau^2$  y donde  $X_i = \{i - \frac{(m+1)}{2}\}w$ , con  $w$  el ancho del intervalo usado en la primera etapa.

En segundo término, propone aproximar la varianza posterior esperada (al final de las dos etapas) por  $G_2(w, \tau)/2mk$  donde

$$G_2(w, \tau) = \int G_1^0(t(w, \theta, \tau)) \phi(\theta | \tau)$$

con  $\phi(\theta | \tau)$  la distribución posterior de la primera etapa. Por último encuentra el valor de  $w$  tal que  $G_2(w, \tau)$  es mínimo.

En un trabajo posterior, Tsutakawa (1980), considera una situación similar a la anterior con la variante de que ahora  $P(X)$  se modela mediante una distribución Logística de dos parámetros definida por

$$P(X) = \frac{1}{1 + \exp\{-(\theta_1 + \theta_2 X)\}} \quad \begin{array}{l} -\infty < X < \infty \\ -\infty < \theta_1 < \infty \\ 0 < \theta_2 < \infty \end{array}$$

con  $\underline{\theta} = (\theta_1, \theta_2)$  desconocido.

Al igual que en el otro caso, el autor propone el diseño de experimentos óptimos, en el sentido de que produzcan una varianza posterior mínima. Sin embargo, para el caso de esta  $P(X)$ ,

la distribución apriori que se asigne a  $\underline{\theta}$  involucrará, ahora, una matriz de varianzas y covarianzas. Bajo los mismos argumentos del trabajo anterior, Tsutakawa propone aproximar la matriz de varianzas y covarianzas posterior de  $\underline{\theta}$  mediante una función matricial H que, en este caso, estará basada en la matriz de información de Fisher y en la matriz de varianzas y covarianzas apriori de  $\underline{\theta}$ . Naturalmente es posible calcular el valor esperado de la matriz H, pero ya no tiene sentido minimizar esta expresión. Alternativamente, el autor considera como objetivo del experimento la estimación de una función real del vector de parámetros  $\underline{\theta}$ . Esta modificación resulta particularmente afortunada en el caso de bioensayos en donde muy a menudo el parámetro de interés es, en realidad, la dosis  $\gamma$ -efectiva,  $X_\gamma$ , esto es, la mínima dosis que produzca al menos una respuesta esperada de magnitud  $\gamma$ . Específicamente, en el caso de que  $P(X)$  sea la Logística con dos parámetros, se tiene que

$$X_\gamma = -\ln\left(\frac{1-\gamma}{\gamma}\right) \left(\frac{1}{\theta_1 + \theta_2}\right)$$

y puesto que la matriz de información de Fisher se puede calcular muy fácilmente bajo transformaciones uno a uno de los parámetros y ya que  $X_\gamma$  es tal que  $P(X_\gamma) = \gamma$ , el autor reparametriza a  $P(X)$  como sigue

$$X_\gamma = \frac{(\ell - \theta_1)}{\theta_2}, \quad \eta = \theta_2^2 \quad \text{con} \quad \ell = \ln\left(\frac{\gamma}{1-\gamma}\right) = \text{logit } \gamma$$

De manera que

$$P_{X_\gamma, \eta}(X) = \frac{1}{1 + c \exp\{-(X - X_\gamma)\eta^{1/2}\}}$$

$$\text{con } c = \frac{1-\gamma}{\gamma}$$



Debe notarse que para hacer inferencias sobre la dosis  $\gamma$ -letal basta hacer inferencias sobre  $X_\gamma$ .

Con esta reparametrización, Tsutakawa desarrolla el siguiente procedimiento:

Tomando  $N$  observaciones en un conjunto de niveles  $X=(X_1, X_2, \dots, X_m)$  distribuidas como  $\underline{n}=(n_1, n_2, \dots, n_m)$ , con  $n_i$  observaciones independientes en el  $i$ -ésimo nivel; la matriz de información de Fisher con respecto a  $(X_\gamma, \eta)$  está dada por

$$\Lambda = \Lambda(X, \underline{n}, X_\gamma, \eta) = \begin{bmatrix} \eta \sum n_i P_i Q_i & -\sum (X_i - X_\gamma) n_i P_i Q_i / 2 \\ -\sum (X_i - X_\gamma) n_i P_i Q_i / 2 & \sum (X_i - X_\gamma)^2 n_i P_i Q_i / 4\eta \end{bmatrix}$$

en donde  $P_i = 1 - Q_i = P_{X_\gamma, \eta}(X)$ ,  $i=1, \dots, m$  y  $\sum_{i=1}^m n_i = N$

