13
2ej

UNIVERSIDAD ANAHUAC

ESCUELA DE ACTUARIA

CON ESTUDIOS INCORPORADOS A LA U.N.A.M.

EL PROBLEMA DE LA PARADA OPTIMA EN PRUEBAS SECUENCIALES BAYESIANAS

TESIS QUE PARA OBTENER EL TITULO DE

A C T U A R I O

PRESENTA

E R N E S T O H. M O N R O Y Y U R R I E T A

MEXICO, D.F. 1986

"TEORIA DE DECISIONES ESTADISTICAS: EL PROBLEMA DE LA PARADA OPTIMA EN PRUEBAS SECUENCIALES BAYESIANAS"

INDICE

I N T R O D U C C I O N

Se puede mencionar que en general, la estadística matemática se aboca a resolver problemas en los cuales las observaciones de algún fenómeno contienen elementos de variabilidad fuera del control del observador, y además fuera de cualquier otro control que se pudiera establecer; es decir, son observaciones de procesos estocásticos.

Una de las principales preocupaciones de la estadística consiste en la estimación o inferencia de parámetros desconocidos, los cuales puedan ayudar a determinar las verdaderas características del fenómeno en estudio.

El proceso de construcción de modelos probabilísticos a través de la experimentación, la observación y el registro metódico de los resultados, es un ejemplo de la inferencia estadística; la cual puede llevarse a cabo de muy diversas formas -algunas de las más comunes están desarrolladas en la teoría clásica o de Neyman-Pearson; con el uso de las pruebas de hipótesis.

Además de las inferencias, el estadístico está interesado en establecer criterios que le permitan conocer los grados de confianza que pueda tener en las inferencias y que también le brinden la oportunidad de comparar varios resultados, obtenidos de diversas formas, con el objeto de decidir cuales son los mas adecuados y por qué lo son. Estos criterios se pueden obtener a través del empleo de técnicas sobre pruebas de hipótesis o teorías de decisiones.

Sin embargo, para llevar a cabo inferencias y pruebas de hipótesis, es necesario que el estadístico determine, a priori, las principales características tanto de muestreo como de la técnica que va a utilizar. Esto es, el estadístico debe contar con un 'Plan Muestral' por medio del cual queden completamente determinados el tipo de muestra a estudiar, el número de observaciones que se desean tomar, la técnica para el procedimiento inferencial, el control de los errores* y, en general, la estructura del estudio a realizar.

---

* Error tipo I: rechazar una hipótesis cuando ésta es verdadera.

Error tipo II: aceptar una hipótesis cuando ésta es falsa.

En esta tesis se pretende presentar las pruebas secuenciales bayesianas como una técnica alternativa de solución a los problemas de la inferencia estadística y las pruebas de hipótesis, a través del uso de nuevos elementos no considerados en la estadística clásica.

De acuerdo con la teoría de Neyman-Pearson, en el momento en que se realizan las pruebas de hipótesis, el estadístico está interesado en minimizar el error tipo II dada una cota superior fija al error tipo I; sea esta cota igual a τ; esto se debe a que, en este punto de vista clásico, se considera más importante evitar el rechazo de una hipótesis cuando ésta es verdadera, que evitar aceptarla cuando ésta es falsa; por lo que, con base en este razonamiento, el estadístico, finalmente, concluye, nada más, si en la muestra tiene la evidencia suficiente o no para rechazar una hipótesis.

Es decir, en su decisión no se determina si tiene o no la suficiente evidencia para aceptar una hipótesis. En contraste con esto las pruebas secuenciales, bajo un punto de vista clásico, le permiten al estadístico controlar ambos tipos de error con cotas τ, β respectivamente; y como consecuencia de esto, poder decidir si en su muestra existe o no la suficiente evidencia para aceptar una hipótesis o rechazarla.

Esta teoría de las pruebas secuenciales se ve enriquecida con la estadística bayesiana por medio de la incorporación de probabilidades a priori y de probabilidades subjetivas; y por elementos de la teoría de decisiones estadísticas, que son importantes para la optimización que el estadístico debe hacer sobre su plan muestral; tales elementos son: el riesgo, la pérdida y la función de decisión.

La técnica de las pruebas secuenciales bayesianas establece criterios que indican cuál es el mejor plan muestral a través de la resolución del problema de la parada óptima, tema central del presente trabajo, cuyo objetivo primordial es: formular reglas de decisión sólidas, con base en las cuales poder determinar en qué momento se debe detener el muestreo, y tomar una decisión que garantice tener un riesgo estadístico, o pérdida esperada, mínimo.

Esto es, con base en las pruebas secuenciales bayesianas, el problema de la parada óptima pretende proveer al estadístico de los elementos necesarios para determinar cuál es el mejor plan muestral, tomando en cuenta el riesgo y la pérdida.

Esto brinda la posibilidad de reducir costos por muestreo en muchos casos, y conocer, si en algunos otros, es mejor tomar una decisión sin llevar a cabo ninguna observación.

Es, por todas las razones mencionadas, que esta técnica se presenta más atractiva al estadístico y le brinda un campo de decisión más amplio y pleno de elementos importantes, como el riesgo y la pérdida, que le permiten fundamentar en cimientos más sólidos su decisión.

Como se notará en el transcurso de ésta tesis, la metodología que se utiliza pretende brindar no sólo una nueva perspectiva de solución para problemas estadísticos sino que también una nueva filosofía para la óptima aplicación y utilización de los recursos económicos para la resolución de éstos problemas.

Por otro lado, se debe subrayar que ésta filosofía es cimiento fundamental para la presentación, en ésta tesis, de una metodología que proporciona una alternativa para el estadístico interesado en optimizar tanto los resultados como los recursos que se aplican en la solución de problemas de decisión.

En el primer capítulo de la tesis se describen las principales propiedades de las pruebas secuenciales desde un punto de vista clásico; teoría que se concretiza por medio de un ejemplo en la última sección del mismo capítulo.

El segundo capítulo expone los conceptos básicos de la estadística bayesiana y la teoría de decisiones; para, en el capítulo III presentar formalmente las pruebas secuenciales bayesianas y el problema de la parada óptima, tema central del presente trabajo.

Con el objeto de entender perfectamente la aplicación práctica del método de las pruebas secuenciales bayesianas, en el capitulo IV se presenta un ejemplo tipo y su resolución. Además, en el capítulo V, se analiza un problema práctico que presenta un algoritmo de solución que es aplicable a otros problemas con características similares.

Finalmente se resumen las conclusiones sobresalientes del trabajo en general. Además existen dos anexos con información sobre los programas de computadora utilizados y la distribución de probabilidades beta; respectivamente.

C A P I T U L O   I

L A S   P R U E B A S   S E C U E N C I A L E S
D E S D E   U N   P U N T O   D E
V I S T A   C L A S I C O

1.1 INTRODUCCION

La teoría de las pruebas secuenciales formaliza la siguiente idea intuitiva: desarrollar una técnica por medio de la cual se realicen observaciones en forma secuencial y, después de cada una de ellas, se evalúe un criterio por medio del cuál se determine si se tiene la evidencia suficiente en la muestra para tomar una decisión, ya sea aceptar una hipótesis o rechazarla, o si es necesario hacer uso de una nueva observación. Todo ésto sujeto al control de los dos tipos de error.

Desde el punto de vista de la teoría de Neyman-Pearson, se minimiza el error tipo II dada una cota superior para el error tipo I y un número fijo de observaciones; y con la prueba de la razón de verosimilitud, que es la más potente (ver: Mood, Graybill, Boes.), se puede concluir si en la muestra existe la evidencia suficiente o no para rechazar una hipótesis.

Si ahora, se desea analizar cuándo se tiene suficiente evidencia o no para aceptar una hipótesis como estadísticamente verdadera, sería necesario aumentar el número de observaciones o crear un diseño de experimento que permita controlar los dos tipos de error. Para la primer solución sería posible que el costo por muestreo se incrementara de manera significativa y para la otra respuesta, sería necesario, bajo la teoría de Neyman-Pearson, disminuir la potencia de la prueba con su consecuente pérdida de control sobre los errores.

Sin embargo, bajo el punto de vista de las pruebas secuenciales, es posible controlar los dos tipos de error sin necesidad, ni de disminuir la potencia de la prueba ni, de hacer crecer la muestra a altos niveles; es decir, es posible, fijar cotas superiores para ambos tipos de error y, debido a la naturaleza de la prueba, controlar el tamaño de la muestra; que en muchos casos sería menor que el utilizado bajo la teoría clásica.

Esta metodología reduce los gastos por muestreo, en muchos casos, y es muy útil cuando el estadístico está sujeto a un presupuesto y las observaciones tienen un costo.

Por ejemplo: si se sabe que en un lote de artículos la mayoría son defectuosos o no lo son pero no se sabe con certeza cuál de estas afirmaciones es verdadera, es más ventajoso contar con un plan muestral que indique, por ejemplo:

> "Tomar una muestra de dos artículos; si ambos son defectuosos entonces la decisión es que la mayoría son defectuosos; si los dos están en buen estado entonces la decisión es que la mayoría también lo está; de otra forma, sacar otra muestra de dos artículos y repetir la prueba."

que con otro que establezca:

> "Tomar una muestra de n observaciones y hacer la prueba de la razón de verosimilitud (o cualquier otra)"

y para el cual n sea suficientemente grande.

En este capítulo se desarrolla la teoría de las pruebas secuenciales desde un punto de vista clásico y se analizan sus limitaciones. Debido a la naturaleza de la tésis, no se presentan pruebas rigurosas de todos los teoremas y conclusiones; sin embargo, se indican las referencias bibliográficas para las pruebas respectivas.

Siguiendo el planteamiento teórico y después de un análisis del mismo, se podrá dar respuesta afirmativa a las siguientes interrogantes: 1) ¿ Es posible tener una prueba que permita aceptar casi seguramente una hipótesis ?, 2) ¿ Es posible controlar ambos tipos de error en una prueba ? y 3) ¿ No es necesario fijar a priori el número de observaciones en un plan muestral ?

Se recomienda como bibliografía adicional el texto de Mood, Graybill, Boes.

1.2 PRUEBAS SECUENCIALES, UN ENFOQUE CLASICO

Como se mencionó anteriormente, en esta sección se presenta el desarrollo de las pruebas secuenciales desde el punto de vista clásico. Para tales efectos se define la prueba de la razón de la probabilidad secuencial y se estudian sus propiedades. Sin embargo, antes de entrar en detalle en el estudio de estas pruebas se establece la notación que será usada a lo largo del capítulo.

Sean

$X1, X2, X3, \ldots$ variables aleatorias independientes e idénticamente distribuidas (iid).

Supóngase que se tienen dos hipótesis simples sobre la estructura probabilística (distribución de probabilidades) de las Xi's, denotadas por

$$P0(Xi) \quad y \quad P1(Xi) \;;$$

es decir, las Xi's tienen alguna de estas dos estructuras probabilísticas; pero no se conoce con certeza cuál de éstas es la verdadera.

Asociadas a éstas estructuras probabilísticas, sean las funciones de densidad de probabilidad

$$f0(Xi) \quad y \quad f1(Xi) \,.$$

Es importante notar que, debido a que las variables aleatorias son independientes e idénticamente distribuidas, se tiene que:

$$fi(X1)=fi(X2)=\ldots\ldots=fi(Xj)$$

y $\qquad\qquad i=0,1$

$$Pi(X1)=Pi(X2)=\ldots\ldots=Pi(Xj)$$

Con estos elementos, defínase a

$$lm = \frac{f0(X1,X2,\ldots,Xm)}{f1(X1,X2,\ldots,Xm)} \qquad (1)$$

como la razón de verosimilitud, donde m es el número de observaciones y $f0(X1,X2,\ldots,Xm)$ y $f1(X1,X2,\ldots,Xm)$ son las funciones de densidad de probabilidades conjuntas de $X1,X2,\ldots,Xm$.

Como las Xi's son idependientes, se concluye que

$$fi(X1,X2,\ldots,Xm)=fi(X1)fi(X2)\ldots fi(Xm) \quad i=0,1,$$

de modo que

$$lm = \prod_{i=1}^{m} \frac{f0(Xi)}{f1(Xi)} , \qquad (2)$$

que es, como se mencionó, la razón de verosimilitud que se utiliza para la prueba del mismo nombre que es la más potente (ver: Mood, Grayhill, Boes).

Habiendo establecido la notación necesaria, es posible continuar con el desarrollo de las pruebas secuenciales de acuerdo con la idea intuitiva escrita en la introducción de este capitulo.

Uno de los principales objetivos de las pruebas secuenciales es acotar los dos tipos de error con niveles preasignados; sean estos τ, β para los tipos de error I y II respectivamente; de tal forma que, en términos matemáticos, lo que se desea es tener

Pr{rechazar una hipótesis dado que ésta es verdadera} $\leq \tau$

Pr{aceptar una hipótesis dado que ésta es falsa} $\leq \beta$.

Si se considera que la hipótesis nula es la estructura probabilistica P0 (podria considerarse también P1), las ecuaciones anteriores podrian formularse de la siguiente manera

$$\Pr\{\text{rechazar } P0/P0 \text{ es verdadera}\} \leq \tau$$

y (3)

$$\Pr\{\text{aceptar } P0/P1 \text{ es verdadera}\} \leq \beta$$

con las cuales queda satisfecho el requerimiento del control de ambos errores y se puede definir la prueba secuencial de la razón de probabilidades.

PRUEBA SECUENCIAL DE LA RAZON DE PROBABILIDADES:

DESCRIPCION 1 Sean, A, B dos constantes tales que A<1 y B>1. Supóngase que se tienen j observaciones aleatorias e independientes entre si del fenómeno en estudio (j=1,2,...). Evalúese l_j (definida por la ecuación (2)) y:

si $l_j<A$ acéptese P1

si $l_j>B$ acéptese P0

si $A<l_j<B$ tómese otra observación y evalúese l_{j+1}.

Nótese que l_j no es sino el cociente de las funciones de verosimilitud, y que si

$l_j<A$ entonces $l_j<1$, (4)

de modo que $\prod_{i=1}^{j} f1(Xi) > \prod_{i=1}^{j} f0(Xi)$,

por lo que es razonable aceptar a P1 como la estructura probabilística del fenómeno. Esto es, se acepta P1 cuando el divisor del cociente l_j es mayor.

Intuitivamente, y debido a la naturaleza de la prueba, es lógico pensar que las constantes A, B están relacionadas con los valores de las cotas superiores de los errores, es decir; se espera que

$$A = G1(\tau,\beta) \qquad y \qquad B = G2(\tau,\beta)$$

por lo cual es importante determinar estas constantes como funciones explícitas de τ, β.

Esta relación funcional puede derivarse de las restricciónes de los dos tipos de error. Sea

$$Cn=\{(x1,x2,...xn)/A<lj<B \ j=1,2...n-1;ln<A\} \quad ;$$

es decir, Cn es el conjunto de vectores (x1,x2,...xn) tales que la prueba secuencial termina con la aceptación de P1, después de n observaciones y no antes.

Supóngase ahora que se tiene una prueba secuencial con errores τ, β exactamente. Entonces se tiene que:

$$\tau = Pr\{rechazar\ P0/P0\ verdadera\}=$$

$$= \sum_{n=1}^{\infty} \int_{Cn} \prod_{i=1}^{n} f0(Xi) \leq \sum_{n=1}^{\infty} \int_{Cn} \prod_{i=1}^{n} f1(Xi)=$$

$$= A \sum_{n=1}^{\infty} \int_{Cn} \prod_{i=1}^{n} f1(Xi) \quad =$$

$$= A\ Pr\{rechazar\ P0/P1\ es\ veradera\} = A(1-\beta) \quad ,$$

lo que implica que

$$A \geq \tau/(1-\beta) \quad ; \qquad (5)$$

y de manera similar puede obtenerse la relación

$$B \leq (1-\tau)/\beta \quad , \tag{6}$$

y, por aproximación, puede tomarse:

$$A' = \tau/(1-\beta) \quad y \quad B' = (1-\tau)/\beta; \tag{7}$$

debido a que, por ejemplo, la decisión 'PO es la verdadera distribución de probabilidades' se toma si se cumple que $lm > B$ mas que cuando $lm = B$ (para alguna m que cumpla con esto); por lo cual se puede suponer que para que se tome esta decisión debe de cumplirse que: $B < lm \leq B + \epsilon$; en donde ϵ es un número mayor que cero. Es claro ver que esto se cumple para cualquier $\epsilon > 0$ de tal forma que las constantes A, B pueden ser definidas correctamente por la ecuación (7); y, además, se cumple que:

$$A' \leq A < B \leq B'.$$

[Nota: sean τ', β' los errores de la prueba secuencial definida por las ecuaciones (7); entonces se puede demostrar que:

$$\tau' + \beta' < \tau + \beta;$$

por lo cual las restricciones se siguen cumpliendo (ver: Mood, Graybill, Boes).]

De esta forma han quedado detrminadas las relaciones funcionales de A y B con τ y β de manera explícita, y se cumple la restricción sobre el control de los tipos de error, que era uno de los principales objetivos de la prueba.

Por otro lado, como el muestreo es secuencial, lo que constituye otro de los objetivos formulados, es necesario analizar algunas de las características más importantes acerca del número de observaciones en la muestra (j) que, como se subrayó, es una variable aleatoria.

Antes de proceder a la obtención del valor esperado $E[j]$, es necesario anotar que el interés de estas pruebas está centrado solamente en aquellos procesos que cumplen con

$$Pr\{j<\infty\}=1; \tag{8}$$

lo cual denota que se consideran sólo procesos que terminan con probabilidad 1.

Esta condición se cumple si se cumple también que las variables aleatorias son independientes, lo que se va a suponer acontinuación; ya que, por el contrario, se podría dar el caso en el que toda la información obtenible este contenida en las primeras m observaciones; y si en ese momento no se ha tomado una decisión, entonces ésta nunca se llegará a tomar y el muestreo se haría infinito, lo que significa que si:

$$lm=lm+i \qquad \text{para } i=0,1,2,\ldots.$$

entonces

$$Pr\{j<\infty\}\neq 1$$

y si, además:

$$A<lm<B$$

entonces:

$$A<lm+i<B \qquad i=1,2,3,\ldots.$$

lo que implicaría que el muestreo nunca terminaría y la prueba perdería sentido.

Es pues, por esto que la condición de independencia entre las variables aleatorias debe cumplirse ya que en otro caso no se podría asegurar que la prueba termina con probabilidad uno.

Finalmente, se presenta la expresión de la esperanza del número de observaciones en la muestra; siempre y cuando esta esperanza exista y sea finita.

