

Estadística Clásica e Inferencia

1.- Introducción

En la actualidad, la inferencia estadística clásica, desarrollada en el primer tercio del siglo XX, da respuesta suficiente para tratar de forma adecuada la problemática suscitada.

Tomar el valor de una medida (población), pasa por realizar unas concretas determinaciones (muestra) y a partir de ella, dar el resultado final de la medida y su incertidumbre.

La incertidumbre es un parámetro asociado al resultado de una medición que caracteriza la dispersión de los valores que pueden atribuirse a la medición.

De esta forma, la muestra y las técnicas de muestreo se convierten en el fundamento de la inferencia estadística clásica.

2.- Estadística Clásica

La estadística tiene muchos significados, siendo una técnica que debe establecer una guía para determinar situaciones de incertidumbre.

La información y la incertidumbre participan de forma decisiva. Es, por lo tanto, la estadística una ligazón entre pasado y futuro, de especial trascendencia en la mayor parte de los procesos.

Todo procedimiento estadístico que emplea información para obtener conocimiento, está en el ámbito de la inferencia y cuando hay que determinar acciones, se aborda la denominada decisión estadística.

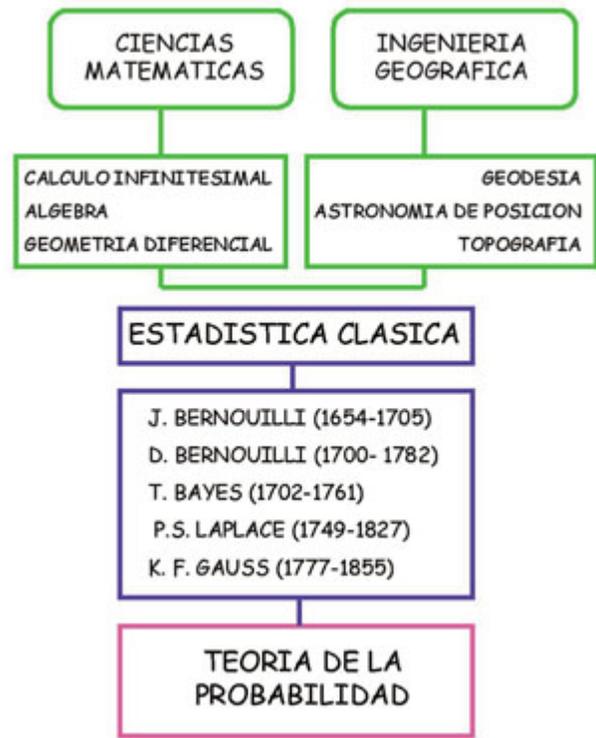


Figura 1.- Marco de la estadística clásica

Los primeros nombres ligados a esta etapa significativa son los Bernoulli. Originarios de los Países Bajos, huyeron a Suiza por la persecución que en el reinado de Felipe II hicieron los españoles a los no católicos.

3.- La Teoría de la Probabilidad

En el primer tercio del siglo XX irrumpen con fuerza en el ámbito científico-matemático con corrientes que tratan de configurar un marco propio: la teoría de la probabilidad.

Crear una axiomática propia y establecer las bases para difundir una nueva estadística, ocupó un amplio período de tiempo donde científicos de gran renombre fueron plasmando sus teorías.

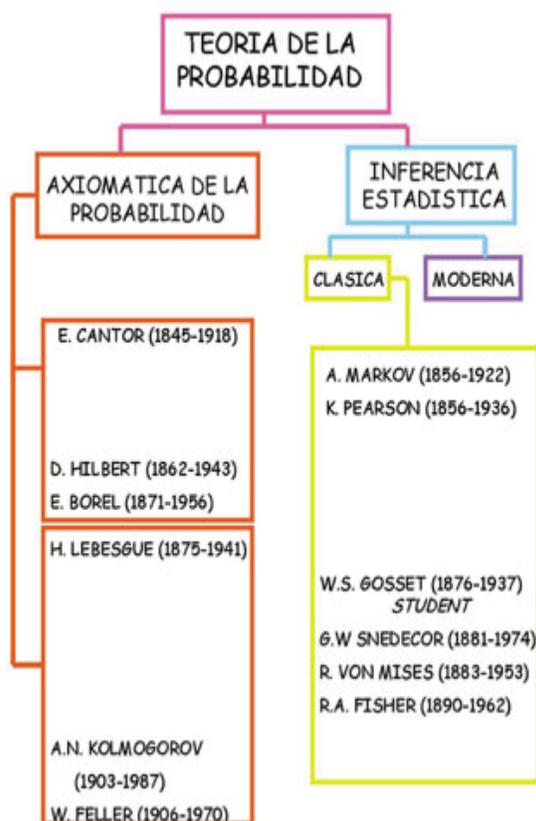


Figura N 2.- Flujo significativo

Es necesario destacar los tres grandes bloques de tratamiento en esta etapa tan interesante que trata de dar contenidos a una nueva forma en el ámbito del conocimiento matemático.

