UNIVERSIDAD AUTÓNOMA METROPOLITANA UNIDAD -IZTAPALAPA
DEPARTAMENTO DE MATEMÁTICAS

Optimalidad de Pruebas Secuenciales para Dos Hipótesis Simples

Tesis que presenta:
Efrén Francisco Pérez

Para obtener el grado de:
Maestro en Ciencias (Matemáticas)

Asesor

Dr. Andrey Novikov

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA
México, julio de 2012

ÍNDICE GENERAL

INTRODUCCIÓN

El análisis secuencial forma parte de la teoría estadística desde hace casi 70 años. El resultado principal y más famoso del análisis secuencial es la optimalidad de la prueba secuencial de razón de probabilidades (sequential probability ratio test, SPRT), cuya primera demostración aparece en 1948 (ver A.Wald y J.Wolfowitz [17]). La descripción de la prueba y sus aplicaciones se pueden encontrar actualmente en muchos libros de texto (por ejemplo, en [9], [1], [6], [5]), monografías especializadas como [18], [7], [8], por mencionar algunas. Sin embargo, en muchas ocasiones el resultado formal carece de una demostración rigurosa, o no tiene ninguna. Algunos pocos sí la tienen, pero a veces sin detalles esenciales, como por ejemplo, en [6]. Desde los trabajos de Chow y Robbins (culminados en [3]) quedó claro que atrás de la optimalidad de la SPRT está la teoría general de paro ótimo ([3], [15]), sin embargo aún en estos dos últimos no se llega a ver todos los detalles de la optimalidad de la SPRT.

Lo anterior motivó el desarrollo de esta tesis, en la cual se ven todos los detalles del desarrollo de las pruebas óptimas, entre las cuales se encuentra la SPRT clásica.

Otro motivo fue el desarrollo de la teoría de pruebas óptimas para procesos estocásticos a tiempo discreto generales. Es que la teoría de paro óptimo clásica abarca los procesos generales, mientras la mayoría de los resultados sobre las pruebas secuenciales son acerca de observaciones independientes e identicamente distribuidas. Algunas excepciones son los trabajos [10] y [14]. En [10] se considera el caso de observaciones independientes pero no identicamente distribuidas, y no se usa la teoría general de paro óptimo (como la de [3]). Aunque en general el resultado parece estar bien (ver [13] al respecto), en la demostración de la optimalidad faltan algunos detalles esenciales. En [14] se considera la aplicación de la SPRT a los procesos de Markov a tiempo discreto (el uso de la SPRT a los procesos generales propuso Wald en su tiempo). La historia atrás es que existe un resultado previo de Schmitz (la referencia exacta se puede encontrar en [14]) en donde aparece un bosquejo de la forma de la prueba secuencial óptima en el caso de procesos de Markov, misma que, según los autores queda demasiado compleja para ser aplicable. Por fin, existe un resultado muy general [4], basado en la teoría de paro óptimo de Chow y Robbins [3], sin embargo es puro resultado de existencia, ya que sólo se demuestra que existe una prueba óptima, sin embargo su forma, generalmente hablando, queda desconocida. En este trabajo desarrollamos la teoría general de pruebas óptimas para dos hipótesis simples suponiendo que las observaciones están generadas por un proceso estocástico a tiempo discreto de forma muy general.

El tercer motivo para este trabajo es el hecho que, al parecer, hasta este momento no se conoce la forma de *todas* las pruebas óptimas, en el problema con restricciones (sobre las probabilidades de error tipo uno y dos). El hecho es que en [2] y en [11] se demuestra que no solo la SPRT, sino también las llamadas SPRT extendidas poseen la misma propiedad de optimalidad que la prueba SPRT original de Wald (ver [17]). Del trabajo de Lorden [11] quedó claro que este resultado se puede obtener usando el método de Lagrange aplicado al problema con restricciones. Sin embargo, parece que hasta este momento no existen resultados que digan que solo las SPRT extendidas (de [2]) son las óptimas en el problema con restricciones. En esta tesis, caracterizamos, en caso general de procesos estocásticos a tiempo discreto, la estructura de *todas* las pruebas óptimas en el problema con restricciones.

CAPÍTULO 1

MODELO MATEMÁTICO Y CONCEPTOS BÁSICOS

Nuestro objetivo en este capítulo es dar una descripción formal de lo que es un experimento estadístico secuencial, así como la definición de los elementos básicos que lo conforman, definición de pruebas de hipótesis y sus características numéricas.

1.1. Experimento estadístico secuencial

En cualquier experimento estadístico se recopilan datos (medidas, pesos, características importantes, etc.) el modelo más frecuente de un experimento estadístico, es una *muestra de una distribución de probabilidad*, lo que significa que se observan de manera independiente, cierto número n de datos, los cuales vienen de una misma distribución. Al resultado de estas observaciones se les llama muestra aleatoria de esta distribución o más específicamente, *muestra aleatoria de tamaño n.*

La idea del análisis secuencial, es que en lugar de tomar un número determinado de observaciones en una sóla toma, esta va formando la muestra *secuencialmente*, (es decir, tomando una observación a la vez), en función de las observaciones mismas terminando en un momento que sea adecuado o conveniente.

Ahora daremos el formalismo necesario al concepto del experimento estadístico secuencial.

Sea $X_1, X_2, \cdots, X_n, \ldots$ un proceso estocástico a tiempo discreto (con valores en un espacio medible $(\mathfrak{X}, \mathcal{X})$), cuya distribución depende de un parámetro desconocido θ. Primero trataremos el problema en general, suponiendo que existen las densidades conjuntas

$$f_\theta^n(x_1, x_2, \cdots, x_n) \tag{1.1}$$

con respecto a la medida producto $\mu^n = \mu \otimes \mu \otimes \cdots \otimes \mu$ sobre la σ-algebra producto $\mathcal{X}^n$. En el capítulo 3 se verá el caso particular, cuando las variables aleatorias $X_1, X_2, \cdots, X_n, \ldots$ son independientes e idénticamente distribuidas (copias independientes de una variable "genérica" X).

Uno de los problemas más fundamentales en estadística es el problema de pruebas de hipótesis. Una prueba de hipótesis es una afirmación o conjetura acerca de la distribución de una o más variables aleatorias. Aquí se tratará el caso más básico de las pruebas: el caso de *dos hipótesis simples.* A una hipótesis se le llama *simple* si la distribución hipotetizada es única, por consiguiente, debe especificar, no sólo la forma funcional de la distribución subyacente, sino también los valores de todos los parámetros.

Así que el problema, con base en los datos observados será, decidir cuál de las dos hipótesis, $H_0 : \theta = \theta_0$ ó $H_1 : \theta = \theta_1$ es la verdadera.

El objetivo de este esrito será caracterizar la estructura de la prueba secuencial óptima para este problema.

Para indicar cuál es el momento en que el experimento termina, es necesario aplicar algunas reglas a las observaciones que se están obteniendo de manera sucesiva. Y esto deberá llevarse a cabo en cada etapa del experimento, es decir, teniendo un número n dado de observaciones recopiladas $(X_1, X_2, \ldots, X_n)$, una regla nos tiene que indicar si este es el momento de terminar el experimento, o aún no, y entonces será necesario tomar algunas observaciones más.

Una regla de paro es una sucesión de funciones $\varphi = (\varphi_1(x_1), \varphi_2(x_1, x_2), \ldots)$. Donde la función $\varphi_n(x_1, \ldots, x_n)$ representa la probabilidad condicional de que el estadístico termine con la muestra dado que él ha obtenido n observaciones y $X_1 = x_1, \cdots, X_n = x_n$. Una manera alternativa de definir la regla de paro es por la sucesión de funciones $s^\varphi(x) = (s_1^\varphi(x_1), s_2^\varphi(x_1, x_2), \cdots)$ donde $s_n^\varphi(x_1, \ldots, x_n)$ representa la probabilidad condicional, dadas observaciones $x_1, \ldots, x_n$, de no parar durante las primeras $n-1$ observaciones y entonces parar cuando la $n-$ésima observación ha sido tomada. Estas dos funciones se relacionan mediante las fórmulas siguientes:

$$s_1^\varphi(x_1) = \varphi_1(x_1),$$

y

$$s_n^\varphi(x_1, \ldots, x_n) = (1 - \varphi_1(x_1))(1 - \varphi_2(x_1, x_2)) \ldots (1 - \varphi_{n-1}(x_1, \ldots, x_{n-1}))\varphi_n(x_1, \cdots, x_n),$$

para cada $n = 2, 3, \ldots$

Analogamente podemos representar la probabilidad condicional de continuar después de las primeras $n-1$ observaciones como $t^\varphi(x) = (t_1^\varphi, t_2^\varphi(x_1), \ldots)$, donde

$$t_1^\varphi = 1,$$

y

$$t_n^\varphi(x_1, \cdots, x_{n-1}) = (1 - \varphi_1(x_1)) \ldots (1 - \varphi_{n-1}(x_1, \ldots, x_{n-1})),$$

para cada $n = 2, 3, \ldots$.

Así, por la fórmula de probabilidad total, la probabilidad de parar al tiempo n está dada por la siguente expresión:

$$E_\theta s_n^\varphi(X_1, X_2, \cdots, X_n). \tag{1.2}$$

El experimento puede terminar en cualquier número de pasos $n = 1, 2, 3, \ldots$, así que tenemos que pensar en una manera de decidir a favor de H_0 ó H_1 cualquiera que sea el número terminal n de observaciones.

Una vez que el experimento se detiene, una decisión *aceptar o rechazar* H_0 es hecha, en este sentido se define a la siguiente función:

Una *regla de decisión (terminal)* es una sucesión de funciones $\delta = (\delta_1(x_1), \delta_2(x_1, x_2), \cdots)$, donde la función $\delta_n(x_1, x_2, \ldots, x_n)$ es interpretada como la probabilidad condicional de rechazar H_0 dado que el experimento termina en la etapa n y el vector $(x_1, x_2, \ldots, x_n)$ ha sido observado.

En resumen, una regla de decisión secuencial es un par (φ, δ) en el cual φ es la regla de paro y δ es una regla de decisión terminal. Supondremos que φ y δ son funciones medibles con valores en $[0, 1]$, para $n = 1, 2, \ldots$.

Las pruebas clásicas, basadas en una muestra aleatoria de tamaño fijo n es un caso particular de esta definición, más específicamente, el que corresponde a $\varphi_1(x_1) \equiv \varphi_2(x_1, x_2) \equiv \cdots \equiv \varphi_{n-1}(x_1, \cdots, x_{n-1}) \equiv 0$ y $\varphi_n(x_1, \cdots, x_n) = 1$ y por lo tanto pueden ser identificadas con tan sólo una regla de decisión δ_n misma que al terminar el experimento con datos $X_1, X_2, \ldots, X_n$, significa la aceptación de la hipótesis H_1 con probabilidad $\delta_n(x_1, x_2 \cdots, x_n)$.

1.2. Características de las pruebas

Una prueba proporciona un algoritmo de ir observando los datos y finalmente concluir, una vez terminado el experimento.

Los datos vienen de manera aleatoria por lo que resulta posible que el algoritmo se equivoque. Esto origina los errores que a continuación se detallan: Siendo verdadera la H_0, la decisión terminal puede ser $\ll$rechazar$H_0 \gg$. Por el contrario, cuando H_1 es válida, la decisión tomada puede ser $\ll$aceptar$H_0 \gg$, al primer caso se le conoce como error $\ll$tipo I$\gg$ y al segundo como error $\ll$tipo II$\gg$.

Si nos preocupa la calidad de la inferencia estadística tenemos que procurar que estos dos tipos de error no sean frecuentes. Como medida de la frecuencia de estos errores se usa la probabilidad, así que trabajando con las pruebas tenemos que tomar en cuenta las probabilidades de error, esto da lugar a la definición de la *probabilidad de error tipo I.*

$$\alpha(\varphi, \delta) = P_{\theta_0}(\text{rechazar} H_0) \tag{1.3}$$

y de la *probabilidad de error tipo II*

$$\beta(\varphi, \delta) = P_{\theta_1}(\text{aceptar} H_0). \tag{1.4}$$

Otra característica importante de una prueba es el número final de datos observados, ya que en la práctica nos interesa terminar lo más pronto posible, empleando menos observaciones, tiempo, recursos, etc.

La regla de paro φ, genera una nueva variable aleatoria τ_φ que es el número final de datos observados a la que se le llama *tiempo de paro.* Formalmente el tiempo de paro τ_φ puede ser definido, partiendo del sentido de aleatorización, de la siguiente manera: Sea $U_1, U_2, \ldots$ (proceso de aleatorización) una sucesión de variables aleatorias e idénticamente distribuidas que tienen distribución uniforme en el intervalo $[0, 1]$, y tal que el proceso $(U_1, U_2, \ldots)$ es independiente del proceso de observaciones $(X_1, X_2, \ldots)$. Entonces el tiempo de paro τ_φ es definido como

$$\tau_\varphi = \text{mín}\{n : U_n \leq \varphi_n(X_1, \ldots, X_n)\}$$

(por definición, $\tau_\varphi = \infty$ si $U_n > \varphi_n(X_1, \ldots, X_n)$ para cualquier n natural). Sea $\varphi_i = \varphi_i(X_1, \ldots, X_i)$ cuando φ_i se encuentra bajo el signo de probabilidad o de la esperanza matemática, y $\varphi_i =$

$\varphi_i(x_1, \ldots, x_i)$ en otro caso, $i = 1, 2, \ldots$. Es obvio, que

$$P_\theta(\tau_\varphi = n) = P_\theta(U_1 > \varphi_1, U_2 > \varphi_2, \ldots, U_{n-1} > \varphi_{n-1}, U_n \leq \varphi_n) = E_\theta s_n^\varphi$$

(ver 1.2), y

$$P_\theta(\tau_\varphi \geq n) = P_\theta(U_1 > \varphi_1, U_2 > \varphi_2, \ldots, U_{n-1} > \varphi_{n-1}) = E_\theta t_n^\varphi.$$

Debido a que el tiempo de paro τ_φ es una variable aleatoria, es natural tomar en cuenta su valor promedio como una característica de la duración del experimento (número promedio muestral). Pero el promedio puede ser calculado tanto bajo la hipótesis H_0, como bajo H_1, así que tenemos dos características

$$N(\theta_0; \varphi) = E_{\theta_0}\tau_\varphi \quad \text{y} \quad N(\theta_1; \varphi) = E_{\theta_1}\tau_\varphi, \tag{1.5}$$

mismas que tienen que ver con la regla de paro pero no con la regla de decisión.

Antes que nada vamos a expresar las características (1.3)-(1.5) en términos de los elementos del modelo y de la prueba.

De manera que, la probabilidad de *rechazar* H_0, suponiendo que el valor verdadero del parámetro es θ_0, se define como:

$$\alpha(\varphi, \delta) = P_{\theta_0}(\{\text{rechazar} \quad H_0\} \cap \{\tau_\varphi < \infty\}) = \sum_{n=1}^{\infty} E_{\theta_0} s_n^\varphi(X_1, \cdots, X_n)\delta_n(X_1, \cdots, X_n)$$

$$= \sum_{n=1}^{\infty} \int f_{\theta_0}^n(x_1, x_2, \ldots, x_n) s_n^\varphi(x_1, x_2, \ldots, x_n)\delta_n(x_1, x_2, \ldots, x_n) d\mu^n, \tag{1.6}$$

y por (1.4) se tiene que

$$\beta(\varphi, \delta) = P_{\theta_1}(\{\text{aceptar} \quad H_0\} \cap \{\tau_\varphi < \infty\}) = \sum_{n=1}^{\infty} E_{\theta_1} s_n^\varphi(X_1, \cdots, X_n)\left(1 - \delta_n(X_1, \cdots, X_n)\right)$$

$$= \sum_{n=1}^{\infty} \int f_{\theta_1}^n(x_1, x_2, \ldots, x_n) s_n^\varphi(x_1, x_2, \ldots, x_n)\left(1 - \delta_n(x_1, x_2, \ldots, x_n)\right) d\mu^n. \tag{1.7}$$

De manera muy parecida a (1.6) y (1.7),

$$N(\theta; \varphi) = E_\theta \tau_\varphi = \sum_{n=1}^{\infty} n E_\theta s_\varphi^n(X_1, \cdots, X_n)$$

$$= \sum_{n=1}^{\infty} \int n f_\theta^n(x_1, x_2, \ldots, x_n) s_n^\varphi(x_1, x_2, \ldots, x_n) d\mu^n. \tag{1.8}$$

Las integrales como en (1.6), (1.7) y (1.8) serán muy frecuentes a lo largo de este escrito, y aunque esta establecido que $\varphi_n = \varphi(X_1, X_2, \ldots, X_n)$ y $s_n^\varphi = s_n^\varphi(X_1, X_2, \ldots, X_n)$, en ocasiones como ésta resulta conveniente entender que la misma letra, digamos φ_n puede significar también $\varphi_n = \varphi_n(x_1, x_2, \ldots, x_n)$. Pero esto no debe causar problemas si se toma en cuenta el contexto.

También la aparición del vector $(x_1, x_2, \ldots, x_n)$ es muy recurrente, es por eso que es necesario cambiar de notación para hacer mas accesible el manejo de las ecuaciones, por lo tanto sin pérdida de generalidad se omitirá de las ecuaciones (1.6) y (1.8). Así estas se transforman en,

$$f_{\theta_0}^n(x_1, x_2, \ldots, x_n) = f_{\theta_0}^n, \tag{1.9}$$

$$\alpha(\varphi, \delta) = \sum_{n=1}^{\infty} \int f_{\theta_0}^n s_n^{\varphi} \delta_n d\mu^n, \tag{1.10}$$

$$\beta(\varphi, \delta) = \sum_{n=1}^{\infty} \int f_{\theta_1}^n s_n^{\varphi} (1 - \delta_n) \, d\mu^n \tag{1.11}$$

y

$$N(\theta; \varphi) = \sum_{n=1}^{\infty} \int n f_{\theta}^n s_n^{\varphi} d\mu^n, \tag{1.12}$$

respectivamente y con esto finalizamos el cálculo de las características de una prueba secuencial.

CAPÍTULO 2

OPTIMALIDAD DE PRUEBAS SECUENCIALES

En este capítulo se plantea el problema de optimalidad de las pruebas secuenciales como un problema de optimización con restricciones (Método variacional de Lagrange) y se va a proporcionar la solución completa del mismo.

2.1. Planteamiento del problema

El problema clásico del análisis secuencial es hallar la forma de la prueba que minimiza el número promedio de observaciones en la clase de todas las pruebas cuyas probabilidades de error no excedan ciertas cotas dadas, es decir se considera la clase $\Delta(\alpha, \beta)$ de las pruebas (φ, δ) que satisfacen las desigualdades siguientes:

$$\alpha(\varphi, \delta) \leq \alpha \text{ y } \beta(\varphi, \delta) \leq \beta, \tag{2.1}$$

siendo α y β números tales que $0 \leq \alpha < 1$ y $0 \leq \beta < 1$. El objetivo es hallar una prueba en $\Delta(\alpha, \beta)$ con el valor de $N(\theta_0, \varphi)$ (y/o $N(\theta_1, \varphi)$) mínimo en esta clase. De antemano no queda claro si existe una prueba que minimiza los dos valores promedio, $N(\theta_0, \varphi)$ y $N(\theta_1, \varphi)$ por eso se trata un problema menos ambicioso: encontrar una prueba con el valor mínimo de $N(\theta_0, \varphi)$ en $\Delta(\alpha, \beta)$. En realidad, la prueba que vamos a encontrar también va a minimizar $N(\theta_1, \varphi)$. Pero no se considera de mucha importancia, por lo menos no la teórica, por la siguiente razón: Como θ_0, θ_1 son intercambiables, al resolver el problema de minimización de $N(\theta_0, \varphi)$ tendremos también una prueba que minimiza $N(\theta_1, \varphi)$ en $\Delta(\alpha, \beta)$, como se demuestra al final de este escrito.