Se sabe que:

$$A < \prod_{i=1}^{j} \frac{f0(Xi)}{f1(Xi)} < B$$

entonces si

$$Zi = \mathrm{Log} \frac{f0(Xi)}{f1(Xi)}$$

se obtiene

$$\mathrm{Log}\ A < \sum_{i=1}^{j} Zi < \log\ B \ .$$

Obsérvese que, como las Xi's son independientes e identicamente distribuidas, las Zi's también lo son. Ahora supongase que para toda Zi se cumple que:

$$E|Zi|<\infty \qquad y \qquad E[Zi] \neq 0;$$

entonces por la desigualdad de Wald (ver: Feller) se concluye que:

$$E[j] = \frac{E[Z1+Z2+...+Zj]}{E[Zi]} \qquad \text{para toda i;}$$

con lo que

$$E[j/P1] = \frac{\tau\ \mathrm{Log}\ A + (1-\tau)\ \mathrm{Log}\ B}{E[Zi/P1]} \qquad (9)$$

y

$$E[j/P0] = \frac{(1-\beta)\ \mathrm{Log}\ A + \beta\ \mathrm{Log}\ B}{E[Zi/P0]} \qquad (10)$$

Para una prueba formal de las igualdades (9) y (10) consúltese el texto de Silvey.

En estos momentos ya se cuenta con los elementos necesarios para las pruebas secuenciales, los que se resumen y enumeran adelante y se concretizan con un ejemplo presentado en la siguiente sección.

RESUMEN DEL METODO

Sean: $X1, X2, \ldots\ldots\ldots$ variables aleatorias i.i.d.;

supóngase que se quiere efectuar una prueba de hipótesis:

$$H0: P0 \quad vs \quad H1: P1$$

en donde P0 y P1 son estructuras probabilísticas diferentes, con funciones de densidad de probabilidades f0 y f1, respectivamente; defínase:

$$l_j = \sum_{i=1}^{j} \frac{f0(Xi)}{f1(Xi)} ,$$

$$A = \alpha/(1-\beta) \qquad y \qquad B = (1-\alpha)/\beta ,$$

en donde α, β son las cotas superiores de los errores tipo I y II, respectivamente.

Tómense observaciones secuencialmente, y después de cada una de ellas pruébese:

si $l_j < A$ acéptese P1

si $l_j > B$ acéptese P0

si $A < l_j < B$ obsérvese X_{j+1}.

La esperanza del número de observaciones a muestrear está dada por las ecuaciones (9) y (10).

1.3 EJEMPLO

Supóngase que se sabe que las variables aleatorias Xi's tienen una distribución normal con media desconocida y varianza igual a 16; además, se sabe que estas variables son independientes e idénticamente distribuidas.

Supóngase, también, que es conocido que la media puede tomar los valores 11 ó 12, pero no se sabe con exactitud cuál de éstos es el verdadero. Y el estadístico desea hacer una prueba de hipótesis sobre la media de la distribución, pero controlando los tipos de error con niveles .05 para ambos.

Es decir; se tiene:

$$X_i \text{ con distribución } N(\omega,16)$$

$$\tau = .05$$

$$\beta = .05$$

y se quiere probar:

$$H0: \omega=11 \quad \text{vs} \quad H1=: \omega=12$$

En primer lugar, el estadístico quiere determinar el valor esperado del número de observaciones a tomar. Sin embargo, antes es necesario evaluar las constantes A, B.

$$A=\tau/(1-\beta)=.05/.95=.0526358$$

$$B=(1-\tau)/\beta=.95/.05=19$$

lo que implica:

$$E[j/\omega=11)=\frac{\tau \log A + (1-\tau) \log B}{E[Zi/\omega=11]}$$

y

$$E[j/w=12]=\frac{(1-\beta)\log A+\beta\log B}{E[Zi/w=12]}$$

pero:

$$E[Zi/w=11]=(w0-w1)^2/2s^2$$

$$E[Zi/w=12]=-(w0-w1)^2/2s^2$$

en donde $s^2=16 \quad w0=11 \quad w1=12$

entonces:

$$E[Zi/w=11]=1/32 \quad y \quad E[Zi/w=12]=-1/32;$$

por lo que:

$$E[j/w=11]=84=E[j/w=12]$$

es decir, el valor esperado del número de observaciones a muestrear es aproximadamente 84 para ambos casos. En contraste con el número que se pide para pruebas clásicas, el cual es de 100.

Ahora, el estadístico procede a determinar la prueba de hipótesis, para la cual debe de determinar el cociente de la razón de verosimilitud:

$$l_j=\frac{\exp\{(-1/32)\sum_{i=1}^{j}(xi-12)^2\}}{\exp\{(-1/32)\sum_{i=1}^{j}(xi-11)^2\}}$$

lo que implica que:

$$l_j=\exp\{(-1/32)\sum_{i=1}^{j}(23-2xi)\}=\exp\{(-1/32)(23j-2\sum_{i=1}^{j}xi)\}$$

y se obtiene la prueba

$$.05263158 < l_j < 19 \quad ,$$

o, equivalentemente

$$A_j=-47.11102+(23/2)j < \sum_{i=1}^{j} x_i < 47.11102+(23/2)j=B_j$$

que determina la prueba, y el estadístico puede llevar a cabo el muestreo secuencial.

Supóngase que el estadístico realiza el muestreo y observa los valores que se muestran en la tabla 1.1 (valores simulados con una computadora, cuyo programa se adjunta en el Anexo 1, programa uno, por medio de un generador de números aleatorios de una distribución normal con media 12 y varianza 16).

Después de 32 observaciones, el estadístico concluye los siguientes resultados:

$$A32=320.889 \quad , \quad B32=415.111$$

$$\sum_{i=1}^{32} x_i = 415.689 > B32 \quad ,$$

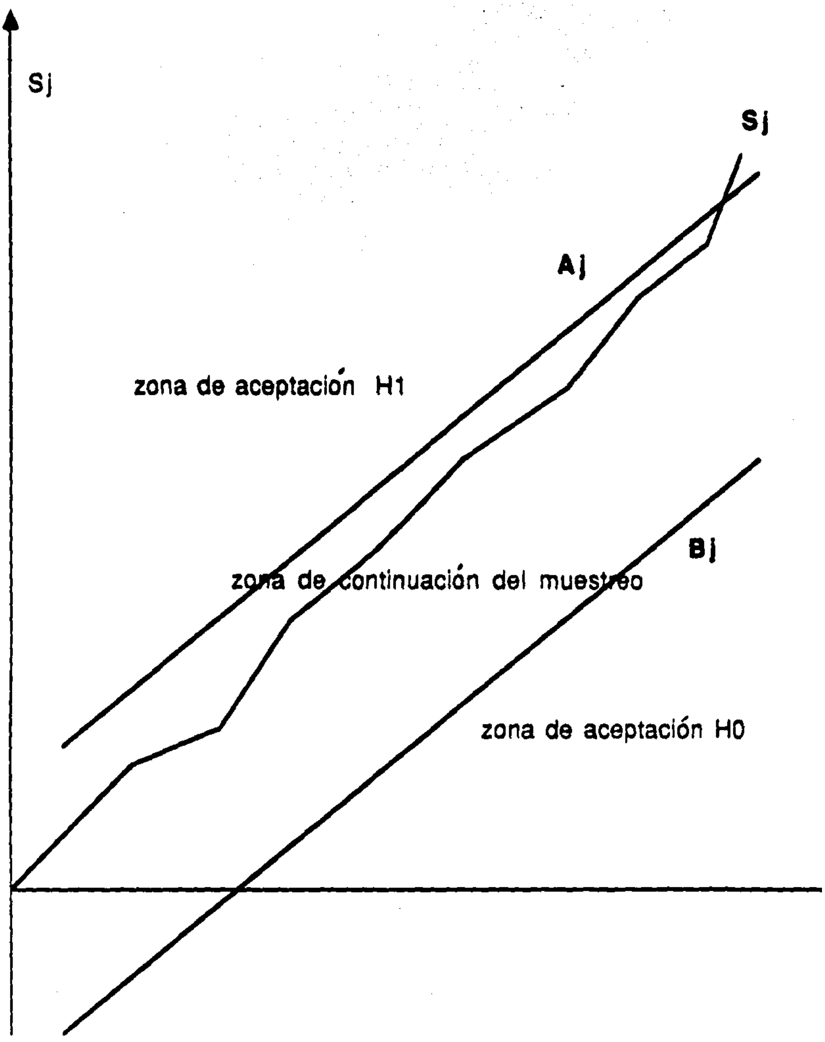
lo que indica que el estadístico debe rechazar H0: w=11 y aceptar H1: w=12 (que es el valor verdadero de w).

Se presenta, en la tabla 1.2, un gráfico con el comportamiento de la sumatoria de las xi's, las zonas de aceptación y rechazo, y las bandas que las determinan.

TABLA 1.1

j	xj	Σ xi	Aj	Bj
1	9.597	9.597	-35.611	58.611
2	12.404	22.002	-24.111	70.111
3	11.330	33.332	-12.611	81.611
4	12.981	46.314	-1.111	93.111
5	12.562	58.876	10.389	104.611
6	13.368	72.245	21.889	116.111
7	16.423	88.668	33.389	127.611
8	10.002	100.671	44.889	139.111
9	14.426	115.097	56.389	150.611
10	10.928	126.025	67.889	162.111
11	15.560	141.586	79.389	173.611
12	18.422	160.009	90.889	185.111
13	12.423	172.432	102.389	196.611
14	11.568	184.000	113.889	208.111
15	9.425	193.426	125.389	219.611
16	10.623	204.049	136.889	231.111
17	12.032	216.082	148.389	242.611
18	16.034	232.116	159.889	254.111
19	10.932	243.048	171.389	265.611
20	14.322	257.371	182.889	277.111
21	15.432	272.806	194.389	288.611
22	12.423	285.227	205.889	300.111
23	8.639	293.866	217.389	311.611
24	16.325	310.192	228.889	323.111
25	11.569	321.761	240.389	334.611
26	17.424	339.185	251.889	346.111
27	10.623	349.809	263.389	357.611
28	11.562	361.371	274.889	369.111
29	13.623	374.985	286.389	380.611
30	12.423	387.419	297.889	392.111
31	15.628	403.047	309.389	403.611
32	12.642	415.689	320.889	415.111

tabla 1.2


Sj
Sj
Aj
zona de aceptación H1
Bj
zona de continuación del muestreo
zona de aceptación H0

1.4 LIMITACIONES

Hasta estos momentos se ha desarrollado la teoría estadística de las pruebas secuenciales desde el punto de vista clásico, y se ha dado una respuesta ventajosa a las interrogantes planteadas en la introducción de este capítulo.

Es decir, se han logrado los dos objetivos principales, a saber:

-Controlar ambos tipos de error y hacerlo en forma óptima,

-muestrear secuencialmente.

Sin embargo, surgen otras dos interrogantes acerca de la optimización del "Plan Muestral":

¿En qué momento es óptimo detener una prueba de carácter secuencial?

¿En qué casos no es relevante muestrear, debido a la escasa posibilidad de aumentar la información sobre el fenómeno en estudio?

Todo esto es debido a que, en muchos casos, el estadístico está sujeto a un presupuesto y el muestreo tiene un costo; por lo cual desea optimizar sus estrategias inferenciales y sus técnicas muestrales. Esto es, sería útil contar con un criterio que evaluara el riesgo de continuar muestreando y, con base en éste, determinar cuándo se debe detener una prueba; y también, en particular, cuándo se debe iniciar el muestreo, y cuándo las observaciones no enriquecen en forma importante o significativa la información que hasta el momento se tiene, por lo cual no es óptimo hacer uso de una muestra.

Con la teoría de las pruebas secuenciales clásicas no es posible dar respuesta a estas interrogantes debido a que no se cuenta con los elementos necesarios para ello. Y, por otro lado, en esta teoría no se toma en cuenta la información que el estadístico pudiera tener, a priori, sobre el fenómeno que desea analizar; por lo que se hace necesario estudiar alguna otra teoría que brinde la posibilidad de superar los obstáculos mencionados.

Una técnica que provee una solución adecuada a este problema es el análisis de las pruebas secuenciales bayesianas, las cuales hacen uso de nuevas herramientas, desde otro punto de vista, de tal forma que permiten al estadístico llevar a cabo una optimización de su plan muestral a través de la solución del problema de la parada óptima.

Y no solamente se consigue esta optimización; sino que también, dentro de esta técnica, se puede tomar en cuenta y evaluar el conocimiento a priori que el estadístico tenga; experiencia que, en la mayoría de los casos, resulta relevante e importante.

Por otro lado, y como principal característica de la prueba, es la única manera de resolver varios tipos de problemas para los que en la teoría clásica se podrían tener varios resultados iguales sin llegar a alguno en concreto.

CAPITULO II

ESTADISTICA BAYESIANA

Y

TEORIA DE DECISIONES

2.1 INTRODUCCION

En la sección final del capítulo anterior se enumeraron las principales limitaciones de las pruebas secuenciales clásicas, motivo por el cual en las secciones subsecuentes se desarrollarán técnicas y criterios de las pruebas secuenciales bayesianas y del problema de la parada óptima.

Sin embargo, antes de abordar de lleno estos temas, es necesario presentar los conceptos básicos de la estadística bayesiana y de la teoría de decisiones, técnicas por medio de las cuales se pueden combatir exitosamente tales limitaciones.

En este capítulo se desarrollan y presentan estos conceptos, en forma no muy profunda, ya que esto no se considera tema central de esta tesis. Como bibliografía complementaria, para una formalización de los temas tratados, se recomiendan los textos de: O. Berger, Box and Tiao y De Groot.

Es muy importante analizar la relevancia de la estadística bayesiana dentro de la inferencia y las pruebas de hipótesis, y de la teoría de decisiones, para una nueva presentación de muchos problemas. Esta nueva técnica está basada en la misma idea intuitiva del análisis estadístico, pero incorpora y formaliza elementos no considerados anteriormente.

Fundamentada en el teorema de Bayes, la estadística bayesiana trabaja con elementos nuevos que permiten hacer uso de información importante para la resolución de problemas inferenciales. Uno de estos elementos es la incorporación de una distribución a priori del parámetro a inferir, la cual refleja el grado de creencia que el estadístico pudiera tener sobre el fenómeno en estudio; antes de llevar a cabo algún muestreo. Es decir, la experiencia del estadístico es tomada en cuenta para la inferencia.

En la mayoría de los casos, los parámetros de estas probabilidades apriori son completamente subjetivos y rompen con la idea frecuentista clásica; por lo que la interpretación de una probabilidad es ampliada para aquellos experimentos y sucesos que no pueden ser repetidos. Por ejemplo: Si se menciona que la probabilidad de que se case X con Y es un medio, bajo la idea frecuentista no es posible concebir el hecho de que se llegó a este número al repetir el experimento una infinidad de veces; por lo que este número pierde su interpretación práctica. Ahora bien, bajo el punto de vista bayesiano, el suceso de que X se case con Y se asocia a un experimento auxiliar, sea este experimento, por ejemplo: el de lanzar una moneda al aire, y la interpretación que este número tiene es que es tan probable que se case X con Y como que al lanzar una moneda al aire salga águila.

Nótese que el experimento auxiliar, cualquiera que éste sea, sirve para 'pasar' de creencias y experiencias a una distribución de probabilidades númerica y concreta.

Por otro lado, y haciendo uso de la teoría de decisiones estadísticas, se pueden evaluar los riesgos y las pérdidas para una prueba de hipótesis. Ideas que en alguna forma conceptualizan y evalúan los procedimientos inferenciales adecuados para poder optimizar el plan muestral y la toma de decisiones.

Con base en estas ideas y conceptos, y después de su presentación en este capitulo; ya será posible, en los siguientes, hacer un análisis formal de la teoria de las pruebas secuenciales bayesianas y del problema de la parada óptima.

2.2 PROBABILIDAD SUBJETIVA

Como se mencionó en la introducción, por medio de la metodología de la probabilidad subjetiva es posible no sólo asignar probabilidades a aquellos eventos repetibles, sino que también a sucesos que no se pueden repetir para obtener sus frecuencias de ocurrencia.

No se desarrollará un análisis profundo al respecto ya que este tema, por sí mismo, podría representar un estudio individual; por el contrario, se presenta solamente con suficiente detalle para llevar a cabo ciertas consideraciones en la justificación del uso de la estadística bayesiana.

El objetivo principal de la probabilidad subjetiva es: interpretar la factibilidad de la ocurrencia de un evento como probabilidad, por medio de un experimento auxiliar; de tal forma que esta regla de asignación de probabilidades determine un ordenamiento completo del conjunto de eventos.

Sea

S --------------- el espacio muestral,
A --------------- el conjunto de eventos asociados con S;

y defínase una relación dentro de A de la suguiente manera, para cualesquiera dos eventos $a, b \in A$,

$a < b$ sí y sólo sí b es mas factible que a;

$a \equiv b$ sí y sólo sí b es tan factible como a;

$a \leq b$ sí y sólo sí b es, al menos, tan factible como a.

Nótese que esta relación determina un ordenamiento completo del conjunto A ya que para cualquier pareja de elementos de A esta relación se cumple.

Teniendo ya un ordenamiento de éste conjunto, se persigue encontrar una función numérica que este de acuerdo con este orden, es decir:

$$\text{si} \quad P(a)<P(b) \quad \text{entonces} \quad a < b.$$

Ahora supóngase que la relación $<$ cumple con las siguientes propiedades:

1) para todo $a, b \in A$, se cumple sólo una de las expresiones siguientes:

   $a < b$ ó $a \equiv b$ ó $b < a$;

2) si $a \in A$, entonces:

   $\emptyset \leq a$ y en particular $\emptyset \leq S$.

Con ayuda de estos dos supuestos y la definición de un experimento auxiliar, se estará en posibilidades de asociar una probabilidad única para cada elemento de A.

Supóngase la existencia de una clase β de eventos que cumplen con:

1) Cada elemento de β tiene una probabilidad de ocurrencia conocida.

2) Para todo número p $(0 \leq p \leq 1)$ existe un elemento de β cuya probabilidad de ocurrencia es p.

Por ejemplo, para algunos casos, de manera natural se puede definir esta clase de eventos β, con una variable aleatoria uniformemente distribuida en el intervalo [0;1]. Esta variable aleatoria cumple con las dos propiedades anteriores, por lo que se puede usar como experimento auxiliar.

Entonces, con ayuda de este experimento, es posible asignar probabilidades a los elementos de A. Para cada elemento $a \in A$, se define su probabilidad asignándole la misma de un elemento $b \in \beta$ y que cumplan que:

$$a \equiv b$$

y, a partir de aquí se puede construir la distribución de probabilidades de A.

Nótese que, para este ejemplo, la clase β la integran los intervalos

$$[0,X];$$

donde X es una variable aleatoria con distribución $U[0;1]$; por lo que se tiene que

$$P(a)=m \quad \text{si y sólo si} \quad (0,m) \equiv a$$

donde $\qquad a \in A \quad y \quad 0 \le m \le 1;$

ó, de otra forma: $a \equiv (0,P(a))$.

