



UNIVERSIDAD AUTÓNOMA METROPOLITANA

Unidad Iztapalapa

División de Ciencias Básicas e Ingeniería

**Pruebas de hipótesis
secuenciales óptimas con
grupos de tamaño aleatorio**

Tesis que presenta
XÓCHITL ITXEL POPOCA JIMÉNEZ

Para obtener el grado de
Doctora en Ciencias
Matemáticas

Asesor:
Dr. Andrey Novikov

Jurado calificador:

Presidente:	Dr. José Andrés Christen Gracia,	CIMAT
Secretaría:	Dra. Blanca Rosa Pérez Salvador,	UAM-I
Vocal:	Dr. José Raúl Montes de Oca Machorro,	UAM-I
Vocal:	Dr. Hugo Adán Cruz Suárez,	BUAP
Vocal:	Dr. Andrey Novikov,	UAM-I

Ciudad de México

01 de febrero 2023



Agradecimientos

Antes que nada agradezco enormemente a las personas trabajadoras del pueblo de México, porque debido a su esfuerzo recibí por parte del Consejo Nacional de Ciencia y Tecnología (CONACYT) una beca para la realización de la presente investigación. También por dicho esfuerzo y el de muchos investigadores existe la Universidad Autónoma Metropolitana-Iztapalapa (UAM-I), la cual apoyó mis estudios de doctorado al aceptarme como alumna y, en su momento como profesora.

A los profesores del departamento de matemáticas de la UAM-I que creyeron en mí, por su confianza, conocimiento, enseñanza, aceptación, apoyo y orientación en este camino. En particular, muchas gracias a los doctores: Andrey Novikov y Julio César García Corte, porque fueron pilares para mí. A los doctores Elsa Omaña, Rogelio Fernández, Raúl Montes de Oca y Blanca Rosa Pérez por su apoyo y por confiar en mi trabajo.

A mi padre *Leonardo Daniel Popoca T.* por sus consejos, apoyo y cariño. A mi hermano Ismael Popoca J. por enseñarme el gusto a la matemática, y a mis hermanas Verónica Popoca y Mariana Popoca por su enorme cariño, apoyo e impulso.

A mis verdaderos amigos, porque no dudaron de mí, por alentarme cuando estaba en el ocaso, por su cariño y amor, para: Soledad, José Luis, Teresita, Leonardo, Carlos, Hugo Pérez, Hugo Váldez, Alfonso, Jaime y Gi.

A todos GRACIAS.

ÍNDICE GENERAL

1. Análisis secuencial con grupos de tamaño aleatorio	1
1.1. Introducción.	1
1.2. Descripción del modelo y antecedentes	3
1.2.1. Antecedentes	11
1.3. Problemas de investigación.	13
2. Optimalidad de pruebas secuenciales con grupos aleatorios	17
2.1. Prueba secuencial óptima con grupos de tamaño aleatorio	17
2.1.1. Regla de decisión óptima	21
2.1.2. Regla de paro óptima: caso truncado.	27
2.1.3. Regla de paro óptima: caso no truncado	47
2.2. Problema Kiefer-Weiss con grupos de tamaño aleatorio	67
2.2.1. Kiefer-Weiss con grupos aleatorios aplicado a la familia exponencial	69
3. Prueba secuencial óptima con grupos aleatorios y observaciones independientes	75
3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades	75
3.1.1. Regla de decisión óptima bajo H_0	78
3.1.2. Regla paro óptima bajo H_0 : horizonte finito	81
3.1.3. Regla de paro óptima bajo H_0 : horizonte infinito	88
3.1.4. Región de continuación en términos de la razón de probabilidades	99
3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)	114
4. Conclusiones y Perspectivas	127
Bibliografía.	131

RESUMEN

El presente trabajo da respuesta a una pregunta abierta sobre la optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT, siglas en inglés) introducida en la investigación de [Mukhopadhyay-De Silva, (2008)]; además se exponen algunas generalizaciones de problemas clásicos desarrollados en el análisis secuencial (como el problema Kiefer-Weiss aplicado al nuevo contexto).

Primeramente, para el contexto de la estadística secuencial se demuestra el método generalizado de los multiplicadores de Lagrange, el cual es la base de la metodología de esta investigación. Segundo, con base en un experimento estadístico secuencial y dada una prueba de hipótesis simple con H_0 versus H_1 se deducen las estructuras de las pruebas secuenciales que minimizan los costos totales promedio del experimento. Se exponen condiciones necesarias y suficientes para que dichas estructuras sean óptimas, tanto para pruebas truncadas como para pruebas con horizonte infinito. Finalmente, se consideran casos cuando las observaciones (variables aleatorias) que se presentan en el experimento vienen en grupos de tamaño aleatorio (segundas variables aleatorias) y donde las observaciones pueden no ser, necesariamente, independientes, y a su vez, pueden tener o no la misma función de probabilidad, aún cuando los tamaños de grupos son independientes entre ellos y de las observaciones, y con iguales o diferentes distribuciones de probabilidad para cada grupo.

CAPÍTULO 1

ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

1.1. Introducción.

En contraste con los métodos clásicos de la estadística matemática, en los cuales el número de observaciones en la muestra está fijo; el *análisis secuencial* (un campo también de la estadística), trabaja, dado un experimento estadístico, con métodos que permiten la recolección de datos por etapas, típicamente de uno en uno. Otra característica de estos métodos estriba en que el tiempo del experimento para el cual se determina el número de observaciones (*tiempo de paro*) es aleatorio (véanse [Novikov, (2007)], [Novikov, (2008)], [Novikov, (2009)], [Wald, (1945)], [Weiss, (1962)], etc.).

Existen relatos que marcan que quizás “*el enfoque secuencial ha sido una forma natural de proceder a lo largo de la historia de la experimentación. Quizás el primer propónete fue Noé, quien en sucesivos días soltó una paloma del Arca para probar la presencia de tierra seca durante el hundimiento del Diluvio*”¹ ([Jennison-Turnbull, (1999)], p. 22). La aplicación formal de los procedimientos secuenciales comenzó a fines de la década de 1920 en el área de control estadístico de la calidad en la producción manufacturera. En donde se definieron planes de muestreo de aceptación en dos etapas para los componentes que podrían ser probados y clasificados como efectivos o defectuosos.

La teoría moderna formalizada del análisis secuencial surgió en respuesta a las demandas de toma de muestras más eficientes en procesos de inspección durante la segunda guerra mundial. Los primeros desarrollos del nuevo campo fueron introducidos en 1947 por Abraham Wald²

¹Traducción propia.

²(1902-1950)

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

en Estados Unidos de Norteamérica y por George Barnard en 1946 en Gran Bretaña, quienes participaban en grupos de asesores industriales para la producción bélica. Wald mostró que el utilizar métodos secuenciales en problemas de pruebas de dos hipótesis simples (con observaciones independientes) conlleva a un número de observaciones más pequeño, en promedio, que aquellos métodos que utilizan una muestra de tamaño fijo. De hecho en 1948 Wald y Wolfowitz demostraron teóricamente la propiedad de optimalidad de la prueba secuencial de la razón de probabilidades (SPRT ³, siglas en inglés) [Wald-Wolfowitz, (1948)].

En consecuencia, una ventaja de utilizar procedimientos secuenciales es que, en promedio, estos requieren de un número menor de observaciones con la misma confiabilidad que en el procedimiento con muestra fija (no secuencial) [Lehmann, (1959)]. Lo cual conlleva aplicaciones con menores tamaños de muestra, menores costos y/o menor tiempo. Entre algunas de las aplicaciones se encuentran: análisis clínicos, control de calidad, técnicas de confiabilidad, economía, finanzas, psicología, psicología social, análisis de supervivencia (bioestadística), entre otras ([Bhattachajee-Mukhopadhyay, (2012)], [Denne-Jennison, (2000)], [Ghosh, *et al* (1997)], [Weiss, (1962)]). Con los procedimientos secuenciales se establece una herramienta clave en la construcción de pruebas de hipótesis secuenciales simples y forma la base de métodos diseñados para manejar tamaños de grupos impredecibles. Ahora debido también al avance computacional actual, muchos de estos métodos pueden ser implementados de manera numérica a través de un procedimiento computable.

En la actualidad en algunas aplicaciones se está incrementando el uso de procedimientos secuenciales por grupos donde cada agrupación pueda ser de cualquier tamaño, generalmente dichos tamaños son diferentes de etapa en etapa más no se consideran necesariamente de tamaño aleatorio. Aplicaciones como: procesos bioquímicos para nuevos tratamientos y en los análisis clínicos de la industria farmacéutica (véanse, Bartroff, *et al* (2013), Jennison & Turnbull, (1999)), en donde debido a las condiciones de los grupos de experimentación y control, el tamaño de muestra va cambiando en cada etapa, aunque es predeterminado el tamaño para cada grupo, por las diferentes condiciones (o poca homogeneidad) de las unidades experimentales. Entonces se está recurriendo al análisis secuencial mediante grupos tamaño fijo pero diferente, y por fases, donde primero se inicia con animales y después con seres humanos (para más detalle sobre el proceso véase *Ídem*). Por tanto, el tener grupos de tamaño aleatorio, podría en principio dar una ventaja fuerte a dichos experimentos, y más si se tuviera un procedimiento óptimo que conlleve un costo total promedio mínimo del mismo, y a su vez un tiempo de paro y observaciones menores. Actualmente, debido a la crisis sanitaria que se generó en los años pasados por la pandemia, el estudio óptimo de tratamientos farmacéuticos es de suma importancia para la investigación. Mas ésta no sería la única aplicación en la cual se pudiera trabajar de manera secuencial y con grupos de tamaño aleatorio.

³Sequential Probability Ratio Test

1.2. Descripción del modelo y antecedentes

Sea un proceso estocástico a tiempo discreto

$$\{X_i : i \in \mathbb{N}\}$$

cuya distribución es P , y donde las variables aleatorias (v.a.) X_i toman sus valores en un espacio medible $(\mathcal{X}, \mathfrak{X})$ ⁴ cualquiera.

Supongamos que se tiene un experimento estadístico secuencial (en grupos) con un problema de prueba de hipótesis simples $H_0 : P = P_0$ vs. $H_1 : P = P_1$ y $P_0 \neq P_1$.

En este contexto se establece y denota **un problema de prueba de hipótesis simples** mediante la terna:

$$\mathcal{M} = (P_0, P_1, P). \quad (1.1)$$

Esta terna será fija durante todo el desarrollo, a menos que se indique lo contrario. La mayoría de los elementos trabajados en esta investigación estarán en dependencia de este problema de prueba \mathcal{M} , y tendrán marcada dicha relación. Las primeras dos distribuciones P_0 y P_1 hacen referencia a las hipótesis a contrastar H_0 y H_1 , respectivamente; y P es una tercera distribución correspondiente al proceso estocástico, bajo la cual se determinará la función a optimizar en esta investigación.

Típicamente, las tres distribuciones en \mathcal{M} pertenecen a una familia paramétrica específica (binomial, normal, exponencial, etc.); sin embargo, en este contexto esto no necesariamente se aplica. Nosotros, en principio, tomaremos a P o bien distinta de las primeras dos distribuciones o inclusive podría suceder que fuese igual a alguna de ellas o bien una combinación convexa de ellas, de la forma: $P = \pi P_0 + (1 - \pi)P_1$, con $\pi \in (0, 1)$. Pero, en general, P es cualquier distribución. Lo único que sí es necesario es que el problema de prueba tratado y sus elementos sean fijos durante todo el proceso. En este capítulo y el siguiente P será cualquiera, pero fija, a menos que se especifique otra cosa. En caso de variar algún elemento de la terna, se tendrá por tanto otro problema de prueba, y por ende otra notación.

Para cada $n \geq 1$ las v.a. tienen su función de “densidad” de probabilidad conjunta

$$f^n = f^n(x_1, \dots, x_n),$$

(en el sentido de la derivada de Radon-Nikodym de su distribución) con respecto a la medida producto

$$\mu^n = \underbrace{\mu \otimes \dots \otimes \mu}_{n\text{-veces}},$$

donde μ es una medida σ -finita sobre \mathcal{X} .

⁴ \mathcal{X} es cualquier espacio y \mathfrak{X} es la σ -álgebra de subconjuntos medibles en dicho espacio.

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

En el experimento estadístico secuencial por grupos se tienen en cuenta, al menos, dos tipos diferentes de sucesiones de variables aleatorias; las que representan a los tamaños de los grupos y las correspondientes a las observaciones de cada grupo.

Tomemos un subconjunto $\mathcal{G} \subseteq \{0, 1, 2, \dots\}$ y sea $\nu_1, \dots, \nu_k \dots$ una sucesión de v.a. no negativas, discretas e independientes (no necesariamente idénticamente distribuídas (i.d.)), donde dichas v.a. toman sus valores en \mathcal{G} . Los valores de éstas representan los tamaños de grupos consecutivos y se considera que pueden existir grupos de tamaño nulo durante el experimento. La función de probabilidad del tamaño del i -ésimo grupo es $P(\nu_i = n) = p_i(n) \geq 0$ con $\sum_{n \in \mathcal{G}} p_i(n) = 1$ para cada $i \geq 1$. Además, se toma en cuenta que los tamaños de los grupos son independientes de las observaciones.

Dado un vector aleatorio $\boldsymbol{\nu} = (\nu_1, \dots, \nu_k)$, para cualquier $k \geq 1$, una realización de éste es el vector $\boldsymbol{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$, donde las k componentes son los respectivos tamaños de k grupos consecutivos. Entonces si $\boldsymbol{n} \in \mathcal{G}^k$, se hablará de **los tamaños de k grupos consecutivos**. El total de observaciones en dichos k grupos consecutivos es $|\boldsymbol{n}| = n_1 + n_2 + \dots + n_k$. Sea $1 \leq s < k$ se define $(\boldsymbol{n} : s) = (n_1, n_2, \dots, n_s)$ el subvector de los primeros s elementos del vector \boldsymbol{n} .

Dados k tamaños de grupos consecutivos representados en $\boldsymbol{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$, se definen los espacios de probabilidad para las observaciones en dichos grupos. Debido a que las observaciones vienen en grupos y pueden existir grupos de tamaño nulo ($n = 0$), entonces para estos es importante señalar sus correspondientes “observaciones”, las cuales serán representadas mediante la notación “ $()$ ”.

1. Sea el espacio $\mathcal{X}^0 := \{()\}$ con su σ -álgebra trivial $\mathfrak{X}_0 := \{\emptyset, \mathcal{X}^0\}$.

Este espacio medible será utilizado para **grupos sin observaciones** que provienen de tamaños nulos.

Sobre este espacio también se considera la medida de probabilidad μ^0 con respecto a la σ -álgebra trivial \mathfrak{X}_0 .

2. Para cualquier $n \geq 2$ se denota el espacio $\mathcal{X}^n := \underbrace{\mathcal{X} \times \dots \times \mathcal{X}}_{n\text{-veces}}$, con su respectiva σ -álgebra producto (generada por el producto de subconjuntos medibles).

3. Dado el vector $\boldsymbol{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$ con $k \geq 2$ se denota el espacio

$$\mathcal{X}^{\boldsymbol{n}} := \mathcal{X}^{n_1} \times \mathcal{X}^{n_2} \times \dots \times \mathcal{X}^{n_k}.$$

Este espacio tiene su correspondiente σ -álgebra producto.

4. Dado $\boldsymbol{n} \in \mathcal{G}^k$, sobre el espacio $\mathcal{X}^{\boldsymbol{n}}$ se define la medida producto:

$$\mu^{\boldsymbol{n}} := \mu^{n_1} \otimes \mu^{n_2} \otimes \dots \otimes \mu^{n_k}$$

donde cada $\mu^{n_i} := \underbrace{\mu \otimes \cdots \otimes \mu}_{n_i\text{-veces}}$ para $n_i \geq 2$ e $i = 1, \dots, k$. Para $n_i = 0, 1$, μ^0 es medida de probabilidad trivial y μ es medida σ -finita sobre \mathcal{X} .

Dado un vector aleatorio $\nu = (\nu_1, \dots, \nu_k)$, para cada $k \geq 1$, al proceso estocástico se le puede ver, también, como un proceso seccionado mediante vectores aleatorios de tamaño aleatorio (ν_i) , $X_1^\nu, X_2^\nu, \dots, X_k^\nu$, donde cada X_i^ν es un vector aleatorio con las v.a. del i -ésimo grupo para $1 \leq i \leq k$:

$$\left(\underbrace{(X_1, \dots, X_{\nu_1})}_{X_1^\nu}, \underbrace{(X_{\nu_1+1}, \dots, X_{\nu_1+\nu_2})}_{X_2^\nu}, \underbrace{(X_{\nu_1+\nu_2+1}, \dots, X_{\nu_1+\nu_2+\nu_3})}_{X_3^\nu}, \dots, \underbrace{(X_{|\nu:k-1|+1}, \dots, X_{|\nu|})}_{X_k^\nu} \right)$$

Entonces dado el vector $\mathbf{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$ de tamaños de k grupos consecutivos, se representa al vector de sus observaciones:

$$x^\mathbf{n} = \left(\underbrace{(x_1, \dots, x_{n_1})}_{x_1^\mathbf{n}}, \underbrace{(x_{n_1+1}, \dots, x_{n_1+n_2})}_{x_2^\mathbf{n}}, \underbrace{(x_{n_1+n_2+1}, \dots, x_{n_1+n_2+n_3})}_{x_3^\mathbf{n}}, \dots, \underbrace{(x_{|\mathbf{n}:k-1|+1}, \dots, x_{|\mathbf{n}|})}_{x_k^\mathbf{n}} \right).$$

$$x^\mathbf{n} := (x_1^\mathbf{n}, \dots, x_k^\mathbf{n}) \in \mathcal{X}^\mathbf{n}.$$

Cada vector $x_i^\mathbf{n}$ representará el vector de observaciones del i -ésimo grupo. Obsérvese que cada grupo tiene la dependencia del vector \mathbf{n} porque éste contiene los respectivos tamaños de los grupos anteriores, y a su vez se marca la “dependencia” entre las observaciones.

En el caso particular de que alguna(s) v.a. ν_i sea(n) degenerada(s) nula(s), por ejemplo supongamos que la primer y la tercer v.a. son degeneradas nulas ($P(\nu_1 = 0) = P(\nu_3 = 0) = 1$), entonces

$$x^\mathbf{n} = \left(\underbrace{()}_{x_1^\mathbf{n}}, \underbrace{(x_1, \dots, x_{n_2})}_{x_2^\mathbf{n}}, \underbrace{()}_{x_3^\mathbf{n}}, \underbrace{(x_{n_2+1}, \dots, x_{n_2+n_4})}_{x_4^\mathbf{n}}, \dots, \underbrace{(x_{|\mathbf{n}:k-1|+1}, \dots, x_{|\mathbf{n}|})}_{x_k^\mathbf{n}} \right).$$

O simplemente, si el tamaño de algún grupo es nulo ($n_i = 0$), por ejemplo, sean 6 grupos consecutivos de observaciones con sus respectivos tamaños, $\mathbf{n} = (0, 3, 0, 2, 0, 4) \in \mathcal{G}^6$, entonces su correspondiente vector de observaciones será:

$$x^\mathbf{n} = \left(\underbrace{()}_{x_1^\mathbf{n}}, \underbrace{(x_1, x_2, x_3)}_{x_2^\mathbf{n}}, \underbrace{()}_{x_3^\mathbf{n}}, \underbrace{(x_4, x_5)}_{x_4^\mathbf{n}}, \underbrace{()}_{x_5^\mathbf{n}}, \underbrace{(x_6, x_7, x_8, x_9)}_{x_6^\mathbf{n}} \right) \in \mathcal{X}^\mathbf{n}.$$

Dada la función de densidad conjunta $f^\mathbf{n}$ con $n \in \mathbb{N}$ y dado el vector $\mathbf{n} \in \mathcal{G}^k$ se define:

$$f^\mathbf{n}(x^\mathbf{n}) := \begin{cases} f^{|\mathbf{n}|}(x_1, \dots, x_{n_1}, x_{n_1+1}, \dots, x_{|\mathbf{n}|}), & \text{si } |\mathbf{n}| \geq 1, \\ 1, & \text{si } |\mathbf{n}| = 0. \end{cases} \quad (1.2)$$

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

En particular, cuando se trabaje con la función de densidad f_0^n y/o f_1^n , se tendrá, respectivamente la distribución P_0 o P_1 . Supongamos, por ejemplo, que en el experimento secuencial solo tenemos un grupo de tamaño nulo ($n_1 = 0$), entonces el vector de observaciones será $x^n = ()$ con $\mathbf{n} \in \mathcal{G}$; nótese que al evaluar la correspondiente función de densidad en el grupo, ésta valdrá la unidad, i.e. $f^n(x^n) = f^n(()) = 1$.

Definición 1.

Una *regla de paro* ψ aleatorizada es una familia de funciones medibles $\{\psi_{\mathbf{n}}\}_k$ donde

$$\psi_{\mathbf{n}} : \mathcal{X}^{\mathbf{n}} \rightarrow [0, 1],$$

para cada $k \geq 1$ y cada $\mathbf{n} \in \mathcal{G}^k$.

Definición 2.

Una *regla de decisión* ϕ aleatorizada es una familia de funciones medibles $\{\phi_{\mathbf{n}}\}_k$ donde

$$\phi_{\mathbf{n}} : \mathcal{X}^{\mathbf{n}} \rightarrow [0, 1],$$

para cada $k \geq 1$ y cada $\mathbf{n} \in \mathcal{G}^k$.

Definición 3.

Se define una ***prueba estadística secuencial*** como una pareja (ψ, ϕ) donde ψ es una *regla de paro* y ϕ es una *regla de decisión*.

Dada una prueba estadística secuencial (ψ, ϕ) fija, el experimento comienza con el primer grupo de observaciones, dado $\mathbf{n} = (n_1) \in \mathcal{G}$ se tiene $x_1^n = (x_1, x_2, \dots, x_{n_1})$ (primera etapa, $k = 1$). A este grupo se le aplica la correspondiente función de paro $\psi_{\mathbf{n}}(x_1^n)$, y a su valor se le interpreta como *la probabilidad condicional de parar el experimento* dadas las observaciones en la primera etapa. Si el experimento se detiene en $k = 1$, se toma una decisión; si no, se *continúa* y se toma el siguiente grupo de observaciones, de lo cual se genera $\mathbf{n} = (n_1, n_2) \in \mathcal{G}^2$ y por lo tanto $x^n = (x_1^n, x_2^n)(k = 2)$, donde $x_2^n = (x_{n_1+1}, x_{n_1+2}, \dots, x_{n_1+n_2})$ y nuevamente se analizará el valor $\psi_{\mathbf{n}}(x_1^n, x_2^n)$.

En general, al valor de $\psi_{\mathbf{n}}(x_1^n, \dots, x_k^n)$ se le interpreta como ***la probabilidad condicional de parar el experimento dadas las observaciones*** en x^n , para cualquier etapa $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$. En el caso de no parar se continúa a la siguiente etapa ($k + 1$) y se añade al análisis el siguiente grupo de observaciones. Entonces se aplica la componente de la regla de paro $\psi_{\mathbf{n}}$ con $\mathbf{n} \in \mathcal{G}^{k+1}$

a los $k + 1$ grupos de la misma manera descrita anteriormente, hasta que el experimento eventualmente tenga un paro.

Si existe alguna etapa de paro $k \geq 1$ ⁵, entonces ahora corresponde preguntarse sobre la decisión de *aceptar o rechazar* la hipótesis nula H_0 . Para lo cual se aplica la correspondiente función de decisión ϕ_n con $\mathbf{n} \in \mathcal{G}^k$ a los k grupos de observaciones obtenidos. El valor de la función $\phi_n(x_1^n, \dots, x_k^n)$ se interpreta como **la probabilidad condicional de rechazar la hipótesis** H_0 dadas las observaciones en x^n , dada la etapa de paro k y $\mathbf{n} \in \mathcal{G}^k$.

Convención: Cuando se trabaje con alguna función bajo el signo de probabilidad P o valor esperado E se sobreentenderá que se tienen v.a. X_1, \dots, X_n , por ejemplo $P(X_1, \dots, X_n)$. En otro caso se entenderá que se trabaja con los valores de dichas v.a. e.g. $F(x_1, \dots, x_n)$, $\psi(x^n)$ o como en la siguiente definición.

Definición 4.

Dada cualquier regla de paro ψ , para cualquier $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$ fijos se definen:

$$t_n^\psi := \begin{cases} 1, & \text{si } k = 1; \\ (1 - \psi_{(\mathbf{n}:1)})(1 - \psi_{(\mathbf{n}:2)}) \cdots (1 - \psi_{(\mathbf{n}:k-1)}), & \text{si } k > 1. \end{cases} \quad (1.3)$$

$$s_n^\psi := t_n^\psi \psi_n. \quad (1.4)$$

El valor de la v. a. t_n^ψ con $\mathbf{n} \in \mathcal{G}^k$ indicará **la probabilidad condicional de no parar el experimento antes de la etapa k** dadas las observaciones en x^n . O dicho de otra manera t_n^ψ con $\mathbf{n} \in \mathcal{G}^k$ será la indicadora de parar el experimento en la etapa k o después.

Por otro lado, el valor de s_n^ψ con $\mathbf{n} \in \mathcal{G}^k$ indicará **la probabilidad condicional de que la regla de paro ψ detenga el experimento en etapa k (y no antes)**, dadas las observaciones en x^n .

Es importante mencionar que la presente investigación siempre iniciará con el primer grupo de observaciones (o primera etapa del experimento, $k \geq 1$) independientemente del número de observaciones que en él haya (incluyendo tamaño nulo). Este hecho incluye aquellas investigaciones que analizan pruebas secuenciales que tienen etapa cero o etapa sin observaciones (e.g. véase [Wald, (1947)]).

⁵Con k grupos observados.

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

Ahora, con base en que los tamaños de grupos son v.a. independientes si $\mathbf{n} = (n_1, \dots, n_k)$ representa los respectivos k tamaños de grupos consecutivos, su probabilidad conjunta es

$$p_{\mathbf{n}} = \prod_{i=1}^k P(\nu_i = n_i) = \prod_{i=1}^k p_i(n_i), \quad (1.5)$$

donde el subíndice i hace referencia al número de distribución del correspondiente i -ésimo grupo, para cualquier $k \geq 1$. Recuerde que la distribución P de la fórmula (1.5) es de los tamaños de los grupos y no de las observaciones.

Sea $n_i \geq 0$ el tamaño del i -ésimo grupo de observaciones, con $i \geq 1$, el costo que se tiene que pagar por este grupo de datos será una función no negativa $c_i(n_i)$ que depende del número (n_i) de observaciones en ese grupo y del grupo mismo i . Este costo puede ser proporcional o no al número de observaciones, en general es cualquier función no negativa. Como casos particulares de funciones de costo se pueden tener: $c_i(n_i) = n_i$, ejemplo clásico en el campo del análisis secuencial (véase [Novikov, (2007)], [Novikov, (2008)] y [Wald-Wolfowitz, (1948)]), en donde el tamaño de cada grupo puede ser $n_i = 1$, $n_i = k$ (constante) o variable para cada grupo. Otro caso de una función de costo sería $c_i(n_i) = \kappa n_i + l$ donde $\kappa > 0$ es un costo unitario por unidad y l es un costo por lote. En general, el costo por grupo es cualquier función no negativa. Entonces para una realización de la sucesión de v.a. ν_1, ν_2, \dots (correspondientes a una sucesión de tamaños de grupos), se obtendrá otra sucesión de los costos de cada grupo con elementos no negativos $c_1(n_1), c_2(n_2), \dots, c_i(n_i), \dots$. Debido a que las v.a. ν_i son independientes entonces las variables $c_i(\nu_i)$ también tendrán las mismas propiedades. Para dichas v.a. tomaremos su valor promedio de la siguiente forma.

Para toda $i \geq 1$, el **costo promedio del i -ésimo grupo** se denota como:

$$\bar{c}_i = \sum_{n_i \in \mathcal{G}} c_i(n_i) p_i(n_i). \quad (1.6)$$

Se supondrá que para toda $i \geq 1$,

$$0 \leq \bar{c}_i < \infty. \quad (1.7)$$

Dado $\mathbf{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$ se define el **costo total de k etapas**:

$$c(\mathbf{n}) := c_1(n_1) + c_2(n_2) + \dots + c_k(n_k) = \sum_{i=1}^k c_i(n_i). \quad (1.8)$$

El costo total será una serie no decreciente, ya que si se aumenta el número de etapas y dado que $c_i(n_i)$ es no negativa, entonces el costo total de dicho experimento será mayor o igual que el costo de la etapa anterior.

Definición 5.

Dada una regla de paro ψ cualquiera, ésta genera la variable aleatoria τ^ψ llamada *tiempo de paro*, tal que $\tau^\psi \in \{1, 2, \dots\} \cup \{\infty\}$ y cuya distribución para $k \in \mathbb{N}$ está dada por:

$$P(\tau^\psi = k) = \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E s_{\mathbf{n}}^\psi. \quad (1.9)$$

Donde $E s_{\mathbf{n}}^\psi$ es el valor esperado (condicionado a los tamaños de grupos) de parar el experimento en la etapa k . Además la probabilidad de no parar antes de la etapa k se determina como:

$$P(\tau^\psi \geq k) = \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E t_{\mathbf{n}}^\psi. \quad (1.10)$$

Cabe mencionar que pueden existir reglas ψ que conlleven a un tiempo de paro infinito. Por ejemplo, la prueba secuencial (ψ, ϕ) cuya regla de paro tiene componentes $\psi_{\mathbf{n}} \equiv 0, \forall \mathbf{n}$, lo cual indica que en ninguna etapa habrá paro (i.e. el experimento continúa indefinidamente); y por tanto $\tau^\psi = \infty$. Si esto sucede, se pueden tener costos no finitos. Entonces y por razones prácticas para garantizar el término del experimento secuencial (en grupos) se trabajará con las reglas ψ que tengan un tiempo de paro finito con probabilidad uno.

Definición 6.

Se define la clase de reglas de paro que tienen un tiempo de paro finito con probabilidad uno, i. e.

$$\mathcal{F}_{\mathcal{M}} := \{\psi : P(\tau^\psi < \infty) = 1\}. \quad (1.11)$$

Nótese que las reglas $\psi \in \mathcal{F}_{\mathcal{M}}$ tienen la propiedad de que $P(\tau^\psi = \infty) = 0$.

Téngamos presente que la estructura de la clase $\mathcal{F}_{\mathcal{M}}$ está en dependencia del problema de prueba $\mathcal{M} = (P_0, P_1, P)$. Si la distribución P cambiara, fuese, por ejemplo, $P = P_0$ o $P = P_1$, entonces el problema de prueba de hipótesis simples se denotaría \mathcal{M}_0 o \mathcal{M}_1 , respectivamente, y a su vez la clase $\mathcal{F}_{\mathcal{M}}$ cambiaría de definición y notación a: \mathcal{F}_0 o \mathcal{F}_1 .

Sabemos que en las pruebas de hipótesis estadísticas existen dos tipos de errores que se pueden cometer. El error de rechazar la hipótesis H_0 siendo ésta verdadera (error tipo I) y el error de aceptar H_0 , siendo ella falsa (error tipo II). Este tipo de errores se pueden medir, hasta cierto punto, mediante la determinación de sus probabilidades. Entonces para una prueba estadística secuencial (ψ, ϕ) se define y determina la probabilidad de *error tipo I*:

$$\alpha(\psi, \phi) := P_0\left(\{\text{rechazar } H_0\} \cap \{\tau^\psi < \infty\}\right) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} (E_0 s_{\mathbf{n}}^\psi \phi_{\mathbf{n}}); \quad (1.12)$$

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

y la probabilidad de *error tipo II* como:

$$\beta(\psi, \phi) := P_1\left(\{\text{aceptar } H_0\} \cap \{\tau^\psi < \infty\}\right) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}}(E_1 s_{\mathbf{n}}^\psi (1 - \phi_{\mathbf{n}})). \quad (1.13)$$

Nótese que cada probabilidad de error está asociada, respectivamente, a la hipótesis que se rechaza. Estas probabilidades se “controlan” manteniéndolas debajo de ciertos niveles específicos $\alpha, \beta \in [0, 1)$, para ello supongamos que:

$$\alpha(\psi, \phi) \leq \alpha \quad \text{y} \quad \beta(\psi, \phi) \leq \beta. \quad (1.14)$$

En el análisis secuencial clásico (grupos de tamaño uno) para caracterizar el número de observaciones por analizar hasta obtener una decisión final, se trabaja con el tiempo promedio de paro, bajo la hipótesis H_0 o H_1 :

$$\mathcal{N}(i; \psi) = E_i \tau^\psi, \quad i = 0, 1. \quad (1.15)$$

En nuestro contexto, se trabajará con el costo promedio total del experimento, denotado como $K_{\mathcal{M}}(\psi)$. Este costo promedio es muy general y tiene como casos particulares a los costos promedios del experimento determinados bajo las distribuciones P_0 o P_1 , denotados $K_i(\psi)$, con $i = 0, 1$, respectivamente; y estos a su vez contienen a los tiempos promedios (1.15) trabajados en investigaciones anteriores. Esto es, nuestro costo promedio total del experimento recupera los problemas clásicos trabajados.

Ahora, si $\psi \in \mathcal{F}_{\mathcal{M}}$ entonces el **costo total promedio** del experimento se define:

$$K_{\mathcal{M}}(\psi) := E\{\text{Costo total}\} = \sum_{k=1}^{\infty} E\left((c_1(\nu_1) + c_2(\nu_2) + \cdots + c_k(\nu_k)) I_{\{\tau^\psi = k\}}\right).$$

Por tanto, si $\psi \in \mathcal{F}_{\mathcal{M}}$, el costo promedio del experimento será:

$$K_{\mathcal{M}}(\psi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^\psi. \quad (1.16)$$

Cabe mencionar que el hecho de que ψ tenga un tiempo de paro finito con probabilidad uno ($\psi \in \mathcal{F}_{\mathcal{M}}$) no garantiza que el costo promedio del experimento $K_{\mathcal{M}}(\psi)$ sea finito, de hecho se puede dar el caso en que sea infinito, aún cuando $\psi \in \mathcal{F}_{\mathcal{M}}$.

1.2.1. Antecedentes

En 1948 Wald y Wolfowitz [Wald-Wolfowitz, (1948)] bajo las hipótesis de:

- (a) una sola observación en cada etapa ($\nu_1 = 1, \nu_2 = 1, \dots$);
- (b) observaciones independientes e idénticamente distribuidas (X_i i.i.d.); y
- (c) el tiempo de paro promedio ⁶ de k etapas proporcional al número de observaciones (v. g. $E\tau^\psi = n$),

demonstraron que existe una prueba secuencial (llamada SPRT ⁷, por sus siglas en inglés) que conlleva a un tiempo promedio de paro mínimo tanto para H_0 ($E_0\tau^\psi$) como para H_1 , comparada con cualquier otra prueba con errores tipo I y II menores. Esto es la SPRT es una prueba óptima y termina el experimento en el momento en que la razón de verosimilitud $z_n \notin (A, B)$ para ciertas A y B constantes tales que $0 < A \leq 1 \leq B < \infty$. Aceptando la hipótesis nula H_0 , si $z_n < A$, o rechazándola, a favor de H_1 , si $z_n > B$.

Sin embargo, existen muchos casos prácticos para los cuales no se cumplen las hipótesis de la SPRT y se sabe que al variar estas hipótesis, la SPRT pierde su optimalidad. De hecho, para casos específicos con distribución normal o binomial y con $\alpha = \beta$, la SPRT en el punto $\frac{\theta_0 + \theta_1}{2}$ alcanza un tiempo de paro promedio mayor, comparada con otras pruebas (incluida la prueba Neyman-Pearson). El mismo Wald hizo variar la condiciones iniciales para la SPRT, como por ejemplo el truncar el procedimiento para cierto tamaño de muestra, y para el cual mucho después se determinó a través de integración numérica, la probabilidad de aceptar H_0 en función de la distribución con parámetro θ .

A partir de la década de los 50's del siglo pasado, se publicaron una serie de artículos que expresaban algún tipo de variante de la SPRT.

En 1953 Weiss definió una clase de pruebas llamadas GSPRT (Generalized Sequential Probability Ratio Test). Para las cuales no se utilizan, necesariamente, las mismas constantes A y B en todo el muestreo, sino que ellas variaban en dependencia de cada etapa, teniendo así para cada período el análisis de $A_n < z_n < B_n$, con A_n, B_n constantes predeterminadas que dependen del tamaño de muestra (acumulado); bajo hipótesis simples H_0 y H_1 . Con las GSPRT se tiene la posibilidad de tomar etapas no con un solo dato sino con observaciones en grupos cuyo tamaño está predeterminado [Weiss, (1953)]. Además, la clase de GSPRT tiene propiedades atractivas aún sin los supuestos sobre independencia o distribución idéntica de las observaciones [Eisenberg, *et al*, (1976)].

En el proceso de estudio de las propiedades de la GSPRT se planteó el problema conocido como Kiefer-Weiss (original) ([Kiefer-Weiss, (1957)]) que consistía en: minimizar $\sup_{\theta_0 \leq \theta \leq \theta_1} E_\theta \tau^\psi$, bajo ciertas condiciones, sobre la clase de todas las pruebas secuenciales; donde θ hace referencia

⁶O el costo promedio.

⁷Sequential Probability Ratio Test.

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

a una tercera distribución P_θ distinta de las distribuciones hipotetizadas P_{θ_0} o P_{θ_1} . La solución de este último problema no minimiza el promedio del tiempo de paro ni bajo θ_0 ni en θ_1 , sino que minimiza al máximo de $E_\theta \tau^\psi$ con $\theta_0 < \theta < \theta_1$ (véase por ejemplo [Weiss, (1962)]); dicha solución se dio mediante la resolución del problema “Kiefer-Weiss” modificado ([Lorden, (1980)], [Novikov, (2009)]).

En el 2008 Mukhopadhyay & De Silva [Mukhopadhyay-De Silva, (2008)] presentaron la *prueba secuencial aleatoria de la razón de probabilidades (RSPRT: Random Sequential Probability Ratio Test)*. La RSPRT fue aplicada a experimentos secuenciales con observaciones independientes y en pruebas de hipótesis simples H_0 y H_1 . Experimentos en los cuales: el tamaño de los grupos de observaciones es aleatorio, la distribución de probabilidad de dichos tamaños no es fija; y además se tienen constantes fijas A y B con $0 < A < 1 < B < \infty$ para todo el proceso. En dicha investigación (*ídem*) se dejaron abiertas las propiedades de optimalidad de la RSPRT. Sin embargo, en el 2012 se demostró (véase [Popoca, (2012)]) que para variables discretas (correspondientes a las observaciones) independientes, tamaños de grupos aleatorios, con distribución fija y costos de grupos generales, la RSPRT es óptima tanto para θ_0 como para θ_1 . Además de esta última investigación también se desprende, implícitamente, que la RSPRT no puede ser óptima para el caso en que los tamaños de grupos tienen diferente distribución y A, B son las mismas constantes fijas para todo el proceso. Se observa que para la optimalidad se necesita cierta modificación de la RSPRT en el sentido de variar las constantes en cada etapa, similar a la GSPRT.

Las pruebas secuenciales anteriormente descritas (SPRT, GSPRT, RSPRT), entre otras, han sido pruebas que el análisis secuencial ha investigado debido a su importancia en la aplicación de modelos ([Bhattachajee-Mukhopadhyay, (2012)], [Mukhopadhyay-De Silva, (2008)], [Weiss, (1962)], etc). La SPRT y las pruebas con procedimientos relacionados se ocupan principalmente de un problema de prueba de hipótesis simples. Sin embargo, el problema de elegir entre más de dos hipótesis es más difícil y ha llamado menos la atención. Por tanto son antecedentes de lo que se podría o no explorar en nuestro trabajo.

1.3. Problemas de investigación.

En esta sección se plantearán los problemas de investigación que se desarrollarán en este trabajo para lo cual definiremos el contexto.

Definición 7.

Sea $\mathfrak{S}_{\mathcal{M}}$ la clase de pruebas secuenciales tal que su regla de paro tiene un tiempo de paro finito con probabilidad uno. Esto es

$$\mathfrak{S}_{\mathcal{M}} := \{(\psi, \phi) : \psi \in \mathcal{F}_{\mathcal{M}}\}. \quad (1.17)$$

La clase $\mathfrak{S}_{\mathcal{M}}$ es no vacía, ya que por ejemplo se puede tomar una prueba secuencial trivial (ψ, ϕ) tal que detiene el experimento en la primera etapa. Esta prueba trivial tendrá un tiempo de paro finito ($\tau^\psi = 1$), por tanto $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$. O en general, dado un $N \geq 1$ entero, se puede definir la clase de pruebas secuenciales (denotada como \mathfrak{S}^N) cuyas reglas de paro detienen el experimento en la etapa N o antes ($\tau^\psi \leq N$), esta clase es una subclase de la definición anterior (1.17) (i.e. $\mathfrak{S}^N \subseteq \mathfrak{S}_{\mathcal{M}}$).

Dado un problema de pruebas de hipótesis simples $\mathcal{M} = (P_0, P_1, P)$, y dados un proceso estocástico a tiempo discreto $\{X_i\}_{i \in \mathbb{N}}$, con distribución P , y una sucesión ν_1, ν_2, \dots v.a. no negativas, discretas e independientes (representantes de tamaños de grupos), nuestro primer y principal problema de investigación consiste en minimizar el costo promedio total (1.16) del experimento, sobre todas las pruebas secuenciales $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ cuyas probabilidades de error estén por debajo de ciertas cotas fijas (α y β). En símbolos, el problema principal es:

$$\min_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}} K_{\mathcal{M}}(\psi) \quad \text{restringido a} \quad \alpha(\psi, \phi) \leq \alpha \quad \text{y} \quad \beta(\psi, \phi) \leq \beta. \quad (1.18)$$

donde $\alpha, \beta \in [0, 1)$ son constantes predeterminadas fijas; y la distribución P es cualquiera, pero fija también. Cuando la distribución P es diferente a las distribuciones P_0 o P_1 , este problema es conocido como el problema de Kiefer-Weiss modificado.

Al planteamiento dado en (1.18) le llamaremos el problema con restricciones.

El problema clásico del análisis secuencial (1.15), donde las observaciones son v.a. independientes, tienen la misma distribución $P = P_0$ o $P = P_1$, todos grupos son de tamaño uno ($\nu_i = 1$) para toda $i \geq 1$, y el costo de cada grupo es proporcional al número de observaciones en él ($c(n) = n$), resulta un caso particular de nuestro problema de investigación.

El proceso que se llevará a cabo para resolver el problema de optimización con restricciones planteado es utilizar el método variacional de Langrange, el cual permite reescribir el problema (1.18) como un problema sin restricciones de la forma:

$$L_{\mathcal{M}}(\psi, \phi; \lambda_0, \lambda_1) = K_{\mathcal{M}}(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi)$$

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

donde $\lambda_0 > 0, \lambda_1 > 0$ son los multiplicadores de Lagrange. Dichos multiplicadores son finitos y no pueden ser cero. Debido a que si, por ejemplo, $\lambda_0 = 0$, entonces el problema original no tendría la restricción sobre $\alpha(\psi, \phi)$; y el problema se reducirá a un problema con una sola restricción, análogamente sucedería con $\lambda_1 = 0$. Del hecho de que los multiplicadores son constantes finitas cualesquiera, pero fijas en el proceso entonces no consideraremos la dependencia de estas constantes en la función de Lagrange a minimizar, sin embargo, sí es importante marcar la dependencia con el problema de hipótesis en cuestión, por lo cual:

$$L_{\mathcal{M}}(\psi, \phi) := L_{\mathcal{M}}(\psi, \phi; \lambda_0, \lambda_1) = K_{\mathcal{M}}(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi). \quad (1.19)$$

Utilizando los resultados (1.12), (1.13) y (1.16) se determina:

$$L_{\mathcal{M}}(\psi, \phi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} \left(c(\mathbf{n}) f^{\mathbf{n}} + \phi_{\mathbf{n}} \lambda_0 f_0^{\mathbf{n}} + (1 - \phi_{\mathbf{n}}) \lambda_1 f_1^{\mathbf{n}} \right) d\mu^{\mathbf{n}}. \quad (1.20)$$

Cabe mencionar que el método de los multiplicadores de Lagrange ha sido utilizado frecuentemente en el contexto secuencial, de manera implícita, como por ejemplo en Kiefer y Weiss (1957), Weiss (1962), Lorden (1980), Schmitz (1993), entre otros. Incluso también en el contexto “clásico”, *e.g.* en la prueba del Lema Fundamental de Neyman-Pearson [Lehmann, (1959)]. En esta investigación daremos la demostración explícita (dentro de nuestro contexto secuencial) de cómo el método de Lagrange resuelve nuestro problema de minimización.

“De cierta manera el enfoque Bayesiano sobre pruebas de hipótesis puede ser considerado como una variante del método de multiplicadores de Lagrange. Si en particular, $P = \pi P_0 + (1 - \pi) P_1$, con algún $0 < \pi < 1$, entonces la función de los multiplicadores de Lagrange (1.19) no es otra cosa que el riesgo Bayesiano...”⁸ [Novikov, 2009]. De lo cual el resolver (1.19) implica también la solución del problema de *riesgo Bayesiano*.

Grosso modo, para resolver el problema de la optimización de $L_{\mathcal{M}}(\psi, \phi)$ se buscará la estructura de la prueba secuencial aleatorizada, llamémosla (ψ^*, ϕ^*) , que alcance la cota mínima (si ésta existe) de la función (1.20). Entonces, se buscarán las condiciones *necesarias y suficientes* tanto para la regla de paro, como para la regla de decisión que minimicen a la función de Lagrange; y solo entonces se obtendrá la prueba óptima. Para encontrar la estructura de la regla óptima ψ^* se trabajarán dos casos: el caso truncado y el no truncado. En el primero se fija un número máximo (N) de etapas⁹ para el experimento y se busca la estructura de la prueba óptima (este caso se puede ver también como el usado en la estadística clásica, cuando se analiza la prueba con base en muestras de tamaño fijo). En el segundo caso se buscará la prueba óptima, cuando el número de grupos por observar crece indefinidamente.

Cabe mencionar que para el primer problema de investigación las observaciones del experimento no se considerarán v.a. independientes, se trabajarán solamente con su función de

⁸Traducción del autor.

⁹O número máximo de grupos de observaciones.

distribución conjunta; pero los tamaños de grupos sí serán v.a. independientes. Aunque en general en toda nuestra investigación las observaciones y los tamaños de grupos serán v.a. independientes entre sí.

Dada la importancia en la investigación, en general, de modelos cuyas observaciones son independientes (*e.g.* [Barber, (2003)], [Bhattachajee-Mukhopadhyay, (2012)], [Mukhopadhyay-De Silva, (2008)], entre otros), como parte de las aplicaciones de nuestro problema principal (1.18), se investigará también: *la(s) estructura(s) de las pruebas óptimas con observaciones independientes y en términos de la razón de probabilidad $Z_n = \frac{f_1^n}{f_0^n}$.*

Una segunda investigación a desarrollar será: dado un problema de pruebas secuenciales de hipótesis simples H_0 vs. H_1 (contexto \mathcal{M}_0), donde las observaciones son independientes e idénticamente distribuidas (i.d.) (con distribución P_0), los tamaños de grupos ν_1, ν_2, \dots v.a. discretas e independientes (no idénticamente distribuidas), y se considera la razón $Z_n = \frac{f_1^n}{f_0^n}$, nos interesará examinar la estructura de la prueba óptima (ψ^*, ϕ^*) , solución de (1.18), en términos de la razón de verosimilitud Z_n , determinando con esto las regiones de continuidad y/o rechazo para la prueba secuencial óptima.