2.1.1. Reducción a un problema de optimización sin restricciones. Método variacional de Lagrange

Es conveniente representar el problema planteado en la sección anterior en términos del costo del experimento. En nuestro caso consideramos este costo como constante $c > 0$. Así el costo total promedio del experimento secuencial, gobernado por una regla de paro φ, será:

$$E_{\theta_0}(c\tau_\varphi) = cN(\theta_0;\varphi) = \sum_{n=1}^{\infty} \int cnf_{\theta_0}^n s_n^\varphi d\mu^n,$$

por la ecuación (1.12). Entonces el problema planteado en el segundo párrafo de este capítulo es equivalente a minimizar el costo promedio del experimento $cN(\theta_0;\varphi)$, dado que las probabilidades de error del procedimiento de prueba satisfacen (2.1). El método que se utiliza para la solución de problemas de optimización con restricciones se conoce como *el método de Lagrange.*

El método consiste en la reducción del problema con restricciones a un problema de optimización sin restricciones en una nueva función objetivo llamada función de Lagrange. En el caso que nos interesa, la función de Lagrange adquiere la forma,

$$R(\varphi,\delta) = cN(\theta_0;\varphi) + \pi_0\alpha(\varphi,\delta) + \pi_1\beta(\varphi,\delta), \tag{2.2}$$

donde $\pi_0 > 0$ y $\pi_1 > 0$ son dos constantes cualesquiera llamadas multiplicadores. Así el objetivo será encontrar (φ^*,δ^*) tales que

$$R(\varphi^*,\delta^*) = \inf_{(\varphi,\delta)} \{cN(\theta_0;\varphi) + \pi_0\alpha(\varphi,\delta) + \pi_1\beta(\varphi,\delta)\}. \tag{2.3}$$

El siguiente Teorema representa la reducción del problema con restricciones a un problema sin restricciones en la que propiamente consiste el método de Lagrange.

Teorema 2.1. *Sean $\pi_0 > 0$, $\pi_1 > 0$ constantes y (φ^*,δ^*) una prueba tal que para cualquier otra prueba secuencial (φ,δ) se cumple*

$$R(\varphi,\delta) \geq R(\varphi^*,\delta^*), \tag{2.4}$$

Entonces para cualquier prueba (φ,δ) que satisface:

$$\alpha(\varphi,\delta) \leq \alpha(\varphi^*\delta^*) \ \text{y} \ \beta(\varphi,\delta) \leq \beta(\varphi^*\delta^*), \tag{2.5}$$

se tiene,

$$N(\theta_0;\varphi) \geq N(\theta_0;\varphi^*). \tag{2.6}$$

La desigualdad en (2.6) es estricta, si alguna de las desigualdades en (2.5) lo es.

Observación 2.1. *Es obvio que (2.5) garantiza que $(\varphi^*,\delta^*) \in \Delta(\alpha,\beta)$, así que (2.6) representa la optimalidad de (φ^*,δ^*) en la clase $\Delta(\alpha,\beta)$, en el sentido que se describe al principio de esta sección.*

Demostración. Sea (φ^*,δ^*) que satisface (2.4). Entonces para cualquier otra prueba (φ,δ) que satisface (2.5) tenemos,

$$cN(\theta_0;\varphi) + \pi_0\alpha + \pi_1\beta \geq cN(\theta_0;\varphi) + \pi_0\alpha(\varphi,\delta) + \pi_1\beta(\varphi,\delta) \tag{2.7}$$

$$= R(\varphi,\delta) \geq R(\varphi^*,\delta^*) = cN(\theta_0;\varphi^*) + \pi_0\alpha(\varphi^*,\delta^*) + \pi_1\beta(\varphi^*,\delta^*) \tag{2.8}$$

$$= cN(\theta_0;\varphi^*) + \pi_0\alpha + \pi_1\beta. \tag{2.9}$$

Comparando el lado izquierdo de la cadena de desigualdades (2.7)-(2.8) con su lado derecho en (2.9) concluimos que:

$$N(\theta_0;\varphi) \geq N(\theta_0;\varphi^*),$$

lo que demuestra la primera parte del Teorema. Además es claro que si alguna de las desigualdades en (2.5) es estricta, entonces la desigualdad en (2.7) es estricta, es decir,

$$N(\theta_0;\varphi) > N(\theta_0;\varphi^*).$$

Y esto demuestra el Teorema. □

2.2. Reducción al problema de paro óptimo.

Del Teorema 2.1 vemos que para hallar la forma de la prueba óptima, necesitamos minimizar la función de Lagrange $R(\varphi,\delta)$. En esta sección este problema recibe una resolución parcial, resulta posible dar una regla de decisión δ^* óptima en el sentido de que para cualquier otra regla de decisión δ se cumple la desigualdad,

$$R(\varphi,\delta) \geq R(\varphi,\delta^*). \tag{2.10}$$

para cualquier φ. La demostración formal de esta aseveración se hace en el siguiente:

Teorema 2.2. *Para cada $\pi_0 \geq 0$, $\pi_1 \geq 0$ y cada regla de decisión δ,*

$$\pi_0\alpha(\varphi,\delta) + \pi_1\beta(\varphi,\delta) \geq \sum_{n=1}^{\infty} \int \min\left\{\pi_0 f_{\theta_0}^n, \pi_1 f_{\theta_1}^n\right\} s_n^{\varphi} d\mu^n. \tag{2.11}$$

La igualdad en (2.11) se alcanza si y sólo si,

$$I_{\{\pi_0 f_{\theta_0}^n < \pi_1 f_{\theta_1}^n\}} \leq \delta_n^* \leq I_{\{\pi_0 f_{\theta_0}^n \leq \pi_1 f_{\theta_1}^n\}}, \tag{2.12}$$

μ^n-casi dondequiera en $S_n^{\varphi} = \{(x_1,\ldots,x_n) : s_n^{\varphi}(x_1,\ldots,x_n) > 0\}$ para cada $n = 1, 2, \ldots$.

Demostración. Son necesarias ciertas explicaciones para dar comienzo a la demostración de este Teorema, pues su claridad depende de ciertos conceptos que necesariamente se deben destacar. La clave de esta demostración (y de algunas otras más adelante), es el siguiente sencillo:

Lema 2.1. *Sean ϕ, F_1, F_2 funciones medibles tales que*

$$0 \leq \phi(x) \leq 1, \;\; F_1(x) \geq 0, \;\; F_2(x) \geq 0, \tag{2.13}$$

y

$$\int \min\{F_1(x), F_2(x)\} d\mu(x) < \infty.$$

Entonces

$$\int (\phi(x)F_1(x) + (1-\phi(x))F_2(x))d\mu(x) \geq \int \text{mín}\{F_1(x), F_2(x)\}d\mu(x), \tag{2.14}$$

siendo esta una igualdad si y sólo si

$$I_{\{F_1(x)<F_2(x)\}} \leq \phi(x) \leq I_{\{F_1(x)\leq F_2(x)\}} \tag{2.15}$$

$\mu-$casi dondequiera.

Demostración. Comenzaremos demostrando la desigualdad (2.14). Para esto es suficiente mostrar lo siguiente:

$$\int [(\phi(x)F_1(x) + (1-\phi(x))F_2(x)) - \text{mín}\{F_1(x), F_2(x)\}]d\mu(x) \geq 0. \tag{2.16}$$

Pero por propiedades de la integral será suficiente mostrar que la función bajo el signo de integral es no negativa, es decir,

$$(\phi(x)F_1(x) + (1-\phi(x))F_2(x)) - \text{mín}\{F_1(x), F_2(x)\} \geq 0. \tag{2.17}$$

Para una x cualquiera veamos los dos casos posibles,

1. Si $F_1(x) > F_2(x)$ entonces la desigualdad (2.17) se transforma en $(\phi(x)(F_1(x)-F_2(x)) \geq 0$ y este valor obviamente es no negativo por como se definieron las funciones en (2.13).

2. Si $F_2(x) \geq F_1(x)$, entonces la desigualdad (2.17) se transforma en $(1-\phi(x))(F_2(x) - F_1(x)) \geq 0$ que obviamente es no negativo por (2.13) por lo tanto la desigualdad en (2.17) se cumple.

Así la desigualdad (2.14) queda demostrada.

Ahora se demuestra el caso de la igualdad en la ecuación (2.14).

Supongamos que en (2.14) se da la igualdad. Esto implica que la integral en (2.16) es nula, donde la función bajo el signo de la integral es no negativa. Entonces esta es nula $\mu-$ casi dondequiera, es decir

$$(\phi(x)F_1(x) + (1-\phi(x))F_2(x)) - \text{mín}\{F_1(x), F_2(x)\} = 0, \tag{2.18}$$

se cumple $\mu-$ casi dondequiera.

Definamos

$$A = \{x : F_1(x) > F_2(x)\},$$
$$B = \{x : F_1(x) < F_2(x)\},$$
$$C = \{x : F_1(x) = F_2(x)\},$$
$$D = \{x : \phi(x)F_1(x) + (1-\phi(x))F_2(x)) - \text{mín}\{F_1(x), F_2(x)\} = 0\},$$

entonces sabemos que $\mu(D^c) = 0$.

Sea $x \in A \cap D$, se tiene que (2.18) es equivalente a $\phi(x)(F_1(x) - F_2(x)) = 0$, lo que implica que $\phi(x) = 0$.

Análogamente si $x \in B \cap D$, entonces (2.18) es equivalente a $(1 - \phi(x))(F_1(x) - F_2(x)) = 0$ o sea $\phi(x) = 1$.

De esto se sigue que $\phi(x)$ cumple la ecuación (2.15) para cualquier $x \in D$. Efectivamente si $x \in A \cap D$ entonces vimos que $\phi(x) = 0$ y las dos indicadoras en (2.15) también son nulas; si $x \in B \cap D$ entonces $\phi(x) = 1$ y las dos indicadoras en (2.15) también valen 1; finalmente si $x \in C \cap D$ entonces $I_{\{F_1(x)<F_2(x)\}} = 0$, $I_{\{F_1(x)\leq F_2(x)\}} = 1$ y por lo tanto (2.15) se cumple trivialmente.

Dado que para cualquier $x \in D$ se cumple $x \in A \cap D$ ó $x \in B \cap D$ ó $x \in C \cap D$, tenemos que (2.15) se cumple para cualquier $x \in D$, es decir (2.15) se cumple $\mu-$casi dondequiera (recordemos que $\mu(D^c) = 0$). Esto demuestra que la igualdad en (2.14) se cumple si y sólo si $\phi(x)$ cumple (2.15) y por lo tanto el Lema queda demostrado. □

Regresando a la demostración del Teorema 2.2 se trata de demostrar la desigualdad (2.11), en principio podemos expresar el lado izquierdo de (2.11) utilizando las ecuaciones (1.10), (1.11) y (1.12) de la siguiente manera:

$$\pi_0\alpha(\varphi, \delta) + \pi_1\beta(\varphi, \delta) = \pi_0 \sum_{n=1}^{\infty} \int f_{\theta_0}^n s_n^\varphi \delta_n d\mu^n + \pi_1 \sum_{n=1}^{\infty} \int f_{\theta_1}^n s_n^\varphi (1 - \delta_n) d\mu^n$$

$$= \sum_{n=1}^{\infty} \int s_n^\varphi [\delta_n(\pi_1 f_{\theta_0}^n) + (1 - \delta_n)(\pi_1 f_{\theta_1}^n)] d\mu^n. \tag{2.19}$$

Ahora aplicando el Lema 2.1 a cada sumando de la función (2.19) con $\phi = \delta_n$, $F_1 = \pi_0 f_{\theta_0}^n$ y $F_2 = \pi_1 f_{\theta_1}^n$, es válida la siguiente desigualdad:

$$\sum_{n=1}^{\infty} \int s_n^\varphi [\delta_n(\pi_1 f_{\theta_0}^n) + (1 - \delta_n)(\pi_1 f_{\theta_1}^n)] d\mu^n \geq \sum_{n=1}^{\infty} \int s_n^\varphi \min\{\pi_0 f_{\theta_o}^n, \pi_1 f_{\theta_1}^n\} d\mu^n,$$

siendo esta una igualdad si y sólo si

$$I_{\{\pi_0 f_{\theta_0}^n < \pi_1 f_{\theta_1}^n\}} \leq \delta_n \leq I_{\{\pi_0 f_{\theta_0}^n \leq \pi_1 f_{\theta_1}^n\}},$$

$\mu-$casi dondequiera en S_n^φ para $n = 1, 2, 3 \ldots$. Por lo tanto el Teorema queda demostrado. □

Corolario 2.1. *Para el valor de $R(\varphi, \delta^*)$ se tiene la siguiente expresión:*

$$R(\varphi, \delta^*) = \sum_{n=1}^{\infty} \int s_n^\varphi \left[c n f_{\theta_0}^n + l_n\right] d\mu^n, \tag{2.20}$$

si $\sum_{n=1}^{\infty} E_{\theta_o} s_n^\varphi = 1$, donde

$$l_n = \min\left\{\pi_0 f_{\theta_0}^n, \pi_1 f_{\theta_1}^n\right\}. \tag{2.21}$$

En otro caso $R(\varphi, \delta^*) = \infty$.

Demostración. Por la ecuación (2.2) y (2.10) se tiene que

$$\begin{aligned} R(\varphi,\delta^*) &= cN(\theta_0,\varphi)+\sum_{n=1}^{\infty}\int s_n^{\varphi}\min\left\{\pi_0 f_{\theta_0}^n,\pi_1 f_{\theta_1}^n\right\}d\mu^n \\ &= \sum_{n=1}^{\infty}\int s_n^{\varphi} cn f_{\theta_0}^n d\mu^n+\sum_{n=1}^{\infty}\int s_n^{\varphi}\min\left\{\pi_0 f_{\theta_0}^n,\pi_1 f_{\theta_1}^n\right\}d\mu^n \\ &= \sum_{n=1}^{\infty}\int s_n^{\varphi}\left[cn f_{\theta_0}^n+\min\left\{\pi_0 f_{\theta_0}^n,\pi_1 f_{\theta_1}^n\right\}\right]d\mu^n, \end{aligned} \tag{2.22}$$

haciendo, $l_n=\min\left\{\pi_0 f_{\theta_0}^n,\pi_1 f_{\theta_1}^n\right\}$ en (2.22) se obtiene la ecuación (2.20) y esto demuestra el corolario. □

Según el Teorema 2.2 y el Corolario 2.1 $R(\varphi,\delta^*)$ es el mínimo de $R(\varphi,\delta)$ sobre *todas las reglas de decisión* δ. De esta manera el problema de minimización de $R(\varphi,\delta)$ se convierte en un problema de paro óptimo, ya que $R(\varphi,\delta^*)$ depende únicamente de la regla de paro y por lo tanto se puede denotar como

$$R(\varphi,\delta^*)=R(\varphi).$$

El objetivo, a partir de ahora será hallar la regla de paro φ^* tal que

$$R(\varphi)\geq R(\varphi^*),$$

para cualquier regla de paro φ. Y esto pondrá fin al problema de minimización ya que para cualquier prueba (φ,δ) por el Teorema 2.2 se cumple que

$$R(\varphi,\delta)\geq R(\varphi,\delta^*)=R(\varphi)\geq R(\varphi^*)=R(\varphi^*,\delta^*),$$

es decir, la prueba óptima será (φ^*,δ^*) por el Teorema 2.2, (con $\alpha(\varphi^*,\delta^*)=\alpha$ y $\beta(\varphi^*,\delta^*)=\beta$).

2.3. Paro óptimo. Caso truncado.

En esta sección vamos a resolver el problema de paro óptimo planteado en la sección anterior sobre la clase de tiempos de paro truncados.

Sea N un número natural tal que es el número máximo de pasos que puede realizar una prueba, es decir, sólamente se permiten las reglas de paro φ con $\varphi_N\equiv 1$ en lo demás siendo éstas arbitrarias. Entonces la única parte variable de las reglas de paro será $\varphi_1,\varphi_2,\ldots\varphi_{N-1}$, lo que hace posible la solución completa del problema. Empezaremos por darle a la función $R(\varphi)$ una forma más conveniente.

Se define para el problema truncado en N

$$R_N(\varphi)=\sum_{n=1}^{N}\int s_n^{\varphi}\left[cn f_{\theta_0}^n+l_n\right]d\mu^n. \tag{2.23}$$

Esta es la función de Lagrange definida en (2.2) que corresponde a la regla φ truncada al nivel N, es decir, la regla con las componentes $(\varphi_1,\varphi_2,\ldots\varphi_{N-1},1,\ldots)$ más explícitamente, se puede expresar (2.23) como

$$R_N(\varphi) = \sum_{n=1}^{N} \int (1-\varphi_1)\cdots(1-\varphi_{n-1})\varphi_n \left[cnf_{\theta_0}^n + \text{mín}\left\{\pi_0 f_{\theta_0}^n, \pi_1 f_{\theta_1}^n\right\}\right] d\mu^n.$$

y tomando en cuenta que $\varphi_N \equiv 1$ la forma final para $R_N(\varphi)$ será

$$R_N(\varphi) = \sum_{n=1}^{N-1} \int s_n^{\varphi} \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int (1-\varphi_1)\cdots(1-\varphi_{N-1}) \left[cNf_{\theta_0}^N + l_N\right] d\mu^N, \tag{2.24}$$

$$R_N(\varphi) = \sum_{n=1}^{N-1} \int s_n^{\varphi} \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_N^{\varphi} \left[cNf_{\theta_0}^N + l_N\right] d\mu^N.$$

El siguiente Lema absorbe la mayor parte de la carga técnica en el desarrollo de la regla truncada óptima.

Lema 2.2. *Sea $r \geq 2$ y $v_r = v_r(x_1, x_2, \ldots, x_r)$ una función medible arbitraria entonces*

$$\sum_{n=1}^{r-1} \int s_n^{\varphi} \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_r^{\varphi} \left[crf_{\theta_0}^r + v_r\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s_n^{\varphi} \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^{\varphi} \left[c(r-1)f_{\theta_0}^{r-1} + v_{r-1}\right] d\mu^{r-1}, \tag{2.25}$$

donde

$$v_{r-1} = \text{mín}\{l_{r-1}, cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\},$$

aqui $d\mu(x_r)$ significa que se esta integrando con respecto de la variable x_r.

La igualdad en (2.25) se alcanza si y sólo si

$$I_{\{l_{r-1} < cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}} < \varphi_{r-1} \leq I_{\{l_{r-1} \leq cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}}.$$

μ^{r-1}-casi dondequiera en $T_r^{\varphi} = \{(x_1, x_2, \ldots, x_r) : t_r^{\varphi}(x_1, x_2, \ldots, x_{r-1}) > 0\}$.

Demostración. Para demostrar la desigualdad en (2.25) es suficiente con mostrar que

$$\int s_{r-1}^{\varphi} \left[c(r-1)f_{\theta_0}^{r-1} + l_{r-1}\right] d\mu^{r-1} + \int t_r^{\varphi} \left[crf_{\theta_0}^r + v_r\right] d\mu^r \tag{2.26}$$

$$\geq \int t_{r-1}^{\varphi} \left[c(r-1)f_{\theta_0}^{r-1} + v_{r-1}\right] d\mu^{r-1},$$

porque los sumandos restantes $\left(\sum_{n=1}^{r-2} \int s_n^{\varphi} \left[cnf_{\theta_0}^n + l_n\right] d\mu^n\right)$ en la desigualdad (2.25) quedan intactos.