Preparación de una base topológica y de las estructuras asociadas necesarias para establecer una nueva axiomática (Cantor, Hilbert, Borel).

Inicio y consolidación de la axiomática de la teoría de la probabilidad (Lebesgue, Kolmogorov, Feller).

Desarrollo de la inferencia estadística clásica y su aplicación a Ingeniería, Industria (Markov, Pearson, Gosset, Snedecor, Von Mises, Fisher).

Los tres bloques fueron desarrollándose a la vez, consiguiendo en tan solo cincuenta años desgajar unos contenidos que han formado una estructura propia en el ámbito de las ciencias matemáticas.

4.- Inferencia Estadística Moderna

A partir de mediados del siglo XX, tras la implantación de la denominada **inferencia clásica**, surgen investigadores que intentan reforzar nuevas formas, tratando de implementar aplicaciones dirigidas hacia las ciencias sociales y ámbito económico.

La **moderna inferencia** tiene dos pilares fundamentales: la teoría de la decisión y los métodos bayesianos.

A mediados de 1.800 ya se distinguen aspectos como utilidad marginal y utilidad total, renovándose el primitivo concepto de Bernoulli.

Con posterioridad Von Neumann y Oscar Morgenstern, hacia 1.945, desarrollan la moderna teoría probabilística de la utilidad, definiendo una axiomática común para la probabilidad y para la utilidad.

En 1.954 aparece la obra de Savage, **iniciador de la línea bayesiana**, en el marco de la teoría de la utilidad.

En los últimos años el empleo de la teoría ha sido predictivo, aplicado en muchas ocasiones a las ciencias sociales y económicas, destacando la utilidad colectiva de Arrow (problema de bienestar social) y hace poco tiempo Black y Rethenberg.

Recientemente trabajos de McCrimmon (1968) han contrastado la axiomática de Savage, empleando experimentación adecuada, obteniendo interesantes resultados.

Fundamentalmente la **inferencia estadística moderna** sigue dos sentidos que, aunque convergentes, tienen enfoque diferenciado:

Teoría bayesiana. Admite conocimientos a priori, así como observaciones nuevas. Las a priori son modificadas por las verosimilitudes.

Teoría de la decisión. Punto de partida en los trabajos de Wald (1.950). Pretende establecer una teoría, unas reglas de acción en situación de incertidumbre o riesgo. El valor de cualquier decisión es medido por su pérdida.

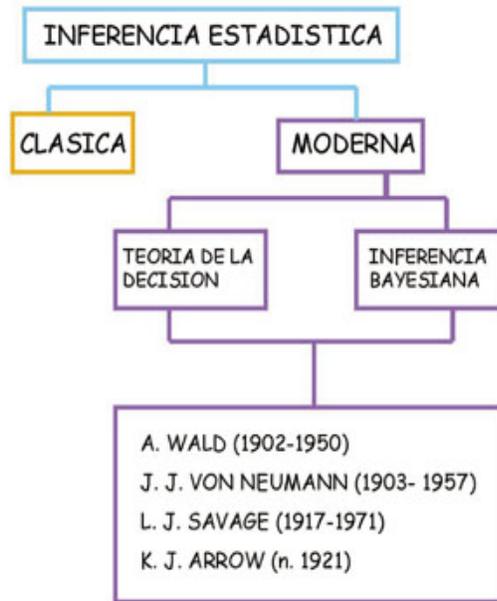


Figura 3.- Inferencia estadística moderna

5.- Inferencia Estadística Robusta

La inferencia estadística clásica establece mecanismos de actuación ampliamente desarrollados como la estimación genérica y los contrastes de hipótesis.

Ambos constituyen una parte esencial del trabajo habitual en la estadística actual. En ellos impera el dogma de la normalidad, quedando invalidadas las observaciones discordantes con esta norma.