Suponiendo que la distribución apriori de  $(X_\gamma, \eta)$  tiene matriz de varianzas y covarianzas  $\Sigma$ , Tsutakawa considera la matriz

$$\Sigma_2 = (\Lambda + \Sigma^{-1})^{-1}$$

y propone como criterio de diseño encontrar el punto  $(X, \underline{n})$  que minimice la función

$$G(X, \underline{n}) = \iint H(X, \underline{n}, X_\gamma, \eta) P(X_\gamma, \eta) dX_\gamma d\eta$$

en donde  $P(X_\gamma, \eta)$  es la densidad apriori de  $(X_\gamma, \eta)$  y

$$H(X, \underline{\eta}, X_{\gamma}, \eta) = \frac{(\lambda_{22} + \tau_{22})}{(\lambda_{11} + \tau_{11})(\lambda_{22} + \tau_{22}) - (\lambda_{12} + \tau_{12})^2}$$

con  $\lambda_{ij}$  y  $\tau_{ij}$  los elementos  $ij$ -ésimos de  $\Lambda$  y  $\Sigma^{-1}$  respectivamente.

Es interesante notar que el autor no hace mas que generalizar el procedimiento planteado en su trabajo anterior tomando ahora a  $B(\lambda)$  como  $\Sigma^{-1}$ .

Al igual que en el trabajo anterior, obtiene soluciones numéricas suponiendo que se toman  $k$  observaciones en cada uno de  $m$  niveles igualmente espaciados de la forma

$$X_i = Z + \left\{ i - \frac{(m+1)}{2} \right\} w \quad i=1, 2, \dots, m$$

con  $Z$  el punto medio del intervalo.

La distribución inicial para  $(X_{\gamma}, \eta)$  que considera el autor está dada por

$$P(X_{\gamma}, \eta) \propto \exp\left\{-\frac{(X_{\gamma} - \mu)^2}{2\tau^2}\right\} \exp\left\{\frac{1}{2} \nu \eta\right\} \eta^{(\nu-2)/2}$$

que coincide con una distribución  $\mathcal{N}(\mu, \tau^2)$  para  $X_{\gamma}$  y una distribución Gamma con parámetros  $\nu/2, \nu/2$  para  $\eta$ , suponiendo que  $\eta$  y  $X_{\gamma}$  son independientes.

De la expresión para la distribución inicial puede notarse que  $\mu$  es un parámetro de localización y  $\nu$  un parámetro de escala. El autor sugiere tomar  $\mu=0$  y  $\nu=1$ , es decir, el caso estándar, y dejar a  $(\tau, \nu)$  por especificar. En caso de que  $\mu \neq 0$  y/o

$\nu \neq 1$ , la solución óptima se puede encontrar determinando la solución  $(X_1, X_2, \dots, X_m)$  cuando  $\mu=0$ ,  $\nu=1$ ,  $\tau^* = \tau/\sqrt{\nu}$  y estableciendo los niveles óptimos  $(X_1^*, X_2^*, \dots, X_m^*)$  mediante las relaciones

$$X_i^* = \mu + X_i \sqrt{\nu}$$

La varianza posterior esperada de  $X^*$  resulta ser  $\nu G$  con  $G$  la varianza posterior esperada de  $X$ .

Como en su trabajo anterior Tsutakawa considera minimizar la función

$$G_1(Z, w) = mk G(X, \underline{n})$$

dados  $m, k, \tau$  y  $\nu$ , con el fin de poder comparar diseños con diferentes tamaños muestrales.

De acuerdo a sus resultados, el autor comenta que para diferentes valores de  $\gamma, \nu, n$  y  $k$ ,  $G_1^0$  se incrementa en forma similar con respecto a  $\tau$ , siendo este incremento más notorio para valores pequeños de  $\nu$  y valores grandes de  $\gamma$ . Si  $\tau < 2.5$  y  $N$  está fija,  $G_1^0$  se comporta prácticamente constante con respecto a  $m$ . Sin embargo si  $\tau > 3$ , a medida que  $m$  crece  $G_1^0$  disminuye. Además, Tsutakawa menciona que el valor óptimo de  $Z$  depende principalmente de  $\tau$  y prácticamente no depende de  $\nu, m$  y  $k$ . Por otra parte, el valor óptimo de  $w$  ( $w^0$ ) depende muy poco de  $\nu$  y  $\gamma$  pero sí depende de  $m$ . El autor indica que la poca dependencia de  $Z$  y  $w$  en  $\nu$  se debe a que  $X_\gamma$  y  $\eta$  se supusieron independientes a priori.

El trabajo de Tsutakawa aunado al de Freeman parecen sugerir que el problema de analizar resultados de un bioensayo es complicado aún en el caso de tener una estructura para la función

$P(X)$ . Parecería razonable esperar una complicación aún mayor si se renuncia a ese elemento estructural. Sin embargo, Ramsey (1972), ha publicado algunos resultados más generales que son muy interesantes.