Aplicando el Teorema de Fubini a la función

$$\int t_r^{\varphi} \left[crf_{\theta_0}^r + v_r\right] d\mu^r, \tag{2.27}$$

recordemos que esta es una integral múltiple definida en la ecuacion (1.1), dado que la función $t_r^{\varphi}(x_1, \ldots, x_{r-1})$ no depende de x_r, se puede ver a (2.27) como sigue:

$$\int t_r^{\varphi} \left[\int \left(crf_{\theta_0}^r + v_r \right) d\mu(x_r) \right] d\mu^{r-1}. \tag{2.28}$$

Ahora es posible calcular la integral interior en (2.28), integrando con respecto de la variable x_r de la siguiente manera:

$$\int \left(crf_{\theta_0}^r + v_r \right) d\mu(x_r) = \int crf_{\theta_0}^r d\mu(x_r) + \int v_r d\mu(x_r),$$

así que por ser $f_{\theta_0}^r$ la densidad conjunta de $X_1, X_2, ...X_r$ la última expresión es igual a

$$crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r),$$

entonces se puede ver a la ecuación (2.26) como sigue:

$$\int s_{r-1}^{\varphi} \left[c(r-1) f_{\theta_0}^{r-1} + l_{r-1} \right] d\mu^{r-1} + \int t_{r-1}^{\varphi} (1 - \varphi_{r-1}) \left[crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r) \right] d\mu^{r-1}, \tag{2.29}$$

tomando en cuenta que $s_{r-1}^{\varphi} = t_{r-1}^{\varphi} \varphi_{r-1}$ la ecuación (2.29) se transforma en lo siguiente:

$$= \int t_{r-1}^{\varphi} \left[\varphi_{r-1}(c(r-1) f_{\theta_0}^{r-1} + l_{r-1}) + (1 - \varphi_{r-1})[crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)] \right] d\mu^{r-1} \tag{2.30}$$

Ahora aplicando el Lema 2.1 a la ecuación (2.30) con las funciones

$$\phi = \varphi_{r-1},$$

$$F_1 = c(r-1) f_{\theta_0}^{r-1} + l_{r-1}$$

$$F_2 = crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r),$$

es válida la siguiente desigualdad:

$$\begin{aligned} &\int t_{r-1}^{\varphi} \left[\varphi_{r-1}(c(r-1) f_{\theta_0}^{r-1} + l_{r-1}) + (1 - \varphi_{r-1})[crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)] \right] d\mu^{r-1} \\ &\geq \int t_{r-1}^{\varphi} \min \left\{ c(r-1) f_{\theta_0}^{r-1} + l_{r-1}, crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r) \right\} d\mu^{r-1}, \end{aligned} \tag{2.31}$$

siendo esta una igualdad si y sólo si

$$I_{\{c(r-1)f_{\theta_0}^{r-1} + l_{r-1} < crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}} \leq \varphi_{r-1} \leq I_{\{c(r-1)f_{\theta_0}^{r-1} + l_{r-1} \leq crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}},$$

o equivalentemente

$$I_{\{l_{r-1} < cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}} \leq \varphi_{r-1} \leq I_{\{l_{r-1} \leq cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\}},$$

por el mismo Lema 2.1. Ahora desarrollando el lado derecho de la desigualdad (2.31) tomando en cuenta que $s_{r-1}^{\varphi} = t_{r-1}^{\varphi}\varphi_{r-1}$ y $t_r^{\varphi} = t_{r-1}^{\varphi}(1-\varphi_{r-1})$ esta se transforma en lo siguiente:

$$\int s_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + l_{r-1}\right] d\mu^{r-1} + \int t_r^{\varphi}\left[crf_{\theta_0}^{r} + v_r\right] d\mu^r$$

$$\geq \int t_{r-1}^{\varphi} \min\left\{c(r-1)f_{\theta_0}^{r-1} + l_{r-1}, crf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\right\} d\mu^{r-1}, \qquad (2.32)$$

Por otra parte utilizando el hecho de que $\min\{a+b, a+d\} = a + \min\{b, d\}$ con $a = c(r-1)f_{\theta_0}^{r-1}$, $b = l_{r-1}$ y $d = cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)$ el lado derecho de la desigualdad (2.32) se puede ver como sigue:

$$\int t_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + \min\left\{l_{r-1}, cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\right\}\right] d\mu^{r-1}. \qquad (2.33)$$

Y finalmente la ecuación (2.33) se puede escribir como

$$= \int t_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + v_{r-1}\right] d\mu^{r-1}$$

donde

$$v_{r-1} = \min\left\{l_{r-1}, cf_{\theta_0}^{r-1} + \int v_r d\mu(x_r)\right\}.$$

Por lo tanto la desigualdad en (2.32) se puede ver como sigue:

$$\int s_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + l_{r-1}\right] d\mu^{r-1} + \int t_r^{\varphi}\left[crf_{\theta_0}^{r} + v_r\right] d\mu^r$$

$$\geq \int t_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + v_{r-1}\right] d\mu^{r-1}.$$

Y esta es igual a la desigualdad (2.26) por lo tanto el Lema esta demostrado. □

El Lema 2.2 se aplica inmediatamente a la función $R_N(\varphi)$ (ver (2.24) definiendo $v_N \equiv l_N$), y $V_N^N = v_N$. Por el Lema 2.2,

$$R_N(\varphi) \geq \sum_{n=1}^{N-2} \int s_n^{\varphi}\left[cnf_{\theta_0}^{n} + l_n\right] d\mu^n + \int t_{N-1}^{\varphi}\left[c(N-1)f_{\theta_0}^{N-1} + v_{N-1}\right] d\mu^{N-1}, \qquad (2.34)$$

donde ,$v_{N-1} = \min\{l_{N-1}, cf_{\theta_0}^{N-1} + \int V_N^N d\mu(x_N)\}$. Sea también $V_{N-1}^N = v_{N-1}$. Por el Lema 2.2 la desigualdad en (2.34) es en realidad una igualdad si y sólo si,

$$I_{\{l_{N-1} < cf_{\theta_0}^{N-1} + \int V_N^N d\mu(x_N)\}} \leq \varphi_{N-1} \leq I_{\{l_{N-1} \leq cf_{\theta_0}^{N-1} + \int V_N^N d\mu(x_N)\}}. \qquad (2.35)$$

$\mu^{N-1}-$casi dondequiera en T_{φ}^N.

Aplicando al lado derecho de (2.34) nuevamente el Lema 2.2 vemos que,

$$R_N(\varphi) \geq \sum_{n=1}^{N-3} \int s_n^{\varphi}\left[cnf_{\theta_0}^{n} + l_n\right] d\mu^n + \int t_{N-2}^{\varphi}\left[c(N-2)f_{\theta_0}^{N-2} + v_{N-2}\right] d\mu^{N-2}, \qquad (2.36)$$

donde,

$$v_{N-2} = \text{mín}\{l_{N-2}, cf_{\theta_0}^{N-2} + \int V_{N-1}^N d\mu(x_{N-1})\}.$$

Definamos también $V_{N-2}^N = v_{N-2}$. La igualdad en (2.36) se alcanza si se cumple (2.35) y

$$I_{\{l_{N-2} < cf_{\theta_0}^{N-2} + \int V_{N-1}^N d\mu(x_{N-1})\}} \leq \varphi_{N-2} \leq I_{\{l_{N-2} \leq cf_{\theta_0}^{N-2} + \int V_{N-1}^N d\mu(x_{N-1})\}}. \tag{2.37}$$

$\mu^{N-2}-$casi dondequiera en T_φ^{N-1}.

Si aplicamos nuevamente el Lema 2.2 al lado derecho de (2.36), se obtiene otra cota aun más baja, y una nueva φ_{N-3} para que ésta sea alcanzada, hasta obtener la cota de la forma,

$$R_N(\varphi) \geq \int \left[cf_{\theta_0} + V_1^N\right] d\mu(x_1) = c + \int V_1^N d\mu(x_1),$$

y unas condiciones sobre φ empezando por (2.35), (2.37), suficientes para alcanzarla, lo que significa que estas condiciones describen la regla de paro óptima.

El resultado formal se muestra en el siguiente:

Teorema 2.3. *Sea φ cualquier regla de paro truncada en N ($\varphi_N \equiv 1$). Entonces para cualquier $2 \leq r \leq N$ se cumplen las siguientes desigualdades:*

$$R_N(\varphi) \geq \sum_{n=1}^{r-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_r^\varphi \left[c(r)f_{\theta_0}^r + V_r^N\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^\varphi \left[c(r-1)f_{\theta_0}^{r-1} + V_{r-1}^N\right] d\mu^{r-1}, \tag{2.38}$$

donde $V_N^N = l_N$, y para $r < N$

$$V_r^N = \text{mín}\left\{l_r, Q_r^N + cf_{\theta_0}^r\right\},$$

y se define,

$$Q_r^N = \int V_{r+1}^N d\mu(x_{r+1}). \tag{2.39}$$

La cota inferior en (2.38) se alcanza si y sólo si,

$$I_{\{l_k < cf_{\theta_0}^k + Q_k^N\}} \leq \varphi_k \leq I_{\{l_k \leq cf_{\theta_0}^k + Q_k^N\}}, \tag{2.40}$$

μ^k-casi dondequiera en $T_r^\varphi = \{(x_1, x_2, \ldots, x_k) : t_k^\varphi(x_1, x_2, \ldots, x_{k-1}) > 0\}$, para cada $k = N-1, N-2, \ldots, r, r-1$. En particular para $k = 1$, las fórmulas en (2.40) describen de manera completa la regla de paro truncada óptima (dado que $\varphi_N \equiv 1$)

Demostración. Se demostrará la desigualdad (2.38) mediante un método de inducción hacia atrás aplicado a r.

Por definición

$$R_N(\varphi) = \sum_{n=1}^{N-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_N^\varphi \left[cNf_{\theta_0}^N + l_N\right] d\mu^N \tag{2.41}$$

Para $r = N$ aplicando el Lema 2.2 al lado derecho de la ecuación (2.41) con $v_N = l_N = V_N^N$ se cumple la siguiente desigualdad:

$$R_N(\varphi) = \sum_{n=1}^{N-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n \right] d\mu^n + \int t_N^\varphi \left[cNf_{\theta_0}^N + l_N \right] d\mu^N$$

$$\geq \sum_{n=1}^{N-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n \right] d\mu^n + \int t_{N-1}^\varphi \left[c(N-1)f_{\theta_0}^{N-1} + V_{N-1}^N \right] d\mu^{N-1}$$

con,

$$v_{N-1} = V_{N-1}^N = \text{mín} \left\{ l_{N-1}, \int V_N^N d\mu(x_N) + cf_{\theta_0}^{N-1} \right\} = \text{mín} \left\{ l_{N-1}, Q_{N-1}^N + cf_{\theta_0}^{N-1} \right\}.$$

Además, la igualdad se tiene si y sólo si,

$$I_{\{l_{N-1} < cf_{\theta_0}^{N-1} + Q_N^N\}} \leq \varphi_{N-1} \leq I_{\{l_{N-1} \leq cf_{\theta_0}^{N-1} + Q_{N-1}^N\}},$$

esto demuestra el Teorema para $r = N$.

Ahora supóngase que es válida la desigualdad (2.38) para $r = k$, esto es,

$$R_N(\varphi) \geq \sum_{n=1}^{k-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n \right] d\mu^n + \int t_k^\varphi \left[ckf_{\theta_0}^k + V_k^N \right] d\mu^k, \tag{2.42}$$

donde,

$$V_k^N = \text{mín} \left\{ l_k, cf_{\theta_0}^k + \int V_{k+1}^N d\mu(x_{k+1}) \right\} = \text{mín} \left\{ l_k, cf_{\theta_0}^k + Q_k^N \right\}$$

queda por demostrar la desigualdad (2.38) para $r = k - 1$.

Pero otra vez por el Lema 2.2 aplicado al lado derecho de la desigualdad en (2.42) con,

$$v_k = V_k^N,$$

se tiene,

$$\sum_{n=1}^{k-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n \right] d\mu^n + \int t_k^\varphi \left[ckf_{\theta_0}^k + V_k^N \right] d\mu^k$$

$$\geq \sum_{n=1}^{k-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n \right] d\mu^n + \int t_{k-1}^\varphi \left[c(k-1)f_{\theta_0}^{k-1} + V_{k-1}^N \right] d\mu^{k-1}$$

$$V_{k-1}^N = \text{mín} \left\{ l_{k-1}, cf_{\theta_0}^{k-1} + \int V_k^N d\mu(x_k) \right\} = \text{mín} \left\{ l_{k-1}, cf_{\theta_0}^{k-1} + Q_{k-1}^N \right\}.$$

Y la igualdad se alcanza si y sólo si se cumple,

$$I_{\{l_{k-1} < cf_{\theta_0}^{k-1} + Q_{k-1}^N\}} \leq \varphi_{k-1} \leq I_{\{l_{k-1} \leq cf_{\theta_0}^{k-1} + Q_{k-1}^N\}},$$

para cada $k - 1 < N$ Así la desigualdad (2.38) se cumple para cualquier $r < N$ y esto concluye la demostración. □

En particular, al final de este desarrollo se tiene:

$$R_N(\varphi) \geq \int \left[cf_{\theta_0} + V_1^N\right] d\mu = c + \int V_1^N d\mu(x_1) = c + Q_0^N.$$

De este desarrollo podemos concluir que $c + Q_0^N$ es una cota inferior para todas las reglas de paro truncadas, una vez que se alcanza esta cota, la regla respectiva será optima. El Teorema 2.3 describe todas las reglas truncadas que de esta manera seran optimas. El siguiente paso es analizar lo que ocurre con las reglas de paro no truncadas.

2.4. Paro óptimo. Caso no truncado.

En esta sección se le da la forma de la regla de paro óptima en la clase de todas las pruebas secuenciales.

La base para el desarrollo son los resultados de la sección anterior aplicados a, $R_N(\varphi)$, definida en (2.24).

La idea de lo que sigue es pasar al límite de estas desigualdades, cuando $N \to \infty$ para obtener las cotas inferiores para $R(\varphi)$.

Y la primer pregunta en este sentido es, ¿qué pasa con $R_N(\varphi)$ cuando $N \to \infty$?. La respuesta a esta pregunta, se da en el siguiente:

Teorema 2.4. *Para cualquier regla de paro* $\varphi = (\varphi_1, \varphi_2, \cdots,)$ *tal que* $P_{\theta_o}(\tau_\varphi < \infty) = 1$ *se cumple,*

$$\lim_{N\to\infty} R_N(\varphi) = R(\varphi). \tag{2.43}$$

Demostración. Para demostrar la igualdad (2.43) se tienen los dos casos siguientes:

1. $E_{\theta_0}\tau_\varphi < \infty$
2. $E_{\theta_0}\tau_\varphi = \infty$.

Primero se demuestra el caso en que,

$$E_{\theta_0}\tau_\varphi < \infty \tag{2.44}$$

Y lo que debemos demostrar es lo siguiente:

$$R(\varphi) - \lim_{N\to\infty} R_N(\varphi) = 0. \tag{2.45}$$

Por definición se tiene que,

$$\begin{aligned}
R(\varphi) - R_N(\varphi) &= \sum_{n=1}^{\infty} \int s_n^\varphi (ncf_{\theta_0}^n + l_n) d\mu^n - \sum_{n=1}^{N-1} \int s_n^\varphi (ncf_{\theta_0}^n + l_n)\psi_n d\mu^n \\
&- \int t_N^\varphi (Ncf_{\theta_0}^N + l_N) d\mu^N \\
&= \sum_{n=N}^{\infty} \int s_n^\varphi (ncf_{\theta_0}^n + l_n) d\mu^n - \int t_N^\varphi (Ncf_{\theta_0}^N + l_N) d\mu^N.
\end{aligned} \tag{2.46}$$

Como el primer sumando de (2.46) es la cola de una serie convergente, a saber la serie (2.44), entonces

$$\lim_{N\to\infty}\sum_{n=N}^{\infty}\int s_n^{\varphi}(ncf_{\theta_0}^n + l_n)d\mu^n = 0,$$

por lo que ahora sólo es necesario verificar que el segundo sumando en la ecuación (2.46) se va a cero.

Como sabemos, $l_n = \text{mín}\left\{\pi_0 f_{\theta_0}^n, \pi_1 f_{\theta_1}^n\right\}$, y entonces es válida la siguiente desigualdad:

$$\begin{aligned}\int t_N^{\varphi}[Ncf_{\theta_0}^N + l_N]d\mu^N &\leq \int t_N^{\varphi}[Ncf_{\theta_0}^N + \pi_0 f_{\theta_0}^N]d\mu^N \\ &= \pi_0\int t_N^{\varphi} f_{\theta_0}^N d\mu^N + cN\int t_N^{\varphi} f_{\theta_0}^N d\mu^N \\ &= \pi_0 E_{\theta_0} t_N^{\varphi} + cNE_{\theta_0} t_N^{\varphi}. \end{aligned} \tag{2.47}$$

Para demostrar que los sumandos en (2.47) se van a cero es necesario el siguiente:

Lema 2.3.

$$E_{\theta_0} t_N^{\varphi} = P_{\theta_0}(\tau_\varphi \geq N). \tag{2.48}$$

Demostración. Por definición se tiene,

$$P_{\theta_0}(\tau_\varphi \geq N) = 1 - P_{\theta_0}(\tau_\varphi < N) = 1 - \sum_{n=1}^{N-1} P_{\theta_0}(\tau_\varphi = n) = 1 - \sum_{n=1}^{N-1} E_{\theta_0} s_n^{\varphi}$$

$$= 1 - E_{\theta_0}[\varphi_1 + (1-\varphi_1)\varphi_2 + (1-\varphi_1)(1-\varphi_2)\varphi_3 + \ldots + (1-\varphi_1)(1-\varphi_2)\ldots(1-\varphi_{N-2})\varphi_{N-1}]$$
$$= E_{\theta_0}[1 - \varphi_1 - (1-\varphi_1)\varphi_2 - (1-\varphi_1)(1-\varphi_2)\varphi_3 - \ldots - (1-\varphi_1)(1-\varphi_2)\ldots(1-\varphi_{N-2})\varphi_{N-1}],$$

ahora factorizando, $(1-\varphi_1)$,

$$= E_{\theta_0}[(1-\varphi_1)[1 - \varphi_2 - (1-\varphi_2)\varphi_3 - \ldots - (1-\varphi_2)\ldots(1-\varphi_{N-2})\varphi_{N-1}]],$$

y luego factorizando $(1-\varphi_2)$ y así sucesivamente se llega a lo siguiente:

$$P_{\theta_0}(\tau_\varphi \geq N) = E_{\theta_0}[(1-\varphi_1)(1-\varphi_2)\ldots(1-\varphi_{N-1})] = E_{\theta_0} t_N^{\varphi}.$$

Y esta es la igualdad (2.48) que se quería probar. Por lo tanto el Lema queda demostrado. □

Regresando a la demostración del Teorema 2.4, aplicamos el Lema 2.3 a los sumandos de la ecuación (2.47), entonces,

$$\pi_0 E_{\theta_0} t_N^{\varphi} + cNE_{\theta_0} t_N^{\varphi} = \pi_0 P_{\theta_0}(\tau_\varphi \geq N) + cNP_{\theta_0}(\tau_\varphi \geq N). \tag{2.49}$$

De aquí inmediatamente se tiene que, $\text{lím}_{N\to\infty}\pi_0 P_{\theta_0}(\tau_\varphi \geq N) = 0$, por ser la cola de una serie convergente pues como sabemos $P_{\theta_0}(\tau_\varphi < \infty) = 1$. Además también,

$$\lim_{N\to\infty} cNP_{\theta_0}(\tau_\varphi > N) = 0, \tag{2.50}$$

porque,

$$E_{\theta_0}\tau_\varphi = \sum_{n=1}^{\infty} nP_{\theta_0}(\tau_\varphi = n) < \infty,$$

entonces,

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} nP_{\theta_0}(\tau_\varphi = n) = 0,$$

de aquí se sigue que:

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} nP_{\theta_0}(\tau_\varphi = n) \geq \lim_{N\to\infty} N \sum_{n=N}^{\infty} P_{\theta_0}(\tau_\varphi = n) = \lim_{N\to\infty} NP_{\theta_0}(\tau_\varphi \geq N) = 0,$$

y por lo tanto se cumple la igualdad en (2.50). Así, los sumandos en (2.49) tienden a cero, y esto demuestra la igualdad (2.45) con lo que queda demostrado el Teorema, para el primer caso.