Y P es la función de densidad de A;y ahora lo único que queda por demostrar es que esta distribución de probabilidades cumple con:

1) $P(\emptyset)=0$

2) $P(S)=1$

3) $\int_A P(a)da = 1 \quad$ ó $\quad \sum_A P(a) = 1$

(ver: De Groot.)

2.3 ESTADISTICA BAYESIANA

En esta sección se desarrollan los principales característicos y los principales criterios de la estadística bayesiana, todo ello con ayuda de la probabilidad subjetiva para la determinación de las distribuciones a priori de los parámetros.

Supóngase que se tiene un problema inferencial sobre un determinado parámetro θ numérico o vectorial que pertenece a un conjunto de estos mismos Θ.

Se denomina distribución de probabilidades a priori de θ al grado de creencia que el estadístico tiene sobre la posibilidad de que cierto parámetro θ ∈ Θ sea el verdadero; este grado de creencia puede quedar determinado por la experiencia del estadístico, o puede ser que, a falta de ésta, por ejemplo se determine que: todos los elementos de Θ tengan la misma probabilidad.

Es decir, el propio parámetro desconocido puede tener una distribución de probabilidades que se busca determinar.

Sea $X=(x1,x2,x3,...xn)$ un vector de observaciones, las cuales se distribuyen con dependencia de un parámetro θ ∈ Θ desconocido, sobre el cual se quiere hacer inferencia.

Se tiene entonces que:

1) $P(\theta)$ ------- la distribución a priori del parámetro θ ∈ Θ,

2) $P(X/\theta)$ ----- la distribución de las observaciones dado θ,

Y con base en ésta información se desea determinar:

3) $P(\theta/X)$ ----- la distribución del parámetro dadas las observaciones.

A $P(\theta/X)$ se le conoce como distribución de probabilidades a posteriori de $\theta \in \Theta$.

Por el teorema de Bayes se sabe que:

$$P(X/\theta)P(\theta)=P(X,\theta)=P(\theta/X)P(X) \quad ,$$

en donde $P(X,\theta)$ es la distribución conjunta de (X,θ) sobre el espacio $S \times \Theta$. (S es el espacio muestral)

Dada la observación X, es posible obtener la distribución condicional del parámetro θ, que es la distribución a posteriori; ésta está dada por la siguiente expresión:

$$P(\theta/X)=\frac{P(X/\theta)P(\theta)}{P(X)}$$

en donde se tiene que:

1) $P(X)=E(P(X/\theta))$ es la esperanza de $P(X/\theta)$; para la cual X constituye un vector de observaciones conocidas y como se obtiene esta esperanza sobre el espacio Θ, entonces esta expresión es una constante;

2) $P(\theta)$ es la distribución a priori de $\theta \in \Theta$;

3) $P(X/\theta)$ es la influencia que tiene la observación X sobre la creencia que se tiene apriori del parámetro y

4) $P(\theta/X)$ es la distribución a posteriori de θ.

Nótese que $P(X/\theta)$ se trata como función de θ y no de X, debido a que X es un vector que en el momento de la inferencia es conocido; a esta función se le denomina: función de verosimilitud.

A $P(\theta)$, distribución a priori del parámetro a inferir, se le denotará en una forma diferente con el objeto de estar acordes con la literatura general, por lo que: $P(\theta)=\pi(\theta)$.

Por lo que se tiene:

$$P(\theta/X)= \frac{\pi(\theta)P(X/\theta)}{P(X)} =$$

$$P(\theta/X)= \frac{\pi(\theta)P(X/\theta)}{E(P(X/\theta))}$$

pero como $E(P(X/\theta))$ es tan sólo una constante que normaliza $P(\theta/X)$, no es necesario ponerla en su forma explícita; por lo tanto la expresión anterior finalmente queda de la siguiente forma:

$$P(\theta/X) \propto \pi(\theta)P(X/\theta)$$

o bien $\quad P(\theta/X) = k\pi(\theta)P(X/\theta)$

en donde k es un número real que se determina en tal forma que se cumpla que

$$\int_{\Theta} P(\theta/X)\, d\theta = 1 \qquad \text{en el caso continuo;}$$

$$\sum_{\Theta} P(\theta/X) = 1 \qquad \text{en el caso discreto.}$$

En el caso en que se tengan dos observaciones, o conjunto de observaciones secuenciales, sean estas X1 y X2, a la distribución a posteriori P(θ/X1) se le puede usar como distribución a priori para la información contenida en la observación X2, si X1 y X2 dos son independientes e idénticamente distribuidas; y así para el caso de n observaciones

$$P(\theta/X1,X2) \propto \pi(\theta)P(X1,X2/\theta)$$

pero: $$P(X1,X2/\theta) = P(X1/\theta)P(X2/\theta)$$

si X1, X2 son independientes dado θ, por tanto

$$P(\theta/X1,X2) \propto \pi(\theta)P(X1/\theta)P(X2/\theta)$$

pero $$\pi(\theta)P(X1/\theta) \propto P(\theta/X1)$$

de tal forma que:

$$P(\theta/X1,X2) \propto P(\theta/X1)P(X2/\theta)$$

en donde puede notarse que P(θ/X1) desempeña el papel de la distribución a priori de θ para la segunda observación X2.

Finalmente, ya teniendo la expresión para P(θ/X), se puede llevar a cabo la inferencia del parámetro de acuerdo con los criterios elegidos por el estadístico. También es posible determinar intervalos de confianza y realizar pruebas de hipótesis acerca del parámetro inferido.

En resumen, es posible mencionar que π(θ) es la información que el estadístico posee del parámetro apriori, ya sea por experiencia o por cualquier otro factor, y que P(X/θ) es la información que la observación X proporciona sobre el parámetro; si π(θ) es, al final, muy diferente a P(θ/X) se puede concluir que las observaciónes tuvieron gran influencia (proporcionaron mucha información) para la inferencia, y que en cambio, la información a priori no tuvo gran relevancia para la resolución del problema.

2.4 LOS CONCEPTOS DE PERDIDA Y RIESGO

Los conceptos de pérdida y riesgo son parte de la teoría de decisiones estadísticas, y juegan un papel preponderante en el desarrollo del problema de la parada óptima; es por esto que, en esta sección, se establecen sus definiciones formales y se comentan sus propiedades mas importantes.

Como bibliografía complementaria se sugiere el texto de Ferguson, en el que se presenta un extenso tratado no sólo de estos conceptos, sino también de lo que en general trata la teoría de decisiones.

Cuando se trabaja con herramientas estadísticas y, en particular, con técnicas inferenciales; es porque, como se ha mencionado, existe algún parámetro o estado de la naturaleza que resulta desconocido para el estadístico. En la gran mayoría de los casos, la inferencia de dicho estado es necesaria para la toma de acciones que permitan optimizar los resultados deseados; es por ello que la teoría de decisiones estadísticas busca desarrollar criterios indicadores de acciones estadísticas en el momento de la inferencia.

Se sabe que el estado de la naturaleza pertenece a un conjunto Θ de entre los cuales uno es el verdadero y se tiene, también, un conocimiento a priori acerca del parámetro; con estos datos, el problema, en general, consiste en determinar la mejor aproximación al parámetro verdadero.

Sin embargo, es muy posible que en ciertos casos tanto la observación como la inferencia en sí tengan un costo determinado, por lo tanto el estadístico no solo está interesado en una buena estimación del parámetro, sino también en el control de los gastos en los que se pueda incurrir.

Este tipo de restricciones dan lugar a consideraciones adicionales sobre la estimaciones.

Se puede entonces definir un espacio de decisiones D, cuyos elementos sean acciones a tomar; sin embargo, a cada decisión $d \in D$ le corresponde un riesgo determinado y una pérdida (ganancia en el caso negativo) dado por alguna función que tome en cuenta el verdadero valor del parámetro.

Esta función de pérdida se define a continuación.

DEFINICION 1 A una función $L: \Theta \times D \longrightarrow \mathbb{R}$ que cumpla con la siguiente propiedad:

$L(\theta,d)$ está acotada por abajo; para todo $(\theta,d) \in \Theta \times D$, se le llama función de pérdida.

Esta función de pérdida cuantifica la pérdida de tomar la decisión d cuando θ es el verdadero parámetro.

Ahora, debido a que el parámetro θ no es conocido y sobre él es la inferencia, es natural pensar en obtener la esperanza matemática de esta función, que no es más que el riesgo que el estadístico corre al tomar la decisión d.

Esto es, si $\pi(\theta)$ es la distribución a priori del parámetro, entonces el riesgo de tomar la acción d está dado por:

$$\sigma(\pi,d) = \int_{\Theta} L(\theta,d)\, \pi(\theta)\, d\theta$$

en donde se supone que la integral existe y es finita para toda $d \in D$; entendiendo a la integral como un símbolo de sumatoria en el caso discreto.

Se puede dar ya una definición formal de riesgo para toda $d \in D$, solamente con el objeto de resumir los conceptos del parráfo anterior, como sigue:

DEFINICION 2 Para una función de pérdida $L(\theta,d)$, la función de riesgo, denotada por $\sigma(\theta,d)$, está definida por la expresión siguiente:

$$\sigma(\theta,d) = E(L(\theta,d)) \qquad (\theta,d) \in \Theta \times D$$

en donde la esperanza se toma en el espacio X.

DEFINICION 3 Para cualquier distribución π del parámetro θ, el riesgo bayesiano se define como la cota inferior de la función de riesgo $\sigma(\pi,d)$, para toda $d \in D$ se le denota por $\sigma^*(\pi)$. Esto es:

$$\sigma^*(\pi) = \inf_{d \in D} \sigma(\pi,d)$$

y a la d que cumple con esta condición se le conoce como decisión de Bayes en contra de π.

Estas definiciones conforman toda la base sobre la cual se trabajará; sin embargo, es necesario extenderlas un poco para los casos en que se tengan observaciones.

Es lógico pensar que, en los casos en los que se pueda hacer uso de un conjunto de observaciones, la decisión sea una función de este conjunto; y en lugar de tener y denotar esta decisión como d, se debe hacer como función explícita de la muestra; es decir, si se han muestreado $X_1,X_2,\ldots,X_n$; entonces a la función que vincula estas observaciones con el espacio D se le conoce como regla o procedimiento de decisión y se define como:

DEFINICION 4 A una función $\delta: S \longrightarrow D$ se le llama procedimiento de decisión si, cuando se ha observado X, entonces se toma la decisión $\delta(X) \in D$.

Finalmente, se modifican las definiciones 1, 2 y 3 para los casos en los que se tenga δ en lugar de d; es muy importante darse cuenta de las diferencias básicas que estas definiciones tienen entre sí, ya que los conceptos que están en ellos son la base teórica para el desarrollo del tema central del presente trabajo.

DEFINICION 5 La función $L(\theta,a(X))$ es la pérdida de tomar la decisión $a(X)$ cuando el parámetro verdadero es θ y se ha observado X.

DEFINICION 6 Para cualquier distribución a priori π del parámetro $\theta \in \Theta$, y cualquier procedimiento de decisión $a(X) \in D$:

$$\sigma(\pi,a) = E(L(\theta,a(X))) = \int_{\Theta \times S} L(\theta,a(X))P(X,\theta)\, dx\, d\theta$$

ó bien

$$\sigma(\pi,a) = \int_{\Theta} \int_{S} L(\theta,a(X))f(X/\theta)\pi(\theta)\, dx\, d\theta$$

es el riesgo esperado, o pérdida promedio, del procedimiento respectivo.

DEFINICION 7 El riesgo de Bayes con observaciones se define como:

$$\sigma^*(\pi) = \inf_{a \in D} \sigma(\pi,a) = \sigma^*(\pi,a^*)$$

y a a^* se le conoce como la decisión de Bayes en contra de π.

Con estos conceptos perfectamente bien entendidos es posible entrar de lleno al tema de las pruebas secuenciales bayesianas y al problema de la parada óptima. Sin embargo es importante que antes de seguir adelante se haga una revisión general de todo el material que, hasta ahora, se ha presentado.

CAPITULO III

PRUEBAS SECUENCIALES BAYESIANAS: EL PROBLEMA DE LA PARADA OPTIMA

3.1 INTRODUCCION

Dentro del problema general de la teoría de decisiones estadísticas, la parte referente a las pruebas secuenciales bayesianas y, específicamente, al problema de la parada óptima tienen una gran importancia para la optimización del plan muestral del estadístico.

Como se ha mencionado en capítulos anteriores, esta optimización del plan muestral no es sino el establecimiento y formulación adecuados de la regla de decisión muestral.

En este capítulo se consideran problemas para los cuales el estadístico tiene la posibilidad de tomar las observaciones una por una, es decir, secuencialmente, de alguna distribución con un parámetro o vector de parámetros θ desconocido. Después de cada observación Xi, el estadístico puede evaluar la información que ha obtenido, a través de las muestras, y con base en ello decidir si continua muestreando o toma una decisión en ese momento.

Todo esto se hace suponiendo que el mustreo tiene un costo -suposición que se cumple en la mayoría de los casos.

Con este fin, el estadístico cuenta con un plan muestral, que es la estrategia de muestreo que será optimizada; y con una regla de decisión. Regla que indica qué acción se debe tomar, de acuerdo a la información que se tenga, es decir, es una función del espacio muestral S al campo de decisiones D.

Para los efectos de la optimización antes mencionada, el estadístico debe de evaluar los riesgos que tiene en el momento que se quiere tomar la decisión y determinar, si la esperanza matemática del riesgo de la continuación es menor, o el riesgo de tomar una acción inmediata, que se debe elegir otra muestra que esté en posibilidades de aportar información adicional para una elección posterior que minimice un poco más los riesgos.

Es muy importante conocer y manejar los conceptos de riesgo y de pérdida, definidos en la última sección del capitulo anterior, así como sus propiedades; puesto que con base en ellos se desarrollará toda la teoría en el presente capítulo.

En principio, se formula el problema para casos que están acotados en el número de observaciones, esto es, para casos en los que a lo más se puedan tomar n observaciones ($n \in \mathbb{N}$), haciendo uso de una técnica llamada inducción retrospectiva (sección 3.2.)

En la sección 3.3 se extienden y se amplían los conceptos para casos en los que se puedan considerar hasta un número infinito de observaciones.

Finalmente, se termina con algunos comentarios y conclusiones generales para la aplicación práctica de los conceptos e ideas presentados.

Como bibliografía adicional se recomiendan los textos de: Ferguson, Silvey, Lehmann y De Groot.

La notación usada es, prácticamente, la misma; excepto por algunas nuevas inclusiones que se definirán conforme se vaya haciendo necesario; y, en su totalidad, está de acuerdo con la usada en los textos sobre la materia.

3.2 INDUCCION RETROSPECTIVA

Como fue mencionado en la introducción de este capitulo, se inicia con la presentación formal y rigurosa del caso acotado; por lo que, en esta sección, el análisis se centra, solamente, en los casos que cumplen esta restricción.

DEFINICION 1 Un proceso de decisión secuencial acotado en observaciones es aquél que cumple con:

$$\Pr\{N \leq n\} = 1$$

en donde N es el número de observaciones y n es un entero positivo.

Es decir, un proceso de decisión secuencial acotado en observaciones es aquél para el cual, a lo más, se pueden tomar n muestras $X1, X2, \ldots, Xn$.

Por otro lado, y como fue señalado en la sección 2.4, se sabe que $\pi(\theta)$ es la distribución a priori del parámetro $\theta \in \Theta$; pero, después de j observaciones, a $P(\theta/X1,X2,\ldots,Xj)$ se le puede considerar como distribución a priori para la observación j+1 por lo que, con el objeto de facilitar la notación, se define el concepto de distribución a priori después de j observaciones.

DEFINICION 2 Sea $\pi(\theta)$ la distribución a priori del parámetro $\theta \in \Theta$; entonces, a la distribución de probabilidades a posteriori:

$$\pi_j(\theta) = P(\theta/X1,X2,\ldots,Xj)$$

para $j=1,2,3,\ldots,n$ se le denominará distribución a priori después de j observaciones, y a $\pi(\theta)$ se le denotará por

$$\pi_0(\theta) = \pi(\theta)$$

Ahora, antes de continuar, es importante anotar cuál es el objetivo primordial que el estadístico persigue con el uso de esta técnica de la inducción retrospectiva, y hacer algunos comentarios acerca de la idea intutiva que sustenta esta metodología.

El fin último que se persigue es el de encontrar un procedimiento óptimo para problemas de decisión secuencial acotados, entendiéndoce por procedimiento a la estructura del plan muestral y a la regla de decisión respectiva, que minimice el riesgo esperado del problema.

La técnica de la inducción retrospectiva se emplea, como su nombre lo indica, considerando en un principio el último paso del muestreo, y de ahí se continúa trabajando hacia atrás. Esto es así debido a que la idea intuitiva que sustenta esta metodología señala que, para que el estadístico pueda tomar una decisión, en un determinado paso del proceso, es necesario que analice y considere en cuánto más puede ayudarle una nueva observación, si es conveniente tomarla, y cómo la utilizaría' en tal caso. Como sólo se pueden tomar n muestras, resulta claro que el primer paso es tan solo el análisis, determinista, de la información que se podría tener, y de la pérdida o ganancia que en cada caso se obtendría. Para el segundo paso se hace una evaluación similar, y se compara el riesgo de detener el estudio en ese momento contra el riesgo esperado de tomar una muestra adicional, más el costo por muestreo; y así sucesivamente, hasta llegar al último paso, en el cual todavía no se tienen observaciones.

Debido a la complejidad y cantidad de cálculos numéricos que se deben de llevar a cabo para un estudio de esta índole, es deseable contar con la ayuda de una computadora.

La construcción de este procedimiento de decisión secuencial óptimo por el método de la inducción retrospectiva es llamado: "Principio de Optimalidad" (ver: Bellman) y se enuncia de la siguiente manera:

> "El procedimiento secuencial óptimo tiene que satisfacer el requerimiento de que, si en algún paso del procedimiento los valores $x1,x2,\ldots,xj$ $(j<n)$ se han observado, entonces la continuación de éste tiene que ser la decisión secuencial óptima del nuevo procedimiento para el cual la distribución apriori de θ está dado por $\pi j(\theta) = \pi(\theta/x1,x2,\ldots,xj)$; y para el que, a lo más, $n-j$ observaciones se pueden tomar."

Esto es, después de cada observación que se toma se considera un nuevo problema, pero con una distribución a priori del parámetro, o estado de la naturaleza, nueva también, que está dado por la distribución a posteriori del paso anterior y con un número de observaciones por tomar menor a los que se podían tomar en el paso inmediato anterior.

Basado en el principio y en la idea intuitiva mencionados, en adelante se presenta el desarrollo matemático para la construcción del procedimiento de decisión secuencial óptimo, a través del uso de la inducción retrospectiva.

Por razones de notación, en este desarrollo matemático se usarán signos integrales; los cuales, en los casos discretos, deberán interpretarse como sumatorias.