Un tercer problema (continuación del anterior) será investigar la estructura de la prueba secuencial óptima con observaciones i.i.d. y tamaños de grupos aleatorios también i.i.d., y en términos de la razón de verosimilitud Z_n . Este problema es el originalmente planteado en la investigación [Mukhopadhyay-De Silva, (2008)], conocido como la prueba RSPRT. Con esto se podrán establecer algunas propiedades de la RSPRT y sus alcances, como por ejemplo su optimización bajo H_0 y H_1 .

Finalmente un último problema a investigar sería trabajar el conocido problema de Kiefer-Weiss original (véase [Kiefer-Weiss, (1957)]), en el cual se exploran las posibilidades de minimizar el supremo del costo promedio del experimento¹⁰ con un parámetro θ cualquiera, pero fijo *i.e.*

$$\min_{(\psi, \phi)} \sup_{\theta} K_{\mathcal{M}}(\psi). \tag{1.21}$$

Este planteamiento es un problema alterno que se puede investigar como aplicación de la solución de nuestro problema principal (1.18) llamado de Kiefer-Weiss modificado, aplicado a nuestro contexto de grupos de tamaño aleatorio. La solución del problema modificado (1.18) resuelve, con base en investigaciones recientes, el problema Kiefer-Weiss original (1.21) (véanse [Kiefer-Weiss, (1957)], [Weiss, (1962)], [Lorden, (1980)], [Novikov, (2009)], [Novikov, *et al*, (2022)]). En este problema las tres distribuciones (P_0, P_1, P) sí pertenecen a una misma familia paramétrica fija, y además solo se aborda para el caso truncado. Entonces, para este problema solo se buscarán algunas condiciones y estructuras semejantes, pero más

¹⁰En el contexto secuencial clásico, se minimiza al supremo del tiempo de paro promedio, $\sup_{\theta} E_{\theta} \tau$, del experimento.

CAPÍTULO 1. ANÁLISIS SECUENCIAL CON GRUPOS DE TAMAÑO ALEATORIO

generales, a las trabajos en Weiss (1962), Novikov (2009) y Novikov, *et al* (2022). Para esto nos basaremos en la estructura empleada en [Lorden, (1980)] sobre el problema de Kiefer-Weiss modificado. Además, solo para la optimalidad del caso truncado consideraremos a las observaciones como v.a. i.i.d. y a los tamaños de grupos solo independientes; a su vez se analizará la estructura de la prueba secuencial óptima (solución de (1.18)), en términos de las tres funciones de distribución trabajadas, resumidas en las razones de probabilidad $Z_0 = \frac{f_0^n}{f^n}$ y $Z_1 = \frac{f_1^n}{f^n}$, donde f es una distribución cualquiera, pero fija y diferente de las primeras dos, y donde las tres distribuciones pertenecen a la misma familia paramétrica.

CAPÍTULO 2

OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

En este apartado, primero se justificará la metodología utilizada en esta investigación: el método de Lagrange (aplicado al contexto secuencial) y después se buscará la prueba secuencial óptima. Para lo cual introduciremos una clase de pruebas secuenciales que será utilizada a lo largo del trabajo.

El siguiente teorema está basado en lo que Wald llamaba *admisibilidad* ([Wald, (1971)], p. 15), pero aquí está aplicado a nuestro contexto secuencial con grupos aleatorios. Se puede decir que una prueba secuencial es admisible, si NO se puede mejorar respecto a sus componentes de costo (K^*) y probabilidades de error (α^*, β^*). Esto es, si no existe otra prueba secuencial que tenga tanto su costo como sus probabilidades de error menores que ella, con al menos una desigualdad estricta (esto se expresará en (2.2)).

Para que una prueba secuencial sea admisible es suficiente que las “combinaciones lineales” de sus componentes con coeficientes no negativos sean el ínfimo respecto a todas las “combinaciones lineales” de otras pruebas (esto último se expresa en la relación (2.1)).

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Teorema 1.

Sean $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ con al menos una positiva. Si existe una prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{M}}$ tal que

$$b_1 K_{\mathcal{M}}(\psi^*) + b_2 \alpha(\psi^*, \phi^*) + b_3 \beta(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}} \{b_1 K_{\mathcal{M}}(\psi) + b_2 \alpha(\psi, \phi) + b_3 \beta(\psi, \phi)\}. \quad (2.1)$$

Entonces no existe otra prueba $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ tal que

$$K_{\mathcal{M}}(\psi) \leq K_{\mathcal{M}}(\psi^*), \quad \alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (2.2)$$

donde la desigualdad que corresponde a la b_i positiva es estricta en (2.2).

Demostración del Teorema 1:

Supongamos que existe otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ que es mejor que (ψ^*, ϕ^*) , es decir que satisface las tres desigualdades en (2.2) con al menos una desigualdad estricta, la correspondiente con su b_i positiva.

Por hipótesis se tiene que al menos una de las constantes “ b ’s” es positiva. Sin pérdida de generalidad (SPG) supongamos que $b_2 > 0$, entonces su componente asociada es α . Para b_2 la desigualdad estricta correspondiente de (2.2) es $\alpha(\psi, \phi) < \alpha(\psi^*, \phi^*)$. De esto se sigue que:

$$b_1 K_{\mathcal{M}}(\psi) + b_2 \alpha(\psi, \phi) + b_3 \beta(\psi, \phi) < b_1 K_{\mathcal{M}}(\psi^*) + b_2 \alpha(\psi^*, \phi^*) + b_3 \beta(\psi^*, \phi^*),$$

con $b_1 \geq 0$ y $b_3 \geq 0$. Lo cual es una contradicción ya que la combinación lineal de las componentes de (ψ, ϕ) con coeficientes no negativos es más pequeña que el ínfimo, *i. e.* se contradice (2.1) con $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$. Análogamente se genera una contradicción para los casos $b_1 > 0$ y $b_3 > 0$. Por lo tanto, no existe una prueba $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ que satisfaga (2.2) con al menos una desigualdad estricta, la correspondiente a su b_i *i.e.* no puede darse que ($b_1 > 0$ y $K_{\mathcal{M}}(\psi) < K_{\mathcal{M}}(\psi^*)$), ($b_2 > 0$ y $\alpha(\psi, \phi) < \alpha(\psi^*, \phi^*)$) o ($b_3 > 0$ y $\beta(\psi, \phi) < \beta(\psi^*, \phi^*)$). ■

En el teorema anterior la expresión de la derecha de (2.1) puede verse como una “generalización” de la función de Lagrange $L_{\mathcal{M}}(\psi, \phi) = K_{\mathcal{M}}(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi)$, salvo porque el costo en (2.1) está ponderado por una constante $b_1 \geq 0$. En el caso de que sea $b_1 > 0$, se tendrá exactamente la función de Lagrange con $\lambda_0 = \frac{b_2}{b_1} \geq 0$ y $\lambda_1 = \frac{b_3}{b_1} \geq 0$. Las siguientes afirmaciones expresan casos particulares del Teorema 1.

El siguiente corolario señala que cuando exista una prueba secuencial que satisfaga el ínfimo de la función de Lagrange, entonces la prueba será admisible. Es decir, si para cualquier otra prueba que tenga ambas probabilidades de error mejores (por debajo de la primera), entonces

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

necesariamente la prueba admisible tendrá un costo menor. Esta afirmación justifica propiamente el “método de Lagrange” aplicado a nuestro contexto estadístico secuencial (por grupos); lo cual nos llevará a desarrollar nuestro problema con restricciones a uno sin ellas.

Corolario 1.

Para algunas $\lambda_0 > 0$, $\lambda_1 > 0$ fijas, sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{M}}$ tal que

$$L_{\mathcal{M}}(\psi^*, \phi^*; \lambda_0, \lambda_1) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}} L_{\mathcal{M}}(\psi, \phi; \lambda_0, \lambda_1). \quad (2.3)$$

Entonces para cualquier prueba $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ que satisfice

$$\alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (2.4)$$

$$\text{se tiene que} \quad K_{\mathcal{M}}(\psi) \geq K_{\mathcal{M}}(\psi^*). \quad (2.5)$$

La desigualdad (2.5) es estricta, si al menos una de las desigualdades en (2.4) lo es.

Demostración del Corolario 1:

Sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{M}}$ tal que satisfice (2.3), esto es

$$K_{\mathcal{M}}(\psi^*) + \lambda_0 \alpha(\psi^*, \phi^*) + \lambda_1 \beta(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}} \{K_{\mathcal{M}}(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi)\}$$

entonces por Teorema 1 (con $b_1 > 0$, $b_2 > 0$ y $b_3 > 0$) no existe otra prueba que mejore en las tres componentes a (ψ^*, ϕ^*) . Sea $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ cualquier otra prueba tal que satisfice (2.4), entonces

$$K_{\mathcal{M}}(\psi) \geq K_{\mathcal{M}}(\psi^*).$$

Dado que las constantes de la combinación son todas positivas ($\lambda_0 > 0$, $\lambda_1 > 0$), entonces si al menos una de las desigualdades en (2.4) es estricta, la desigualdad (2.5) también lo será. ■

El Corolario 1 nos permite dedicarnos a la solución del problema de minimización de la función de Langrange $L_{\mathcal{M}}(\psi, \phi)$ (ver la condición (2.3)). De cierta forma, una vez que exista la prueba (ψ^*, ϕ^*) que cumpla (2.3) para algunas $\lambda_0, \lambda_1 > 0$ fijas, se toman en automático sus respectivas probabilidades de error $\alpha(\psi^*, \phi^*)$ y $\beta(\psi^*, \phi^*)$ como las cotas máximas de error. De hecho, es precisamente el sentido que se le da a la optimización de la SPRT (véase [Wald-Wolfowitz, (1948)]).

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Corolario 2.

Para alguna $\lambda > 0$ fija, sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{M}}$ una prueba secuencial tal que

$$\alpha(\psi^*, \phi^*) + \lambda\beta(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}} \{ \alpha(\psi, \phi) + \lambda\beta(\psi, \phi) \} \quad (2.6)$$

Entonces para cualquier otra prueba $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$

$$\text{si } \alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*), \quad (2.7)$$

$$\text{se tiene que } \beta(\psi, \phi) \geq \beta(\psi^*, \phi^*). \quad (2.8)$$

La desigualdad (2.8) es estricta, si (2.7) lo es.

Demostración del Corolario 2:

Sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{M}}$ tal que satisface (2.6), entonces por Teorema 1 (con $b_1 = 0$, $b_2 > 0$ y $b_3 > 0$) no existe otra prueba que mejore en las dos componentes de error a (ψ^*, ϕ^*) . Sea $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ cualquier otra prueba tal que satisface (2.7), entonces

$$\beta(\psi, \phi) \geq \beta(\psi^*, \phi^*).$$

Como en el corolario anterior, dado que $\lambda > 0$, si la desigualdad en (2.7) es estricta, entonces (2.8) también lo será. ■

La optimización de la función de Lagrange (1.20) consistirá primero en: dada $\psi \in \mathcal{F}_{\mathcal{M}}$ arbitraria, pero fija, encontrar la cota inferior de $L_{\mathcal{M}}(\psi, \phi)$ sobre todas las reglas de decisión ϕ . A este proceso se le llama el problema de **decisión óptima**, i.e.

$$L_{\mathcal{M}}(\psi) := L_{\mathcal{M}}(\psi, \phi^*) := \inf_{\phi} L_{\mathcal{M}}(\psi, \phi). \quad (2.9)$$

Después de encontrar las condiciones necesarias y suficientes para la regla de decisión óptima (ϕ^*), como segundo paso se encontrará la estructura de la regla de paro óptima. Entonces el problema se reduce a un problema de **paro óptimo**; en éste se buscará la cota inferior para la función (2.9), obteniendo

$$L_{\mathcal{M}}(\psi^*) := \inf_{\psi \in \mathcal{F}_{\mathcal{M}}} L_{\mathcal{M}}(\psi). \quad (2.10)$$

Este problema, a su vez, se divide en dos casos: el caso de paro truncado y el caso no truncado, en cada uno de ellos se buscará la estructura de la regla de paro óptima. Finalmente, con las reglas de paro ψ^* y de decisión ϕ^* óptimas y fijas (o familias de reglas óptimas), se tendrá la mínima cota de la función de Lagrange $L_{\mathcal{M}}(\psi^*, \phi^*)$. Es decir,

$$L_{\mathcal{M}}(\psi^*, \phi^*) := L_{\mathcal{M}}(\psi^*) \leq L_{\mathcal{M}}(\psi) = L_{\mathcal{M}}(\psi, \phi^*) \leq L_{\mathcal{M}}(\psi, \phi). \quad (2.11)$$

2.1.1. Regla de decisión óptima

En esta sección se caracterizará a la regla de decisión ϕ^* que minimiza a la función de Lagrange $L_{\mathcal{M}}(\psi, \phi)$ con $\psi \in \mathcal{F}_{\mathcal{M}}$ fija. Pero, primero tengamos en cuenta lo siguiente.

Notación:

- (1) I_A denota la función indicadora del evento A .
- (2) Si una función ρ está acotada por funciones indicadoras de la siguiente manera

$$I_{\{F_1(x) < F_2(x)\}} \leq \rho(x) \leq I_{\{F_1(x) \leq F_2(x)\}}, \quad (2.12)$$

entonces (2.12) será denotada como:

$$\rho(x) \gtrsim I_{\{F_1(x) \lesssim F_2(x)\}}. \quad (2.13)$$

El símbolo \gtrsim indica las dos desigualdades no estrictas para ρ , esto es la función $\rho(x)$ es mayor o igual (\geq) que $I_{\{F_1(x) < F_2(x)\}}$ y, a su vez, igual o menor (\leq) que $I_{\{F_1(x) \leq F_2(x)\}}$.

Por otro lado, en la indicadora de (2.13) el símbolo \lesssim marcará que la primera indicadora NO tiene igualdad, es decir se tiene solo la desigualdad estricta entre las funciones ($F_1(x) < F_2(x)$), mas la segunda indicadora sí contendrá la desigualdad no estricta ($F_1(x) \leq F_2(x)$), por este motivo se pone \sim .

El siguiente lema se aplicará a la demostración de algunos resultados importantes.

Lema 1.

Sean ϕ, F_1, F_2 algunas funciones medibles sobre un espacio de medida con su medida μ , tales que:

$$0 \leq \phi(x) \leq 1, \quad F_1(x) \geq 0, \quad F_2(x) \geq 0, \quad y \quad (2.14)$$

$$\int \min \{F_1(x), F_2(x)\} d\mu(x) < \infty. \quad (2.15)$$

Entonces

$$\int \left(\phi(x)F_1(x) + (1 - \phi(x))F_2(x) \right) d\mu(x) \geq \int \min \{F_1(x), F_2(x)\} d\mu(x). \quad (2.16)$$

Además, la igualdad en (2.16) se obtiene si y solo si

$$\phi(x) \gtrsim I_{\{F_1(x) \lesssim F_2(x)\}} \quad \mu\text{-c.t.p.} \quad (2.17)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Demostración del Lema 1:

Dadas ϕ, F_1, F_2 funciones medibles que cumplen (2.14) se tiene que: para todo x

$$\phi(x)F_1(x) + (1 - \phi(x))F_2(x) - \min \{F_1(x), F_2(x)\} \geq 0, \quad (2.18)$$

y si, además, se cumple (2.15), entonces se tiene

$$\int \phi(x)F_1(x) + (1 - \phi(x))F_2(x)d\mu(x) \geq \int \min \{F_1(x), F_2(x)\}d\mu(x). \quad (2.19)$$

Ahora, sea el conjunto de x tales que:

$$F_1(x) < F_2(x) \implies \phi(x) = 1 \quad \text{y} \quad \phi(x)F_1(x) + (1 - \phi(x))F_2(x) - \min \{F_1(x), F_2(x)\} = 0,$$

$$F_2(x) < F_1(x) \implies \phi(x) = 0 \quad \text{y} \quad \phi(x)F_1(x) + (1 - \phi(x))F_2(x) - \min \{F_1(x), F_2(x)\} = 0,$$

$$F_1(x) = F_2(x) \implies 0 \leq \phi(x) \leq 1 \quad \text{y} \quad \phi(x)F_1(x) + (1 - \phi(x))F_2(x) - \min \{F_1(x), F_2(x)\} = 0,$$

Por lo tanto

$$\phi(x) \geq I_{\{F_1(x) \lesssim F_2(x)\}} \iff \phi(x)F_1(x) + (1 - \phi(x))F_2(x) - \min \{F_1(x), F_2(x)\} = 0.$$

μ -casi en todas partes (c.t.p.). Por tanto, la igualdad en (2.16) se da si y solo si (2.17) se tiene. ■

Convenciones:

- (a) En el desarrollo de este trabajo no se marcarán, en general, el espacio y/o las variables de integración respectivas, a menos que éste y/o éstas se requieran. Esto es, se trabajarán las funciones sin sus respectivos argumentos.
- (b) Por definición, el valor esperado de una función $g(x)$ es

$$Eg(X) = \int g(x)f(x)d\mu(x),$$

donde f es la función de densidad de X . Sin embargo, ésta generalmente es trabajada como

$$Eg(X) = \int g(x)f(x)I_{\{f(x) \neq 0\}}d\mu(x). \quad (2.20)$$

Entonces de aquí en adelante al trabajar con valores esperados $Eg(X)$ se sobreentenderá que se trabaja sobre el conjunto $\{x : f(x) \neq 0\}$.

- (c) Se trabajarán sobre conjuntos respecto a una medida μ casi en todas partes, esto se denotará μ -c.t.p. También es importante señalar que la mayoría de las desigualdades, límites e integrales trabajadas en esta investigación son válidas, respectivamente μ -casi en todas partes (c.t.p.), aunque no se especifique este punto.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Notación:

(1) Si $\mathbf{n} = (n_1, \dots, n_k) \in \mathcal{G}^k$ son los k tamaños de grupos consecutivos y el vector de sus correspondientes observaciones es $x^{\mathbf{n}} = (x_1^{\mathbf{n}}, \dots, x_k^{\mathbf{n}}) \in \mathcal{X}^{\mathbf{n}}$; en ocasiones, este vector de observaciones $x^{\mathbf{n}}$ será necesario reescribirlo de diferente manera, aunque con las mismas observaciones haciendo énfasis en algún grupo de observaciones en concreto.

Por ejemplo si ahora $\mathbf{n} = (n_1, \dots, n_{k-1}) \in \mathcal{G}^{k-1}$, entonces se reescribe $(\mathbf{n}, n_k) \in \mathcal{G}^k$ y $x^{\mathbf{n}} = (x^{\mathbf{n}}, x^{n_k}) \in \mathcal{X}^{(\mathbf{n}, n_k)}$, donde $x^{n_k} = (x_{|\mathbf{n}|+1}, x_{|\mathbf{n}|+2}, \dots, x_{|\mathbf{n}|+n_k}) \in \mathcal{X}^{n_k}$ es el vector de las observaciones correspondientes al grupo de tamaño n_k .

De lo cual $(x^{\mathbf{n}}, x^{n_k})$ con $\mathbf{n} \in \mathcal{G}^{k-1}$ representará nuevamente al primer vector de todas las observaciones en los k grupos, pero ahora se separan las observaciones del grupo de tamaño n_k , ya que esto nos permitirá trabajar exclusivamente con dicho grupo.

Definición 8.

Sean $\lambda_0, \lambda_1 > 0$. Para cualesquiera $k = 1, 2, \dots$ y $\mathbf{n} \in \mathcal{G}^k$ sea

$$l_{\mathbf{n}}(x^{\mathbf{n}}) := \min \{ \lambda_0 f_0^{\mathbf{n}}(x^{\mathbf{n}}), \lambda_1 f_1^{\mathbf{n}}(x^{\mathbf{n}}) \}. \quad (2.21)$$

Por cuestiones prácticas, a lo largo del trabajo la función anterior (2.21) será expresada como “ $l_{\mathbf{n}}$ ” sin su argumento $x^{\mathbf{n}}$, a menos que éste sea necesario.

El siguiente corolario demuestra que $l_{\mathbf{n}}$ es una función integrable.

Corolario 3.

Sean $\lambda_0 > 0$, $\lambda_1 > 0$, $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$ cualesquiera,

$$\int l_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \min \{ \lambda_0, \lambda_1 \}, \quad \mu^{\mathbf{n}}\text{-c.t.p.} \quad (2.22)$$

Demostración del Corolario 3:

Sean $k \geq 1$, $\mathbf{n} \in \mathcal{G}^k$.

$$\int_{\mathcal{X}^{\mathbf{n}}} l_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \lambda_0 \int_{\mathcal{X}^{\mathbf{n}}} f_0^{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \lambda_0. \quad (2.23)$$

Análogamente,

$$\int_{\mathcal{X}^{\mathbf{n}}} l_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \lambda_1 \int_{\mathcal{X}^{\mathbf{n}}} f_1^{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \lambda_1. \quad (2.24)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Por tanto, de (2.23) y (2.24) se obtiene:

$$\int_{\mathcal{X}^n} l_n(x^n) d\mu^n(x^n) \leq \min \{ \lambda_0, \lambda_1 \}, \quad \mu^n\text{-c.t.p.} \quad \blacksquare$$

Definición 9.

Sean $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$, y ψ cualquiera se define el conjunto,

$$S_n^\psi := \{ x^n = (x_1^n, \dots, x_k^n) \in \mathcal{X}^n : s_n^\psi(x^n) > 0 \}. \quad (2.25)$$

Definición 10.

Sea $k \geq 1$ se define el conjunto

$$\mathcal{P}_k := \{ \mathbf{n} = (n_1, \dots, n_k) \in \mathcal{G}^k : p_n > 0 \}. \quad (2.26)$$

Condición Δ_ψ (Regla de decisión óptima)

Dada una regla de paro ψ fija, una regla de decisión ϕ **cumple la condición Δ_ψ** , si para cada $k \geq 1$ y cualquier $\mathbf{n} \in \mathcal{P}_k$ se tiene:

$$\phi_n \stackrel{\geq}{\leq} I_{\{ \lambda_0 f_0^n \lesssim \lambda_1 f_1^n \}}, \quad (2.27)$$

sobre el conjunto S_n^ψ μ^n -c.t.p., para $\lambda_0 > 0$ y $\lambda_1 > 0$.

La condición Δ_ψ marca la estructura que tienen las componentes ϕ_n y algunas características para ellas de una regla de decisión ϕ . En lo subsiguiente se demostrará que dicha condición es necesaria y suficiente para alcanzar el ínfimo de la función de Lagrange $L_{\mathcal{M}}(\psi, \phi)$ sobre todas las reglas de decisión y para una regla de paro ψ fija.

El nombre “ Δ_ψ ” se utiliza por mnemotecnica del término “decisión” utilizando la letra griega mayúscula “ Δ ”; y el subíndice “ ψ ” marca la dependencia de la regla de paro, la cual está presente en el conjunto S_n^ψ .

El siguiente lema marca las condiciones necesarias y suficientes para que la regla de decisión alcance cierta cota.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Lema 2.

Para cualesquiera $\lambda_0 > 0$, $\lambda_1 > 0$ y para cualquier prueba secuencial (ψ, ϕ) se tiene

$$\lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) \geq \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}}. \quad (2.28)$$

La igualdad de (2.28) se obtiene si y solo si ϕ cumple la **condición** Δ_{ψ} .

Demostración del Lema 2:

Tomemos el lado izquierdo de (2.28), con base en (1.12) y (1.13) se tiene:

$$\begin{aligned} & \lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) = \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \left(\lambda_0 \int s_{\mathbf{n}}^{\psi} \phi_{\mathbf{n}} f_0^{\mathbf{n}} d\mu^{\mathbf{n}} + \lambda_1 \int s_{\mathbf{n}}^{\psi} (1 - \phi_{\mathbf{n}}) f_1^{\mathbf{n}} d\mu^{\mathbf{n}} \right); \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} \left(\phi_{\mathbf{n}} \lambda_0 f_0^{\mathbf{n}} + (1 - \phi_{\mathbf{n}}) \lambda_1 f_1^{\mathbf{n}} \right) d\mu^{\mathbf{n}}. \end{aligned}$$

Entonces por Lema 1, (2.16),

$$\begin{aligned} \lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) & \geq \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} \min \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \} d\mu^{\mathbf{n}}, \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}}, \quad \text{por (2.21)}. \end{aligned}$$

Lo cual demuestra (2.28).

Ahora, nuevamente, con base en el Lema 1 (segunda parte) se tendrá la igualdad entre las integrales presentes en (2.28) si y solo si se cumple cierta estructura para ϕ .

Sin embargo, si existe algún $k \geq 1$ con sus tamaños de grupos $\mathbf{n} \in \mathcal{G}^k$ tal que su probabilidad conjunta es nula ($p_{\mathbf{n}} = 0$), entonces en (2.28) se tendrá la igualdad $0 = 0$ independientemente de la estructura de (ψ, ϕ) . De lo cual, para este caso *no es necesaria la estructura* (2.17). Por lo tanto, tomemos aquellos \mathbf{n} tales que $p_{\mathbf{n}} > 0$, (*i.e.* $\mathbf{n} \in \mathcal{P}_k$), para cada $k \geq 1$.

Ahora sí por Lema 1 (2.17) se tiene que para cualesquiera $\lambda_0 > 0$, $\lambda_1 > 0$ y (ψ, ϕ) una

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

prueba secuencial

$$\lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{P}_k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}}, \quad (2.29)$$

si y solo si

$$\phi_{\mathbf{n}} \geq I_{\{s_{\mathbf{n}}^{\psi} \lambda_0 f_0^{\mathbf{n}} \lesssim s_{\mathbf{n}}^{\psi} \lambda_1 f_1^{\mathbf{n}}\}}, \quad \mu^{\mathbf{n}}\text{-c.t.p. para } \mathbf{n} \in \mathcal{P}_k. \quad (2.30)$$

Sin embargo, si $s_{\mathbf{n}}^{\psi}(x_1, \dots, x_n) = 0$ entonces $\phi_{\mathbf{n}}$ no tendrá una caracterización, ya que

$$I_{\{s_{\mathbf{n}}^{\psi} \lambda_0 f_0^{\mathbf{n}} < s_{\mathbf{n}}^{\psi} \lambda_1 f_1^{\mathbf{n}}\}} = 0 \quad \text{y} \quad I_{\{s_{\mathbf{n}}^{\psi} \lambda_0 f_0^{\mathbf{n}} \leq s_{\mathbf{n}}^{\psi} \lambda_1 f_1^{\mathbf{n}}\}} = 1.$$

Por tanto, tomemos el caso cuando $s_{\mathbf{n}}^{\psi}(x^{\mathbf{n}}) > 0$. Sea el conjunto $S_{\mathbf{n}}^{\psi}$, entonces (2.30) se reescribe:

$$\phi_{\mathbf{n}} \geq I_{\{\lambda_0 f_0^{\mathbf{n}} \lesssim \lambda_1 f_1^{\mathbf{n}}\}}, \quad (2.31)$$

sobre el conjunto $S_{\mathbf{n}}^{\psi}$, $\mu^{\mathbf{n}}$ -c.t.p., para cualquier $k = 1, 2, \dots$ y $\mathbf{n} \in \mathcal{P}_k$. Por lo tanto, la igualdad en (2.28) se alcanza si y solo si ϕ cumple la condición Δ_{ψ} . ■

El lema anterior indica cuándo se obtiene la cota inferior del lado derecho de (2.28). Ahora como el lado izquierdo de (2.28) está presente en la función de Lagrange (1.19), entonces se tendrán las condiciones necesarias y suficientes para alcanzar una cota inferior de esta función.

Teorema 2.

Para cualquier prueba secuencial $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ se tiene que:

$$L_{\mathcal{M}}(\psi, \phi) \geq \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \quad (2.32)$$

La igualdad en (2.32) se alcanza si y solo si ϕ **cumple la condición Δ_{ψ}** .

Demostración del Teorema 2:

Sea $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$ cualquiera, entonces con base en (1.16), (1.19) y (2.28) se tiene:

$$\begin{aligned} L_{\mathcal{M}}(\psi, \phi) &= K_{\mathcal{M}}(\psi) + \lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) \geq \\ &\geq \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^{\psi} + \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{P}_k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}} = \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \end{aligned}$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Finalmente, por la segunda parte del Lema 2 la desigualdad anterior se alcanza si y solo si ϕ cumple la condición Δ_ψ . ■

El resultado anterior indica que para cualquier $(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}}$:

$$\inf_{\phi} L_{\mathcal{M}}(\psi, \phi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \quad (2.33)$$

Más aún este ínfimo se alcanza si y solo si ϕ cumple la condición Δ_ψ . Por otro lado, como el término derecho de (2.33) ya no depende de alguna regla de decisión ϕ , solo de las reglas de paro $\psi \in \mathfrak{F}_{\mathcal{M}}$, sea:

$$L_{\mathcal{M}}(\psi) := \inf_{\phi} L_{\mathcal{M}}(\psi, \phi). \quad (2.34)$$

Es importante señalar que a través de esta investigación se abusará un poco de la notación, debido a que en ocasiones se utilizará la “misma” letra L para diferentes funciones, o en algunos otros casos para la misma función. Sin embargo, éstas tendrán algún tipo de elemento distintivo, como por ejemplo diferentes argumentos o algún detalle que distinga su uso.

2.1.2. Regla de paro óptima: caso truncado.

La mayoría de los experimentos tienen recursos limitados (por ejemplo, tiempo, dinero, espacio, personal, etc.), entonces es natural hacer un análisis para estos casos. Por lo cual se puede, de antemano, determinar un número máximo de grupos por observar (en dependencia de los recursos) para el experimento, de tal manera que el proceso se detenga al llegar a dicho límite como máximo, si es que no se había detenido el análisis antes. Esto es muy común en el contexto estadístico clásico, cuando se establece una muestra de tamaño fijo para el experimento. Sin embargo, en este contexto secuencial también se puede establecer una etapa máxima o número máximo de grupos a analizar, sin necesariamente llegar a ella.

En esta sección se caracterizará a la regla de paro que minimiza a la función $L_{\mathcal{M}}(\psi)$ (2.34) en la clase de las pruebas de paro truncadas.

Supongamos que el problema de prueba de hipótesis \mathcal{M} se puede truncar. En este caso se toma un número máximo de etapas de paro en el experimento. Sea N la etapa de paro máxima en un experimento secuencial con grupos de tamaño aleatorio, *i.e.* se considerarán a lo más N grupos de observaciones. A lo largo del trabajo se tomará a $N \geq 1$ como un número entero fijo.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Definición 11.

Se define la clase de las reglas de paro truncadas:

$$\mathcal{F}^N := \left\{ \psi : (1 - \psi_{(\mathbf{n}:1)}) \cdots (1 - \psi_{(\mathbf{n}:N-1)}) (1 - \psi_{(\mathbf{n}:N)}) = 0 \quad \mu^n\text{-c.t.p.} \right. \\ \left. \text{para cualquier } \mathbf{n} \in \mathcal{G}^N \right\}. \quad (2.35)$$

De la definición anterior se tiene que $\psi \in \mathcal{F}^N$ si y solo $t_{\mathbf{n}}^{\psi} = 0$ para toda $\mathbf{n} \in \mathcal{G}^{N+1}$ μ^n -c.t.p. De lo cual $s_{\mathbf{n}}^{\psi} = 0$ para toda $\mathbf{n} \in \mathcal{G}^k$ y para toda $k \geq N + 1$. Ahora de lo anterior y con base en (1.9) se deduce que si $\psi \in \mathcal{F}^N$ entonces $P(\tau^{\psi} \leq N) = 1$. Además, dado que $P(\tau^{\psi} < \infty) \geq P(\tau^{\psi} \leq N)$ entonces $P(\tau^{\psi} < \infty) = 1$. Por tanto, $\mathcal{F}^N \subset \mathcal{F}_{\mathcal{M}}$, es decir

$$\text{si } \psi \in \mathcal{F}^N \text{ entonces } \psi \in \mathcal{F}_{\mathcal{M}}. \quad (2.36)$$

Ahora para cada regla $\psi \in \mathcal{F}^N$ y utilizando (2.33) se determina la función de Lagrange de dicha regla como:

$$\begin{aligned} L_{\mathcal{M}}(\psi) &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} \\ &= \sum_{k=1}^N \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} + \sum_{k=N+1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} \\ &= \sum_{k=1}^N \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \end{aligned}$$

El último paso se debe a (1.3) ya que $s_{\mathbf{n}}^{\psi} = 0$ para toda $k \geq N + 1$ y para toda $\mathbf{n} \in \mathcal{G}^k$, cuando $\psi \in \mathcal{F}^N$, por lo tanto a partir de la etapa $k = N + 1$ se tienen sumandos nulos. Cabe mencionar que $\psi \in \mathcal{F}^N$ tiene definidas todas sus componentes $\psi_{\mathbf{n}}$ para cualesquiera $\mathbf{n} \in \mathcal{G}^k$ y $k \geq 1$, solo que a partir de $k > N$ dichas componentes valen cero.

Además, como $\psi \in \mathcal{F}^N$ entonces para $\mathbf{n} \in \mathcal{G}^N$ se tiene:

$$t_{\mathbf{n}}^{\psi} (1 - \psi_{\mathbf{n}}) = 0 \quad \implies \quad t_{\mathbf{n}}^{\psi} = t_{\mathbf{n}}^{\psi} \psi_{\mathbf{n}} = s_{\mathbf{n}}^{\psi}.$$

Por tanto la función de Lagrange truncada para $\psi \in \mathcal{F}^N$ se expresa como:

$$L_{\mathcal{M}}^N(\psi) = \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} + \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \quad (2.37)$$

Dado que para $\psi \in \mathcal{F}^N$ la función $L_{\mathcal{M}}^N(\psi)$ tiene un número finito de sumandos a diferencia de (2.34), entonces como un referente de que se utilizan reglas de la clase \mathcal{F}^N se señala el subíndice

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

N en (2.37). La función L^N será llamada la función de Lagrange truncada.

Por último, de las propiedades de la clase \mathcal{F}^N se demuestra fácilmente que

$$\mathcal{F}^1 \subset \mathcal{F}^2 \subset \dots \subset \mathcal{F}^N \subset \mathcal{F}^{N+1} \subset \dots \subset \mathcal{F}_{\mathcal{M}}.$$

Ahora se buscará la regla de paro truncada $\psi \in \mathcal{F}^N$ que minimice a la función $L_{\mathcal{M}}^N(\psi)$ (2.37).

Definición 12.

Para $r \geq 1$, sea

$$\mathcal{V}_r := \left\{ v = \{v_{\mathbf{n}}\}_{\mathbf{n} \in \mathcal{G}^r} : v_{\mathbf{n}} = v_{\mathbf{n}}(x^{\mathbf{n}}) \text{ son medibles, no negativas y tales que} \right. \\ \left. v_{\mathbf{n}} \leq l_{\mathbf{n}} \quad \forall \mathbf{n} \in \mathcal{G}^r \right\}. \quad (2.38)$$

Tómese en cuenta que v denota a una familia cuyos elementos $v_{\mathbf{n}}$ son funciones medibles, no negativas y acotadas por $l_{\mathbf{n}}$ para toda $\mathbf{n} \in \mathcal{G}^r$, por lo cual $v_{\mathbf{n}} \in v \in \mathcal{V}_r$. Un ejemplo de estas familias es $\{l_{\mathbf{n}}\}_{\mathbf{n} \in \mathcal{G}^r} \in \mathcal{V}_r$, para cualquier $r \geq 1$. Entonces los elementos determinan a su familia y viceversa. Cabe mencionar que debido a que $v_{\mathbf{n}}$ es medible, no negativa y acotada por $l_{\mathbf{n}}$ función integrable (véase (2.22)), por lo tanto $v_{\mathbf{n}}$ es μ^n -integrable para toda $n \in \mathbb{N}$, i.e. $\int v_{\mathbf{n}} d\mu^n < \infty$.

Convenciones: Es importante tomar en cuenta los siguientes convenios.

- (a) Si para algún $k \geq 1$, se tiene que $n_k = 0$, (i.e. existe algún grupo que no tiene observaciones) y se quiere “integrar” respecto a las variables (que no hay) de este grupo, entonces dicha “integración” no realizará ningún efecto. En este caso se tendrá el espacio medible \mathcal{X}^0 y se considera la medida de probabilidad μ^0 respecto a su σ -álgebra trivial. Entonces si $\mathbf{n} = (n_1, \dots, n_{k-1}, n_k = 0) \in \mathcal{G}^k$, $x^{\mathbf{n}} = (x_1^n, \dots, x_{k-1}^n, ())$, y $v_{\mathbf{n}}(x^{\mathbf{n}})$ es una función medible, entonces para $\mathbf{n} \in \mathcal{G}^{k-1}$,

$$\int v_{(\mathbf{n}, n_k)}(x^{\mathbf{n}}, ()) d\mu^0(()) = v_{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.39)$$

- (b) Si se tiene una suma vacía, ésta tendrá un valor nulo, i.e. $\sum_{k=1}^0 _ = 0$. (2.40)

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Lema 3.

Para cualquier $k \geq 1$, y cualesquiera $\mathbf{n} \in \mathcal{G}^k$ y $m \in \mathcal{G}$ se tiene

$$\int_{\mathcal{X}^m} l_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq l_{\mathbf{n}}(x^{\mathbf{n}}), \quad \mu^{\mathbf{n}}\text{-c.t.p.} \quad (2.41)$$

Demostración del Lema 3:

Sean $k \geq 1$, $\mathbf{n} \in \mathcal{G}^k$. Sea $m \in \mathcal{G}$, entonces para (\mathbf{n}, m) se tiene:

$$\int_{\mathcal{X}^m} l_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq \lambda_0 \int_{\mathcal{X}^m} f_0^{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq \lambda_0 f_0^{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.42)$$

Análogamente a (2.42),

$$\int_{\mathcal{X}^m} l_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq \lambda_1 \int_{\mathcal{X}^m} f_1^{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq \lambda_1 f_1^{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.43)$$

Por tanto,

$$\int_{\mathcal{X}^m} l_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq \min \{ \lambda_0 f_0^{\mathbf{n}}(x^{\mathbf{n}}), \lambda_1 f_1^{\mathbf{n}}(x^{\mathbf{n}}) \} = l_{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.44)$$

Las integrales anteriores están definidas aún cuando $m = 0$, debido a la convección (2.39). ■

Definición 13.

Sea $r \geq 2$ cualquiera, se define el operador $I_r : \mathcal{V}_r \rightarrow \mathcal{V}_{r-1}$ mediante lo siguiente: para cualesquiera $v \in \mathcal{V}_r$ y $\mathbf{n} \in \mathcal{G}^{r-1}$

$$(I_r v)_{\mathbf{n}}(x^{\mathbf{n}}) := \sum_{m \in \mathcal{G}} p_r(m) \int v_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m). \quad (2.45)$$

El operador I_r (2.45) está bien definido debido al Teorema de Fubini y al Lema 3. Ya que, por un lado las funciones $v_{\mathbf{n}}$ son $\mu^{\mathbf{n}}$ -integrables para toda $\mathbf{n} \in \mathbb{N}$, por lo cual la integración de m iteraciones de funciones medibles y no negativas resulta en una función medible y no negativa. Por otro lado, debido a que $v_{\mathbf{n}} \leq l_{\mathbf{n}}$ para toda $\mathbf{n} \in \mathcal{G}^r$ con $r \geq 2$ entonces por Lema 3 se tiene que $\int v_{(\mathbf{n},m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \leq l_{\mathbf{n}}$, ahora para $\mathbf{n} \in \mathcal{G}^r$ con $r \geq 1$, por lo tanto $I_r v \in \mathcal{V}_{r-1}$.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Entonces el operador I_r se aplica a los elementos de una familia v de dimensión r y a su vez se genera una nueva familia $I_r v$ de dimensión $r - 1$ (esto es $(I_r v)_n \in I_r v \in \mathcal{V}_{r-1}$).

Definición 14.

Sea $r \geq 2$ cualquiera, se define el operador $M_r : \mathcal{V}_r \rightarrow \mathcal{V}_{r-1}$ mediante lo siguiente: para cualesquiera $v \in \mathcal{V}_r$ y $\mathbf{n} \in \mathcal{G}^{r-1}$

$$(M_r v)_n := \min \left\{ l_n, \bar{c}_r f^n + (I_r v)_n \right\}. \quad (2.46)$$

Claramente para $\mathbf{n} \in \mathcal{G}^{r-1}$ se tiene que $(M_r v)_n$ son funciones mediables y no negativas; y también $(M_r v)_n \leq l_n$ con $\mathbf{n} \in \mathcal{G}^{r-1}$. Por lo tanto, $(M_r v) \in M_r v \in \mathcal{V}_{r-1}$, está bien definido.

Definición 15.

Sean ψ y $r \geq 1$ cualesquiera. Se define el funcional $B_r^\psi : \mathcal{V}_r \rightarrow \mathbb{R}$ para cualquier $v \in \mathcal{V}_r$

$$B_r^\psi(v) := \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + l_n) d\mu^n + \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + v_n) d\mu^n. \quad (2.47)$$

El funcional (2.47) está bien definido debido al Corolario 4. Además, el funcional está en dependencia directa de la regla de paro ψ y de la dimensión r de la familia de funciones v a la cual se aplica.

Ahora en particular, la función de Lagrange truncada $L_{\mathcal{M}}^N(\psi)$ (2.37) puede escribirse en términos del funcional B . Si $\psi \in \mathcal{F}^N$ entonces

$$\begin{aligned} L_{\mathcal{M}}^N(\psi) &= \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + l_n) d\mu^n + \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + l_n) d\mu^n, \\ &= B_N^\psi(\{l_n\}). \end{aligned} \quad (2.48)$$

Claramente $\{l_n\}_{\mathbf{n} \in \mathcal{G}^N} \in \mathcal{F}^N$.

Es importante señalar la siguiente propiedad de B_r^ψ para $r = 1$. Dada cualquier $v \in \mathcal{F}_{\mathcal{M}}$, el operador B_1 no depende de la regla de paro ψ , debido a que dicha dependencia desaparece en

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

este caso, como se describe en lo siguiente:

$$\begin{aligned}
 B_1^\psi(v) &:= \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + v_{\mathbf{n}}) d\mu^{\mathbf{n}} = \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int (c(\mathbf{n})f^n + v_{\mathbf{n}}) d\mu^{\mathbf{n}}, \\
 &= \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} c(\mathbf{n}) \int f^n(x^n) d\mu^{\mathbf{n}}(x^n) + \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int v_{\mathbf{n}}(x^n) d\mu^{\mathbf{n}}(x^n). \\
 &= \bar{c}_1 + \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int v_{\mathbf{n}}(x^n) d\mu^{\mathbf{n}}(x^n) = B_1(v).
 \end{aligned} \tag{2.49}$$

La segunda igualdad del primer paso se sigue por (1.3), ya que $t_{\mathbf{n}}^\psi = 1$ cuando $\mathbf{n} \in \mathcal{G}$.

Corolario 4.

El valor $B_r^\psi(v)$ es finito para cualesquiera ψ , $r \geq 1$ y $v \in \mathcal{V}_r$. (2.50)

Demostración del Corolario 4:

Por definición (2.47),

$$B_r^\psi(v) = \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + l_{\mathbf{n}}) d\mu^{\mathbf{n}} + \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + v_{\mathbf{n}}) d\mu^{\mathbf{n}}.$$

Sea $1 \leq k \leq r-1$ cualquiera, pero fija. Tomemos el k -ésimo miembro de la primera suma

$$\begin{aligned}
 \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + l_{\mathbf{n}}) d\mu^{\mathbf{n}} &\leq \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \left(c(\mathbf{n}) \int f^n(x^n) d\mu^{\mathbf{n}}(x^n) + \int l_{\mathbf{n}} d\mu^{\mathbf{n}}(x^n) \right) \\
 &\leq \sum_{\mathbf{n} \in \mathcal{G}^k} (p_{\mathbf{n}} c(\mathbf{n}) + \lambda_0 p_{\mathbf{n}}) = \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) + \lambda_0 \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} = S_k + \lambda_0 < \infty.
 \end{aligned} \tag{2.51}$$

Donde $S_k := \sum_{i=1}^k \bar{c}_i < \infty$, debido a (1.7) para cada $k \geq 1$. Sumando (2.51) para $1 \leq k \leq r-1$

se obtiene el primer sumando de B_r^ψ , el cual es menor o igual a $\sum_{k=1}^{r-1} (S_k + \lambda_0) < \infty$ para cada $r \geq 2$.

Análogamente, el segundo sumando:

$$\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi (c(\mathbf{n})f^n + v_{\mathbf{n}}) d\mu^{\mathbf{n}} \leq \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int (c(\mathbf{n})f^n + l_{\mathbf{n}}) d\mu^{\mathbf{n}} \leq S_r + \lambda_0 < \infty. \tag{2.52}$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Por lo tanto, de (2.51) y (2.52) se sigue $B_r^\psi(v) \leq \sum_{k=1}^r (S_k + \lambda_0) < \infty$, para cualesquiera ψ , $r \geq 1$ y $v \in \mathcal{V}_r$. ■

Definición 16.

Sea ψ regla de paro cualquiera, se define

$$T_{\mathbf{n}}^\psi := \{x^n = (x_1^n, \dots, x_k^n) \in \mathcal{X}^n : t_{\mathbf{n}}^\psi(x^n) > 0\}. \quad (2.53)$$

Con base en el Corolario 4 se pueden hacer comparaciones entre los funcionales B_r^ψ , por lo cual el siguiente lema es posible.

Lema 4.

Sean ψ y $r \geq 1$ cualesquiera. Si $v \in \mathcal{V}_{r+1}$ entonces

$$B_{r+1}^\psi(v) \geq B_r^\psi(M_{r+1}v). \quad (2.54)$$

La igualdad en (2.54) se alcanza si y solo si ,

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1}f^n + (I_{r+1}v)_{\mathbf{n}}\}}, \quad (2.55)$$

sobre el conjunto $T_{\mathbf{n}}^\psi$ μ^n -c.t.p. para toda $\mathbf{n} \in \mathcal{P}_r$.

Demostración del Lema 4:

Sea $r \geq 1$ y $v \in \mathcal{V}_{r+1}$. Por definiciones (2.47), (2.45) y (2.46), y por Corolario 4, la desigualdad (2.54) es equivalente a demostrar que:

$$\begin{aligned} & \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi(c(\mathbf{n})f^n + l_{\mathbf{n}}) d\mu^n(x^n) + \sum_{\mathbf{n} \in \mathcal{G}^{r+1}} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi(c(\mathbf{n})f^n + v_{\mathbf{n}}) d\mu^n(x^n) \\ & \geq \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi(c(\mathbf{n})f^n + (M_{r+1}v)_{\mathbf{n}}) d\mu^n(x^n). \end{aligned} \quad (2.56)$$

Donde $(M_{r+1}v)_{\mathbf{n}} = \min \{l_{\mathbf{n}}, \bar{c}_{r+1}f^n + (I_{r+1}v)_{\mathbf{n}}\}$ con $\mathbf{n} \in \mathcal{G}^r$.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Desarrollemos el lado izquierdo de (2.56), utilizando el teorema de Fubini, se obtiene:

$$\begin{aligned}
& \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi}(c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}})d\mu^{\mathbf{n}} + \sum_{(\mathbf{n}, m) \in \mathcal{G}^{r+1}} p_{(\mathbf{n}, m)} \int t_{(\mathbf{n}, m)}^{\psi}(c((\mathbf{n}, m))f^{(\mathbf{n}, m)} + v_{(\mathbf{n}, m)})d\mu^{(\mathbf{n}, m)} \\
&= \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(\psi_{\mathbf{n}}(c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) + \right. \right. \\
&\quad \left. \left. + \sum_{m \in \mathcal{G}} p_{r+1}(m)(1 - \psi_{\mathbf{n}}) \int \left((c(\mathbf{n}) + c(m))f^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) + v_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) \right) d\mu^m(x^m) \right) d\mu^{\mathbf{n}} \right] \\
&= \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(\psi_{\mathbf{n}}(c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) + (1 - \psi_{\mathbf{n}}) \left[\sum_{m \in \mathcal{G}} p_{r+1}(m)c(\mathbf{n}) \int f^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) + \right. \right. \right. \\
&\quad \left. \left. \left. + \sum_{m \in \mathcal{G}} p_{r+1}(m) \int (c(m)f^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) + v_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m)) d\mu^m(x^m) \right] \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \right], \\
&= \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(\psi_{\mathbf{n}}(c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) + \right. \right. \\
&\quad \left. \left. + (1 - \psi_{\mathbf{n}}) \left[c(\mathbf{n})f^{\mathbf{n}} + \sum_{m \in \mathcal{G}} p_{r+1}(m)c(m)f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}} \right] \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \right], \\
&= \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n})f^{\mathbf{n}} + \psi_{\mathbf{n}}l_{\mathbf{n}} + (1 - \psi_{\mathbf{n}}) \left[\bar{c}_{r+1}f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}} \right] \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \right].
\end{aligned}$$

En la tercera igualdad se aplicó que $\int f^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) = f^{\mathbf{n}}(x^{\mathbf{n}})$ y $\sum_{m \in \mathcal{G}} p_{r+1}(m) = 1$.