Dadas  $M$  dosis experimentales ordenadas  $X_1 \leq X_2 \leq \dots \leq X_M$  y  $n_i$  ensayos independientes en  $X_i$ ,  $i = 1, 2, \dots, M$  se tiene que

$$p(Y|X) = \prod_{i=1}^M \prod_{j=1}^{n_i} [P(X_i)]^{y_{ij}} [1 - P(X_i)]^{1-y_{ij}} \quad \begin{matrix} i=1,2,\dots,M \\ j=1,2,\dots,n_i \end{matrix}$$

$$= \prod_{i=1}^M [P(X_i)]^{s_i} [1 - P(X_i)]^{n_i - s_i}$$

con  $s_i = \sum_{j=1}^{n_i} y_{ij}$

Si no se adopta una estructura particular para  $P(X)$  excepto la hipótesis de monotonía ( $P(X_i) \leq P(X_j)$   $i \leq j$ ), el problema consiste en encontrar una función  $\hat{P}(X)$  no decreciente que aproxime a  $P(X)$ .

En el caso  $M=1$ , si se hace  $P(X_1) \equiv P$ , se tiene que

$$p(Y|X) = P^{s_1} (1 - P)^{n_1 - s_1}$$

Esto es, la verosimilitud se reduce a la de una distribución Bernoulli con parámetro  $P$ , en donde el problema es estimar  $P$ . Esta situación ha sido ampliamente tratada en la literatura (DeGroot 1970 y Winkler 1972, por ejemplo) utilizando como densidad inicial para  $P$  la conjugada natural Beta ( $\alpha, \beta$ ).

El caso  $M > 1$  podría reducirse al anterior si se efectúan las inferencias para cada  $P(X_i)$   $i=1,2,\dots,M$  por separado utilizando sólo la información  $y_{i1}, y_{i2}, \dots, y_{in_i}$  para estimar  $P(X_i)$ .

Este tratamiento equivale a suponer a priori que  $P(X_i)$  es independiente de  $P(X_j)$  si  $i \neq j$ . Sin embargo, esta hipótesis no es razonable puesto que se cuenta con la restricción estructural

$$P(X_{i-1}) \leq P(X_i) \leq P(X_{i+1}) \quad i=1,2,\dots,M$$

con  $P(X_0) \equiv 0$  y  $P(X_{M+1}) \equiv 1$

Así, es necesario considerar una distribución inicial  $p(P(X_1), \dots, P(X_M))$  que incorpore esta restricción. Puede verificarse fácilmente que si  $P(X_i)$  dados  $P(X_{i-1}), P(X_{i-2}), \dots, P(X_1)$  depende sólo de  $P(X_{i-1})$  entonces

$$p(P(X_1), P(X_2), \dots, P(X_M)) = \prod_{i=1}^M p(P(X_i) | P(X_{i-1}))$$

con  $P(X_0) \equiv 0$

En estas condiciones, basta con elegir la forma de las funciones  $p(P(X_i) | P(X_{i-1}))$ , de manera que describa el conocimiento sobre  $P(X_i)$  dado que se conoce  $P(X_{i-1})$  preferentemente a través de una función que sea matemáticamente conveniente.

$P(X_i)$  es una probabilidad y por tanto  $0 \leq P(X_i) \leq 1$ . Si adicionalmente se conoce el valor de  $P(X_{i-1})$ , por la restricción estructural se sabe que  $P(X_{i-1}) \leq P(X_i) \leq 1$ .

Por otra parte, las ventajas técnicas de utilizar una densidad Beta para describir el conocimiento sobre una probabilidad han sido ampliamente establecidas. Resulta entonces natural la proposición de Ramsey de asignar a  $P(X_i) | P(X_{i-1})$  una densidad Beta truncada al intervalo  $[P(X_{i-1}), 1]$ ; ésto es

$$p(P(X_1)|P(X_{1-1})) \propto [P(X_1)]^{\alpha_1} [1 - P(X_1)]^{b_1} \quad \alpha_1, b_1 > 0$$

$$P(X_1) \in [P(X_{1-1}), 1]$$

de manera que

$$p(P(X_1), P(X_2), \dots, P(X_M)) \propto \prod_{i=1}^M [P(X_i)]^{\alpha_i} [1 - P(X_i)]^{b_i}$$

Se puede verificar que esta densidad conjunta puede reexpresarse como

$$p(P(X_1), P(X_2), \dots, P(X_M)) = D(\alpha_1, \alpha_2, \dots, \alpha_{M+1}, \beta)$$

$$\text{donde } D(\alpha_1, \alpha_2, \dots, \alpha_{M+1}, \beta) \propto \left[ \prod_{i=1}^{M+1} [P(X_i) - P(X_{i-1})]^{\alpha_i} \right]^\beta$$

para  $0 \equiv P(X_0) \leq P(X_1) \leq \dots \leq P(X_M) \leq P(X_{M+1}) \equiv 1$  y  $\beta, \alpha_i > 0$   
 $i=1, \dots, M+1$

Esta distribución ha sido estudiada en la literatura (Wilks 1962, por ejemplo) y recibe el nombre de distribución Dirichlet (en las diferencias) o Dirichlet ordenada (en las  $P(X_i)$ 's).