Ahora para el segundo caso se tiene,

$$E_{\theta_0}\tau_\varphi = \sum_{n=1}^{N-1} \int s_n^\varphi (ncf_{\theta_0}^n + l_n) d\mu^n = \infty$$

la igualdad se cumple de manera trivial ya que,

$$\lim_{N\to\infty} R_N(\varphi) \geq \lim_{N\to\infty} \sum_{n=1}^{N-1} \int s_n^\varphi (ncf_{\theta_0}^n + l_n) d\mu^n = \infty,$$

por ser una serie divergente y por lo tanto la igualdad (2.43) se cumple en ambos casos y esto demuestra el Teorema.

□

Ahora nos enfocaremos en el comporamiento de las funciones V_r^N que aparecen en el Teorema 2.3 para tener una idea mas clara de lo que pasara cuando $N \to \infty$, porque dichas funciones participan de manera fundamental en el desarrollo de las desigualdades y la caracterización de la regla de paro óptima para el caso truncado.

Resulta que esta sucesión de funciones es no creciente con respecto a N para cada n fija, esta aseveración se demuestra en el siguiente:

Lema 2.4.

$$V_n^N \geq V_n^{N+1} \tag{2.51}$$

para cada $n = 1, 2, \ldots N$.

Demostración. La demostración de la desigualdad (2.51) se hace por inducción hacia atrás.

Por definición para $n = N$ se tiene,

$$V_N^N = l_N$$

y

$$V_N^{N+1} = \min \left\{ l_N,\ \ cf_{\theta_0}^N + Q_N^{N+1} \right\} \leq l_N = V_N^N$$

y entonces es claro que
$$V_N^N \geq V_N^{N+1},$$
es decir, se cumple la base de la inducción.

Supongase ahora que la desigualdad es válida para $n = r$, con $1 \leq r \leq N-1$ entonces,
$$V_r^N \geq V_r^{N+1},$$
esto implica que
$$Q_{r-1}^N \geq Q_{r-1}^{N+1}, \tag{2.52}$$
con $Q_r^N = \int V_{r+1}^N d\mu(x_{r+1})$ definido en (2.39). Entonces por la propiedad de monotonía en la integral queda por demostrar la desigualdad para $n = r-1$, es decir
$$V_{r-1}^N \geq V_{r-1}^{N+1}. \tag{2.53}$$

Pero por definición se tiene que
$$V_{r-1}^{N+1} = \text{mín}\left\{l_{r-1},\ \ cf_{\theta_0}^{r-1} + Q_{r-1}^{N+1}\right\}$$
y
$$V_{r-1}^{N+1} = \text{mín}\left\{l_{r-1},\ \ cf_{\theta_0}^{r-1} + Q_{r-1}^{N}\right\},$$
entonces por (2.52) se sigue la desigualdad (2.53), y por lo tanto el Lema queda demostrado. $\square$

Por el Lema 2.4 se sigue que para cualquier $n \geq 1$, la sucesión $V_n^N \geq 0$ es decreciente. Entonces su límite existe y le vamos a llamar
$$\lim_{N\to\infty} V_n^N = V_n = V_r = \text{mín}\left\{l_r, cf_{\theta_0}^r + \int V_{r+1} d\mu(x_r)\right\} \tag{2.54}$$
análogamente,
$$\lim_{N\to\infty} Q_r^N = Q_r = \int V_{r+1} d\mu(x_r). \tag{2.55}$$
La pregunta ahora es, ¿que pasa con el límite cuando $N \to \infty$, en todas las desigualdades e igualdades del Teorema 2.3, la respuesta a esta pregunta se da en el siguiente:

Lema 2.5. *Sea φ cualquier regla de paro entonces se cumplen las siguientes desigualdades:*
$$R(\varphi) \geq \sum_{n=1}^{r-1} \int s_n^{\varphi}\left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_r^{\varphi}\left[crf_{\theta_0}^r + V_r\right] d\mu^r$$
$$\geq \sum_{n=1}^{r-2} \int s_n^{\varphi}\left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^{\varphi}\left[c(r-1)f_{\theta_0}^{r-1} + V_{r-1}\right] d\mu^{r-1}$$
en donde,
$$V_r = \text{mín}\left\{l_r, cf_{\theta_0}^r + \int V_{r+1} d\mu(x_r)\right\} = \text{mín}\left\{l_r, cf_{\theta_0}^r + Q_r\right\}$$
para cualquier $r \geq 2$. En particular, para $r = 2$, se tiene la siguiente cota inferior:
$$R(\varphi) \geq c + Q_0.$$

Demostración. Del Teorema 2.3 se tiene que se cumplen las siguientes desigualdades:

$$R_N(\varphi) \geq \sum_{n=1}^{r-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_r^\varphi \left[crf_{\theta_0}^r + V_r^N\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^\varphi \left[c(r-1)f_{\theta_0}^{r-1} + V_{r-1}^N\right] d\mu^{r-1}, \tag{2.56}$$

ahora pasando al límite en las desigualdades (2.56), se tiene

$$\lim_{N\to\infty} R_N(\varphi) \geq \left\{\sum_{n=1}^{r-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n\right\} + \lim_{N\to\infty} \left\{\int t_r^\varphi \left[crf_{\theta_0}^r + V_r^N\right] d\mu^r\right\}$$

$$\geq \left\{\sum_{n=1}^{r-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n\right\} + \lim_{N\to\infty} \left\{\int t_{r-1}^\varphi \left[c(r-1)f_{\theta_0}^{r-1} + V_{r-1}^N\right] d\mu^{r-1}\right\}, \tag{2.57}$$

porque las unicas funciones que dependen de N son V_r^N *y* V_{r-1}^N. Ahora por el Teorema de Convergencia Monótona aplicado a las funciones $V_r^N,\ V_{r-1}^N$ y por el Teorema 2.4, $\lim_{N\to\infty} R_N(\varphi) = R(\varphi)$, las desigualdades en (2.57) se pueden reescribir como sigue:

$$R(\varphi) \geq \sum_{n=1}^{r-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^\varphi \left[crf_{\theta_0}^r + \lim_{N\to\infty} V_r^N\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^\varphi \left[c(r-1)f_{\theta_0}^{r-1} + \lim_{N\to\infty} V_{r-1}^N\right] d\mu^{r-1},$$

y finalmente,

$$R(\varphi) \geq \sum_{n=1}^{r-1} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_r^\varphi \left[crf_{\theta_0}^r + V_r\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s_n^\varphi \left[cnf_{\theta_0}^n + l_n\right] d\mu^n + \int t_{r-1}^\varphi \left[c(r-1)f_{\theta_0}^{r-1} + V_{r-1}\right] d\mu^{r-1},$$

por la ecuación (2.54). Por lo tanto, el Teorema queda demostrado. □

Lo que falta al Teorema 2.5, a comparación con el Teorema 2.4, es un elemento muy esencial y esta es la regla de paro para alcanzar la cota inferior, misma que así resulta ser óptima. Aunque pasando al límite de manera informal en las ecuaciones (2.40) (empezando desde la última de ellas), cuando $N \to \infty$, uno puede imaginar que la regla tiene que ser,

$$I_{\{l_n < cf_{\theta_0}^n + Q_n\}} \leq \varphi_n^* \leq I_{\{l_n \leq cf_{\theta_0}^n + Q_n\}}, \qquad n = 1, 2, 3, \ldots \tag{2.58}$$

El objetivo principal de esta sección es demostrar que la regla definida por (2.58) es realmente óptima, y la prueba de ello se puede ver en el siguiente:

Teorema 2.5. *Sea Q_r definida en (2.55) para cualquier regla de paro φ se tiene*

$$R(\varphi) \geq c + Q_0, \tag{2.59}$$

se alcanza la igualdad en (2.59) si y sólo si

$$I_{\{l_r < cf^r_{\theta_0} + Q_r\}} \leq \varphi_r \leq I_{\{l_r \leq cf^r_{\theta_0} + Q_r\}}, \tag{2.60}$$

μ^r-casi dondequiera en T^φ_r para cada $r = 1, 2, \ldots$

Demostración. Sea φ una regla de paro. Por el Lema 2.5 para cualquier $r \geq 1$ fijo se cumplen las desigualdades siguientes:

$$R(\varphi) \geq \sum_{n=1}^{r-1} \int s^\varphi_n \left[cnf^n_{\theta_0} + l_n\right] d\mu^n + \int t^\varphi_r \left[crf^r_{\theta_0} + V_r\right] d\mu^r$$

$$\geq \sum_{n=1}^{r-2} \int s^\varphi_n \left[cnf^n_{\theta_0} + l_n\right] d\mu^n + \int t^\varphi_{r-1} \left[c(r-1)f^{r-1}_{\theta_0} + V_{r-1}\right] d\mu^{r-1} \tag{2.61}$$

$$\geq \cdots$$

$$\geq \int \varphi_1 \left[cnf^1_{\theta_0} + l_1\right] d\mu(x_1) + \int (1-\varphi_1)\left[2cf^2_{\theta_0} + V_2\right] d\mu^2 \geq c + \int V_1 d\mu(x_1) = c + Q_0.$$

Esto demuestra la desigualdad (2.59).

Ahora sea φ^*_{r-1} como en (2.58) para cada $r = 2, 3, ...$ y un proceso truncado en r es decir ($\varphi_r \equiv 1$), aplicando el Lema 2.2 utilizando $v_r = V_r$ se tiene una igualdad para la segunda desigualdad en (2.61), con $I_{\{l_{r-1} < cf^{r-1}_{\theta_0} + Q_{r-1}\}} \leq \varphi^*_{r-1} \leq I_{\{l_{r-1} \leq cf^{r-1}_{\theta_0} + Q_{r-1}\}}$, $\mu^{r-1}-casi$ dondequiera en T^φ_r de manera análoga es fácil ver que todas las desigualdades exceptuando tal vez la primera van a ser igualdades para la regla

$$\varphi^* = (\varphi^*_1, \varphi^*_2, \ldots \varphi^*_{r-2}, \varphi^*_{r-1}, 1 \ldots).$$

Entonces se tiene

$$\sum_{n=1}^{r-1} \int s^{\varphi^*}_n \left[cnf^n_{\theta_0} + l_n\right] d\mu^n + \int t^{\varphi^*}_r \left[crf^r_{\theta_0} + V_r\right] d\mu^r = c + Q_0, \tag{2.62}$$

tomando en cuenta que los sumandos del lado izquierdo de la ecuación (2.62) son todos positivos se sigue

$$cr \int t^{\varphi^*}_r f^r_{\theta_0} d\mu^r = crP_{\theta_o}(\tau_{\varphi^*} \geq r) \leq c + Q_0,$$

entonces

$$P_{\theta_o}(\tau_{\varphi^*} \geq r) \leq \frac{c + Q_0}{cr}. \tag{2.63}$$

Pasando al límite cuando $r \to \infty$ en (2.63) se tiene

$$\lim_{r\to\infty} P_{\theta_o}(\tau_{\varphi^*} \geq r) \leq \lim_{r\to\infty} \frac{c + Q_0}{cr} = 0,$$

y por lo tanto $P_{\theta_0}(\tau_{\varphi^*} < \infty) = 1$.

Por otra parte teniendo igualdades en (2.61) significa que existe el límite entonces

$$\lim_{r\to\infty}\left\{\sum_{n=1}^{r-1}\int s_n^{\varphi}\left[cnf_{\theta_0}^n + l_n\right]d\mu^n + \int t_r^{\varphi}\left[c(r)f_{\theta_0}^r + V_r\right]d\mu^r\right\} = c + Q_0,$$

por lo tanto

$$\lim_{r\to\infty}\left\{\sum_{n=1}^{r-1}\int s_n^{\varphi}\left[cnf_{\theta_0}^n + l_n\right]d\mu^n\right\} \le c + Q_0, \tag{2.64}$$

pero el lado izquierdo de la desigualdad (2.64) es $R(\varphi)$, ya que $P_{\theta_0}(\tau_\varphi < \infty) = 1$ por lo tanto $R(\varphi) \le c + Q_0$.

Además en virtud de las desigualdades en (2.61) $R(\varphi) \ge c + Q_0$. Por lo tanto,

$$R(\varphi) = c + Q_0.$$

Esto demuestra la condición suficiente del Teorema.

Ahora se demuestra la condición necesaria.

Supongamos que $R(\varphi) = c + Q_0$. Por el Lema 2.2 se puede observar de las desigualdades en (2.61) que la igualdad en la última desigualdad se alcanza sólo si

$$I_{\{l_1 < cf_{\theta_0} + Q_1\}} \le \varphi_1 \le I_{\{l_1 \le cf_{\theta_0} + Q_1\}},$$

$\mu-$casi dondequiera en T_2^{φ}.

La igualdad en la penultima desigualdad se alcanza solo si

$$I_{\{l_2 < cf_{\theta_0}^2 + Q_2\}} \le \varphi_2 \le I_{\{l_2 \le cf_{\theta_0}^2 + Q_2\}}$$

μ^2-casi dondequiera en T_3^{φ}, y asi sucesivamente hasta llegar a la primer desigualdad, se alcanza igualdad sólo si

$$I_{\{l_{r-1} < cf_{\theta_0}^{r-1} + Q_{r-1}\}} \le \varphi_{r-1} \le I_{\{l_{r-1} \le cf_{\theta_0}^{r-1} + Q_{r-1}\}}$$

μ^{r-1}-casi dondequiera en T_r^{φ}. Es decir para cualquier r existe una φ_r tal que se alcanza la igualdad en (2.59) y esto prueba el Teorema. □

Finalmente el siguiente corolario resume los resultados mas importantes obtenidos en este capítulo.

Corolario 2.2. *Sean l_n y Q_n definidas en (2.21),(2.55), y sea (φ^*, δ^*) una prueba que satisface*

$$I_{\{l_n < cf_{\theta_0}^n + Q_n\}} \le \varphi_n^* \le I_{\{l_n \le cf_{\theta_0}^n + Q_n\}}, \tag{2.65}$$

μ^n-casi dondequiera en T_n^{φ}, $n = 1, 2, \ldots$

$$I_{\{\pi_0 f_{\theta_0}^n < \pi_1 f_{\theta_1}^n\}} \le \delta_n^* \le I_{\{\pi_0 f_{\theta_0}^n \le \pi_1 f_{\theta_1}^n\}}, \tag{2.66}$$

μ^n-casi dondequiera en S_n^{φ} para cada $n = 1, 2, \ldots$

Entonces para cualquier prueba que satisface

$$\alpha(\varphi,\delta) \leq \alpha(\varphi^*\delta^*) \;\; y \;\; \beta(\varphi,\delta) \leq \beta(\varphi^*\delta^*) \tag{2.67}$$

se cumple,

$$N(\theta_o;\varphi) \geq N(\theta_o;\varphi^*). \tag{2.68}$$

La desigualdad en (2.68) es estricta, si alguna de las desigualdades en (2.67) lo es. Si se tienen igualdades en (2.68) y (2.67) entonces para las componentes de φ se cumple (2.65) μ^n-casi dondequiera en T_n^φ, $n = 1, 2, \ldots$, y para las componentes de δ se cumple (2.66) μ^n-casi dondequiera en S_n^φ para cada $n = 1, 2, \ldots$.

Demostración. Para $\varphi^* = (\varphi_1^*, \varphi_2^*, \ldots \varphi_{r-2}^*, \varphi_{r-1}^*, \varphi_r^* \ldots)$ y $\delta^* = (\delta_1^*, \delta_2^*, \ldots \delta_{n-2}^*, \delta_{n-1}^*, \delta_n^* \ldots)$ como en las ecuaciones (2.65) y (2.66) se cumple la desigualdad

$$R(\varphi,\delta) \geq R(\varphi^*,\delta^*) = c + Q_0,$$

para cualquier otra prueba (φ,δ) por los Teoremas 2.2 y 2.5, con $R(\varphi^*,\delta^*)$ definida en (2.3).

Por lo tanto, se cumple la desigualdad $N(\theta_o;\varphi) \geq N(\theta_o;\varphi^*)$, por el Teorema 2.1. Además si se tienen igualdades en (2.68) y (2.67) entonces para las componentes de φ se cumple (2.65) μ^n-casi dondequiera en T_n^φ, $n = 1, 2, \ldots$, y para las componentes de δ se cumple (2.66) μ^n-casi dondequiera en S_n^φ para cada $n = 1, 2, \ldots$, por los Teoremas 2.2 y 2.5, debido a que en este caso se cumple la igualdad $R(\varphi,\delta) = R(\varphi^*,\delta^*) = c + Q_0$. □

CAPÍTULO 3

LA PRUEBA SECUENCIAL DE LA RAZÓN DE PROBABILIDADES

En esta sección se demuestra que la prueba (SPRT) Sequential Probability Ratio Test, es una de muchas que describe el Teorema 2.5.

Para esto es necesario darle a la prueba óptima una forma más específica, es decir poner dicha prueba en función de un estadístico más conocido, y se verá que el más adecuado es la llamada razón de probabilidades que es un estadístico muy conocido, y expresa la regla de paro óptima en forma más explícita.

Definición 3.1. Razón de probabilidades. *Se define como,*

$$Z_n = \prod_{r=1}^{n} \frac{f_{\theta_1}(X_r)}{f_{\theta_0}(X_r)} = \frac{\prod_{r=1}^{n} f_{\theta_1}(X_r)}{\prod_{r=1}^{n} f_{\theta_0}(X_r)} = \frac{f_{\theta_1}^n(X^n)}{f_{\theta_0}^n(X^n)}$$

esta es, la llamada razón de probabilidades, con $n = 1, 2, \ldots$.

Y entonces,

$$z_n = \begin{cases} \frac{f_{\theta_1}^n(x^n)}{f_{\theta_0}^n(x^n)} & \text{si } f_{\theta_0}^n(x^n) > 0 \\ \infty & \text{si } f_{\theta_0}^n(x^n) = 0 \ y \ f_{\theta_1}^n(x^n) > 0 \\ 0 & \text{si } f_{\theta_0}^n(x^n) = 0 \ y \ f_{\theta_1}^n(x^n) = 0. \end{cases} \tag{3.1}$$

Ahora para expresar los elementos de la regla de paro óptima en función de Z_n resulta conveniente definir, con base en las funciones V_n^N unas nuevas $\rho_n(Z_n)$ y $g(Z_n)$ como sigue:

Definición 3.2. *Se define*

$$g(z) = \text{mín}\left\{\pi_0, \pi_1 z\right\}, \tag{3.2}$$

$$\rho_n(z) = \text{mín}\left\{g(z), \ \ c + \int f_{\theta_0}(x)\rho_{n-1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x)\right\} \tag{3.3}$$

para $n = 1, 2, \ldots,$ *y*

$$\rho_0(z) = g(z).$$

Se observa en el siguiente Lema que la forma de definir estas funciones surge de manera natural, para expresar las funciones V_n^N en términos de Z_N, y con ello demostrar que una de las pruebas óptimas obtenidas en el capítulo anterior (Teorema 2.5) es equivalente a la SPRT.