Ahora utilizando el Lema 1, (2.16) con $\mathbf{n} \in \mathcal{G}^r$ y

$$\phi(x) \equiv \psi_{\mathbf{n}}, \quad F_1 \equiv l_{\mathbf{n}} \quad \text{y} \quad F_2 \equiv \bar{c}_{r+1}f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}}.$$

Donde además por (2.22),

$$\int \min \{F_1, F_2\} d\mu^{\mathbf{n}} = \int \min \{l_{\mathbf{n}}, \bar{c}_{r+1}f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}}\} d\mu^{\mathbf{n}} \leq \int l_{\mathbf{n}} d\mu^{\mathbf{n}} < \infty.$$

Por tanto, para toda $r \geq 1$ y cualquier $\mathbf{n} \in \mathcal{G}^r$,

$$\begin{aligned}
& \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n})f^{\mathbf{n}} + \psi_{\mathbf{n}}l_{\mathbf{n}} + (1 - \psi_{\mathbf{n}}) \left[\bar{c}_{r+1}f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}} \right] \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \right] \\
&\geq \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n})f^{\mathbf{n}} + \min \{l_{\mathbf{n}}, \bar{c}_{r+1}f^{\mathbf{n}} + (I_{r+1}v)_{\mathbf{n}}\} \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \right] \\
&= \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n})f^{\mathbf{n}} + (M_{r+1}v)_{\mathbf{n}} \right) d\mu^{\mathbf{n}}(x^{\mathbf{n}}). \tag{2.57}
\end{aligned}$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Lo cual demuestra (2.54).

Ahora, nuevamente, con base en el Lema 1 se tendrá la igualdad entre algunas integrales presentes en (2.57) si y solo si se cumple cierta estructura para $\psi_{\mathbf{n}}$.

Sin embargo, si existe algún \mathbf{n} tal que $p_{\mathbf{n}} = 0$, entonces se tendrá la igualdad en (2.57) independientemente de la regla ψ . Entonces para este caso *no es necesaria la estructura* (2.55). Por lo tanto, tomemos \mathbf{n} tal que $p_{\mathbf{n}} > 0$ (*i.e.* sea $\mathbf{n} \in \mathcal{P}_r$).

Entonces, por Lema 1 (2.17) la igualdad en (2.57) se alcanza si y solo si

$$\psi_{\mathbf{n}} \stackrel{\cong}{\approx} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1} f^{\mathbf{n} + (I_{r+1} v)_{\mathbf{n}}}\}}, \quad (2.58)$$

sobre el conjunto $T_{\mathbf{n}}^{\psi}$ $\mu^{\mathbf{n}}$ -c.t.p., para toda $\mathbf{n} \in \mathcal{P}_r$. ■

Definición 17.

Se define para $r = N$

$$V_r^N := \{l_{\mathbf{n}}\}_{\mathbf{n} \in \mathcal{G}^N}, \quad (2.59)$$

y recursivamente para $r = N - 1, N - 2, \dots, 1$,

$$V_r^N := M_{r+1} V_{r+1}^N. \quad (2.60)$$

Obviamente por propiedades de la familia $\{l_{\mathbf{n}}\}_{\mathbf{n} \in \mathcal{G}^N}$ y por la Definición 14 del operador M_r se tiene que para toda $r = N, \dots, 1$:

$$V_r^N \in \mathcal{V}_r. \quad (2.61)$$

El siguiente corolario nos permitirá acotar por debajo a la función de Langrange truncada $L_{\mathcal{M}}^N$.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Corolario 5.

Sea $\psi \in \mathcal{F}^N$ cualquier regla de paro (truncada). Entonces para cualquier $1 \leq r \leq N-1$ se tiene la siguiente desigualdad:

$$B_{r+1}^\psi(V_{r+1}^N) \geq B_r^\psi(V_r^N). \quad (2.62)$$

Además, se alcanza la igualdad en (2.62) si y solo si ψ satisface para toda $\mathbf{n} \in \mathcal{G}^r$:

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1} f^{\mathbf{n}+}(I_{r+1} V_{r+1}^N)_{\mathbf{n}}\}}, \quad (2.63)$$

sobre el conjunto $T_{\mathbf{n}}^\psi$ $\mu^{\mathbf{n}}$ -c.t.p.

Demostración del Corolario 5:

Se demostrará por inducción sobre $r = N-1, \dots, 1$.

Sea $r = N-1$, por (2.47), (2.48) y aplicando la desigualdad (2.54) del Lema 4, se tiene que:

$$L_{\mathcal{M}}^N(\psi) = B_N^\psi(\{l_{\mathbf{n}}\}) = B_N^\psi(V_N^N) \geq B_{N-1}^\psi(M_N V_N^N) = B_{N-1}^\psi(V_{N-1}^N). \quad (2.64)$$

Las últimas igualdades se tienen por las definiciones (2.59) y (2.60) con $V_{N-1}^N = M_N V_N^N$. Además por (2.55) se obtiene la igualdad en (2.64) si y solo si

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_N f^{\mathbf{n}+}(I_N V_N^N)_{\mathbf{n}}\}}, \quad (2.65)$$

sobre $T_{\mathbf{n}}^\psi$ $\mu^{\mathbf{n}}$ -c.t.p., para toda $\mathbf{n} \in \mathcal{P}_{N-1}$.

Ahora supongamos que (2.62) se cumple para alguna $r = N-1, N-2, \dots, 2$. Entonces por (2.60) y (2.61) se sabe que $V_r^N = M_{r+1} V_{r+1}^N \in \mathcal{V}_r$. Aplicando, nuevamente, (2.54) del Lema 4

$$B_r^\psi(V_r^N) \geq B_{r-1}^\psi(M_r V_r^N) = B_{r-1}^\psi(V_{r-1}^N). \quad (2.66)$$

donde $V_{r-1}^N = M_r V_r^N \in \mathcal{F}_{r-1}$. Además por (2.55) se obtiene la igualdad en (2.66) si y solo si

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_r f^{\mathbf{n}+}(I_r V_r^N)_{\mathbf{n}}\}}, \quad (2.67)$$

sobre el conjunto $T_{\mathbf{n}}^\psi$ $\mu^{\mathbf{n}}$ -c.t.p., para toda $\mathbf{n} \in \mathcal{P}_{r-1}$.

Por lo tanto (2.62) se cumple para toda $1 \leq r \leq N-1$. Finalmente, también se demuestra que se alcanza la igualdad en (2.62) si y solo si ψ satisface para toda $\mathbf{n} \in \mathcal{G}^r$ (2.63) sobre el conjunto $T_{\mathbf{n}}^\psi$ $\mu^{\mathbf{n}}$ -c.t.p. ■

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Condición Π^N . (Regla de paro truncada óptima)

Una regla de paro ψ satisface la condición Π^N , si cumple las siguientes dos condiciones:

Para toda $1 \leq r \leq N - 1$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$

$$\psi_{(\mathbf{n}:r)} \geq I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1} f^{\mathbf{n}} + (I_{r+1} V_{r+1}^N)_{\mathbf{n}}\}}, \quad (2.68)$$

sobre el conjunto $T_{\mathbf{n}}^{\psi} \mu^{\mathbf{n}}$ -c.t.p., para $\lambda_0 > 0$ y $\lambda_1 > 0$.

Y para toda $\mathbf{n} \in \mathcal{P}_N$,

$$\psi_{\mathbf{n}} = 1, \quad \text{sobre el conjunto } T_{\mathbf{n}}^{\psi} \mu^{\mathbf{n}}\text{-c.t.p.} \quad (2.69)$$

La condición Π^N caracteriza a una regla de paro truncada ($\psi \in \mathcal{F}^N$), ya que necesariamente dicha regla detiene el experimento en la etapa N , si no lo hizo antes.

El nombre “ Π^N ” hace referencia al nombre “paro” (regla), por lo cual se utiliza la letra griega mayúscula “ Π ” y el superíndice “ N ” marca el nivel de truncamiento (horizonte finito).

Teorema 3.

Para cualquier $\psi \in \mathcal{F}^N$,

$$L_{\mathcal{M}}^N(\psi) \geq B_1(V_1^N). \quad (2.70)$$

Donde V_1^N está definido en (2.60).

Se obtiene la igualdad en (2.70) si y solo si ψ **satisface la condición Π^N** .

Demostración del Teorema 3:

Aplicando sucesivamente la desigualdad (2.62) del Corolario 5 se tiene que:

$$L_{\mathcal{M}}^N(\psi) = B_N^{\psi}(V_N^N) \geq B_{N-1}^{\psi}(V_{N-1}^N) \geq B_{N-2}^{\psi}(V_{N-2}^N) \geq \cdots \geq B_1(V_1^N). \quad (2.71)$$

Por tanto, por (2.49)

$$L_{\mathcal{M}}^N(\psi) \geq B_1(V_1^N) = \bar{c}_1 + \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int (V_1^N)_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.72)$$

Lo cual demuestra (2.70), donde V_1^N está definido en (2.60).

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Por otro lado, también del Corolario 5 se tiene la igualdad en cada una de las desigualdades de la cadena (2.71) si y solo si ψ satisface la condición Π^N . Por tanto para dicha ψ se tiene:

$$L_{\mathcal{M}}^N(\psi) = B_1(V_1^N). \quad (2.73)$$

■

El resultado anterior nos da las condiciones necesarias y suficientes para que una regla de paro truncada alcance la cota inferior $B_1(V_1^N)$ de la función de Langrange (2.37). Por lo tanto, **la regla** (o familia) **de paro truncada óptima** es toda aquella que satisface la condición Π^N .

Un ejemplo de este tipo de regla de paro truncada óptima fue desarrollada en la investigación [Popoca, (2012)]. Aunque dicha regla es óptima cuando tanto las observaciones como los tamaños de los grupos son v.a. independientes e idénticamente distribuidas.

Otra manera de expresar el Teorema 3 es la siguiente:

$$L_{\mathcal{M}}^N(\psi) = \inf_{\psi' \in \mathcal{F}^N} L_{\mathcal{M}}^N(\psi') \quad \text{si y solo si } \psi \text{ satisface la condición } \Pi^N. \quad (2.74)$$

Definición 18.

Se define la clase de pruebas secuenciales cuya regla de paro detiene el experimento en la etapa N o antes, *i.e.*

$$\mathfrak{S}^N := \{(\psi, \phi) : \psi \in \mathcal{F}^N\}. \quad (2.75)$$

Debido a la propiedad (2.36) entre las clases \mathcal{F}^N y $\mathcal{F}_{\mathcal{M}}$ se tiene que $\mathfrak{S}^N \subset \mathfrak{S}_{\mathcal{M}}$.

Condición (Π^N, Δ_{ψ}) . (Prueba secuencial truncada óptima)

Una prueba secuencial (ψ, ϕ) satisface la condición (Π^N, Δ_{ψ}) , si ψ cumple la condición Π^N y, para dicha regla de paro, ϕ satisface la condición Δ_{ψ} , para algunas $\lambda_0 > 0$ y $\lambda_1 > 0$.

El siguiente teorema nos da condiciones necesarias y suficientes para que la prueba secuencial truncada alcance el ínfimo de la función de Lagrange truncada $L_{\mathcal{M}}^N$ (resolución con horizonte finito de nuestro problema sin restricciones).

Teorema 4.

Si una prueba secuencial (ψ^*, ϕ^*) satisface la condición (Π^N, Δ_ψ) entonces

$$L_{\mathcal{M}}^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_{\mathcal{M}}^N(\psi, \phi). \quad (2.76)$$

Además, si existe alguna prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}^N$ tal que cumple (2.76) entonces la prueba (ψ^*, ϕ^*) satisface la condición (Π^N, Δ_ψ) .

Demostración del Teorema 4:

Suficiencia: Sea (ψ^*, ϕ^*) tal que satisface la condición (Π^N, Δ_{ψ^*}) . Debido a que ϕ^* satisface la condición Δ_{ψ^*} , entonces por Teorema 2 (segunda parte) y a (2.33) se tiene:

$$L_{\mathcal{M}}(\psi, \phi^*) = \inf_{\phi} L_{\mathcal{M}}(\psi, \phi), \quad \text{para cualquier } \psi \in \mathcal{F}_{\mathcal{M}}.$$

Por otro lado, ya que ψ^* cumple Π^N , entonces $\psi^* \in \mathcal{F}^N$ y por (2.37) $L_{\mathcal{M}}(\psi^*, \phi^*) = L_{\mathcal{M}}^N(\psi^*)$. De lo cual y por (2.74) se tiene que:

$$L_{\mathcal{M}}^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_{\mathcal{M}}^N(\psi, \phi).$$

Esto demuestra la primera parte del teorema.

Necesidad: Supongamos que existe una prueba secuencial (ψ^*, ϕ^*) con $\psi^* \in \mathcal{F}^N$ y tal que cumple (2.76). Entonces (ψ^*, ϕ^*) minimiza a la función de Lagrange (2.37), por tanto por Teoremas 2 y 3 (segunda parte) se tiene que ψ^* y ϕ^* cumplen, respectivamente, las condiciones Π^N y Δ_{ψ^*} . Por lo tanto, (ψ^*, ϕ^*) satisface la condición (Π^N, Δ_{ψ^*}) . ■

El siguiente teorema nos da condiciones suficientes de cuándo una prueba secuencial truncada minimiza el costo promedio del experimento $K_{\mathcal{M}}(\psi)$ (resolución con horizonte finito de nuestro problema con restricciones).

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Teorema 5.

Si una prueba secuencial truncada (ψ^*, ϕ^*) satisface la condición (Π^N, Δ_{ψ^*}) , entonces para cualquier otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}^N$ tal que

$$\alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (2.77)$$

se tiene que

$$K_{\mathcal{M}}(\psi) \geq K_{\mathcal{M}}(\psi^*). \quad (2.78)$$

La desigualdad (2.78) es estricta, si al menos una de las desigualdades en (2.77) lo es.

Demostración del Teorema 5: Como (ψ^*, ϕ^*) satisface la condición (Π^N, Δ_{ψ^*}) entonces, por Teorema 4, se tiene que $L_{\mathcal{M}}^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_{\mathcal{M}}^N(\psi, \phi)$.

Ahora sea $(\psi, \phi) \in \mathfrak{S}^N$ cualquier otra prueba secuencial que cumple (2.77). Además se sabe que $\mathfrak{S}^N \subset \mathfrak{S}_{\mathcal{M}}$ entonces por Corolario 1 y (2.4) se sigue (2.78).

La desigualdad (2.78) es estricta, si al menos una de las desigualdades en (2.77) lo es. ■

Con el Teorema 5 se concluye la demostración de la optimalidad de la prueba secuencial truncada, dando **condiciones suficientes** para que ella tenga un costo promedio del experimento mínimo, sobre cualquier otra prueba truncada. Esta suficiencia se realizó a través del método de Lagrange (generalizado) o Teorema 1 (de admisibilidad). Sin embargo, también es posible demostrar, solo para el caso truncado, el teorema “recíproco” de la generalización de Lagrange. Esta sección se terminará con la demostración de este recíproco, demostrando así que para el caso truncado la manera de encontrar la optimalidad de una prueba secuencial truncada es a través del método generalizado de Lagrange y viceversa.

Lema 5.

Para $N \geq 1$ cualquiera, el siguiente conjunto es convexo

$$\mathcal{B}^N := \left\{ (K, \alpha, \beta) \in \mathbb{R}^3 : \text{existe una prueba secuencial } (\psi, \phi) \in \mathfrak{S}^N \text{ tal que} \right. \\ \left. K = K_{\mathcal{M}}(\psi), \alpha = \alpha(\psi, \phi), \beta = \beta(\psi, \phi) \right\}. \quad (2.79)$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Demostración del Lema 5:

Sean (K_1, α_1, β_1) y $(K_2, \alpha_2, \beta_2) \in \mathcal{B}^N$ cualesquiera, entonces existen pruebas secuenciales (ψ_1, ϕ_1) y $(\psi_2, \phi_2) \in \mathfrak{S}^N$ tales que sus costos promedios y probabilidades de error son, respectivamente:

$$K_1 = K_1(\psi_1), \alpha_1 = \alpha_1(\psi_1, \phi_1), \beta_1 = \beta_1(\psi_1, \phi_1) \text{ y } K_2 = K_2(\psi_2), \alpha_2 = \alpha_2(\psi_2, \phi_2), \beta_2 = \beta_2(\psi_2, \phi_2).$$

Las pruebas (ψ_1, ϕ_1) y (ψ_2, ϕ_2) generan a su vez las probabilidades condicionadas $t_n^{\psi_1}$, $s_n^{\psi_1}$ y $t_n^{\psi_2}$, $s_n^{\psi_2}$, respectivamente, para toda $\mathbf{n} \in \mathcal{G}^k$ y toda $k \geq 1$.

Sea $\gamma \in [0, 1]$ y sea la combinación lineal $\gamma(K_1, \alpha_1, \beta_1) + (1 - \gamma)(K_2, \alpha_2, \beta_2) = (K, \alpha, \beta)$.

Se demostrará que (K, α, β) pertenece al conjunto \mathcal{B}^N , para cualquier $\gamma \in [0, 1]$. Para lo cual, para toda $\gamma \in [0, 1]$ se construirá una prueba secuencial $(\psi, \phi) \in \mathfrak{S}^N$ tal que $K = \gamma K_1 + (1 - \gamma)K_2$, $\alpha = \gamma \alpha_1 + (1 - \gamma)\alpha_2$ y $\beta = \gamma \beta_1 + (1 - \gamma)\beta_2$.

Entonces para $\gamma = 1$ y sea la prueba secuencial $(\psi_1, \phi_1) \in \mathfrak{S}^N$ con su respectivo $(K_1, \alpha_1, \beta_1) \in \mathcal{B}^N$. Análogamente, para $\gamma = 0$, sea la prueba secuencial $(\psi_2, \phi_2) \in \mathfrak{S}^N$ con su respectivo $(K_2, \alpha_2, \beta_2) \in \mathcal{B}^N$.

Sea $\gamma \in (0, 1)$ cualquiera, se definen para toda $k \geq 1$ y para toda $\mathbf{n} \in \mathcal{G}^k$.

$$\psi_{\mathbf{n}} := \begin{cases} \frac{\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2}}{\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma) t_{\mathbf{n}}^{\psi_2}}, & \text{si } t_{\mathbf{n}}^{\psi_1} > 0 \text{ o } t_{\mathbf{n}}^{\psi_2} > 0; \\ 1, & \text{en otro caso.} \end{cases} \quad (2.80)$$

$$\phi_{\mathbf{n}} := \begin{cases} \frac{\gamma s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}}}{\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2}}, & \text{si } s_{\mathbf{n}}^{\psi_1} > 0 \text{ o } s_{\mathbf{n}}^{\psi_2} > 0; \\ 1, & \text{en otro caso.} \end{cases} \quad (2.81)$$

Como $t_n^{\psi_i}$ y $s_n^{\psi_i}$ son medibles para $i = 1, 2$, se sigue que $\psi_{\mathbf{n}}$ y $\phi_{\mathbf{n}}$ son también funciones medibles.

Por otro lado, sean $\gamma_1 = \frac{\gamma t_{\mathbf{n}}^{\psi_1}}{\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma) t_{\mathbf{n}}^{\psi_2}}$ y $(1 - \gamma_1) = \frac{(1 - \gamma) t_{\mathbf{n}}^{\psi_2}}{\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma) t_{\mathbf{n}}^{\psi_2}}$, entonces $\gamma_1 \in [0, 1]$ y $\psi_{\mathbf{n}} = \gamma_1 \psi_{1\mathbf{n}} + (1 - \gamma_1) \psi_{2\mathbf{n}}$. Además, se sabe que $\min\{\psi_{1\mathbf{n}}, \psi_{2\mathbf{n}}\} \leq \gamma_1 \psi_{1\mathbf{n}} + (1 - \gamma_1) \psi_{2\mathbf{n}} \leq \max\{\psi_{1\mathbf{n}}, \psi_{2\mathbf{n}}\}$. Por tanto $0 \leq \psi_{\mathbf{n}} \leq 1$.

Análogamente, sean $\gamma_2 = \frac{\gamma s_{\mathbf{n}}^{\psi_1}}{\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2}}$ y $(1 - \gamma_2) = \frac{(1 - \gamma) s_{\mathbf{n}}^{\psi_2}}{\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2}}$ entonces $\gamma_2 \in [0, 1]$ y $\phi_{\mathbf{n}} = \gamma_2 \phi_{1\mathbf{n}} + (1 - \gamma_2) \phi_{2\mathbf{n}}$. Además, $\min\{\phi_{1\mathbf{n}}, \phi_{2\mathbf{n}}\} \leq \gamma_2 \phi_{1\mathbf{n}} + (1 - \gamma_2) \phi_{2\mathbf{n}} \leq \max\{\phi_{1\mathbf{n}}, \phi_{2\mathbf{n}}\}$. De lo cual $0 \leq \phi_{\mathbf{n}} \leq 1$. Por tanto, para todo $\mathbf{n} \in \mathcal{G}^k$ y toda $k \geq 1$, $\psi = \{\psi_{\mathbf{n}}\}_k$ y $\phi = \{\phi_{\mathbf{n}}\}_k$ son sucesiones de funciones medibles en $[0, 1]$.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Ahora se demostrará por inducción que $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}$ para toda $k \geq 1$ y toda $\mathbf{n} \in \mathcal{G}^k$, utilizando las definiciones (2.80), (1.3) y (1.4).

Sea $k = 1$ y $\mathbf{n} \in \mathcal{G}^1$ cualquiera, entonces por definición (1.3) $t_{\mathbf{n}}^{\psi} := 1$ y $t_{\mathbf{n}}^{\psi_1} := 1 = t_{\mathbf{n}}^{\psi_2}$, de lo cual $\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2} = 1$. Por lo tanto para toda $\mathbf{n} \in \mathcal{G}^1$, $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2} = 1$.

Supongamos que para $k > 1$ y cualquier $\mathbf{n} \in \mathcal{G}^k$ se cumple que $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}$. Ahora tomemos $k + 1$ y $\mathbf{n} \in \mathcal{G}^{k+1}$ cualquiera, entonces por (1.3)

$$t_{\mathbf{n}}^{\psi} = (1 - \psi_{(\mathbf{n}:1)})(1 - \psi_{(\mathbf{n}:2)}) \cdots (1 - \psi_{(\mathbf{n}:k-1)})(1 - \psi_{(\mathbf{n}:k)}) = t_{(\mathbf{n}:k)}^{\psi}(1 - \psi_{(\mathbf{n}:k)}). \quad (2.82)$$

Si $t_{(\mathbf{n}:k)}^{\psi_1} = 0$ y $t_{(\mathbf{n}:k)}^{\psi_2} = 0$, entonces por hipótesis de inducción $t_{(\mathbf{n}:k)}^{\psi} = \gamma t_{(\mathbf{n}:k)}^{\psi_1} + (1 - \gamma)t_{(\mathbf{n}:k)}^{\psi_2} = 0$. De esto y de (2.82) se tiene que $t_{\mathbf{n}}^{\psi} = t_{(\mathbf{n}:k)}^{\psi}(1 - \psi_{(\mathbf{n}:k)}) = 0$.

Por otro lado $t_{\mathbf{n}}^{\psi_1} = t_{(\mathbf{n}:k)}^{\psi_1}(1 - \psi_{1(\mathbf{n}:k)}) = 0$ y $t_{\mathbf{n}}^{\psi_2} = t_{(\mathbf{n}:k)}^{\psi_2}(1 - \psi_{2(\mathbf{n}:k)}) = 0$. Por lo tanto para cualquier $\mathbf{n} \in \mathcal{G}^{k+1}$, $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2} = 0$.

Si $t_{(\mathbf{n}:k)}^{\psi_1} > 0$ o $t_{(\mathbf{n}:k)}^{\psi_2} > 0$, entonces por (2.82)

$$\begin{aligned} t_{\mathbf{n}}^{\psi} &= t_{(\mathbf{n}:k)}^{\psi}(1 - \psi_{(\mathbf{n}:k)}) = t_{(\mathbf{n}:k)}^{\psi} \left(1 - \frac{\gamma s_{(\mathbf{n}:k)}^{\psi_1} + (1 - \gamma)s_{(\mathbf{n}:k)}^{\psi_2}}{\gamma t_{(\mathbf{n}:k)}^{\psi_1} + (1 - \gamma)t_{(\mathbf{n}:k)}^{\psi_2}} \right), \quad \text{por (2.80),} \\ &= t_{(\mathbf{n}:k)}^{\psi} \left(\frac{\gamma t_{(\mathbf{n}:k)}^{\psi_1}(1 - \psi_{1\mathbf{n}}) + (1 - \gamma)t_{(\mathbf{n}:k)}^{\psi_2}(1 - \psi_{2\mathbf{n}})}{t_{(\mathbf{n}:k)}^{\psi}} \right), \quad \text{por hipótesis de inducción.} \\ &= \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}. \end{aligned}$$

Ahora se demostrará que $s_{\mathbf{n}}^{\psi} = \gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma)s_{\mathbf{n}}^{\psi_2}$ para toda $k \geq 1$ y toda $\mathbf{n} \in \mathcal{G}^k$, utilizando lo anterior y las definiciones (2.80), (1.3), (1.4).

Sea $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$ cualesquiera. Por la definición (1.4) se tiene que $s_{\mathbf{n}}^{\psi} = t_{\mathbf{n}}^{\psi} \psi_{\mathbf{n}}$.

Si $t_{\mathbf{n}}^{\psi_1} = 0$ y $t_{\mathbf{n}}^{\psi_2} = 0$, entonces de lo anterior $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2} = 0$, y por tanto $s_{\mathbf{n}}^{\psi} = 0$. Por otro lado, también se tiene que $s_{\mathbf{n}}^{\psi_1} = 0$ y $s_{\mathbf{n}}^{\psi_2} = 0$, y de lo cual $\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma)s_{\mathbf{n}}^{\psi_2} = 0$. Por tanto $s_{\mathbf{n}}^{\psi} = \gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma)s_{\mathbf{n}}^{\psi_2} = 0$.

Si $t_{\mathbf{n}}^{\psi_1} > 0$ o $t_{\mathbf{n}}^{\psi_2} > 0$, entonces

$$s_{\mathbf{n}}^{\psi} = t_{\mathbf{n}}^{\psi} \psi_{\mathbf{n}} = (\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}) \left(\frac{\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma)s_{\mathbf{n}}^{\psi_2}}{\gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}} \right) = \gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma)s_{\mathbf{n}}^{\psi_2}.$$

Ahora se demostrará que $\psi \in \mathcal{F}^N$. Sea $N \geq 1$ y $\mathbf{n} \in \mathcal{G}^N$ cualesquiera, sabemos que $\psi_1, \psi_2 \in \mathcal{F}^N$, entonces $t_{\mathbf{n}}^{\psi_1} = 0$ y $t_{\mathbf{n}}^{\psi_2} = 0$ para $\mathbf{n} \in \mathcal{G}^{N+1}$ $\mu^{\mathbf{n}}$ -c.t.p. Entonces como $t_{\mathbf{n}}^{\psi} = \gamma t_{\mathbf{n}}^{\psi_1} + (1 - \gamma)t_{\mathbf{n}}^{\psi_2}$,

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

por tanto $t_n^\psi = 0$ para toda $\mathbf{n} \in \mathcal{G}^{N+1}$. De lo cual y por definición (2.35), $\psi \in \mathcal{F}^N$.

Finalmente como $\psi \in \mathcal{F}_{\mathcal{M}}$ (ya que $\mathcal{F}^N \subset \mathcal{F}_{\mathcal{M}}$, (2.36)), entonces existe su costo promedio $K(\psi)$. Utilizando las definiciones de $K(\psi)$ (1.16), $\alpha(\psi, \phi)$ (1.12) y $\beta(\psi, \phi)$ (1.13) se tiene que:

$$\begin{aligned} K(\psi) &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^\psi = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E (\gamma s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2}) \\ &= \gamma \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^{\psi_1} + (1 - \gamma) \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^{\psi_2}. \end{aligned}$$

De lo cual se deduce que $K(\psi) = \gamma K_1(\psi_1) + (1 - \gamma) K_2(\psi_2)$.

$$\begin{aligned} \alpha(\psi, \phi) &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 \left(s_{\mathbf{n}}^\psi \phi_{\mathbf{n}} I_{\{s_{\mathbf{n}}^{\psi_1} > 0 \cup s_{\mathbf{n}}^{\psi_2} > 0\}} + s_{\mathbf{n}}^\psi \phi_{\mathbf{n}} I_{\{s_{\mathbf{n}}^{\psi_1} = 0 \cap s_{\mathbf{n}}^{\psi_2} = 0\}} \right) = \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi \left(\frac{\gamma s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}}}{s_{\mathbf{n}}^\psi} \right) I_{\{s_{\mathbf{n}}^{\psi_1} > 0 \cup s_{\mathbf{n}}^{\psi_2} > 0\}} = \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 (\gamma s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}}) \left(1 - I_{\{s_{\mathbf{n}}^{\psi_1} = 0 \cap s_{\mathbf{n}}^{\psi_2} = 0\}} \right) = \\ &= \gamma \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}}. \end{aligned}$$

Lo anterior implica que $\alpha(\psi, \phi) = \gamma \alpha_1(\psi_1, \phi_1) + (1 - \gamma) \alpha_2(\psi_2, \phi_2)$.

$$\begin{aligned} \beta(\psi, \phi) &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 \left(s_{\mathbf{n}}^\psi (1 - \phi_{\mathbf{n}}) I_{\{s_{\mathbf{n}}^{\psi_1} > 0 \cup s_{\mathbf{n}}^{\psi_2} > 0\}} + s_{\mathbf{n}}^\psi (1 - \phi_{\mathbf{n}}) I_{\{s_{\mathbf{n}}^{\psi_1} = 0 \cap s_{\mathbf{n}}^{\psi_2} = 0\}} \right) \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 s_{\mathbf{n}}^\psi \left(\frac{s_{\mathbf{n}}^\psi - (\gamma s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}})}{s_{\mathbf{n}}^\psi} \right) I_{\{s_{\mathbf{n}}^{\psi_1} > 0 \cup s_{\mathbf{n}}^{\psi_2} > 0\}} = \\ &= \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 \left((\gamma s_{\mathbf{n}}^{\psi_1} - (1 - \gamma) s_{\mathbf{n}}^{\psi_2}) - (\gamma s_{\mathbf{n}}^{\psi_1} \phi_{1\mathbf{n}} + (1 - \gamma) s_{\mathbf{n}}^{\psi_2} \phi_{2\mathbf{n}}) \right) \\ &\hspace{25em} \left(1 - I_{\{s_{\mathbf{n}}^{\psi_1} = 0 \cap s_{\mathbf{n}}^{\psi_2} = 0\}} \right) \\ &= \gamma \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 s_{\mathbf{n}}^{\psi_1} (1 - \phi_{1\mathbf{n}}) + (1 - \gamma) \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 s_{\mathbf{n}}^{\psi_2} (1 - \phi_{2\mathbf{n}}). \end{aligned}$$

Se concluye que $\beta(\psi, \phi) = \gamma \beta_1(\psi_1, \phi_1) + (1 - \gamma) \beta_2(\psi_2, \phi_2)$.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Por tanto, para cada $\gamma \in (0, 1)$ se define la prueba secuencial (ψ, ϕ) con componentes en (2.80) y (2.81) con $\psi \in \mathcal{F}^N$ y $K = \gamma K_1 + (1 - \gamma)K_2$, $\alpha = \gamma\alpha_1 + (1 - \gamma)\alpha_2$ y $\beta = \gamma\beta_1 + (1 - \gamma)\beta_2$; de lo cual $(K, \alpha, \beta) \in \mathcal{B}^N$. Por lo tanto, \mathcal{B}^N es convexo y además es no vacío. ■

El siguiente teorema nos marca condiciones necesarias para que una prueba secuencial truncada sea admisible.

Teorema 6.

Sea una prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}^N$. Supongamos que no existe otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}^N$ tal que

$$K_{\mathcal{M}}(\psi) \leq K_{\mathcal{M}}(\psi^*), \quad \alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (2.83)$$

con al menos una desigualdad estricta. Entonces

$$b_1 K_{\mathcal{M}}(\psi^*) + b_2 \alpha(\psi^*, \phi^*) + b_3 \beta(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} \{b_1 K_{\mathcal{M}}(\psi) + b_2 \alpha(\psi, \phi) + b_3 \beta(\psi, \phi)\}, \quad (2.84)$$

para $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ con al menos una positiva.

La hipótesis asociada a (2.83) está basada en la idea de “admisibilidad” de Wald, salvo porque aquí se aplica a una prueba estadística secuencial truncada. En este teorema estamos suponiendo que no existe otra prueba secuencial truncada que mejore en sus tres componentes a la prueba (ψ^*, ϕ^*) truncada. El teorema marca que para que una prueba no se pueda mejorar es *necesario* que sea el ínfimo de la “generalización” de la función de Lagrange.

Demostración del Teorema 6:

Sea una prueba secuencial (ψ^*, ϕ^*) con $\psi^* \in \mathcal{F}^N$. Sean $K^* = K_{\mathcal{M}}(\psi^*)$, $\alpha^* = \alpha(\psi^*, \phi^*)$ y $\beta^* = \beta(\psi^*, \phi^*)$. Se define el conjunto

$$\mathcal{A} := \{(K, \alpha, \beta) \in \mathbb{R}^3 : K < K^*, \alpha < \alpha^* \text{ y } \beta < \beta^*\}. \quad (2.85)$$

Sean (K_1, α_1, β_1) y $(K_2, \alpha_2, \beta_2) \in \mathcal{A}$.

Sea $\gamma \in [0, 1]$ cualquiera y sean

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

$K := \gamma K_1 + (1 - \gamma)K_2$, $\alpha := \gamma\alpha_1 + (1 - \gamma)\alpha_2$ y $\beta := \gamma\beta_1 + (1 - \gamma)\beta_2$. Entonces

$$\begin{aligned} K &= \gamma K_1 + (1 - \gamma)K_2 < \gamma K^* + (1 - \gamma)K^* = K^*. \\ \alpha &= \gamma\alpha_1 + (1 - \gamma)\alpha_2 < \gamma\alpha^* + (1 - \gamma)\alpha^* = \alpha^*. \\ \beta &= \gamma\beta_1 + (1 - \gamma)\beta_2 < \gamma\beta^* + (1 - \gamma)\beta^* = \beta^*. \end{aligned}$$

Esto es $(K, \alpha, \beta) \in \mathcal{A}$. Por tanto \mathcal{A} es convexo. Además es no vacío.

Por Lema 5 el conjunto \mathcal{B}^N es convexo. Supongamos que $\mathcal{B}^N \cap \mathcal{A} \neq \emptyset$, sea $(K, \alpha, \beta) \in \mathcal{B}^N \cap \mathcal{A}$, entonces existe una prueba secuencial $(\psi, \phi) \in \mathfrak{S}^N$ tal que $K = K_{\mathcal{M}}(\psi, \phi)$, $\alpha = \alpha(\psi, \phi)$ y $\beta = \beta(\psi, \phi)$ y además $K < K^*$ y $\alpha < \alpha^*$, $\beta < \beta^*$, pero esto es una contradicción porque por hipótesis no existe una prueba que cumpla (2.83) con al menos una desigualdad estricta. Por lo tanto $\mathcal{B}^N \cap \mathcal{A} = \emptyset$.

Ahora utilizando el Teorema de Separación para conjuntos convexos (Theorem 11.3, pp. 97 en [Rockafellar, (1972)]) existe un hiperplano que separa propiamente a los conjuntos \mathcal{A} y \mathcal{B}^N . Por Teorema 11.1 (Theorem 11.1, pp. 95 en [Rockafellar, (1972)]) existe un $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$ tal que:

$$\inf \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{B}^N\} \geq \sup \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{A}\}. \quad (2.86)$$

Utilizando otra desigualdad del teorema citado (*ídem*) se deduce que $(b_1, b_2, b_3) \neq (0, 0, 0)$.

Por un lado, se tiene que el ínfimo de (2.86) es finito debido a lo siguiente. Tomemos una prueba secuencial (ψ', ϕ') tal que ψ' detiene el experimento en la etapa uno, *i.e.* $\psi_{\mathbf{n}} = 1$ para toda $\mathbf{n} \in \mathcal{P}_1$. Como dicha regla es truncada entonces $\psi' \in \mathcal{F}_{\mathcal{M}}^1 \subset \mathcal{F}^N$. De lo cual, $(\psi', \phi') \in \mathfrak{S}^N$. Además, sus respectivas probabilidades de error son finitas ($\alpha' \leq 1$ y $\beta' \leq 1$) y debido a que $s_{\mathbf{n}}^{\psi'} = 1$ para toda $\mathbf{n} \in \mathcal{G}^1$ entonces

$$K' = K_{\mathcal{M}}(\psi') = \sum_{\mathbf{n} \in \mathcal{G}^1} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^{\psi'} = \sum_{\mathbf{n} \in \mathcal{G}} p_1(n) c_1(n) = \bar{c}_1 < \infty.$$

La finitud de K' se sigue por (1.7). Entonces $(K', \alpha', \beta') \in \mathcal{B}^N$. Por tanto $\inf \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{B}^N\} \leq b_1 K' + b_2 \alpha' + b_3 \beta' < \infty$.

Sea $M := \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{A}\}$.

Sin pérdida de generalidad (SPG), supongamos que $b_2 < 0$. Sean $b_1 K$ y $b_3 \beta$ fijos, tomemos una sucesión $\alpha_n \rightarrow -\infty$ cuando $n \rightarrow \infty$, entonces $b_2 \alpha_n \xrightarrow{n \rightarrow \infty} \infty$. De lo cual $\sup M = \infty$ y de la desigualdad (2.86) se tiene

$$\inf \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{B}^N\} \geq \infty.$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Lo cual es una contradicción, porque el ínfimo es finito. Por lo tanto, $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ con al menos una positiva.

Ahora como $b_i \geq 0$ para cada $i = 1, 2, 3$, con al menos una desigualdad estricta. Entonces sean las sucesiones: $K_n \rightarrow K^*$, $\alpha_n \rightarrow \alpha^*$ y $\beta_n \rightarrow \beta^*$ cuando $n \rightarrow \infty$, con $0 < K_n < K^*$, $0 < \alpha_n < \alpha^*$ y $0 < \beta_n < \beta^*$ para toda $n \in \mathbb{N}$. Entonces $(K_n, \alpha_n, \beta_n) \in \mathcal{A}$ para toda $n \in \mathbb{N}$. Por tanto, $0 \leq \lim_{n \rightarrow \infty} (b_1 K_n + b_2 \alpha_n + b_3 \beta_n) \leq \sup M$. De lo cual $0 \leq b_1 K^* + b_2 \alpha^* + b_3 \beta^* \leq \sup M$. Por lo que de (2.86) se tiene:

$$\inf \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{B}^N\} \geq b_1 K^* + b_2 \alpha^* + b_3 \beta^*.$$

Sin embargo, como $(K^*, \alpha^*, \beta^*) \in \mathcal{B}^N$ entonces

$$\inf \{b_1 K + b_2 \alpha + b_3 \beta : (K, \alpha, \beta) \in \mathcal{B}^N\} = b_1 K^* + b_2 \alpha^* + b_3 \beta^*. \quad (2.87)$$

Donde $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ con al menos una positiva. \blacksquare

Observaciones: Existen relaciones importantes que se deducen del teorema anterior. Supongamos que existe una prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{G}^N$ que es admisible, i.e. no se puede mejorar al mismo tiempo respecto a sus tres componentes (su costo promedio y sus probabilidades de error). Entonces por Teorema 6, dicha prueba es el ínfimo “ponderado” (satisface (2.84)) para algunas $b_1 \geq 0$, $b_2 \geq 0$ y $b_3 \geq 0$ con al menos una positiva.

- (a) Si $b_1 > 0$ entonces dicho ínfimo “ponderado” (los elementos del conjunto izquierdo de (2.84)) se puede reescribir como $K^* + \lambda_0 \alpha^* + \lambda_1 \beta^*$ para algunas $\lambda_0 = \frac{b_2}{b_1} \geq 0$ y $\lambda_1 = \frac{b_3}{b_1} \geq 0$. Lo cual es nuestra conocida función de Lagrange truncada aplicada a la prueba admisible $L_{\mathcal{M}}^N(\psi^*, \phi^*)$. Por lo cual, la igualdad (2.84) se convierte en

$$\inf_{(\psi, \phi) \in \mathfrak{G}^N} L_{\mathcal{M}}^N(\psi, \phi) = K^* + \lambda_0 \alpha^* + \lambda_1 \beta^*. \quad (2.88)$$

Por tanto, la prueba secuencial (ψ^*, ϕ^*) admisible satisface (2.76) y por Teorema 4 (segunda parte) (ψ^*, ϕ^*) satisface también la condición (Π^N, Δ_{ψ^*}) . Es decir una prueba admisible satisface la estructura de la prueba óptima.

Este resultado nos marca las condiciones necesarias del método de Lagrange para encontrar la prueba truncada óptima truncada.

- (b) Si $b_1 = 0$ entonces supongamos SPG que $b_2 > 0$ y sea $\lambda = \frac{b_3}{b_2} \geq 0$, entonces (2.84) se reescribe:

$$\inf_{(\psi, \phi) \in \mathfrak{G}^N} \{\alpha + \lambda \beta\} = \alpha^* + \lambda \beta^*. \quad (2.89)$$

En este caso se tiene la prueba secuencial cuyo costo promedio $K_{\mathcal{M}}$ no se considera y solo se trabajan sus probabilidades de error. Pero, debido al Corolario 2, si existe otra prueba (ψ, ϕ) tal que $\alpha(\psi, \phi) \leq \alpha^*$ entonces $\beta(\psi, \phi) \geq \beta^*$. La prueba admisible sigue siendo óptima.

2.1.3. Regla de paro óptima: caso no truncado

Con base en las reglas (o familia) de paro truncadas óptimas ($\psi \in \mathcal{F}^N$) que satisfacen la condición Π^N para N fija, en esta sección se caracterizará a la regla de paro óptima cuando el nivel de truncamiento vaya creciendo (*i.e.* $N \rightarrow \infty$). Para lo cual se trabajará con la clase de reglas de paro cuya función de Lagrange truncada $L_{\mathcal{M}}^N(\psi)$ converja a la correspondiente función de Lagrange no truncada $L_{\mathcal{M}}(\psi)$; además se examinarán algunas condiciones que determinen cuándo se tendrá esta convergencia $L_{\mathcal{M}}^N(\psi) \rightarrow L_{\mathcal{M}}(\psi)$, cuando $N \rightarrow \infty$. A este caso se le conoce como horizonte infinito o caso no truncado.

Para cualquier regla de paro ψ se *define* la regla de paro truncada asociada a ella como ψ^{TN} (donde T indica el truncamiento en la etapa N) con $\psi_{\mathbf{n}} \equiv 1$, $\mathbf{n} \in \mathcal{G}^N$ y $N \geq 1$ fija. Por tanto toda regla de paro se puede truncar. Por lo cual a toda ψ se le puede aplicar la función de Lagrange truncada (2.37), *i.e.* para cualquier ψ y para cualquier $N \geq 1$ entero fijo se tiene:

$$L_{\mathcal{M}}^N(\psi) := L_{\mathcal{M}}^N(\psi^{TN}) \quad (2.90)$$

Por otro lado es importante mencionar que como nos encontramos en el caso no truncado y el interés en este contexto es analizar el proceso, cuando el número de etapas (o grupos) crece indefinidamente ($k \rightarrow \infty$), entonces es por ello que se consideraría que, a su vez, crezca el número de observaciones conforme crecen las etapas, porque, en general es posible que existan grupos de tamaño cero, entonces es trascendente no tener sucesiones de tamaños de grupos cuyas colas sean todas nulas.

Recordemos que estamos trabajando con los multiplicadores λ_0 y λ_1 fijos también en el proceso. Entonces en este contexto de horizonte infinito vamos a trabajar con todas las reglas de paro que cumplen la propiedad.

$$\lim_{N \rightarrow \infty} L_{\mathcal{M}}^N(\psi; \lambda_0, \lambda_1) = L_{\mathcal{M}}(\psi; \lambda_0, \lambda_1). \quad (2.91)$$

Sin embargo, es importante marcar que esta propiedad está en dependencia de la estructura del costo $c(\mathbf{n})$ de los grupos y de los multiplicadores λ 's. Por lo que definimos la siguiente clase.

Definición 19.

Para $\lambda_0 > 0$ y $\lambda_1 > 0$ fijas, se define la clase

$$\mathcal{T}_{\mathcal{M}} = \mathcal{T}_{\mathcal{M}}(\lambda_0, \lambda_1) := \{\psi \in \mathcal{F}_{\mathcal{M}} : \psi \text{ satisface (2.91)}\}. \quad (2.92)$$

Esta clase la llamaremos **clase de reglas de paro truncables**. Por definición se tiene que $\mathcal{T}_{\mathcal{M}} \subset \mathcal{F}_{\mathcal{M}}$.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Los siguientes lemas nos darán condiciones suficientes y/o necesarias para que una regla de paro sea truncable (i.e. pertenezca a la clase $\mathcal{T}_{\mathcal{M}}$).

Lema 6.

Sea $k \geq 1$ un entero. Para que una regla de paro $\psi \in \mathcal{T}_{\mathcal{M}}$ es suficiente que para cualesquiera $\lambda_0 > 0$, $\lambda_1 > 0$ se cumpla

$$\lim_{k \rightarrow \infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} \min \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \} d\mu^{\mathbf{n}} = 0. \quad (2.93)$$

Además, para que se cumpla (2.93) es necesario que $\psi \in \mathcal{T}_{\mathcal{M}}$ y $K_{\mathcal{M}}(\psi) < \infty$.

Demostración del Lema 6:

Suficiencia: Supongamos que (2.93) se cumple para cualesquiera $\lambda_0 > 0$ y $\lambda_1 > 0$. Se demostrará que $\psi \in \mathcal{T}_{\mathcal{M}}$.

Supongamos que $L_{\mathcal{M}}(\psi) < \infty$ entonces $K_{\mathcal{M}}(\psi) < \infty$. De esto se sigue que $\psi \in \mathcal{F}_{\mathcal{M}}$, i.e. $P(\tau^{\psi} < \infty) = 1$. Entonces, por (2.34)-(2.33) y (2.37),

$$L_{\mathcal{M}}(\psi) - L_{\mathcal{M}}^N(\psi) = \sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} \int p_{\mathbf{n}} s_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} - \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}. \quad (2.94)$$

El primer sumando de la ecuación anterior converge a cero, cuando $N \rightarrow \infty$, por ser la cola de una serie convergente (ya que $L_{\mathcal{M}}(\psi) < \infty$).