Esta distribución tiene, entre otras, las siguientes propiedades:

i) Todas las distribuciones marginales pertenecen a la misma familia, en particular  $p(P(X_1), \dots, P(X_{M-1})) = D(\alpha'_1, \dots, \alpha'_M, \beta)$  donde  $\alpha'_i = \alpha_i$   $i=1, 2, \dots, M-1$ ,  $\alpha'_M = \alpha_M + \alpha_{M+1}$  y  $p(P(X_i))$  es una Beta con parámetros  $(A_i \beta, (1-A_i) \beta)$  en  $[0, 1]$  donde  $A_i = \sum_{j=1}^i \alpha_j$ .

ii) El punto  $(A_1, A_2, \dots, A_M)$  es la moda de la distribución conjunta.

iii)  $\beta$  puede considerarse un parámetro de precisión en el sentido de que si  $\beta \rightarrow 0$  la distribución tiende a una Uniforme en  $[0,1]$  mientras que si  $\beta \rightarrow \infty$  la distribución se concentra alrededor de la moda.

Así, para proceder a la inferencia bayesiana basta seleccionar los valores de  $\alpha_1, \alpha_2, \dots, \alpha_{M+1}$  y  $\beta$ . La propiedad ii) junto con la definición de  $A_i$  sugiere que si a priori se cree que el valor más probable de  $P(X_i)$  es  $P^*(X_i)$   $i=1, 2, \dots, M$  entonces se utilice

$$\alpha_i = P^*(X_i) - P(X_{i-1}) \quad (P^*(X_0) \equiv 0, P^*(X_{M+1}) \equiv 1)$$

lo cual garantiza que  $P^*(X_i)$   $i=1, 2, \dots, M$  es la moda a priori.

$\beta$  por su parte, en opinión de Ramsey debe elegirse de manera que refleje la credibilidad en la asignación de las  $\alpha_i$ 's ( $\beta \rightarrow 0$  poca credibilidad,  $\beta \rightarrow \infty$  alta credibilidad).

En particular, si  $\beta \rightarrow \infty$  la distribución final estará dominada por la inicial mientras que si  $\beta \rightarrow 0$  lo es por la verosimilitud.

Así,

$$p(P(X_1), \dots, P(X_M) | Y) \propto \prod_{i=1}^M [P(X_i)]^{s_i} \left[1 - P(X_i)\right]^{n_i - s_i} \prod_{i=1}^{M+1} \left[ [P(X_i) - P(X_{i-1})]^{\alpha_i} \right]^\beta$$

Esta distribución debe ser utilizada para efectuar

inferencias sobre la función  $P(X)$ . Es importante hacer notar que en todo el proceso sólo se ha hecho referencia a  $P(X_1), \dots, P(X_M)$ , un conjunto de valores de  $P(X)$  y no a toda la función. Adicionalmente, se puede verificar que la distribución inicial no es conjugada, ésto es, la final no pertenece a la familia Dirichlet.

De cualquier manera, con la final obtenida se puede utilizar una función de pérdida adecuada para, en particular, obtener una estimación de  $P(X_i)$   $i = 1, 2, \dots, M$ .

La elección de la función de pérdida depende de los objetivos específicos del estudio y de las consecuencias que se esperan al elegir un estimador particular y una opción es la pérdida cuadrática. Cuando esta elección no es muy clara una alternativa para estimar  $P(X_i)$   $i=1, 2, \dots, M$  consiste en utilizar la moda posterior. Ramsey verifica que el máximo de la distribución final existe, es único y se puede obtener por diferenciación

$$\frac{d(P(X_1), \dots, P(X_M) | Y)}{dP(X_i)} = \frac{d}{dP(X_i)} \left\{ \prod_{i=1}^M [P(X_i)]^{s_i} [1 - P(X_i)]^{n_i - s_i} \right. \\ \left. \prod_{i=1}^{M+1} [P(X_i) - P(X_{i-1})]^{\alpha_i} \right\}$$

Derivando e igualando a cero se obtiene el siguiente sistema de ecuaciones no lineales en  $\hat{P}(X_i)$ ,  $i=1, 2, \dots, M$



$$\beta \left[ \frac{P^*(X_{i+1}) - P^*(X_i)}{\hat{P}(X_{i+1}) - \hat{P}(X_i)} - \frac{P^*(X_i) - P^*(X_{i-1})}{\hat{P}(X_i) - \hat{P}(X_{i-1})} \right] =$$

$$= \frac{n_i}{\hat{P}(X_i) (1 - \hat{P}(X_i))} - \left[ \frac{s_i}{n_i} - P(\hat{X}_i) \right] \quad (3.3)$$

$i=1,2,\dots,M$  ;  $P(X_0) \equiv 0$  y  $P(X_{M+1}) \equiv 1$

que se puede resolver por métodos numéricos.