Etapas de transición han intentado establecer mecanismos fundados en incluir los datos anómalos y diseñar funciones de distribución mixtas que modelen ambos comportamientos.

Los métodos robustos son complementarios de los métodos estadísticos clásicos y configuran una teoría más afinada en aspectos tan interesantes como la parte de estimación y de contraste.

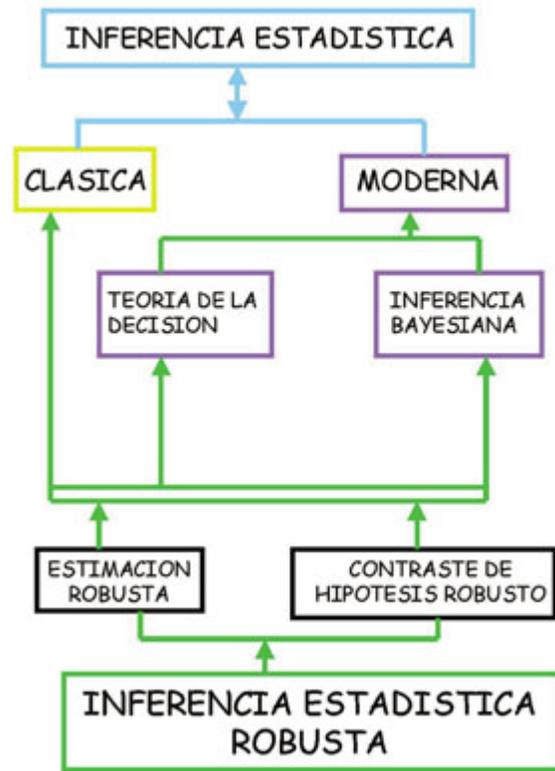


Figura 4.- Ámbito de la inferencia estadística robusta

El término robusto fue establecido por Box (1.953), tratando de forzar hacia una modelización más completa que la establecida. Trabajos posteriores fueron reforzando la teoría, destacando:

Quenouille (1.956), con técnicas que permiten reducir el sesgo y establecer nuevos entornos para la estimación.

Anscombel (1.960), estimulando la investigación teórica y experimental en datos extremos.

Huber (1.964), diseñando los procedimientos estadísticos robustos.

Hampel (1.968), introduciendo el uso de curvas de influencia para analizar la sensibilidad de los estimadores.

Jaekel (1.971) y Berger (1.976), sobre aspectos de admisibilidad de los estimadores.

En la actualidad, y una vez establecidos los elementos teóricos necesarios, es preciso aplicar el potencial de la nueva herramienta a las múltiples utilidades que diferentes áreas tienen establecidas.

BAYES, Thomas (1702 – 1761)

Teólogo y matemático británico. Hijo de uno de los primeros ministros presbiterianos ordenados en Inglaterra. También fue ministro y asistió a su padre.

Fue el primero en utilizar la probabilidad inductivamente, estableciendo un fundamento matemático para la inferencia probabilística.

En un ensayo póstumo quedó patente el intento de establecer las probabilidades tras el conocimiento de causas. Laplace, a los once años de su muerte, dio a conocer definitivamente sus revolucionarias teorías.

La estadística bayesiana le debe su nombre al trabajo pionero del reverendo Thomas Bayes titulado *“An Essay towards solving a Problem in the Doctrine of Chances”* (*Una composición para solucionar un problema en la doctrina de las oportunidades*), publicado póstumamente en 1764 en la *“Philosophical Transactions of the Royal Society of London”*.

El artículo fue enviado a la Real Sociedad de Londres por Richard Price, amigo de Bayes, en 1763, quien escribió:

“Yo ahora le mando un ensayo que he encontrado entre los papeles de nuestro fallecido amigo Thomas Bayes, y el cual, en mi opinión, tiene un gran mérito, y bien merece ser preservado... En una introducción que él ha escrito para este ensayo, él dice, que su objetivo en un principio fue, descubrir un método por el cual se pueda juzgar la probabilidad de que un evento tenga que ocurrir bajo circunstancias dadas, y bajo la suposición de que nada es conocido sobre dicho evento, salvo que, bajo las mismas circunstancias, éste ha ocurrido un cierto número de veces y fallado otro tanto... Cualquiera persona juiciosa verá que el problema aquí mencionado no es de ninguna manera una

simple especulación producto de la curiosidad, sino un problema que se necesita resolver para contar con un fundamento seguro para todos nuestros razonamientos concernientes a hechos pasados y a lo que probablemente ocurra de ahí en adelante... El propósito a mí me parece es, mostrar qué razones nosotros tenemos para creer que en la constitución de las cosas existen leyes fijas de acuerdo con las cuales las cosas pasan, y que, por lo tanto, el funcionamiento del mundo debe ser el efecto de la sabiduría y el poder de una causa inteligente, y así, confirmar el argumento tomado desde las causas finales para la existencia de la deidad.”