Ahora analicemos el segundo término derecho de (2.94):

$$\sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} = \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} c(\mathbf{n}) \int t_{\mathbf{n}}^{\psi} f^{\mathbf{n}} d\mu^{\mathbf{n}}(x^{\mathbf{n}}) + \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.95)$$

Por hipótesis se cumple (2.93), entonces el segundo término de la parte derecha de (2.95)

$$\sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} l_{\mathbf{n}} d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \rightarrow 0, \text{ cuando } N \rightarrow \infty.$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Por otro lado, dado $\mathbf{n} \in \mathcal{G}^k$ con $k = 1, 2, \dots$

$$\begin{aligned}
\sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E s_{\mathbf{n}}^{\psi} &\geq \sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} c(\mathbf{n} : N) p_{\mathbf{n}} E s_{\mathbf{n}}^{\psi} && \text{ya que } c(\mathbf{n}) \geq c(\mathbf{n} : N), \\
&= \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} E s_{\mathbf{n}}^{\psi} + \sum_{k=N+1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n} : N) E s_{\mathbf{n}}^{\psi} \\
&= \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} E s_{\mathbf{n}}^{\psi} + \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} \sum_{k=1}^{\infty} \sum_{m \in \mathcal{G}^k} p_m E s_{\mathbf{n},m}^{\psi} \\
&= \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} E s_{\mathbf{n}}^{\psi} + \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} P(\tau^{\psi} > N | n_1, n_2, \dots, n_N) \\
&= \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} E t_{\mathbf{n}}^{\psi} = \sum_{\mathbf{n} \in \mathcal{G}^N} c(\mathbf{n}) p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} f^{\mathbf{n}} d\mu^{\mathbf{n}}(x^{\mathbf{n}}).
\end{aligned}$$

Se sabe que el lado izquierdo de la desigualdad anterior converge a cero cuando $N \rightarrow \infty$ (por ser cola de $K_{\mathcal{M}}(\psi) < \infty$). Por tanto el primer término del lado derecho de (2.95) también es cero cuando $N \rightarrow \infty$.

$$\text{Por lo tanto } L_{\mathcal{M}}(\psi) - L_{\mathcal{M}}^N(\psi) \xrightarrow{N \rightarrow \infty} 0.$$

Ahora supongamos que $L_{\mathcal{M}}(\psi) = \infty$, i.e.

$$L_{\mathcal{M}}(\psi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) = \infty$$

Entonces de (2.37) se tiene que

$$L_{\mathcal{M}}^N(\psi) \geq \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} \xrightarrow{N \rightarrow \infty} \infty.$$

Por lo tanto, $\lim_{N \rightarrow \infty} L_{\mathcal{M}}^N(\psi; \lambda_0, \lambda_1) = L_{\mathcal{M}}(\psi; \lambda_0, \lambda_1)$, para cualesquiera $\lambda_0 > 0$ y $\lambda_1 > 0$. Por tanto $\psi \in \mathcal{T}_{\mathcal{M}}$.

Necesidad: Supongamos que $\psi \in \mathcal{T}_{\mathcal{M}}$ y que $K_{\mathcal{M}}(\psi) < \infty$.

Si $K_{\mathcal{M}}(\psi) < \infty$, se sigue que $L_{\mathcal{M}}(\psi) < \infty$, ya que $\alpha(\psi, \phi), \beta(\psi, \phi) \in [0, 1]$, y $\lambda_0, \lambda_1 < \infty$. Además, por hipótesis $\lim_{N \rightarrow \infty} L_{\mathcal{M}}^N(\psi; \lambda_0, \lambda_1) = L_{\mathcal{M}}(\psi; \lambda_0, \lambda_1) \implies L_{\mathcal{M}}(\psi) - L_{\mathcal{M}}^N(\psi) \xrightarrow{N \rightarrow \infty} 0$.

Entonces

$$\sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} \int p_{\mathbf{n}} s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} - \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}} \xrightarrow{N \rightarrow \infty} 0.$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Dado que el primer término anterior es convergente a cero, cuando $N \rightarrow \infty$ porque $L_{\mathcal{M}}(\psi) < \infty$, entonces el segundo también es convergente y por tanto $\sum_{n \in \mathcal{G}^N} p_n \int t_n^\psi l_n d\mu^n \xrightarrow{N \rightarrow \infty} 0$. ■

Definición 20.

Sean $k \geq 1$ y $\mathbf{n} \in \mathcal{G}^k$ cualesquiera se define la siguiente clase:

$$\mathcal{K}^{\mathbf{n}} := \left\{ x^{\mathbf{n}} = (x_1^{\mathbf{n}}, \dots, x_k^{\mathbf{n}}) \in \mathcal{X}^{\mathbf{n}} : f^{\mathbf{n}}(x^{\mathbf{n}}) > 0 \right\}. \quad (2.96)$$

En el caso de trabajar con las funciones $f_0^{\mathbf{n}}$ o $f_1^{\mathbf{n}}$ se definirán las clases $\mathcal{K}_0^{\mathbf{n}}$ o $\mathcal{K}_1^{\mathbf{n}}$, respectivamente.

Definición 21.

Para $\mathbf{n} \in \mathcal{G}^k$ y $k \geq 1$ se define la *razón*

$$Z_{\mathbf{n}} = \frac{f_1^{\mathbf{n}}(X_1^{\mathbf{n}}, \dots, X_k^{\mathbf{n}})}{f_0^{\mathbf{n}}(X_1^{\mathbf{n}}, \dots, X_k^{\mathbf{n}})} = \frac{f_1^{\mathbf{n}}}{f_0^{\mathbf{n}}}. \quad (2.97)$$

Lema 7.

Una regla $\psi \in \mathcal{T}_{\mathcal{M}}$, si cualesquiera de las siguientes condiciones se cumple:

(i) Para cualquier regla de paro $\psi \in \mathcal{F}_{\mathcal{M}}$, se tiene que

$$P_0(\tau^\psi < \infty) = 1 \quad \text{o} \quad P_1(\tau^\psi < \infty) = 1. \quad (2.98)$$

(ii) Para cualquier $c > 0$

$$\sum_{n \in \mathcal{G}^k} p_n P_0 \left(\frac{f_1^n(X_1^n, \dots, X_k^n)}{f_0^n(X_1^n, \dots, X_k^n)} > c \right) \rightarrow 0, \quad \text{cuando } k \rightarrow \infty. \quad (2.99)$$

(iii) Para algunos $\lambda_0 > 0$ y $\lambda_1 > 0$, se cumple

$$\lim_{k \rightarrow \infty} \sum_{n \in \mathcal{G}^k} p_n \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n(x^n) = 0. \quad (2.100)$$

Demostración del Lema 7:

(i) Sea $\psi \in \mathcal{F}_{\mathcal{M}}$ cualquier regla de paro, supongamos que para algún $i = 0, 1$ se tiene que $P_i(\tau^\psi < \infty) = 1$.

Sea $k \geq 1$ cualquiera, entonces para dicho i se cumple que:

$$\begin{aligned} \sum_{n \in \mathcal{G}^k} p_n \int t_n^\psi \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n &\leq \lambda_i \sum_{n \in \mathcal{G}^k} p_n \int t_n^\psi f_i^n d\mu^n(x^n) = \\ &= \lambda_i P_i(\tau^\psi \geq k) \xrightarrow[k \rightarrow \infty]{} 0. \end{aligned}$$

Por tanto, se cumple la condición (2.93) y por Lema 6 se sigue que $\psi \in \mathcal{T}_{\mathcal{M}}$.

(ii) Supongamos que para cualquier $c > 0$, $\sum_{n \in \mathcal{G}^k} p_n P_0(Z_n > c) \xrightarrow[k \rightarrow \infty]{} 0$. Entonces para cualesquiera $k \geq 1$, $\lambda_0 > 0$ y $\lambda_1 > 0$, fijos,

$$\begin{aligned} \sum_{n \in \mathcal{G}^k} p_n \int t_n^\psi \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n &\leq \sum_{n \in \mathcal{G}^k} p_n \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n, \quad \text{porque } t_n^\psi \leq 1, \\ &= \sum_{n \in \mathcal{G}^k} p_n \int_{\mathcal{K}_0^n} \min \left\{ \lambda_0, \lambda_1 \frac{f_1^n}{f_0^n} \right\} f_0^n d\mu^n(x^n) = \sum_{n \in \mathcal{G}^k} p_n E_0 \min \{ \lambda_0, \lambda_1 Z_n \} \xrightarrow[k \rightarrow \infty]{} 0. \end{aligned}$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Esto último se sigue de lo siguiente: $x \mapsto \min\{\lambda_0, \lambda_1 x\}$, con $x \geq 0$, es una función real continua, no decreciente, no negativa y, además, acotada por λ_0 ; entonces ¹

$$\begin{aligned} \sum_{n \in \mathcal{G}^k} p_n Z_n \xrightarrow{P_0} 0, & \implies \sum_{n \in \mathcal{G}^k} p_n \min\{\lambda_0, \lambda_1 Z_n\} \xrightarrow{P_0} 0, \quad \text{cuando } k \rightarrow \infty. \\ \implies \sum_{n \in \mathcal{G}^k} p_n \min\{\lambda_0, \lambda_1 Z_n\} \xrightarrow{d} 0 & \implies \sum_{n \in \mathcal{G}^k} p_n E_0 \min\{\lambda_0, \lambda_1 Z_n\} \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Por tanto, nuevamente, por Lema 6 se sigue que $\psi \in \mathcal{T}_{\mathcal{M}}$.

(iii) Sea $\psi \in \mathcal{F}_{\mathcal{M}}$ tal que $K_{\mathcal{M}}(\psi) < \infty$ y supongamos que se cumple (2.100) para algunos $\lambda_0 > 0$ y $\lambda_1 > 0$. Entonces para estos λ 's se tiene:

$$\begin{aligned} \sum_{n \in \mathcal{G}^k} p_n E_0 \min\{\lambda_0, \lambda_1 Z_n\} &= \sum_{n \in \mathcal{G}^k} p_n \int_{\mathcal{K}_0^n} \min\left\{\lambda_0, \lambda_1 \frac{f_1^n}{f_0^n}\right\} f_0^n d\mu^n(x^n), \\ &\leq \sum_{n \in \mathcal{G}^k} p_n \int \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\} d\mu^n(x^n) \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Ahora por la desigualdad de Chebyshev se tiene que para cualquier $c > 0$:

$$P_0\left(\min\{\lambda_0, \lambda_1 Z_n\} > c\right) \leq \frac{E_0 \min\{\lambda_0, \lambda_1 Z_n\}}{c} \xrightarrow{k \rightarrow \infty} 0.$$

De lo cual, $\min\{\lambda_0, \lambda_1 Z_n\} \xrightarrow{P_0} 0$ cuando $k \rightarrow \infty$. Ahora sea $0 < \epsilon < \lambda_0$ cualquiera:

$$\begin{aligned} \sum_{n \in \mathcal{G}^k} p_n P_0\left(Z_n > \frac{\epsilon}{\lambda_1}\right) &= \sum_{n \in \mathcal{G}^k} p_n P_0(\lambda_0 > \epsilon, \lambda_1 Z_n > \epsilon) \\ &= \sum_{n \in \mathcal{G}^k} p_n P_0\left(\min\{\lambda_0, \lambda_1 Z_n\} > \epsilon\right) \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

En el caso de que $\epsilon \geq \lambda_0 > \epsilon$ entonces $P(Z_n > \epsilon) \leq P(Z_n > \epsilon)$ y también se obtiene la convergencia anterior.

Por lo tanto para cualquier $\epsilon > 0$ se tiene:

$$\sum_{n \in \mathcal{G}^k} p_n P_0(Z_n > \epsilon) \rightarrow 0 \quad \text{cuando } k \rightarrow \infty.$$

Ahora por condición (ii), (2.99), de este lema se sigue que $\psi \in \mathcal{T}_{\mathcal{M}}$. ■

¹Recordemos: una sucesión $\{Z_n\}$ converge a Z en probabilidad P (denotado $Z_n \xrightarrow{P} Z$), si $\forall \epsilon > 0$, $P(|Z_n - Z| > \epsilon) \rightarrow 0$, cuando $n \rightarrow \infty$.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

El siguiente resultado indica que la esperanza del mínimo de ciertas funciones es menor cuando el número de observaciones es mayor.

Lema 8.

Sean $m \in \mathbb{N}$ y cualesquiera $\lambda_0 > 0$, $\lambda_1 > 0$ entonces

$$E_0 \min \{ \lambda_0, \lambda_1 Z_{m+1} \} \leq E_0 \min \{ \lambda_0, \lambda_1 Z_m \}. \quad (2.101)$$

Demostración Lema 8:

Sea $m \in \mathbb{N}$ cualquiera, sabemos que la función $x \mapsto \min \{ \lambda_0, \lambda_1 x \}$ con $x \geq 0$ es cóncava y continua, entonces utilizando propiedades de la esperanza condicional y la desigualdad de Jensen $E[g(X)|Y] \leq g[E(X|Y)]$ se tiene:

$$\begin{aligned} E_0 \min \{ \lambda_0, \lambda_1 Z_{m+1} \} &= E_0 \left[E_0 \min \{ \lambda_0, \lambda_1 Z_{m+1} \mid X_1, \dots, X_m \} \right] \\ &\leq E_0 \min \left\{ \lambda_0, \lambda_1 E_0(Z_{m+1} \mid X_1, \dots, X_m) \right\} \\ &\leq E_0 \min \{ \lambda_0, \lambda_1 Z_m \} \end{aligned}$$

La última desigualdad sucede porque:

$$\begin{aligned} E_0(Z_{m+1} \mid X_1, \dots, X_m) &= \int_{\mathcal{K}_0} \frac{f_1^{m+1}(x_1, \dots, x_{m+1})}{f_0^{m+1}(x_1, \dots, x_{m+1})} f_0^{m+1}(x_{m+1} \mid x_1, \dots, x_m) d\mu(x_{m+1}) \\ &= \int_{\mathcal{K}_0} \frac{f_1^{m+1}(x_1, \dots, x_{m+1})}{f_0^{m+1}(x_1, \dots, x_{m+1})} \frac{f_0^{m+1}(x_1, \dots, x_{m+1})}{f_0^m(x_1, \dots, x_m)} d\mu(x_{m+1}) \\ &= \int_{\mathcal{K}_0} \frac{f_1^{m+1}(x_1, \dots, x_{m+1})}{f_0^m(x_1, \dots, x_m)} d\mu(x_{m+1}) \leq \int_{\mathcal{X}} \frac{f_1^{m+1}(x_1, \dots, x_{m+1})}{f_0^m(x_1, \dots, x_m)} d\mu(x_{m+1}); \\ &= \frac{1}{f_0^m(x_1, \dots, x_m)} \int_{\mathcal{X}} f_1^{m+1}(x_1, \dots, x_{m+1}) d\mu(x_{m+1}) \\ &= \frac{f_1^m(x_1, \dots, x_m)}{f_0^m(x_1, \dots, x_m)} = Z_m. \end{aligned}$$

Por lo tanto, $E_0 \{ Z_{m+1} \mid X_1, \dots, X_m \} \leq Z_m$. ■

El siguiente lema muestra la equivalencia entre dos condiciones muy útiles en la estadística matemática. Esta equivalencia, implícitamente, indica que *conforme se incrementa el número*

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

de observaciones en un experimento ($n \rightarrow \infty$), y alguna de las probabilidades de error tipo I o II tiende a cero (ponderadas por λ_0 y λ_1 respectivamente), entonces la otra probabilidad también tenderá a cero; esto es debido a que mientras más observaciones se tengan, más precisión en el experimento habrá respecto a la decisión sobre éste, y por ende menor probabilidad de cometer error en la decisión tomada.

Lema 9.

Las siguientes condiciones son equivalentes:

(i) Para algunos $\lambda_0 > 0$ y $\lambda_1 > 0$, se cumple

$$\lim_{n \rightarrow \infty} \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n = 0, \quad (2.102)$$

(ii) Para cualquier $\epsilon > 0$

$$P_0 \left(\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} > \epsilon \right) \rightarrow 0, \quad \text{cuando } n \rightarrow \infty. \quad (2.103)$$

Demostración Lema 9:

- (i) \implies (ii) : Supongamos que se cumple (2.102) para algunos $\lambda_0 > 0$ y $\lambda_1 > 0$. Entonces para estos λ 's se tiene:

$$E_0 \min \{ \lambda_0, \lambda_1 Z_n \} = \int_{\mathcal{K}_0} \min \left\{ \lambda_0, \lambda_1 \frac{f_1^n}{f_0^n} \right\} f_0^n d\mu^n = \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n(x^n) \xrightarrow{n \rightarrow \infty} 0.$$

Ahora sea $0 < \epsilon < \lambda_0$ cualquiera, por la desigualdad de Markov se tiene que:

$$P_0 \left(Z_n > \frac{\epsilon}{\lambda_1} \right) \leq \frac{E_0 \min \{ \lambda_0, \lambda_1 Z_n \}}{\min \{ \lambda_0, \lambda_1 x \}} \xrightarrow{n \rightarrow \infty} 0.$$

Entonces, para cualquier $\epsilon > 0$

$$P_0(Z_n > \epsilon) \rightarrow 0 \quad \text{cuando } n \rightarrow \infty.$$

Por lo tanto $Z_n \xrightarrow{P_0} 0$, cuando $n \rightarrow \infty$.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

- (ii) \implies (i) : Supongamos que se cumple (2.103), entonces

$$\begin{aligned} Z_n \xrightarrow{P_0} 0, & \implies \min \{ \lambda_0, \lambda_1 Z_n \} \xrightarrow{P_0} 0, \quad \text{cuando } n \rightarrow \infty. \\ \implies \min \{ \lambda_0, \lambda_1 Z_n \} \xrightarrow{d} 0 & \implies E_0 \min \{ \lambda_0, \lambda_1 Z_n \} \rightarrow 0, \quad \text{cuando } n \rightarrow \infty. \end{aligned}$$

Esto último porque $x \mapsto \min \{ \lambda_0, \lambda_1 x \}$ es una función real, continua para toda $x \geq 0$ y acotada. Si $Z_n \xrightarrow{P} Z$ (convergencia en probabilidad) y ρ es una función continua en \mathbb{R} , entonces $\rho(Z_n) \xrightarrow{P} \rho(Z)$. Por otro lado, se sabe también que $Z_n \xrightarrow{P} Z \implies Z_n \xrightarrow{d} Z$ (ésta es convergencia en distribución); y $Z_n \xrightarrow{d} Z$ si y solo si $E\rho(Z_n) \rightarrow E\rho(Z)$ para toda función ρ real, continua y acotada.

Por lo tanto, $E_0 \min \{ \lambda_0, \lambda_1 Z_n \} = \int_{\mathcal{K}_0} \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n(x^n) \xrightarrow{n \rightarrow \infty} 0$. ■

Corolario 6.

Supongamos que:

$$\nu_1 + \nu_2 + \cdots + \nu_k \rightarrow \infty \quad \text{casi seguramente cuando } k \rightarrow \infty. \quad (2.104)$$

Si se cumple alguna de las siguientes condiciones:

- (i) Para algunos $\lambda_0 > 0$ y $\lambda_1 > 0$ se tiene

$$\lim_{n \rightarrow \infty} \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n = 0. \quad (2.105)$$

- (ii) Para cualquier $c > 0$,

$$P_0 \left(\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} > c \right) \rightarrow 0, \quad \text{cuando } n \rightarrow \infty. \quad (2.106)$$

Entonces $\psi \in \mathcal{T}_{\mathcal{M}}$.

Demostración Corolario 6:

Supongamos que se tiene (2.104).

- (i) Sean algunos $\lambda_0 > 0$ y $\lambda_1 > 0$ fijos tales que se cumple (2.105). Sea $\varepsilon > 0$ cualquiera,

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

aplicando probabilidad condicional y total se obtiene:

$$\begin{aligned}
& \sum_{n \in \mathcal{G}^k} p_n \int \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} d\mu^n(x^n) = \sum_{n \in \mathcal{G}^k} p_n \int_{\mathcal{K}_0^n} \min \left\{ \lambda_0, \lambda_1 \frac{f_1^n}{f_0^n} \right\} f_0^n d\mu^n(x^n), \\
& = E_0 \min \left\{ \lambda_0, \lambda_1 \frac{f_1^{\nu_1, \dots, \nu_k}}{f_0^{\nu_1, \dots, \nu_k}} \right\} = E_0 \min \{ \lambda_0, \lambda_1 Z_{\nu_1, \dots, \nu_k} \} \\
& = E_0 \min \{ \lambda_0, \lambda_1 Z_{\nu_1, \dots, \nu_k} \} I_{\{\nu_1 + \dots + \nu_k < m\}} + E_0 \min \{ \lambda_0, \lambda_1 Z_{\nu_1, \dots, \nu_k} \} I_{\{\nu_1 + \dots + \nu_k \geq m\}} \\
& \leq \lambda_0 P_0(\nu_1 + \dots + \nu_k < m) + \sum_{n \in \mathcal{G}^k, n_1 + \dots + n_k \geq m} p_n E_0 \min \{ \lambda_0, \lambda_1 Z_n \} I_{\{n_1 + \dots + n_k \geq m\}} \\
& \leq \lambda_0 P(\nu_1 + \dots + \nu_k \leq m) + E_0 \min \{ \lambda_0, \lambda_1 Z_m \} \quad \text{por Lema 8} \\
& = \lambda_0 P(\nu_1 + \dots + \nu_k \leq m) + \int \min \{ \lambda_0 f_0^m, \lambda_1 f_1^m \} d\mu^m(x^m) < 2\varepsilon
\end{aligned}$$

En el último paso sea m fija tan grande que $\int \min \{ \lambda_0 f_0^m, \lambda_1 f_1^m \} d\mu^m(x^m) < \varepsilon$, esto es posible por la hipótesis (2.105); y con dicha m fija sea k tan grande tal que $P(\nu_1 + \dots + \nu_k \leq m) < \frac{\varepsilon}{\lambda_0}$, esto por (2.104). Por tanto, por Lema (7)-(iii) se sigue que $\psi \in \mathcal{T}_{\mathcal{M}}$.

(ii) Supongamos que se tiene (2.106), entonces por Lema 9 se sigue (2.105) y de lo cual $\psi \in \mathcal{T}_{\mathcal{M}}$. ■

Debido a la importancia del siguiente resultado, vamos a cambiar momentáneamente un poco el contexto del problema de pruebas de este capítulo, y vamos a introducir ahora como base a **un problema de pruebas de hipótesis Bayesiano**, denotado mediante la terna:

$$\mathcal{M}_B = (P_0, P_1, \pi P_0 + (1 - \pi)P_1), \quad (2.107)$$

con $\pi \in [0, 1]$ fijo. En esta terna, la tercera distribución es una combinación convexa de las distribuciones hipotetizadas en H_0 y H_1 , esto es $P = \pi P_0 + (1 - \pi)P_1$ con $\pi \in [0, 1]$ fijo. En general, cuando $\pi \in (0, 1)$, el problema de prueba está asociado al contexto del riesgo Bayesiano R , donde R se puede asociar con nuestra función de Lagrange, bajo la tercera distribución dada en (2.107). Pero cuando $\pi = 0$ o $\pi = 1$, le llamaremos problema Bayesiano degenerado, porque en estos casos la tercera distribución coincide con alguna de las distribuciones hipotetizadas. De hecho, si $\pi = 0$, coincide con $P = P_1$ y podríamos trabajar con la terna $\mathcal{M}_1 = (P_0, P_1, P_1)$, o si $\pi = 1$, coincide con $P = P_0$ y tendríamos la terna $\mathcal{M}_0 = (P_0, P_1, P_0)$.

Entonces para este problema de pruebas Bayesiano \mathcal{M}_B , también se definen las clases:

$$\mathcal{F}_B = \{ \psi : (\pi P_0 + (1 - \pi)P_1)(\tau^\psi < \infty) = 1 \}, \quad y \quad (2.108)$$

$$\mathcal{T}_B = \{ \psi \in \mathcal{F}_B : L_B^N(\psi) \rightarrow L_B(\psi), \text{ cuando } N \rightarrow \infty \}. \quad (2.109)$$

Considerando que la función de Lagrange es $L_B(\psi) = K_B(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi)$, para algunas $\lambda_0 > 0$ y $\lambda_1 > 0$ y fijas. En esta función L_B se considera al costo promedio total K_B del

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

experimento bajo la distribución $\pi P_0 + (1 - \pi)P_1$, y a su vez $L_B^N(\psi)$ es la función de Lagrange truncada que tiene al costo total promedio y las probabilidades de error solo para N etapas.

El siguiente resultado establece que toda regla de paro ψ con un tiempo de paro finito en un **problema de pruebas de hipótesis Bayesiano \mathcal{M}_B es truncable** ($\psi \in \mathcal{T}_B$), y viceversa.

Corolario 7.

Si \mathcal{M}_B es un problema de pruebas de hipótesis Bayesiano, representado en (2.107), entonces $\mathcal{T}_B = \mathcal{F}_B$.

Demostración de Corolario 7.

Sea $\mathcal{M}_B = (P_0, P_1, \pi P_0 + (1 - \pi)P_1)$ un problema de pruebas de hipótesis Bayesiano con $\pi \in [0, 1]$ fijo. De la definición de \mathcal{T}_B (2.109) se sigue que $\mathcal{T}_B \subseteq \mathcal{F}_B$.

Por otro lado, si $\psi \in \mathcal{F}_B$, entonces $(\pi P_0 + (1 - \pi)P_1)(\tau^\psi < \infty) = 1$. De lo cual se tiene que:

$$\pi P_0(\tau^\psi < \infty) + (1 - \pi)P_1(\tau^\psi < \infty) = 1. \quad (2.110)$$

De (2.110) sigue que para cualquier $\pi \in [0, 1]$, al menos se cumple una de las siguientes afirmaciones:

$$P_0(\tau^\psi < \infty) = 1 \quad \text{o} \quad P_1(\tau^\psi < \infty) = 1.$$

Por tanto, por Lema (7)-(i), se tiene que $\psi \in \mathcal{T}_B$ y por lo tanto $\mathcal{F}_B = \mathcal{T}_B$. ■

Para cerrar con el resultado del corolario anterior, en éste también se marca, de manera implícita, que en los problemas de pruebas de hipótesis \mathcal{M}_0 y \mathcal{M}_1 (casos Bayesianos “degenerados”) se tendrá, respectivamente, que $\mathcal{T}_0 = \mathcal{F}_0$ y $\mathcal{T}_1 = \mathcal{F}_1$. Estos dos resultados son importantes porque en uno de los siguientes capítulos precisamente se trabajarán a las observaciones de los grupos, bajo las distribuciones P_0 y/o P_1 . Y para estos contextos todas las reglas ψ con tiempo de paro finito, bajo la distribución respectiva, serán truncables, y viceversa. Resultado que en el contexto \mathcal{M} no necesariamente es cierto, ya que se pueden tener reglas de paro $\psi \in \mathcal{F}$, pero que no son truncables ($\psi \notin \mathcal{T}_\mathcal{M}$).

Nuevamente regresamos a nuestro contexto general con un problema de pruebas de hipótesis simples representado mediante $\mathcal{M} = (P_0, P_1, P)$.

Observación: En todo lo subsiguiente, de esta sección, vamos a trabajar con reglas de paro truncables, i.e. $\psi \in \mathcal{T}_\mathcal{M}$, a menos que se diga lo contrario. En general vamos a suponer este hecho y ya no será marcado en cada resultado.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Ahora sí vamos a caracterizar a la familia de reglas de paro óptimas, para lo cual primero vamos a buscar la cota inferior de la función de Lagrange $L_{\mathcal{M}}(\psi)$ y las condiciones (si las hay) para alcanzar dicha cota.

Tomemos en cuenta que debido a que una familia $v \in \mathcal{V}_r$ está caracterizada por sus elementos, entonces la “comparación” entre dos familias $v_1, v_2 \in \mathcal{V}_r$ será a nivel de sus elementos. De la misma manera se definirá la “convergencia”.

Definición 22.

Sean $r \geq 1$, $N \geq 1$ y sean $v_1, v_2, v^N, v \in \mathcal{V}_r$,

1. Se define $v_1 \leq v_2$, si para cada $\mathbf{n} \in \mathcal{G}^r$ se cumple que

$$(v_1)_{\mathbf{n}} \leq (v_2)_{\mathbf{n}} \quad \mu^{\mathbf{n}}\text{-c.t.p.} \quad (2.111)$$

2. Se define $v^N \xrightarrow{N \rightarrow \infty} v$, si para cada $\mathbf{n} \in \mathcal{G}^r$ se cumple que

$$(v^N)_{\mathbf{n}} \xrightarrow{N \rightarrow \infty} (v)_{\mathbf{n}} \quad \mu^{\mathbf{n}}\text{-c.t.p.} \quad (2.112)$$

El siguiente lema muestra el comportamiento de las familias V_r^N , en particular se muestra que dichas familias son decrecientes conforme aumenta el nivel de truncación N . Es decir, si se tienen más etapas, estas familias V_r^N son iguales o más pequeñas que en una etapa anterior.

Lema 10.

Sea $N \geq 1$ fija, para cada $1 \leq r \leq N$ se tiene:

$$V_r^N \geq V_r^{N+1}. \quad (2.113)$$

Demostración del Lema 10.

Sea $N \geq 1$ fija, se utilizará inducción inversa sobre $r = N, N-1, \dots, 1$.²

Sea $r = N$, por definición (2.60) se tiene que $V_N^{N+1} = M_N V_{N+1}^{N+1}$. Sea $\mathbf{n} \in \mathcal{G}^N$ entonces

$$(V_N^{N+1})_{\mathbf{n}} = (M_{N+1} V_{N+1}^{N+1})_{\mathbf{n}} = \min \{l_{\mathbf{n}}, \bar{c}_{N+1} f^{\mathbf{n}} + (I_{N+1} V_{N+1}^{N+1})_{\mathbf{n}}\} \leq l_{\mathbf{n}} = (V_N^N)_{\mathbf{n}}.$$

²La inducción inversa consiste demostrar la propiedad para $r = N$, después se supone la verdad de la propiedad para $r = k$ y se demuestra para $r = k-1$; demostrando así la veracidad de la propiedad para todos los enteros desde $r = N$ hasta 1.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Por lo cual, para toda $\mathbf{n} \in \mathcal{G}^N$ $(V_N^{N+1})_{\mathbf{n}} \leq (V_N^N)_{\mathbf{n}}$, $\mu^{\mathbf{n}}$ -c.t.p.

$$\therefore \text{ por (2.111) } \quad V_N^{N+1} \leq V_N^N.$$

Hipótesis de Inducción: Supongamos que para cada $1 < r \leq N$ con $\mathbf{n} \in \mathcal{G}^r$ se cumple que

$$(V_r^N)_{\mathbf{n}} \geq (V_r^{N+1})_{\mathbf{n}}. \quad (2.114)$$

Sea $\mathbf{n} \in \mathcal{G}^{r-1}$ cualquiera entonces por (2.60) y (2.114) se tiene:

$$\begin{aligned} (V_{r-1}^N)_{\mathbf{n}} = (M_r V_r^N)_{\mathbf{n}} &= \min \left\{ l_{\mathbf{n}}, \bar{c}_r f^{\mathbf{n}} + (I_r V_r^N)_{\mathbf{n}} \right\} \\ &\geq \min \left\{ l_{\mathbf{n}}, \bar{c}_r f^{\mathbf{n}} + (I_r V_r^{N+1})_{\mathbf{n}} \right\} = (M_r V_r^{N+1})_{\mathbf{n}} = (V_{r-1}^{N+1})_{\mathbf{n}}. \end{aligned}$$

Lo anterior se tiene $\mu^{\mathbf{n}}$ -c.t.p. De lo cual $V_{r-1}^{N+1} \leq V_{r-1}^N$.

Por lo tanto para toda $1 \leq r \leq N$ se tiene (2.113). ■

Con base en el Lema 10 podemos decir que la sucesión $\{V_r^N\}_N$ es decreciente con respecto al nivel de truncamiento N , además la sucesión es acotada ($V_r^N > 0$). Sea $V_r \in \mathcal{V}_r$ tal que:

$$V_r^N \xrightarrow{N \rightarrow \infty} V_r. \quad (2.115)$$

Lema 11.

Para cualquier $r \geq 1$ se cumple que:

$$V_r = M_{r+1} V_{r+1}. \quad (2.116)$$

Demostración del Lema 11:

Por Definición (2.60), se tiene que $V_r^N := M_{r+1} V_{r+1}^N$ para cualquier $r = N - 1, \dots, 1$. Por Lema 10

$$(V_{r+1}^N)_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) \geq (V_{r+1}^{N+1})_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m), \quad \mu^{(\mathbf{n}, m)}\text{-c.t.p.}$$

Entonces para $1 \leq r \leq N - 1$, para cualquier $m > 1$ y cada $\mathbf{n} \in \mathcal{G}^r$ las funciones $(V_{r+1}^N)_{(\mathbf{n}, m)}$ están acotadas por funciones integrables, entonces por (2.115) y aplicando el Teorema de Convergencia Dominada (TCD):

$$\int_{\mathcal{X}^{(\mathbf{n}, m)}} (V_{r+1}^N)_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) d\mu^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) \xrightarrow{N \rightarrow \infty} \int_{\mathcal{X}^{(\mathbf{n}, m)}} (V_{r+1})_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) d\mu^{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m). \quad (2.117)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Ahora por Teorema de Fubini y utilizando la desigualdad de Markov, sea $\varepsilon > 0$ cualquiera

$$\begin{aligned} & \mu^n \left\{ x^n : \left(\int_{\mathcal{X}^m} (V_{r+1}^N)_{(n,m)}(x^n, x^m) d\mu^m(x^m) - \int_{\mathcal{X}^m} (V_{r+1})_{(n,m)}(x^n, x^m) d\mu^m(x^m) \right) > \varepsilon \right\} \leq \\ & \leq \frac{1}{\varepsilon} \int_{\mathcal{X}^n} \left(\int_{\mathcal{X}^m} (V_{r+1}^N)_{(n,m)}(x^n, x^m) d\mu^m(x^m) - \int_{\mathcal{X}^m} (V_{r+1})_{(n,m)}(x^n, x^m) d\mu^m(x^m) \right) d\mu^n \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

La convergencia se tiene por (2.117). Por tanto,

$$\int (V_{r+1}^N)_{(n,m)}(x^n, x^m) d\mu^m(x^m) \xrightarrow{N \rightarrow \infty} \int (V_{r+1})_{(n,m)}(x^n, x^m) d\mu^m(x^m), \quad \mu^n\text{-c.t.p.}$$

También, como $\sum_{m \in \mathcal{G}} p_{r+1}(m) \leq 1$ y tomando la medida de conteo entonces para $\mathbf{n} \in \mathcal{G}^r$, por (2.115) y TCD:

$$\begin{aligned} (I_{r+1} V_{r+1}^N)_{\mathbf{n}} \xrightarrow{N \rightarrow \infty} (I_{r+1} V_{r+1})_{\mathbf{n}}, \quad \mu^n\text{-c.t.p.} & \implies (M_{r+1} V_{r+1}^N)_{\mathbf{n}} \xrightarrow{N \rightarrow \infty} (M_{r+1} V_{r+1})_{\mathbf{n}}, \quad \mu^n\text{-c.t.p.} \\ & \implies V_r^N = M_{r+1} V_{r+1}^N \xrightarrow{N \rightarrow \infty} M_{r+1} V_{r+1} = V_r. \end{aligned}$$

Lo cual demuestra (2.116). ■

Al igual que en el caso truncado, en esta sección se trabajará una condición llamada Π^∞ . Nuevamente, dicho nombre hace referencia a “paro” (regla), por lo que se utiliza la letra mayúscula griega “ Π ”, pero ahora con horizonte infinito (con superíndice “ ∞ ”) ya que se aplica a toda etapa $r \geq 1$.

Condición Π^∞ . (Regla de paro óptima)

Una regla de paro ψ **satisface la condición Π^∞** , si para toda $r \geq 1$ y cualquier $\mathbf{n} \in \mathcal{P}_r$:

$$\psi_{\mathbf{n}} \geq I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1} f^n + (I_{r+1} V_{r+1})_{\mathbf{n}}\}}, \quad (2.118)$$

sobre $T_{\mathbf{n}}^\psi \mu^n\text{-c.t.p.}$

Lema 12.

Para para cualquier $r \geq 1$ se cumple la siguiente desigualdad:

$$B_{r+1}^\psi(V_{r+1}) \geq B_r^\psi(V_r). \quad (2.119)$$

Además, se alcanza la igualdad en (2.119) si y solo si ψ satisface la condición Π^∞ .

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Demostración del Lema 12:

Sea $N > 1$ cualquiera, entonces por Corolario 5 se tiene que para cualquier $1 \leq r \leq N - 1$ se cumple:

$$B_{r+1}^\psi(V_{r+1}^N) \geq B_r^\psi(V_r^N). \quad (2.120)$$

donde V_r^N está definido en (2.60).

Ahora, por (2.115), (2.116) y TCD se tiene que $B_{r+1}^\psi(V_{r+1}^N) \rightarrow B_{r+1}^\psi(V_{r+1})$, cuando $N \rightarrow \infty$. Análogamente, $B_r^\psi(V_r^N) \rightarrow B_r^\psi(V_r)$, cuando $N \rightarrow \infty$. Por tanto se demuestra (2.119).

Finalmente, nuevamente por Corolario 5 y tomando el límite cuando $N \rightarrow \infty$, se alcanza la igualdad en (2.119) si y solo si

$$\psi_{\mathbf{n}} \stackrel{\geq}{=} I_{\{l_{\mathbf{n}} \lesssim \bar{c}_{r+1} f^{\mathbf{n}} + (I_{r+1} V_{r+1})_{\mathbf{n}}\}},$$

sobre el conjunto $T_{\mathbf{n}}^\psi \mu^{\mathbf{n}}$ -c.t.p., para cada $r \geq 1$ y para toda $\mathbf{n} \in \mathcal{P}_r$. Por lo tanto, es necesario y suficiente que ψ cumpla la condición $\mathbf{\Pi}^\infty$ para que se alcance la igualdad en (2.119). ■

Un caso particular de la desigualdad (2.62) es cuando $r = N - 1$, entonces debido a que se está suponiendo que las reglas ψ son truncables ($\psi \in \mathcal{T}_{\mathcal{M}}$) entonces tomando el límite cuando $N \rightarrow \infty$ para este caso se obtiene:

$$L_{\mathcal{M}}^N(\psi) = B_N^\psi(V_N^N) \geq B_{N-1}^\psi(V_{N-1}^N) \implies L_{\mathcal{M}}(\psi) \geq B_{N-1}^\psi(V_{N-1}).$$

Aplicando sucesivamente la desigualdad (2.119) se obtiene la siguiente cadena:

$$L_{\mathcal{M}}(\psi) \geq B_{N-1}^\psi(V_{N-1}) \geq \dots \geq B_2^\psi(V_2) \geq B_1(V_1). \quad (2.121)$$

De lo cual para cualquier regla de paro $\psi \in \mathcal{T}_{\mathcal{M}}$ su función de Lagrange está acotada inferiormente por:

$$L_{\mathcal{M}}(\psi) \geq B_1(V_1) = \bar{c}_1 + \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \int (V_1)_{\mathbf{n}} d\mu^{\mathbf{n}}. \quad (2.122)$$

Donde, además, dicha cota es independiente de la regla de paro. Asimismo, esta cota es finita, por (1.7) y porque de (2.22),

$$0 \leq \int (V_1)_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \int l_{\mathbf{n}}(x^{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}) \leq \min\{\lambda_0, \lambda_1\}.$$

También $\sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \leq 1$. Por lo tanto

$$B_1(V_1) < \infty. \quad (2.123)$$

La expresión del lado derecho de (2.122) es una cota menor finita de la función de Lagrange, mas el siguiente lema demuestra cuál es la mayor cota inferior de $L_{\mathcal{M}}(\psi)$. Después se demostrará cuáles reglas de paro alcanzan dicho ínfimo.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Lema 13.

$$\inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi) = B_1(V_1). \quad (2.124)$$

Demostración del Lema 13:

Sean:

$$U = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi), \quad U_N = \inf_{\psi \in \mathcal{F}^N} L_{\mathcal{M}}^N(\psi).$$

Por Teorema 3 se tiene que:

$$U_N = \bar{c}_1 + \sum_{n \in \mathcal{G}} p_n \int (V_1^N)_n(x^n) d\mu^n(x^n).$$

Debido a que $\mathcal{F}^N \subset \mathcal{F}_{\mathcal{M}}$ entonces $U_N \geq U$ para cualquier $N = 1, 2, \dots$, entonces

$$\lim_{N \rightarrow \infty} U_N \geq U. \quad (2.125)$$

Se demostrará que existe una igualdad en (2.125).

Supongamos que no hay igualdad en (2.125), *i.e.* supongamos que $\lim_{N \rightarrow \infty} U_N = U + 3\varepsilon$ para algún $\varepsilon > 0$. De aquí se sigue que para una N suficientemente grande

$$U_N \geq U + 3\varepsilon. \quad (2.126)$$

Por otro lado, por la definición de U existe una $\psi \in \mathcal{T}_{\mathcal{M}}$ tal que $U \leq L_{\mathcal{M}}(\psi) \leq U + \varepsilon$.

Dado que la regla es truncable, entonces $L_{\mathcal{M}}^N(\psi) \xrightarrow{N \rightarrow \infty} L_{\mathcal{M}}(\psi)$, entonces de aquí se tiene que para un N suficientemente grande:

$$L_{\mathcal{M}}^N(\psi) \leq U + 2\varepsilon. \quad (2.127)$$

Por definición de ínfimo $U_N \leq L_{\mathcal{M}}^N(\psi)$, de donde se tiene que para N suficientemente grande

$$U_N \leq U + 2\varepsilon. \quad (2.128)$$

Lo cual contradice (2.126). Por lo tanto $\lim_{N \rightarrow \infty} U_N = U$.

Finalmente, aplicando el teorema de convergencia dominada se tiene que:

$$\inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi) = \lim_{N \rightarrow \infty} U_N = \bar{c}_1 + \lim_{N \rightarrow \infty} \sum_{n \in \mathcal{G}} p_n \int (V_1^N)_n(x^n) d\mu^n(x^n) = B_1(V_1).$$

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Lo cual demuestra (2.124). ■

El siguiente teorema nos marca las condiciones necesarias para que una regla de paro ψ alcance a la cota inferior de la función de Lagrange.

Teorema 7.

Si existe $\psi^ \in \mathcal{T}_{\mathcal{M}}$ tal que*

$$L_{\mathcal{M}}(\psi^*) = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi) \quad (2.129)$$

entonces ψ^ cumple la condición Π^∞ .*

Demostración del Teorema 7:

Supongamos que existe $\psi^* \in \mathcal{T}_{\mathcal{M}}$ tal que cumple (2.129). Por Lema 12 para cualquier $r \geq 1$ se tienen las siguientes desigualdades:

$$L_{\mathcal{M}}(\psi^*) \geq B_{r+1}^{\psi^*}(V_{r+1}) \geq B_r^{\psi^*}(V_r) \geq \cdots \geq B_2^{\psi^*}(V_2) \geq B_1(V_1) \quad (2.130)$$

Dado que se tiene (2.129) entonces por Lema 13 existen igualdades en todas las desigualdades de (2.130). Nuevamente, por Lema 12 como la igualdad se alcanza para cada $r \geq 1$ y toda $\mathbf{n} \in \mathcal{P}_r$, entonces ψ^* satisface la condición Π^∞ . ■

El siguiente teorema marca las condiciones suficientes para alcanzar la cota inferior de la función $L_{\mathcal{M}}(\psi)$.

Teorema 8.

Sea ψ^ tal que cumple la condición Π^∞ . Si $\psi^* \in \mathcal{T}_{\mathcal{M}}$, entonces*

$$L_{\mathcal{M}}(\psi^*) = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi). \quad (2.131)$$

Demostración del Teorema 8:

Si $\psi^* \in \mathcal{T}_{\mathcal{M}}$, por (2.34) y (2.33):

$$L_{\mathcal{M}}(\psi^*) = \lim_{r \rightarrow \infty} \sum_{k=1}^r \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi^*} (c(\mathbf{n})f^{\mathbf{n}} + l_{\mathbf{n}}) d\mu^{\mathbf{n}}(x^{\mathbf{n}}). \quad (2.132)$$

Debido a que ψ^* cumple la condición Π^∞ y por el Lema 12, para cualquier $k \geq 1$ se tiene la siguiente cadena de igualdades:

$$B_k^{\psi^*}(V_k) = B_{k-1}^{\psi^*}(V_{k-1}) = \cdots = B_2^{\psi^*}(V_2) = B_1(V_1) = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi). \quad (2.133)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

La última igualdad se cumple por Lema 13. Entonces de (2.132) y (2.133) se tiene $L_{\mathcal{M}}(\psi^*) \leq \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi)$. Por otro lado, por ser ínfimo se tiene que $L_{\mathcal{M}}(\psi^*) \geq \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi)$, por tanto $L_{\mathcal{M}}(\psi^*) = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi)$.

Por tanto, lo anterior demuestra este teorema, lo cual nos da las condiciones suficientes para alcanzar el ínfimo. ■

Por lo tanto, los Teoremas 7 y 8 caracterizan (con condiciones necesarias y suficientes) a la familia de reglas de paro ψ^* que **alcanzan** la cota inferior de la función de Lagrange.

Condición $(\Pi^\infty, \Delta_\psi)$. (Prueba secuencial óptima)

Una prueba secuencial (ψ, ϕ) satisface la condición $(\Pi^\infty, \Delta_\psi)$, si ψ cumple la condición Π^∞ y ϕ cumple la condición Δ_ψ .

El siguiente teorema determina condiciones suficientes y necesarias para nuestro problema de optimización sin restricciones (1.19) (minimización de la función de Lagrange).

Dado que en esta sección estamos trabajando con la clase de reglas de paro truncables, entonces vamos a tomar la siguiente subclase de $\mathfrak{S}_{\mathcal{M}}$.

$$\mathfrak{S}_{\mathcal{T}} := \{(\psi, \phi) \in \mathfrak{S}_{\mathcal{M}} : \psi \in \mathcal{T}_{\mathcal{M}}\}. \quad (2.134)$$

Teorema 9.

Si una prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$ entonces

$$L_{\mathcal{M}}(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{T}}} L_{\mathcal{M}}(\psi, \phi). \quad (2.135)$$

Además, si existe alguna prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ tal que cumple (2.135) entonces (ψ^*, ϕ^*) satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$.

Demostración del Teorema 9:

Sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ la prueba secuencial que satisface $(\Pi^\infty, \Delta_{\psi^*})$. Ahora, por Teorema 8 se sigue que $L_{\mathcal{M}}(\psi^*) = \inf_{\psi \in \mathcal{T}_{\mathcal{M}}} L_{\mathcal{M}}(\psi)$; pero sabemos que $L_{\mathcal{M}}(\psi^*) = \inf_{\psi \in \mathcal{F}} L_{\mathcal{M}}(\psi) = L_{\mathcal{M}}(\psi^*, \phi)$.

Después dado que ϕ^* satisface la condición Δ_{ψ^*} , entonces por Teorema 2 (segunda parte) y de (2.9) se sigue que $L_{\mathcal{M}}(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{T}}} L_{\mathcal{M}}(\psi, \phi)$.

Esto demuestra la primera parte del teorema.

2.1. Prueba secuencial óptima con grupos de tamaño aleatorio

Ahora, supongamos que existe una prueba secuencial (ψ^*, ϕ^*) con $\psi^* \in \mathcal{T}_{\mathcal{M}}$ y tal que se cumple (2.135). Esto significa que (ψ^*, ϕ^*) es la regla que minimiza a la función de Lagrange $L_{\mathcal{M}}(\psi, \phi)$, por tanto por Teorema 7 se sigue que ψ^* satisface la condición Π^∞ , y debido a (2.33) y a Teorema 2 (segunda parte) se tiene que ϕ^* cumple la condición Δ_{ψ^*} . Por lo tanto, (ψ^*, ϕ^*) satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$. ■

El siguiente teorema determina las condiciones suficientes para resolver nuestro primer problema de investigación (con restricciones) (1.18).