Lema 3.1. *Todas las funciones V_n^N dependen de las observaciones $x_1, x_2, \dots x_n$ sólo a travéz de z_n como sigue:*

$$V_n^N(z_n) = f_{\theta_0}^n \rho_{N-n}(z_n), \tag{3.4}$$

para toda $n \leq N$. μ^n-casi dondequiera.

Demostración. La demostración se hace por inducción hacia atrás sobre la n.

Por definición se tiene para $n = N$,

$$V_N^N = l_N = \min\left\{\pi_0 f_{\theta_0}^N, \pi_1 f_{\theta_1}^N\right\} = f_{\theta_0}^N \min\left\{\pi_0, \pi_1 \frac{f_{\theta_1}^N}{f_{\theta_0}^N}\right\} = f_{\theta_0}^N \min\left\{\pi_0, \pi_1 z_N\right\},$$

y entonces,

$$V_N^N(z_N) = f_{\theta_0}^N \min\left\{\pi_0, \pi_1 z_N\right\} = f_{\theta_0}^N g(z_N) = f_{\theta_0}^N \rho_0(z_N),$$

por la ecuación (3.3), por lo tanto la igualdad es válida para $n = N$, es decir, se cumple la base de la inducción.

Ahora supongamos que (3.4) es válida para $n = r$, esto es,

$$V_r^N(z_r) = f_{\theta_0}^r \rho_{N-r}(z_r) \tag{3.5}$$

entonces queda por demostrar que la igualdad se cumple para $n = r-1$, es decir,

$$V_{r-1}^N(z_{r-1}) = f_{\theta_0}^{r-1} \rho_{N-r+1}(z_{r-1}) \tag{3.6}$$

$$= f_{\theta_0}^{r-1} \min\left\{g(z_{r-1}),\ c + \int f_{\theta_0}(x_r)\rho_{N-r}\left(z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)}\right) d\mu(x_r)\right\}.$$

Pero por definición se tiene,

$$V_{r-1}^N = \min\left\{l_{r-1},\ c f_{\theta_0}^{r-1} + \int V_r^N d\mu(x_r)\right\} \tag{3.7}$$

$$= \min\left\{\min\left\{\pi_0 f_{\theta_0}^{r-1}, \pi_1 f_{\theta_1}^{r-1}\right\},\ c f_{\theta_0}^{r-1} + \int V_r^N d\mu(x_r)\right\}, \tag{3.8}$$

además por la hipótesis (3.5), la ecuación (3.8) se puede escribir como sigue:

$$= \min\left\{\min\left\{\pi_0 f_{\theta_0}^{r-1}, \pi_1 f_{\theta_1}^{r-1}\right\},\ c f_{\theta_0}^{r-1} + \int f_{\theta_0}^r \rho_{N-r}(z_r) d\mu(x_r)\right\}. \tag{3.9}$$

Ahora por las ecuacion (3.1), las ecuaciones $f_{\theta_0}^r = f_{\theta_0}^{r-1} f_{\theta_0}(x_r)$ y $z_r = z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)}$, entonces la ecuación (3.9) se transforma en,

$$= \min\left\{\min\left\{\pi_0 f_{\theta_0}^{r-1}, \pi_1 f_{\theta_1}^{r-1}\right\},\ c f_{\theta_0}^{r-1} + \int f_{\theta_0}^{r-1} f_{\theta_0}(x_r)\rho_{N-r}\left(z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)}\right) d\mu(x_r)\right\}. \tag{3.10}$$

Factorizando $f_{\theta_0}^{r-1}$ en la ecuación (3.10), resulta,

$$= \text{mín}\left\{ f_{\theta_0}^{r-1} \text{mín}\left\{ \pi_0, \pi_1 \frac{f_{\theta_1}^{r-1}}{f_{\theta_0}^{r-1}} \right\}, f_{\theta_0}^{r-1} \left[c + \int f_{\theta_0}(x_r)\rho_{N-r}\left(z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)} \right) d\mu(x_r) \right] \right\}$$

$$= f_{\theta_0}^{r-1} \text{mín}\left\{ \text{mín}\left\{ \pi_0, \pi_1 z_{r-1} \right\},\ c + \int f_{\theta_0}(x_r)\rho_{N-r}\left(z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)} \right) d\mu(x_r) \right\}$$

$$= f_{\theta_0}^{r-1} \text{mín}\left\{ g(z_{r-1}), c + \int f_{\theta_0}(x_r)\rho_{N-r}\left(z_{r-1}\frac{f_{\theta_1}(x_r)}{f_{\theta_0}(x_r)} \right) d\mu(x_r) \right\}$$

$$= f_{\theta_0}^{r-1}\rho_{N-r+1}(z_{r-1}), \tag{3.11}$$

comparando la cadena de igualdades (3.7)-(3.11), se obtiene (3.6), que faltaba por demostrar, por lo tanto, el Lema esta demostrado. □

Para evitar un poco la carga en la notación que hasta el momento se tiene es de gran ayuda definir,

$$h_n(z) = \int f_{\theta_0}(x_{n+1})\rho_n\left(z\frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})} \right) d\mu(x_{n+1}), \tag{3.12}$$

y entonces,

$$\rho_{n+1}(z) = \text{mín}\left\{ g(z),\ \ c + h_n(z) \right\}.$$

Ahora se demuestra un Lema sobre el comportamiento de las $\rho_n(z)$ definidas en (3.3) que nos servirá para encontrar el límite de dichas funciones.

Lema 3.2.

$$\rho_n(z) \geq \rho_{n+1}(z), \tag{3.13}$$

para toda n y para toda $z \geq 0$.

Demostración. La demostración se hace por inducción sobre la n.

Para $n = 0$ por definición se tiene,

$$\rho_0(z) = g(z)$$

y

$$\rho_1(z) = \text{mín}\left\{ g(z),\ \ c + h_0(z) \right\} \leq g(z) = \rho_0(z)$$

y por lo tanto $\rho_0(z) \geq \rho_1(z)$, es decir, se cumple la base de la inducción.

Ahora supongamos que la desigualdad (3.13) es válida para $n = r$, es decir,

$$\rho_r(z) \geq \rho_{r+1}(z), \tag{3.14}$$

entonces queda por demostrar,

$$\rho_{r+1}(z) \geq \rho_{r+2}(z), \tag{3.15}$$

pero por definición:

$$\rho_{r+1}(z) = \text{mín}\left\{ g(z),\ \ c + h_r(z) \right\}$$

$$\rho_{r+2}(z) = \text{mín}\left\{g(z),\ \ c + h_{r+1}(z)\right\},$$

por lo tanto, es suficiente con mostrar que:

$$h_r(z) \geq h_{r+1}(z), \tag{3.16}$$

porque el resto de los sumandos que involucra la desigualdad (3.15) son los mismos en ambos lados de la desigualdad.

Pero esto es inmediato ya que por definición la desigualdad (3.16), implica

$$\int f_{\theta_0}(x)\rho_r\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x) \geq \int f_{\theta_0}(x)\rho_{r+1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x), \tag{3.17}$$

y la desigualdad (3.17) se cumple por la hipótesis (3.14) y por la propiedad de monotonía de la integral, entonces la desigualdad (3.15) es válida y por lo tanto el Lema queda demostrado. □

El Lema 3.2 implica que para todo $z \geq 0$ la sucesión $\rho_n(z)$ $n = 0, 1, 2, 3, ...$ es decreciente, y por lo tanto, existe el límite y lo denotamos de la siguiente manera:

$$\lim_{n\to\infty} \rho_n(z) = \rho(z). \tag{3.18}$$

Pasando al límite en (3.3) cuando $n \to \infty$ se tiene,

$$\lim_{n\to\infty} \rho_n(z) = \text{mín}\left\{g(z),\ \ c + \lim_{n\to\infty}\int f_{\theta_0}(x)\rho_{n-1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x)\right\}$$

a y aplicando el Teorema de Convergencia Dominada de Lebesgue a la función, $\rho_{n-1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)$ se tiene que para la función $\rho(z)$ se cumple la siguiente igualdad:

$$\rho(z) = \lim_{n\to\infty} \rho_n(z) = \text{mín}\left\{g(z),\ \ c + \int f_{\theta_0}(x)\lim_{n\to\infty} \rho_{n-1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x)\right\}$$

$$\rho(z) = \text{mín}\left\{g(z),\ \ c + \int f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x)\right\}. \tag{3.19}$$

Finalmente

$$\rho(z) = \text{mín}\left\{g(z),\ \ c + h(z)\right\},$$

donde,

$$h(z) = \int f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x). \tag{3.20}$$

Por otra parte, podemos representar la cota inferior Q_0 definida en (2.55) en terminos de las nuevas funciones $g(z)$ y $\rho(z)$ de la siguiente manera:

$$Q_0 = \int f_{\theta_0}\rho(\frac{f_{\theta_1}(x_1)}{f_{\theta_0}(x_1)})d\mu(x_1) \tag{3.21}$$

A la ecuación (3.19) se le conoce como *ecuación de Bellman.*

Ahora es posible expresar la regla de decisión optima del Teorema 2.2 y la regla de paro óptima del Teorema 2.5 en términos de la razón de probabilidades z_n, expresando los elementos de las formulas (2.12) y (2.60) en términos de z_n, esta idea se puede formalizar en el siguiente:

Teorema 3.1. *Sea Q_0 definido en (3.21), $\pi_0 > 0$, $\pi_1 > 0$ y las funciones $g(z)$, $h(z)$ definidas en (3.2) y (3.20) respectivamente, entonces para cualquier prueba secuencial (φ, δ) se cumple,*

$$R(\varphi, \delta) = c + Q_0, \tag{3.22}$$

si y sólo si,

$$I_{\{g(z_n)<c+h(z_n)\}} \leq \varphi_n \leq I_{\{g(z_n)\leq c+h(z_n)\}}, \tag{3.23}$$

μ^n-casi dondequiera sobre $T_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots$ y

$$I_{\{z_n>\frac{\pi_0}{\pi_1}\}} \leq \delta_n \leq I_{\{z_n\geq\frac{\pi_0}{\pi_1}\}} \tag{3.24}$$

μ^n-casi dondequiera en $S_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots$

Demostración. Por los Teoremas 2.5 y 2.2, sabemos que la igualdad (3.22) se cumple para

$$I_{\{l_n<cf_{\theta_0}^n+Q_n\}} \leq \varphi_n \leq I_{\{l_n\leq cf_{\theta_0}^n+Q_n\}},$$

μ^n-casi dondequiera en T_n^{φ} para cada $n = 1, 2, \ldots$ con Q_n definida en (2.55) y

$$I_{\{\pi_0 f_{\theta_0}^n<\pi_1 f_{\theta_1}^n\}} \leq \delta_n \leq I_{\{\pi_0 f_{\theta_0}^n\leq\pi_1 f_{\theta_1}^n\}},$$

μ^n-casi dondequiera en S_n^{φ} para cada $n = 1, 2, \ldots$.

Entonces lo primero que se debe demostrar es que las desigualdades siguientes:

$$g(z_n) < c + h(z_n) \quad y \quad l_n < cf_{\theta_0}^n + \int V_{n+1} d\mu(x_{n+1}), \tag{3.25}$$

son equivalentes.

Por definición de l_n, y por el Lema 3.1 se tiene lo siguiente:

$$l_n = V_n^n = f_{\theta_0}^n g(z_n). \tag{3.26}$$

Por otra parte, por la ecuación (2.54) sabemos que

$$V_{n+1} = \lim_{N\to\infty} V_{n+1}^N = \lim_{N\to\infty} f_{\theta_0}^{n+1} \rho_{N-n+1}(z_{n+1}), \tag{3.27}$$

por el Lema 3.1, entonces sustituyendo (3.27) en la ecuación (3.25) resulta,

$$cf_{\theta_0}^n + \int V_{n+1} d\mu(x_{n+1}) = cf_{\theta_0}^n + \int \lim_{N\to\infty} f_{\theta_0}^{n+1} \rho_{N-n+1}(z_{n+1}) d\mu(x_{n+1}), \tag{3.28}$$

además por las ecuaciones (**??**) y (3.1) la ecuación (3.28), se puede reescribir como sigue:

$$= cf_{\theta_0}^n + \int f_{\theta_0}^n f_{\theta_0}(x_{n+1}) \lim_{N\to\infty} \rho_{N-n+1}(z_n \frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})}) d\mu(x_{n+1}),$$

y por la ecuación (3.18).

$$= cf_{\theta_0}^n + \int f_{\theta_0}^n f_{\theta_0}(x_{n+1}) \rho(z_n \frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})}) d\mu(x_{n+1}). \tag{3.29}$$

Ahora factorizando $f_{\theta_0}^n$ en (3.29), se tiene

$$= f_{\theta_0}^n \left[c + \int f_{\theta_0}(x_{n+1}) \rho(z_n \frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})}) d\mu(x_{n+1}) \right], \tag{3.30}$$

porque la integral solo depende de x_{n+1}. Así por las ecuaciones (3.26) y (3.30) la desigualdad del lado derecho de (3.25) se puede expresar como sigue:

$$f_{\theta_0}^n g(z_n) < f_{\theta_0}^n \left[c + \int f_{\theta_0}(x_{n+1}) \rho(z_n \frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})}) d\mu(x_{n+1}) \right],$$

o lo que es lo mismo,

$$g(z_n) < c + \int f_{\theta_0}(x_{n+1}) \rho(z_n \frac{f_{\theta_1}(x_{n+1})}{f_{\theta_0}(x_{n+1})}) d\mu(x_{n+1}) = c + h(z_n).$$

Esta última desigualdad es equivalente al lado izquierdo de la ecuación (3.25), por lo tanto se ha demostrado que estas dos desigualdades son equivalentes, y esto demuestra la primera parte del Teorema puesto que la desigualdad del lado derecho de (3.23), es la misma que (3.25), sólo cambia el símbolo de la desigualdad $<$, por $\leq$. Por ota parte, para las componentes de δ se debe probar lo siguiente:

$$I_{\{z_n > \frac{\pi_0}{\pi_1}\}} \leq \delta_n \leq I_{\{z_n \geq \frac{\pi_0}{\pi_1}\}},$$

pero por el Teorema 2.2,

$$I_{\{\pi_0 f_{\theta_0}^n < \pi_1 f_{\theta_1}^n\}} \leq \delta_n \leq I_{\{\pi_0 f_{\theta_0}^n \leq \pi_1 f_{\theta_1}^n\}},$$

ahora, desarrollando el lado izquierdo de la desigualdad anterior se tiene,

$$I_{\{f_{\theta_1}^n \pi_1 > f_{\theta_0}^n \pi_0\}} = I_{\{\frac{f_{\theta_1}^n}{f_{\theta_0}^n} > \frac{\pi_0}{\pi_1}\}} = I_{\{z_n > \frac{\pi_0}{\pi_1}\}}$$

analogamente

$$I_{\{f_{\theta_1}^n \pi_1 \geq f_{\theta_0}^n \pi_0\}} = I_{\{z_n \geq \frac{\pi_0}{\pi_1}\}}$$

por lo tanto se cumple la desigualdad (3.24). □

Teorema 3.2. *Sean $\pi_0 > 0$, $\pi_1 > 0$ y las funciones $g(z)$ y $h(z)$ definidas en (3.2) y (3.20) respectivamente, y sea una prueba secuencial (φ^*, δ^*), que satisface*

$$I_{\{g(Z_n) < c + h(Z_n)\}} \leq \varphi_n^* \leq I_{\{g(Z_n) \leq c + h(Z_n)\}}, \tag{3.31}$$

μ^n-casi dondequiera sobre $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, 3 \ldots$ y

$$I_{\{Z_n > \frac{\pi_0}{\pi_1}\}} \leq \delta_n^* \leq I_{\{Z_n \geq \frac{\pi_0}{\pi_1}\}} \tag{3.32}$$

μ^n-casi dondequiera en $S_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots$.

Entonces para cualquier prueba que satisface

$$\alpha(\varphi, \delta) \leq \alpha(\varphi^*, \delta^*) \text{ y } \beta(\varphi, \delta) \leq \beta(\varphi^*, \delta^*) \tag{3.33}$$

se cumple,

$$N(\theta_o;\varphi) \geq N(\theta_o;\varphi^*). \tag{3.34}$$

La desigualdad en (3.34) es estricta, si alguna de las desigualdades en (3.33) lo es. Si existe igualdad en (3.33) y (3.34) entonces para las componentes de φ *se cumple (3.31)* μ^n*-casi dondequiera sobre* $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ *para cada* $n = 1, 2, 3 \ldots$, *y para las componentes de* δ *se cumple (3.32)* μ^n*-casi dondequiera en* $S_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ *para cada* $n = 1, 2, \ldots$.

Demostración. Por el Teorema 3.1 se cumple la desigualdad

$$R(\varphi,\delta) \geq R(\varphi^*,\delta^*) = c + Q_0$$

con $R(\varphi^*,\delta^*)$ definida en la ecuación (2.3) por lo tanto $N(\theta_o;\varphi) \geq N(\theta_o;\varphi^*)$. Por otro lado, la desigualdad es estricta si alguna de las desigualdades en (3.33) lo es, por el Teorema 2.1. Además si existe igualdad en (3.33) y (3.34) entonces para las componentes de φ se cumple (3.31) μ^n-casi dondequiera sobre $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, 3 \ldots$, y para las componentes de δ se cumple (3.32) μ^n-casi dondequiera en $S_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots$, por el mismo Teorema 3.1 porque se tiene la igualdad $R(\varphi,\delta) = R(\varphi^*,\delta^*) = c + Q_0$.

□

Este resultado es ya mucho mas explícito pero es posible darle una forma mas específica, demostrando que la desigualdad que define el paro en (3.31),

$$g(z) \leq c + h(z)$$

equivale simplemente a

$$z \notin (A, B),$$

con ciertas constantes A y B, tales que $0 < A < B < \infty$, por lo tanto quedando (3.31) de la siguiente forma:

$$I_{\{Z_r \notin [A,B]\}} \leq \varphi_r^* \leq I_{\{Z_r \notin (A,B)\}}.$$

y

$$I_{\{Z_r > B\}} \leq \delta_r^* \leq I_{\{Z_r \geq B\}}$$

Para demostrarlo necesitamos el siguiente:

Lema 3.3. *Las funciones* $\rho_n(z)$, *definidas en el Lema 3.1 y la función* $\rho(z)$ *todas poseen las siguientes propiedades:*

1. $\lim_{z\to 0} \rho_n(z) = \lim_{z\to 0} \rho(z) = 0$ *y* $\lim_{z\to\infty} \rho_n(z) = \lim_{z\to\infty} \rho(z) = \pi_0$, *para toda n=0,1,2,. . . .*
2. *son cóncavas y continuas para* $z \in [0, \infty)$,
3. *son crecientes para* $z \in [0, \infty)$.