Aunque la obra de Thomas Bayes data ya de hace más de dos siglos, la estadística bayesiana es relativamente nueva, y actualmente ostenta un gran desarrollo aunque no ajeno a también grandes controversias.

Los métodos bayesianos, con una interpretación diferente del concepto de probabilidad, constituyen una alternativa a la estadística tradicional centrada en el contraste de hipótesis, denominada por contraposición estadística frecuentista, y están siendo motivo actual de debate.

En esencia se diferencian en que incorporan información externa al estudio para con ella y los propios datos observados estimar una distribución de probabilidad para la magnitud -efecto- que se está investigando.

Se denomina método bayesiano por basarse originalmente en el teorema de Bayes, publicación póstuma de Thomas Bayes en 1764, que en esencia nos permite, si conocemos la probabilidad de que ocurra un suceso, modificar su valor cuando disponemos de nueva información.

6.- Teorema de Bayes

Recordando la terminología y el teorema de Bayes. Vamos a llamar $P(A)$ a la probabilidad de que ocurra el suceso A .

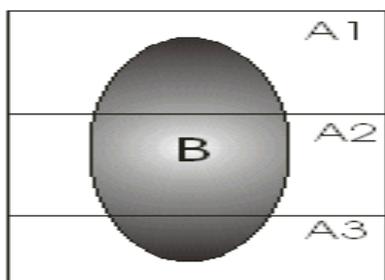
$P(AB)$ a la probabilidad de que ocurran los sucesos A y B (ambos).

$P(A / B)$ a la probabilidad de que ocurra A cuando sabemos que ha ocurrido B (se denomina **probabilidad condicionada**).

La probabilidad de que ocurra A y B es igual a la probabilidad de B multiplicada por la probabilidad de A condicionada a que haya ocurrido B.

$$P(AB) = P(B) \times P(A / B) = P(A) \times P(B / A)$$

Por simetría es obvio que se cumple la tercera igualdad.



Si tenemos un conjunto de posibles sucesos A_i ($A_1 \dots A_n$), mutuamente excluyentes (no puede ocurrir dos de ellos a la vez) y que constituyen todas las posibles situaciones (o lo que es lo mismo $P(A_1) + P(A_2) + \dots + P(A_n) = 1$, el que ocurra alguno de los sucesos A tiene probabilidad 1, suceso seguro). Lo representamos gráficamente en la figura. El cuadrado corresponde a todas las situaciones posibles, que en este caso pueden dividirse en tres: A_1, A_2, A_3 . El suceso B se puede producir en cualquiera de las tres situaciones.

Si reescribimos ahora la anterior ecuación por ejemplo para A_1 tenemos

$$P(A_1 \cdot B) = P(A_1 / B) \times P(B) = P(B / A_1) \times P(A_1)$$

Con un poco de álgebra tenemos

$$P(A_1 / B) = \frac{P(B / A_1) \cdot P(A_1)}{P(B)}$$

$$P(B) = P(B \cdot A_1) + P(B \cdot A_2) + P(B \cdot A_3) = \sum P(B \cdot A_i) = \sum P(B / A_i) \cdot P(A_i)$$

$$P(A_1 / B) = \frac{P(B / A_1) \cdot P(A_1)}{\sum P(B / A_i) \cdot P(A_i)}$$

que constituye el famoso **teorema de Bayes**. Para cualquiera de las otras situaciones (A_2, A_3) la fórmula es similar.

7.- Metodología bayesiana

En la metodología estadística clásica -frecuentista- se calcula la probabilidad de observar un resultado suponiendo que la realidad sea de una manera determinada (hipótesis nula), sin embargo en la práctica necesitamos los conocimientos para tomar decisiones, y lo que realmente nos interesa es conocer la probabilidad de que las cosas sean de una manera determinada dados los datos (condicionado a..) que hemos observado.