Teorema 10.

Si una prueba secuencial $(\psi^, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$, entonces para cualquier otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}_{\mathcal{T}}$ tal que*

$$\alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (2.136)$$

se tiene que

$$K_{\mathcal{M}}(\psi) \geq K_{\mathcal{M}}(\psi^*). \quad (2.137)$$

La desigualdad (2.137) es estricta, si al menos una de las desigualdades en (2.136) lo es.

Demostración del Teorema 10:

Sea $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ una prueba secuencial tal que ψ^* y ϕ^* cumplen, respectivamente, las condiciones Π^∞ y Δ_{ψ^*} . Dado que $\psi^* \in \mathcal{T}_{\mathcal{M}}$ entonces por Teorema 9, se tiene que $L_{\mathcal{M}}(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_{\mathcal{T}}} L_{\mathcal{M}}(\psi, \phi)$. Ahora sea (ψ, ϕ) cualquier otra prueba secuencial con $\psi \in \mathcal{T}_{\mathcal{M}} \subseteq \mathcal{F}_{\mathcal{M}}$ y tal que cumple (2.136). Entonces por Corolario 1 se obtiene (2.137). Además, por el mismo corolario la desigualdad (2.137) es estricta, si al menos una de las desigualdades en (2.136) lo es.

Por lo tanto, una prueba secuencial $(\psi^*, \phi^*) \in \mathfrak{S}_{\mathcal{T}}$ que satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$ es la que minimiza el costo promedio $K_{\mathcal{M}}(\psi)$ del experimento estadístico secuencial. ■

Finalmente, a manera de resumen, pondremos la estructura expandida de las componentes de la prueba secuencial óptima (ψ^*, ϕ^*) (para el caso infinito) que satisface la condición $(\Pi^\infty, \Delta_{\psi^*})$. Para cada $r = 1, 2, \dots$ y $\mathbf{n} \in \mathcal{P}_r$:

$$\begin{aligned} I_{\{l_{\mathbf{n}} < \bar{c}_{r+1}f_{\mathbf{n}} + (I_{r+1}V_{r+1})_{\mathbf{n}}\}} &\leq \psi_{\mathbf{n}}^* \leq I_{\{l_{\mathbf{n}} \leq \bar{c}_{r+1}f_{\mathbf{n}} + (I_{r+1}V_{r+1})_{\mathbf{n}}\}} \\ I_{\{\lambda_0 f_0^{\mathbf{n}} < \lambda_1 f_1^{\mathbf{n}}\}} &\leq \phi_{\mathbf{n}}^* \leq I_{\{\lambda_0 f_0^{\mathbf{n}} \leq \lambda_1 f_1^{\mathbf{n}}\}}, \end{aligned} \quad (2.138)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

sobre los conjuntos $S_n^{\psi^*}$ y $T_n^{\psi^*}$, μ^n -c.t.p.

Cabe mencionar que en todo el proceso anterior se trabajaron implícitamente las dos constantes $\lambda_0 > 0$ y $\lambda_1 > 0$ cualesquiera, pero fijas. Estos multiplicadores de Lagrange estuvieron presentes (implícita o explícitamente) en cada una de las definiciones y teoremas que hacían referencia a la función de Lagrange $L_{\mathcal{M}}(\psi, \phi; \lambda_0, \lambda_1)$ y por tanto dichas constantes son específicas para la regla secuencial óptima encontrada (ψ^*, ϕ^*) . Sin embargo, para finalizar esta sección queda pendiente la siguiente pregunta ¿se pueden variar estas constantes de tal manera que se encuentren aquéllas para las cuales la condición $(\Pi^\infty, \Delta_\psi)$ sea necesaria para que minimicen a $K_{\mathcal{M}}$?

2.2. Problema Kiefer-Weiss con grupos de tamaño aleatorio

En el contexto del análisis secuencial clásico con grupos de tamaño uno y observaciones independientes e idénticamente distribuidas, el problema de Kiefer-Weiss consiste en encontrar una prueba secuencial (ψ, ϕ) que mínimize al supremo del tiempo promedio de paro del experimento $(\sup_{\theta} E_{\theta}\tau)$, sobre todas las pruebas secuenciales cuyas probabilidades de error tipo I y II estén acotadas por debajo de ciertas constantes α y β (véase, [Kiefer-Weiss, (1957)]). No obstante, con base en la mayoría de la literatura al respecto, se sabe que la solución a este problema para modelos particulares se obtiene a través de la resolución de otro problema más sencillo, conocido como el problema de Kiefer-Weiss (K-W) modificado (véanse, [Lorden, (1980)], [Novikov, (2009)], [Novikov, *et al*, (2022)]), entre otros). El cual supone que existe una prueba secuencial (ψ^*, ϕ^*) cuyas probabilidades de error tipo I y II son fijas, y que minimiza al tiempo de paro promedio $E_{\theta^*}\tau^*$, para un parámetro θ^* fijo que es el menos favorable para el tiempo de paro τ^* en el sentido de que $\sup_{\theta} E_{\theta}\tau^* = E_{\theta^*}\tau^*$. Debido a la complejidad del problema existen varias investigaciones que resuelven, bajo ciertas condiciones, el problema Kiefer-Weiss modificado donde consideran a las tres distribuciones pertenecientes a la misma familia paramétrica (generalmente exponencial), y algunas otras, a su vez, hacen aproximaciones de la prueba óptima de manera computacional para una distribución en particular, por ejemplo: en Novikov, *et al* (2022) se presenta un método numérico al caso clásico de K-W para una población con distribución Bernoulli.

En el contexto del análisis secuencial con grupos de tamaño aleatorio, vamos a abordar también el problema de K-W modificado; que resuelve el problema original, bajo los supuestos clásicos del mismo, i.e. con base en un problema de hipótesis simples, con observaciones i.i.d. y con las tres distribuciones de probabilidad en la misma familia paramétrica. Cabe mencionar que en lo tradicional de K-W, el tercer parámetro se encuentra entre los parámetros hipotetizados, de la forma $\theta_0 < \theta < \theta_1$, y bajo este supuesto, la prueba óptima generalmente es truncada. Por lo cual, solo trabajaremos el K-W para la prueba óptima truncada.

En secciones anteriores hemos presentado (secciones 2.1.2 y 2.1.3) las pruebas ³ secuenciales (ψ^*, ϕ^*) óptimas, que minimizan el costo promedio del experimento $K_{\mathcal{M}}(\psi^*)$ sobre todas las pruebas secuenciales en cierto conjunto, tanto para el caso truncado como para el caso infinito (véanse Teoremas 5 y 10). Pero, hasta ahora, en el presente trabajo no se habían considerado a las observaciones de los grupos como independientes e idénticamente distribuidas (i.i.d.), con distribución P . Entonces, en esta sección se expresará de manera analítica a la prueba secuencial óptima (ψ^*, ϕ^*) truncada, determinada en la sección 2.1.2, considerando ahora a las observaciones como variables aleatorias i.i.d. y con grupos de observaciones de tamaño aleatorios.

³En realidad son familias de pruebas secuenciales óptimas.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Para comenzar tomemos los siguientes supuestos que serán considerados a lo largo de este apartado.

Supongamos que se tiene un proceso estocástico $\{X_i : i \in \mathbb{N}\}$ con v.a. i.i.d., bajo la distribución P sobre un espacio $(\mathcal{X}, \mathfrak{X})$ medible. Supongamos que se tiene un problema de pruebas secuenciales con hipótesis simples $H_0 : P = P_0$ vs. $H_1 : P = P_1$, donde las observaciones vienen en grupos de tamaño aleatorio representados por la sucesión ν_1, ν_2, \dots v.a. independientes, donde para cada $r \geq 1$, el r -ésimo grupo tiene $P(\nu_r = n) = p_r(n) \geq 0$, para cada $n \in \mathcal{G}$ con $\sum_{n \in \mathcal{G}} p_r(n) = 1$. Nuevamente, estamos en un contexto de un problema de pruebas de hipótesis simples $\mathcal{M} = (P_0, P_1, P)$ donde ahora las tres distribuciones pertenecen a la misma familia paramétrica.

Supongamos que

$$\mu\{x \in \mathcal{X} : f_0(x) \neq f_1(x)\} > 0. \quad (2.139)$$

Esta suposición indica que la existe una diferencia entre las distribuciones a contrastar en las hipótesis; ya que si existiera, se tendría la igualdad entre las distribuciones y por ende no tendría sentido el “contraste” de hipótesis.

El objetivo es expresar dicha prueba óptima truncada en términos de dos razones de verosimilitud diferentes:

$$Z_n^0 = \frac{f_0^n}{f^n} \quad \text{y} \quad Z_n^1 = \frac{f_1^n}{f^n},$$

donde las tres funciones de distribución pertenecen a la misma familia paramétrica, pero la tercera distribución f es diferente a f_0 y f_1 , y con $\theta_0 < \theta < \theta_1$ (versión del problema de K-W modificado).

Definición 23.

Se define el conjunto:

$$\mathcal{K} := \{x \in \mathcal{X} : f(x) > 0\}. \quad (2.140)$$

Análogamente a la definición (2.140), al utilizar \mathcal{K}_0 o \mathcal{K}_1 nos referiremos a los conjuntos definidos, respectivamente, bajo las funciones f_0 y f_1 , e.g. $\mathcal{K}_0 := \{x : f_0(x) > 0\}$.

En este contexto con v.a. independientes es importante notar que si $x_1, x_2, \dots, x_k, \dots$ son observaciones y $\mathbf{n} \in \mathcal{G}^k$, entonces

$$x^n \in \mathcal{K}^n \iff x_i \in \mathcal{K} \text{ para toda } i = 1, \dots, |\mathbf{n}|.$$

2.2.1. Kiefer-Weiss con grupos aleatorios aplicado a la familia exponencial

Debido a que en la mayoría de las investigaciones que abordan el problema de K-W trabajan con las funciones de densidad que pertenecen a la familia exponencial, y ya que este problema no es el objetivo principal de nuestra investigación, entonces nosotros también trabajaremos dicho problema K-W modificado aplicado a dicha familia.

Cabe mencionar que debido a las características de las distribuciones pertenecientes a la familia exponencial, en donde las funciones de densidad son positivas, entonces consideramos que vamos a trabajar con el conjunto donde las observaciones de las tres funciones de densidad son positivas, y no nos encontraremos con casos nulos para alguna densidad. Por tanto, se tiene la siguiente definición.

Definición 24.

Para cualesquiera $k \geq 1$, $\mathbf{n} \in \mathcal{G}^k$ y $x^{\mathbf{n}} \in \mathcal{K}^{\mathbf{n}}$, sea

$$z_{\mathbf{n}}^i := z_{\mathbf{n}}^i(x^{\mathbf{n}}) = \frac{f_i^{\mathbf{n}}(x^{\mathbf{n}})}{f^{\mathbf{n}}(x^{\mathbf{n}})}, \quad \text{con } i = 0, 1. \quad (2.141)$$

Definición 25.

Sea $r \geq 1$ natural cualquiera. Para cualquier función $v(z_0, z_1)$, $z_0, z_1 \geq 0$, medible y no negativa, se define

$$J_r v(z_0, z_1) := \sum_{m \in \mathcal{G}} p_r(m) E v \left(z_0 \frac{f_0^m(X^m)}{f^m(X^m)}, z_1 \frac{f_1^m(X^m)}{f^m(X^m)} \right). \quad (2.142)$$

Cabe notarse que debido a las propiedades de la esperanza E , y del hecho de que $z_0, z_1 \geq 0$ y $\frac{f_i^m(X^m)}{f^m(X^m)} > 0$ para $i = 1, 0$ se obtiene que J es un operador donde su imagen es a su vez una función medible y no negativa. Entonces debido a esto podemos aplicar iteradamente el operador J en la siguiente definición.

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Definición 26.

Dadas $\lambda_0 > 0$ y $\lambda_1 > 0$ fijas, para cualesquiera $z_0, z_1 \geq 0$ y para $N \geq 1$ sea

$$\rho_N^N(z_0, z_1) = \rho_N^N(z_0, z_1; \lambda_0, \lambda_1) := g(z_0, z_1; \lambda_0, \lambda_1) = \min\{\lambda_0 z_0, \lambda_1 z_1\}, \quad (2.143)$$

y, de manera recursiva, para $r = N, N - 1, \dots, 2$,

$$\rho_{r-1}^N(z_0, z_1) := \min \left\{ g(z_0, z_1), \bar{c}_r + J_r \rho_r^N(z_0, z_1) \right\}. \quad (2.144)$$

Lema 14.

Sea N un entero positivo. Para cualesquiera $\mathbf{n} \in \mathcal{G}^r$ y $r = N, \dots, 1$, se tiene

$$V_r^N = \rho_r^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) f^{\mathbf{n}}, \quad (2.145)$$

donde V_r^N se definió en (2.59)-(2.60).

Demostración del Lema 14:

Se demostrará por inducción inversa empezando por $r = N$ y luego se demostrará para $r = N - 1, \dots, 1$. Supongamos que $\mathbf{n} \in \mathcal{G}^N$, entonces por (2.59) se tiene que

$$\begin{aligned} V_N^N &= l_{\mathbf{n}} = \min \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \} \\ &= \min \{ \lambda_0 z_{\mathbf{n}}^0, \lambda_1 z_{\mathbf{n}}^1 \} f^{\mathbf{n}} I_{\{f^{\mathbf{n}} > 0\}} + \min \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \} I_{\{f^{\mathbf{n}} = 0\}}. \end{aligned} \quad (2.146)$$

Dado que el segundo sumando en (2.146) es nulo, entonces debido a (2.143) para $\mathbf{n} \in \mathcal{G}^N$,

$$V_N^N = \min \{ \lambda_0 z_{\mathbf{n}}^0, \lambda_1 z_{\mathbf{n}}^1 \} f^{\mathbf{n}} = \rho_N^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) f^{\mathbf{n}}.$$

(H.I.)⁴: Supongamos que (2.145) se cumple para algún $r = N, \dots, 2$.

Debido a (2.60) y (2.46) para $\mathbf{n} \in \mathcal{G}^{r-1}$ se tiene que

$$\begin{aligned} (V_{r-1}^N)_{\mathbf{n}} &= (M_r V_r^N)_{\mathbf{n}} = \min \left\{ l_{\mathbf{n}}, \bar{c}_r f^{\mathbf{n}} + (I_r V_r^N)_{\mathbf{n}} \right\} \\ &= \min \left\{ \min \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \}, \bar{c}_r f^{\mathbf{n}} + \sum_{m \in \mathcal{G}} p_r(m) \int (V_r^N)_{(\mathbf{n}, m)}(x^{\mathbf{n}}, x^m) d\mu^m(x^m) \right\}. \end{aligned}$$

⁴Hipótesis de Inducción.

2.2. Problema Kiefer-Weiss con grupos de tamaño aleatorio

Ahora como $\mathbf{n} \in \mathcal{G}^{r-1}$ entonces $(\mathbf{n}, m) \in \mathcal{G}^r$ para $2 \leq r \leq N$, entonces aplicando la hipótesis de inducción se obtiene que:

$$(V_{r-1}^N)_{\mathbf{n}} = \min \left\{ g(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1), \bar{c}_r + \sum_{m \in \mathcal{G}} p_r(m) \int \rho_r^N \left(z_{(\mathbf{n}, m)}^0(x^n, x^m), z_{(\mathbf{n}, m)}^1(x^n, x^m) \right) f^m(x^m) d\mu^m(x^m) \right\} f^n.$$

Entonces para $\mathbf{n} \in \mathcal{G}^{r-1}$,

$$\begin{aligned} (V_{r-1}^N)_{\mathbf{n}} &= \min \left\{ g(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1), \bar{c}_r + \sum_{m \in \mathcal{G}} p_r(m) \int_{\{f^m(x^m) > 0\}} \rho_r^N \left(z_{\mathbf{n}}^0 \frac{f_0^m(x^m)}{f^m(x^m)}, z_{\mathbf{n}}^1 \frac{f_1^m(x^m)}{f^m(x^m)} \right) f^m(x^m) d\mu^m \right\} f^n \\ &= \min \left\{ g(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1), \bar{c}_r + J_r \rho_r^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) \right\} f^n = \rho_{r-1}^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) f^n. \end{aligned} \quad (2.147)$$

Lo cual demuestra (2.145) para toda $r = N, \dots, 1$. ■

El siguiente lema demuestra que la sucesión $\{\rho_r^N\}_N$ es decreciente con respecto al nivel N (de truncamiento).

Lema 15.

Para cualesquiera $N \geq 1$ y $r = N, \dots, 1$ se tiene que

$$\rho_r^N(z_0, z_1) \geq \rho_r^{N+1}(z_0, z_1), \quad (2.148)$$

para todas $z_0, z_1 \geq 0$.

Demostración del Lema 15:

Nuevamente aplicando inducción sobre r . Sean $N \geq 1$ fijo. Sea $r = N$,

$$\rho_N^N(z_0, z_1) = g(z_0, z_1) \geq \min \left\{ g(z_0, z_1), \bar{c}_{N+1} + J_{N+1} \rho_{N+1}^{N+1}(z_0, z_1) \right\} = \rho_N^{N+1}(z_0, z_1),$$

por (2.144).

H.I.: Supongamos que (3.21) se cumple para algún $r = N, N-1, \dots, 2$. Entonces, cualesquiera que sean $z_0, z_1 \geq 0$,

$$\begin{aligned} \rho_{r-1}^N(z_0, z_1) &= \min \left\{ g(z_0, z_1), \bar{c}_r + J_r \rho_r^N(z_0, z_1) \right\} \\ &\geq \min \left\{ g(z_0, z_1), \bar{c}_r + J_r \rho_r^{N+1}(z_0, z_1) \right\} = \rho_{r-1}^{N+1}(z_0, z_1), \end{aligned} \quad (2.149)$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

ya que por la H.I.,

$$\begin{aligned} J_r \rho_r^N(z_0, z_1) &= \sum_{m \in \mathcal{G}} p_r(m) E \rho_r^N \left(z_0 \frac{f_0^m(X^m)}{f^m(X^m)}, z_1 \frac{f_1^m(X^m)}{f^m(X^m)} \right) \\ &\geq \sum_{m \in \mathcal{G}} p_r(m) E \rho_r^{N+1} \left(z_0 \frac{f_0^m(X^m)}{f^m(X^m)}, z_1 \frac{f_1^m(X^m)}{f^m(X^m)} \right) = J_r \rho_r^{N+1}(z_0, z_1). \end{aligned}$$

Entonces de (2.149), $\rho_{r-1}^N(z_0, z_1) \geq \rho_{r-1}^{N+1}(z_0, z_1)$. ■

Ahora, vamos a recordar la estructura de la prueba secuencial óptima (ψ^*, ϕ^*) truncada que satisface la condición (Π^N, Δ_{ψ^*}) , encontrada en la sección 2.1.2.

Las componentes de la regla de paro ψ^* son: para toda $1 \leq r \leq N-1$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$

$$\begin{aligned} I_{\{l_{\mathbf{n}} < \bar{c}_{r+1} f^{\mathbf{n}} + (I_{r+1} V_{r+1}^N)_{\mathbf{n}}\}} &\leq \psi_{\mathbf{n}}^* \leq I_{\{l_{\mathbf{n}} \leq \bar{c}_{r+1} f^{\mathbf{n}} + (I_{r+1} V_{r+1}^N)_{\mathbf{n}}\}} \\ \text{y } \psi_{\mathbf{n}} &\equiv 1 \quad \text{para toda } \mathbf{n} \in \mathcal{P}_N, \end{aligned} \tag{2.150}$$

sobre el conjunto $T_{\mathbf{n}}^{\psi^*} := \{x^{\mathbf{n}} : t_{\mathbf{n}}^{\psi^*}(x^{\mathbf{n}}) > 0\}$, $\mu^{\mathbf{n}}$ -c.t.p.

Las componentes de la regla de decisión ϕ^* son: para toda $1 \leq r \leq N$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$

$$I_{\{\lambda_0 f_0^{\mathbf{n}} < \lambda_1 f_1^{\mathbf{n}}\}} \leq \phi_{\mathbf{n}}^* \leq I_{\{\lambda_0 f_0^{\mathbf{n}} \leq \lambda_1 f_1^{\mathbf{n}}\}}, \tag{2.151}$$

sobre el conjunto $S_{\mathbf{n}}^{\psi^*} := \{x^{\mathbf{n}} : s_{\mathbf{n}}^{\psi^*}(x^{\mathbf{n}}) > 0\}$ $\mu^{\mathbf{n}}$ -c.t.p.

Entonces la prueba secuencial truncada óptima (ψ^*, ϕ^*) se reescribirá en términos de $z_{\mathbf{n}}^0$ y $z_{\mathbf{n}}^1$.

Corolario 8.

Dada la prueba secuencial óptima truncada (ψ^, ϕ^*) con componentes en (2.150) y (2.151) esta se reescribe de la siguiente forma.*

Para cada $1 \leq r \leq N-1$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$

$$\begin{aligned} I_{\{g(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) < \bar{c}_{r+1} + J_{r+1} \rho_{r+1}^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1)\}} &\leq \psi_{\mathbf{n}}^* \leq I_{\{g(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) \leq \bar{c}_{r+1} + J_{r+1} \rho_{r+1}^N(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1)\}}, \\ \text{y } \psi_{\mathbf{n}} &\equiv 1 \quad \text{para toda } \mathbf{n} \in \mathcal{P}_N. \end{aligned} \tag{2.152}$$

sobre el conjunto $T_{\mathbf{n}}^{\psi^}$. Para cada $1 \leq r \leq N$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$*

$$I_{\{\lambda_0 z_{\mathbf{n}}^0 < \lambda_1 z_{\mathbf{n}}^1\}} \leq \phi_{\mathbf{n}}^* \leq I_{\{\lambda_0 z_{\mathbf{n}}^0 \leq \lambda_1 z_{\mathbf{n}}^1\}}. \tag{2.153}$$

sobre el conjunto $S_{\mathbf{n}}^{\psi^}$ $\mu^{\mathbf{n}}$ -c.t.p., respectivamente. Donde $g(z_0, z_1)$ y $J_r \rho_r^N(z_0, z_1)$ están definidas en (2.143) y (2.142), y $z_{\mathbf{n}}^0, z_{\mathbf{n}}^1$ son las razones definidas en (2.141).*

2.2. Problema Kiefer-Weiss con grupos de tamaño aleatorio

Demostración del Corolario 8:

Tomemos primero las componentes de la regla de decisión ϕ^* . Sean $1 \leq r \leq N$, $\mathbf{n} \in \mathcal{P}_r$, si $x^n \in \mathcal{K}^n$ entonces $I_{\{f^n > 0\}} = 1$.

$$I_{\{\lambda_0 f_0^n < \lambda_1 f_1^n\}} \leq \phi_n^* \leq I_{\{\lambda_0 f_0^n \leq \lambda_1 f_1^n\}} \implies I_{\{\lambda_0 z_n^0 < \lambda_1 z_n^1\}} \leq \phi_n^* \leq I_{\{\lambda_0 z_n^0 \leq \lambda_1 z_n^1\}}.$$

Ahora para ψ^* por (2.150), para $1 \leq r \leq N - 1$, $\mathbf{n} \in \mathcal{P}_r$ y $x^n \in \mathcal{K}^n$ se tiene que

$$I_{\{l_n < \bar{c}_{r+1} f^n + (I_{r+1} V_{r+1}^N)_n\}} \leq \psi_n^* \leq I_{\{l_n \leq \bar{c}_{r+1} f^n + (I_{r+1} V_{r+1}^N)_n\}}$$

Sin embargo, para $\mathbf{n} \in \mathcal{P}_r$ con $1 \leq r \leq N - 1$

$$\begin{aligned} l_n &= \min \{ \lambda_0 z_n^0, \lambda_1 z_n^1 \} f^n I_{\{f^n > 0\}} + \min \{ \lambda_0 f_0^n, \lambda_1 f_1^n \} I_{\{f^n = 0\}} \\ &= \min \{ \lambda_0 z_n^0, \lambda_1 z_n^1 \} f^n = g(z_n^0, z_n^1) f^n. \end{aligned}$$

Por otro lado,

$$\begin{aligned} \bar{c}_{r+1} f^n + (I_{r+1} V_{r+1}^N)_n &= \bar{c}_{r+1} f^n + \sum_{m \in \mathcal{G}} p_{r+1}(m) \int (V_{r+1}^N)_{(n,m)}(x^n, x^m) d\mu^m(x^m) \\ &= \left\{ \bar{c}_{r+1} + \sum_{m \in \mathcal{G}} p_{r+1}(m) \int_{\mathcal{K}^m} \rho_{r+1}^N \left(z_n^0 \frac{f_0^m(x^m)}{f^m(x^m)}, z_n^1 \frac{f_1^m(x^m)}{f^m(x^m)} \right) f^m(x^m) d\mu^m \right\} f^n \\ &= \left\{ \bar{c}_{r+1} + J_{r+1} \rho_{r+1}(z_n^0, z_n^1) \right\} f^n. \end{aligned} \tag{2.154}$$

Dado que para toda $x^n \in \mathcal{K}^n$ se tiene $I_{\{f^n > 0\}} = 1$, entonces para toda $1 \leq r \leq N - 1$, $\mathbf{n} \in \mathcal{P}_r$ se sigue que:

$$\begin{aligned} I_{\{g(z_n^0, z_n^1) < \bar{c}_{r+1} + J_{r+1} \rho_{r+1}^N(z_n^0, z_n^1)\}} &\leq \psi_n^* \leq I_{\{g(z_n^0, z_n^1) < \bar{c}_{r+1} + J_{r+1} \rho_{r+1}^N(z_n^0, z_n^1)\}} \\ &\text{y para toda } \mathbf{n} \in \mathcal{P}_N \quad \psi_n \equiv 1. \quad \blacksquare \end{aligned}$$

Corolario 9.

Para la prueba secuencial óptima (ψ^*, ϕ^*) con componentes en (2.152) y (2.153), respectivamente, se tiene que:

$$L_{\mathcal{M}}^N(\psi^*, \phi^*) = \bar{c}_1 + J_1 \rho_1^N(1, 1). \tag{2.155}$$

CAPÍTULO 2. OPTIMALIDAD DE PRUEBAS SECUENCIALES CON GRUPOS ALEATORIOS

Demostración del Corolario 9:

Sabemos ϕ^* satisface la condición Δ_{ψ^*} y además que ψ^* satisface la condición Π^N , entonces por Teorema 4 y (2.73) se tiene que:

$$L_{\mathcal{M}}^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_{\mathcal{M}}^N(\psi, \phi) = \bar{c}_1 + \sum_{n \in \mathcal{G}} p_n \int (V_1)_n(x^n) d\mu^n(x^n).$$

Entonces por Corolario 10, (2.158) y (2.142),

$$L_{\mathcal{M}}^N(\psi^*) = \bar{c}_1 + \sum_{n \in \mathcal{G}} p_n \int \rho_1(z_n^0, z_n^1) f^n(x^n) d\mu^n(x^n) = \bar{c}_1 + J_1 \rho_1(1, 1). \quad \blacksquare$$

Debido a que en esta sección solo vamos a trabajar con la prueba óptima truncada, no sería necesario analizar el límite cuando $N \rightarrow \infty$, sin embargo, es importante mencionar que en el caso en el que el tercer parámetro θ se encuentre fuera del intervalo (θ_0, θ_1) , aún cuando toque los extremos, es posible analizar el problema K-W para la prueba óptima con horizonte infinito. Por lo cual, solo terminaremos esta sección con algunos otros detalles de las funciones ρ 's, dejando precedente de que también se podría trabajar el problema K-W para el caso no truncado.

Debido a las características de la sucesión de las funciones ρ_r^N , por Lema 15 para cualesquiera $r \geq 1$ y $z_0, z_1 \geq 0$, la sucesión $\{\rho_r^N(z_0, z_1)\}_{N \geq r}$ es decreciente y no negativa, por lo tanto su límite existe cuando $N \rightarrow \infty$. Para cualquier $r \geq 1$,

$$\boxed{\rho_r(z_0, z_1) := \lim_{N \rightarrow \infty} \rho_r^N(z_0, z_1).} \quad (2.156)$$

Donde para cualquier $r \geq 2$:

$$\rho_{r-1}(z_0, z_1) = \min \left\{ g(z_0, z_1), \bar{c}_r + J_r \rho_r(z_0, z_1) \right\}. \quad (2.157)$$

Corolario 10.

Para cualesquiera $\mathbf{n} \in \mathcal{G}^r$ y $r \geq 1$, se tiene

$$V_r = \rho_r(z_{\mathbf{n}}^0, z_{\mathbf{n}}^1) f^{\mathbf{n}}, \quad (2.158)$$

donde V_r se definió en (2.116).

Demostración del Corolario 10:

Tomando el límite a (2.145), cuando $N \rightarrow \infty$ y con base en (2.116) y (2.156) se sigue (2.158). \blacksquare

CAPÍTULO 3

PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

En este capítulo se determinará de manera analítica la prueba secuencial óptima (ψ^*, ϕ^*) en términos de la razón de verosimilitud $Z_n = \frac{f_1^n}{f_0^n}$, donde tanto las observaciones como los tamaños de grupos serán variables aleatorias independientes (como en el capítulo anterior), pero las observaciones, además, serán idénticamente distribuidas (i.d.). A su vez, en la siguiente sección se trabajará la prueba óptima considerando también a los tamaños de los grupos como v.a. i.i.d., y esto nos llevará demostrar la optimalidad de la RSPRT.

Supongamos que se tiene un proceso estocástico $\{X_i : i \in \mathbb{N}\}$ con v.a. i.i.d., el cual tiene una distribución P sobre un espacio $(\mathcal{X}, \mathfrak{X})$ medible. Estas variables representan las observaciones en un experimento aleatorio secuencial. Supongamos que se tiene un problema de pruebas secuenciales con hipótesis simples $H_0 : P = P_0$ *v.s.* $H_1 : P = P_1$, donde las observaciones vienen en grupos de tamaños aleatorios ν_1, ν_2, \dots y éstos son v.a. discretas, no negativas e independientes (más no idénticamente distribuidas).

Con base en esto, consideraremos un problema de pruebas de hipótesis simples como:

$$\mathcal{M}_0 = (P_0, P_1, P_0). \tag{3.1}$$

Esta terna nos indica que las observaciones del experimento tiene la distribución P_0 , y ésta será

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

a su vez con la que se determinará la minimización del costo promedio total del experimento. Para esta sección supongamos lo siguiente:

$$\mu\{x \in \mathcal{X} : f_0(x) \neq f_1(x)\} > 0. \quad (3.2)$$

La suposición (3.2) indica que existen observaciones (o puntos) para las cuales las funciones de distribución no son iguales.

Retomemos algunos de los elementos trabajados en capítulos anteriores, **pero con sus respectivos cambios para este contexto**. Cabe resaltar que dichos elementos tendrán la misma notación que en los Capítulos 1 y 2, aunque ahora habrá “pequeñas modificaciones” como el cambio de la función de densidad f por f_0 , entre otras.

Consideramos un experimento estadístico secuencial por grupos con pruebas secuenciales (ψ, ϕ) , donde ψ es una regla de paro y ϕ una regla de decisión aplicadas a las etapas de observaciones. Sea ν_1, ν_2, \dots una sucesión de v.a. discretas, no negativas e independientes que representan los tamaños de los grupos de observaciones donde sus valores $n_1, n_2, \dots \in \mathcal{G}$ con $\mathcal{G} \subseteq \{0, 1, 2, \dots\}$. Es importante mencionar que en esta investigación los tamaños de grupos son aleatorios, sin embargo, implícitamente también se consideran sucesiones de tamaños fijos: como grupos de tamaño uno (contexto secuencial clásico) o cualquier tamaño constante (contexto secuencial generalizado).

La distribución de los tamaños de grupos es diferente para cada grupo, así para el r -ésimo grupo, $P(\nu_r = n) = p_r(n) \geq 0$, para cada $n \in \mathcal{G}$ con $\sum_{n \in \mathcal{G}} p_r(n) = 1$. Además, para los tamaños de r grupos consecutivos, denotado como $\mathbf{n} = (n_1, n_2, \dots, n_r) \in \mathcal{G}^r$, se tiene su probabilidad conjunta, que debido a la independencia entre los grupos $p_{\mathbf{n}} = \prod_{i=1}^r P(\nu_i = n_i) = \prod_{i=1}^r p_i(n_i)$. Dado $n_r \geq 0$ el tamaño del r -ésimo grupo de observaciones, se tiene su función de costo $c_r(n_r) \geq 0$. A partir esto se sigue que para $\mathbf{n} \in \mathcal{G}^r$, el *costo total de r etapas del experimento* es

$$c(\mathbf{n}) = \sum_{i=1}^r c_i(n_i).$$

A continuación se enlistan los elementos a utilizarse:

- El costo promedio del r -ésimo grupo,

$$\bar{c}_r = \sum_{m \in \mathcal{G}} c_r(m) p_r(m), \quad (3.3)$$

donde $0 \leq \bar{c}_r < \infty$ para toda $r \geq 1$.

- Se define la clase de reglas de paro cuyo tiempo de paro τ^ψ es finito, bajo P_0 como:

$$\mathcal{F}_0 := \{\psi : P_0(\tau^\psi < \infty) = 1\}. \quad (3.4)$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

- La clase de pruebas secuenciales cuya regla de paro detiene el experimento (bajo f_0):

$$\mathfrak{S}_0 := \{(\psi, \phi) : \psi \in \mathcal{F}_0\}.$$

- Dada ψ cualquiera, y $r \geq 1$ se define la variable aleatoria que determina la probabilidad condicional de no parar el experimento antes de la etapa r ,

$$\text{Para } \mathbf{n} \in \mathcal{G}^r, \quad t_{\mathbf{n}}^{\psi} := \begin{cases} 1, & \text{si } r = 1; \\ (1 - \psi_{(\mathbf{n}:1)}) \cdots (1 - \psi_{(\mathbf{n}:r-1)}), & \text{si } r > 1. \end{cases} \quad (3.5)$$

Además para $r \geq 1$ se define la v.a. que determina la probabilidad condicional de parar el experimento exactamente en la etapa r .

$$\text{Para } \mathbf{n} \in \mathcal{G}^r, \quad s_{\mathbf{n}}^{\psi} = t_{\mathbf{n}}^{\psi} \psi_{\mathbf{n}}. \quad (3.6)$$

Recordemos que las probabilidades anteriormente presentadas están condicionadas a las observaciones $x^{\mathbf{n}}$ del experimento en la respectiva etapa, pero que para no hacer más densa la notación, omitiremos esta escritura ($s_{\mathbf{n}}^{\psi}(x^{\mathbf{n}})$).

- Si $\psi \in \mathcal{F}_0$, se define el **costo total promedio** del experimento, bajo f_0 como:

$$K_0(\psi) := \sum_{r=1}^{\infty} E_0 (c_1(\nu_1) + \cdots + c_r(\nu_r) I_{\{\tau^{\psi}=r\}}) = \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} c(\mathbf{n}) E_0 s_{\mathbf{n}}^{\psi}.$$

Este costo es un caso particular del costo $K_{\mathcal{M}}(\psi)$ ((1.16) del Capítulo 2), pero en este capítulo se trabajará bajo el problema de pruebas de hipótesis \mathcal{M}_0 donde la tercera distribución es P_0 y de lo cual también se considerará a E_0 .

- Dada una prueba estadístico secuencial $(\psi, \phi) \in \mathfrak{S}_0$ se consideran sus probabilidades de error tipo I, $\alpha(\psi, \phi)$ y tipo II, $\beta(\psi, \phi)$ como:

$$\begin{aligned} \alpha(\psi, \phi) &= P_0(\text{Rechazar } H_0) = \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 s_{\mathbf{n}} \phi_{\mathbf{n}}, \\ \beta(\psi, \phi) &= P_1(\text{Aceptar } H_0 \cap \{\tau^{\psi} < \infty\}) = \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_1 s_{\mathbf{n}} (1 - \phi_{\mathbf{n}}). \end{aligned}$$

Recordemos que el objetivo principal es encontrar la prueba secuencial (ψ, ϕ) que minimice el costo promedio $K_0(\psi)$ del experimento, restringido a $\alpha(\psi, \phi) \leq \alpha$ y $\beta(\psi, \phi) \leq \beta$ para ciertas constantes dadas $\alpha, \beta \in [0, 1]$. Por tanto, para resolver este problema se aplicará el método de Lagrange en este contexto secuencial.

Recordemos que el Capítulo 2 se demostró, Teorema 1 y Corolario 1 la aplicación del método de Lagrange en el contexto secuencial para encontrar la prueba secuencial óptima, aplicado al problema de pruebas de hipótesis \mathcal{M} . Entonces ahora esto se aplicará al contexto \mathcal{M}_0 .

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

- Sean $\lambda_0 > 0$ y $\lambda_1 > 0$ cualesquiera, pero fijos. Para cualquier prueba secuencial $(\psi, \phi) \in \mathfrak{S}_0$ se define su función de Lagrange,

$$L_0(\psi, \phi) := L_0(\psi, \phi; \lambda_0, \lambda_1) = K_0(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi) \quad (3.7)$$

Las constantes de Lagrange λ_0 y λ_1 estarán siempre presentes en la mayoría de las funciones, pero por cuestiones prácticas omitiremos dicha dependencia, y las consideraremos fijas (como parámetros) a través del trabajo, a menos que se diga lo contrario.

Grosso modo, el desarrollo que aplicaremos para encontrar la prueba secuencial que minimiza la función de Lagrange se describe en la siguiente línea. Para cualquier prueba secuencial $(\psi, \phi) \in \mathfrak{S}_0$,

$$L_0(\psi, \phi) \geq \inf_{\phi} L_0(\psi, \phi) = L_0(\psi, \phi^*) = L_0(\psi) \geq \inf_{\psi \in \mathcal{F}_0} L_0(\psi) = L_0(\psi^*).$$

Ahora con base en las definiciones del costo promedio del experimento y las probabilidades de error tipo I y tipo II se desprende que:

$$L_0(\psi, \phi) = \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} \left((c(\mathbf{n}) + \phi_{\mathbf{n}} \lambda_0) f_0^{\mathbf{n}} + (1 - \phi_{\mathbf{n}}) \lambda_1 f_1^{\mathbf{n}} \right) d\mu^{\mathbf{n}}. \quad (3.8)$$

3.1.1. Regla de decisión óptima bajo H_0

Dada un regla de paro $\psi \in \mathcal{F}_0$ cualquiera, pero fija, en esta parte se determinará la cota inferior de la función de Lagrange $L_0(\psi, \phi)$ para las reglas de decisión ϕ .

- Para $r \geq 1$ y $\mathbf{n} \in \mathcal{G}^r$,

$$l_{\mathbf{n}} := \text{mín} \{ \lambda_0 f_0^{\mathbf{n}}, \lambda_1 f_1^{\mathbf{n}} \}. \quad (3.9)$$

- Para $\lambda_0 > 0$ y $\lambda_1 > 0$ cualesquiera, y $0 \leq z < \infty$ se define la función

$$g(z) = g(z; \lambda_0, \lambda_1) := \text{mín} \{ \lambda_0, \lambda_1 z \}. \quad (3.10)$$

Definición 27.

Para cualesquiera $\mathbf{n} \in \mathcal{G}^r$ y $r \geq 1$, se define la razón (de verosimilitud):

$$z_{\mathbf{n}} = z_{\mathbf{n}}(x^{\mathbf{n}}) := \begin{cases} \frac{f_1^{\mathbf{n}}(x^{\mathbf{n}})}{f_0^{\mathbf{n}}(x^{\mathbf{n}})}, & \text{si } f_0^{\mathbf{n}}(x^{\mathbf{n}}) > 0; \\ \infty, & \text{si } f_0^{\mathbf{n}}(x^{\mathbf{n}}) = 0 \text{ y } f_1^{\mathbf{n}}(x^{\mathbf{n}}) > 0; \\ 0, & \text{en otro caso.} \end{cases} \quad (3.11)$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Algunas observaciones respecto a la definición de la razón de verosimilitud anterior. Por un lado, debido a que en esta investigación estamos considerando grupos de tamaño nulo ($n = 0$) entonces para dichos grupos no habría observaciones $x^n = ()$, y dado que en este contexto las observaciones son v.a. independientes e i.d.; entonces en estos casos al aplicar las funciones de densidad f_1 y f_0 , respectivamente, con base en la medida de probabilidad trivial μ^0 , se tendrá por definición (1.2) que $f_1() = 1$ y $f_0() = 1$ y por tanto $z_n() = 1$. Esto es, para grupos de tamaño nulo, su razón de verosimilitud valdrá uno.

Por otro lado, es importante señalar que en el caso de que ninguna de las dos densidades “aportará información” sobre las observaciones x^n , *i.e.* ambas densidades fueran nulas $f_0^n(x^n) = 0$ y $f_1^n(x^n) = 0$, entonces las reglas de paro y decisión no estarían escritas en términos de la razón z_n ; en este caso se aplicaría para ambas reglas una aleatorización en términos solo de las observaciones:

$$0 \leq \psi_n^*(x^n) \leq 1 \quad \text{y} \quad 0 \leq \phi_n^*(x^n) \leq 1. \quad (3.12)$$

Esto es, se puede hacer lo que sea, tanto para la regla de paro ψ^* como para la decisión ϕ^* óptimas. Esta misma aleatorización de las reglas de paro y decisión sucederá para el caso $z_n = \infty$, pero esto se analizará más adelante. Por tanto, en lo subsiguiente nos centraremos en el caso cuando $f_0^n(x^n) > 0$.

Para $r \geq 1$ y $\mathbf{n} \in \mathcal{G}^r$ cualesquiera se definen los siguientes conjuntos.

- $\mathcal{K}_0^n := \{x^n \in \mathcal{X}^n : f_0^n(x^n) > 0\}$. Análogamente se define el conjunto \mathcal{K}_1^n .
- Sean $S_n^\psi := \{x^n \in \mathcal{X}^n : s_n^\psi(x^n) > 0\}$ y $T_n^\psi := \{x^n : t_n^\psi(x^n) > 0\}$.
- También, se tiene el conjunto $\mathcal{P}_r := \{\mathbf{n} \in \mathcal{G}^r : p_n > 0\}$.

Recordemos que en este trabajo estamos considerando a los multiplicadores de Lagrange λ_0 y λ_1 ambos positivos, entonces tiene sentido la siguiente condición.

Condición Δ_ψ (Estructura de la regla de decisión óptima)

Dada una regla de paro ψ fija, una regla de decisión ϕ cumple la condición Δ_ψ , si para cada $r \geq 1$ y cualquier $\mathbf{n} \in \mathcal{P}_r$ se tiene:

$$\phi_n \stackrel{\geq}{\leq} I_{\{\lambda_0/\lambda_1 \lesssim z_n\}}, \quad \text{sobre el conjunto } S_n^\psi \cap \mathcal{K}_0^n, \quad \mu^n\text{-c.t.p.}, \quad (3.13)$$

para $\lambda_0 > 0$ y $\lambda_1 > 0$.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

La condición anterior está basada en la condición (2.27) del Capítulo 2, salvo lo siguiente:

$$\begin{aligned}
 I_{\{\lambda_0 f_0^n \lesssim \lambda_1 f_1^n\}} &= I_{\{\lambda_0 f_0^n \lesssim \lambda_1 f_1^n\}} I_{\{f_0^n > 0\}} + I_{\{\lambda_0 f_0^n \lesssim \lambda_1 f_1^n\}} I_{\{f_0^n = 0\}}, \\
 &= I_{\{\lambda_0/\lambda_1 \lesssim f_1^n/f_0^n\}} I_{\{f_0^n > 0\}} \\
 &= I_{\{\lambda_0/\lambda_1 \lesssim z_n\}}, \quad \text{sobre el conjunto } S_n^\psi \cap \mathcal{K}_0^n > 0, \quad \mu^n\text{-c.t.p.}
 \end{aligned}$$

El siguiente teorema determina la cota inferior para la función de Lagrange general $L_0(\psi, \phi)$, y además marca condiciones necesarias y suficientes para que una regla de decisión alcance dicha cota.

Teorema 11.

Para cualquier prueba secuencial $(\psi, \phi) \in \mathfrak{S}_0$ se tiene que:

$$L_0(\psi, \phi) \geq \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi (c(\mathbf{n}) + g(z_{\mathbf{n}})). \quad (3.14)$$

La igualdad en (3.14) se alcanza si y solo si ϕ cumple la condición Δ_ψ .

Demostración del Teorema 11:

La demostración de este teorema es la misma que la del Teorema 2 al utilizar f_0 en lugar de f , solo que ahora, además se trabajará en términos de la función $g(z)$. Sabemos por (2.32), Teorema 2, que:

$$\begin{aligned}
 L_0(\psi, \phi) &\geq \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int s_{\mathbf{n}}^\psi (c(\mathbf{n}) f_0^n + l_{\mathbf{n}}) d\mu^n, \\
 &= \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int_{\mathcal{K}_0} s_{\mathbf{n}}^\psi \left(c(\mathbf{n}) + \min \left\{ \lambda_0, \lambda_1 \frac{f_1^n}{f_0^n} \right\} \right) f_0^n d\mu^n, \\
 &= \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi (c(\mathbf{n}) + g(z_{\mathbf{n}})).
 \end{aligned}$$

Por tanto se tiene (3.14). Nuevamente por Teorema 2 se sigue que la igualdad en (3.14) se obtiene si y solo si la regla de decisión ϕ cumple la condición Δ_ψ . ■

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Con base en el teorema anterior se tiene que la condición Δ_ψ da la estructura de la regla de decisión óptima. Recordemos la notación presente en esta condición (3.13),

$$\phi_{\mathbf{n}} \geq I_{\{\lambda_0/\lambda_1 \lesssim z_{\mathbf{n}}\}} \quad \text{significa} \quad I_{\{\lambda_0/\lambda_1 < z_{\mathbf{n}}\}} \leq \phi_{\mathbf{n}} \leq I_{\{\lambda_0/\lambda_1 \leq z_{\mathbf{n}}\}}.$$

Observemos que en el caso de que para alguna etapa con su grupo de observaciones x^n se tenga que $f_0^n(x^n) = 0$ y $f_1^n(x^n) > 0$, entonces $z_{\mathbf{n}} = \infty$. De lo cual la regla de decisión óptima será $\phi_{\mathbf{n}}^* = 1$ (se rechaza H_0 , a favor de H_1). Lo cual tiene sentido porque la única distribución que está aportando información sobre las observaciones es f_1 .

Continuando con el análisis, debido a que se tiene una cota inferior alcanzable para la función de Lagrange $L_0(\psi, \phi)$, cuando la regla de decisión ϕ cumple ciertas condiciones, entonces dejaremos fija a la regla ϕ .

- Para cualquier $\psi \in \mathcal{F}_0$ se tiene:

$$L_0(\psi) := L_0(\psi, \phi^*) = \inf_{\phi} L_0(\psi, \phi) \quad \text{si y solo si } \phi^* \text{ cumple la condición } \Delta_\psi; \quad (3.15)$$

donde

$$L_0(\psi) := \sum_{r=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) + g(z_{\mathbf{n}})). \quad (3.16)$$

3.1.2. Regla para óptima bajo H_0 : horizonte finito

Ahora, con base en que ya se tienen las condiciones para alcanzar el ínfimo de la función de Lagrange escrita en (3.16) sobre todas las reglas de decisión ϕ , entonces se buscará la cota inferior para las reglas de paro ψ . Primero vamos a trabajar el caso en el que se tiene un número fijo de etapas N (horizonte finito o caso truncado) y después haremos el análisis, cuando el número de etapas crece indefinidamente (horizonte infinito o no finito).