Así pues supóngase que se tiene un problema en el que se desconoce el parámetro θ que pertenece a un conjunto de parámetros Θ, para el que se tiene la función de probabilidades a priori $\pi 0$. Se sabe que el riesgo bayesiano, del procedimiento óptimo, para una decisión inmediata está dado por

$$\sigma(\pi 0(\theta)) = \inf_{d \in D} \int_{\Theta} L(\theta,d)\, \pi(\theta)\, d\theta \quad ; \quad (1)$$

de la misma manera, después de j observaciones, el riesgo bayesiano para una decisión inmediata está dado por:

$$\sigma(\pi_j(\theta)) = \inf_{d \in D} \int_{\Theta} L(\theta,d)\, \pi_j(\theta)\, d\theta \qquad (2)$$

Es importante aclarar, antes de continuar, que sólo se consideran riesgos y decisiones bayesianos, ya que el procedimiento correspondiente es el que tiene riesgo mínimo, lo cual es uno de los objetivos mencionados.

Con objeto de facilitar aún más la notación, se presenta la siguiente definición. Para ello, y para el resto del proceso siguiente, supóngase además que cada observación tiene un costo C_i por muestreo.

DEFINICION 3 El riesgo de la parada, después de j observaciones ($j<n$), está dado por:

$$U_j = \sigma(\pi_j(\theta)) + \sum_{i=1}^{j} C_j$$

en donde $\sigma(\pi_j(\theta))$ está dado por la ecuación (2). Para el caso $j=n$, la definición pierde su variabilidad estocástica, y se tiene la utilidad determinista y se escribe simplemente como U_n.

En estos momentos ya se cuenta con los elementos necesarios para poder establecer los criterios de la técnica de la inducción retrospectiva. Siguiendo la idea intuitiva, supóngase que se tiene una población de n elementos.

Comenzando por el último elemento de la muestra, es decir, suponiendo que ya se observaron los n elementos , entonces se puede llevar a cabo un análisis determinista para la toma de decisiones; y, con base en él, elegir la decisión d que haga óptima la función de pérdida.

Valdría la pena aclarar que este desarrollo retrospectivo de la prueba solo se hará analíticamente. Una vez teniendo este desarrollo analítico no es necesario hacer todo esto en la práctica. Es decir, practicamente no se lleva acabo el estudio en forma retrospectiva; si, en cambio, en forma normal.

Ahora supóngase que se tienen n-1 observaciones. Lo que se quiere decidir es si se muestrea al último elemento de la población, o se toma una decisión que garantice la minimización del riesgo esperado. Se sabe que el riesgo de detener la prueba en esos momentos está dado por la expresión siguiente:

$$U_{n-1} = \sigma(\pi_{n-1}(\theta)) + \sum_{i=1}^{n-1} C_i \tag{3}$$

y se quiere comparar este valor contra el riesgo esperado de tomar la n-ésima muestra; lo cual queda determinado por:

$$E[U_n] = E[\sigma(\pi_{n-1}(\theta/X_n))] + \sum_{i=1}^{n} C_i \tag{4}$$

en donde la esperanza se toma sobre S (espacio muestral) debido a que se pretende determinar el beneficio en promedio que se obtendría de muestrear Xn; lo cual se puede expresar de la siguiente forma también:

$$E[U_n] = \int_S \sigma(\pi_{n-1}(\theta/X_n))\, f(X_n/\theta)\, dx \tag{5}$$

para la cual f es la función de probabilidades de la variable aleatoria Xn, dado el parámetro θ.

Nótese que $\sigma(\pi(\theta/x_1,x_2,\ldots,x_{n-1},X_n))$ es una función de Xn exclusivamente, por lo que tanto la esperanza como la integral, o sumatoria en su caso, se definen sobre el espacio S como se aclaró anteriormente; y lo único que indica es cómo utilizaría el estadístico la información dada por la variable aleatoria Xn.

El procedimiento óptimo y la regla de parada (stopping), en este caso estaría dado por la siguiente desigualdad:

$$\text{si} \qquad U_{n-1} < E[U_n]$$

entonces, debido a que el riesgo o la pérdida en promedio por parar el procedimiento es menor, se debe de dejar de muestrear; si esto no sucede, entonces hay que seguir muestreando. Es decir, con esto se determina si en el paso n-1 del proceso se encuentra la parada óptima.

En vista de lo anterior, el riesgo del procedimiento óptimo V_{n-1} está dado por la siguiente expresión:

$$V_{n-1} = \min \{U_{n-1}, E[U_n]\}. \tag{6}$$

Para una muestra anterior, se podría considerar el mismo razonamiento; pero en lugar de tomar a la continuación óptima como a $E[U_{n-1}]$ se debe tomar a $E[V_{n-1}]$, ya que en el (n-1)-ésimo momento de la muestra el riesgo estará dado por V_{n-1} y no por U_{n-1}. Por lo que en el paso n-2 el riesgo del procedimiento óptimo estará dado por:

$$V_{n-2} = \min \{U_{n-2}, E[V_{n-1}]\}; \tag{7}$$

y se debe de continuar muestreando el n-1 ésimo elemento de la población si:

$$V_{n-2} = E[V_{n-1}] < U_{n-2} \tag{8}$$

Ahora bien antes de la estructuración del procedimiento de la parada óptima o, lo que es lo mismo, con la determinación del momento de la parada óptima; se define formalmente a V_j como el riesgo del procedimiento secuencial óptimo.

DEFINICION 4 .- Supóngase que se han observado j elementos de la población ($j=0,1,2 ...,n-1$), entonces el riesgo del procedimiento óptimo está dado por:

$$V_j = \min \{U_j, E[V_{j+1}]\} ;$$

y para $j=n$, defínase a:

$$V_n = \min \{U_n, U_n\} = U_n .$$

Entonces, en cualquier paso del procedimiento óptimo, el riesgo de éste mismo está dado por la definición anterior y si:

$$U_j \leq E[V_{j+1}] , \qquad (9)$$

entonces debe detenerse el muestreo, (obviamente para la primer j que cumpla esta restricción) y, en ese momento, se tiene la parada óptima. Es importante señalcar que, aunque en el caso en el que $U_j=E[V_{j+1}]$ se podría continuar el muestreo sin perder condiciones óptrimas, se debe detener la prueba ya que el tomar una observación más no brinda información extra considerable.

Se ha establecido el criterio de la inducción retrospectiva; y lo único que resta en estos momentos es hacer la anotación que resume el proceso y los indicadores generales.

INDUCCION RETROSPECTIVA

Entre todos los procedimientos de decisión secuencial, el óptimo esta dado por:

"Si se han muestreado j observaciones ($j=0,1, 2,...,n$) y se cumple que:

$U_j \leq E[V_{j+1}]$

y $\qquad U_i > E[V_{i+1}]$ para todo $i<j$,

entonces deténgase el muestreo, y tómese la decisión d^* tal que cumpla con

$$\sigma(\pi(\theta),d^*)= \inf_{d \in D} \int_{\Theta} L(\theta,d)\, \pi_j(\theta)\, d\theta \;;$$

en otro caso tómese la muestra $j+1$."

El riesgo bayesiano del procedimiento de decisión secuencial óptimo, según el método de la inducción retrospectiva, está dado por la definición 4.

En la siguiente sección se generalizan los conceptos, y se determina el procedimiento óptimo para procesos no acotados en el número de observaciones.

3.3 CASO GENERAL

Habiendo establecido el criterio de la inducción retrospectiva, reviste mucho interés el contar con un procedimiento para casos que no estén acotados en el número de observaciones. De manera intuitiva se puede observar que la metodología de la inducción retrospectiva no se puede aplicar directamente a estos casos; sin embargo, después de una serie de restricciones, se podrá ver en qué forma resulta posible llevar a cabo esta adaptación.

Es importante aclarar que, si bien se considerarán procesos no acotados en observaciones, es decir procesos generales, se tendrá que tomar en cuenta una restricción para ellos: se considerarán procesos para los cuales se cumpla que la probabilidad de que terminen sea uno, de la misma forma que se hizo en el capítulo I. Así los procesos considerados cumplen

$$\Pr\{N<\infty\}=\lim_{n\to\infty} \Pr\{N<n\}=1$$

Debido a que ya no sólo se trata de una decisión, sino a que ésta es todo un procedimiento de decisión, se modifica la notación utilizada en la sección anterior, referente al campo de decisiones y a sus elementos por la presentada a continuación.

DEFINICION 1 A una función $\alpha : S \longrightarrow @$ se le conoce como procedimiento de decisión secuencial. En donde $@$ es el espacio de procedimientos de decisión y S el espacio muestral; es decir:

$$\alpha(X) \in @ \qquad \text{para toda } X \in S.$$

Por otro lado, se entenderá como variable de alto (stopping) a aquella variable que indique cuántas observaciones se deben tomar en cada procedimiento. Nótese la estrecha relación que existe entre lo que es un procedimiento de decisión secuencial $\delta \in \mathscr{D}$ y la que es una variable de alto.

Debido a esta estrecha relación, se puede ver que, de hecho, un procedimiento δ está completamente determinado por la variable de alto respectiva, y viceversa; por esto, δ no solamente denotará al procedimiento, sino que también será utilizado como variable de alto, en la forma indicada en la siguiente definición.

DEFINICION 2 Para cualquier sucesión infinita X1,X2,X3, ... de observaciones en el espacio muestral S, la ecuación

$$\delta(X1,X2,X3,.....)=n$$

denota el número de muestras que se deben tomar antes de una decisión; de modo que el procedimiento muestral debe detenerse después de n observaciones, y no antes

Con estas breves modificaciones notacionales, es posible iniciar la determinación y demostración de la existencia del procedimiento óptimo.

Inicialmente, se anotan y definen algunos conceptos importantes para el cumplimiento del objetivo deseado; posteriormente, se establece la existencia de un procedimiento de decisión secuencial óptimo en $\mathscr{D}$; y, finalmente, se desarrolla un método por medio del cual se puede aproximar este procedimiento óptimo a través de la aplicación de la inducción retrospectiva.

DEFINICION 3 A una función de decisión δ definida en el espacio S^{∞}=SxSxSxSxSx.... con la siguiente propiedad

si δ(X1,X2,X3,...)=n

para alguna secuencia infinita de

observaciones y si Y1,Y2,Y3,... es otra sucuencia infinita de observaciones tales que:

$$X_i=Y_i \quad \text{para todo } i=1,2,3,\ldots n$$

entonces se tiene que

$$\Omega(X_1,X_2,X_3,\ldots)=\Omega(Y_1,Y_2,Y_3,\ldots)=n$$

y a esta última se le denomina variable de alto.

DEFINICION 4 El conjunto $\{\Omega=n\}$ es aquel conjunto que contiene a todas las sucesiones infinitas de observaciones X1,X2,X3,... tales que terminan su procedimiento de decisión secuencial después de n observaciones, y no antes.

DEFINICION 5 A un procedimiento de decisión secuencial $\Omega \in \mathcal{Q}$ se le llama regular si para todo $(x_1,x_2,x_3,\ldots x_n) \in \{\Omega>n\}$ se cumple que:

$$E\{\sigma(\pi,\Omega)/x_1,x_2,\ldots x_n,X_{n+1}\}<\sigma\{\pi_n(\theta)\}+nc$$

para n=1,2,3,...

en donde c es el costo por muestreo.

Esta último definición indica, solamente, que un procedimiento es regular cuando el riesgo de continuar es menor que el riesgo de detener la prueba, si Ω indica que hay que tomar más observaciones.

Debido a que estos procedimientos son congruentes y naturales, el presente trabajo se restringe sólo a aquellos que cumplen con esta definición

DEFINICION 6 Para un procedimiento de decisión secuencial regular $\Omega \in \mathcal{Q}$, y para cualquier entero positivo n, $\Omega' \in \mathcal{Q}$ será el procedimiento acotado tal que $\Omega=\Omega'$ en las primeros n observaciones, pero bajo Ω' la prueba se termina después de n observaciones nada más. El procedimiento Ω' se dice que se obtiene de truncar Ω en n observaciones. Si en particular $\Omega \leq n$ entonces $\Omega = \Omega'$.

DEFINICION 7 Si $\Omega 1, \Omega 2, \ldots \Omega K$ son K procedimientos de decisión secuencial, entonces el procedimiento definido por:

$$\text{máx}\{\Omega 1, \Omega 2, \ldots, \Omega K\} \in @$$

es el procedimiento que indica que otra observación se debe tomar cuando al menos un Ωi asi lo indique $(i=1,2,\ldots K)$; de manera análoga se define el procedimiento

$$\sup\{\Omega 1, \Omega 2, \ldots, \Omega K\} \in @.$$

Ya con estas definiciónes se puede continuar con la demostración de la existencia de un procedimiento óptimo en @; para luego, y como se anotó anteriormente, determinar éste mismo y, finalmente, aplicar el método de la inducción retrospectiva.

LEMA 1 Sean $\Omega 1 \in @$ y $\Omega 2 \in @$ procedimientos de decisón secuencial regulares; entonces

$\Omega = \text{máx}\ \{\Omega 1, \Omega 2\}$ es también regular y

$$\sigma(\pi, \Omega) \leq \sigma(\pi, \Omega i) \quad i=1,2.$$

DEMOSTRACION Supóngase que $\Omega 1 > \Omega 2$; entonces, por la definicón 7, $\Omega = \Omega 1$; y como $\Omega 1$ es regular entonces también Ω lo es; de la misma forma para $\Omega 2$.

Ahora, para un punto $(x1, x2, \ldots, xn)$ (considérese sin pérdida de generalidad que $\Omega 1 > n$), entonces se cumple que:

$$E[\sigma(\pi, \Omega)/X1, X2, \ldots Xn] = E[\sigma(\pi, \Omega 1)/X1, X2, \ldots Xn]$$

$<E[\sigma(\pi n(\theta)]+nc$

$=E[\sigma(\pi,\Omega 2)/X1,X2,...Xn]$

lo que implica que:

$\sigma(\pi,\Omega)\leq\sigma(\pi,\Omega 2)$

y de manera similar puede verse que:

$\sigma(\pi,\Omega)\leq\sigma(\pi,\Omega 1)$

L.Q.Q.D.

LEMA 2 Sea $\Omega i \in \Theta$ $(i=1,2,...)$ una secuencia de procedimientos de decisión regular; y sea:

$Gn=\text{máx}\ \{\Omega 1,\Omega 2,...\Omega n\}$

para $n=1,2,3,.....$ entonces, para todo n, Gn es regular y se cumple que:

$\sigma(\pi,Gn)\leq\sigma(\pi,Gi)\quad i=1,2,3...n$

y $\quad\sigma(\pi,Gn)\geq\sigma(\pi,Gn+1)$

DEMOSTRACION Se puede observar que:

$G1=\Omega 1$

$G2=\text{máx}\ \{\Omega 1,\Omega 2\}$

y por el lema 1; G1 y G2 son regulares.

Ahora:

$G3=\text{máx}\ \{\Omega 1,\Omega 2,\Omega 3\}=\text{máx}\ \{\Omega 1,G2\}$

por lo que G3 también es regular, y así, inductivamente:

$Gn=\text{máx}\ \{\Omega 1,Gn-1\}$

es regular para $n=1,2,3...$

Por otro lado, la primera relación se cumple para $n=1,2$, es decir:

$\sigma(\pi,\Omega 1) \leq \sigma(\pi,G1)$

$\sigma(\pi,G2) \leq \sigma(\pi,\Omega i) \quad i=1,2$ (lema 1)

y la segunda relación se cumple para $n=1$

$\sigma(\pi,G1) \geq \sigma(\pi,mG2)$ (lema 1)

y finalmente, definiendo a $Gn=\text{máx}\{\Omega 1,Gn-1\}$, se prueban por inducción las dos relaciones para $n=1,2,3,\ldots$

L.Q.Q.D.

LEMA 3 Sean Ωi y Gi ($i=1,2,3,\ldots$) como se definieron en el lema 2 y sea $G=\sup\{\Omega 1,\Omega 2,\ldots\}$, entonces:

$$Pr\{G<\infty\}=1$$

DEMOSTRACION como $\Omega i \in \mathcal{Q}$ para todo i; entonces:

$\sup\{\Omega 1,\Omega 2,\Omega 3,\ldots\} \in \mathcal{Q}$; entonces:

$G \in \mathcal{Q}$ y $Pr\{G<\infty\}=1$

L.Q.Q.D.

LEMA 4 Sean Ωi Gi ($i=1,2,3,\ldots$) como se definieron en los lemas anteriores; entonces, para $i=1,2,3\ldots$, se cumple que:

$\sigma(\pi,G) \leq \sigma(\pi,Gi) \leq \sigma(\pi,\Omega i)$

DEMOSTRACION Por la aplicación directa del lema 2, relación 1, se sabe que:

$\sigma(\pi,Gi) \leq \sigma(\pi,\Omega i) \quad i=1,2,3\ldots$

y por la relación 2 del mismo lema:

$\sigma(\pi,Gn+1) \leq \sigma(\pi,Gn)$, lo que implica que:

$\sigma(\pi,G) \leq \sigma(\pi,Gi) \quad i=1,2,3,\ldots$

y, en general:

$\sigma(\pi,G) \leq \sigma(\pi,Gi) \leq \sigma(\pi,\Omega i) \quad$ para $\quad i=1,2,3\ldots$

L.Q.Q.D.

En estos momentos, y con la ayuda de los lemas anteriores por su aplicación directa; el siguiente teorema demuestra la existencia de un procedimiento óptimo en $\mathcal{C}$.

TEOREMA 1 Existe un procedimiento de decisión secuencial óptimo en $\mathcal{C}$.

DEMOSTRACION Sea $\Omega 1,\Omega 2,\ldots$ una secuencia de procedimientos en la clase $\mathcal{C}$ tales que:

$$\lim_{i\to\infty} \sigma(\pi,\Omega i) = \inf_{\Omega \in \mathcal{C}} \sigma(\pi,\Omega)$$

se puede asumir que Ωi $(i=1,2,\ldots)$ es regular.

Ahora sea $\Omega^*=\sup \{\Omega 1,\Omega 2,\ldots\}$, entonces, por los lemas 2 y 3: $\Omega^* \in \mathcal{C}$ y:

$\sigma(\pi,\Omega^*) \leq \sigma(\pi,\Omega i)$ para $i=1,2,3,\ldots$ y

$$\sigma(\pi,\Omega^*) \leq \inf_{\Omega \in \mathcal{C}} \sigma(\pi,\Omega)$$

y como $\Omega^* \in \mathcal{C}$, entonces también se cumple que:

$$\sigma(\pi,\Omega^*) = \inf_{\Omega \in \mathcal{C}} \sigma(\pi,\Omega) \qquad y$$

Ω^* es óptimo y pertenece a $\mathcal{C}$.

L.Q.Q.D.