Demostración. 1.)Se demostrará por inducción sobre n que

$$\lim_{z\to 0} \rho_n(z) = 0. \tag{3.35}$$

Para $n = 0$ se tiene, $\lim_{z\to 0}\rho_0(z) = \lim_{z\to 0} g(z) = g(0) = 0$, por lo tanto se cumple la base de la inducción.

Ahora supongamos que el resultado es válido para $n = r$, es decir, $\lim_{z\to 0}\rho_r(z) = 0$, entonces queda por demostrar que $\lim_{z\to 0}\rho_{r+1}(z) = 0$, pero por definición,

$$\lim_{z\to 0}\rho_{r+1}(z) = \lim_{z\to 0}\min\left\{g(z),\ \ c + \int f_{\theta_0}(x)\lim_{z\to 0}\rho_r\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x)\right\} = \min\{0, c+0\} = 0,$$

por el Teorema de convergencia dominada de Lebesgue, por lo tanto es válida la ecuación (3.35).

Por otra parte $\rho(z) = \lim_{n\to\infty}\rho_n(z)$, por la ecuación (3.18), y por lo tanto $\lim_{z\to 0}\rho(z) = \lim_{n\to\infty}\rho_n(0) = 0$, entonces en cualquier caso se cumple la igualdad (3.35).

Por otra parte, se demostrará también que

$$\lim_{z\to\infty}\rho_n(z) = \pi_0 \tag{3.36}$$

por inducción y además $\lim_{z\to\infty}\rho(z) = \pi_0$. Para $n = 0$, se tiene,

$$\lim_{z\to\infty}\rho_0(z) = \lim_{z\to\infty} g(z) = \pi_0,$$

es decir, se cumple la base de la inducción.

Ahora supóngase que es válida la igualdad (3.36) para $n = r$, es decir, $\lim_{z\to\infty}\rho_r(z) = \pi_0$ entonces queda por demostrar que $\lim_{z\to\infty}\rho_{r+1}(z) = \pi_0$. Pero por definición

$$\lim_{z\to\infty}\rho_{r+1}(z) = \min\left\{\lim_{z\to\infty} g(z),\ \ c + \lim_{z\to\infty}\int f_{\theta_0}(x)\rho_r\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)d\mu(x)\right\}$$

aplicando el Teorema de Convergencia Monótona a la función, $\rho_r\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)$ se tiene, $\lim_{z\to\infty}\rho_{r+1}(z) = \min\left\{\pi_0, c + \pi_0\int f_{\theta_0}(x)d\mu(x)\right\} = \pi_0$. Por lo tanto, se cumple la ecuación (3.36) en cualquier caso. Y otra vez por la ecuación (3.18),

$$\lim_{z\to\infty}\rho(z) = \lim_{z\to\infty}\lim_{n\to\infty}\rho_n(z) = \lim_{n\to\infty}\lim_{z\to\infty}\rho_n(z) = \pi_0.$$

con esto queda demostrado el punto número uno del Lema 3.3. 2.) Se demuestra por inducción que las ρ_n son cóncavas y continuas en $[0,\infty)$. Para esto será de mucha ayuda probar primero el siguiente Lema:

Lema 3.4. *El mínimo de dos funciones cóncavas es una función cóncava.*

Demostración. Sean G_1 y G_2 funciones cóncavas en $[0,\infty)$ entonces queremos demostrar que $G(x) = \min\{G_1(x), G_2(x)\}$ es una función cóncava en $[0,\infty)$, es decir

$$G(\eta x + (1-\eta)y) \geq \eta G(x) + (1-\eta)G(y) \tag{3.37}$$

Pero esto es sencillo puesto que como G_1 y G_2 son cóncavas, entonces se cumplen las siguientes desigualdades para cualesquiera $x, y \in [0,\infty)$ y $\eta \in [0,1]$ para G_1

$$G_1(\eta x + (1-\eta)y) \geq \eta G_1(x) + (1-\eta)G_1(y)$$

$$\geq \eta \min\{G_1(x), G_2(x)\} + (1-\eta)\min\{G_1(y), G_2(y)\} = \eta G(x) + (1-\eta)G(y), \tag{3.38}$$

y para G_2,

$$G_2(\eta x + (1-\eta)y) \geq \eta G_2(x) + (1-\eta)G_2(y)$$

$$\geq \eta \min\{G_1(x), G_2(x)\} + (1-\eta)\min\{G_1(y), G_2(y)\} = \eta G(x) + (1-\eta)G(y), \tag{3.39}$$

como se ve en las desigualdades (3.38), (3.39), se cumple (3.37) en cualquier caso. □

Regresando a la demostración del Lema 3.3, aplicando inducción sobre n.

Para $n = 0$ se tiene $\rho_0(z) = g(z) = \min\{\pi_0, \pi_1 z\}$ y esta función es cóncava por ser el mínimo de dos funciones cóncavas por el Lema 3.4, es decir, se cumple la base de la inducción.

Ahora supóngase que es válida la afirmación de que $\rho_n(z)$ es cóncava para $n = r$ es decir $\rho_r(z)$ es cóncava entonces queda por demostrar que $\rho_{r+1}(z)$ es cóncava, pero por definición:

$$\rho_{r+1}(z) = \min\left\{g(z),\ \ c + h_r(z)\right\}, \tag{3.40}$$

y por el Lema 3.4 el mínimo de dos funciones cóncavas es cóncava y la suma de funciones cóncavas es cóncava entonces, por la ecuación (3.40) sólo queda por demostrar que la función $h_r(z)$, es cóncava, y esto se demuestra en el siguiente:

Lema 3.5. *Suponiendo que la función $\rho_r(z)$ es cóncava entonces, $h_r(z)$ definida en la ecuación (3.12) es una función cóncava.*

Demostración. Por la definición de concavidad, lo que se debe demostrar es que la desigualdad es válida para cualesquiera z_1, $z_2 \in [0, \infty)$ y $\eta \in [0, 1]$, o equivalentemente

$$\eta h_r(z_1) + (1-\eta)h_r(z_2) \leq h_r(\eta z_1 + (1-\eta)z_2)$$

$$\eta \int f_{\theta_0}(x)\rho_r\left(z_1 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x) + (1-\eta)\int f_{\theta_0}(x)\rho_r\left(z_2 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x)$$

$$\leq \int f_{\theta_0}(x)\rho_r\left([\eta z_1 + (1-\eta)z_2]\,\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x)$$

desarrollando la desigualdad anterior mediante propiedades de suma y multiplicación por escalar de la integral se tiene

$$\int \left(\eta f_{\theta_0}(x)\rho_r\left(z_1 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) + (1-\eta)f_{\theta_0}(x)\rho_r\left(z_2 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)\right) d\mu(x)$$

$$\leq \int f_{\theta_0}(x)\rho_r\left([\eta z_1 + (1-\eta)z_2]\,\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x). \tag{3.41}$$

Pero por la propiedad de monotonía de las integrales es suficiente con demostrar lo siguiente:

$$\eta f_{\theta_0}(x)\rho_r\left(z_1 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) + (1-\eta)f_{\theta_0}(x)\rho_r\left(z_2 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) \leq f_{\theta_0}(x)\rho_r\left([\eta z_1 + (1-\eta)z_2]\,\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right).$$

Y esto equivale a mostrar que:

$$\eta \rho_r\left(z_1 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) + (1-\eta)\rho_r\left(z_2 \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) \leq \rho_r\left([\eta z_1 + (1-\eta)z_2]\,\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)},\right) \tag{3.42}$$

pero la desigualdad (3.42) es válida porque $\rho_r(z)$ es una función cóncava por hipótesis, por lo tanto la desigualdad (3.41) es válida y el Lema 3.5 se cumple.

□

Así, la función $\rho_{r+1}(z)$ es cóncava. Por otra parte, el límite puntual de funciones cóncavas es una función cóncava y entonces $\rho(z)$ es cóncava. Además la cóncavidad implica continuidad (ver [12] section 10 Continuity of Convex Functions.), y por lo tanto las funciones $\rho_n(z)$ y $\rho(z)$ son todas continuas en $(0, \infty)$, y también en cero ya que por el inciso 1, $\lim_{z\to 0} \rho(z) = 0$ y por lo tanto $\rho(z)$ es continua en $[0, \infty)$, análogamente para cada ρ_n se tiene la continuidad en $z = 0$, con esto queda demostrado el punto número dos del Lema 3.3.

3.)Se demuestra por inducción sobre n que $\rho(z)$ y las $\rho_n(z)$ son crecientes en $[0, \infty)$.

Para $n = 0$ es obvio púes $\rho_0(z) = g(z) = \min\{\pi_0, \pi_1 z\}$, es el mínimo de dos funciones lineales con pendiente positiva, por lo tanto se cumple la base de la inducción.

Ahora supóngase que la afirmación es válida para $n = r$, es decir $\rho_r(z)$ es creciente, entonces queda por demostrar que,

$$\rho_{r+1}(z) = \min \{g(z), \;\; c + h_r(z)\}\,, \tag{3.43}$$

es creciente.

Pero sólo será necesario demostrar que la función $h_r(z)$ es creciente, y esto se prueba en el siguiente:

Lema 3.6. *Suponiendo que la función $\rho_r(z)$ es creciente, entonces $h_r(z)$ es creciente con $h_r(z)$ definida por la ecuación (3.12).*

Demostración. Por la definición de una función creciente se debe probar la siguiente desigualdad: $h_r(z_1) \le h_r(z_2)$, para cualesquiera z_1 y z_2 en $[0, \infty)$ con $z_1 < z_2$ o equivalentemente,

$$\int f_{\theta_0}(x)\rho_r\left(z_1\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x) \le \int f_{\theta_0}(x)\rho_r\left(z_2\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x).$$

Por la propiedad de monotonía de la integral, será suficiente con probar la siguiente desigualdad,

$$\rho_r\left(z_1\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) \le \rho_r\left(z_2\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right), \tag{3.44}$$

pero la desigualdad (3.44) se cumple trivialmente porque $\rho_r(z)$ es una función creciente en $[0, \infty)$ por hipótesis. □

Regresando a la demostración del punto número tres del Lema 3.3 hemos demostrado que las funciones $\rho_n(z)$ son crecientes en $[0, \infty)$. Además el límite de funciones crecientes es una función creciente, así que $\rho(z)$ es creciente, por lo tanto el Lema 3.3 queda demostrado.

□

Definición 3.3. *Se definen*

$$A = \sup\{z : 0 \le z \le \pi_0/\pi_1, \;\; c + h(z) \ge g(z)\} \tag{3.45}$$

$$B = \inf\{z : \pi_0/\pi_1 \le z, \;\; c + h(z) \ge g(z)\}\,. \tag{3.46}$$

Por el Lema 3.3 anterior $h(0) = 0$ y entonces $c + h(0) = c > g(0) = 0$, por lo tanto $A > 0$.

Por otra parte, por el mismo Lema 3.3,

$$\lim_{z\to\infty} h(z) = \pi_0,$$

por lo que,

$$c + \pi_0 = \lim_{z\to\infty}[c + h(z)] > \lim_{z\to\infty} g(z) = \pi_0,$$

por lo tanto resulta que $B < \infty$. Además si $A = B = \frac{\pi_0}{\pi_1}$ entonces $c+h(z) \geq g(z)$ para cualquier $z \geq 0$, así el tiempo de paro óptimo es $\tau_\varphi \equiv 1$. En cualquier otro caso $A < B$. El siguiente Lema muestra la relación que existe entre las constantes A, B y la desigualdad que define el paro en (3.31).

Lema 3.7. *$A < z < B$ si y solamente si $g(z) > c + h(z)$, con A y B definidas por las ecuaciones (3.45), (3.46) respectivamente y las funciones $g(z)$ y $h(z)$ definidas en (3.3), (3.12) respectivamente.*

Demostración. Por continuidad de todas las funciones,

$$g(A) = c + h(A) \tag{3.47}$$

$$g(B) = c + h(B). \tag{3.48}$$

Además por definición $c+h(z) < g(z)$ para $z \in (A, \pi_0/\pi_1]$. También por definición $c+h(z) < g(z)$ para $z \in [\pi_0/\pi_1, B)$. Ahora si $z \in [0, A)$, entonces existe $1 \geq a > 0$, tal que $z = (1-a)A$. Como $h(0) = 0$ por concavidad de $h(z)$ se tiene:

$$h(z) = h(a0 + (1-a)A) \geq ah(0) + (1-a)h(A) = (1-a)(g(A) - c),$$

por (3.47), entonces,

$$c + h(z) \geq c + (1-a)(g(A) - c) = ac + (1-a)g(A) > (1-a)g(A) = g((1-a)A) = g(z),$$

porque $g(z) = \pi_1 z$ para $z \in [0, \pi_0/\pi_1]$, entonces, $c + h(z) > g(z)$ en $[0, A)$.

Por otro lado, si $z \in (B, \infty)$, entonces sabemos por el Lema 3.3 que el límite

$$\lim_{z\to\infty}[c + h(z)] = c + \pi_0,$$

por lo tanto existe $z^{'}$, tal que,

$$c + h(z) > \pi_0 + c/2, \tag{3.49}$$

para toda $z \geq z^{'}$. Así, la desigualdad $c + h(z) > g(z) = \pi_0$ se cumple trivialmente para $z \geq z^{'}$. Ahora para $z < z^{'}$, existe $1 > a \geq 0$ tal que $z = aB + (1-a)z^{'}$ y por la concavidad de $c + h(z)$ se cumple la siguiente desigualdad:

$$c + h(z) \geq a(c + h(B)) + (1-a)(c + h(z^{'})), \tag{3.50}$$

pero por la ecuación (3.49) el lado derecho de (3.50) es mayor que,

$$a(c + h(B)) + (1-a)(\pi_0 + c/2),$$

que a su vez por (3.48) es igual a,

$$= ag(B) + (1-a)(\pi_0 + c/2) = \pi_0 + (1-a)c/2 > \pi_0 = g(z)$$

y por lo tanto $c + h(z) > g(z)$ en (B, ∞), y con esto queda demostrado el Lema. $\square$

Ahora será posible caracterizar tanto a las componentes de la regla de paro φ, como a las de la regla de decisión δ, en terminos de A y de B, y la prueba de este hecho se ve en el siguiente:

Teorema 3.3. *Sean Q_0 definida en la ecuación (3.21), $\pi_0 > 0$, $\pi_1 > 0$, y A, B definidas en las ecuaciones (3.45), (3.46), respectivamente, entonces para cada prueba (φ, δ) se cumple:*

$$R(\varphi, \delta) = c + Q_0,$$

para

$$I_{\{z_n \notin [A,B]\}} \leq \varphi_n \leq I_{\{z_n \notin (A,B)\}} \tag{3.51}$$

μ^n-casi dondequiera sobre $T_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ $n = 1, 2, 3 \ldots,$ y

$$I_{\{z_n > B\}} \leq \delta_n \leq I_{\{z_n \geq B\}} \tag{3.52}$$

μ^n-casi dondequiera en $S_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots,$

Demostración. Por el Teorema 3.2 sabemos que:

$$I_{\{g(z_n) < c + h(z_n)\}} \leq \varphi_n \leq I_{\{g(z_n) \leq c + h(z_n)\}}$$

μ^n-casi dondequiera sobre $T_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ $n = 1, 2, 3 \ldots$

Por el Lema 3.7 $A < z < B$, si y sólo si, $g(z) > c + h(z)$, o lo que es lo mismo, $z \in (A, B)$, si y sólo si, $g(z) > c + h(z)$, por lo tanto, $g(z_n) < c + h(z_n)$, si y sólo si, $z_n \notin [A, B]$. Además también se cumple que $g(z_n) \leq c + h(z_n)$, si y sólo si, $z_n \notin (A, B)$. Esto demuestra la ecuación (3.51). Y por el mismo Teorema 3.2

$$\delta^* = I_{\{z_n \geq \frac{\pi_0}{\pi_1}\}}$$

Así que por el Lema 3.7 $z_n \geq \frac{\pi_0}{\pi_1}$ y $g(z_n) \leq c + h(z_n)$ si y sólo si $z_n \geq B$, analogamente $g(z_n) > c + h(z_n)$ si y sólo si $z_n > B$, por lo tanto,

$$I_{\{z_n > B\}} \leq \delta \leq I_{\{z_n \geq B\}}.$$

Esto demuestra la ecuación (3.51) y por lo tanto el Teorema. □

Teorema 3.4. *Sean $\pi_0 > 0$, $\pi_1 > 0$, y A, B definidas en las ecuaciones (3.45) y (3.46) respectivamente, entonces para cada prueba (φ^*, δ^*) que satisface*

$$I_{\{z_n \notin [A,B]\}} \leq \varphi_n^* \leq I_{\{z_n \notin (A,B)\}}, \tag{3.53}$$

μ^n-casi dondequiera sobre $T_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$ $n = 1, 2, 3 \ldots,$ y

$$\delta_n^* = I_{\{z_n \geq B\}}, \tag{3.54}$$

μ^n-casi dondequiera en $S_n^{\varphi} \bigcap \{f_{\theta_0}^n > 0\}$, para cada $n = 1, 2, \ldots$. Entonces para cualquier prueba (φ, δ), que satisface

$$\alpha(\varphi, \delta) \leq \alpha(\varphi^*, \delta^*) \ \text{y} \ \beta(\varphi, \delta) \leq \beta(\varphi^*, \delta^*) \tag{3.55}$$

se cumple,

$$N(\theta_o; \varphi) \geq N(\theta_o; \varphi^*). \tag{3.56}$$

La desigualdad en (3.56) es estricta, si alguna de las desigualdades en (3.55) lo es.

Si las dos desigualdades en (3.55) son igualdades, y, además, (3.56) es igualdad, entonces para las componentes de φ se cumple (3.53) μ^n-casi dondequiera en $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$, $n = 1, 2, \ldots,$ y para las componentes de δ se cumple (3.54) μ^n-casi dondequiera en $S_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$, $n = 1, 2, \ldots$.

Demostración. Por el Teorema 3.3 se cumple la desigualdad

$$R(\varphi, \delta) \geq R(\varphi^*, \delta^*)$$

por lo tanto $N(\theta_o; \varphi) \geq N(\theta_o; \varphi^*)$, y la desigualdad es estricta si lo es alguna de las desigualdades en (3.55) por el Teorema 2.1.

Además si existe igualdad en las desigualdades (3.55) y (3.56), entonces para las componentes de φ se cumple (3.53) μ^n-casi dondequiera sobre $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, 3 \ldots,$ y para las componentes de δ se cumple (3.54) μ^n-casi dondequiera en $S_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}$ para cada $n = 1, 2, \ldots,$ por el mismo Teorema 3.3, porque en este caso se tiene la igualdad $R(\varphi, \delta) = R(\varphi^*, \delta^*) = c + Q_0$. □

El Teorema 3.4 trata la optimalidad de pruebas secuenciales en el problema con restricciones (sobre las probabilidades de error tipo uno y tipo dos). A su vez, la forma de la prueba óptima se obtuvo de la solución del problema de minimización de la función de Lagrange. En otras palabras, la prueba que resulta óptima en el Teorema 3.4 se obtuvo del problema de paro óptimo correspondiente al mínimo de la función de Lagrange. Es decir, si usamos el método de Lagrange para resolver el problema con restricciones, la solución está dada en el Teorema 3.4. Por lo tanto no queda claro si existen otras soluciones del problema condicional que no sean obtenidas por el método de Lagrange.

En esta parte de la tesis vamos a demostrar que todas soluciones del problema con restricciones (si éstas existen) que son de interés práctico, pueden ser obtenidas a través del método de Lagrange. De esta manera va a resultar que todas soluciones al problema con restricciones en este caso se dan en el Teorema 3.4.