- Sea $N \geq 1$ entero, se define la clase de reglas de paro que detienen el experimento en la etapa N o antes como $\mathcal{F}^N = \{\psi : t_{\mathbf{n}}^{\psi} = 0, \text{ para cualquier } \mathbf{n} \in \mathcal{G}^{N+1}\} \subset \mathcal{F}_0$. De esta definición, se puede demostrar fácilmente la siguiente cadena:

$$\mathcal{F}^1 \subset \mathcal{F}^2 \subset \dots \subset \mathcal{F}^N \subset \mathcal{F}^{N+1} \subset \dots \subset \mathcal{F}_0.$$

- Para cualquier $\psi \in \mathcal{F}^N$, con base en (3.16) y dado que $t_{\mathbf{n}}^{\psi} = 0$ para cualquier $\mathbf{n} \in \mathcal{G}^{N+1}$ se obtiene la función de Langrange truncada:

$$L_0^N(\psi) = \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} \int s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f_0^n + l_{\mathbf{n}}) d\mu^n + \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} (c(\mathbf{n}) f_0^n + l_{\mathbf{n}}) d\mu^n.$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

La cual con los términos de este contexto, resulta en

$$L_0^N(\psi) = \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} \left(c(\mathbf{n}) + g(z_{\mathbf{n}}) \right) + \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n}) + g(z_{\mathbf{n}}) \right). \quad (3.17)$$

Ahora corresponderá encontrar una cota inferior para (3.17).

Como la función (3.17) está en términos de E_0 y $g(z)$, entonces al igual que como se realizó en el capítulo 2, vamos a trabajar las funciones llamadas V_r^N , pero ahora éstas estarán en términos de la función f_0 y de la esperanza E_0 , a diferencia del capítulo 2 en donde se trabajó en términos de operadores llamados I_r y M_r .

En la teoría desarrollada hasta este punto no habíamos necesitado de la independencia de las observaciones x^n explícitamente, sin embargo a continuación se define un operador que trabaja con la razón de verosimilitud de un grupo en particular y para poder trabajar las observaciones de dicho grupo, es necesario que ellas sean independientes.

Definición 28.

Sea $r \geq 1$ cualquiera y $z \geq 0$. Para cualquier función $v(z)$ medible y no negativa. Se define el operador:

$$J_r v(z) := \sum_{m \in \mathcal{G}} p_r(m) E_0 v \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) = \sum_{m \in \mathcal{G}} p_r(m) E_0 v(z z_m). \quad (3.18)$$

Definición 29.

Dadas $\lambda_0 > 0$ y $\lambda_1 > 0$ fijas, y para cualquier $z \in \mathbb{R}^+$. Para $r = N$ sea

$$V_r^N(z) = V_r^N(z; \lambda_0, \lambda_1) := g(z; \lambda_0, \lambda_1) \equiv \min\{\lambda_0, \lambda_1 z\}, \quad (3.19)$$

y, de manera recursiva para $r = N - 1, N - 2, \dots, 1$,

$$\begin{aligned} V_r^N(z) = V_r^N(z; \lambda_0, \lambda_1) &:= \min \left\{ g(z; \lambda_0, \lambda_1), \bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z; \lambda_0, \lambda_1) \right\} \\ &= \min \left\{ g(z), \bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z) \right\}. \end{aligned} \quad (3.20)$$

Por simplicidad trabajaremos a las funciones V_r^N con $1 \leq r \leq N$ y asociadas solo en dependencia de la variable z , ya que λ_0 y λ_1 serán constantes fijas, a menos que se necesite la

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

dependencia estos multiplicadores, pero en este caso será marcado explícitamente.

El siguiente lema establece que la sucesión de funciones $V_r^N(z)$ es decreciente respecto al nivel de truncamiento N .

Lema 16.

Para cualesquiera entero $N \geq 1$ y $r = N, \dots, 1$ se tiene que

$$V_r^N(z) \geq V_r^{N+1}(z), \quad (3.21)$$

para cualquier $z \geq 0$.

Demostración del Lema 16:

Aplicaremos inducción inversa ¹ sobre r . Sea $N \geq 1$ fijo pero cualquiera, y sea $r = N$,

$$V_N^N(z) = g(z) \geq \min \left\{ g(z), \bar{c}_{N+1} + J_{N+1} V_{N+1}^{N+1}(z) \right\} = V_N^{N+1}(z),$$

la primera igualdad se obtiene de (3.19) y la segunda por (3.20).

Hipótesis de Inducción: Supongamos que (3.21) se cumple para algún $r = N, N-1, \dots, 2$. Para dicha r , por propiedades de la esperanza y debido a que $\sum_{m \in \mathcal{G}} p_r(m) \leq 1$ se obtiene:

$$J_r V_r^N(z) = \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^N \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) \geq \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^{N+1} \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) = J_r V_r^{N+1}(z).$$

Por tanto, para cualquier $z \geq 0$,

$$V_{r-1}^N(z) = \min \left\{ g(z), \bar{c}_r + J_r V_r^N(z) \right\} \geq \min \left\{ g(z), \bar{c}_r + J_r V_r^{N+1}(z) \right\} = V_{r-1}^{N+1}(z).$$

Entonces $V_{r-1}^N(z) \geq V_{r-1}^{N+1}(z)$. Por lo tanto se tiene (3.21) para todo $r = N, \dots, 1$. ■

¹La inducción inversa consiste demostrar la propiedad para $r = N$, después se supone la verdad de la propiedad para $r = k$ y se demuestra para $r = k - 1$; demostrando así la veracidad de la propiedad para todos los enteros desde $r = N$ hasta 1.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Condición Π_0^N . (Regla de paro truncada óptima, con f_0)

Una regla de paro ψ satisface la condición Π_0^N , si:

Para toda $r = 1, \dots, N - 1$ y para cualquier $\mathbf{n} \in \mathcal{P}_r$

$$\psi_{(\mathbf{n};r)} \geq I_{\{g(z_{\mathbf{n}}) \lesssim \bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z_{\mathbf{n}})\}}, \quad \text{sobre el conjunto } T_{\mathbf{n}}^\psi \cap \mathcal{K}_0^{\mathbf{n}}, \mu^{\mathbf{n}}\text{-c.t.p.} \quad (3.22)$$

Y para toda $\mathbf{n} \in \mathcal{P}_N$,

$$\psi_{\mathbf{n}} = 1, \quad \text{sobre el conjunto } T_{\mathbf{n}}^\psi \cap \mathcal{K}_0^{\mathbf{n}}, \mu^{\mathbf{n}}\text{-c.t.p.} \quad (3.23)$$

El siguiente teorema es equivalente al Teorema 3 del Capítulo 2, y determina la cota inferior de la función de Lagrange truncada.

Teorema 12.

Dado $N \geq 1$ entero. Para cualquier $\psi \in \mathcal{F}^N$ se cumple que

$$L_0^N(\psi) \geq \bar{c}_1 + J_1 V_1^N(1). \quad (3.24)$$

Se obtiene la igualdad en (3.24) si y solo si ψ satisface la condición Π_0^N .

Demostración del Teorema 12:

Sea $N \geq 1$ entero fijo y sea $\psi \in \mathcal{F}^N$ una regla de paro truncada cualquiera.. Para toda $z \geq 0$ y para cualquier $r = 1, 2, \dots, N$, se define

$$Q_r^N(z) := \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi(c(\mathbf{n}) + g(z_{\mathbf{n}})) + \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^\psi(c(\mathbf{n}) + V_r^N(z z_{\mathbf{n}})). \quad (3.25)$$

Es importante mencionar que el funcional Q_r^N definido anteriormente depende también de la regla de paro ψ , sin embargo para no saturar la notación, esta dependencia no se expresa. De hecho este funcional es equivalente a $B_r^\psi(v)$ definido en (2.47) (Capítulo 2).

Ahora, por un lado con base en (3.17), en particular se tiene que $Q_N^N(1) = L_0^N(\psi)$. Por otro lado

$$\begin{aligned} Q_1^N(1) &= \sum_{n \in \mathcal{G}} p_1(n) E_0 t_n^\psi(c(n) + V_1^N(z_n)), \\ &= E_0 \sum_{n \in \mathcal{G}} p_1(n) c(n) + \sum_{n \in \mathcal{G}} p_1(n) E_0 V_1^N(z_n) = \bar{c}_1 + J_1 V_1^N(1). \end{aligned}$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Dado que en (2.50) se demuestra que $B_r^\psi(v)$ es finito; entonces también consideraremos que $Q_r^N(z) < \infty$ para todo $z \in [0, \infty)$. Por lo tanto $\bar{c}_1 + J_1 V_1^N(1) < \infty$.

Vamos a demostrar que

$$L_0^N(\psi) = Q_N^N(1) \geq Q_{N-1}^N(1) \geq \cdots \geq Q_1^N(1) = \bar{c}_1 + J_1 V_1^N(1). \quad (3.26)$$

y determinaremos cómo deben ser las componentes de $\psi \in \mathcal{F}^N$ sucesivamente para que todas las desigualdades en (3.26) se conviertan en igualdades.

Primero vamos a demostrar que para $1 \leq r \leq N - 1$ se sigue

$$Q_{r+1}^N(1) \geq Q_r^N(1). \quad (3.27)$$

De la definición (3.25), la desigualdad (3.27) es equivalente a

$$\begin{aligned} \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi(c(\mathbf{n}) + g(z_{\mathbf{n}})) &+ \sum_{\mathbf{n} \in \mathcal{G}^{r+1}} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^\psi(c(\mathbf{n}) + V_{r+1}^N(z_{\mathbf{n}})) \\ &\geq \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^\psi(c(\mathbf{n}) + V_r^N(z_{\mathbf{n}})). \end{aligned} \quad (3.28)$$

Por el Teorema de Fubini, el lado izquierdo de (3.28) es igual a

$$\begin{aligned} &\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^\psi \psi_{\mathbf{n}}(c(\mathbf{n}) + g(z_{\mathbf{n}})) + \sum_{(\mathbf{n}, m) \in \mathcal{G}^{r+1}} p_{(\mathbf{n}, m)} E_0 t_{(\mathbf{n}, m)}^\psi(c((\mathbf{n}, m)) + V_{r+1}^N(z_{(\mathbf{n}, m)})) \\ = &\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^\psi(\psi_{\mathbf{n}}(c(\mathbf{n}) + g(z_{\mathbf{n}}))) \right. \\ &\quad \left. + \sum_{m \in \mathcal{G}} p_{r+1}(m)(1 - \psi_{\mathbf{n}}) \int (c(\mathbf{n}) + c(m) + V_{r+1}^N(z_{(\mathbf{n}, m)})) f_0^{(m)} d\mu^m \right] f_0^{(\mathbf{n})} d\mu^{\mathbf{n}}. \\ = &\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \left[\int t_{\mathbf{n}}^\psi(\psi_{\mathbf{n}}(c(\mathbf{n}) + g(z_{\mathbf{n}}))) + \sum_{m \in \mathcal{G}} p_{r+1}(m)(1 - \psi_{\mathbf{n}}) c(\mathbf{n}) \int f_0^{(m)} d\mu^m \right. \\ &\quad \left. + \sum_{m \in \mathcal{G}} p_{r+1}(m)(1 - \psi_{\mathbf{n}}) \int (c(m) + V_{r+1}^N(z_{(\mathbf{n}, m)})) f_0^{(m)} d\mu^m \right] f_0^{(\mathbf{n})} d\mu^{\mathbf{n}}. \\ = &\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi \left[c(\mathbf{n}) + \psi_{\mathbf{n}} g(z_{\mathbf{n}}) + (1 - \psi_{\mathbf{n}}) \left(\sum_{m \in \mathcal{G}} p_{r+1}(m) c(m) + \right. \right. \\ &\quad \left. \left. + \sum_{m \in \mathcal{G}} p_{r+1}(m) \int V_{r+1}^N(z_{\mathbf{n}} z_m) f_0^{(m)} d\mu^m \right) \right] f_0^{(\mathbf{n})} d\mu^{\mathbf{n}}. \\ = &\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^\psi \left[c(\mathbf{n}) + \psi_{\mathbf{n}} g(z_{\mathbf{n}}) + (1 - \psi_{\mathbf{n}}) \left(\bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z_{\mathbf{n}}) \right) \right] f_0^{(\mathbf{n})} d\mu^{\mathbf{n}}. \end{aligned} \quad (3.29)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Aplicando a (3.29) la desigualdad (2.16) del Lema 1 (capítulo 2) con $\phi \equiv \psi_{\mathbf{n}}$, $F_1 \equiv g(z_{\mathbf{n}})$ y $F_2 \equiv \bar{c}_{r+1} + J_{r+1}V_{r+1}^N(z_{\mathbf{n}})$, se obtiene que una cota inferior de (3.29), sobre todas las $\psi_{\mathbf{n}}$, $\mathbf{n} \in \mathcal{G}^r$.

$$\begin{aligned} & \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} \int t_{\mathbf{n}}^{\psi} \left[c(\mathbf{n}) + \psi_{\mathbf{n}} g(z_{\mathbf{n}}) + (1 - \psi_{\mathbf{n}}) \left(\bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z_{\mathbf{n}}) \right) \right] f_0^{(\mathbf{n})} d\mu^{\mathbf{n}} \\ & \geq \sum_{\mathbf{n} \in \mathcal{G}^r} p(\mathbf{n}) \int t_{\mathbf{n}}^{\psi} \left(c(\mathbf{n}) + \min \left\{ g(z_{\mathbf{n}}), \bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z_{\mathbf{n}}) \right\} \right) f_0^{(\mathbf{n})} d\mu^{\mathbf{n}} \\ & = \sum_{\mathbf{n} \in \mathcal{G}^r} p(\mathbf{n}) E_0 t_{\mathbf{n}}^{\psi} (c(\mathbf{n}) + V_r^N(z_{\mathbf{n}})). \end{aligned} \quad (3.30)$$

Ahora aplicando (2.17) del mismo Lema 1, la desigualdad (3.30) se alcanza si y solo si para toda $\mathbf{n} \in \mathcal{P}_r$

$$\psi_{\mathbf{n}} \simeq I_{\{g(z_{\mathbf{n}}) \lesssim \bar{c}_{r+1} + J_{r+1} V_{r+1}^N(z_{\mathbf{n}})\}}, \quad (3.31)$$

μ^n -c.t.p. sobre $T_{\mathbf{n}}^{\psi} \cap \mathcal{K}_0^{\mathbf{n}}$. Lo cual demuestra (3.27).

Ahora, aplicando recursivamente (3.27) para $r = 1, 2, \dots, N - 1$ se obtiene la prueba de la sucesión de desigualdades en (3.26). Además, se establece una sucesión de igualdades para todas las desigualdades en (3.26) si y solo si (3.31) satisface $\mathbf{n} \in \mathcal{P}_k$ μ^n -c.t.p. sobre $T_{\mathbf{n}}^{\psi} \cap \mathcal{K}_0^{\mathbf{n}}$, para toda $r = 1, 2, \dots, N - 1$, lo cual coincide con la condición Π_0^N , y demuestra el Teorema 12. ■

Puede notarse que si ψ^* satisface la condición Π_0^N , entonces $\psi^* \in \mathcal{F}^N$. Una manera de expresar el anterior teorema es

$$L_0^N(\psi^*) = \inf_{\psi' \in \mathcal{F}^N} L_0^N(\psi') \quad \text{si y solo si } \psi^* \text{ cumple la condición } \Pi_0^N, \quad (3.32)$$

donde

$$L_0^N(\psi^*) = \bar{c}_1 + J_1 V_1^N(1). \quad (3.33)$$

Condición (Π_0^N, Δ_{ψ}) (Prueba secuencial truncada óptima)

Una prueba secuencial (ψ, ϕ) satisface la condición (Π_0^N, Δ_{ψ}) , si ψ cumple la *condición* Π_0^N y, para esta regla de paro la función ϕ cumple la *condición* Δ_{ψ} .

- Para $N \geq 1$, se define la clase de pruebas secuenciales cuya regla de paro detiene el experimento en la etapa N o antes como: $\mathfrak{S}^N = \{(\psi, \phi) : \psi \in \mathcal{F}^N\}$.

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Debido a la propiedad $\mathcal{F}^N \subset \mathcal{F}_0$ se sigue que $\mathfrak{S}^N \subset \mathfrak{S}_0$, para cualquier $N \geq 1$.

El siguiente teorema establece condiciones necesarias y suficientes para que una prueba secuencial truncada $\psi \in \mathcal{F}^N$ alcance el ínfimo de la función de Lagrange truncada L_0^N (3.17).

Teorema 13.

Si una prueba secuencial (ψ^, ϕ^*) satisface la condición $(\Pi_0^N, \Delta_{\psi^*})$ entonces*

$$L_0^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_0^N(\psi, \phi). \quad (3.34)$$

Además, si existe alguna prueba secuencial $(\psi^, \phi^*) \in \mathfrak{S}^N$ tal que satisface (3.34), entonces (ψ^*, ϕ^*) cumple la condición $(\Pi_0^N, \Delta_{\psi^*})$.*

Demostración del Teorema 13:

Suficiencia: Sea (ψ^*, ϕ^*) tal que satisface la condición $(\Pi_0^N, \Delta_{\psi^*})$. Como ϕ^* satisface la condición Δ_{ψ^*} , entonces por (3.15) se tiene que:

$$L_0(\psi, \phi^*) = \inf_{\phi} L_0(\psi, \phi), \quad \text{para cualquier } \psi \in \mathcal{F}_0.$$

Por otro lado, como ψ^* cumple Π_0^N , entonces $\psi^* \in \mathcal{F}^N$ y de esto $L_0^N(\psi^*) = L_0(\psi^*, \phi^*)$. Por tanto, de lo anterior y de (3.32) se tiene que:

$$L_0^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_0^N(\psi, \phi).$$

Esto demuestra la primera parte del teorema.

Necesidad: Supongamos que existe una prueba secuencial (ψ^*, ϕ^*) con $\psi^* \in \mathcal{F}^N$ y tal que cumple (3.34). Entonces (ψ^*, ϕ^*) minimiza a la función de Lagrange (3.17), por tanto por (3.32) y (3.15) (segunda parte) se tiene que ψ^* y ϕ^* cumplen, respectivamente, las condiciones Π_0^N y Δ_{ψ^*} . Por lo tanto, (ψ^*, ϕ^*) satisface la condición $(\Pi_0^N, \Delta_{\psi^*})$. ■

El siguiente corolario marca las condiciones suficientes que debe cumplir una regla secuencial truncada para que ella conlleve a un costo promedio mínimo del experimento, comparada con cualquier otra prueba aún cuando sus probabilidades de error sean menores.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Cololario 11.

Si una prueba secuencial truncada (ψ^, ϕ^*) satisface la condición $(\Pi_0^N, \Delta_{\psi^*})$, entonces para cualquier otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}^N$ tal que*

$$\alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (3.35)$$

se tiene que

$$K_0(\psi) \geq K_0(\psi^*). \quad (3.36)$$

La desigualdad (3.36) es estricta, si al menos una de las desigualdades en (3.35) lo es.

Demostración del Corolario 11:

Como (ψ^*, ϕ^*) satisface la condición $(\Pi_0^N, \Delta_{\psi^*})$ entonces, por Teorema 13, se tiene que $L_0^N(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}^N} L_0^N(\psi, \phi)$. Ahora sea $(\psi, \phi) \in \mathfrak{S}^N$ cualquier otra prueba secuencial truncada tal que cumple (3.35). Debido a que $\mathfrak{S}^N \subseteq \mathfrak{S}_0$, aplicando por Corolario 1 del Capítulo 2 y considerando que estamos en el contexto \mathcal{M}_0 , con la función de distribución f_0 , entonces se sigue (3.36). Del mismo corolario se sigue que la desigualdad (3.36) es estricta si al menos una de las desigualdades en (3.35) lo es. ■

3.1.3. Regla de paro óptima bajo H_0 : horizonte infinito

Nuevamente vamos a extender los resultados anteriores para el caso cuando el número de etapas crece indefinidamente ($N \rightarrow \infty$) conocido como horizonte infinito, adaptado a nuestro contexto de observaciones i.i.d. con (f_0) y tamaños de grupos independientes.

Primero vamos a establecer condiciones que determinan que las reglas de paro que trabajaremos son reglas convergentes a su respectiva función de Lagrange, es decir, se cumple $\lim_{N \rightarrow \infty} L_0^N(\psi) = L_0(\psi)$. Los siguientes tres lemas nos darán las condiciones suficientes para esto.

Cabe resaltar que para los siguientes dos lemas se considera un parámetro θ cualquiera, pero fijo.

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Lema 17.

Si para algún $\mathbf{n} \in \mathcal{G}^k$, $k \geq 1$ y alguna regla de paro ψ es tal que

$$\sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} \rightarrow 0 \text{ cuando } r \rightarrow \infty, \quad (3.37)$$

entonces

$$E_{\theta} s_{\mathbf{n}}^{\psi} + \sum_{r=1}^{\infty} \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{n}, \mathbf{m}}^{\psi} = E_{\theta} t_{\mathbf{n}}^{\psi}. \quad (3.38)$$

Demostración del Lema 17. Sea r cualquier número natural, entonces

$$\begin{aligned} & \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^{r+1}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} = \\ &= \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^r} \sum_{i \in \mathcal{G}} p_{\mathbf{m}} p(i) E_{\theta} t_{\mathbf{n}, \mathbf{m}, i}^{\psi} \\ &= \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} \left(E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{i \in \mathcal{G}} p(i) E_{\theta} t_{\mathbf{n}, \mathbf{m}, i}^{\psi} \right) = \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} \left(E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - E_{\theta} t_{\mathbf{n}, \mathbf{m}, i}^{\psi} \right). \end{aligned} \quad (3.39)$$

La última igualdad se debe a que $\sum_{i \in \mathcal{G}} p(i) = 1$ por la definición de la distribución del grupo. Por otro lado, por definición (3.5) y dado que $\mathbf{n} \in \mathcal{G}^k$, $\mathbf{m} \in \mathcal{G}^r$ e $i \in \mathcal{G}$:

$$\begin{aligned} t_{\mathbf{n}, \mathbf{m}}^{\psi} - t_{\mathbf{n}, \mathbf{m}, i}^{\psi} &= (1 - \psi_{n_1}) \cdots (1 - \psi_{n_r}) (1 - \psi_{n, m_1}) \cdots (1 - \psi_{n, m_{r-1}}) - \\ &\quad - (1 - \psi_{n_1}) \cdots (1 - \psi_{n_r}) (1 - \psi_{n, m_1}) \cdots (1 - \psi_{n, m_{r-1}}) (1 - \psi_{n, m}) \\ &= (1 - \psi_{n_1}) \cdots (1 - \psi_{n_r}) (1 - \psi_{n, m_1}) \cdots (1 - \psi_{n, m_{r-1}}) (\psi_{n, m}) \\ &= t_{\mathbf{n}, \mathbf{m}}^{\psi} \psi_{n, m} = s_{\mathbf{n}, \mathbf{m}}^{\psi}. \end{aligned} \quad (3.40)$$

Esto último es por la definición (3.6). Por tanto regresando a (3.39) y aplicando propiedades del valor esperado E_{θ} se tiene que

$$\sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{n}, \mathbf{m}}^{\psi} = \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^{r+1}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi}. \quad (3.41)$$

Al aplicar la suma desde $r = 1$ hasta $r = k$ en ambos lados de (3.41), se obtiene

$$\begin{aligned} \sum_{r=1}^k \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{n}, \mathbf{m}}^{\psi} &= \sum_{r=1}^k \left(\sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^{r+1}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} \right) \\ &= \sum_{\mathbf{m} \in \mathcal{G}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^{k+1}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi}. \end{aligned} \quad (3.42)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Pasando al límite en (3.42) cuando $k \rightarrow \infty$ y con base en la hipótesis (3.37) entonces en el lado derecho de la última expresión tenemos,

$$\sum_{r=1}^{\infty} \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{n}, \mathbf{m}}^{\psi} = \sum_{\mathbf{m} \in \mathcal{G}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi}. \quad (3.43)$$

Por otro lado, aplicando las definiciones (3.5) y (3.6) a la igualdad (3.40) y propiedades del valor esperado para cualquier $\mathbf{n} \in \mathcal{G}^k$ se obtiene que:

$$E_{\theta} s_{\mathbf{n}}^{\psi} = E_{\theta} t_{\mathbf{n}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{n}, \mathbf{m}}^{\psi}. \quad (3.44)$$

Finalmente, sumando las expresiones (3.44) y (3.43) respectivamente en ambos lados obtenemos (3.38). ■

Lema 18.

Para una regla de paro ψ

$$\sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} t_{\mathbf{m}}^{\psi} \rightarrow 0, \quad \text{cuando } r \rightarrow \infty, \quad (3.45)$$

si y solo si

$$\sum_{r=1}^{\infty} \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{m}}^{\psi} = 1. \quad (3.46)$$

Demostración del Lema 18.

Supongamos que (3.45) se cumple. Ahora, repitiendo los pasos desde (3.39) a (3.42) del Lema 17, pero con $t_{\mathbf{m}}^{\psi}$ en lugar de $t_{\mathbf{n}, \mathbf{m}}^{\psi}$, sustituyendo en (3.43) obtendremos:

$$\sum_{r=1}^{\infty} \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{m}}^{\psi} = \sum_{\mathbf{m} \in \mathcal{G}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{m}}^{\psi} = 1. \quad (3.47)$$

Esto último se debe a que $t_{\mathbf{m}}^{\psi} = 1$ para todo $\mathbf{m} \in \mathcal{G}$ por (3.5). Por tanto (3.46) está demostrada.

Ahora supongamos que (3.46) se satisface. Nuevamente, repitiendo los pasos de la demostración del Lema 17 se obtiene:

$$\sum_{r=1}^k \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_{\theta} s_{\mathbf{m}}^{\psi} = \sum_{\mathbf{m} \in \mathcal{G}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{m}}^{\psi} - \sum_{\mathbf{m} \in \mathcal{G}^{k+1}} p_{\mathbf{m}} E_{\theta} t_{\mathbf{m}}^{\psi}. \quad (3.48)$$

El lado izquierdo de (3.48) tiende a 1, cuando $k \rightarrow \infty$, por hipótesis (3.46). El primer término del lado derecho también es igual a 1, por (3.5). Por lo tanto, el segundo término del lado derecho de (3.48) tiende a 0, así (3.45) se satisface. ■

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Lema 19.

Sea $\psi \in \mathcal{F}_0$ tal que $K_0(\psi) < \infty$. Entonces

$$\sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} c(\mathbf{n}) E_0 t_{\mathbf{n}}^{\psi} \rightarrow 0, \text{ cuando } N \rightarrow \infty. \quad (3.49)$$

Demostración del Lema 19.

Sea $\psi \in \mathcal{F}_0$ tal que $K_0(\psi) < \infty$. Por definición e hipótesis se tiene que

$$K_0(\psi) = \sum_{k=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E_0 s_{\mathbf{n}}^{\psi} < \infty,$$

así,

$$\sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E_0 s_{\mathbf{n}}^{\psi} \rightarrow 0, \quad (3.50)$$

cuando $N \rightarrow \infty$. Sea $c(\mathbf{n}|N) = c(n_1) + \dots + c(n_N)$ para cualquier $\mathbf{n} \in \mathcal{G}^k$ y $k \geq N$.

Se sigue del Lema 17 que

$$\begin{aligned} \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} c(\mathbf{n}) E_0 t_{\mathbf{n}}^{\psi} &= \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} c(\mathbf{n}) \left(E_0 s_{\mathbf{n}}^{\psi} + \sum_{r=1}^{\infty} \sum_{\mathbf{m} \in \mathcal{G}^r} p_{\mathbf{m}} E_0 s_{\mathbf{n}, \mathbf{m}}^{\psi} \right) \\ &\leq \sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} c(\mathbf{n}|N) p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} \leq \sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} c(\mathbf{n}) E_0 s_{\mathbf{n}}^{\psi} \rightarrow 0, \end{aligned}$$

cuando $N \rightarrow \infty$, esto último en virtud de (3.50). ■

Antes de continuar recordemos del Capítulo 2 que cualquier regla de paro $\psi \in \mathcal{F}_0$ puede ser truncada en el nivel N , de la siguiente manera: $\psi^N = (\psi_{n_1}, \psi_{n_2}, \dots, \psi_{n_{N-1}}, 1, \dots)$ para toda $\mathbf{n} = (n_1, n_2, \dots, n_{N-1}) \in \mathcal{G}^{N-1}$. Debido a esto a cualquier regla de paro ψ se le puede aplicar la función de Lagrange Truncada (3.17) con $L_0^N(\psi) = L_0^N(\psi^N)$. De esta manera es fácil determinar (3.17) sin importar si la regla de paro ψ es truncada o no.

Ahora sí con base en los tres lemas anteriores estamos en posibilidad de demostrar que, en este contexto de un problema de hipótesis simples \mathcal{M}_0 , toda regla de paro con tiempo de paro finito es una regla truncable, o sea su función de Lagrange truncada converge a la función de Lagrange general, cuando el nivel de truncamiento crece indefinidamente.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Teorema 14.

Para cualquier regla de paro $\psi \in \mathcal{F}_0$ se tiene que

$$\lim_{N \rightarrow \infty} L_0^N(\psi) = L_0(\psi), \quad (3.51)$$

donde $L_0(\psi)$ se definió en (3.16).

Demostración del Teorema 14: Supongamos primero que $L_0(\psi) < \infty$. Con base en las definiciones de las funciones de Lagrange no truncada (3.16) y truncada (3.17) se tiene:

$$\begin{aligned} & L_0(\psi) - L_0^N(\psi) = \\ &= \sum_{k=N}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) + g(z_{\mathbf{n}})) - \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} c(\mathbf{n}) E_0 t_{\mathbf{n}}^{\psi} - \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} g(z_{\mathbf{n}}). \end{aligned} \quad (3.52)$$

Debido a que $L_0(\psi)$ es convergente, ya que $L_0(\psi) < \infty$, el primer término de (3.52) tiende a cero, cuando $N \rightarrow \infty$. Además, esto implica que $K_0(\psi) < \infty$, entonces por Lema 19 (véase (3.49)), el segundo término de (3.52) también tiende a cero cuando $N \rightarrow \infty$.

Finalmente, $g(z) \leq \lambda_0$ para toda $z \geq 0$. Así,

$$\sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} g(z_{\mathbf{n}}) \leq \lambda_0 \sum_{\mathbf{n} \in \mathcal{G}^N} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} \rightarrow 0, \quad \text{cuando } N \rightarrow \infty.$$

Esta última convergencia se sigue por el Lema 18, porque por hipótesis $\psi \in \mathcal{F}_0$ implica (3.46). Así se ha demostrado (3.51) en el caso de que $L(\psi) < \infty$.

Ahora supongamos que $L_0(\psi) = \infty$, entonces

$$L_0^N(\psi) \geq \sum_{k=1}^{N-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) + g(z_{\mathbf{n}})) \rightarrow L_0(\psi) = \infty, \quad \text{cuando } N \rightarrow \infty,$$

entonces (3.51) se satisface en este caso también. ■

En este contexto de horizonte infinito estamos trabajando con reglas de paro truncables ($L_0^N(\psi) \rightarrow L_0(\psi)$) bajo los multiplicadores de Lagrange fijos, entonces ahora ya podemos dedicarnos a determinar la cota inferior de la función $L_0(\psi)$, pero para esto es necesario definir los límites de las funciones previamente truncadas.

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Con base en el Lema 16 se tiene que para cualesquiera $r \geq 1$ y $z \geq 0$ las funciones $V_N^N(z)$ y $V_r^N(z)$ definidas en (3.19)-(3.20) forman una sucesión $\{V_r^N(z)\}_{N \geq r}$ acotada y decreciente (respecto al nivel de truncamiento) de funciones $V_r^N(z) \geq V_r^{N+1}(z) \geq 0$ (por (3.21)), para $N \geq 1$. Por lo tanto existe su límite cuando $N \rightarrow \infty$. Entonces para cualquier $r \geq 1$,

$$V_r(z) := \lim_{N \rightarrow \infty} V_r^N(z). \quad (3.53)$$

Con base en (3.53) y tomando el límite a (3.20) cuando $N \rightarrow \infty$ se obtiene que para cualquier $r \geq 1$;

$$V_r(z) = \min \left\{ g(z), \bar{c}_{r+1} + J_{r+1} V_{r+1}(z) \right\} \quad (3.54)$$

Cabe mencionar que las relaciones (3.53) y (3.54), al igual que las anteriores, traen consigo la dependencia de la distribución del grupo de observaciones r en cuestión, debido a que $\nu_1, \dots, \nu_r \dots$ son v.a. solo independientes. En el caso de tener, además, tamaños de grupo idénticamente distribuidos (i.d.), dicha dependencia no será necesaria, como será trabajada en la siguiente sección.

En el Lema 13 (capítulo 2) o la fórmula (2.124) se encontró el ínfimo de la función de $L_{\mathcal{M}}(\psi)$, ahora tendremos la cota respectiva para nuestra función $L_0(\psi)$, cuya manera equivalente en nuestro contexto \mathcal{M}_0 es la siguiente.

Lema 20.

$$\inf_{\psi \in \mathcal{F}_0} L_0(\psi) = \bar{c}_1 + J_1 V_1(1). \quad (3.55)$$

Demostración del Lema 20: Sean $U = \inf_{\psi \in \mathcal{F}_0} L_0(\psi)$ y $U_N = \inf_{\psi \in \mathcal{F}^N} L_0^N(\psi)$.

Debido al Teorema 12 y a (3.33), $U_N = \bar{c}_1 + J_1 V_1^N(1)$ para cualquier $N = 1, 2, \dots$. Debido a que $\mathcal{F}^N \subset \mathcal{F}_0$, entonces $U_N \geq U$ para cualquier $N = 1, 2, \dots$, así que $\lim_{N \rightarrow \infty} U_N \geq U$. Ahora se demostrará que $\lim_{N \rightarrow \infty} U_N = U$.

Supongamos lo contrario, *i.e* supongamos que $\lim_{N \rightarrow \infty} U_N = U + 4\varepsilon$ para algún $\varepsilon > 0$; entonces para alguna N suficientemente grande

$$U_N \geq U + 3\varepsilon, \quad (3.56)$$

Por definición de U existe una regla de paro $\psi \in \mathcal{F}_0$, tal que $U \leq L_0(\psi) \leq U + \varepsilon$. Por el Teorema 14 $L_0^N(\psi) \rightarrow L_0(\psi)$, cuando $N \rightarrow \infty$, entonces para un N suficientemente grande $L_0^N(\psi) \leq U + 2\varepsilon$. Porque por definición $U_N \leq L_0^N(\psi)$ tenemos que $U_N \leq U + 2\varepsilon$ para toda N grande, lo cual contradice (3.56).

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Por lo tanto, $U = \lim_{N \rightarrow \infty} U_N = \bar{c}_1 + \lim_{N \rightarrow \infty} J_1 V_1^N(1) = \bar{c}_1 + J_1 V_1(1)$. ■

Es importante marcar dos cosas sobre el ínfimo anterior (parte derecha de (3.55)), por un lado esta cota no depende de ninguna regla de paro ψ , mas para llegar a ella sí se requiere analizar las condiciones necesarias y/o suficientes de la estructura de la regla que la alcanza. Por otro lado, como $\bar{c}_1 < \infty$ (por hipótesis los costos promedios de los grupos son finitos), y además $J_1 V_1(1) = \sum_{m \in \mathcal{G}} p_1(m) E_0 V_1(Z_m) \leq E_0 V_1(Z_m) \leq \min \{\lambda_0, \lambda_1\} < \infty$. Por lo tanto

$$\bar{c}_1 + J_1 V_1(1) < \infty. \quad (3.57)$$

Condición Π_0^∞ . (Estructura de la regla de paro óptima, bajo f_0)

Una regla de paro ψ satisface la condición Π_0^∞ , si para toda $r \geq 1$ y cualquier $\mathbf{n} \in \mathcal{P}_r$:

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{g(z_{\mathbf{n}}) \lesssim \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_{\mathbf{n}})\}}, \quad \text{sobre el conjunto } T_{\mathbf{n}}^\psi \cap \mathcal{K}_0^{\mathbf{n}}, \quad \mu^{\mathbf{n}}\text{-c.t.p.} \quad (3.58)$$

Recordemos por (3.1) que estamos en el contexto de un problema de pruebas de hipótesis simples $\mathcal{M}_0 = (P_0, P_1, P_0)$, donde las observaciones de los grupos están bajo H_0 . En el Corolario 7 del capítulo 2 se demostró, como un caso particular, que todas las reglas de paro truncables, en este contexto ($\psi \in \mathcal{T}_0$), tienen un tiempo de paro finito ($\psi \in \mathcal{F}_0$) y viceversa, i.e. $\mathcal{T}_0 = \mathcal{F}_0$. Por lo tanto, otra cosa importante que marcar sobre las reglas de paro de \mathcal{M}_0 , es que además de reglas truncables tienen un tiempo de paro finito con probabilidad uno, bajo P_0 , en particular la definida en (3.58).

Teorema 15.

Para cualquier $\psi \in \mathcal{F}_0$ se cumple que

$$L_0(\psi) \geq \bar{c}_1 + J_1 V_1(1). \quad (3.59)$$

Se obtiene la igualdad en (3.59), si ψ satisface la condición Π_0^∞ .

Recíprocamente, si existe una igualdad en (3.59) para alguna $\psi \in \mathcal{F}_0$, entonces ψ satisface la condición Π_0^∞ .

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Demostración del Teorema 15:

Suficiencia: Sea $\psi \in \mathcal{F}_0$ cualquier regla de paro. Para cualesquiera $r = 1, 2, \dots$ y $z \geq 0$ se define

$$Q_r(z) = \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi}(c(\mathbf{n}) + g(z_{\mathbf{n}})) + \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi}(c(\mathbf{n}) + V_r(z z_{\mathbf{n}})), \quad (3.60)$$

(compárese con (3.25)). Ahora, por el Teorema de Convergencia Dominada (TCD) de Lebesgue, se sigue que

$$\lim_{N \rightarrow \infty} Q_r^N(z) = Q_r(z). \quad (3.61)$$

Por la misma razón (TCD), aplicando el límite cuando $r \rightarrow \infty$ a (3.27) obtenemos que para cualquier $r \geq 1$

$$Q_{r+1}(1) \geq Q_r(1). \quad (3.62)$$

Recordemos que el funcional $Q_r(z)$, depende de la regla de paro ψ , y además este es finito.

Análogamente a la obtención de (3.31) a partir de (3.27), a través de los pasos (3.28) (3.29) y (3.30), podemos obtener que existe una igualdad en (3.62), si para toda $\mathbf{n} \in \mathcal{P}_r$

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{g(z_{\mathbf{n}}) \lesssim \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_{\mathbf{n}})\}} \quad (3.63)$$

μ^n -c.t.p. sobre $T_{\mathbf{n}}^{\psi} \cap \mathcal{K}_0^n$. Al aplicar (3.62) repetidamente para $r = 1, 2, \dots$, obtenemos

$$Q_r(1) \geq Q_{r-1}(1) \geq \dots \geq Q_1(1) = \bar{c}_1 + J_1 V_1(1). \quad (3.64)$$

Se obtiene una sucesión de igualdades en (3.64), para cualquier natural r , en el caso de que se cumpla la condición $\Pi_0^{\infty}(\psi)$. En particular, si ψ cumple la condición $\Pi_0^{\infty}(\psi)$, entonces para todo natural $r \geq 1$ se tiene que:

$$\begin{aligned} Q_r(1) &= \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi}(c(\mathbf{n}) + g(z_{\mathbf{n}})) + \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi}(c(\mathbf{n}) + V_r(z_{\mathbf{n}})) = \\ &= \bar{c}_1 + J_1 V_1(1). \end{aligned} \quad (3.65)$$

De esto se sigue que $K_0(\psi) < \infty$, porque de lo contrario

$$\lim_{r \rightarrow \infty} \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi}(c(\mathbf{n}) + g(z_{\mathbf{n}})) \geq K_0(\psi)$$

sería infinito, y esto sería una contradicción con (3.65). Así si $K_0(\psi) < \infty$, entonces en virtud del Lema 19,

$$\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} c(\mathbf{n}) E_0 t_{\mathbf{n}}^{\psi} \rightarrow 0, \text{ cuando } r \rightarrow \infty. \quad (3.66)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Por lo tanto,

$$\sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} V_r(z_{\mathbf{n}}) \leq \lambda_0 \sum_{\mathbf{n} \in \mathcal{G}^r} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^{\psi} \rightarrow 0, \text{ cuando } r \rightarrow \infty, \quad (3.67)$$

porque por suposición $\psi \in \mathcal{F}_0$.

Por otro lado se sigue de (3.65), (3.66) y (3.67) que

$$\lim_{r \rightarrow \infty} \sum_{k=1}^{r-1} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^{\psi} (c(\mathbf{n}) + g(z_{\mathbf{n}})) = L_0(\psi) = \bar{c}_1 + J_1 V_1(1).$$

Esto demuestra la suficiencia del Teorema 15.

Necesidad: Ahora, pasando al límite (3.26) cuando $N \rightarrow \infty$, se tiene que para cualquier $\psi \in \mathcal{F}_0$

$$L_0(\psi) \geq Q_r(1) \geq Q_{r-1}(1) \geq \cdots \geq Q_1(1) = \bar{c}_1 + J_1 V_1(1), \quad (3.68)$$

para toda r número natural. Ahora supongamos que existe una igualdad en (3.59) para alguna regla de paro $\psi \in \mathcal{F}_0$. Se sigue de (3.68) que todas las desigualdades se convierten en igualdades, entonces $Q_{r+1}(1) = Q_r(1)$ para todo natural r . Esto implica nuevamente que se satisface la condición $\Pi_0^{\infty}(\psi)$. ■

En el teorema anterior se ha demostrado que es necesario y suficiente que una regla de paro cumpla la condición Π_0^{∞} para que ella alcance el ínfimo de la función de Lagrange L_0 . Ahora recordemos la notación presente en la estructura de las reglas de paro definidas en dicha condición (3.58). Cada regla de paro se expande como:

$$I_{\{g(z_{\mathbf{n}}) < \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_{\mathbf{n}})\}} \leq \psi_{\mathbf{n}} \leq I_{\{g(z_{\mathbf{n}}) \leq \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_{\mathbf{n}})\}}, \quad (3.69)$$

para toda $r \geq 1$ y cualquier $\mathbf{n} \in \mathcal{P}_r$.

Se reitera que en esta investigación el experimento estadístico secuencial siempre parte (o toma en cuenta) **el primer grupo de observaciones** de cualquier tamaño, o sea, el experimento siempre empieza con la etapa uno ($r = 1$). Es por ello que la estructura de las reglas de paro en la condición (3.58) tienen sentido porque toman en comparación, en cada etapa, funciones que dependen de las observaciones en el grupo r y el siguiente ($r + 1$). Por ejemplo, para la primera etapa $r = 1$, se analiza a las observaciones en este primer grupo, a través de la razón $z_{\mathbf{n}}$ mediante la función $g(z_{\mathbf{n}})$ con $\mathbf{n} \in \mathcal{G}$. Se evalúa nuevamente a $z_{\mathbf{n}}$, del primer grupo, en $\bar{c}_2 + J_2 V_2(z_{\mathbf{n}})$ (obsérvese que en esta segunda función ya se consideran elementos del segundo grupo como su costo promedio \bar{c}_2), y el resultado de esta última evaluación se comparará con $g(z_{\mathbf{n}})$. Al comparar ambas evaluaciones se tienen tres alternativas.

(1) Si se da la desigualdad estricta $g(z_{\mathbf{n}}) < \bar{c}_2 + J_2 V_2(z_{\mathbf{n}})$, las dos indicadoras en (3.69) valdrán uno, y la regla $\psi_{\mathbf{n}} \equiv 1$ con $\mathbf{n} \in \mathcal{G}$; lo cual detendrá al experimento, ya que el valor del primer

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

grupo es menor que el segundo valor. (2) Si se da la igualdad entre $g(z_n)$ y $\bar{c}_2 + J_2 V_2(z_n)$, entonces se tendrá una aleatorización de la regla de paro en esta etapa con $0 \leq \psi_n \leq 1$, lo cual indica que se puede continuar o parar el experimento. Finalmente, (3) si $\bar{c}_2 + J_2 V_2(z_n)$ es menor que $g(z_n)$, entonces ambas indicadores en (3.69) valen cero y $\psi_n \equiv 0$ con $n \in \mathcal{G}$, y por tanto se continua el experimento, ya que el segundo valor es menor que el del primero, por lo cual es mejor continuar y tomar más observaciones. Así sucesivamente se continua para las siguientes etapas $r = 2, 3, \dots$, si no hubo paro en la etapa previa.

Condición $(\Pi_0^\infty, \Delta_\psi)$

Una prueba secuencial (ψ, ϕ) satisface la condición $(\Pi_0^\infty, \Delta_\psi)$, si ψ cumple la *condición* Π_0^∞ y para esta regla de paro, la regla ϕ cumple la *condición* Δ_ψ .

El siguiente teorema establece condiciones necesarias y suficientes para que una prueba secuencial satisfaga la condición $(\Pi_0^\infty, \Delta_\psi)$.

Teorema 16.

Si una prueba secuencial (ψ^, ϕ^*) satisface la condición $(\Pi_0^\infty, \Delta_{\psi^*})$ entonces:*

$$\psi^* \in \mathcal{F}_0 \quad y \tag{3.70}$$

$$L_0(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_0} L_0(\psi, \phi). \tag{3.71}$$

Además, si existe alguna prueba secuencial $(\psi^, \phi^*) \in \mathfrak{S}_0$ y tal que cumple (3.71), entonces (ψ^*, ϕ^*) satisface la condición $(\Pi_0^\infty, \Delta_{\psi^*})$.*

Demostración del Teorema 16:

Suficiencia: Sea (ψ^*, ϕ^*) tal que cumple la condición $(\Pi_0^\infty, \Delta_{\psi^*})$, entonces ψ^* satisface la condición Π_0^∞ . De lo cual, debido a la Corolario 7 (Capítulo 2 o (2.107) con $\pi = 1$) aplicado a nuestro contexto \mathcal{M}_0 se sigue que $\psi^* \in \mathcal{F}_0$.

Por otro lado, debido que se cumple lo anterior (3.70), a (3.15), al Teorema 8 (Capítulo 2) y a (3.16), se sigue que $L_0(\psi^*, \phi^*) = \inf_{\psi \in \mathcal{F}_0} L_0(\psi, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_0} L_0(\psi, \phi)$.

Necesidad: Sea (ψ^*, ϕ^*) una prueba con $\psi^* \in \mathcal{F}_0$ y tal que satisface (3.71). Entonces por Teoremas 2, 6 (Capítulo 2) y (3.16) se sigue que ψ^* y ϕ^* cumplen las condiciones Π_0^∞ y Δ_{ψ^*} respectivamente. Por tanto, (ψ^*, ϕ^*) satisface la condición $(\Pi_0^\infty, \Delta_{\psi^*})$. ■

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Finalmente, en el siguiente lema se establece que es suficiente que una prueba secuencial (ψ^*, ϕ^*) cumpla la condición $(\Pi_0^\infty, \Delta_{\psi^*})$ para que ella minimice el costo promedio total K_0 del experimento. Con esto se resuelve nuestro problema de optimización restringido a las probabilidades de error.

Corolario 12.

Si una prueba secuencial (ψ^, ϕ^*) satisface la condición $(\Pi_0^\infty, \Delta_{\psi^*})$, entonces para cualquier otra prueba secuencial $(\psi, \phi) \in \mathfrak{S}_0$ tal que*

$$\alpha(\psi, \phi) \leq \alpha(\psi^*, \phi^*) \quad \text{y} \quad \beta(\psi, \phi) \leq \beta(\psi^*, \phi^*), \quad (3.72)$$

se tiene que

$$K_0(\psi) \geq K_0(\psi^*). \quad (3.73)$$

La desigualdad (3.73) es estricta, si al menos una de las desigualdades en (3.72) lo es.