Finalmente, y ya habiendo establecido formalmente la existencia de un procedimiento óptimo en la clase θ, se presentan un teorema y dos corolarios que indican que con el uso de procedimientos truncados, se puede aproximar este procedimiento óptimo; utilizando para dicha aproximación la metodología de la inducción retrospectiva presentada en la sección anterior.

TEOREMA 2 Para $n=1,2,...$ sea Qn como se define adelante:

$$Qn=\int_{\{\Omega^*>n\}} \sigma(\pi n(\theta))fn(x1,x2,...xn/\theta)\pi(\theta)d\theta \ ,$$

ahora, para $n=1,2,3,...$ sea Ωn el procedimiento que se obtiene de truncar el procedimiento óptimo Ω^* después de n observaciones.

Si $\lim_{n\to\infty} Qn=0$ entonces $\lim_{n\to\infty} \sigma(\pi,\Omega n)=\sigma(\pi,\Omega^*)$

Para su prueba se puede hacer referencia al texto de De Groot; además, el teorema dice que si el procedimiento óptimo tiene riesgo finito puede ser aproximado por Ωn para n suficientemente grande.

En este teorema nótese que Qn es la eperanza del riesgo de no continuar y de tomar una decisión después de n observaciones; y que, finalmente, el riesgo de tomar una decisión se aproxima a cero cuando n tiende a infinito.

COROLARIO 1 Sea Ω^* y Ωn $(n=1,2,....)$ como se definieron en el teorema 2; entonces:

$$\lim_{n\to\infty} \sigma(\pi,\Omega n)=\sigma(\pi,\Omega^*)$$

se cumple, si se cumple alguna de las dos condiciones siguientes:

1) Existe un número K tal que, para todo $(x1,x2,x3,...xn) \in \{\Omega^*>n\}$ y para n=1,2,...

$$\sigma[\pi n(\theta)] < nK;$$

2) $\lim_{n\to\infty} E\{\sigma[\pi n(\theta)]\} = 0$

Y, finalmente, concluye esta sección con la presentación de un corolario y su demostración; el que indica que el procedimiento óptimo Ω^* puede ser aproximado por el procedimiento óptimo acotado; y éste se puede encontrar con la aplicación del método de la inducción retrospectiva.

COROLARIO 2 Si se satisface alguna de las condiciones del corolario 1; entonces se cumple que:

$$\lim_{n\to\infty} E[\sigma(\pi/X1,X2,...Xn)]=\sigma^*(\pi)$$

DEMOSTRACION Sea Ωn, como se define anteriormente, el procedimiento que se obtiene de truncar el procedimiento óptimo Ω^* después de n observaciones; se cumple, por las propiedades óptimas de Ω^* y de $E[\sigma(\pi/X1,X2,...Xn)]$, la relación:

$$\sigma^*(\pi) \leq E[\sigma(\pi/X1,X2,...Xn)] \leq \sigma(\pi,\Omega n)$$

entonces:

$$\sigma^*(\pi) = \lim_{n\to\infty} E[\sigma(\pi/X1,X2,X3,,,Xn)]$$

L.Q.Q.D.

Con esta última demostración, se dispone de una metodología para determinar el proceso de decisión secuencial óptimo para problemas generales.

3.4 DISCUSION Y COMENTARIOS

Se han visto y analizado la formalización y aplicación de la técnica de las pruebas secuenciales bayesianas y, particularmente, el problema de la parada óptima, que no es más que la elección del procedimiento de decisión secuencial que optimiza el plan muestral del estadístico.

No obstante es necesario concretar, de una manera palpable, los usos que a esta técnica se le pueden dar. Es por ello que, en el siguiente capítulo, se presenta un ejemplo práctico.

Debido a las características generales de esta técnica, el lector puede darse cuenta de que los procesos que se han presentado proporcionan un método para la resolución de problemas inferenciales y de pruebas de hipótesis.

Principalmente, esta metodología es útil para los casos en los que se deba de llevar a cabo una optimización de los presupuestos a los que pudiera estar sujeto el estadístico; ya que, como sucede en la realidad, los costos por muestreo y análisis de información son muy altos.

Sin embargo, no para todos los casos es ésta técnica la más adecuada. Esto es, existen infinidad de problemas que pueden ser resueltos, con mayor ventaja, por medio de la aplicación de alguna otra teoría inferencial.

Es pues, por esta razón, por la que el estadístico debe estar familiarizado con otras metodologías alternativas de solución -para poder discernir entre ellas y la presentada en el texto, evaluando las ventajas y desventajas que cada una de ellas presente.

Por otro lado, de la filosofía que sustenta a las pruebas secuenciales bayesianas surgen nuevas inquietudes, algunas ya teorizadas, que seguramente el estadístico y los estudiosos de la materia encontrarán interesantes para análisis y aplicación práctica.

El problema de la parada óptima en casos mas complejos -por ejemplo, para martingalas o procesos de

Markov- es fuente de investigaciones teóricas como de aplicaciones prácticas hoy en día.

Adicionalmente a lo anotado en secciones anteriores a la presente, la técnica de la inducción retrospectiva está limitado a casos en los que exista un número fijo de observaciones muestreables. Aun cuando se llega a generalizar para todos los casos, inclusive los no acotados en observaciones, se puede ver que se pierde eficiencia para éstos y que, en tal caso, los cálculos numéricos se hacen tediosos y largos.

Pero, por otro lado, el control de ambos tipos de error suple y supera las desventajas antes mencionadas, en términos de resultados; ya que con ello se consiguen decisiones más acertadas y razonables, estadísticamente hablando.

Sin embargo, pese a todos los comentarios y conclusiones que se pudieran hacer al respecto, es necesario que, para cada caso, el estadístico evalúe cuál es el mejor procedimiento inferencial, de acuerdo con sus objetivos y restricciones.

CAPITULO IV

EJEMPLO TIPO

4.1 INTRODUCCION Y CONSIDERACIONES

Habiendo desarrollado todo el marco teórico conceptual de las pruebas secuenciales bayesianas, se presenta, en este capítulo, el planteamiento y solución de un problema práctico con el objeto de concretar los conceptos expuestos en capítulos anteriores.

Es importante mencionar que el problema tratado es, en su planteamiento, realista; aun cuando la solución propuesta no se haya puesto en práctica.

Antes de iniciar el planteamiento del problema, se tienen que hacer algunas consideraciones para lograr un mejor entendimiento del mismo -además de una comprensión propia de la filosofía de la solución propuesta.

Una de estas consideraciones, en el momento del análisis de la solución presentada, es el hecho de que se adopta un punto de vista nuevo en la concepción de la función de pérdida; y más aún, en la interpretación que de ésta se dá.

El estadístico debe, en el contexto de la optimización en la toma de decisiones, considerar no sólo la pérdida como tal, sino que es de suma importancia el hecho de considerar, dentro de esta pérdida, el rechazo de una oportunidad que, en el caso de ser ventajosa, hubiese dejado un margen de utilidad considerable.

Esto es, el costo de oportunidad debe de jugar un papel preponderante en la toma de decisiones. De hecho, desde el planteamiento de la teoría, en el capítulo I, se menciona que una de las ventajas que ofrece el método de las pruebas secuenciales es, justamente, que en la estimación de un parámetro se pueden controlar ambos tipos de error; es decir, tanto el rechazo de una hipótesis verdadera (tipo I), como la aceptación de una hipótesis falsa (tipo II).

Con esta manera de pensar, como principio fundamental, se da solución al problema en cuestión.

El cálculo numérico de la solución se llevó a cabo con la ayuda de un computador, y se presentan los programas que se hicieron en el Anexo 1 (Programas 2,3 y 4).

Finalmente, cabe mencionar que, por el planteamiento y solución del problema, es posible extender el uso de esta solución y de los programas mencionados en el párrafo anterior, a una extensa variedad de nuevos problemas a los que se hará referencia en el último capitulo.

Se recomiendan, como bibliografía adicional, los ejemplos presentados en los textos de: Ferguson, Silvey y DeGroot.

4.2 PLANTEAMIENTO

La industria X desea adquirir un lote de autos para su departamento de promoción en ocasión del lanzamiento de un nuevo producto.

Debido a que el estudio de mercado previo al lanzamiento requiere de esta información, tal empresa desea determinar, de la mejor manera posible, los costos y el rendimiento de esta inversión. Adicionalmente se desea diseñar una estrategia de compra que optimice los resultados del desembolso económico que, en su caso, podría realizar la empresa.

Los autos serán proporcionados al personal del departamento de "nuevos productos" para la promoción, la cual durará solamente un año. Se establece que estos autos darán servicio a la empresa durante este periodo y, al término del año, se pondrán a la venta. Debido a esto, el director de compras de la empresa determinó que un lote de 10 autos usados, en buen estado, reúne las características necesarias para el logro de los objetivos propuestos.

Para los efectos de la compra, la compañía recibe la oferta de tres lotes diferentes de diez autos cada uno. Sin embargo, tales lotes se venden en paquetes (de los diez autos respectivos) y no pueden ser adquiridos parcialmente. Es decir, o se compra un lote completo de diez autos, o no se compra ningún auto de ese lote.

El director de compras de la compañía X considera atractivas las ofertas, y designa a un responsable de proyecto para que realice un estudio minucioso de dichas ofertas y analice la factibilidad de adquirir los lotes.

En una reunión que tiene el director de compras con el responsable del proyecto se concluye que, debido al precio tan atractivo de los lotes, se puede pensar en considerar como pérdida para la empresa tanto la pérdida que resulte de los autos malos, como la utilidad que se deje de realizar en caso de rechazar una buena oferta.

Antes de tomar una decisión, el responsable del proyecto desea realizar algunas pruebas a los coches, y pide a un mecánico que le proporcione un presupuesto por cada prueba. Además, le indica a este mecánico que, en los resultados de cada prueba, se deberá especificar: si el auto en turno está en condiciones de cumplir los objetivos deseados, dentro del año de servicio, sin fallas considerables que impliquen un desembolso significativo para la empresa X; o que el auto tiene fallas que lo hacen inútil para todo el año.

El mecánico decide que, por cada prueba, cobrará tres mil pesos y que ésta le tomará una hora hábil. Por otro lado, añade que determinará si el coche cumple con los requerimientos establecidos; además de que no dará un reporte detallado del estado físico del automovil; pero que se compromete y garantiza los autos que él mismo determine aptos contra cualquier reparación mayor, si es que éstos se llegan a comprar.

Por otro lado, el director de compras pone al servicio del responsable del proyecto a un estadistico, con el objeto de que determine una estrategia de muestreo y análisis de información que proporcione los elementos necesarios para poder tomar la decisión mas ventajosa para la estrategia de compra de la empresa.

Después de un análisis minucioso del caso, el estadistico propone, con base en la filosofia de perdidas propuesta por el director de compras y los costos por prueba, que la mejor alternativa la brinda el análisis de prubas secuenciales bayesianas; y que, en tal caso, se podria usar como función de pérdida a

$$L(\theta,a)=K(\theta-a)^2$$

en donde θ representa el porcentaje de autos buenos en el lote en estudio, dato desconocido; y a, el procedimiento de decisión y variable de alto, seria la estimación del parámetro θ, es decir:

$$a(X)=\hat{\theta}$$

en donde X es el muestreo; y, finalmente, k sería una constante que indica la pérdida por auto debida a una mala decisión.

Como justificación para el uso de esta función de pérdida, el estadístico argumenta que ésta mide ambos tipos de pérdida propuesta; de acuerdo con el responsable del proyecto, decide que el valor de la constante k puede ser el costo por unidad en cada lote, ya que se estima que la productividad por automóvil podría quedar, más o menos, determinada por el doble del valor de adquisición del mismo auto; de éste modo, restando su costo, la utilidad estimada sería igual a su costo de compra.

Con base en estas consideraciones, se deciden los siguientes valores para la constante k:

k=1'250,000.00	lote 1
k=1'500,000.00	lote 2
k=1'750,000.00	lote 3.

Se acepta la propuesta del estadístico y 'se dá luz verde' al estudio.

Para la determinación de las funciones de probabilidad a priori de θ el estadístico piensa que, ya que el primer grupo tiene una pérdida unitaria igual al costo unitario, más pequeño, es posible que en este lote exista una probabilidad más baja de que un auto esté en buen estado; de igual forma, una probabilidad un poco más alta para el segundo grupo, y la más alta para el tercero. En tal forma que las funciones a priori son

$\pi(\theta)=\beta e(1,2)$ para el lote 1

$\pi(\theta)=\beta e(1,1)$ para el lote 2

$\pi(\theta)=\beta e(2,1)$ para el lote 3

en donde βe es la función de distribución de probabilidades beta:

$$Be(a,b)=\frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\,\theta^{a-1}(1-\theta)^{b-1}$$

con parámetros a y b (para una consulta teórica de esta función así como de sus principales propiedades, refiérase al Anexo 2).

Finalmente, como las pruebas de los coches pueden tener solamente dos valores, a saber:

$$X_i=\begin{cases}1 & \text{si el coche está en buen estado}\\ 0 & \text{si el coche está en mal estado;}\end{cases}\qquad i=1,2,..10$$

se consideraron, en términos estadísticos, ensayos de Bernoulli con parámetro θ esto para cada grupo por separado.

4.3 DESARROLLO DE UNA SOLUCION

Es importante, antes de establecer y desrrollar una solución al problema de la compañía X, que se revise todo el material incluido en el Anexo 2 sobre la función de distribución beta y se considere la siguiente definición:

DEFINICION 1 Supóngase que se han observado j automóviles de un determinado lote, entonces defínase a:

$$S_j = \sum_{i=1}^{j} x_i \; ;$$

claramente S_j es el número de autos buenos después de j observaciones.

Para una distribución de probabilidades beta, su esperanza y varianza quedan determinadas por:

$$E[\theta]=a/(a+b)$$

$$Var[\theta]=ab/(a+b)^2(a+b+1)$$

(consúltese Anexo 2); y por el teorema 4 del Anexo 2 se sabe que la distribución a posteriori, después de j observaciones para cada grupo es

$$\pi(\theta/S_j)= \frac{\Gamma(a'+b')}{\Gamma(a')\,\Gamma(b')}\;\theta^{a'-1}\,(1-\theta)^{b'-1} \quad ,$$

con: $a'=a+S_j$ y $b'=b+j-S_j$.

Como se anotó en su oportunidad, en la sección anterior, el objetivo principal para el estadístico es la determinación del procedimiento secuencial óptimo que incluye la elección de la variable de alto y la decisión que minimice el riesgo.

Tal meta se resume, en términos matemáticos, en la expresión siguiente, que compara los riesgos tanto de detener la prueba y de tomar una decisión, como de continuar muestreando una unidad más; es decir:

$$V_j = \min \{U_j, E[V_{j+1}]\} \ , \qquad j=0,1,2,3...9 \ ,$$

$$y \qquad V_{10}=U_{10} \ ;$$

para lo cual si $V_j=U_j$, entonces el proceso muestral debe de detenerse y estimarse θ; de otro modo se debe checar el siguiente auto. Todo el procedimiento se hace para cada grupo de coches por separado y, al final, se comparan los riesgos iniciales para determinar con qué lote se inicia la muestra.

Es importante aclarar que la elección de los coches, en un determinado lote, se debe hacer aleatoriamente; ya que de otra manera toda la prueba pierde su sentido.

Así, resulta que lo primero que se debe de determinar es el valor de U_j, dadas j observaciones; para lo cual se procede de la siguiente manera:

$$U_j = \sigma(\pi(\theta/S_j)) = \min \int_\theta K(\theta-\hat{\theta})^2 \pi(\theta/S_j) \, d\theta$$

esta expresión se minimiza cuando:

$$U_j = Var(\theta) = ab/(a+b)^2(a+b+1)$$

para

$$\hat{\theta} = E[\theta] = a/(a+b)$$

de modo que para cada grupo se tiene, respectivamente:

Grupo 1 $\hat{\theta}=(1+S_j)/(j+3)$

$$V_j=(1+S_j)(2+j-S_j)/(j+3)^2(j+4)$$

Grupo 2 $\hat{\theta}=(1+S_j)/(j+2)$

$$V_j=(1+S_j)(1+j-S_j)/(j+2)^2(j+3)$$

Grupo 3 $\hat{\theta}=(2+S_j)/(j+3)$

$$V_j=(2+S_j)(1+j-S_j)/(j+3)^2(j+4)$$

valores que deben ser comparados contra $E[V_{j+1}]$.

Pero, para conocer $E[V_{j+1}]$ es necesario antes haber determinado las probabilidades condicionales de X_{j+1} dado S_j; las cuales se obtienen, directamente, del estimador de θ después de j observaciones que por lotes son:

Grupo 1 $P(X_{j+1}=1/S_j)=(1+S_j)/(j+3)$

$$P(X_{j+1}=0/S_j)=(j+2-S_j)/(j+3)$$

Grupo 2 $P(X_{j+1}=1/S_j)=(1+S_j)/(j+2)$

$$P(X_{j+1}=0/S_j)=(j+1-S_j)/(j+2)$$

Grupo 3 $P(X_{j+1}=1/S_j)=(2+S_j)/(j+3)$

$$P(X_{j+1}=0/S_j)=(j+1-S_j)/(j+3)$$

Con estos resultados se puede obtener $E[V_{j+1}]$; con lo que finalmente, queda completamente determinado el procedimiento óptimo para el que, si el estimador de θ es muy bajo y se debe detener la prueba, entoces no se debe comprar; si el estimador es alto, también se debe de detener la prueba; entonces, el lote debe ser adquirido; y, de otra forma, continuar la prueba.

4.4 CALCULO DE LA SOLUCION

Ya teniendo determinado el procedimiento óptimo, lo único que resta por hacer es realizar los cálculos numéricos que, por su complejidad y cantidad, se corren en una computadora, cuyos programas se presentan en el anexo uno (programas dos, tres y cuatro); para los riesgos de parar después de j observaciones, de continuar después de j observaciones, y probabilidades condicionales de X_{j+1} dado S_j, respectivamente.

Debido a que el estadístico todavía no realiza las pruebas de campo, los programas son exhaustivos; esto es, se toman en cuenta todas las posibilidades para S_j (de 0 a j) y para j (de 0 a 10).