Un punto que falta por considerar es la estimación de  $P(Z)$  para un valor  $Z=X_i$   $i=1,2,\dots,M$ .

Si es de interés estimar  $P(Z)$ , se debe incluir el correspondiente conocimiento sobre esta cantidad en una distribución inicial como (3.3). Con la convención  $X_0 \equiv 0$  y  $X_{M+1} \equiv \infty$  es claro que  $X_i \leq Z \leq X_{M+1}$  para alguna  $i=1,2,\dots,M$ . La correspondiente inicial queda dada por

$$p(P(X_1), P(X_2), \dots, P(X_i), P(Z), P(X_{i+1}), \dots, P(X_{M+1})) =$$

$$\left\{ \prod_{j=1}^i [P(X_j) - P(X_{j-1})]^{\alpha_j \beta} [P(Z) - P(X_i)]^{\alpha_Z \beta} [P(X_{i+1}) - P(Z)]^{\alpha_{i+1}} \right\}$$

$$\left\{ \prod_{j=i+2}^{M+1} [P(X_j) - P(X_{j+1})]^{\alpha_j \beta} \right\}$$

Al combinar esta expresión con la verosimilitud se obtiene la distribución final conjunta para las dosis experimentales y para la dosis no experimental  $Z$ .

Derivando e igualando a cero, de nuevo se obtiene la expresión para la moda posterior. Es importante notar que dado que

$Z$  es una dosis no experimental no aparece en la verosimilitud y por tanto el resultado no es el que ya se había obtenido en (3.3).  
 Explícitamente, se tiene

$$\frac{P^*(X_{i+1}) - P^*(Z)}{\hat{P}(X_{i+1}) - \hat{P}(Z)} = \frac{P^*(Z) - P^*(X_i)}{\hat{P}(Z) - \hat{P}(X_i)}$$

de donde

$$[\hat{P}(Z) - \hat{P}(X_i)][P^*(X_{i+1}) - P^*(Z)] = [\hat{P}(X_{i+1}) - \hat{P}(Z)][P^*(Z) - P^*(X_i)]$$

$$\rightarrow \hat{P}(Z)P^*(X_{i+1}) - \hat{P}(Z)P^*(Z) - \hat{P}(X_i)P^*(X_{i+1}) + \hat{P}(X_i)P^*(Z) = \\ \hat{P}(X_{i+1})P^*(Z) - \hat{P}(X_{i+1})P^*(X_i) - \hat{P}(Z)P^*(Z) + \hat{P}(Z)P^*(X_i)$$

$$\rightarrow \hat{P}(Z)[P^*(X_{i+1}) - P^*(Z)] + \hat{P}(X_i)[P^*(Z) - P^*(X_{i+1})] = \\ \hat{P}(Z)[P^*(X_i) - P^*(Z)] + \hat{P}(X_{i+1})[P^*(Z) - P^*(X_i)]$$

$$\rightarrow \hat{P}(Z)[P^*(X_{i+1}) - P^*(Z)] - \hat{P}(Z)[P^*(X_i) - P^*(Z)] = \\ \hat{P}(X_{i+1})[P^*(Z) - P^*(X_i)] - \hat{P}(X_i)[\hat{P}(Z) - P^*(X_{i+1})]$$

$$\rightarrow \hat{P}(Z)[P^*(X_{i+1}) - P^*(Z) - P^*(X_i) + P^*(Z)] = \\ \hat{P}(X_{i+1})[P^*(Z) - P^*(X_i)] + \hat{P}(X_i)[P^*(X_{i+1}) - P^*(Z)]$$

1.º

$$\hat{P}(Z) = \frac{\hat{P}(X_{i+1})[P^*(Z) - P^*(X_i)] + \hat{P}(X_i)[P^*(X_{i+1}) - P^*(Z)]}{P^*(X_{i+1}) - P^*(X_i)}$$

Este resultado es extremadamente interesante puesto que se puede interpretar como una interpolación lineal en  $\hat{P}(Z)$ .