Esta es la diferencia que radica en el enfoque bayesiano. En el caso de las pruebas diagnósticas lo que nos interesa en la práctica es el valor predictivo, positivo o negativo, de la prueba no la sensibilidad o especificidad de éstas.

Aunque estamos habituados a la presentación de los métodos bayesianos con sucesos binarios o dicotómicos (enfermo o sano), también son aplicables cuando los resultados son continuos (por ejemplo proporción de pacientes que sobreviven).

En el análisis estadístico clásico para evaluar por ejemplo la eficacia de un nuevo tratamiento frente al tratamiento anterior se utiliza exclusivamente la información obtenida en el estudio.

El teorema de Bayes es bastante importante, pues relaciona las probabilidades “**a priori**” $P(A_i)$ con las probabilidades “**aposteriori**” $P(A_i / B)$ (Probabilidad de A_i después que ha ocurrido el evento B).

Supongamos que se está comparando la tasa de mortalidad. Para utilizar la terminología habitual vamos a llamar a esa magnitud que interesa calcular θ .

Lo primero que hay que determinar es la distribución de probabilidad de esa magnitud con la información externa de la que se dispone, es lo que se denomina **probabilidad a priori** y vamos a representar como $P(\theta)$. Seguidamente se cuantifica

la información que aportan los datos observados en nuestro estudio mediante lo que se denomina **función de verosimilitud** (likelihood), que denotaremos como $P(\text{Datos} / \theta)$. La verosimilitud representa la probabilidad de los datos observados para cualquier valor del parámetro θ . Podemos ahora utilizar el teorema de Bayes para actualizar el valor a priori $P(\theta)$ a la luz de los datos obtenidos y calcular $P(\theta / \text{Datos})$, la denominada función de **probabilidad a posteriori**, es decir cómo de probables son los diferentes valores posibles de θ una vez obtenidos nuestros datos. Según el teorema de Bayes tenemos que

$$P(\theta / \text{Datos}) \propto P(\theta) \cdot P(\text{Datos} / \theta)$$

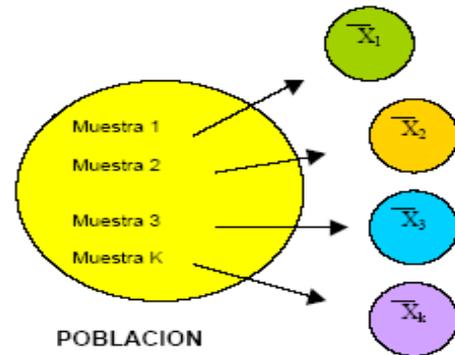
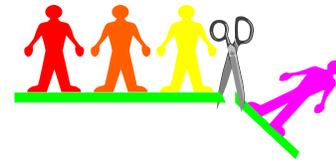
donde el símbolo \propto indica que el lado de la izquierda es proporcional al lado de la derecha, es decir que son iguales salvo por un término constante (el denominador del teorema de Bayes) que no depende del parámetro de interés θ .

Así pues los resultados se expresan como una función de la probabilidad a posteriori de los diferentes valores de θ .

8.- Inferencia Estadística

Inferencia Estadística es la ciencia que permite hacer conclusiones acerca de una población a partir de la información de una muestra. Esta definición deja algunas preguntas por responder tales como:

- ¿Qué es una población?
- ¿Qué es una muestra?
- ¿Cómo se debe extraer la muestra?
- ¿Cómo se relacionan la muestra y la población?

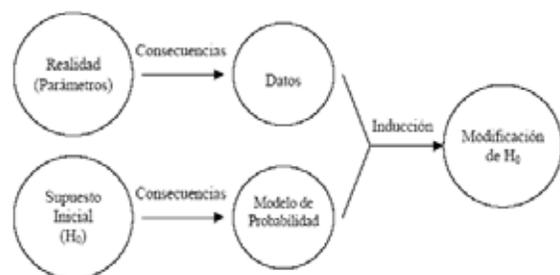


Una respuesta a esta última pregunta es “a través de probabilidades” (condicionales). En pruebas de hipótesis clásicas se definen dos tipos de error cuyas probabilidades están dadas por:

α = Probabilidad de rechazar H_0 / H_0 es verdadera.

β = Probabilidad de aceptar H_0 / H_0 es falsa.