En las tablas 4.1, 4.6 y 4.11 se presentan los valores de la matriz de riesgos de la parada después de j observaciones; en las tablas 4.2, 4.7 y 4.12 los valores de la matriz de riesgos esperados de la continuación después de j observaciones; en las tablas 4.3, 4.8 y 4.13 los valores del procedimiento óptimo, que no es mas que los mínimos entre los valores de los dos conceptos anteriores; y en las tablas 4.4 y 4.5, 4.9 y 4.10, 4.14 y 4.15, los valores de las condicionales de X_{j+1} dadas j observaciones, para 1 y 0, respectivamente; todas ellas para los tres lotes. Esto es:

Tablas 4.1, 4.6, 4.11

$$U_j$$

Tablas 4.2, 4.7, 4.12

$$E[V_{j+1}]$$

Tablas 4.3, 4.8, 4.13

$$V_j=\min \{U_j, E[V_{j+1}]\}$$

Tablas 4.4, 4.9, 4.14

$$P(X_{j+1}=1/S_j)$$

Tablas 4.5, 4.10, 4.15

$$P(X_{j+1}=0/S_j)$$

todas ellas para:

$$S_j=0,1,2,3,\ldots,j$$

$$j=0,1,2,3,\ldots,10.$$

Con todos estos valores, el estadístico está en la posibilidad de hacer el análisis final, y de presentar los resultados y estrategias de muestreo y toma de decisiones que optimizan el procedimiento de las pruebas secuenciales bayesianas para este problema en particular. Información que se presenta en la siguiente sección.

TABLA 4.1

GRUPO 1

RIESGO DE LA PARADA DESPUES DE J OBSERVACIONES

AUTOS OBSERVADOS

		0	1	2	3	4	5	6	7	8	9	10
B	0	69444	49875	39333	33802*	31133	30191	30346	31227	32609	34345	36340
U	1		65500	56000	48683	43888	41042	39605	39182	39496	40355	41623
E	2			56000	53643	50265	47552	45778	44864	44661	45029	45850
N	3				48683	50265	49722	48864	48273	48105	48368	49019
O	4					43888	47552	48864	49409	49826	50371	51133
S	5						41042	45778	48273	49826	51038	52189
	6							39605	44864	48105	50371	52189
	7								39182	44661	48368	51133
	8									39496	45029	49019
	9										40355	45850
	10											41623

TABLA 4.2

GRUPO 1

RIESGO DE LA CONTINUACION DESPUES DE J OBSERVACIONES

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	46026	44423	42821	41447	40302	39348	38547	37867	37284	36780	36340
U	1		49231	49231	48315	47170	46026	44957	43986	43112	42327	41623
E	2			49231	50604	50604	50032	49231	48357	47483	46642	45850
N	3				48315	50604	51368	51368	50979	50396	49724	49019
O	4					47170	50032	51368	51853	51853	51573	51133
S	5						46026	49231	50979	51853	52189	52189
	6							44957	48357	50396	51573	52189
	7								43986	47483	49724	51133
	8									43112	46642	49019
	9										42327	45850
	10											41623

TABLA 4.3

GRUPO 1

RIESGO DEL PROCEDIMIENTO OPTIMO

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	46026	44423	39333	33802	31133	30191	30346	31227	32609	34345	36340
U	1		49231	49231	48315	43888	41042	39605	39182	39496	40355	41623
E	2			49231	50604	50265	47552	45778	44864	44661	45029	45850
N	3				48315	50265	49722	48864	48273	48105	48368	49019
O	4					43888	47552	48864	49409	49826	50371	51133
S	5						41042	45778	48273	49826	51038	52189
	6							39605	44864	48105	50371	52189
	7								39182	44661	48368	51133
	8									39496	45029	49019
	9										40355	45850
	10											41623

TABLA 4.4

GRUPO 1

$P(X_{j+1}=1/S_j)$ CONDICIONAL DE $X_{j+1}=1$ DADO S_j

AUTOS		OBSERVADOS										
		0	1	2	3	4	5	6	7	8	9	10
B	0	0.333333	0.250000	0.200000	0.166667	0.142857	0.125000	0.111111	0.100000	0.090909	0.083333	0.076923
U	1		0.500000	0.400000	0.333333	0.285714	0.250000	0.222222	0.200000	0.181818	0.166667	0.153846
E	2			0.600000	0.500000	0.428571	0.375000	0.333333	0.300000	0.272727	0.250000	0.230769
N	3				0.666667	0.571429	0.500000	0.444444	0.400000	0.363636	0.333333	0.307692
O	4					0.714286	0.625000	0.555556	0.500000	0.454545	0.416667	0.384615
S	5						0.750000	0.666667	0.600000	0.545455	0.500000	0.461538
	6							0.777778	0.700000	0.636364	0.583333	0.538462
	7								0.800000	0.727273	0.666667	0.615385
	8									0.818182	0.750000	0.692308
	9										0.833333	0.769231
	10											0.846154

TABLA 4.5

GRUPO 1

$P(X_{j+1}=0/S_j)$ CONDICIONAL DE $X_{j+1}=0$ DADO S_j

AUTOS		OBSERVADOS										
		0	1	2	3	4	5	6	7	8	9	10
B	0	0.666667	0.750000	0.800000	0.833333	0.857143	0.875000	0.888889	0.900000	0.909091	0.916667	0.923077
U	1		0.500000	0.600000	0.666667	0.714286	0.750000	0.777778	0.800000	0.818182	0.833333	0.846154
E	2			0.400000	0.500000	0.571429	0.625000	0.666667	0.700000	0.727273	0.750000	0.769231
N	3				0.333333	0.428571	0.500000	0.555556	0.600000	0.636364	0.666667	0.692308
O	4					0.285714	0.375000	0.444444	0.500000	0.545455	0.583333	0.615385
S	5						0.250000	0.333333	0.400000	0.454545	0.500000	0.538462
	6							0.222222	0.300000	0.363636	0.416667	0.461538
	7								0.200000	0.272727	0.333333	0.384615
	8									0.181818	0.250000	0.307692
	9										0.166667	0.230769
	10											0.153846

TABLA 4.6

GRUPO 2
RIESGO DE LA PARADA DESPUES DE J OBSERVACIONES

AUTOS		OBSERVADOS										
		0	1	2	3	4	5	6	7	8	9	10
B	0	125000	86333	62250	49000	41762	37959	36229	35815	36273	37331	38814
U	1		86333	81000	69000	59619	53265	49250	46926	45818	45595	46026
E	2			62250	69000	65571	60918	57063	54333	52636	51793	51635
N	3				49000	59619	60918	59667	58037	56727	55926	55641
O	4					41762	53265	57063	58037	58091	57992	58045
S	5						37959	49250	54333	56727	57992	58046
	6							36229	46926	52636	55926	58045
	7								35815	45818	51793	55641
	8									36273	45595	51635
	9										37331	46026
	10											38814

TABLA 4.7

GRUPO 2

RIESGO DE LA CONTINUACION DESPUES DE J OBSERVACIONES

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	50833	50833	48750	46667	44881	43393	42153	41111	40227	39470	38814
U	1		50833	55000	55000	53810	52321	50833	49444	48182	47045	46026
E	2			48750	55000	56786	56786	56042	55000	53864	52727	51635
N	3				46667	53810	56786	57778	57778	57273	56515	55641
O	4					44881	52321	56042	57778	58409	58409	58045
S	5						43393	50833	55000	57273	58409	58846
	6							42153	49444	53864	56515	58045
	7								41111	48182	52727	55641
	8									40227	47045	51635
	9										39470	46026
	10											38814

TABLA 4.8

GRUPO 2
RIESGO DEL PROCEDIMIENTO OPTIMO

AUTOS		OBSERVADOS										
		0	1	2	3	4	5	6	7	8	9	10
B	0	50833	50833	48750	46667	41762	37959	36229	35815	36273	37331	38814
U	1		50833	55000	55000	53810	52321	49250	46926	45818	45595	46026
E	2			48750	55000	56786	56786	56042	54333	52636	51793	51635
N	3				46667	53810	56786	57778	57778	56727	55926	55641
O	4					41762	52321	56042	57778	58091	57992	58045
S	5						37959	49250	54333	56727	57992	58846
	6							36229	46926	52636	55926	58045
	7								35815	45818	51793	55641
	8									36273	45595	51635
	9										37331	46026
	10											38814

TABLA 4.9

GRUPO 2

$P(X_{j+1}=1/S_j)$ CONDICIONAL DE $X_{j+1}=1$ DADO S_j

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	0.500000	0.333333	0.250000	0.200000	0.166667	0.142857	0.125000	0.111111	0.100000	0.090909	0.083333
U	1		0.666667	0.500000	0.400000	0.333333	0.285714	0.250000	0.222222	0.200000	0.181818	0.166667
E	2			0.750000	0.600000	0.500000	0.428571	0.375000	0.333333	0.300000	0.272727	0.250000
N	3				0.800000	0.666667	0.571429	0.500000	0.444444	0.400000	0.363636	0.333333
O	4					0.833333	0.714286	0.625000	0.555556	0.500000	0.454545	0.416667
S	5						0.857143	0.750000	0.666667	0.600000	0.545455	0.500000
	6							0.875000	0.777778	0.700000	0.636364	0.583333
	7								0.888889	0.800000	0.727273	0.666667
	8									0.900000	0.818182	0.750000
	9										0.909091	0.833333
	10											0.916667

TABLA 4.10

GRUPO 2

$P(X_{j+1}=0/S_j)$ CONDICIONAL DE $X_{j+1}=0$ DADO S_j

AUTOS OBSERVADOS

		0	1	2	3	4	5	6	7	8	9	10
B	0	0.500000	0.666667	0.750000	0.800000	0.833333	0.857143	0.875000	0.888889	0.900000	0.909091	0.916667
U	1		0.333333	0.500000	0.600000	0.666667	0.714286	0.750000	0.777778	0.800000	0.818182	0.833333
E	2			0.250000	0.400000	0.500000	0.571429	0.625000	0.666667	0.700000	0.727273	0.750000
N	3				0.200000	0.333333	0.428571	0.500000	0.555556	0.600000	0.636364	0.666667
O	4					0.166667	0.285714	0.375000	0.444444	0.500000	0.545455	0.583333
S	5						0.142857	0.250000	0.333333	0.400000	0.454545	0.500000
	6							0.125000	0.222222	0.300000	0.363636	0.416667
	7								0.111111	0.200000	0.272727	0.333333
	8									0.100000	0.181818	0.250000
	9										0.090909	0.166667
	10											0.083333

TABLA 4.11

GRUPO 3

RIESGO DE LA PARADA DESPUES DE J OBSERVACIONES

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	97222	90500	76000	64556	56643	51458	48247	46455	45694	45697	46272
U	1		68625	76000	71500	65571	60573	56889	54409	52926	52240	52189
E	2			52667	64556	65571	63611	61210	59182	57747	56915	56627
N	3				43722	56643	60573	61210	60773	60157	59719	59586
O	4					38786	51458	56889	59182	60157	60654	61065
S	5						36267	48247	54409	57747	59719	61065
	6							35284	46455	52926	56915	59586
	7								35318	45694	52240	56627
	8									36052	45697	52189
	9										37283	46272
	10											38876

TABLA 4.12

GRUPO 3

RIESGO DE LA CONTINUACION DESPUES DE J OBSERVACIONES

AUTOS		OBSERVADOS										
		0	1	2	3	4	5	6	7	8	9	10
B	0	52436	56923	56923	55641	54038	52436	50940	49580	48357	47258	46272
U	1		50192	56923	58846	58846	58045	56923	55699	54476	53299	52189
E	2			47949	55641	58846	59915	59915	59371	58555	57613	56627
N	3				46026	54038	58045	59915	60594	60594	60202	59586
O	4					44423	52436	56923	59371	60594	61065	61065
S	5						43088	50940	55699	58555	60202	61065
	6							41966	49580	54476	57613	59586
	7								41014	48357	53299	56627
	8									40198	47258	52189
	9										39492	46272
	10											38876

TABLA 4.13

GRUPO 3
RIESGO DEL PROCEDIMIENTO OPTIMO

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	52436	56923	56923	55641	54038	51458	48247	46455	45694	45697	46272
U	1		50192	56923	58846	58846	58045	56889	54409	52926	52240	52189
E	2			47949	55641	58846	59915	59915	59182	57747	56915	56627
N	3				43722	54038	58045	59915	60594	60157	59719	59586
O	4					38786	51458	56889	59182	60157	60654	61065
S	5						36267	48247	54409	57747	59719	61065
	6							35284	46455	52926	56915	59586
	7								35318	45694	52240	56627
	8									36052	45697	52189
	9										37283	46272
	10											38876

TABLA 4.14

GRUPO 3

$P(X_{j+1}=1/S_j)$ CONDICIONAL DE $X_{j+1}=1$ DADO S_j

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	0.666667	0.500000	0.400000	0.333333	0.285714	0.250000	0.222222	0.200000	0.181818	0.166667	0.153846
U	1		0.750000	0.600000	0.500000	0.428571	0.375000	0.333333	0.300000	0.272727	0.250000	0.230769
E	2			0.800000	0.666667	0.571429	0.500000	0.444444	0.400000	0.363636	0.333333	0.307692
N	3				0.833333	0.714286	0.625000	0.555556	0.500000	0.454545	0.416667	0.384615
O	4					0.857143	0.750000	0.666667	0.600000	0.545455	0.500000	0.461538
S	5						0.875000	0.777778	0.700000	0.636364	0.583333	0.538462
	6							0.888889	0.800000	0.727273	0.666667	0.615385
	7								0.900000	0.818182	0.750000	0.692308
	8									0.909091	0.833333	0.769231
	9										0.916667	0.846154
	10											0.923077

TABLA 4.15

GRUPO 3

$P(X_{j+1}=0/S_j)$ CONDICIONAL DE $X_{j+1}=0$ DADO S_j

AUTOS OBSERVADOS

		0	1	2	3	4	5	6	7	8	9	10
B	0	0.333333	0.500000	0.600000	0.666667	0.714286	0.750000	0.777778	0.800000	0.818182	0.833333	0.846154
U	1		0.250000	0.400000	0.500000	0.571429	0.625000	0.666667	0.700000	0.727273	0.750000	0.769231
E	2			0.200000	0.333333	0.428571	0.500000	0.555556	0.600000	0.636364	0.666667	0.692308
N	3				0.166667	0.285714	0.375000	0.444444	0.500000	0.545455	0.583333	0.615385
O	4					0.142857	0.250000	0.333333	0.400000	0.454545	0.500000	0.538462
S	5						0.125000	0.222222	0.300000	0.363636	0.416667	0.461538
	6							0.111111	0.200000	0.272727	0.333333	0.384615
	7								0.100000	0.181818	0.250000	0.307692
	8									0.090909	0.166667	0.230769
	9										0.083333	0.153846
	10											0.076923

4.5 CONCLUSIONES Y PRESENTACION DE RESULTADOS

Después de analizar las tablas presentadas en la sección anterior, el estadístico conforma las tablas 4.1, 4.3 y 4.5; que determinan, de manera única, el procedimiento óptimo que presenta al responsable del proyecto, junto con las gráficas de continuación anexadas en las tablas 4.2, 4.4 y 4.6, para cada lote de autos, respectivamente.

Para la interpretación de las tablas, el estadístico informa que los valores tienen la interpretación siguiente:

1 seguir checando autos;

2 detener la prueba y comprar el lote;

3 detener la prueba y rechazar la oferta.

Sin embargo, antes de iniciar las pruebas, es necesario que se indique con qué grupo se debe iniciar el muestreo, con cuál se debe de seguir en caso de no comprar el primer lote, y con cuál se debe de finalizar en caso de no comprar lo dos lotes iniciales.

Se sabe que los valores iniciales respectivos de riesgo son:

69,444.44

125,000.00

97,222.22,

por lo que el estadístico aconseja iniciar el muestreo con el grupo uno (riesgo inicial menor), continuar con el grupo tres, en caso de ser necesario, y finalizar con el dos.

Con todo esto queda debidamente determinado el procedimiento óptimo, y se puede continuar con las pruebas de campo.

Junto con esta información, el estadístico anota que, a lo más, se checarán, para cada grupo, cuatro, ocho y siete coches, respectivamente antes de tomar una decisión.

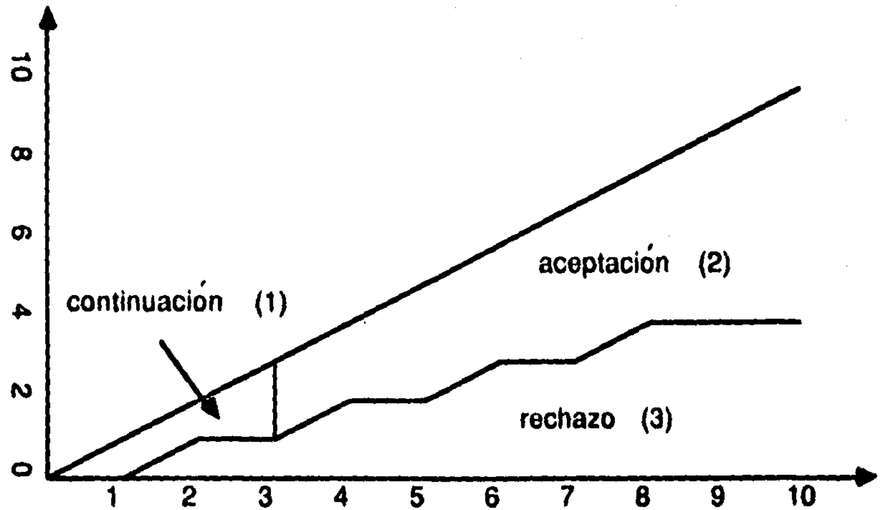
TABLA DE RESULTADOS 4.1

MATRIZ DE DECISION DEL GRUPO UNO

Autos Observados

		0	1	2	3	4	5	6	7	8	9	10
A	0	1	1	3	3	3	3	3	3	3	3	3
u	1		1	1	1	3	3	3	3	3	3	3
t	2			1	1	3	3	3	3	3	3	3
o	3				1	2	2	3	3	3	3	3
s	4					2	2	2	2	3	3	3
	5						2	2	2	2	2	2
b	6							2	2	2	2	2
u	7								2	2	2	2
e	8									2	2	2
n	9										2	2
o	10											2
s												

grupo 1



continuación (1)
aceptación (2)
rechazo (3)
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10

TABLA DE RESULTADOS 4.3

MATRIZ DE DECISION DEL GRUPO DOS

Autos Observados		0	1	2	3	4	5	6	7	8	9	10
A	0	1	1	1	1	3	3	3	3	3	3	3
u	1		1	1	1	1	1	3	3	3	3	3
t	2			1	1	1	1	1	3	3	3	3
o	3				1	1	1	1	1	3	3	3
s	4					2	1	1	1	2	3	3
	5						2	2	2	2	2	3
b	6							2	2	2	2	2
u	7								2	2	2	2
e	8									2	2	2
n	9										2	2
o	10											2
s												