Otro objetivo que se persigue en un problema de

Bioensayo, como ya se ha mencionado, es la estimación de la dosis  $X$  tal que  $P(X)=\gamma$  con un valor de  $\gamma$  prestablecido. En este punto, Ramsey se aleja del procedimiento Bayesiano usual que consistiría en asignar una distribución inicial para  $X$  y encontrar su correspondiente distribución final, y propone el siguiente tratamiento:

Si  $\hat{P}(X_i)=\gamma$  para alguna  $i=1,2,\dots,M$  entonces utilícese  $\hat{X} = X_i$ , en caso contrario, debe ocurrir que  $X_i < X < X_{i+1}$  para alguna  $i=1,2,\dots,M$  y  $X$  puede entonces tratarse como una dosis no experimental. Así para  $X$  vale la expresión

$$\frac{P^*(X_{i+1}) - P^*(X)}{\hat{P}(X_{i+1}) - \hat{P}(X)} = \frac{P^*(X) - P^*(X_i)}{\hat{P}(X) - \hat{P}(X_i)}$$

Ramsey sugiere que si se busca  $X$  tal que  $P(X)=\gamma$  se exija que  $\hat{P}(\hat{X})=\gamma$  de modo que

$$P^*(\hat{X}) = P^*(X_i) + \frac{P^*(X_{i+1}) - P^*(X_i)}{\gamma - \hat{P}(X_i)} [\hat{P}(X_{i+1}) - \hat{P}(X_i)]$$

de esta manera  $X$  queda estimado por el valor unívocamente definido,  $\hat{X}$ , que satisface esta expresión.

Finalmente, Ramsey analiza algunos ejemplos simulados para mostrar el empleo de su método y obtiene conclusiones interesantes. Los ejemplos sugieren, en opinión del autor, que el mejor diseño para estimar la dosis media es aquel que asigna una observación por dosis (ésto es,  $n_i=1$ ,  $i=1,2,\dots,M$ ) situación que no puede ser tratada por los métodos logit y probit que, desde una

perspectiva frecuentista, también se han empleado en problemas relacionados. Esta conclusión la obtiene después de aplicar su propuesta de análisis a varios diseños con distinto número de observaciones por dosis, en los cuales se conocía el verdadero valor de la dosis media. Analizando el sesgo, la desviación estándar y la raíz cuadrada del error cuadrático medio del estimador de la dosis media en cada diseño, concluye que entre menos observaciones por dosis se tenga, mejores son los estimadores.

El autor comenta que también se puede usar un método basado en frecuencias relativas, como regresión isotónica (Brunk, 1970), pero se presenta el inconveniente de que no se acompaña de un método de interpolación. En su opinión, el suavizamiento de una curva que está asociado al valor de  $\beta$ , y el enfoque bayesiano del procedimiento de inferencia, parece producir estimaciones superiores a las de regresión isotónica aún cuando en la distribución inicial se cuente con poca información.

Por otra parte, si se cuenta con información previa, resultado de estudios experimentales anteriores, la elección de los parámetros en la distribución inicial debe ser consistente con tal información. Por supuesto, el caso más simple en el tratamiento bayesiano se obtiene cuando la distribución inicial pertenece a una familia conjugada. En el procedimiento de Ramsey, sin embargo, debe notarse que no es ésta la situación y por tanto el autor propone utilizar la siguiente regla: "Si un experimento previo basado en  $n_0$  observaciones por dosis arrojó una moda posterior de  $P^*(X)$ , entonces adóptese  $P^*(X)$  como la moda inicial y use  $\beta = n_0$ ", la cual, dice Ramsey, resulta ser una forma conservadora pero técnicamente conveniente de incorporar el conocimiento inicial.

Una objeción a este trabajo es que, como ya se indicó, la distribución inicial conjugada no es simplemente que reduzca los cálculos, sino que ella misma es una distribución final y por tanto permite incorporar toda la evidencia de experimentos previos. En un sentido estricto, si todos los resultados de un experimento previo se conocen, entonces no debe de usarse la distribución inicial propuesta por el autor, sino que debería de usarse la distribución final de esos experimentos que a su vez conduciría, al combinarse con los nuevos resultados experimentales, a la nueva distribución final correspondiente. Sin embargo, la proposición de Ramsey constituye una alternativa simple y atractiva que parece proporcionar estimaciones razonables.

Es importante notar que aún cuando los resultados obtenidos por Ramsey están basados en métodos bayesianos, la forma de estimar la dosis no observacional  $\hat{X}$  y  $P(\hat{X})$  no es estrictamente bayesiana sino el resultado de aplicar interpolaciones lineales.

**ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA**

## **CAPITULO 4**

### **CONCLUSIONES**

## CONCLUSIONES

Los métodos presentados en este trabajo forman parte de un conjunto mayor de técnicas para analizar datos experimentales provenientes de bioensayo.

En los procedimientos que aquí se discuten, vale la pena mencionar que tanto desde el punto de vista frecuentista como bayesiano se sugiere modelar la función potencia  $P(X)$  mediante alguna función de distribución, exceptuando el trabajo de Ramsey (1972).

En los métodos presentados dentro del enfoque frecuentista, la mayor parte de ellos tienden a reconstruir la situación que se presenta en Regresión Lineal, sin embargo, la naturaleza dicotómica de la variable de respuesta introduce violaciones especialmente a los supuestos de normalidad y homocedasticidad. En tales condiciones el ajuste de los modelos suele llevarse a cabo mediante aplicaciones sucesivas del procedimiento de mínimos cuadrados ponderados, lo que en general implica que se pierda la simplicidad en el ajuste del modelo, que es característica de los modelos lineales de Regresión.