Teorema 3.5. *Supongamos que existe una prueba secuencial (φ^*, δ^*) tal que $N(\theta_0, \varphi^*) < \infty$ y que es óptima en el siguiente sentido (problema con restricciones): para cualquier prueba (φ, δ) con*

$$\alpha(\varphi, \delta) \leq \alpha(\varphi^*, \delta^*), \quad y \quad \beta(\varphi, \delta) \leq \beta(\varphi^*, \delta^*) \tag{3.57}$$

se cumple

$$N(\theta_0, \varphi) \geq N(\theta_0, \varphi^*). \tag{3.58}$$

Entonces existen $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ (no todas iguales a cero) tales que

$$b_1 N(\theta_0, \varphi) + b_2 \alpha(\varphi, \delta) + b_3 \beta(\varphi, \delta) \geq b_1 N(\theta_0, \varphi^*) + b_2 \alpha(\varphi^*, \delta^*) + b_3 \beta(\varphi^*, \delta^*). \tag{3.59}$$

para cualquier prueba (φ, δ).

Demostración. Denotemos

$$K = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \; x_1 = \alpha(\varphi, \delta), \, x_2 = \beta(\varphi, \delta), \, x_3 = N(\theta_0, \varphi)$$

para alguna prueba $(\varphi, \delta)\}$.

Mostremos primero que K es un subconjunto convexo en $\mathbb{R}^3$.

Sean $x, y \in K$ dos puntos cualesquiera. Esto implica, por definición que existen dos pruebas (φ_1, δ_1) y (φ_2, δ_2) tales que

$$x = (\alpha(\varphi_1, \delta_1), \beta(\varphi_1, \delta_1), N(\theta_0, \varphi_1))$$

y

$$y = (\alpha(\varphi_2, \delta_2), \beta(\varphi_2, \delta_2), N(\theta_0, \varphi_2)).$$

Demostremos que para cualquier $\lambda \in [0, 1]$ el punto $\lambda x + (1 - \lambda)y \in K$. Para esto construimos una prueba (φ, δ) tal que

$$\alpha(\varphi, \delta) = \lambda\alpha(\varphi_1, \delta_1) + (1 - \lambda)\alpha(\varphi_2, \delta_2) \tag{3.60}$$

$$\beta(\varphi, \delta) = \lambda\beta(\varphi_1, \delta_1) + (1 - \lambda)\beta(\varphi_2, \delta_2) \tag{3.61}$$

$$N(\theta_0, \varphi) = \lambda N(\theta_0, \varphi) + (1 - \lambda)N(\theta_0, \varphi_2). \tag{3.62}$$

Definamos consecutivamente para $n = 1, 2, \ldots$: $t_n^{\varphi} = \lambda t_n^{\varphi_1} + (1-\lambda)t_n^{\varphi_2}$. Debido a que $t_1^{\varphi_1}$, $t_2^{\varphi_1}, \ldots$ y $t_1^{\varphi_2}$, $t_2^{\varphi_2}, \ldots$ son dos sucesiones decrecientes, se tiene que t_1^{φ}, $t_2^{\varphi}, \ldots$ también lo es. Partiendo de dos valores consecutivos t_k^{φ} y t_{k+1}^{φ}, si $t_k^{\varphi} > 0$, se determina unívocamente el valor de φ_k, ya que $t_{k+1}^{\varphi}/t_k^{\varphi} = 1 - \varphi_k$, por definición de t_k^{φ}. Si $t_k^{\varphi} = 0$ entonces todos los valores posteriores de t_n^{φ}, $n > k$, son ceros, y φ_n para $n > k$ se definen arbitrariamente. Es fácil ver que $s_n^{\varphi} = t_n^{\varphi} - t_{n+1}^{\varphi}$ y por lo tanto también

$$s_n^{\varphi} = \lambda s_n^{\varphi_1} + (1 - \lambda)s_n^{\varphi_2},\ n = 1, 2, \ldots. \tag{3.63}$$

Además, sea definido δ_n como

$$\frac{1}{s_n^{\varphi}}(\lambda s_n^{\varphi_1}\delta_{1n} + (1 - \lambda)s_n^{\varphi_2}\delta_{2n}). \tag{3.64}$$

De (3.63) y (3.64) se siguen (3.60), (3.61) y (3.62) de inmediato ya que, por ejemplo,

$$\alpha(\varphi, \delta) = \sum_{n=1}^{\infty} E_{\theta_0} s_n^{\varphi}\delta_n = \lambda \sum_{n=1}^{\infty} E_{\theta_0} s_n^{\varphi_1}\delta_{1n} + (1 - \lambda) \sum_{n=1}^{\infty} E_{\theta_0} s_n^{\varphi_2}\delta_{2n},$$

etc.

Sean $x^* = (x_1^*, x_2^*, x_3^*) = (\alpha(\varphi^*, \delta^*), \beta(\varphi^*, \delta^*), N(\theta_0, \varphi^*))$ y $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3 :\ x_1 \leq x_1^*,\ x_2 \leq x_2^*,\ x_3 < x_3^*\}$. Notemos que M también es un subconjunto convexo de $\mathbb{R}^3$, y tal que, por la condición del Teorema, $K \bigcap M = \emptyset$.

Entonces por el teorema de separación por hiperplanos (ver, por ejemplo, Teorema 11.3 en [12]) existe $b = (b_1, b_2, b_3) \neq 0$ tal que

$$\langle y, b\rangle \geq \langle x, b\rangle \tag{3.65}$$

para cualquier $y \in K$ y cualquier $x \in M$. Considerando $x = (x_1^*, x_2^*, x_3) \in M$ (con $x_3 < x_3^*$) y haciendo $x_3 \to x_3^*$ tenemos que

$$\langle y, b\rangle \geq \langle x^*, b\rangle \tag{3.66}$$

para cualquier $y \in K$. Además, de (3.65) se sigue que $b_1 \geq 0$ y $b_2 \geq 0$ (ya que de otra manera (3.65) no se puede cumplir para x_1, x_2 o x_3 negativas grandes en valor absoluto).

Observación. Si $b_1 > 0$ en (3.59) entonces se puede usar $\pi_0 = b_2/b_1$ y $\pi_1 = b_3/b_1$ en el Teorema 3.4 para obtención de la prueba óptima. En caso $b_1 = 0$ el problema de minimización en (3.59) es trivial, ya que la prueba con $\varphi_n^* \equiv 0$, $n = 1, 2, \ldots$ tiene probabilidades de error iguales a cero. De esta manera, queda demostrado que cualquier caso no trivial del problema con restricciones está cubierto por el Teorema 3.4. □

De manera intuitiva y dados los resultados obtenidos podemos pensar que para cualesquiera que sean las constantes $0 < A < B < \infty$, existen c, π_0 y π_1 tales que la SPRT aleatorizada (3.53) y (3.54) minimiza $R(\varphi, \delta)$, con estas constantes, debido a que sólo hay dos restricciones por cumplir, las ecuaciones (3.47) y (3.48) y las variables son tres, c, π_0 y π_1.

De las ecuaciones (3.3) y (3.19) se ve que en realidad las funciones $\rho(z)$, $h(z)$ y $g(z)$ dependen también de las constantes c, π_0 y π_1 por lo que en lo siguiente se denotaran como:

$$\rho(z) = \rho(z; c, \pi_0, \pi_1) = \text{mín}\left\{\text{mín}\{\pi_0, \pi_1 z\}, c + \int f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, \pi_1\right) d\mu(x)\right\},$$

$$h(z) = h(z; c, \pi_0, \pi_1) = \int f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, \pi_1\right) d\mu(x)$$

$$g(z; \pi_0, \pi_1) = \text{mín}\{\pi_0, \pi_1 z\}.$$

Teorema 3.6. *Sean A y B dos constantes tales que $0 < A < B < \infty$, entonces existen constantes $c > 0, \pi_0 > 0, \pi_1 > 0$ tales que A y B son raíces de la ecuación,*

$$c + h(z; c, \pi_0, \pi_1) = g(z, \pi_0, \pi_1).$$

Demostración. Esto da origen al siguiente sistema de ecuaciones:

$$c + h(A; c, \pi_0, \pi_1) = \pi_1 A \quad y \quad c + h(B; c, \pi_0, \pi_1) = \pi_0. \tag{3.67}$$

Y para resolverlo primero que nada recordemos que la prueba óptima (φ, δ) se deriva de minimizar,

$$cN(\theta_o, \varphi) + \pi_0\alpha(\varphi, \delta) + \pi_1\beta(\varphi, \delta). \tag{3.68}$$

Y por el Teorema 2.5,

$$c + \int \rho\left(\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, \pi_1\right) f_{\theta_0}(x)d\mu(x) = \inf_{\varphi}\left\{cN(\theta_0, \varphi) + \pi_0\alpha(\varphi, \delta^*) + \pi_1\beta(\varphi, \delta^*)\right\}. \tag{3.69}$$

Pero es necesario tener una función que dependa de z del lado izquierdo de (3.69) para tener,

$$c + \int \rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, \pi_1\right) f_{\theta_0}(x)d\mu(x) = \inf_{\varphi}\left\{cN(\theta_0, \varphi) + \pi_0\alpha(\varphi, \delta^*) + z\pi_1\beta(\varphi, \delta^*)\right\}. \tag{3.70}$$

Para ello necesitamos el siguiente:

Lema 3.8.

$$\rho_n(z'; c, \pi_0, z\pi_1) = \rho_n(zz'; c, \pi_0, \pi_1) \tag{3.71}$$

Demostración. La demostración se hace por inducción sobre n.

Para $n = 0$ se tiene:

$$\rho_n(z'; c, \pi_0, z\pi_1) = \text{mín}\{\pi_0, z\pi_1 z'\} = \text{mín}\{\pi_0, \pi_1 zz'\} = \rho_n(zz'; c, \pi_0, \pi_1)$$

Supongamos que (3.71) se cumple para $n = k$, esto es,

$$\rho_k(z'; c, \pi_0, z\pi_1) = \rho_k(zz'; c, \pi_0, \pi_1). \tag{3.72}$$

Queda por demostrar la igualdad (3.71) para $n = k + 1$, por la definición en la ecuación (3.3),

$$\rho_{k+1}(z'; c, \pi_0, z\pi_1) = \text{mín}\left\{\text{mín}\{\pi_0, z\pi_1 z'\}, c + \int \rho_k\left(z'\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, z\pi_1\right) f_{\theta_0}(x) d\mu(x)\right\},$$

sustituyendo (3.72) en la ecuación anterior esta es igual a lo siguiente:

$$= \text{mín}\left\{\text{mín}\{\pi_0, \pi_1 zz'\}, c + \int \rho_k\left(zz'\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi_0, \pi_1\right) f_{\theta_0}(x) d\mu(x)\right\} = \rho_{k+1}(zz'; c, \pi_0, \pi_1),$$

y por lo tanto,

$$\rho_{k+1}(z'; c, \pi_0, z\pi_1) = \rho_{k+1}(zz'; c, \pi_0, \pi_1),$$

que es lo que se quería demostrar. □

Ahora la ecuación (3.70), se puede ver entonces como sigue:

$$c + \int \rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}; c, \pi,\right) f_{\theta_0}(x) d\mu(x) = \inf_{\varphi}\left\{cN(\theta_0, \varphi) + \pi\alpha(\varphi, \delta^*) + z\beta(\varphi, \delta^*)\right\}, \tag{3.73}$$

por el Lema 3.8. Sin pérdida de generalidad haciendo $\pi_1 = 1$ y $\pi_0 = \pi$ en las ecuaciones (3.67) ya que π_1 no influye de manera determinante en el sistema de ecuaciones y al fijar este valor se obtiene un sistema de ecuaciones mas sencillo que es el siguiente:

$$c + h(A; c, \pi) = A \quad y \quad c + h(B; c, \pi) = \pi. \tag{3.74}$$

Estas nuevas ecuaciones son el sistema que ahora se debe resolver, y para ello se prueba el siguiente Lema sobre la continuidad de $h(z; c, \pi)$ con respecto de la varible c.

Lema 3.9. *La función $h(z; c, \pi)$ es continua para toda $c \in [0, \infty)$.*

Demostración. Se puede ver de la ecuación (3.73) que $h(z; c, \pi)$ es cóncava y no decreciente para las variables c, π (púes son el ínfimo de funciones lineales, manteniendo las otras variables fijas). Inmediatamente se sigue de la ecuación (3.73) que $h(z; c, \pi)$ es continua para toda $c \in (0, \infty)$.

Ahora se demuestra que la función (3.73) es continua en $c = 0$ para π , z fijas. Se toma una muestra de tamaño fijo, entonces $\varphi_i \equiv 0$, $i < n$, $\varphi_n \equiv 1$ y aplicando la misma técnica de minimización a (3.73), que la que se aplicó a (3.69) en el Teorema 2.2, se obtiene $\delta_n^* = I_{\{(\pi_0 f_{\theta_o}^n - z\pi_1 f_{\theta_1}^n \leq 0\}}$, y recordando que ahora $\pi_1 = 1$, $\pi_0 = \pi$, entonces, $\delta_n^* = I_{\{(\pi f_{\theta_o}^n - z f_{\theta_1}^n \leq 0\}}$, entonces el lado derecho de (3.73) no excede,

$$nc + \pi P_{\theta_0}(\pi f_{\theta_o}^n - z f_{\theta_1}^n \leq 0) + z P_{\theta_1}(\pi f_{\theta_o}^n - z f_{\theta_1}^n > 0)$$

$$= nc + \pi P_{\theta_0}(f^n_{\theta_1}/f^n_{\theta_o} \geq \pi/z) + zP_{\theta_1}(f^n_{\theta_o}/f^n_{\theta_1} > z/\pi) \tag{3.75}$$

y aquí se puede ver que, $P_{\theta_0}(f^n_{\theta_1}/f^n_{\theta_o} \geq \pi/z) \to 0$ cuando $n \to \infty$.

Efectivamente si

$$\mu\{x : f_{\theta_o}(x) \neq f_{\theta_1}(x)\} > 0, \tag{3.76}$$

para cada $\epsilon > 0$ se tiene:

$$P_{\theta_0}(f^n_{\theta_1}/f^n_{\theta_o} \geq \epsilon) = P_{\theta_0}(Z_n \geq \epsilon) = P_{\theta_0}(Z_n^{1/2} \geq \epsilon^{1/2}) \leq \epsilon^{-1/2} E_{\theta_o}\left(\prod_{i=1}^{n} \frac{f_{\theta_1}(X_i)}{f_{\theta_o}(X_i)}\right)^{1/2},$$

$$= \epsilon^{-1/2}\left(E_{\theta_0}\left(\frac{f_{\theta_1}(X_i)}{f_{\theta_o}(X_i)}\right)^{1/2}\right)^{n} \to 0,$$

cuando $n \to \infty$ por (3.76),

$$q = E_{\theta_0}\left(\frac{f_{\theta_1}(X_i)}{f_{\theta_o}(X_i)}\right)^{1/2} < 1.$$

Análogamente se demuestra que, $P_{\theta_1}(f^n_{\theta_0}/f^n_{\theta_1} > z/\pi_0) \to 0$ cuando $n \to \infty$. Así para cada $\epsilon > 0$, (3.75) no excede

$$nc + 2\epsilon, \tag{3.77}$$

cuando $n \geq N$. Si ahora $c < \epsilon/N$, entonces (3.77), no excede 3ϵ. Esto prueba que el lado derecho de (3.73), tiende a cero cuando $c \to 0$. □

Lema 3.10. *Para cada z fija* $h(z; c, \pi)$, *es continua como una función de* (c, π), *en cada punto* (c_0, π_0), *que satisface* $c_0 \geq 0$, $\pi_0 > 0$.

Demostración. Sean $c_0 \geq 0$ y $\pi_0 > 0$ fijos, entonces para cada $\pi_1 < \pi_0 < \pi_2$, se tiene

$$h(z; c, \pi_1) \leq h(z; c, \pi) \leq h(z; c, \pi_2),$$

si $\pi \in [\pi_1, \pi_2]$.

Ahora,

$$\lim_{c\to c_0} h(z; c, \pi_1) \leq \liminf_{c\to c_0, \pi\to\pi_0} h(z; c, \pi) \leq \limsup_{c\to c_0, \pi\to\pi_0} h(z; c, \pi) \leq \lim_{c\to c_0} h(z; c, \pi_2), \tag{3.78}$$

además se tiene que,

$$\liminf_{c\to c_0} h(z; c, \pi_1) = h(z; c_0, \pi_1), \qquad \limsup_{c\to c_0} h(z; c, \pi_2) = h(z; c_0, \pi_2),$$

entonces la cadena de desigualdades en (3.78) se puede ver como,

$$h(z; c_0, \pi_1) \leq \liminf_{c\to c_0, \pi\to\pi_0} h(z; c, \pi) \leq \limsup_{c\to c_0, \pi\to\pi_0} h(z; c, \pi) \leq h(z; c_0, \pi_2). \tag{3.79}$$

Ahora haciendo tender $\pi_1 \to \pi_0$ y $\pi_2 \to \pi_0$ en las ecuaciones (3.79) se tiene:

$$\lim_{c\to c_0, \pi_1\to\pi_0} h(z; c, \pi_1) = h(z; c_0, \pi_0), \qquad \lim_{c\to c_0, \pi_2\to\pi_0} h(z; c, \pi_2) = h(z; c_0, \pi_0),$$

y por lo tanto se obtiene que

$$\lim_{c\to c_0, \pi\to\pi_0} h(z; c, \pi) = h(z; c_0, \pi_0).$$

Y esto prueba el Lema.

□

Ahora sean A y B dos números con $0 < A < B < \infty$, entonces para cada $\pi \in [A, B]$, se define $c = c(\pi)$ como solución de la ecuación

$$c + h(A; c, \pi) = A \tag{3.80}$$

$c(\pi)$ está bien definida para cualquier $\pi \in [A, B]$, porque el lado izquierdo de (3.80), es una función continua y estrictamente creciente como función de c, (ver Lema 3.5). Además $c(\pi)$ es una función continua como función implícita de π, ver (3.73) es una función que es continua en dos variables (Lema 3.10). También se tiene que $c(\pi) \geq 0$ porque $h(A; c, \pi) \leq g(A; \pi) = \min\{A, \pi\} = A$ para $\pi \geq A$, entonces,

$$c(\pi) = A - h(A; c, \pi) \geq 0.$$

Ahora se define,

$$G(\pi) = \pi - h(B; c(\pi), \pi) - c(\pi). \tag{3.81}$$

Afirmamos que (3.81) es una función continua en función de π como composición de funciones continuas, y para ello se demostrará el siguiente Lema:

Lema 3.11. *Existe π, con $A < \pi < B$ tal que*

1. $G(A) < 0$,
2. $G(B) > 0$, *y*
3. $G(\pi) = 0$.

Demostración. Para el primer caso se tiene lo siguiente:

$$G(A) = A - h(B; c(A), A) - c(A) < A - h(A; c(A), A) - c(A) = 0$$

por (3.80) y porque $c(A) + h(z; c(A), A)$ es estrictamente creciente en una vecindad de A (Lema 3.3).