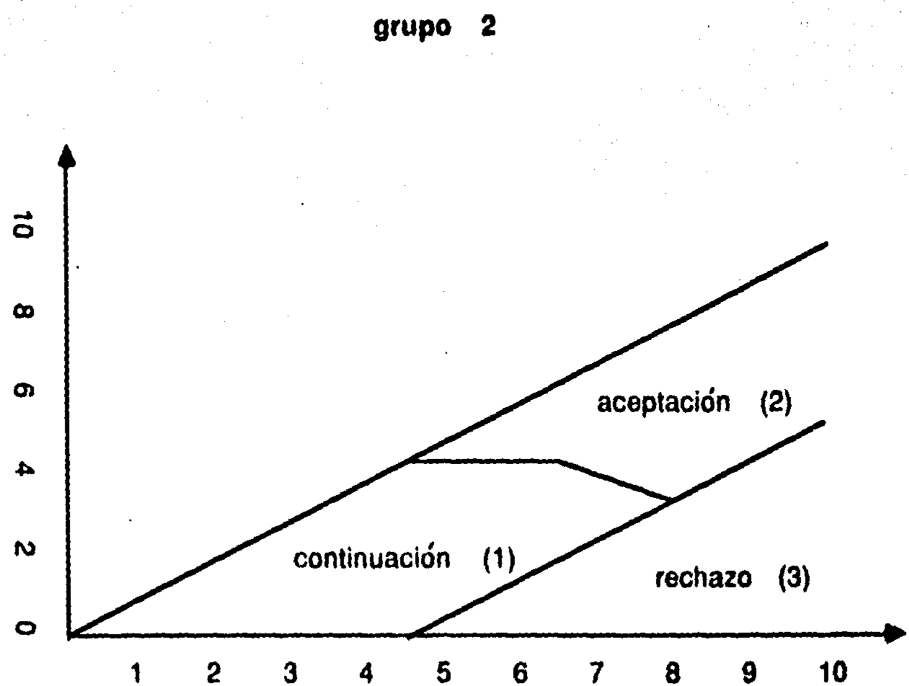

grupo 2
10
8
6
4
2
0
aceptación (2)
continuación (1)
rechazo (3)
1 2 3 4 5 6 7 8 9 10

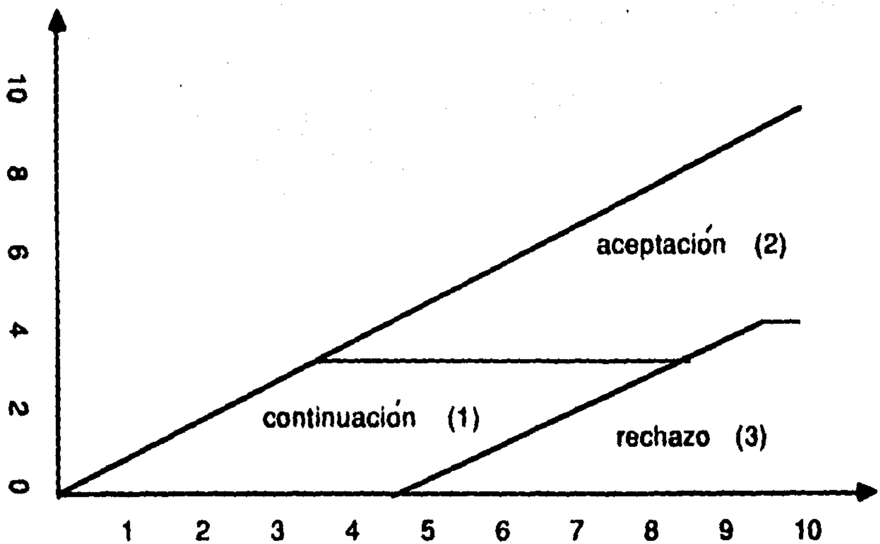
TABLA DE RESULTADOS 4.5

MATRIZ DE DECISION DEL GRUPO TRES

Autos Observados

		0	1	2	3	4	5	6	7	8	9	10
A	0	1	1	1	1	1	3	3	3	3	3	3
u	1		1	1	1	1	1	3	3	3	3	3
t	2			1	1	1	1	1	3	3	3	3
o	3				2	1	1	1	1	3	3	3
s	4					2	2	2	2	2	3	3
	5						2	2	2	2	2	3
b	6							2	2	2	2	2
u	7								2	2	2	2
e	8									2	2	2
n	9										2	2
o	10											2
s												

grupo 3


aceptación (2)
continuación (1)
rechazo (3)
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10

CAPITULO V

PROBLEMA PRACTICO

5.1 INTRODUCCION Y PLANTEAMIENTO

Finalmente, a manera de conclusión, se presenta el análisis y la solución a un nuevo problema cuyas consideraciónes y restricciónes se indican de una manera realista, de tal forma que el algoritmo de solución aquí planteado sea útil en la aplicación y solución de otros problemas.

PLANTEAMIENTO:

"Supóngase que el señor X está interesado en adquirir un lote de autos y desea minimizar la pérdida que este lote le puede proporcionar a través de los años de uso de cada auto. (en el caso de ganancia se puede considerar como pérdida negativa).

Supóngase además que la duración de cada auto es de cinco años, y que en otro caso éste será vendido. Para los efectos de la compra, el señor X tiene la oportunidad de probar los coches; para lo cuál cuenta con los servicios de un mecánico que cobrará un costo determinado por cada prueba.

Es importante aclarar que el resultado de cada prueba que lleve a cabo el mecánico solamente indicará si el auto se encuentra en buen estado o si éste necesita una reparación mayor.

Para la evaluación de la utilidad del lote, el señor X cuenta con:

- una tabla de pérdida/(utilidad) por auto, la cual está en función del tiempo esperado de duración de cada uno de éstos,

- una función de distribución a priori, que indica la probabilidad de que un coche esté en buen estado al momento de la realización de la prueba mecánica y

- una función de probabilidad para la duración en tiempo (años o fracciones) de cada auto.

Con todos estos elementos y con el uso de las pruebas de los coches a que tiene derecho, el señor X desea determinar la estrategia óptima de decisión para la compra o el rechazo del lote en cuestión.

Para tales fines el señor X pide a un estadístico le ayude a determinar el mejor plan muestral y el mejor análisis. El estadístico propone como alternativa de solución el análisis de las pruebas secuenciales bayesianas.*

Antes de seguir adelante con el desarrollo de una solución es necesario subrayar que, debido al nuevo planteamiento del problema, la solución presentada puede ser puesta en práctica cuando las condiciones así lo exijan. Es por ello que en un principio (sección 5.3) se desarrolla un algoritmo general para la resolución del problema y es hasta la sección 5.4 que se analiza un caso numérico.

En general la notación y los elementos estadísticos son los mismos que se utilizaron en el capítulo IV y solamente se establece la función de probabilidades de duración de cada auto como nueva definición, la cual se presenta en la siguiente sección junto con un sumario de la terminología que será utilizada para la presentación de la solución técnica del problema.

Nótese también la nueva elección de la función de pérdida y los resultados que se obtienen a partir de ésta.

5.2 DEFINICIONES Y NOTACION

Como se mencionó en su oportunidad se presenta un sumario de la notación que se ha venido utilizando en capítulos anteriores, la cuál no presenta variaciónes y queda de la forma siguiente:

θ	probabilida de que un auto esté en buen estado en el momento de la prueba mecánica,
$\pi(\theta)$	distribución de probabilidad a priori del parámetro θ,
$\pi_j(\theta)$	distribución de probabilidades a posteriori del parámetro θ después de j observaciones (distribución a priori después de j observaciones para la observación j+1),
$\beta e(a,b)$	distribución de probabilidad beta con parámetros a y b,
X_j	variable aleatoria que indica el resultado de la prueba del j-ésimo coche ($X_j=1$ indica que el coche está en buen estado y $X_j=0$ indica que el coche está en mal estado),
S_j	número de coches en buen estado después de j pruebas (Σ Xi),
L	función de pérdida
U_j	función de utilidad después de j pruebas y
V_j	riesgo de la continuación después de j observaciones.

Sin embargo, debido a la modificación de los términos del problema, es necesario hacer uso de un elemento nuevo para la busqueda de una solución. Este elemento se define en seguida en forma general para poder emplearlo en la determinación del algoritmo de solución.

DEFINICION 1 Al vector $P(t)=(p0,p1,p2,p3,p4,p5)$ se le denomina probabilidad de duración de cada coche y denota la probabilidad de que un coche dure i=0,1,2,3,4,5 años, de tal forma que P debe cumplir la propiedad:

$$\sum_{i=0}^{5} pi = 1 \quad ,$$

Pero P(t) se modifica de acuerdo a los resultados de las pruebas mecánicas de los autos por lo que

$$P(t/Sj)=Pj(t)$$

es la probabilidad de duración de cada coche después de j pruebas.

Como se mencionó en la sección anterior, la función de pérdida tiene una nueva concepción debido a que ahora ya no está ligada a el porcentaje de autos en buen estado solamente, ya que un auto en mal estado puede ser reparado; de tal forma que ahora está intimamente relacionada con la duración de cada auto (o con la esperanza de la pérdida/(utilidad) que cada auto proporciona debido a su duración) por tanto L es ahora un vector de pérdidas/(utilidades) con los valores siguientes:

$$L=(L0,L1,L2,L3,L4,L5)$$

que indican que se tiene una pérdida/(utilidad) Li si un coche dura más de i anos y menos de i+1.

Con ayuda de éstos conceptos que modifican las ideas expuestas hasta el momento, en la siguiente sección se presenta el algoritmo general de solución para problemas con caracteristicas similares a éste y en la sección 5.4 se desarrolla un ejemplo numérico para este caso particular.

Nótese antes de continuar la importancia del cambio en la interpretración de la función de pérdida en comparación con la utilizada en el capitulo IV, la cual era cuadrática y estaba en relación directa con el parámetro θ. Sin embargo esa misma función (la utlizada en la resolución del problema del capitulo anterior) se utilizará en el momento de las pruebas mecánicas por lo que los resultados obtenidos con anterioridad serán importantes para la resolución de este nuevo caso. Aunque para la formulación de la tabla del procedimiento óptimo será utilizada ésta nueva función de pérdida/(utilidad) (la definida en ésta sección) pues en ella se incorpora la duración de los automóviles.

5.3 DESARROLLO DE UNA SOLUCION

Se sabe que el señor X está interesado en el establecimiento del procedimiento que le permita optimizar tanto su decisión de compra ó rechazo del lote en cuestión así como la estrategia de muestreo por medio de la cual se minimice la inversión en pruebas mecánicas, dadas las restricciones que se han considerardo.

El estadístico que está realizando el análisis previo al muestreo, basado en su experiencia y en las estimaciones del señor X y del mecánico, concluye que la forma que la probabilid de tiempo de duración de los coches dada una observación (P(t/X1=x1)) puede quedar expresada por las ecuaciones

$$P(t/X1=1)=P1(t)=(p0-2e,p1-e,p2+2e,p3+.5e,p4+.5e,p5)$$

si el auto está en buen estado y (1)

$$P(t/X1=0)=P1(t)=(p0+3e,p1+e,p2-.5e,p3-1.5e,p4-e,p5-e)$$

si el auto está en mal estado,

donde e es una constante que está en función directa de los valores iniciales de P(t), y que cumpla con

$$P1(t/X1=x1) \geq 0 \quad \text{y} \quad \sum_{i=0}^{5} pi(t/x1)=1 \;.$$

(Nota: se podria definir a e como una función cualqiera, sin embargo esto no se hace así ya que en tal caso, en el momento de la optimización, se tendria que llevar a cabo un análisis funcional más complejo; de tal forma que en este caso e denota tan solo a una constante sujeta sólo al cumplimiento de las restricciones anteriores.)

Se sabe de la definicion 1 de la sección anterior que la función $P_j(t)$ no solamente es la función a posteriori después de j observaciones sino que también juega el papel de función a priori para la muestra $j+1$. Por tanto, si el número de muestras tomadas hasta un cierto momento es j de los cuales K son autos en buen estado se concluye que

$$P_j(t)=P(t/x_1,x_2,\ldots,x_j)=\begin{bmatrix} p0-5Ke+3je \\ p1-2Ke+je \\ p2+2.5Ke-.5je \\ p3+2Ke-1.5je \\ p4+1.5Ke-je \\ p5-je+Ke \end{bmatrix}$$

para e tal que cumpla con

$$P_j(t) \geq 0 \qquad y \qquad \sum_{i=0}^{5} p_{ji} = 1.$$

Por otro lado supóngase que la distribución a priori del estado de los coches en el momento de la prueba es

$\beta e(a,b)$ Beta con parámetros a y b;

por tanto para tal prueba, se conoce que la distribución a posteriori después de j observaciones (a priori para la observación $j+1$) está dada por la expresión

$$\pi_j(\theta)=\beta(a+K,b+j-K)$$

en donde K es el número de autos en buen estado, por lo que las probabilidades de X_{j+1} quedan

$$P(X_{j+1}=1/S_j)=\frac{K+1}{j+2}$$

y

$$P(X_{j+1}=0/S_j)=\frac{j+1-k}{j+2}\quad .$$

Como se puede observar, la metodologia en general no tuvo variaciones significativas, no así la concepción de la función de pérdida; de tal forma que la función de riesgo bayesiano para la determinación del momento de la parada óptima presenta modificaciónes.

Esto es, el riesgo bayesiano para la parada óptima en un momento determinado j está expresado por

$$U_j=jc+10*\sum_{i=1}^{j} L_i P_{ji}$$

en donde c es el costo por prueba mecánica de cada coche, el riesgo de seguir muestreando es

$$\begin{aligned}V_{j+1}=E[U_{j+1}]=&P(X_{j+1}=1/S_j)(U_{j+1}/X_{j+1}=1)\\&+P(X_{j+1}=0/S_j)(U_{j+1}/X_{j+1}=0)\end{aligned}$$

y el procedimiento óptimo queda de la siguiente forma

$$V_j=\min\{U_j,E[V_{j+1}]\}$$

para cualquier momento j; y para j=n

$$V_n=U_n.$$

De esta forma queda completamente determinado el procedimiento de decisión óptimo para el problema en cuestión, nótese que en términos generales la metodología presentada en éste capítulo no presenta grandes diferencias con la desarrollada en el capítulo anterior de no ser por la incorporación de probabilidades de duración por año.

En la sección siguiente se ejemplifica la técnica presentada por medio de una aplicación numérica para este mismo caso; sin embargo, es importante subrayar que el algoritmo aquí desarrollado puede ser empleado para la resolución de muchos otros problemas de características similares al planteado en la sección anterior.

5.4 EVALUACION NUMERICA

Para dejar claros los conceptos presentados, en esta sección se evalúa una aplicación numérica al problema del señor X.

Supóngase que el señor X y el estadístico determinan que los valores para la función de pérdida/(utilidad) quedan como sigue

$$
\begin{aligned}
L0 &= \$\ 1'900{,}000 \\
L1 &= \$\ 1'600{,}000 \\
L2 &= \$(\ \ 550{,}000) \\
L3 &= \$(1'350{,}000) \\
L4 &= \$(1'650{,}000) \\
L5 &= \$(1'925{,}000)
\end{aligned}
$$

(nótese que los valores negativos para L2, L3, L4 y L5 respresentan utilidades); los valores iniciales para las probabilidades anuales son

$$
\begin{aligned}
p0 &= 0.20 \\
p1 &= 0.22 \\
p2 &= 0.24 \\
p3 &= 0.13 \\
p4 &= 0.11 \\
p5 &= 0.10,
\end{aligned}
$$

la distribución a priori del parámetro θ (probabilidad de que un auto esté en buen estado al momento de la prueba) está dada por

$$\pi 0(\theta)=\beta(1,1) \quad ,$$

el costo por chequeo de cada automóvil es

$$c=\$\ 45{,}000.$$

y la constante

$$e=.006$$

Con esta información y con la ayuda de los programas 2, 3, 4 y 5 del Anexo 1 se determinan los resultados presentados en las tablas 5.1 a 5.12 y se encuentra la estratégia óptima para este caso particular.

En la última sección del presente capitulo se anotan las consideraciónes finales y los resultados a los que se llegó en esta aplicación numérica.

TABLA 5.1

RIESGO DE LA PARADA DESPUES DE J OBSERVACIONES (000)

AUTOS	OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B 0	505	1341	2176	3012	3847	4683	5518	6354	7189	8025	8860
U 1		25	861	1696	2532	3367	4203	5038	5874	6709	7545
E 2			-455	381	1216	2052	2887	3723	4558	5394	6229
N 3				-935	-99	736	1572	2407	3243	4078	4914
O 4					-1415	-579	256	1092	1927	2763	3598
S 5						-1895	-1059	-224	612	1447	2283
6							-2375	-1539	-704	132	967
7								-2855	-2019	-1184	-348
8									-3335	-2499	-1664
9										-3815	-2979
10											-4295

TABLA 5.2

RIESGO DE LA CONTINUACION DESPUES DE J OBSERVACIONES (000)

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	205	1738	2683	3584	4463	5330	6189	7043	7893	8740	8860
U	1		-931	936	2005	2929	3827	4709	5581	6446	7305	7545
E	2			-1826	176	1394	2323	3229	4120	4999	5870	6229
N	3				-2494	-518	772	1749	2658	3552	4435	4914
O	4					-2988	-1145	172	1196	2105	3000	3598
S	5						-3356	-1672	-379	657	1565	2283
	6							-3637	-2102	-875	130	967
	7								-3856	-2453	-1305	-348
	8									-4032	-2740	-1664
	9										-4175	-2979
	10											-4295

TABLA 5.3

RIESGO DEL PROCEDIMIENTO OPTIMO (000)

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	205	1341	2176	3012	3847	4683	5518	6354	7189	8025	8860
U	1		-931	861	1696	2532	3367	4203	5038	5874	6709	7545
E	2			-1826	176	1216	2052	2887	3723	4558	5394	6229
N	3				-2494	-518	736	1572	2407	3243	4078	4914
O	4					-2988	-1145	172	1092	1927	2763	3598
S	5						-3356	-1672	-379	612	1447	2283
	6							-3637	-2102	-875	130	967
	7								-3856	-2453	-1305	-348
	8									-4032	-2740	-1664
	9										-4175	-2979
	10											-4295

TÁBLA 5.4

PROCEDIMIENTO OPTIMO

AUTOS BUENOS \ OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
0	1	2	2	2	2	2	2	2	2	2	3
1		1	2	2	2	2	2	2	2	2	3
2			1	1	2	2	2	2	2	2	3
3				1	1	2	2	2	2	2	3
4					1	1	1	2	2	2	3
5						1	1	1	2	2	3
6							1	1	1	1	3
7								1	1	1	3
8									1	1	3
9										1	3
10											3

TABLA 5.5

$P(X_{j+1}=1/S_j)$ CONDICIONAL DE $X_{j+1}=1$ DADO S_j

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	0.500000	0.333333	0.250000	0.200000	0.166667	0.142857	0.125000	0.111111	0.100000	0.090909	0.083333
U	1		0.666667	0.500000	0.400000	0.333333	0.285714	0.250000	0.222222	0.200000	0.181818	0.166667
E	2			0.750000	0.600000	0.500000	0.428571	0.375000	0.333333	0.300000	0.272727	0.250000
N	3				0.800000	0.666667	0.571429	0.500000	0.444444	0.400000	0.363636	0.333333
O	4					0.833333	0.714286	0.625000	0.555556	0.500000	0.454545	0.416667
S	5						0.857143	0.750000	0.666667	0.600000	0.545455	0.500000
	6							0.875000	0.777778	0.700000	0.636364	0.583333
	7								0.888889	0.800000	0.727273	0.666667
	8									0.900000	0.818182	0.750000
	9										0.909091	0.833333
	10											0.916667