Cabe mencionar que las aproximaciones a  $P(X)$  que se discuten, en su mayoría, surgen a partir de situaciones prácticas y no de consideraciones teóricas. Por otra parte, en términos generales, las aproximaciones producen los mismos resultados por lo que no existe una regla que especifique cuál transformación usar. Sin embargo, dado el fácil manejo de las distribuciones Normal y Logística y la gran variedad de situaciones que describen, son las que, en la práctica, más se recomiendan.

Desde la perspectiva bayesiana, el procedimiento de modelado es similar. Sin embargo, la introducción de distribuciones de probabilidad sobre los parámetros conduce a un análisis diferente al de los métodos frecuentistas.

En varios casos el empleo de formas conjugadas tiene aplicaciones restringidas en el sentido de que la correspondiente función de distribución posterior para el parámetro de interés, tiene que ser aproximada.

Específicamente, es importante recalcar que en el procedimiento bayesiano propuesto por Ramsey, la distribución posterior no pertenece a alguna familia conjugada y además la estimación de una dosis no observacional,  $\hat{X}$  y de su correspondiente valor  $P(\hat{X})$ , no se realiza formalmente utilizando los principios bayesianos sino recurriendo a interpolaciones lineales. En cuanto a la propuesta de Freeman, resulta muy restrictiva pues resuelve casos muy particulares que difícilmente pueden extenderse. Por otra parte, en los resultados presentados por Tsutakawa, vale la pena insistir en la dificultad para justificar la existencia de lo que él llama la distribución apriori de inferencia y la apriori del diseño ya que no queda suficientemente claro el manejo que se hace sobre el conocimiento de  $\theta$ .

Tsutakawa propone soluciones para casos particulares, al suponer un número igual de observaciones en cada uno de los niveles a los que considera igualmente espaciados.

En cuanto a la función  $H$  que propone como aproximación a la varianza posterior, sería interesante considerar en ella un factor que involucre el costo asociado al número de piezas de información. Suponiendo que cada pieza de información cuesta lo



mismo, por ejemplo c. tomar n piezas llevaría a un costo final igual a nc. Así, la función H podría plantearse como

$$H(\mathbf{X}, \theta_0, \lambda) = \frac{1}{I(\mathbf{X}, \theta_0) + B(\lambda)} + nc \quad B(\lambda) \geq 0$$

e intentar minimizarla.

Por último, resultaría de interés tratar alguna otra aproximación para  $P(\mathbf{X})$  distinta a la distribución Logística.

Aunque el planteamiento al problema de Bioensayo ha sido presentado en forma muy general, se dan soluciones a casos muy particulares y que sólo son tratables por métodos numéricos.

Para finalizar, convendría explorar si esta situación puede resolverse usando otras técnicas estadísticas tales como Bondad de Ajuste o métodos No Paramétricos.

## BIBLIOGRAFIA

## BIBLIOGRAFIA

- BERKSON, J. (1944). Application of the logistic function to bio-assay. *J. Amer. Statist. Assoc.* 39, 357-368.
- BERNARDO, J. M. (1979). Reference posterior distributions for bayesian inference. *J. Roy. Statist. Soc. B* 41, 113-147. (Con discusión).
- BERNARDO, J. M. (1981). *Bioestadística. Una Perspectiva Bayesiana*. Barcelona: Vines-Vives.
- BLISS, C. I. (1934a). The method of probits. *Science* 79, 38-39.
- BLISS, C. I. (1934b). The method of probits - a correction. *Science* 79, 409-410.
- BLISS, C. I. (1944). The U.S.P. collaborative cat assay for digitalis. *J. Amer. Pharm. Assoc.* 33, 225-245.
- BOX, G. E. P. & COX, D. R. (1964). An analysis of transformations. *J. Roy. Statist. Soc. B* 26, 211-252. (Con discusión).
- BRUNK, H. D. (1970). Estimation of isotonic regressions. En *Nonparametric Techniques in Statistical Inference* (Puri ed.) 177-194, Cambridge: Cambridge University Press.
- De FINETTI, B. (1937). La prévision, ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* 7, 1-68. Reeditado en inglés en *Studies in Subjective Probability* (Kyburg & Smokler eds.) 93-158, New York: Wiley, 1964.
- De FINETTI, B. (1962). Does it make sense to speak of good probability appraisers?. En *The scientist speculates: an anthology of partly baked ideas*. (Good, ed.) 357-364. New York: Basic Books.
- De GROOT, M. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.
- DIXON, W. J. & MOOD, A. M. (1948). A methods of obtaining an