Demostración del Corolario 12:

Dado (ψ^*, ϕ^*) satisface la condición $(\Pi_0^\infty, \Delta_{\psi^*})$, entonces por la primera parte del Teorema 16 se tiene que $L_0(\psi^*, \phi^*) = \inf_{(\psi, \phi) \in \mathfrak{S}_0} L(\psi, \phi)$. Ahora dada otra prueba $(\psi, \phi) \in \mathfrak{S}_0$ cualquiera que satisface las desigualdades (3.72), por Corolario 1* (Capítulo 2) se sigue (3.73). También por este mismo corolario se tiene que la desigualdad (3.73) es estricta, si al menos una de las desigualdades en (3.72) lo es. ■

En resumen, la estructura de las componentes de la prueba secuencial óptima (ψ^*, ϕ^*) es

$$\begin{aligned} I_{\{g(z_n) < \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_n)\}} &\leq \psi_n^* \leq I_{\{g(z_n) \leq \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_n)\}} \\ I_{\{\lambda_0/\lambda_1 < z_n\}} &\leq \phi_n^* \leq I_{\{\lambda_0/\lambda_1 \leq z_n\}}, \end{aligned} \quad (3.74)$$

para cada $r = 1, 2, \dots$ y $\mathbf{n} \in \mathcal{P}_r$. Esta prueba óptima es una generalización (aleatorizada) de la estructura encontrada en la investigación Popoca-Jiménez [Popoca, (2012)].

A continuación vamos a reescribir las estructuras óptimas (3.74) en términos de la región de continuación (o intervalos).

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

3.1.4. Región de continuación en términos de la razón de probabilidades

Como es costumbre en la estadística clásica y en las pruebas de hipótesis (aún en las secuenciales), es necesario determinar y establecer explícitamente la región de rechazo y/o continuación de las pruebas estadísticas secuenciales óptimas. En nuestro caso aquella cuya estructura está en las condiciones Δ_ψ y Π_0^∞ (o su expansión (3.74)). Para determinar dicha región se trabajará primero con la estructura de la regla de paro óptima ψ^* , lo cual nos determinará algunos intervalos, y finalmente con base en estos se establecerá la condición de la decisión óptima ϕ^* .

Primero se desarrollará un análisis de las propiedades de las funciones $g(z)$ y $J_r V_r(z)$ para encontrar los puntos z_n en donde se cumple la relación $g(z_n) \leq \bar{c}_{r+1} + J_{r+1} V_{r+1}(z_n)$ que está presente en la condición Π_0^∞ , que determina a las reglas de paro óptimas. Después se reescribirán las estructuras de la regla de paro y la regla de decisión óptimas con base en dicho análisis.

En toda esta sección y lo que resta del capítulo vamos a suponer que se cumple la siguiente condición.

$$P_0(f_1(X) > 0) = 1. \quad (3.75)$$

La hipótesis (3.76) del siguiente lema indica el crecimiento del costo total del experimento conforme el número de grupos a su vez crece ($k \rightarrow \infty$); implícitamente, supone que no pueden existir subsucesiones de tamaños de grupos con costos nulos.

Lema 21.

Supongamos que se cumple

$$c_1(\nu_1) + c_2(\nu_2) + \cdots + c_k(\nu_k) \longrightarrow \infty \quad \text{casi seguramente, cuando } k \rightarrow \infty. \quad (3.76)$$

Entonces

$$\bar{c}_1 + \bar{c}_2 + \cdots + \bar{c}_k \longrightarrow \infty, \quad \text{cuando } k \rightarrow \infty. \quad (3.77)$$

Demostración del Lema 21:

Sea la variable aleatoria $S_k := c_1(\nu_1) + c_2(\nu_2) + \cdots + c_k(\nu_k)$. Por hipótesis (3.76) se sabe que $S_k \xrightarrow[k \rightarrow \infty]{} \infty$ casi seguramente, de aquí se tiene que $P(S_k > c) \xrightarrow[k \rightarrow \infty]{} 1$ para todo $c > 0$. Sea $\bar{S}_k := ES_k = \bar{c}_1 + \bar{c}_2 + \cdots + \bar{c}_k$. Como $\bar{c}_i \geq 0$ para toda $i \geq 1$ entonces la sucesión $\{\bar{S}_k\}_k$ es

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

monótona creciente. Supongamos que $\lim_{k \rightarrow \infty} \bar{S}_k = L < \infty$.

Aplicando la desigualdad de Markov, sea $c > 0$ cualquiera entonces

$$P(S_k > c) \leq \frac{\bar{S}_k}{c} \leq \frac{L}{c}.$$

Sin embargo, para $c = 2L$, $P(S_k > c) < \frac{1}{2}$, para toda k .

Lo cual contradice que $P(S_k > c) \xrightarrow[k \rightarrow \infty]{} 1$ para toda $c > 0$. Por lo tanto $\bar{S}_k \xrightarrow[k \rightarrow \infty]{} \infty$. ■

Convención:

- En lo subsiguiente se trabajarán funciones medibles $v(z)$ que estarán definidas para todo $z \in [0, \infty)$ y también para $z = \infty$, esto último a través del límite correspondiente cuando $z \rightarrow \infty$, es decir, $v(\infty) := \lim_{z \rightarrow \infty} v(z)$. En general, esto no se explicitará en cada una de las definiciones, a menos que sea necesario.
- En este trabajo consideraremos una función $F(z)$ creciente si y solo si $F(z_1) \leq F(z_2)$, para todo $0 \leq z_1 < z_2$, esto es, función no decreciente.

Lema 22.

Para cualesquiera $N \geq r \geq 1$ enteros, las funciones $V_r^N(z)$ y $V_r(z)$ definidas en (3.19)-(3.20) y (3.53), respectivamente, todas poseen las siguientes propiedades:

- (a) $V_r^N(0) = V_r(0) = 0$;
- (b) *son cóncavas y continuas para todo $z \in [0, \infty)$;*
- (c) *son crecientes en el intervalo $[0, \infty)$; y*
- (d) *Si se cumple que (3.75) entonces $\lim_{z \rightarrow \infty} V_r^N(z) = \lim_{z \rightarrow \infty} V_r(z) = \lambda_0$.*

Demostración del Lema 22:

Sea N un número natural fijo.

- (a) El hecho de que $V_r^N(0) = V_r(0) = 0 \quad \forall N \geq r \geq 1$ es fácilmente demostrable por inducción inversa para $r = N, \dots, 1$ y luego aplicando el límite cuando $N \rightarrow \infty$. Por tanto $V_r^N(0) = V_r(0) = 0$.

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

(b) Demostremos, también, por inducción inversa que $V_r^N(z)$ y $V_r(z)$ son cóncavas.

Sea $r = N$.

$V_N^N(z) = g(z)$ es cóncava por ser el mínimo de dos funciones cóncavas (funciones lineales).

Hipótesis de Inducción: Supongamos que $V_r^N(z)$ es cóncava para cualquier $r = N, \dots, 2$. Es decir, supongamos que $\forall z_1, z_2 \in [0, \infty)$ y $\forall \alpha \in [0, 1]$ se tiene

$$\alpha V_r^N(z_1) + (1 - \alpha)V_r^N(z_2) \leq V_r^N(\alpha z_1 + (1 - \alpha)z_2). \quad (3.78)$$

Primero *demostraremos* que las funciones $J_r V_r^N(z)$ son cóncavas para cualquier $r = N, \dots, 2$. Para cada $m \in \mathcal{G}$ y x^m fija, por (3.78),

$$\alpha V_r^N \left(z_1 \frac{f_1^m(x^m)}{f_0^m(x^m)} \right) + (1 - \alpha)V_r^N \left(z_2 \frac{f_1^m(x^m)}{f_0^m(x^m)} \right) \leq V_r^N \left((\alpha z_1 + (1 - \alpha)z_2) \frac{f_1^m(x^m)}{f_0^m(x^m)} \right).$$

Esta desigualdad se cumple para cualquiera x^m . Ahora debido a que V_r^N son funciones no negativas y medibles μ^m -c.t.p. podemos aplicar la definición de esperanza y sus propiedades, entonces:

$$\alpha E_0 V_r^N \left(z_1 \frac{f_1^m(x^m)}{f_0^m(x^m)} \right) + (1 - \alpha)E_0 V_r^N \left(z_2 \frac{f_1^m(x^m)}{f_0^m(x^m)} \right) \leq E_0 V_r^N \left((\alpha z_1 + (1 - \alpha)z_2) \frac{f_1^m(x^m)}{f_0^m(x^m)} \right).$$

\therefore las funciones $E_0 V_r^N \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right)$ son cóncavas $\forall r = N, \dots, 2$. Además, se sabe (Rockafellar, 1972) que la suma de funciones cóncavas es una función cóncava, por tanto,

$$\begin{aligned} & \alpha J_r V_r^N(z_1) + (1 - \alpha)J_r V_r^N(z_2) = \\ &= \alpha \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^N \left(z_1 \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) + (1 - \alpha) \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^N \left(z_2 \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) \leq \\ &\leq \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^N \left((\alpha z_1 + (1 - \alpha)z_2) \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) = J_r V_r^N(\alpha z_1 + (1 - \alpha)z_2). \end{aligned}$$

Por lo tanto las funciones $J_r V_r^N(z)$ son cóncavas $\forall r = N, \dots, 2$.

Ahora, por definición (3.20) $V_{r-1}^N(z)$ es el mínimo entre funciones cóncavas. Por tanto $V_{r-1}^N(z)$ es también cóncava. De lo cual $V_r^N(z)$ es cóncava para toda $r = N, \dots, 1$.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

El límite puntual de una sucesión de funciones cóncavas es una función cóncava (Rockafellar, 1972), así que $V_r(z) = \lim_{N \rightarrow \infty} V_r^N(z)$ es también cóncava para todo $r \geq 1$ y por lo tanto $J_r V_r(z)$ también es cóncava.

Por tanto, para cualesquiera entero $N \geq r \geq 1$ las funciones $V_r^N(z)$, $V_r(z)$, $J_r V_r^N(z)$ y $J_r V_r(z)$ son cóncavas, y por Rockafellar (1972, p.82) también son continuas para todo $z \in (0, \infty)$.

De **(a)** se sabe que $V_r^N(0) = 0$ para toda $N \geq r \geq 1$, ahora para r fija se sabe que:

$$\begin{aligned} V_r^N(z) &\leq g(z) \\ \Rightarrow \lim_{z \rightarrow 0^+} \sup_{N \geq 1} V_r^N(z) &\leq \lim_{z \rightarrow 0^+} g(z) = 0 \end{aligned}$$

Por otro lado, $V_r^N(z) \geq 0$, así que $\lim_{z \rightarrow 0^+} V_r^N(z) = 0 = V_r^N(0)$.

Por tanto, $V_r^N(z)$ es continua en $z = 0$ para toda $N \geq r \geq 1$, del mismo modo $J_r V_r^N(z)$ es continua en $z = 0$.

Análogamente $V_r(z) \leq g(z) \rightarrow 0 = V_r(0)$, cuando $z \rightarrow 0^+$.

Por lo tanto las funciones $V_r^N(z)$, $J_r V_r^N(z)$, $V_r(z)$ y $J_r V_r(z)$ son continuas en $[0, \infty)$.

(c) Por inducción inversa se demostrará que $V_r^N(z)$ son crecientes para $z \in [0, \infty)$.

Sea $r = N$,

$$V_N^N(z) = \min \{ \lambda_0, \lambda_1 z \} \quad \text{es creciente.}$$

Hipótesis de Inducción: Supongamos que $V_r^N(z)$ es creciente para toda $r = N, \dots, 2$.

Para cualesquiera X^m fija, por hipótesis de inducción y debido a propiedades de la esperanza se tiene que $J_r V_r^N(z)$ es también creciente para toda $r = N, \dots, 2$.

De lo cual y por (3.20) $V_{r-1}^N(z)$ es el mínimo entre dos funciones crecientes. Por tanto, $V_{r-1}^N(z)$ es creciente para toda $r = N, \dots, 1$.

El límite puntual de funciones crecientes es también creciente, por tal $V_r(z)$ es creciente. Por lo tanto las funciones $V_r^N(z)$ y $V_r(z)$ son crecientes en el intervalo $[0, \infty)$ para cualquiera $N \geq r \geq 1$.

(d) Supongamos que se cumple (3.75). Por las propiedades anteriores y utilizando, nuevamente, inducción demostraremos que: $\lim_{z \rightarrow \infty} V_r^N(z) = \lambda_0, \quad \forall N \geq r \geq 1$.

Sea $r = N$,

$$\lim_{z \rightarrow \infty} V_N^N(z) = \lim_{z \rightarrow \infty} g(z) = \lim_{z \rightarrow \infty} (\min \{ \lambda_0, \lambda_1 z \}) = \lambda_0.$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Hipótesis de inducción: Como $N \in \mathbb{N}$ es fijo, supongamos que $\lim_{z \rightarrow \infty} v_r^N(z) = \lambda_0$ para alguna $r = N, \dots, 2$. Para cada $m \in \mathcal{G}$ y X^m fijas se tiene:

$$\begin{aligned}
 \lim_{z \rightarrow \infty} J_r V_r^N(z) &= \lim_{z \rightarrow \infty} \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r^N \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right), \\
 &= \sum_{m \in \mathcal{G}} p_r(m) \lim_{z \rightarrow \infty} E_0 V_r^N \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right), \quad \text{ya que } \sum_{m \in \mathcal{G}} p_r(m) \leq 1 \text{ y TCD,} \\
 &= \sum_{m \in \mathcal{G}} p_r(m) E_0 \lim_{z \rightarrow \infty} V_r^N \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right), \quad \text{debido a (3.75) y TCD,} \\
 &= \lambda_0, \quad \text{por hipótesis de inducción.}
 \end{aligned}$$

Ahora para $r - 1$ y tomando el límite a la ecuación (3.20), cuando $z \rightarrow \infty$ se tiene:

$$\lim_{z \rightarrow \infty} V_{r-1}^N(z) = \lim_{z \rightarrow \infty} \left[\min \{g(z), \bar{c}_r + J_r V_r^N(z)\} \right] = \min \{\lambda_0, \bar{c}_r + \lambda_0\} = \lambda_0.$$

Por tanto, $\forall r = N, \dots, 1$, $\lim_{z \rightarrow \infty} V_r^N(z) = \lambda_0$.

Debido a la monotonía de V_r^N se tiene que $\lim_{z \rightarrow \infty} V_r(z) = a_r$ existe. Ahora aplicando (3.75) y el Teorema de convergencia dominada se sigue que $\lim_{z \rightarrow \infty} J_r V_r(z) = a_r$. Pasando al límite en la ecuación (3.54), cuando $z \rightarrow \infty$ se tiene:

$$\begin{aligned}
 \lim_{z \rightarrow \infty} V_r(z) &= \lim_{z \rightarrow \infty} \min \{g(z), \bar{c}_{r+1} + J_{r+1} V_{r+1}(z)\}. \\
 \Rightarrow a_r &= \min \{\lambda_0, \bar{c}_{r+1} + a_{r+1}\} \leq \lambda_0, \quad \text{para } r \geq 1. \quad (3.79)
 \end{aligned}$$

Si $a_r = \lambda_0$, para alguna $r \geq 1$, entonces $a_{r-1} = \min \{\lambda_0, \bar{c}_{r-1} + \lambda_0\} = \lambda_0$, y obviamente $a_i = \lambda_0$ para cada $i \leq r$.

Supongamos que existe un número finito k tal que para todo $i \geq k$ se tiene que $a_i = \bar{c}_{i+1} + a_{i+1} < \lambda_0$. Esto conlleva a una contradicción porque en el caso de $m > k$ se tendría que:

$$a_k = \bar{c}_{k+1} + a_{k+1} = \bar{c}_{k+1} + \bar{c}_{k+2} + a_{k+2} = \dots = \sum_{j=k+1}^m \bar{c}_j + a_m < \lambda_0.$$

Sin embargo, lo anterior contradice (3.77). Por lo tanto $a_r = \lambda_0 \forall r \geq 1$ i.e.

$$\lim_{z \rightarrow \infty} V_r(z) = \lambda_0, \quad \text{para toda } r \geq 1. \quad \blacksquare$$

Ahora, para alguna $r \geq 1$, el operador $J_r V_r(z)$ está definido para toda $z \in [0, \infty)$ (véase (3.18)). Debido a la definición de dicho operador y la relación entre éste y las funciones V_r^N

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

y V_r , entonces con base en el Lema 22 dicho operador tiene las mismas características de las funciones $V_r(z)$, a saber:

- (1) $J_r V_r(0) = 0$;
- (2) $J_r V_r(z)$ es cóncava y continua para cada $z \in [0, \infty)$,
- (3) $J_r V_r(z)$ es creciente en $[0, \infty)$, y
- (4) Si que se cumple (3.75), entonces $\lim_{z \rightarrow \infty} J_r V_r(z) = \lambda_0$.

(Véase la figura (3.1), gráfica del operador $J_r V_r$). Las demostraciones de las propiedades del operador $J_r V_r$ están internamente en la demostración del Lema 22.

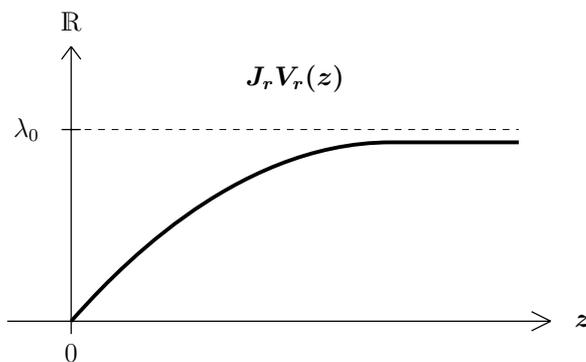


Figura 3.1: Ejemplo de una gráfica de $J_r V_r(z)$

Lema 23.

Sea $r \geq 1$, si $z \in [0, \infty)$ entonces $J_r V_r(z) \leq g(z)$.

Demostración del Lema 23:

Aplicando la desigualdad de Jensen: $E\rho(X) \leq \rho(EX)$ con ρ función cóncava. Sean $m \in \mathcal{G}$ cualquiera, $z \in [0, \infty)$ y X^m fijas, entonces:

$$\begin{aligned} E_0 V_r \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) &\leq V_r \left(z E_0 \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) \\ &= V_r \left(z \int_{\mathcal{K}_0^m} f_1^m(x^m) d\mu^m(x^m) \right) \leq V_r(z). \end{aligned}$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Lo anterior se tiene para cualquier $z \in [0, \infty)$ y X^m fijas entonces:

$$J_r V_r(z) = \sum_{m \in \mathcal{G}} p_r(m) E_0 V_r \left(z \frac{f_1^m(X^m)}{f_0^m(X^m)} \right) \leq \sum_{m \in \mathcal{G}} p_r(m) V_r(z) = V_r(z) \leq g(z).$$

Por lo tanto, $J_r V_r(z) \leq g(z)$ para toda $z \in [0, \infty)$ y cualquier $r \geq 1$ fijo. ■

Solo para puntualizar que, bajo la condición (3.75), es posible considerar a cualesquiera de las funciones anteriores en el punto $z = \infty$ como el valor límite, cuando $z \rightarrow \infty$, en cuyo caso será λ_0 , *i.e.*

$$g(\infty) := \lim_{z \rightarrow \infty} g(z) = \lambda_0 \quad \text{y} \quad \lim_{z \rightarrow \infty} V_r(z) = \lim_{z \rightarrow \infty} J_r V_r(z) = \lambda_0. \quad (3.80)$$

Con base en estas definiciones, si nuevamente consideramos que para alguna etapa r con su grupo de observaciones x^n se tiene que $f_0^n(x^n) = 0$ y $f_1^n(x^n) > 0$ entonces, por definición (3.11), $z_n = \infty$; y por (3.80) se tiene $g(\infty) = J_{r+1} V_{r+1}(\infty) = \lambda_0$, entonces aplicando esto a la regla de paro óptima de (3.74) se tendría que $I_{\{\lambda_0 < \bar{c}_{r+1} + \lambda_0\}} \leq \psi_n^* \leq I_{\{\lambda_0 \leq \bar{c}_{r+1} + \lambda_0\}}$.

Si el costo promedio del siguiente grupo es $\bar{c}_{r+1} > 0$ entonces $\psi_n^* = 1$, lo mejor es parar el experimento. Pero si $\bar{c}_{r+1} = 0$ entonces se tendría una aleatorización para la regla de paro $0 \leq \psi_n^* \leq 1$. Este ejemplo muestra que la definición de la razón de verosimilitud dada en (3.11), es consistente en esta investigación, además de que el tener costos promedios nulos (de grupos), no es muy útil, y en general, en la práctica es extraño que haya costos nulos.

Definición 30.

Para cualesquiera $\lambda_0 > 0$ y $\lambda_1 > 0$,

$$\lambda := \frac{\lambda_0}{\lambda_1}. \quad (3.81)$$

El siguiente lema demuestra que si una función es convexa, entonces ella es creciente, y en el momento en que la función toma un valor positivo, entonces será estrictamente creciente.

Lema 24.

Sea $F(z)$ una función convexa no negativa sobre $\{z \geq 0\}$ tal que $F(0) = 0$. Entonces

$$F(z_1) \leq F(z_2) \quad (3.82)$$

para todo $0 \leq z_1 < z_2$. La desigualdad en (3.82) es estricta cuando $F(z_1) > 0$.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Demostración del Lema 24:

Sean cualesquiera $0 \leq z_1 < z_2$, entonces $0 \leq \frac{z_1}{z_2} < 1$. Sea

$$z_1 = \frac{z_1}{z_2} z_2 + \left(1 - \frac{z_1}{z_2}\right) 0.$$

Dado que $F(z)$ es convexa,

$$0 \leq F(z_1) = F\left(\frac{z_1}{z_2} z_2 + \left(1 - \frac{z_1}{z_2}\right) 0\right) \leq \frac{z_1}{z_2} F(z_2) + \left(1 - \frac{z_1}{z_2}\right) F(0) \leq F(z_2).$$

$\therefore 0 \leq F(z_1) \leq F(z_2)$. Si $F(z_1) > 0$, entonces la desigualdad (3.82) será estricta. ■

Definición 31.

Sean $r \geq 1$ y $z \in [0, \infty)$ cualesquiera se define la función:

$$D_r(z) := g(z) - J_r V_r(z). \tag{3.83}$$

Con base en que $g(z) = \min\{\lambda_0, \lambda_1 z\}$ se puede determinar que :

$$D_r(z) = \begin{cases} \lambda_1 z - J_r V_r(z), & \text{si } z < \lambda; \\ \lambda_0 - J_r V_r(z), & \text{si } z \geq \lambda. \end{cases}$$

Por un lado $g(z)$ es un caso particular de las funciones $V_r(z)$, entonces tiene las mismas propiedades de ellas. Por otro lado, con base en las propiedades de las funciones g y $J_r V_r$ con $r \geq 1$ fijo, la función D_r hereda las siguientes propiedades (véase Figura 3.2):

- (1) $D_r(z)$ está definida para todo $z \in [0, \infty)$.
- (2) Por Lema 23, $D_r(z) \geq 0 \quad \forall z \in [0, \infty)$.
- (3) $D_r(0) = 0$. Además si que se cumple (3.75) entonces $\lim_{z \rightarrow \infty} D_r(z) = 0$.
- (4) $D_r(z)$ es convexa en $[0, \infty)$ y por lo tanto continua para toda z (Rockafellar, 1972, p. 86).
Porque es la suma de dos funciones convexas, una lineal $g(z)$ y otra convexa $-J_{r+1} V_{r+1}(z)$.
- (5) Si $z < \lambda$ entonces $D_r(z)$ es creciente.
Sean $0 \leq z_1 < z_2 < \lambda$. Dado que $D_r(z)$ es convexa y $D_r(0) = 0$, entonces se sigue del Lema 24 que $D_r(z_1) \leq D_r(z_2)$. Por tanto $D_r(z)$ es creciente para todo $z < \lambda$.

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

(6) Si $z \geq \lambda$ entonces $D_r(z)$ es decreciente.

Esto es porque dados $\lambda \leq z_1 < z_2$ y como $J_r V_r(z)$ es creciente, entonces:

$$J_r V_r(z_1) \leq J_r V_r(z_2) \Rightarrow \lambda_0 - J_r V_r(z_1) \geq \lambda_0 - J_r V_r(z_2) \Rightarrow D_r(z_1) \geq D_r(z_2),$$

$$\therefore z_1 < z_2 \Rightarrow D_r(z_2) \leq D_r(z_1).$$

(7) Finalmente, con base en (1), (6) y (5), $D_r(z)$ tiene un valor máximo en el punto $z = \lambda$, *i. e.* $D_r(\lambda)$ es el valor máximo. De lo cual, la imagen de $D_r(z)$ es $[0, D_r(\lambda)]$.

Ahora para un $r \geq 1$ fijo vamos a comparar a la función $D_r(z)$ con el costo promedio del $r + 1$ -ésimo grupo (\bar{c}_{r+1}) (3.3). Al comparar ambas funciones, se tienen tres casos:

$$(1) \bar{c}_{r+1} < D_{r+1}(\lambda), \quad (2) \bar{c}_{r+1} = D_{r+1}(\lambda) \quad \text{o} \quad (3) \bar{c}_{r+1} > D_{r+1}(\lambda). \quad (3.84)$$

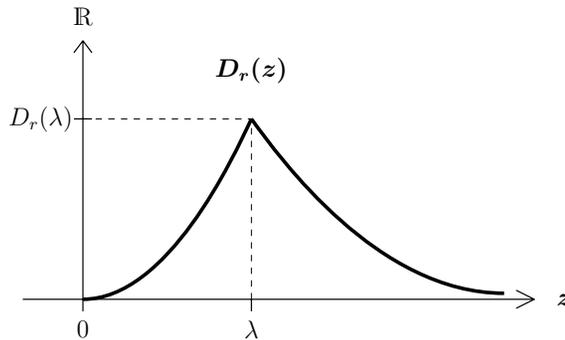


Figura 3.2: Ejemplo de una gráfica de la función $D_r(z)$

Los casos (2) y (3) de (3.84), no serán de interés en esta investigación porque en ambos casos no existirá una región de continuación. En el caso $\bar{c}_{r+1} = D_{r+1}(\lambda)$ solo existe un punto de intersección entre ambas funciones (véase Figura 3.3) y por tanto aquí, debido a la estructura de la regla de paro, se tomará la regla aleatorizada $0 \leq \psi_n \leq 1$. Sin embargo, para el caso $\bar{c}_{r+1} > D_{r+1}(\lambda)$, no existen puntos de intersección entre las funciones porque $D_{r+1}(\lambda)$ es el valor máximo de la función y por tanto no existirá región de continuación y este caso no es de interés. Por tanto, solo nos enfocaremos en el análisis de $\bar{c}_{r+1} < D_{r+1}(\lambda)$.

Por otro lado, también es importante aclarar el caso cuando el costo promedio es nulo del grupo $r + 1$ -ésimo. En general, si $r \geq 1$ y $\bar{c}_{r+1} = 0$, la regla de paro indicará una aleatorización, y en este caso conviene seguir tomando observaciones y pasar a la siguiente etapa, ya que en dicha etapa no hay costo, y al tomar más observaciones esto nos ayudará a su vez a que se reduzcan las probabilidades de error. En la práctica es difícil encontrarnos con situaciones en las cuales no haya un costo por el grupo de observaciones, a menos que en dicho grupo no

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

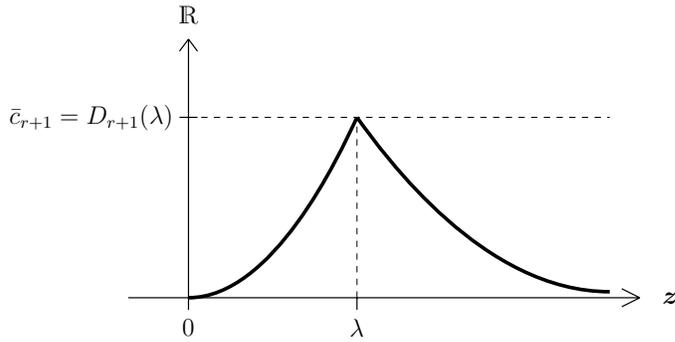


Figura 3.3: $\bar{c}_{r+1} = D_{r+1}(\lambda)$

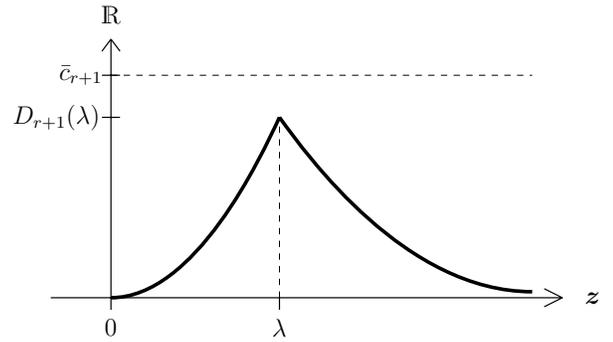


Figura 3.4: $\bar{c}_{r+1} > D_{r+1}(\lambda)$

haya observaciones y esto no genere un gasto. De hecho, en el Lema 21 hipótesis (3.76) estamos suponiendo que no se pueden dar subsucesiones de tamaños de grupos nulos. Por lo cual, no será desarrollado teóricamente el caso $\bar{c}_{r+1} = 0$. Por lo tanto, solo nos centraremos en el caso cuando $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$ (véase figura 3.5).

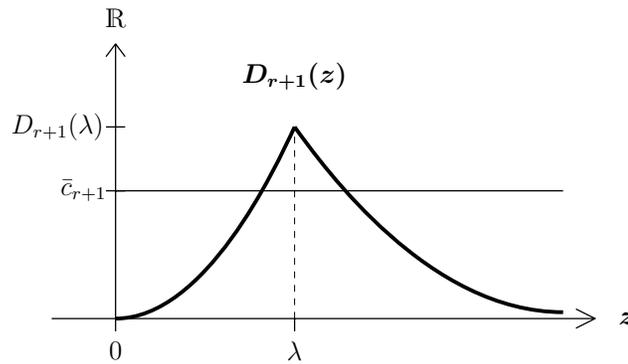


Figura 3.5: $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Lema 25.

Sea $r \geq 1$ fijo. Si se cumple que $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$ entonces existen únicos A_r y B_r tales que:

$$D_{r+1}(A_r) = \bar{c}_{r+1} \quad \text{con} \quad A_r \in (0, \lambda) \quad \text{y} \quad (3.85)$$

$$D_{r+1}(B_r) = \bar{c}_{r+1} \quad \text{con} \quad B_r \in (\lambda, \infty). \quad (3.86)$$

Demostración del Lema 25:

Se sabe que D_{r+1} es creciente en el intervalo $[0, \lambda)$, $D_{r+1}(\lambda)$ es el valor máximo y $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$. Por continuidad de D_{r+1} , existe un punto $A_r \in [0, \lambda)$ tal que $D_{r+1}(A_r) = \bar{c}_{r+1}$. Sin embargo, $A_r \neq 0$ ya que si $A_r = 0$ entonces $D_{r+1}(0) = 0 = \bar{c}_{r+1}$, contradicción porque $0 < \bar{c}_{r+1}$. Por tanto $A_r \in (0, \lambda)$.

Por otro lado, D_{r+1} es decreciente en el intervalo $[\lambda, \infty)$ y $\bar{c}_{r+1} < D_{r+1}(\lambda)$. También, dado que se cumple (3.75), se tiene que $\lim_{z \rightarrow \infty} D_{r+1}(z) = 0$, entonces existe $z \in [\lambda, \infty)$ tal que $D_{r+1}(z) < \bar{c}_{r+1}$. De aquí se tiene:

$$D_{r+1}(z) < \bar{c}_{r+1} < D_{r+1}(\lambda).$$

Por la continuidad de D_{r+1} existe $B_r \in [\lambda, \infty)$ tal que $D_{r+1}(B_r) = \bar{c}_{r+1}$. Dado que $\bar{c}_{r+1} < D_{r+1}(\lambda)$, entonces $B_r \neq \lambda$, por tanto $\lambda < B_r < z < \infty$.

De lo anterior se concluye que existen $0 < A_r < \lambda < B_r < \infty$ tales que $D_{r+1}(A_r) = D_{r+1}(B_r) = \bar{c}_{r+1}$.

Ahora demostrará la unicidad de A_r y B_r respectivamente en su intervalo.

Supongamos que existe otro $A'_r \in (0, \lambda)$ con $D_{r+1}(A'_r) = \bar{c}_{r+1}$ y tal que $A'_r \neq A_r$.

Sin pérdida de generalidad (SPG) sea $A'_r < A_r$. Sabemos que $A'_r \neq 0$ y $A_r \neq 0$ entonces $0 < \frac{A'_r}{A_r} < 1$. Sea

$$A'_r = \left(\frac{A'_r}{A_r}\right)A_r + \left(1 - \frac{A'_r}{A_r}\right)0.$$

Sabemos que $D_{r+1}(z)$ es convexa en el intervalo $[0, \infty)$, entonces

$$D_{r+1}(A'_r) = D_{r+1}\left(\left(\frac{A'_r}{A_r}\right)A_r + \left(1 - \frac{A'_r}{A_r}\right)0\right) \leq \left(\frac{A'_r}{A_r}\right)D_{r+1}(A_r) + \left(1 - \frac{A'_r}{A_r}\right)D_{r+1}(0) \leq \left(\frac{A'_r}{A_r}\right)\bar{c}_{r+1}.$$

Por tanto $\bar{c}_{r+1} \leq \frac{A'_r}{A_r} \bar{c}_{r+1}$, además $\bar{c}_r \neq 0$. Lo cual es una contradicción porque $0 < \frac{A'_r}{A_r} < 1$.

Entonces $A'_r = A_r$. Análogamente se procede, si $A_r < A'_r$. Por lo tanto A_r es único.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Ahora supongamos que existe otro punto $B'_r \in (\lambda, \infty)$ tal que $D_{r+1}(B'_r) = \bar{c}_{r+1}$ y $B'_r \neq B_r$. SPG supongamos que $B'_r < B_r$ (análogamente se procede, si $B_r < B'_r$). Se sabe que $\exists z > B_r$ tal que $D_{r+1}(z) < \bar{c}_{r+1}$ entonces se tiene $B'_r < B_r < z$ con base en esto sean:

$$\alpha = \frac{z - B_r}{z - B'_r} \Rightarrow (1 - \alpha) = \frac{B_r - B'_r}{z - B'_r}$$

donde $0 < \alpha \leq 1$ entonces se puede escribir:

$$B_r = \frac{z - B_r}{z - B'_r} B'_r + \frac{B_r - B'_r}{z - B'_r} z \Rightarrow D_{r+1}(B_r) = D_{r+1} \left(\left(\frac{z - B_r}{z - B'_r} \right) B'_r + \left(\frac{B_r - B'_r}{z - B'_r} \right) z \right)$$

Como D_{r+1} es convexa en el intervalo $[0, \infty)$, entonces

$$\begin{aligned} D_{r+1}(B_r) &\leq \left(\frac{z - B_r}{z - B'_r} \right) D_{r+1}(B'_r) + \left(\frac{B_r - B'_r}{z - B'_r} \right) D_{r+1}(z) \\ \Rightarrow \bar{c}_{r+1} &\leq \left(\frac{z - B_r}{z - B'_r} \right) \bar{c}_{r+1} + \left(\frac{B_r - B'_r}{z - B'_r} \right) D_{r+1}(z) \\ \Rightarrow \left(\frac{B_r - B'_r}{z - B'_r} \right) \bar{c}_{r+1} &\leq \left(\frac{B_r - B'_r}{z - B'_r} \right) D_{r+1}(z) \Rightarrow \bar{c}_{r+1} \leq D_{r+1}(z). \end{aligned}$$

Lo cual es una contradicción porque $D_{r+1}(z) < \bar{c}_{r+1}$.

Entonces $B'_r = B_r$ y por lo tanto el punto B_r es único en (λ, ∞) . ■

En el Lema 25 se demuestra que para $r \geq 1$ tal que $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$ (*i.e.* cuando los costos promedios están acotados), existen $A_r \in (0, \lambda)$ y $B_r \in (\lambda, \infty)$ contantes únicas las cuales satisfacen las siguientes ecuaciones:

$$g(A_r) = \bar{c}_{r+1} + J_{r+1} V_{r+1}(A_r). \quad (3.87)$$

$$g(B_r) = \bar{c}_{r+1} + J_{r+1} V_{r+1}(B_r). \quad (3.88)$$

Véase Figura 3.6.

Lema 26.

Sea $r \geq 1$ fijo. Si A_r y B_r cumplen las ecuaciones (3.87)-(3.88) respectivamente,

$$A_r < z < B_r \text{ si y solamente si } g(z) > \bar{c}_{r+1} + J_{r+1} V_{r+1}(z). \quad (3.89)$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Demostración del Lema 26: Supóngase que A_r y B_r cumplen las ecuaciones (3.87)-(3.88), respectivamente.

Necesidad: Si $A_r < z \leq \lambda$ entonces $D_{r+1}(A_r) \leq D_{r+1}(z)$ por ser D_{r+1} creciente, pero como A_r es único entonces $D_{r+1}(A_r) < D_{r+1}(z)$.

De aquí se sigue que $\bar{c}_{r+1} < g(z) - J_{r+1}V_{r+1}(z)$ y por tanto $\bar{c}_{r+1} + J_{r+1}V_r(z) < g(z)$.

Análogamente, si $\lambda \leq z < B_r$, como D_{r+1} es decreciente, entonces $D_{r+1}(B_{r+1}) \leq D_{r+1}(z)$, sin embargo, por ser B_r único entonces $D_{r+1}(B_{r+1}) < D_{r+1}(z)$. De lo cual $\bar{c}_{r+1} + J_{r+1}V_{r+1}(z) < g(z)$.

Por tanto:

$$\text{si } z \in (A_r, B_r) \text{ entonces } \bar{c}_{r+1} + J_{r+1}V_{r+1}(z) < g(z).$$

Suficiencia: Por otro lado si $z \in [0, A_r]$ entonces sea $z = (1-a)A_r$ para algún $0 \leq a \leq 1$. Se sabe que $J_{r+1}V_{r+1}(0) = 0$, entonces tomemos el segmento que va de 0 a z , como $J_{r+1}V_{r+1}$ es cóncava:

$$J_{r+1}V_{r+1}(z) = J_{r+1}V_{r+1}\left((1-a)A_r\right) \geq (1-a)J_{r+1}V_{r+1}(A_r) = (1-a)\left(g(A_r) - \bar{c}_{r+1}\right)$$

De lo cual

$$\begin{aligned} \bar{c}_{r+1} + J_{r+1}V_{r+1}(z) &\geq \bar{c}_{r+1} + (1-a)\left(g(A_r) - \bar{c}_{r+1}\right) = a\bar{c}_{r+1} + (1-a)g(A_r) \\ &\geq (1-a)g(A_r) = g\left((1-a)A_r\right) = g(z) \end{aligned}$$

$$\therefore \text{ si } z \in [0, A_r] \text{ entonces } \bar{c}_{r+1} + J_{r+1}V_{r+1}(z) \geq g(z).$$

Sea ahora $z \in [B_r, \infty)$, entonces $B \leq z$. Como $J_{r+1}V_{r+1}$ es creciente y cóncava, entonces

$$\begin{aligned} J_{r+1}V_{r+1}(B_r) \leq J_{r+1}V_{r+1}(z) &\Rightarrow 0 \leq J_{r+1}V_{r+1}(z) - J_{r+1}V_{r+1}(B_r) \\ &\Rightarrow \lambda_0 \leq J_{r+1}V_{r+1}(z) - J_{r+1}V_{r+1}(B_r) + \lambda_0. \end{aligned}$$

Por propiedades se sabe que $g(B_r) = \lambda_0 = g(z)$, ya que $\lambda < B_r \leq z$. Entonces:

$$\underbrace{\lambda_0}_{g(z)} \leq J_{r+1}V_{r+1}(z) - J_{r+1}V_{r+1}(B_r) + \underbrace{\lambda_0}_{g(B_r)} \Rightarrow g(z) \leq J_{r+1}V_{r+1}(z) + \bar{c}_{r+1}.$$

$$\therefore \text{ si } z \in [B_r, \infty) \text{ entonces } \bar{c}_{r+1} + J_{r+1}V_{r+1}(z) \geq g(z).$$

Por lo tanto, si $z \notin (A_r, B_r)$ entonces $g(z) \not< \bar{c}_{r+1} + J_{r+1}V_{r+1}(z)$. ■

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

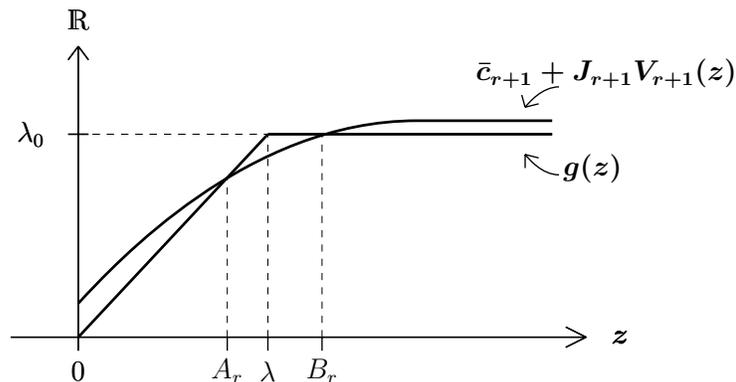


Figura 3.6: Ejemplo de región de continuación: intervalo (A_r, B_r)

Con base en el Lema 26, tenemos su equivalente: si $r \geq 1$ tal que $0 < \bar{c}_{r+1} < D_{r+1}(\lambda)$,

$$z \notin (A_r, B_r) \iff g(z) \leq \bar{c}_{r+1} + J_{r+1}V_{r+1}(z). \quad (3.90)$$

Donde A_r, B_r son constantes tales que $0 < A_r < \lambda < B_r < \infty$ y cumplen las ecuaciones (3.87)-(3.88), respectivamente, véase Figura 3.6.

Corolario 13.

Sea $r \geq 1$ cualquiera. Si A_r, B_r son constantes tales que $0 < A_r < \lambda < B_r < \infty$ y cumplen las ecuaciones (3.87)-(3.88), respectivamente, entonces la prueba secuencial óptima (ψ^*, ϕ^*) del Corolario 12 se reescribe como:

$$I_{\{z_n \in (A_r, B_r)\}} \leq 1 - \psi_n^* \leq I_{\{z_n \in [A_r, B_r]\}} \quad y \quad (3.91)$$

$$I_{\{B_r < z_n\}} \leq \phi_n^* \leq I_{\{B_r \leq z_n\}}. \quad (3.92)$$

Para cualquier $\mathbf{n} \in \mathcal{P}_r$. Donde z_n es la razón de verosimilitud definida en (3.11) y λ está en (3.81).

Demostración del Corolario 13:

Sean $r \geq 1, \mathbf{n} \in \mathcal{G}^r$. De (3.74) se tiene la estructura de la regla de paro óptima

$$I_{\{g(z_n) < \bar{c}_{r+1} + J_{r+1}V_{r+1}(z_n)\}} \leq \psi_n^* \leq I_{\{g(z_n) \leq \bar{c}_{r+1} + J_{r+1}V_{r+1}(z_n)\}}.$$

3.1. Prueba secuencial aleatoria óptima en términos de la razón de probabilidades

Con base en Lema 26 y (3.90) obtiene que:

$$I_{\{z_n \notin [A_r, B_r]\}} \leq \psi_n^* \leq I_{\{z_n \notin (A_r, B_r)\}} \implies I_{\{z_n \in (A_r, B_r)\}} \leq 1 - \psi_n^* \leq I_{\{z_n \in [A_r, B_r]\}}.$$

Ahora, supongamos que $\psi_n^* \equiv 1$ para alguna etapa r fija con $\mathbf{n} \in \mathcal{G}^r$, esto es ya hubo paro en la etapa r , entonces $z_n \notin [A_r, B_r]$. De lo cual se sigue que $z_n < A_r$ o $B_r < z_n$. Además, se sabe que $A_{r+1} < \lambda < B_{r+1}$.

Si $z_n < A_r$, entonces $z_n < \lambda$, por tanto $I_{\{\lambda < z_n\}} = 0 = I_{\{\lambda \leq z_n\}}$ y de lo cual $\phi_n^* = 0$, esto significa que la probabilidad de rechazar H_0 es nula, i.e. en este caso se acepta H_0 . Por tanto, lo anterior también se sigue a través de $I_{\{B_r < z_n\}} \leq \phi_n^* \leq I_{\{B_r \leq z_n\}}$.

Ahora, si $B_r < z_n$, entonces $\lambda < z_n$:

$$I_{\{\lambda < z_n\}} \leq \phi_n^* \leq I_{\{\lambda \leq z_n\}} \implies I_{\{B_r < z_n\}} \leq \phi_n^* \leq I_{\{B_r \leq z_n\}}.$$

de lo cual $\phi_n^* \equiv 1$ (se rechaza H_0 , a favor de H_1).

Finalmente, en los casos “límite” donde $z_n = A_r$ o $z_n = B_r$ se tendrá una aleatorización para la regla de paro $0 \leq \psi_n^* \leq 1$; sin embargo, la regla de decisión en el primero será $\phi_n^* \equiv 0$ (no se rechaza H_0), mas en el segundo caso límite también habrá aleatorización para ϕ_n^* . ■

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

En esta sección se van a aplicar los resultados de la sección anterior al modelo en el cual tanto las observaciones como los tamaños de los grupos son v.a independientes e idénticamente distribuidas. Con base en esto vamos a caracterizar la estructura de las pruebas secuenciales óptimas por grupos en el caso del horizonte infinito. En particular, vamos a probar la optimalidad de la RSPRT (Random Sequential Probability Ratio Test) propuesta por Mukhopadhyay y De Silva (2008) para grupos de tamaño aleatorio. Este resultado se encuentra presentado en el artículo Novikov & Popoca-Jiménez (2022) [Novikov-Popoca, (2022)].

Recordemos que estamos en el contexto de un problema de pruebas de hipótesis simples $\mathcal{M}_0 = (P_0, P_1, P_0)$.

Dado que tenemos tamaños de grupos idénticamente distribuidos. Sea la función de probabilidad común de todos los tamaños de grupos $P(\nu_r = n) = p_r(n) = p(n)$ para cualquier $n \in \mathcal{G}$ y $r = 1, 2, \dots$. Entonces de la sección anterior se modificarán aquéllos elementos que tenían el subíndice r referente a la función de probabilidad de los tamaños de grupos; por ejemplo el costo promedio del r -ésimo grupo será $\bar{c}_r = \sum_{m \in \mathcal{G}} c(m)p(m) = c$, esto es este costo será el mismo valor para cualquier grupo $r \geq 1$.

Existen tres constantes que están involucradas en la construcción de las pruebas óptimas de la sección anterior las cuales son: c , λ_0 y λ_1 . Es fácil ver que solo dos de ellas son suficientes para obtener todas las pruebas secuenciales óptimas del Teorema 3.

Sean $c > 0$, $\lambda = \lambda_0$ y supongamos que $\lambda_1 = 1$. La estructura de la prueba secuencial óptima del Corolario 13 ahora adquirirá una forma más simple. Nuevamente reescribiremos nuestros términos anteriores con los cambios correspondientes.

La sucesión de funciones $\{V_r(z)\}_{r \geq 1}$ definidas en (3.19)-(3.20), ahora serán denotadas mediante las funciones $\rho_k(z)$ con $k = 0, 1, \dots$, y a su vez el operador $J_r V_r(z)$ (3.18) será nombrado con $\bar{\rho}_k(z)$. La idea de esta nueva notación ahora es que el índice de las nuevas funciones irá creciendo, en lugar de decrecer.