TABLA 5.6

$P(X_{j+1}=0/S_j)$ CONDICIONAL DE $X_{j+1}=0$ DADO S_j

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	0.500000	0.666667	0.750000	0.800000	0.833333	0.857143	0.875000	0.888889	0.900000	0.909091	0.916667
U	1		0.333333	0.500000	0.600000	0.666667	0.714286	0.750000	0.777778	0.800000	0.818182	0.833333
E	2			0.250000	0.400000	0.500000	0.571429	0.625000	0.666667	0.700000	0.727273	0.750000
N	3				0.200000	0.333333	0.428571	0.500000	0.555556	0.600000	0.636364	0.666667
O	4					0.166667	0.285714	0.375000	0.444444	0.500000	0.545455	0.583333
S	5						0.142857	0.250000	0.333333	0.400000	0.454545	0.500000
	6							0.125000	0.222222	0.300000	0.363636	0.416667
	7								0.111111	0.200000	0.272727	0.333333
	8									0.100000	0.181818	0.250000
	9										0.090909	0.166667
	10											0.083333

TABLA 5.7

PROBABILIDADES DE DURACION PARA EL PRIMER A*O

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	0.200000	0.218000	0.236000	0.254000	0.272000	0.290000	0.308000	0.326000	0.344000	0.362000	0.380000
U	1		0.188000	0.206000	0.224000	0.242000	0.260000	0.278000	0.296000	0.314000	0.332000	0.350000
E	2			0.176000	0.194000	0.212000	0.230000	0.248000	0.266000	0.284000	0.302000	0.320000
N	3				0.164000	0.182000	0.200000	0.218000	0.236000	0.254000	0.272000	0.290000
O	4					0.152000	0.170000	0.188000	0.206000	0.224000	0.242000	0.260000
S	5						0.140000	0.158000	0.176000	0.194000	0.212000	0.230000
	6							0.128000	0.146000	0.164000	0.182000	0.200000
	7								0.116000	0.134000	0.152000	0.170000
	8									0.104000	0.122000	0.140000
	9										0.092000	0.110000
	10											0.080000

TABLA 5.8

PROBABILIDADES DE DURACION PARA EL SEGUNDO AÑO

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	0.220000	0.226000	0.232000	0.238000	0.244000	0.250000	0.256000	0.262000	0.268000	0.274000	0.280000
U	1		0.214000	0.220000	0.226000	0.232000	0.238000	0.244000	0.250000	0.256000	0.262000	0.268000
E	2			0.208000	0.214000	0.220000	0.226000	0.232000	0.238000	0.244000	0.250000	0.256000
N	3				0.202000	0.208000	0.214000	0.220000	0.226000	0.232000	0.238000	0.244000
O	4					0.196000	0.202000	0.208000	0.214000	0.220000	0.226000	0.232000
S	5						0.190000	0.196000	0.202000	0.208000	0.214000	0.220000
	6							0.184000	0.190000	0.196000	0.202000	0.208000
	7								0.178000	0.184000	0.190000	0.196000
	8									0.172000	0.178000	0.184000
	9										0.166000	0.172000
	10											0.160000

TABLA 5.9

PROBABILIDADES DE DURACION PARA EL TERCER AÑO

AUTOS OBSERVADOS		0	1	2	3	4	5	6	7	8	9	10
B	0	0.240000	0.237000	0.234000	0.231000	0.228000	0.225000	0.222000	0.219000	0.216000	0.213000	0.210000
U	1		0.252000	0.249000	0.246000	0.243000	0.240000	0.237000	0.234000	0.231000	0.228000	0.225000
E	2			0.264000	0.261000	0.258000	0.255000	0.252000	0.249000	0.246000	0.243000	0.240000
N	3				0.276000	0.273000	0.270000	0.267000	0.264000	0.261000	0.258000	0.255000
O	4					0.288000	0.285000	0.282000	0.279000	0.276000	0.273000	0.270000
S	5						0.300000	0.297000	0.294000	0.291000	0.288000	0.285000
	6							0.312000	0.309000	0.306000	0.303000	0.300000
	7								0.324000	0.321000	0.318000	0.315000
	8									0.336000	0.333000	0.330000
	9										0.348000	0.345000
	10											0.360000

TABLA 5.10

PROBABILIDADES DE DURACION PARA EL CUARTO AÑO

AUTOS		OBSERVADOS 0	1	2	3	4	5	6	7	8	9	10
B	0	0.130000	0.121000	0.112000	0.103000	0.094000	0.085000	0.076000	0.067000	0.058000	0.049000	0.040000
U	1		0.133000	0.124000	0.115000	0.106000	0.097000	0.088000	0.079000	0.070000	0.061000	0.052000
E	2			0.136000	0.127000	0.118000	0.109000	0.100000	0.091000	0.082000	0.073000	0.064000
N	3				0.139000	0.130000	0.121000	0.112000	0.103000	0.094000	0.085000	0.076000
O	4					0.142000	0.133000	0.124000	0.115000	0.106000	0.097000	0.088000
S	5						0.145000	0.136000	0.127000	0.118000	0.109000	0.100000
	6							0.148000	0.139000	0.130000	0.121000	0.112000
	7								0.151000	0.142000	0.133000	0.124000
	8									0.154000	0.145000	0.136000
	9										0.157000	0.148000
	10											0.160000

TABLA 5.11

PROBABILIDADES DE DURACION PARA EL QUINTO AÑO

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	0.110000	0.104000	0.098000	0.092000	0.086000	0.080000	0.074000	0.068000	0.062000	0.056000	0.050000
U	1		0.113000	0.107000	0.101000	0.095000	0.089000	0.083000	0.077000	0.071000	0.065000	0.059000
E	2			0.116000	0.110000	0.104000	0.098000	0.092000	0.086000	0.080000	0.074000	0.068000
N	3				0.119000	0.113000	0.107000	0.101000	0.095000	0.089000	0.083000	0.077000
O	4					0.122000	0.116000	0.110000	0.104000	0.098000	0.092000	0.086000
S	5						0.125000	0.119000	0.113000	0.107000	0.101000	0.095000
	6							0.128000	0.122000	0.116000	0.110000	0.104000
	7								0.131000	0.125000	0.119000	0.113000
	8									0.134000	0.128000	0.122000
	9										0.137000	0.131000
	10											0.140000

TABLA 5.12

PROBABILIDADES DE DURACION PARA EL SEXTO A●O

AUTOS	OBSERVADOS	0	1	2	3	4	5	6	7	8	9	10
B	0	0.100000	0.094000	0.088000	0.082000	0.076000	0.070000	0.064000	0.058000	0.052000	0.046000	0.040000
U	1		0.100000	0.094000	0.088000	0.082000	0.076000	0.070000	0.064000	0.058000	0.052000	0.046000
E	2			0.100000	0.094000	0.088000	0.082000	0.076000	0.070000	0.064000	0.058000	0.052000
N	3				0.100000	0.094000	0.088000	0.082000	0.076000	0.070000	0.064000	0.058000
O	4					0.100000	0.094000	0.088000	0.082000	0.076000	0.070000	0.064000
S	5						0.100000	0.094000	0.088000	0.082000	0.076000	0.070000
	6							0.100000	0.094000	0.088000	0.082000	0.076000
	7								0.100000	0.094000	0.088000	0.082000
	8									0.100000	0.094000	0.088000
	9										0.100000	0.094000
	10											0.100000

5.5 CONCLUSIONES Y PRESENTACION DE RESULTADOS

Finalmente el estadístico presenta al señor X las tablas de la sección 5.4 junto con los comentarios siguientes y la tabla de resultados 5.13.

Como conclusión, el estadístico indica que para la interpretación de la tabla de resultados 5.13 se tienen los siguientes indicadores con sus respectivas interpretaciones

0 rechazar la oferta
1 continuar probando autos
2 comprar el lote.

Nótese que aun cuando se han probado por ejemplo, nueve coches y los nueve están en buen estado, la estrategia óptima indica que se siga muestreando; esto se debe a que como es muy bajo el costo por prueba en comparación a los valores de la tabla de pérdida/(utilidad) existen ciertos momentos del muestreo para los que es más importante contar con mayor información aun cuando se pague un poco más por ella.

También es importante notar que en el momento inicial el riesgo es favorable para las pruebas mecánicas de los coches y es por ello que el estadístico recomienda iniciar el estudio.

Ya con todo éste paquete de información, el señor X está en posibilidades de iniciar el muestreo y tomar la decisión que en cada caso sea la mas adecuada.

TABLA DE RESULTADOS 5.13

Autos Observados

		0	1	2	3	4	5	6	7	8	9	10
A	0	1	0	0	0	0	0	0	0	0	0	0
u	1		1	0	0	0	0	0	0	0	0	0
t	2			1	1	0	0	0	0	0	0	0
o	3				1	1	0	0	0	0	0	0
s	4					1	1	1	0	0	0	0
	5						1	1	1	0	0	0
b	6							1	1	1	1	2
u	7								1	1	1	2
e	8									1	1	2
n	9										1	2
o	10											2
s												

C A P I T U L O   V I

C O N C L U S I O N E S

6.1 CONCLUSIONES

Como se pudo notar, a través del desarrollo de toda la tesis, la metodología ha sido poco utilizada, y presenta una nueva perspectiva de solución para los problemas de la inferencia estadística y de las pruebas de hipótesis.

No obstante, como en todos los casos, el éxito de la aplicación práctica de las pruebas secuenciales bayesianas, y del problema de la parada óptima reside, en gran medida, en la interpretación que el estadístico haga de los resultados, así como en la correcta aplicación y utilización de la misma.

Además, en forma importante, para los casos en los que se exige una óptima aplicación de los recursos económicos para la resolución de un problema, esta teoría se presenta como altamente atractiva para el estadístico y para aquel que sea responsable del proyecto en cuestión.

Es importante aclarar que, durante toda la presentación del tema; y, más aún, en el momento de la elección de éste, la filosofía y el propósito principal fueron siempre animados por la idea de desarrollar una metodología nueva que se aplique para obtener soluciónes que optimicen, tanto los resultados como los recursos que se pudieran invertir en los proyectos a solucionar. Prueba palpable de ello fue la elección de la función de pérdida para el ejemplo práctico que se resolvió en el capítulo IV.

Sin embargo, esta técnica no es, en todos los casos, la mejor, ni la mas ventajosa en términos de resultados; pero sí representa una alternativa interesante para el estadístico. Alternativa que, en cada situación, debe de ser evaluada para una correcta aplicación de la misma, en su caso.

Por otro lado, la estadística matemática, e inclusive esta teoría, buscan dar un soporte adecuado a la toma de decisiones, la cual debe de jugar un papel más importante en estos momentos en los que, por la situación económica, se exige la utilización de las mejores técnicas y la aplicación óptima de los recursos para conseguir y determinar los mejores resultados, con base en los cuales se ofrezca una solución de los problemas en forma ventajosa y eficaz.

Así pues, la intención que se persigue con este trabajo, es la de brindar una nueva técnica para la problemática inferencial y de pruebas de hipótesis; además de presentar una breve introducción a la teoría de decisiones.

En la bibliografía anexa se enumeran los principales textos que fueron cimiento de la tésis, y cualquier profundización de los temas tratados; así como la formalización de aquéllos que sólo se tocaron superficialmente, puede ser referida a los libros que, en cada caso, se mencionan en la misma.

A N E X O 1

P R O G R A M A S D E C O M P U T O

PROGRAMA 1

GENERADOR DE NUMEROS ALEATORIOS (NORMAL)

```
    ∇ Z←DES NORMAL MED
[1]   Z←MED+DES×0.5×¯24++/0.01×48?100
    ∇
```

PROGRAMA 2

RIESGO DE LA PARADA DESPUES DE J OBSERVACIONES

```
     ∇ N RIESGOBAY CONST;W;J;SJ
[1]    N←N+1 ⋄ R←(N,N)⍴¯1 ⋄ W←0
[2]   L1:W←W+1 ⋄ →L2×⍳W>N ⋄ J←W-1 ⋄ SJ←0,⍳J
[3]    R[W;⍳W]←(J×C)+((SJ+1)×(J+1-SJ)×CONST)÷(J+2)×
       (J+2)×J+3 ⋄ →L1
[4]   L2:CONDICIONALX ⋄ ESPCOND
[5]    R←⍉R ⋄ V←⍉V ⋄ RES← 1 3 2 ⍉RES ⋄ PXJ← 1 3 2 ⍉
       PXJ ⋄ RES←RES×(2,N,N)⍴R≠(N,N)⍴0
     ∇
```

PROGRAMA 3

RIESGO DE LA CONTINUACION DESPUES DE J OBSERVACIONES

```
    ∇ ESPCOND;J
[1]   V2←(N,N)⍴¯2 ◇ J←N ◇ V2[J;]←R[J;]
[2]   RE← 2 11 11 ⍴0 ◇ RE[1;11;]←R[11;] ◇ J←11
[3]  L4:J←J+1 ◇ →L5×⍳J<1
[4]   RE[1;J;⍳J]←⌊/[1] R[J;⍳J],[0.5] V2[J;⍳J]←(RE[
      1;J+1;1+⍳J]×PXJ[2;J;⍳J])+RE[1;J+1;⍳J]×PXJ[1;
      J;⍳J] ◇ →L4
[5]  L5:RE[2;;]←((11 11 ⍴2)×RE[1;;]=R)+(11 11 ⍴1)×
      RE[1;;]=V2
[6]   V←V×(V≠(N,N)⍴¯2) ◇ R←R×(R≠(N,N)⍴¯1) ◇ V2←V2×
      (V2≠(N,N)⍴¯2) ◇ V←V2 ◇ RES←RE
    ∇
```

PROGRAMA 4

PROBABILIDADES CONDICIONALES DE Xj+1 DADO Sj

```
     ∇ CONDICIONALX;JN;J
[1]    J←0 ⋄ PXJ←(2,N,N)⍴0
[2]   L1:→0×⍳J>N−1×SJ←0,⍳J ⋄ JN←J+1
[3]    PXJ[1;JN;⍳JN]←(JN−SJ)÷(JN+1) ⋄ PXJ[2;JN;⍳JN]
       ←(SJ+1)÷(JN+1) ⋄ J←J+1 ⋄ →L1
     ∇
```

PROGRAMA 5

PROBABILIDADES DE DURACION ANUAL

```
    ∇ DUR←DURACION;I;J;E
[1]   DUR← 6 11 11 ⍴0 ⋄ E←0.006 ⋄ L←1000× 1700
      1600 ¯550 ¯1350 ¯1650 ¯1925
[2]   DUR[;1;1]← 0.2 0.22 0.24 0.13 0.11 0.1 ⋄ I←
      1
[3]  L1:J←0 ⋄ I←I+1 ⋄ →L3×⍳I>11
[4]  L2:J←J+1 ⋄ →L1×⍳J>I
[5]   DUR[;I;J]←DUR[;1;1]+((J-1)×E× ¯5 ¯2 2.5 2
      1.5 1)+((I-1)×E× 3 1 ¯0.5 ¯1.5 ¯1 ¯1 ⋄ →L2
[6]  L3:UTILIDAD←(⍉ 11 11 ⍴45000×0,⍳11)+10×+/[1]
      DUR× 2 3 1 ⍉ 11 11 6 ⍴L
[7]   I←0
[8]  L4:I←I+1 ⋄ →0×⍳I>11 ⋄ UTILIDAD[I;I+⍳11-I]←¯1
      ⋄ →L4
    ∇
```

ANEXO 2

LA FUNCION DE DISTRIBUCION DE PROBABILIDADES BETA

LA DISTRIBUCION DE PROBABILIDADES BETA

Una familia de densidades de probabilidades de variables aleatorias continuas, en el intervalo [0;1], pertenece a la familia de distribuciónes beta; cuya definición y principales caracteristicas se anotan a continuación.

DEFINICION 1 Si una variable aleatoria X tiene una densidad dada por:

$$\beta e(a,b)=f(X)=f(X;a,b)=\frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}X^{a-1}(1-X)^{b-1}I_{[0;1]}(X)$$

donde $a>0$ y $b>0$; entonces se dice que X tiene una distribución de tipo beta. En particular $\beta e(1,1)=U[0;1]$.

TEOREMA 1 Si X es una variable aleatoria con distribución $\beta e(a,b)$, entonces:

$E[X]=a/(a+b)$ y

$Var[X]=ab/(a+b+1)(a+b)^2$.

TEOREMA 2 Supóngase que $X_1,X_2,X_3,\ldots,X_n$; es una muestra aleatoria de ensayos de Bernoulli con parámetro desconocido θ. Supóngase, también, que la distribución a priori del parámetro θ está dada por $\beta e(a,b)$. Entonces, la distribución a posteriori de θ, dadas las n observaciones $X_i=x_i$ $(i=1,2,3,\ldots,n)$, es también beta con parámetros $a'=a+S_n$ y $b'=b+n-S_n$; en donde S_n es la suma de todas las variables $X_i=x_i$ para $i=1,2,3\ldots n$.

BIBLIOGRAFIA

1 Dynamic Programming
Bellman, R.
Princeton University Press
1975

2 Mathematical Statistics
Bickel, P.J. and Doksum, K.A.
Holden Day
1977

3 Bayesian Inference in Statistical Analysis
Box, G. and Tiao, G.
Addison Wesley
1973

4 Mathematical Methods of Statistics
Cramer, H.
Princeton University Press
1946

5 Optimal Statistical Decisions
DeGroot, M. H.
Mc. Graw-Hill
1974

6 Introduccion a la Teoria de Probabilidades y sus Aplicaciones, Vol. I
Feller, W.
Limusa
1973

7 Mathematical Statistics: A Decision Theoretic Approach
Ferguson, T.S.
Academic Press
1965

8 Introduction to Mathematical Statistics
Hoel, P.
Wiley & Sons, Inc.
1974

9 Testing Statistical Hipotheses
Lehmann, E.L.
Wiley & Sons, Inc.
1959

10 Probabilidad y Aplicaciones Estadisticas
Meyer, P.L.
Fondo Educativo Interamericano, S.A.
1973

11 Introduction to the Theory of Statistics
Mood, Graybill, Boes
Mc Graw-Hill
1974

12 Statistical Decision Theory, foundations, concepts and methods
O. Berger, J.
Springer Verlag
1972

13 Statistical Inference
Silvey, S. D.
Chapman and Hall
1975

14 Introducción a la Probabilidad Moderna
Szatzschneider, W.
(por editar)
1986