- analysing sensitivity data. *J. Amer. Statist. Assoc.* 43, 109-126.
- DRAPER, N. R. & SMITH, H. (1981). *Applied Regression Analysis*. New York: Wiley.
- FINNEY, D. J. (1965). The meaning of bioassay. *Biometrics* 21, 785-798.
- FINNEY, D. J. (1971). *Probit Analysis* (3rd. ed.). Cambridge: The University Press.
- FINNEY, D. J. (1978). *Statistical Methods in Biological Assay*. London: Griffin & Company Ltd.
- FREEMAN, P. R. (1970). Optimal bayesian sequential estimation of the median effective dose. *Biometrika* 57, 79-89.
- GREENBERG, B. G. (1982). Biostatistics. En *Encyclopedia of Statistical Sciences Vol.1* (Kotz & Johnson eds.) 251-263 New York: Wiley.
- JEFFREYS, H. (1946). An invariant form for the prior probability in estimation problems. *Proc. Roy. Soc. London A* 186, 453-461.
- KENDALL, M. G. & STUART, A. (1973). *The Advanced Theory of Statistics. Vol.2*. London: Griffin & Company Ltd.
- KNUDSEN, L. F. & CURTIS, J. M. (1947). The use of the angular transformation in biological assays. *J. Amer. Statist. Assoc.* 42, 282-296.
- KOLMOGOROV, A. N. (1933/56). *Foundations of the Theory of Probability*. 2a. Ed. New York: Chelsea Publishing Co.
- LINDLEY, D. V. (1965). *An Introduction to Probability and Statistics from a Bayesian Viewpoint*. (2 vol.). Cambridge: Cambridge University Press.
- LINDLEY, D. V. (1971). *Bayesian Statistics, a Review*. Reg. Conf. Ser. Appl. Math. 2. Philadelphia: SIAM.
- LINDLEY, D. V. (1977). *Principios de la Teoría de la Decisión*. Barcelona: Vives-Vives.

- MENDOZA, M. (1986). On the statistical analysis of biological assay. *Rep. Técnico Depto. Bioestadística*. Universidad de Valencia.
- MIELKE, P. W. (1986). Quantit Analysis. En *Encyclopedia of Statistical Sciences Vol.7* (Kotz & Johnson eds.) 433-435. New York: Wiley.
- MOOD, A.G., GRAYBILL, F.A. & BOES, D.C. (1974). *Introduction to the Theory of Statistics*. Tokyo: McGraw-Hill.
- RAIFFA, H. & SCHLAIFER, R. (1961). *Applied Statistical Decision Theory*. Massachusetts: MIT Press.
- RAMSEY, F. L. (1972). A bayesian approach to bioassay. *Biometrics* 28, 841-858.
- RAMSEY, F. P. (1926). Truth and Probability. Reeditado en *Studies in Subjective Probability* (Kyburg & Smokler eds.) 61-92, New York: Wiley, 1964.
- SAVAGE, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.
- SAVAGE, L. J. (1971). The elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc.* 66, 783-801.
- TSUTAKAWA, R. K. (1972). Design of experiments for bioassay. *J. Amer. Statist. Assoc.* 67, 584-590.
- TSUTAKAWA, R. K. (1980). Selection of the dose levels for estimating a percentage point of a logistic quantal response curve. *Appl. Statist.* 29, 25-33.
- TSUTAKAWA, R. K. (1982). Statistical methods in bioassay. En *Encyclopedia of Statistical Sciences Vol.1* (Kotz & Johnson eds.) 236-243. New York: Wiley.
- URBAN, F. M. (1909). Die psychophysischen Massmethoden als Grundlagen empirischer Messungen. *Arch. ges Psychol.* 15, 261-355.
- URBAN, F. M. (1910). Die psychophysischen Massmethoden als

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

- Grundlagen empirischer Messungen. (continued) *Arch. ges Psychol.* 16, 168-227.
- WILKS, S. S. (1962). *Mathematical Statistics*. New York: Wiley.
- WILSON, E. B. & WOCESTER, J. (1943a). The determination of L.D.50 and its sampling error in bio-assay. *Proc. Natn. Acad. Sci.* 29, 79-85.
- WILSON, E. B. & WOCESTER, J. (1943b). The determination of L.D.50 and its sampling error in bio-assay. II. *Proc. Natn. Acad. Sci.* 29, 114-120.
- WILSON, E. B. & WOCESTER, J. (1943c). Bio-assay on a general curve. *Proc. Natn. Acad. Sci.* 29, 150-154.
- WILSON, E. B. & WOCESTER, J. (1943d). The determination of L.D.50 and its sampling error in bio-assay. III. *Proc. Natn. Acad. Sci.* 29, 257-262.
- WINKLER, R. L. (1967). The assesment of prior distributions in Bayesian analysis. *J. Amer. Statist. Assoc.* 62, 776-800.
- WINKLER, R. L. (1972). *Introduction to Bayesian Inference and Decision*. New York: Holt, Rinehart & Winston.