Para cualquier $z \geq 0$, sea

$$\rho_0(z) = \rho_0(z; c, \lambda) = g(z; \lambda) = \min \{ \lambda, z \}; \quad (3.93)$$

y recursivamente para $k = 1, 2, \dots$ y cualquier $z \geq 0$ sea

$$\rho_k(z) = \rho_k(z; c, \lambda) = \min \left\{ g(z; \lambda), c + \sum_{n \in \mathcal{G}} p(n) E_0 \rho_{k+1}(zz_n; c, \lambda) \right\}. \quad (3.94)$$

También se define

$$\bar{\rho}_k(z) = \bar{\rho}_k(z; c, \lambda) = \sum_{n \in \mathcal{G}} p(n) E_0 \rho_k(zz_n; c, \lambda). \quad (3.95)$$

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

Se sigue de (3.19) y (3.20) que

$$V_k^N(z) = \rho_{N-k}(z; c, \lambda), \quad z \geq 0, \quad (3.96)$$

y de (3.18),

$$J_k V_k^N(z) = \bar{\rho}_{N-k}(z; c, \lambda). \quad (3.97)$$

Se define

$$\rho(z) = \rho(z; c, \lambda) = \lim_{k \rightarrow \infty} \rho_r(z; c, \lambda), \quad z \geq 0. \quad (3.98)$$

Si tomamos el límite, cuando $N \rightarrow \infty$, en (3.96), entonces

$$V_k(z) = \rho(z; c, \lambda), \quad z \geq 0, \quad (3.99)$$

para toda $k = 1, 2, \dots$, y debido a (3.97)

$$J_k V_k(z) = \bar{\rho}(z; c, \lambda), \quad z \geq 0. \quad (3.100)$$

Así, la regla de paro (3.58) en la condición Π_0^∞ ahora se transforma en

$$\psi_{\mathbf{n}} \stackrel{\geq}{\leq} I_{\{g(z_{\mathbf{n}}; \lambda) \lesseqgtr c + \bar{\rho}(z_{\mathbf{n}}; c, \lambda)\}}, \quad (3.101)$$

así la forma de las reglas de paro óptimas dependen únicamente de si la desigualdad

$$g(z; \lambda) \leq c + \bar{\rho}(z; c, \lambda), \quad (3.102)$$

se cumple o no para $z = z_{\mathbf{n}}$, λ y c .

Antes que todo, es fácil ver que si

$$\lambda < c + \bar{\rho}(\lambda; c, \lambda), \quad (3.103)$$

entonces (3.101) implica que $\psi_{\mathbf{n}} \equiv 1$ se detiene con $\mathbf{n} \in \mathcal{G}$ (*i.e.* la prueba óptima se detiene después del primer grupo de observaciones). Por lo tanto, las pruebas secuenciales óptimas no triviales solo se obtendrán si

$$\lambda > c + \bar{\rho}(\lambda; c, \lambda). \quad (3.104)$$

Entonces ahora vamos a estudiar la condición (3.104). Para esto vamos a trabajar primero con la siguiente función.

Definición 32.

Para cualesquiera $z \geq 0$, $c \geq 0$ y $\lambda \geq 0$ se define la función

$$R(z; c, \lambda) := \inf_{(\psi, \phi) \in \mathfrak{S}_0} (cE_0\tau^\psi + \lambda\alpha(\psi, \phi) + z\beta(\psi, \phi)). \quad (3.105)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Se sigue del Teorema 15 y del Teorema 11 que, para toda $c > 0$ y $\lambda > 0$

$$R(1; c, \lambda) = c + \bar{\rho}(1; c, \lambda). \quad (3.106)$$

El siguiente lema muestra que (3.106) es de hecho un caso particular de una relación más general.

Lema 27.

Para toda $z > 0$, $c > 0$ y $\lambda > 0$

$$R(z; c, \lambda) = c + \bar{\rho}(z; c, \lambda). \quad (3.107)$$

Demostración del Lema 27: Se sigue de (3.105) que

$$\begin{aligned} R(z; c, \lambda) &= z \inf_{(\psi, \phi) \in \mathfrak{S}_0} \left(\frac{c}{z} E_0 \tau_\psi + \frac{\lambda}{z} \alpha(\psi, \phi) + \beta(\psi, \phi) \right), \\ &= z(c/z + \bar{\rho}(1; c/z, \lambda/z)) = c + z\bar{\rho}(1; c/z, \lambda/z). \end{aligned} \quad (3.108)$$

Con base en las definiciones (3.93) - (3.95) y (3.100), no es difícil obtener que

$$z\bar{\rho}(1; c/z, \lambda/z) = \bar{\rho}(z; c, \lambda) \quad \text{para toda } z, c, \lambda > 0. \quad (3.109)$$

Al aplicar (3.109) en el lado derecho de (3.108), se obtiene (3.107). ■

Lema 28.

La función $R(z; c, \lambda)$ definida en (3.105) es cóncava y conjuntamente continua sobre $\{z \geq 0, c \geq 0, \lambda \geq 0\}$.

Demostración del Lema 28:

Debido a que R definida en (3.105) es el ínfimo de una familia de funciones lineales, entonces es cóncava sobre $\{z \geq 0, c \geq 0, \lambda \geq 0\}$, y de esto se sigue por el Teorema 10.1 de (Rockafellar, 1972) que es conjuntamente continua sobre $\{z > 0, c > 0, \lambda > 0\}$.

Queda por demostrar que $R(z; c, \lambda)$ es continua para cualquier punto $c = 0, \lambda = 0$ o $z = 0$.

La parte menos trivial es cuando $c = 0$ con $\lambda > 0$ y $z > 0$.

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

Sea k cualquier número natural fijo. Definimos la regla de paro ψ de tal manera que $s_n^\psi = 1$ para toda $\mathbf{n} \in \mathcal{G}^k$, y sea la regla de decisión ϕ con $\phi_{\mathbf{n}} = I_{\{z_n \geq \lambda/z\}}$, $\mathbf{n} \in \mathcal{G}^k$. En lo siguiente vamos a aplicar la desigualdad de Markov: $P(|X| \geq a) \leq E(|X|)/a$ con $a > 0$.

$$\begin{aligned}
\lambda\alpha(\psi, \phi) &= \lambda \sum_{i=1}^{\infty} \sum_{\mathbf{n} \in \mathcal{G}^i} p_{\mathbf{n}} E_0 s_{\mathbf{n}}^\psi \phi_{\mathbf{n}} = \lambda \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 \phi_{\mathbf{n}}, \quad \text{ya que para } i \neq k, s_{\mathbf{n}}^\psi = 0, \\
&= \lambda \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} P_0(Z_{\mathbf{n}} \geq \lambda/z), \\
&\leq \sqrt{\lambda z} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 Z_{\mathbf{n}}^{1/2}, \quad \text{por desigualdad de Markov,} \\
&= \sqrt{\lambda z} \left(\sum_{n_1 \in \mathcal{G}} p(n_1) (E_0 Z_{n_1}^{1/2}) \cdots \sum_{n_k \in \mathcal{G}} p(n_k) (E_0 Z_{n_k}^{1/2}) \right) \quad \text{por independencia de las v.a.} \\
&= \sqrt{\lambda z} r^k, \tag{3.110}
\end{aligned}$$

donde

$$r = \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} E_0 Z_{\mathbf{n}}^{1/2} < 1.$$

Esta última desigualdad se tiene en virtud de la conocida *distancia de Hellinger*: donde para f_0 y f_1 medidas de probabilidad absolutamente continuas respecto a la medida μ ,

$$H^2(f_0, f_1) = \frac{1}{2} \int (\sqrt{f_0(x)} - \sqrt{f_1(x)})^2 d\mu = 1 - \int \sqrt{f_0(x)f_1(x)} d\mu, \quad \text{y } 0 \leq H^2(f_0, f_1) \leq 1.$$

De lo cual $E_0 Z_{\mathbf{n}}^{1/2} = \int f_0^{1/2} f_1^{1/2} d\mu < 1$, dado que se satisface (3.2).

Análogamente a (3.110), se tiene que:

$$z\beta(\psi, \phi) = z \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} P_1(Z_{\mathbf{n}}^{-1} > (\lambda/z)^{-1}) \leq \sqrt{\lambda z} \sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_1 Z_{\mathbf{n}}^{-1/2} = \sqrt{\lambda z} r^k. \tag{3.111}$$

Sea ϵ cualquier entero positivo. Tomando $c = 0$ en (3.105) y

$$k > \left(\ln \epsilon - \ln(\sqrt{\lambda z}) \right) / \ln r \tag{3.112}$$

se obtiene de (3.110) y (3.111) que

$$R(z; 0, \lambda) \leq \lambda\alpha(\psi, \phi) + z\beta(\psi, \phi) < 2\epsilon.$$

Debido a que ϵ puede ser arbitrariamente pequeño, de esto se sigue que $R(z; 0, \lambda) = 0$, para cualesquiera $z > 0, \lambda > 0$.

Ahora, para demostrar la continuidad en cualquier punto $(z, 0, \lambda)$, con $\lambda, z > 0$, primero tomemos algunas sucesiones $\lambda_n \rightarrow \lambda, z_n \rightarrow z, c_n \rightarrow 0$, cuando $n \rightarrow \infty$.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Nuevamente, sea $\epsilon > 0$ cualquier número. Si k es suficientemente grande para satisfacer (3.112), entonces se tiene que

$$R(z_n; c_n, \lambda_n) \leq c_n k + \lambda_n \alpha(\psi, \phi) + z_n \beta(\psi, \phi) < 5\epsilon$$

para toda n tal que $|\lambda_n - \lambda| < \epsilon$, $|z_n - z| < \epsilon$ y $c_n < \epsilon/k$, esto es $\lim_{n \rightarrow \infty} R(z_n; c_n, \lambda_n) = R(z; 0, \lambda) = 0$.

Ahora, si $\lambda = 0$ y $z \geq 0$, $c \geq 0$, tomando la misma prueba secuencial (ψ, ϕ) definida al principio, pero con $k = 1$, se obtiene que $\phi_n = I_{\{z_n \geq 0\}} = 1$, y así $\alpha(\psi, \phi) = 1$ y $\beta(\psi, \phi) = 0$, de este modo, $R(z; c, 0) = \inf_{(\psi, \phi) \in \mathfrak{S}_0} (c E_0 \tau_\psi + z \beta(\psi, \phi)) \leq c$. Por otro lado, $E_0 \tau_\psi \geq 1$, entonces $R(z; c, 0) = c$. Si, ahora $(z_n, c_n, \lambda_n) \rightarrow (z, c, 0)$, as $n \rightarrow \infty$, entonces $R(z_n; c_n, \lambda_n) \leq c_n + \lambda_n \rightarrow c$, cuando $n \rightarrow \infty$. Por otra parte, $R(z_n; c_n, \lambda_n) = \inf_{(\psi, \phi) \in \mathfrak{S}_0} (c_n E_0 \tau_\psi + \lambda_n \alpha(\psi, \phi) + z_n \beta(\psi, \phi)) \geq c_n \rightarrow c$, cuando $n \rightarrow \infty$. Por tanto se obtiene la continuidad sobre $\{z \geq 0, c \geq 0, \lambda = 0\}$.

Finalmente el caso cuando $z = 0$ y $c \geq 0$, $\lambda \geq 0$ puede ser tratado análogamente. Por lo tanto, $R(z; c, \lambda)$ es conjuntamente continua sobre $\{z \geq 0, c \geq 0, \lambda \geq 0\}$. ■

Con base en (3.107), es conveniente extender la definición de $\bar{\rho}(z) = \bar{\rho}(z; c, \lambda)$ (definida inicialmente solo para números positivos c, λ) de tal manera que $R(z; c, \lambda)$ en (3.107) se satisfaga para toda $z \geq 0, c \geq 0, \lambda \geq 0$, esto es definiendo $\bar{\rho}(z; c, \lambda) = 0$ siempre que alguno de sus argumentos sea cero. De lo cual se plantea el siguiente corolario.

Corolario 14.

La función $\bar{\rho}(z; c, \lambda)$ es cóncava y continua sobre $\{z \geq 0, c \geq 0, \lambda \geq 0\}$.

La demostración de este Corolario 14 se sigue de los Lemas 27 y 28.

Ahora, vamos a continuar con el análisis de la condición $\lambda > c + \bar{\rho}(\lambda; c, \lambda)$ (3.104). El siguiente lema es semejante en cuanto a contenido y estructura a los Lemas 25 y 26, solo quitando las distribuciones r de los tamaños de grupos para las constantes.

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

Lema 29.

Si se cumple que la condición (3.75) entonces para cualesquiera números positivos c y λ que satisfacen la condición $\lambda > c + \bar{\rho}(\lambda; c, \lambda)$ (3.104) existen números $0 < A < \lambda$ y $B > \lambda$, tales que

$$g(A; \lambda) = c + \bar{\rho}(A; c, \lambda), \quad g(B; \lambda) = c + \bar{\rho}(B; c, \lambda), \quad (3.113)$$

y

$$g(z; \lambda) < c + \bar{\rho}(z; c, \lambda) \quad \text{para toda } 0 \leq z < A \quad \text{y toda } z > B, \quad (3.114)$$

y

$$g(z; \lambda) > c + \bar{\rho}(z; c, \lambda) \quad \text{para toda } A < z < B. \quad (3.115)$$

Demostración del Lema 29:

Para cualesquiera $c > 0$ y $\lambda > 0$ se definen las funciones:

$$D_1(z) = D_1(z; c, \lambda) = z - \bar{\rho}(z; c, \lambda) \quad \text{y} \quad D_2(z) = D_2(z; c, \lambda) = \lambda - \bar{\rho}(z; c, \lambda), \quad \text{para } z \geq 0.$$

Se sigue obviamente que $D_1(\lambda) = D_2(\lambda)$. En virtud del Corolario 14 y de las propiedades de $\bar{\rho}(z; c, \lambda)$ se tiene que $D_1(z)$ y $D_2(z)$ son funciones convexas y continuas en $[0, \infty)$.

Ahora, aplicando la desigualdad de Jensen a la función cóncava ρ ,

$$\bar{\rho}(z) = \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} E_0 \rho(z z_{\mathbf{n}}) \leq \sum_{\mathbf{n} \in \mathcal{G}} p_{\mathbf{n}} \rho(z) = \rho(z) \leq g(z), \quad (3.116)$$

así $g(z; \lambda) - \bar{\rho}(z; c, \lambda) = \min \{D_1(z), D_2(z)\} \geq 0$.

En virtud del Lema 24, si un número positivo c es tal que $c < D_1(\lambda)$ (lo cual es equivalente a $\lambda > c + \bar{\rho}(\lambda; c, \lambda)$), entonces por la continuidad de D_1 , existe un único $A < \lambda$ tal que $D_1(A) = c$, y $D_1(z) < c$ para toda $z < A$ y $D_1(z) > c$ para toda $z > A$.

No es difícil deducir que (3.75) implica $\lim_{z \rightarrow \infty} D_2(z) = 0$.

De manera análoga a la propiedad de la función D_1 , si $0 < c < D_2(\lambda) = D_1(\lambda)$, entonces existe un único número $B > \lambda$ tal que $D_2(B) = c$, y $D_2(z) > c$ para $z < B$ y $D_2(z) < c$ para $z > B$.

Por tanto, se obtienen (3.113), (3.114) y (3.115). ■

Se sigue del Lema 29 que, si la condición (3.75) se cumple, entonces (3.101) es equivalente a

$$I_{\{z_{\mathbf{n}} \in (A, B)\}} \leq 1 - \psi_{\mathbf{n}} \leq I_{\{z_{\mathbf{n}} \in [A, B]\}}. \quad (3.117)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

Esto es, cualquier prueba óptima, deducida en este trabajo, es una versión aleatorizada de la prueba secuencial de razón de probabilidades (RSPRT) establecida por Mukopadhyay la cual, en nuestros términos, puede ser descrita como (ψ, ϕ) donde

$$\psi_{\mathbf{n}} = I_{\{z_{\mathbf{n}} \notin (A,B)\}} \quad \text{y} \quad \phi_{\mathbf{n}} = I_{\{z_{\mathbf{n}} \geq B\}}, \quad (3.118)$$

para toda $\mathbf{n} \in \mathcal{G}^r$ y $r \geq 1$.

Obviamente, $\psi_{\mathbf{n}}$ en (3.118) es un caso particular de (3.117), y $\phi_{\mathbf{n}}$ satisface la estructura en (3.13) de la condición Δ_{ψ} . Además, en virtud del Teorema 3.1 en Mukhopadhyay, se tiene que $(\psi, \phi) \in \mathfrak{S}_0$ (los detalles se pueden encontrar en la demostración del Teorema 17 que se verá a continuación). Consecuentemente, se sigue que la RSPRT es óptima en el sentido del Corolario 12.

De la misma manera, todas las pruebas secuenciales (ψ, ϕ) tales que ψ y ϕ satisfacen respectivamente (3.117) y (3.118) (para toda $\mathbf{n} \in \mathcal{G}^r$ y $r \geq 1$) comparten la propiedad de optimalidad que la RSPRT, cuando se cumple la condición (3.75). En particular, obviamente, este es el caso cuando las distribuciones hipotetizadas pertenecen a la familia exponencial (o Koopman-Darmois).

Si la condición (3.75) no se satisface, las pruebas óptimas con sus reglas de paro (3.101) no necesariamente son del tipo RSPRT. Esto se puede ver en el siguiente ejemplo.

Consideremos una prueba secuencial con grupos de tamaño uno, donde bajo H_0 , la hipótesis de prueba, las observaciones siguen una distribución uniforme $[0, 1]$, mientras que bajo H_1 ellas siguen una distribución uniforme $[0, 0.5]$. Para esta prueba se tiene que $P_0(f_1(X) > 0) = \int_0^{0.5} 1 dx = 0.5$. Sean $c = 1$ y $\lambda = 2$, con base en la definición (3.93) se tiene que $\rho_0(z; 1, 2) = g(z; 2) = \min\{z, 2\}$.

Ahora de la definición (3.95) para $\bar{\rho}_0$ se sigue que:

$$\begin{aligned} \bar{\rho}_0(z; 1, 2) &= \sum_{n \in \mathcal{G}} p(n) E_0 \rho_0(z z_n; 1, 2) = \sum_{n \in \mathcal{G}} p(n) E_0 \min\{z z_n, 2\}, \\ &= \sum_{n \in \mathcal{G}} p(n) \left(\int_0^{0.5} \min\{2z, 2\} f_0(x) dx + \int_{0.5}^1 0 f_0(x) dx \right), \\ &= \sum_{n \in \mathcal{G}} p(n) \min\{2z, 2\} (0.5) = \min\{z, 1\}. \end{aligned}$$

De (3.94), $\rho_1(z; 1, 2) = \min\{g(z; 2), 1 + \bar{\rho}_0(z; 1, 2)\} = \min\{\min\{z, 2\}, 1 + \min\{z, 1\}\}$.

Pero es fácilmente demostrable que $\min\{z, 2\} \leq 1 + \min\{z, 1\}$ para toda $z \geq 0$. Por tanto $\rho_1(z; 1, 2) = \min\{z, 2\}$. En general, con base en (3.98) para $\rho(z)$ es fácil demostrar que $\rho(z; c, \lambda) = \rho(z; 1, 2) = g(z; 2) = \min\{z, 2\}$ y $c + \bar{\rho}(z; c, \lambda) = 1 + \bar{\rho}(z; 1, 2) = 1 + \min\{z, 1\}$. Ahora consideremos la prueba (óptima) correspondiente a (3.102), con la desigualdad estricta.

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

De aquí, inmediatamente, se sigue que, con las definiciones de arriba de $g(z)$ y $\bar{\rho}(z)$, la condición (3.102) se cumple si y solo si $z < 2$. Pero, consecuentemente, los valores de las razones de probabilidad z_1, z_2, \dots, z_k son, respectivamente, $2, 4, 8, \dots, 2^k$, cuando $X_1, X_2, \dots, X_k \leq 0.5$, así la prueba (que minimiza K_0) solo se detiene cuando, para el primer tiempo, $X_i > 0.5$, de lo cual $z_i = 0$.

Se puede observar de este ejemplo, primero que la prueba que minimiza a K_0 no es una RSPRT (porque una RSPRT debería parar también cuando $z_i \geq B > 0$ lo cual no sucede, bajo H_0 , con una probabilidad positiva, así, habría un error positivo α), y segundo, que, en ninguna forma, esto minimiza a K_1 , porque bajo H_1 la regla nunca se detiene.

Sin embargo, el siguiente teorema muestra que, aún si (3.75) no se satisface, no solo la prueba RSPRT con ψ que satisface (3.118), sino también su versión aleatorizada ψ (3.117) son óptimas en el sentido de Wald-Wolfowitz [Wald-Wolfowitz, (1948)], i.e. ellas minimizan el costo promedio total bajo ambas hipótesis H_0 y H_1 , dadas las restricciones en las probabilidades de error.

Lema 30.

Sean A y B tales que $0 < A < B < \infty$. Entonces existen λ y c , tales que $A \leq \lambda < B$ y $c > 0$ y las ecuaciones en (3.113) se satisfacen.

Demostración del Lema 30.

Sean A y B tales que $0 < A < B < \infty$. Para cualquier $\lambda \in [A, B]$, definimos $c = c(\lambda)$ como una solución de la ecuación

$$c + \bar{\rho}(A; c, \lambda) = A. \quad (3.119)$$

La existencia de una única solución $c = c(\lambda)$ de (3.119) se sigue del hecho de que el lado izquierdo de (3.119) es una función continua y estrictamente creciente de c que toma valores en el intervalo $[0, \infty)$.

Es más, la función $c(\lambda)$ es también continua respecto a λ , como una función implícita de (3.119), definida mediante una función que es continua en todas sus variables (debido al Corolario 14). Además, $c(\lambda) > 0$ para toda $\lambda \in [A, B]$, porque de lo contrario esto implicaría en virtud de (3.119) que $\bar{\rho}(A; 0, \lambda) = A$, i.e. $A = 0$, lo cual es una contradicción.

Ahora, definamos

$$G(\lambda) = \lambda - \bar{\rho}(B; c(\lambda), \lambda) - c(\lambda),$$

la cual es una función continua de λ , ya que es una resta de funciones continuas.

Se demostrará que

$$G(A) \leq 0 \quad \text{y} \quad G(B) > 0. \quad (3.120)$$

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

De hecho,

$$G(A) = A - \bar{\rho}(B; c(A), A) - c(A) \leq A - \bar{\rho}(A; c(A), A) - c(A) = 0, \quad (3.121)$$

(debido a (3.119)).

Ahora se demostrará que

$$G(B) = B - \bar{\rho}(B; c(B), B) - c(B) > 0. \quad (3.122)$$

Tomando en cuenta, por (3.119),

$$c(B) + \bar{\rho}(A; c(B), B) = A,$$

se puede observar que (3.122) es equivalente a

$$B - \bar{\rho}(B; c(B), B) > A - \bar{\rho}(A; c(B), B). \quad (3.123)$$

Debido a que $F(z) = z - \bar{\rho}(z; c(B), B)$ satisface la condición del Lema 24, lo contrario a (3.123) debería implicar que $F(A) = 0$, esto es, $A - \bar{\rho}(A; c(B), B) = 0$, o, en vista de (3.119), $c(B) = 0$, es una contradicción.

Así, (3.120) está demostrada, por tanto existe $\lambda \in [A, B)$ tal que $G(\lambda) = 0$, es decir, se ha encontrado $\lambda \in [A, B)$ y $c = c(\lambda) > 0$ tales que

$$c + \bar{\rho}(A; c, \lambda) = A \quad \text{y} \quad c + \bar{\rho}(B; c, \lambda) = \lambda,$$

lo cual es equivalente a (3.113). ■

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

Teorema 17.

Sean $A < B$ dos constantes positivas. Sea ψ cualquier regla de paro que satisfice que

$$I_{\{z_n \in (A, B)\}} \leq 1 - \psi_n \leq I_{\{z_n \in [A, B]\}}, \quad (3.124)$$

para toda $\mathbf{n} \in \mathcal{G}^k$ y $k = 1, 2, \dots$, y sea ϕ una regla de decisión definida como

$$\phi_n = I_{\{z_n \geq B\}} \quad (3.125)$$

para toda $\mathbf{n} \in \mathcal{G}^k$ y $k = 1, 2, \dots$.

Entonces $(\psi, \phi) \in \mathfrak{S}_0 \cap \mathfrak{S}_1$, y esta prueba es óptima en el siguiente sentido: para cualquier prueba secuencial $(\psi', \phi') \in \mathfrak{S}_0 \cap \mathfrak{S}_1$ tal que

$$\alpha(\psi', \phi') \leq \alpha(\psi, \phi) \quad \text{y} \quad \beta(\psi', \phi') \leq \beta(\psi, \phi) \quad (3.126)$$

se tiene que

$$K_0(\psi) \leq K_0(\psi') \quad \text{y} \quad K_1(\psi) \leq K_1(\psi'). \quad (3.127)$$

Ambas desigualdades en (3.127) son estrictas, si al menos una de las desigualdades en (3.126) es estricta.

Demostración del Teorema 17

Con base en los Lemas 29 y 30 se tiene que la condición $\Pi_0^\infty(\psi)$ se satisface. La condición Δ_ψ también se satisface debido a que $B \geq \lambda$. Para aplicar el Corolario 12 falta demostrar que $(\psi, \phi) \in \mathfrak{S}_0$, i.e. que

$$\sum_{\mathbf{n} \in \mathcal{G}^k} p_{\mathbf{n}} E_0 t_{\mathbf{n}}^\psi = P_0(\tau^\psi \geq k) \rightarrow 0, \quad \text{cuando} \quad k \rightarrow \infty. \quad (3.128)$$

Esto se sigue del Teorema 3.1 de Mukhopadhyay [Mukhopadhyay-De Silva, (2008)] que indica que existen $a > 0$ y $0 < r < 1$ tales que $P_0(\tau^\psi \geq k) \leq ar^k$ para todo número natural k , de modo que (3.128) se satisface. Las condiciones del Teorema de Mukopadhyay [Mukhopadhyay-De Silva, (2008)] se satisfacen por las suposiciones de nuestro modelo y por (3.2).

Aquí, por el Corolario 12, la prueba (ψ, ϕ) tiene un valor mínimo para $K_0(\psi)$ entre todas las pruebas secuenciales (ψ', ϕ') que satisfacen las restricciones (3.127) sobre las probabilidades de error.

Para demostrar que esta misma prueba minimiza a $K_1(\psi)$, podemos aplicar nuevamente el Lema 30, solo intercambiando las distribuciones hipotetizadas. Consideremos dos hipótesis simples H'_0 : “la verdadera distribución está dada por f_1 ” vs H_1 : “la distribución corresponde a f_0 ”.

CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS ALEATORIOS Y OBSERVACIONES INDEPENDIENTES

De una manera natural, cualquier prueba secuencial (ψ, ϕ) que satisfice (3.124) y (3.125) es inmediatamente adaptada al problema de pruebas de hipótesis H'_0 vs. H'_1 :

$$\psi_{\mathbf{n}}^*(x^{\mathbf{n}}) = \psi_{\mathbf{n}}(x^{\mathbf{n}}) \quad \text{y} \quad \phi_{\mathbf{n}}^*(x^{\mathbf{n}}) = 1 - \phi_{\mathbf{n}}(x^{\mathbf{n}}), \quad (3.129)$$

de modo que

$$I_{\{z_{\mathbf{n}}^{-1} \in (B^{-1}, A^{-1})\}} \leq 1 - \psi_{\mathbf{n}}^* \leq I_{\{z_{\mathbf{n}}^{-1} \in [B^{-1}, A^{-1}]\}}, \quad (3.130)$$

y

$$\phi_{\mathbf{n}}^* = I_{\{z_{\mathbf{n}}^{-1} > B^{-1}\}} \quad (3.131)$$

para toda $\mathbf{n} \in \mathcal{G}^k$ y $k = 1, 2, \dots$

Sean α' y β' la notación de las probabilidades de error en el problema de prueba H'_0 vs. H'_1 y K'_0 el costo promedio de observaciones, bajo H'_0 . Entonces, obviamente,

$$\alpha'(\psi^*, \phi^*) = \beta(\psi, \phi), \quad \text{y} \quad \beta'(\psi^*, \phi^*) = \alpha(\psi, \phi), \quad (3.132)$$

y $K'_0(\psi^*) = K_1(\psi)$.

Ahora aplicando el Lema 30 al problema de prueba H'_0 vs. H'_1 en la misma manera en como se aplicó al problema de prueba H_0 vs. H_1 , entonces obtendremos, mediante el Corolario 12, que (ψ, ϕ) minimiza a $K_1(\psi)$ entre todas las pruebas secuenciales $(\psi', \phi') \in \mathfrak{S}_1$ que satisfacen (3.127). ■

La propiedad de optimalidad establecida en el Teorema 17, en el caso de una observación por grupo, es conocida como la optimalidad de la SPRT demostrada por Wald-Wolfowitz (véase [Mukhopadhyay-De Silva, (2008)]). A su vez, Burkholder y Wijsman demostrarán que todas las SPRT “extendidas”, i.e. aquellas que admiten una decisión aleatorizada entre parar y continuar para el caso $z_n = A$ o $z_n = B$, comparten la misma propiedad de optimalidad que la SPRT, [Burkholder-Wijisman, (1963)]. Nuestro Teorema 17 establece lo mismo, pero para el caso de pruebas secuenciales con grupos de tamaño aleatorio: pruebas secuenciales por grupos “extendidas”, i.e. aquéllas con reglas de paro que satisfacen (3.124), minimizando así el costo promedio total del experimento tanto bajo H_0 como bajo H_1 .

De manera muy parecida al caso clásico, de una observación por grupo, no se permite hacer observaciones, cuando $1 \in [A, B]$, solo en el caso significativo para la RSPRT, porque en cualquier otro caso una prueba trivial (la que sin ninguna observación acepta o rechaza H_0 en dependencia si $1 < A$ o $1 > B$) se desempeña mejor que las pruebas óptimas del Teorema 17 (mín $\{1, \lambda\} < c + \bar{\rho}(1; c, \lambda)$ en términos de la regla de paro del Teorema 15).

Finalmente, para esta prueba (ψ^*, ϕ^*) , generalización de la RSPRT, tomemos su tiempo de paro asociado. Sea

$$\tau^{\psi^*} = \text{mín} \left\{ k : Z_{\nu} \notin (A, B) \right\} \quad (3.133)$$

3.2. Optimalidad de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT)

Este tiempo de paro tiene el mismo sentido práctico que el trabajado en ([Mukhopadhyay-De Silva, (2008)] p. 434). Dicho tiempo está directamente asociado a la regla de paro ψ^* , debido a que cuando ella se detiene ($Z_n \notin (A, B)$) se tendrá un tiempo de paro óptimo ².

²O un mínimo de grupos observados en el experimento.

**CAPÍTULO 3. PRUEBA SECUENCIAL ÓPTIMA CON GRUPOS
ALEATORIOS Y OBSERVACIONES INDEPENDIENTES**

CAPÍTULO 4

CONCLUSIONES Y PERSPECTIVAS

En esta última parte vamos a recapitular todos los resultados encontrados en esta extensa exploración, y a su vez establecer algunas de las propuestas y/o perspectivas para continuar con la investigación en cada caso.

Antes que nada, recordemos que en el análisis secuencial clásico, el estudio de las observaciones se realiza a través de etapas usualmente tomadas de una por una o de un tamaño fijo en cada etapa. En el primer contexto, existe la prueba secuencial óptima, llamada SPRT, la cual conlleva a un tiempo de paro (τ), en promedio, menor que cualquier otra prueba aún en el contexto clásico estadístico. En general, los métodos secuenciales conducen a ahorros en tamaño de la muestra, tiempo y costo en comparación con los procedimientos estándar de muestras fijas.

En esta investigación, las observaciones fueron consideradas en grupos de tamaño aleatorio, esto es cada grupo de observaciones puede tener un tamaño indeterminado de etapa en etapa (incluyendo tamaño cero). Anteriormente, se mencionó cómo algunas aplicaciones están considerando en sus investigaciones a grupos de diferente tamaño en cada etapa (no necesariamente aleatorio), tales como: análisis clínicos en la industria farmacéutica (véanse, [Bartroff, *et al*, (2013)] y [Jennison-Turnbull, (1999)]), estudios en control de calidad, entre otros. Una ventaja significativa para estas aplicaciones es el hecho de tener procedimientos teóricos y prácticos con grupos aleatorios que les “garantice” costos promedios mínimos del experimento, y/o también, tiempos de paro y observaciones menores. Es por esto que es muy importante que el resultado teórico aquí presentado sobre las estructuras de las pruebas secuenciales óptimas sean implementadas, al menos numéricamente para alguna distribución en particular (binomial, uniforme, normal, etc.), y a su vez poder determinar características propias de la prueba misma como: probabilidades de error, potencia de la prueba, región de continuación, eficiencia, etc. Actualmente debido a la crisis sanitaria que se generó en los años pasados por la pandemia, el estudio óptimo de tratamientos farmacéuticos es de suma importancia para su investigación.

CAPÍTULO 4. CONCLUSIONES Y PERSPECTIVAS

Otra exploración que surge para las pruebas óptimas encontradas es precisamente investigar las características propias de la misma como la potencia teórica de la prueba, la eficiencia de la misma, etc.

En el capítulo 2, se logró demostrar y dar la estructura de la prueba secuencial óptima tanto para un horizonte finito, como para uno infinito, con grupos de tamaño aleatorio (secciones 2.1.2 y 2.1.3, respectivamente). Estructuras óptimas que conllevan a un costo promedio del experimento menor que cualquier otra prueba secuencial que tenga probabilidades de error menores que la primera. Las estructuras encontradas de la prueba óptima (finita y no finita) consideran, en primera instancia, a las observaciones sin suponer que éstas son idenpendientes. Este hecho da una importancia fuerte al trabajo establecido debido a que, generalmente en las aplicaciones, se consideran observaciones independientes, aún cuando en la “realidad” esta suposición no necesariamente es posible. Por tanto, para futuras investigaciones dicha estructura óptima establecida podría ser aplicada, en particular para algunas distribuciones de probabilidad específicas, tales como la Bernoulli, Binomial y Normal o cualquier otra de interés.

Otro resultado importante del capítulo 2, sección 2.1, fue que para el caso truncado, se demostró que el método de multiplicadores de Lagrange “generalizado” (aplicado a este estudio) es un método necesario y suficiente para demostrar que la función objetivo (costo promedio K) alcanza un mínimo, comparada con aquellas pruebas que tienen respectivas probabilidades de error tipo I y II menores. Usualmente, el método de Lagrange se aplicaba, en la mayoría de las investigaciones clásicas de la estadística (aún en la frecuentista), para demostrar la optimalidad de una función con restricciones (véase Teorema 1, p. 22). Es decir, es suficiente tener el ínfimo de una función sin restricciones para obtener el mínimo de una de sus componentes (o problema con restricciones). Este resultado está basado en lo que Wald llamaba admisibilidad de una prueba. No obstante, en nuestro trabajo, hemos demostrado, para el caso truncado, que también es necesario tener el ínfimo de la función “generalizada” de Lagrange, para demostrar que se obtendrá el mínimo de una de sus componentes (véase Teorema 6, p.48). Aquí surgen algunas preguntas aún sin responder, como por ejemplo, para el caso con horizonte infinito ¿es posible encontrar condiciones necesarias para demostrar la optimalidad del problema con restricciones? o aún para el caso truncado se podría hacer un análisis más profundo de los posibles resultados del Teorema 6, que por falta de tiempo aquí no pudimos realizar. Todo lo relacionado con el método de Lagrange tiene como detalles a sus multiplicadores $\lambda_0 > 0$ y $\lambda_1 > 0$, los cuales fueron valores cualesquiera, pero fijos durante todos los procedimientos. Mas quedan pendientes preguntas de investigación como: ¿es posible variar estos multiplicadores para encontrar también la estructura óptima? ya sea ¿para el caso finito o infinito? Estas preguntas aún necesitan ser analizadas. Tampoco fue posible hacer un análisis sobre cuando el problema solo tiene una restricción, por ejemplo cuando solo se quiere minimizar el error tipo I α restringido al error tipo II β . Esto también es un caso particular del método “generalizado” de Lagrange.

En la sección 2.1.3, (regla de paro óptima: caso no truncado), además de encontrar la estructura óptima de la regla de paro que conlleva a un costo promedio mínimo; se demostró que el problema de Riesgo Bayesiano se puede ver como caso particular de nuestra función

de Lagrange (véase, Corolario 7, p. 61), solo considerando a la tercera distribución como una combinación convexa de las primeras dos distribuciones, es decir para un problema de pruebas de hipótesis Bayesiano $(P_0, P_1, \pi P_0 + (1 - \pi)P_1)$ con $\pi \in [0, 1]$ se tiene que toda regla de paro en este contexto es truncable (i.e. su función de Lagrange truncada converge a la Lagrange no truncada). Este pequeño resultado importante, podría ser también analizado más profundamente para las aplicaciones Bayesianas, de hecho nuestros resultados óptimos podrían ser analizados en dicho contexto. Cabe mencionar que aunque nuestro trabajo está en el contexto frecuentista, el análisis secuencial se aplica tanto a la estadística frecuentista como a la Bayesiana, y con el resultado del corolario 7 se muestra la importancia de ambas áreas para el desarrollo de la propia estadística.

En la sección 2.2, se reescribió la estructura de la prueba óptima truncada para la resolución del problema de Kiefer-Weiss con observaciones i.i.d. y tamaños de grupos aleatorios e independientes, bajo un contexto de un problema de hipótesis simples. Resultado que no se había trabajado en la literatura encontrada. En particular, se consideraron a las tres distribuciones de probabilidad pertenecientes a la misma familia exponencial, donde el tercer parámetro se encuentra entre los otros dos ($\theta_0 < \theta < \theta_1$). Entonces, dado que se trabajaron con las condiciones “clásicas” de este problema, ahora también surgen las preguntas sobre la generalización de los resultados, tales como: si se puede trabajar con observaciones solo independientes (sin que sean i.d.), si se puede analizar la escritura de la prueba óptima para el caso no truncando (i.e. el problema K-W para el horizonte infinito); esto para el caso cuando el tercer parámetro θ esté fuera del intervalo (θ_0, θ_1) . Qué aplicaciones se podrían tener; si se puede resolver el problema para otro tipo de familia paramétrica o en general si las distribuciones pueden pertenecer a diferentes familias paramétricas. Finalmente, es importante realizar aplicaciones de los resultados encontrados de manera numérica o computacional para casos particulares con ciertas distribuciones como: Bernoulli (como se realizó en el trabajo de [Novikov, *et al.*, (2022)]), normal u otras. Para esta pequeña sección 2.2 se pueden tomar una serie de variaciones respecto a las condiciones iniciales que pueden conllevar a investigaciones futuras.

Finalmente, la cereza del pastel de toda esta ardua investigación es el capítulo 3, en particular la sección 3.2 en donde se da respuesta a la investigación planteada en 2008 por Mukhopadhyay & De Silva, pionera del planteamiento teórico del uso de grupos de tamaño aleatorio y de la prueba secuencial aleatoria de la razón de probabilidades (RSPRT). En dicha investigación se deja abierta la pregunta de la optimalidad de la RSPRT. Nosotros en, esta sección, demostramos que una prueba aleatoriza óptima donde la RSPRT es un caso particular. Esta generalización de la RSPRT es óptima en el sentido de que minimiza el costo promedio del experimento tanto para la hipótesis nula H_0 y como para la hipótesis H_1 (véase Teorema 17, p.128). Donde además se presenta la estructura de la prueba óptima en términos de la razón de verosimilitud $z = f_0/f_1$ y de las constantes A y B que definen la región de continuación (o rechazo) para dicha prueba. Para este importante resultado teórico, lo primero y más interesante debería ser el estudiar sus aplicaciones, ya que existen varias áreas de investigación en las cuales se pudiera aplicar, como en la teoría de juegos, en análisis clínicos, en psicología so-

CAPÍTULO 4. CONCLUSIONES Y PERSPECTIVAS

cial, en técnicas de confiabilidad, etc. En dichas aplicaciones se podrían especificar a su vez los parámetros de estudio representados bajo las hipótesis a contrastar. Es decir, estudiar la estructura de la prueba óptima aplicada a distribuciones particulares como por ejemplo, distribuciones Bernoulli, binomial, uniforme, normal, exponencial, entre otras. Es importante mencionar que en esta parte del trabajo no se consideró a las distribuciones en una familia de distribución en particular, de hecho en investigaciones previas [Popoca, (2012)] sí se consideran distribuciones finitas, mas en esta sección solo se tomaron a las observaciones y los tamaños de grupos como i.i.d.

Cabe mencionar, que en la literatura encontrada se trabaja con problemas de hipótesis simples, pero también cabría preguntarse por una investigación futura que involucrase pruebas de hipótesis no simples.

BIBLIOGRAFÍA

- [Anderson, (1959)] Anderson, T. W., (1959), *A modification of the sequential probability ratio test to reduce the sample size*, The Annals of Mathematical Statistics, 165-197.
- [Barber, (2003)] Barber, S., (2003), *An analysis of group size models for optimal group sequential clinical trials*, Statistics group reports, University of Bristol, 1-16.
- [Bartroff, *et al*, (2013)] Bartroff, J., Lai T. L. & Shih, M. C., (2013), *Sequential Experimentation in Clinical Trials: Design and Analysis*, Springer, Series in Statistics.
- [Bhattachajee-Mukhopadhyay, (2012)] Bhattacharjee, D. & Mukhopadhyay, N., (2012), *On SPRT and RSPRT for the unknown mean in a normal distribution with equal mean and variance*, Sequential Analysis, 31: 108-134.
- [Burkholder-Wijisman, (1963)] Burkholder, D. L. & Wijsman, R. A., (1963), *Optimum properties and admissibility of sequential test*, The Annals of Mathematical Statistics, 34: 1-17.
- [Chow *et al*, (1971)] Chow, Y. S., Robbins, H. & Siegmund, D., (1971), *Great Expectations: The Theory of Optimal Stopping*, Boston, Houghton Mifflin Company.
- [Cressie-Morgan, (1993)] Cressie, N. & Morgan, P. B., (1993) *The VRPT: A sequential testing procedure dominating the SPRT*, Econometric Theory 9: 431-450.
- [Denne-Jennison, (2000)] Denne, J. S. & Jennison, C., (2000), *A group sequential test with updating of sample size*, Biometrika: 125-134.
- [Ehrenfeld, (1972)] Ehrenfeld, S., (1972), *On group sequential sampling*, Technometrics, 14-1: 167-174.
- [Eisenberg, *et al*, (1976)] Eisenberg, B., Ghosh, B. K. & Simons, G., (1976), *Properties of generalized sequential probability ratio test*, The Annals of Statistics, 4: 237-251.

BIBLIOGRAFÍA

- [Ghosh, *et al* (1997)] Ghosh, M., Mukhopadhyay, N. & Sen, P., (1997), *Sequential Estimation*, Wiley Series in Probability and Statistics. Probability and Statistics.
- [Jennison-Turnbull, (1999)] Jennison, C. & Turnbull, B. W., (1999), *Group Sequential Methods with Applications to Clinical Trials*, Ed. Chaptam & Hall/CRC, New York.
- [Kiefer-Weiss, (1957)] Kiefer, J. & Weiss, L., (1957), *Some Properties of Generalized Sequential Probability Ratio Test*, Annals of Mathematical Statistics, 28: 57-75.
- [Lehmann, (1959)] Lehmann, E. L., (1959), *Testing Statistical Hypotheses*, Jonh Wiley & Sons, New York; Chapman & Hall, London.
- [Lorden, (1980)] Lorden, G., (1980), *Structure of sequential tests minimizing an expected sample size*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Geviete, 51: 291-302.
- [Mukhopadhyay-De Silva, (2008)] Mukhopadhyay, N. & De Silva, B. M., (2008), *Theory and applications of a new methodology for the random sequential probability ratio test*, Statistical Methodology, 5: 424-453.
- [Novikov, (2007)] Novikov, A., (2007), *Una Introducción elemental al análisis estadístico secuencial*, Notas de curso. Primer coloquio del departamento de Matemáticas UAM-I, la Trinidad, Tlaxcala, México.
- [Novikov, (2008)] Novikov, A., (2008), *Optimal sequential tests for two simple hypothesis based on independent observations*, International Journal of Pure and Applied Mathematics, 25: 291-314.
- [Novikov, (2009)] Novikov, A., (2009), *Optimal sequential tests for two simple hypothesis*, Sequential Analysis, 28: 188-217.
- [Novikov, *et al*, (2022)] Novikov, A., Novikov, A. & Farkhshatov, F. (2022), *A computational approach to the Kiefer-Weiss problem for sampling from a Bernoulli population*, Sequential Analysis, 41:2, 198-219.
- [Novikov-Popoca, (2022)] Novikov, A. & Popoca-Jiménez, X. I., (2022), *Optimal group sequential tests with groups of random size*, Sequential Analysis, 41:2, 220-240.
- [Popoca, (2012)] Popoca-Jiménez, X. I., (2012), *Optimalidad de pruebas de hipótesis secuenciales con grupos de tamaño aleatorio*, Tesis de Maestría, Universidad Autónoma Metropolitana-Iztapalapa (UAM-I), México.
- [Rockafellar, (1972)] Rockafellar, R. Tyrrell, (1972), *Convex Analysis*, Second Printing, Princeton University Press.
- [Schmitz, (1993)] Schmitz, N., (1993), *Optimal sequentially planned decision procedures*, Lecture Notes in Statistics 79, New York: Springer-Verlag, 1993.

-
- [Shiryaev, (2008)] Shiryaev, A. N., (2008), *Optimal Stopping Rules*, Springer-Verlag: Springer series in statistics, New York, (1978), Reimpr.
- [Siegmund, (1985)] Siegmund, D., (1985), *Sequential Analysis: tests and confidence intervals*, Springer-Verlag: Springer Series in Statistics, New York.
- [Wald, (1945)] Wald, A., (1945), *Sequential test of statistical hypotheses*, Annals of Mathematical Statistics 16: 117-186.
- [Wald, (1947)] Wald, A., (1947), *Sequential Analysis*, Wiley, New York.
- [Wald, (1971)] Wald, A., (1971), *Statistical Decision Functions*, 2nd ed., Chelsea Publishing Company, Bronx, New York.
- [Wald-Wolfowitz, (1948)] Wald, A. & Wolfowitz, J., (1948), *Optimum character of the sequential probability ratio test*, Annals of Mathematical Statistics 19: 326-339.
- [Weiss, (1953)] Weiss, L., (1953), *Testing one simple hypothesis against another*, Ann. Math. Stat. 24: 273-281.
- [Weiss, (1962)] Weiss, L., (1962), *On sequential test which minimize the maximum expected sample size*, Journal of American Statistical Association 57: 551-566.

BIBLIOGRAFÍA



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE DISERTACIÓN PÚBLICA

No. 00081

Matrícula: 2123802766

PRUEBAS DE HIPÓTESIS
SECUENCIALES ÓPTIMAS CON
GRUPOS DE TAMAÑO ALEATORIO.

En la Ciudad de México, se presentaron a las 15:00 horas del día 1 del mes de febrero del año 2023 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

- DR. JOSE ANDRES CHRISTEN GRACIA
- DR. JOSE RAUL MONTES DE OCA MACHORRO
- DR. ANDREY NOVIKOV
- DR. HUGO ADAN CRUZ SUAREZ
- DRA. BLANCA ROSA PEREZ SALVADOR




XOCHITL ITXEL POPOCA JIMENEZ
ALUMNA

Bajo la Presidencia del primero y con carácter de Secretaria la última, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

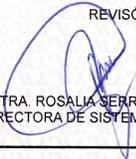
DOCTORA EN CIENCIAS (MATEMATICAS)
DE: XOCHITL ITXEL POPOCA JIMENEZ

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

Aprobar

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

REVISÓ



MTRA. ROSALIA SERRANO DE LA PAZ
DIRECTORA DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI



DR. ROMAN LINARES ROMERO

PRESIDENTE



DR. JOSE ANDRÉS CHRISTEN GRACIA

VOCAL



DR. JOSE RAUL MONTES DE OCA MACHORRO

VOCAL



DR. ANDREY NOVIKOV

VOCAL



DR. HUGO ADAN CRUZ SUAREZ

SECRETARIA



DRA. BLANCA ROSA PEREZ SALVADOR