El proceso de prueba de hipótesis consiste en evaluar cuán probable es la muestra, dada una población desconocida pero definida a partir de una hipótesis nula. Este proceso se puede esquematizar de la siguiente manera:



1. Se tiene una “realidad” desconocida la cual se desea estudiar (parámetros).
2. La realidad desconocida es no observable pero sus consecuencias sí lo son (datos).
3. Aunque la realidad es desconocida, se tiene un supuesto inicial acerca de ella (hipótesis nula).
4. Esta hipótesis inicial conduce, por un proceso de deducción, a ciertas consecuencias necesarias

(modelo de probabilidad) que pueden ser comparadas con los datos.

5. Si las consecuencias de la hipótesis nula y los datos no concuerdan, la discrepancia puede conducir, por un proceso de inducción, al rechazo de la hipótesis nula.

La pregunta ahora es, cómo evaluar esa discrepancia. La respuesta es el “valor de probabilidad”.

El valor de probabilidad (p) mide la probabilidad de obtener datos al menos tan discordantes como los obtenidos en la muestra bajo el modelo de probabilidad supuesto.

Si el valor de probabilidad es alto, los datos obtenidos son muy probables bajo el modelo de probabilidad supuesto (poco discordantes con las consecuencias de H_0) y por lo tanto se concluye en que no hay suficiente evidencia para rechazar H_0 .

Si el valor de probabilidad es bajo, los datos obtenidos son poco probables bajo el modelo de probabilidad supuesto (muy discordantes con las consecuencias de H_0) y por lo tanto se concluye en que sí hay suficiente evidencia para rechazar H_0 .

9.- Inferencia Bayesiana

Un punto importante en la definición clásica de inferencia es que el parámetro θ , el cual es desconocido, es tratado como constante en vez que como variable. Esta es la idea fundamental de la teoría clásica, pero conduce a ciertos problemas de interpretación.

Si $[0.2, 0.3]$ es un intervalo del 95% de confianza para θ , sería cómodo decir que hay un 95% de probabilidad de que θ esté en el intervalo. Sin embargo, esto es incongruente con la idea de que θ no es aleatorio. Dado que θ es fijo, solo existen dos opciones: θ está dentro o está fuera del intervalo. El único elemento aleatorio en este modelo es y , por lo que la interpretación correcta del intervalo consistiría en decir que, si se repite el procedimiento muchas veces, entonces a la larga, el 95% de los intervalos contendrían a θ .

Toda inferencia basada en la estadística clásica es forzada a tener este tipo de interpretación

frecuencial, aunque sin embargo, nosotros solo contamos con un intervalo para interpretar.

El marco teórico en el cual se desarrolla la inferencia bayesiana es idéntico al de la teoría clásica.

Se tiene un parámetro poblacional θ sobre el cual se desea hacer inferencias y se tiene un modelo de probabilidad $p(y|\theta)$ el cual determina la probabilidad de los datos observados y bajo diferentes valores de θ . La diferencia fundamental entre la teoría clásica y la bayesiana está en que θ es tratado como una cantidad aleatoria. Así, la inferencia bayesiana se basa en $p(\theta|y)$ en vez de $p(y|\theta)$, esto es, en la distribución de probabilidades del parámetro dados los datos.

La inferencia bayesiana, se puede resumir como el proceso de ajustar un modelo de probabilidad a un conjunto de datos y resumir los resultados mediante una distribución de probabilidades para los parámetros del modelo y para cantidades desconocidas pero observables tales como predicciones para nuevas observaciones.

La característica esencial de los métodos bayesianos está en su uso explícito de probabilidades para cuantificar la incertidumbre en inferencias basadas en el análisis estadístico de los datos. Esto permite un manejo mucho más natural e intuitivo de la inferencia, salvando por ejemplo el problema de la interpretación frecuencial de los resultados. Sin embargo, para hacer uso de un enfoque bayesiano, es necesario especificar una distribución de probabilidades a priori $p(\theta)$, la cual representa el conocimiento que se tiene sobre la distribución de θ previo a la obtención de los datos.

Esta noción de una distribución a priori para el parámetro constituye el centro del pensamiento bayesiano y, dependiendo de si se es un defensor o un opositor a esta metodología, su principal ventaja sobre la teoría clásica o su mayor vulnerabilidad.

10.- Características de la Aproximación Bayesiana

De acuerdo con O'Hagan (1994), se pueden identificar cuatro aspectos fundamentales que caracterizan la aproximación bayesiana a la inferencia estadística:

- **Información a Priori.** Todos los problemas son únicos y tienen su propio contexto. De tal contexto se deriva información a priori, y es la formulación y uso de esta información a priori la que diferencia la inferencia bayesiana de la estadística clásica.