Ahora se mostrará que:

$$G(B) = B - h(B; c(B), B) - c(B) > 0, \tag{3.82}$$

tomando en cuenta (3.80) se tiene,

$$c(B) + h(A; c(B), B) = A$$

con lo que (3.82) es equivalente a,

$$B - h(B; c(B), B) > A - h(A; c(B), B).$$

Y esta última desigualdad es válida porque $z-h(z;c(B),B)$ es creciente en una vecindad de B, porque de no ser así, como sabemos $h(z;c(B),B)\leq g(z;B)$ y entonces $h(z;c(B),B)\equiv g(z;B)$, de donde se sigue que $c(B)+h(A;c(B),B)=A$, se puede ver como $c(B)+g(A;B)=A$, esto es $c(B)+\text{mín}\{A,B\}=c(B)+A=A$, entonces $c(B)=0$, pero $c(B)+h(A;c(B),B)$ es una función creciente de c y $c(B)+h(A;c(B),B)=0$ en $c=0$, por lo tanto no puede ser nula para $c>0$, así, por el Teorema 1 del capítulo 7 del libro [16] se sigue que existe π, tal que $G(\pi)=0$ y esto prueba el Lema. □

El Lema anterior prueba que existe $\pi\in(A,B)$ y $c(\pi)$ tales que,

$$c+h(A;c,\pi)=A \quad y \quad c+h(B;c,\pi)=\pi.$$

Y esto era lo que se quería demostrar, por lo tanto el Teorema 3.6, queda demostrado. □

Teorema 3.7. *Sean A y B constantes cualesquiera tales que $0<A<B<\infty$ y sea (φ^*,δ^*) con*

$$I_{\{Z_n\notin[A,B]\}}\leq\varphi_n^*\leq I_{\{Z_n\notin(A,B)\}} \tag{3.83}$$

μ^n-casi dondequiera en $T_n^\varphi\bigcap\{f_{\theta_0}^n>0\}$ $n=1,2,3\ldots$ y

$$\delta_n^*=I_{\{Z_n\geq B\}} \tag{3.84}$$

μ^n-casi dondequiera en $S_n^\varphi\bigcap\{f_{\theta_0}^n>0\}$ para cada $n=1,2,\ldots$.

Entonces para cualquier prueba (φ,δ), tal que,

$$\alpha(\varphi,\delta)\leq\alpha(\varphi^*,\delta^*) \quad y \quad \beta(\varphi,\delta)\leq\beta(\varphi^*,\delta^*), \tag{3.85}$$

se cumple

$$N(\theta_0,\varphi)\geq N(\theta_0,\varphi^*), \tag{3.86}$$

la desigualdad en (3.86) es estricta, si lo es alguna de las desigualdades en (3.85).

Si las dos desigualdades en (3.85) son igualdades, y además, (3.86) es igualdad, entonces para las componentes de φ se cumple (3.83) μ^n-casi dondequiera en $C_n^\varphi\bigcap\{f_{\theta_0}^n>0\}$, $n=1,2,\ldots,$ y para las componentes de δ se cumple (3.84) μ^n-casi dondequiera en $S_n^\varphi\bigcap\{f_{\theta_0}^n>0\}$, $n=1,2,\ldots$.

Demostración. La prueba es consecuencia directa de los Teoremas 3.4 y 3.6. □

Teorema 3.8. *Sean A y B dos constantes tales que $0<A<B<\infty$ y sea (φ^*,δ^*) una prueba que satisface*

$$I_{\{Z_n\notin[A,B]\}}\leq\varphi_n^*\leq I_{\{Z_n\notin(A,B)\}} \tag{3.87}$$

μ^n-casi dondequiera en $T_n^\varphi\bigcap\{\{f_{\theta_0}^n>0\}\bigcup\{f_{\theta_1}^n>0\}\}$, $n=1,2,\ldots,$ y

$$\delta_n=I_{\{Z_n\geq B\}} \tag{3.88}$$

μ^n-casi dondequiera en $S_n^\varphi\bigcap\{\{f_{\theta_0}^n>0\}\bigcup\{f_{\theta_1}^n>0\}\}$, $n=1,2,\ldots$.

Entonces para cualquier prueba (φ,δ) tal que

$$\alpha(\varphi,\delta)\leq\alpha(\varphi^*,\delta^*) \quad y \quad \beta(\varphi,\delta)\leq\beta(\varphi^*,\delta^*) \tag{3.89}$$

se tiene que

$$N(\theta_0, \varphi) \geq N(\theta_0, \varphi^*) \quad y \quad N(\theta_1, \varphi) \geq N(\theta_1, \varphi^*). \tag{3.90}$$

Las dos desigualdades en (3.90) son estrictas si alguna de las desigualdades en (3.89) lo es.

Si las dos desigualdades en (3.89) son igualdades, y, además, alguna de las (3.90) es igualdad, entonces para las componentes de φ y δ se cumplen (3.87) y (3.88) μ^n-casi dondequiera en los conjuntos correspondientes: en $T_n^\varphi \bigcap \{f_{\theta_0}^n > 0\}\}$, y, respectivamente, en $S_n^\varphi \bigcap \{\{f_{\theta_0}^n > 0\}$, $n = 1, 2, \ldots$, si se cumple igualdad en la primera de las desigualdades (3.90), y en $T_n^\varphi \bigcap \{f_{\theta_1}^n > 0\}\}$, y, respectivamente, en $S_n^\varphi \bigcap \{\{f_{\theta_1}^n > 0\}$, $n = 1, 2, \ldots$, si se cumple igualdad en la segunda de las desigualdades (3.90).

Nota. *Las pruebas que satisfacen (3.87) y (3.88) son llamadas* extended SPRT *(SPRT extendida) en [2]. En [2] (ver Corollary 2.2) fue demostrada (con métodos completamente distintos) la propiedad de optimalidad de las SPRT extendidas, misma que se afirma en la primer parte del Teorema 3.8. La última parte de nuestro Teorema 3.8 complementa esta propiedad de optimalidad en el sentido de que muestra que si una prueba tiene las mismas probabilidades de error que una SPRT extendida, y el mismo promedio de observaciones (bajo* una *de las dos hipótesis), entonces es una SPRT extendida también, con las mismas constantes A y B.*

Demostración. Por el Teorema 3.8. Sea φ^* tal que se cumplen (3.87) μ^n-casi dondequiera en $T_n^{\varphi^*} \bigcap \{\{f_{\theta_0}^n > 0\} \bigcup \{f_{\theta_1}^n > 0\}\}$ y (3.88) μ^n-casi dondequiera en $S_n^{\varphi^*} \bigcap \{\{f_{\theta_0}^n > 0\} \bigcup \{f_{\theta_1}^n > 0\}\}$, para cualquier $n \geq 1$. Entonces para cualquier prueba (φ, δ) que satisface (3.89) la primera de las desigualdades en (3.90) está demostrada en Teorema 3.4.

Demostremos la segunda desigualdad en (3.90).

Consideremos el problema de prueba de hipótesis $H_0' : \theta = \theta_1$ contra la alternativa $H_1' : \theta = \theta_0$ (estamos intercambiando las hipótesis H_0 y H_1 que teníamos anteriormente). Es fácil ver que cualquier prueba (φ, δ) para el par de hipótesis H_0 y H_1 puede ser considerada como una prueba para el par de hipótesis H_0' y H_1' si aplicamos la regla de decisión terminal $1 - \delta$ cuyas componentes son $1 - \delta_n$, $n = 1, 2, \ldots$ (rechazar la H_0 es aceptar H_0', aceptar H_0 es rechazar H_0'). Además, si $\alpha'(\varphi, 1 - \delta)$ y $\beta'(\varphi, 1 - \delta)$ son las probabilidades de error tipo uno y dos, respectivamente, para las hipótesis H_0' y H_1', entonces

$$\alpha'(\varphi, 1 - \delta) = \beta(\varphi, \delta), \quad \text{y} \quad \beta'(\varphi, 1 - \delta) = \alpha(\varphi, \delta). \tag{3.91}$$

Ahora sea (φ^*, δ^*) una prueba que satisface (3.87) μ^n-casi dondequiera en $T_n^{\varphi^*} \bigcap \{\{f_{\theta_0}^n > 0\} \bigcup \{f_{\theta_1}^n > 0\}\}$ y (3.88), μ^n-casi dondequiera en $S_n^{\varphi^*} \bigcap \{\{f_{\theta_0}^n > 0\} \bigcup \{f_{\theta_1}^n > 0\}\}$, para cualquier $n = 1, 2, \ldots$. La representamos como una prueba para las hipótesis H_0' y H_1' en la siguiente forma equivalente. Es fácil ver que

$$I_{\{Z_n^{-1} \notin [B^{-1}, A^{-1}]\}} \leq \varphi_n^* \leq I_{\{Z_n^{-1} \notin (B^{-1}, A^{-1})\}} \tag{3.92}$$

(suponiendo que $1/0 = \infty$ y $1/\infty = 0$) μ^n-casi dondequiera en $T_n^{\varphi^*} \bigcap \{f_{\theta_1}^n > 0\}$, $n = 1, 2, \ldots$, y

$$1 - \delta_n^* = I_{\{Z_n^{-1} \geq A^{-1}\}} \tag{3.93}$$

μ^n-casi dondequiera en $S_n^{\varphi^*} \bigcap \{f_{\theta_1}^n > 0\}$, $n = 1, 2, \ldots$. Sea ahora (φ, δ) cualquier prueba tal que

$$\alpha(\varphi, \delta) \leq \alpha(\varphi^*, \delta^*), \quad \text{y} \quad \beta(\varphi, \delta) \leq \beta(\varphi^*, \delta^*).$$

De (3.91) se sigue que

$$\alpha'(\varphi, 1-\delta) \leq \alpha'(\varphi^*, 1-\delta^*), \quad \text{y} \quad \beta'(\varphi, 1-\delta) \leq \beta'(\varphi^*, 1-\delta^*).$$

Por lo tanto, del Teorema 3.4 (aplicado a las hipótesis H_0' y H_1') debido a (3.92) y (3.93) se concluye que

$$N(\theta_1, \varphi) \geq N(\theta_1, \varphi^*),$$

es decir, se cumple la segunda de las igualdades (3.90). De la misma manera se obtiene que esta desigualdad es estricta si alguna de las desigualdades (3.89) lo es.

Ahora demostramos la última parte del Teorema. Supongamos que existe otra prueba (φ, δ) para la cual las desigualdades en (3.89) son igualdades, y una de las desigualdades en (3.90) también es igualdad, digamos, la primera:

$$\alpha(\varphi, \delta) = \alpha(\varphi^*, \delta^*), \quad \beta(\varphi, \delta) = \beta(\varphi^*, \delta^*), \quad \text{y} \quad N(\theta_0, \varphi) = N(\theta_0, \varphi^*). \tag{3.94}$$

Primero aplicamos el Teorema 3.5 para encontrar c, π_0 y π_1 tales que A, B satisfacen

$$c + h(z; c, \pi_0.\pi_1) = g(z; \pi_0, \pi_1),$$

y por lo tanto la prueba (φ^*, δ^*) será óptima en el sentido del Teorema 3.3, es decir, para cualquier otra prueba (φ, δ) se tiene

$$cN(\theta_0, \varphi) + \pi_0\alpha(\varphi, \delta) + \pi_1\beta(\varphi, \delta) \geq cN(\theta_0, \varphi^*) + \pi_0\alpha(\varphi^*, \delta^*) + \pi_1\beta(\varphi^*, \delta^*).$$

Ahora de (3.94) se tiene que

$$cN(\theta_0, \varphi) + \pi_0\alpha(\varphi, \delta) + \pi_1\beta(\varphi, \delta) = cN(\theta_0, \varphi^*) + \pi_0\alpha(\varphi^*, \delta^*) + \pi_1\beta(\varphi^*, \delta^*),$$

y el Teorema 3.3 garantiza que para ϕ se cumple (3.87) μ^n-casi dondequiera en $T_n^{\varphi^*} \bigcap \{f_{\theta_0}^n > 0\}$ y para δ se cumple (3.88) μ^n-casi dondequiera en $S_n^{\varphi^*} \bigcap \{f_{\theta_0}^n > 0\}$, para culaquier $n = 1, 2, \ldots$. ■

□

Observación 3.1. *Únicamente el caso* $1 \in (A, B)$ *tiene sentido práctico, ya que en otro caso se puede lograr un mejor resultado sin necesidad de tomar observaciones.*

Demostración. Efectivamente, el hecho de que $1 \notin (A, B)$ por el Lema 3.7 significa que:

$$g(1) \leq c + \int f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) d\mu(x), \tag{3.95}$$

siendo el lado derecho de (3.95, el mínimo absoluto de la función $R(\varphi)$ sobre todos los tiempos de paro que definitivamente toman por lo menos una observación. Pero $g(1) = \text{mín}\{\pi_0, \pi_1\}$ y esta es por decirlo así, la función $R(\varphi_0)$ de la prueba que no toma ninguna observación debido a que el costo del experimento es nulo y aplicando la regla de decisión,

$$\delta_0 = I_{\{\pi_0 \leq \pi_1\}}, \tag{3.96}$$

y dado que $\tau_\varphi \equiv 0$ entonces $E_{\theta_0}\tau_\varphi = 0$ y por lo tanto $N(\theta_0, \varphi) = 0$, entonces:

$$c(N(\theta_0, \varphi)) + \pi_0 P(\theta_0; \varphi, \delta) + \pi_1(1 - P(\theta_0, \varphi, \delta)) = \text{mín}\{\pi_0, \pi_1\} = g(1). \tag{3.97}$$

$$\alpha(\varphi_0, \delta_0) = \begin{cases} 1 & \text{si} \quad \pi_0 \leq \pi_1 \\ 0 & \text{si} \quad \pi_0 > \pi_1 \end{cases}$$

y

$$\beta(\varphi_0, \delta_0) = \begin{cases} 0 & \text{si} \quad \pi_0 \leq \pi_1 \\ 1 & \text{si} \quad \pi_0 > \pi_1 \end{cases}$$

□

Vemos entonces que si se cumple (3.95), entonces hay una prueba trivial (φ_0, δ_0) que da un mejor resultado que la prueba que toma por lo menos una observación. Así que sólamente en el caso $A < 1 < B$ tiene sentido práctico la prueba secuencial del Teorema 3.7.

El caso límite $1 = A < B$ o $A < B = 1$, puede ser incluido en consideración por lo menos para fines teóricos ya que en este caso la prueba trivial da exactamente lo mismo que la SPRT, no algo menor.

CONCLUSIONES

El resultado principal de la tesis es la obtención de la estructura suficiente y necesaria para optimalidad de pruebas secuenciales para dos hipótesis simples sobre distribución de un proceso estocástico a tiempo discreto (Corolario 2.2 del Capítulo 2). Este se puede aplicar en problemas como los considerados en [10], [14], [4], además del problema clásico para observaciones independientes e identicamente distribuidas (ver [17], [11]).

Como caso particular de lo anterior, se estudió detalladamente el caso de observaciones independientes e identicamente distribuidas (capítulo 3), también llegando a la estructura suficiente y necesaria para optimalidad de pruebas secuenciales, considerando el problema de minimación del número promedio de observaciones (bajo una de las dos hipótesis) dado que las probabilidades de error no exceden algunos dos niveles fijos. Se demostró que las pruebas conocidas como las *SPRT extendidas* (en términos de [2]) poseen la misma propiedad de optimalidad que la SPRT clásica de Wald ([17]), es decir, minimizan el número promedio de observaciones bajo *cada una* de las dos hipótesis, dadas restricciones sobre las probabilidades de error. Aunque este resultado no es completamente nuevo (ver [2]), éste se obtuvo aquí usando otro método más general, y junto con una demostración nueva (y completa) del resultado clásico de Wald y Wolfowitz [17] (ver Teorema 3.8 en el Capítulo 3). Más que esto, demostramos (ver el Teorema 3.5) que el método de Lagrange (que usamos nosotros, igual que Lorden en [11]) es esencialmente la única vía para obtención de pruebas óptimas en el problema con restricciones. Por lo tanto, todos los casos importantes del problema con restricciones quedan cubiertos por nuestro Teorema 3.8, mismo que da la estructura suficiente y necesaria para optimalidad de pruebas en problema con restricciones.

BIBLIOGRAFÍA

[1] Borovkov, A.A. *Estadística matemática.* Ed. Mir. Moscú, 1998.

[2] D. L. Burkholder; R. A. Wijsman, Optimum Properties and Admissibility of Sequential Tests. *Ann. Math. Statistics.* 34 (Mar.,1963), pp. 1-17.

[3] Chow, Y.S., Robbins, H.,Siegmund, D. *Great Expectations: The Theory of Optimal Stopping.* Boston. Houghton Mifflin Company, 1971.

[4] Cochlar, J. and Vrana, I. (1978). On the Optimum Sequential Test of Two Hypotheses for Statistically Dependent Observations, *Kybernetika* 14: 57-69.

[5] DeGroot, M.H. *Optimal Statistics Decisions.* McGraw Hill, New York, 1970.

[6] Ferguson, T. S. *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York, 1967.

[7] Ghosh B. K., *Sequential test of statistical hypotheses.* Addison-Wesley, Reading, Mass-London-Don Mills, Ont.,1970.

[8] Govindarajulu, Z. *Sequential Statistics.* World Scientific Publishing, Singapore, 2004.

[9] Lehmann, E. *Testing Statistical Hypoteses.* John Wiley Sons, New York; Chapman Hall, Ltd., London 1959.

[10] Liu, Y. and Blostein, D. (1992). Optimality of the Sequential Probability Ratio Test for Nonstationary Observations, *IEEE Transactions on Information Theory* 28: 177-182.

[11] Lorden, G. Structure of sequential tests minimizing an expected sample size, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **51** (1980), 291-302.

[12] Rockafellar, R.T. *Convex Analysis.* Princeton University Press, New Jersey; 1972.

[13] Schmitz, N. (1995). Comment on: "Optimality of the Sequential Probability Ratio Test for Nonstationary Observations"by Y. Liu and S.D. Blostein, *IEEE Trans. Inform. Theory* 41: 861.

[14] Schmitz, N. and Süselbeck, B. (1983). Sequential Probability Ratio Tests for Homogeneous Markov Chains, *Lecture Notes in Statistics* 20: 181-202.

[15] Shiryaev, A. N. *Statistical Sequential Analysis:Optimal Stopping Rules.* Moscow, Nauka, 1969 (in Russian).

[16] Spivak, M. *Calculo Infinitesimal.* Reverte, Barcelona. 1988.

[17] Wald, A., Wolfowitz, J. Optimum character of the sequential probability ratio test. *Ann. Math. Statistics.* 19 (1948), 326–339.

[18] Zacks, S. *The Theory of Statistical Inference.* Wiley Series in probability and Mathematical Statistics, New York:Wiley, 1971. .

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE EXAMEN DE GRADO

No. 00079
Matrícula: 205384168

OPTIMALIDAD DE LA PRUEBA SECUENCIAL DE RAZON DE PROBABILIDADES.

En México, D.F., se presentaron a las 15:00 horas del día 26 del mes de julio del año 2012 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. JUAN GONZALEZ HERNANDEZ
DR. ANDREY NOVIKOV
DR. JOSE RAUL MONTES DE OCA MACHORRO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron para proceder al Examen de Grado cuya denominación aparece al margen, para la obtención del grado de:

MAESTRO EN CIENCIAS (MATEMÁTICAS)

DE: EFREN FRANCISCO PEREZ

EFREN FRANCISCO PEREZ
ALUMNO

y de acuerdo con el artículo 78 fracción III del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

Aprobar

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

REVISÓ

LIC. JULIO CESAR DE LARA ISASSI
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI

DR. JOSE ANTONIO DE LOS REYES HEREDIA

PRESIDENTE

DR. JUAN GONZALEZ HERNANDEZ

VOCAL

DR. ANDREY NOVIKOV

SECRETARIO

DR. JOSE RAUL MONTES DE OCA MACHORRO