- **Probabilidad Subjetiva.** La estadística bayesiana formaliza la noción de que todas las probabilidades son subjetivas, dependiendo de las creencias individuales y la información disponible. Así, el análisis bayesiano resulta personal, único de acuerdo con las creencias individuales de cada uno.

- **Auto consistente.** Al tratar al parámetro θ como aleatorio, la inferencia bayesiana se basa completamente en la teoría de la probabilidad. Esto tiene muchas ventajas y significa que toda inferencia puede ser tratada en términos de declaraciones probabilísticas para θ .

- **No “adhockery”.** Debido a que la inferencia clásica no puede hacer declaraciones probabilísticas acerca de θ , varios criterios son desarrollados para juzgar si un estimador particular es en algún sentido “bueno”. Esto ha conducido a una proliferación de procedimientos, frecuentemente en conflicto unos con otros. La inferencia bayesiana deja de lado esta tendencia a inventar criterios ad hoc para juzgar y comparar estimadores al basarse exclusivamente en la distribución posterior para expresar en términos exclusivamente probabilísticos toda inferencia referente al parámetro.

11.- Objeciones a la Inferencia Bayesiana.

La principal objeción a la inferencia bayesiana, es que las conclusiones dependen de la selección específica de la distribución a priori. Aunque para otros esto es lo interesante de la aproximación bayesiana, este es un debate aún no cerrado. Sin embargo, antes de dejar esta característica, se debe señalar que inclusive en inferencia clásica, y además en investigación científica en general, estos conocimientos a priori son utilizados implícitamente. Así por ejemplo, el conocimiento a priori es utilizado para formular un modelo de verosimilitud apropiado. En pruebas de hipótesis, las creencias a priori acerca de la plausibilidad de una hipótesis son frecuentemente utilizadas para ajustar el nivel de significancia de la prueba. Así,

si se cree que los datos pueden conducir al rechazo de la hipótesis, esto se puede ajustar escogiendo un nivel de significancia bastante alto. En este sentido entonces, la inferencia bayesiana formaliza la incorporación de la información a priori, la cual es incorporada frecuentemente “debajo de la mesa” en el análisis clásico.

12.- Ampliación de Inferencia bayesiana: distribuciones a priori y a posteriori

Por lo comentado se conocen técnicas y procedimientos de la denominada “inferencia estadística clásica” donde, en los casos en que se exponían métodos basados en parámetros poblacionales, se suponía que estos eran desconocidos pero no aleatorios.

La diferencia fundamental que plantea el “**enfoque inferencial bayesiano**” radica en que considera el parámetro desconocido θ como variable aleatoria,

a la que se asigna, por tanto, una determinada distribución de probabilidad, que recibe el nombre de **distribución a priori** y se representa por $g(\theta)$.

La información suministrada por una muestra aleatoria simple \mathbf{X} bajo la óptica bayesiana, puede cambiar la idea del comportamiento probabilístico que se tuviera sobre el parámetro θ y, en este

sentido, es posible aceptar que exista una **distribución a posteriori** del parámetro θ ,

representa por $g(\theta/X)$, donde se recoge la

modificación de la distribución de probabilidad de dicho parámetro θ cuando se dispone de la información muestral \mathbf{X} .

Si la distribución a priori del parámetro θ se postula

de tipo discreto, aplicando el teorema de Bayes pueden determinarse las correspondientes probabilidades a posteriori para cada uno de los valores que tome el parámetro θ . En efecto

$$P(\theta = \theta_i / X) = \frac{P(X / \theta = \theta_i)P(\theta = \theta_i)}{\sum_{i=1}^K P(X / \theta = \theta_i)P(\theta = \theta_i)}$$

Si la distribución a priori $g(\theta)$ es continua, y la población de donde se extrae la muestra también, puede obtenerse la distribución a posteriori del parámetro θ teniendo en cuenta que:

$f(X; \theta)$ representa la función de densidad conjunta del vector aleatorio muestral \mathbf{X} y del parámetro θ .

$f(X / \theta)$ es la función de densidad de la muestra \mathbf{X} para un θ dado.

$f(X)$ es la distribución marginal de la muestra \mathbf{X}
Como se sabe

$$f(X / \theta) = \frac{f(X; \theta)}{g(\theta)}$$

$$g(\theta / X) = \frac{f(X; \theta)}{f(X)}$$

donde

$$\begin{aligned} f(X; \theta) &= f(X / \theta)g(\theta) \\ f(X; \theta) &= g(\theta / X)f(X) \end{aligned}$$

por lo que $f(X / \theta)g(\theta) = g(\theta / X)f(X)$
y la distribución a posteriori del parámetro θ resulta

$$g(\theta / X) = \frac{f(X / \theta)g(\theta)}{f(X)}$$

obteniéndose la densidad marginal $f(X)$ como

$$f(X) = \int_{\Theta} f(X; \theta)d\theta = \int_{\Theta} f(X / \theta)g(\theta)d\theta$$

llegándose, finalmente, a que

$$g(\theta / X) = \frac{f(X / \theta)g(\theta)}{\int_{\Theta} f(X / \theta)g(\theta)d\theta}$$

versión continua del Teorema de Bayes.

13.- Estimacion de Bayes.

El **estimador de Bayes** es la función de decisión, establecida sobre el espacio paramétrico Θ , que produzca el menor riesgo esperado.

El estimador será $\hat{\theta} = d^*(X)$ tal que minimice la expresión

$$B(d) = \int_{\Theta} \int_X l[\theta; d(X)]f(X / \theta)g(\theta)d\theta d\mathbf{X}$$

Existe un procedimiento operativo para determinar la decisión de Bayes sin tener que describir todo el espacio muestral (y, por tanto, trabajar en él).

En el epígrafe anterior establecimos la distribución a posteriori del parámetro θ , calculada como

$$g(\theta / X) = \frac{f(X / \theta)g(\theta)}{f(X)}$$

por consiguiente, el riesgo de Bayes queda.

$$\begin{aligned} B(d) &= \int_{\Theta} \int_X l[\theta; d(X)]f(X / \theta)g(\theta)d\theta d\mathbf{X} = \\ &= \int_{\Theta} \int_X l[\theta; d(X)]g(\theta / X)f(X)d\theta d\mathbf{X} = \\ &= \int_X f(X) \left[\int_{\Theta} l[\theta; d(X)]g(\theta / X)d\theta \right] d\mathbf{X} = \\ &= \int_X f(X) E_{\theta / X} [l(\theta; d(X))] d\mathbf{X} \end{aligned}$$

siendo $E_{\theta / X} \{ l[\theta; d(X)] \}$, riesgo de Bayes a posteriori, es decir,

$$E_{\theta / X} \{ l[\theta; d(X)] \} = \int_{\Theta} l[\theta; d(X)]g(\theta / X)d(\theta)$$

Dado que la función de decisión $d(X)$ que minimiza

$B(d)$ es la misma que hace mínima la expresión (1),

pues $d(X)$ no depende de $f(X)$, puede utilizarse

esta última para determinar el estimador de Bayes, de manera que dicho estimador será aquel que verifique

$$\frac{\partial E_{\Theta/X} [l(\theta; d(X))] }{\partial d(X)} =$$

$$= \Psi[\theta; d(X)] = 0$$

$$\left[\frac{\partial^2 E_{\Theta/X} [l(\theta; d(X))] }{\partial d^2(X)} \right]_{d(X)=d^*(X)} > 0$$

Referencias Bibliográficas

YAÑEZ DE DIEGO, I.: *Teoría de muestras*. UNED. Madrid, 1.991.

INFANTE MACIAS, R.: *Teoría de la decisión*. UNED. Madrid, 1.992.

CASTILLO RON, E.: *Introducción a la estadística aplicada*. Ed. el autor. Santander, 1.993.

FELLER, W.: *Introducción a la Teoría de probabilidades y sus aplicaciones*. Volumen 1. Editorial Limusa. México, 1.996.

GARCIA PEREZ, A.: *Inferencia estadística robusta*. UNED. Primer Curso sobre métodos de la teoría de la probabilidad y la inferencia estadística. Avila, Otoño 1.997.

GOMEZ VILLEGAS, M.A.: *?R.A. Fisher: el inicio del análisis multivariante?. 100cias@uned*. N?3. 2.000. Pág.: 51-55.

Diversas consultas en Internet y en Enciclopedias.

*Un barco no debería navegar con una sola ancla,
ni la vida con una sola esperanza.*

Epicteto