



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Iztapalapa

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

***“PRUEBAS DE HIPÓTESIS SECUENCIALES CON ETAPAS DE
DURACIÓN FIJA: DISEÑO Y OPTIMIZACIÓN”***

Tesis que presenta

PEDRO REYES PÉREZ

Matrícula: 205181390

Para obtener el grado de
Doctor en Ciencias (Matemáticas)

Director: DOCTOR ANDREY NOVIKOV

JURADO:

Doctor: José Raúl Montes de Oca Machorro

Doctor: Julio César García Corte

Doctor: Luis Antonio Rincón Solís

Doctor: José Andrés Christen Gracia

Doctor: Gabriel Núñez Antonio

Iztapalapa, Ciudad de México, octubre de 2020.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE DISERTACIÓN PÚBLICA

No. 00071

Matrícula: 205181390

Pruebas de hipótesis
secuenciales con etapas de
duración fijas: diseño y
optimización.



Con base en la Legislación de la Universidad Autónoma Metropolitana, en la Ciudad de México se presentaron a las 16:00 horas del día 14 del mes de octubre del año 2020 POR VÍA REMOTA ELECTRÓNICA, los suscritos miembros del jurado designado por la Comisión del Posgrado::

- DR. JOSE RAUL MONTES DE OCA MACHORRO
- DR. JULIO CESAR GARCIA CORTE
- DR. LUIS ANTONIO RINCON SOLIS
- DR. JOSE ANDRES CHRISTEN GRACIA
- DR. GABRIEL NUÑEZ ANTONIO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

DOCTOR EN CIENCIAS (MATEMATICAS)

DE: PEDRO REYES PEREZ

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 al interesado 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. JESUS ALBERTO OCHOA TAPIA

PRESIDENTE

DR. JOSE RAUL MONTES DE OCA
MACHORRO

VOCAL

DR. JULIO CESAR GARCIA CORTE

VOCAL

DR. LUIS ANTONIO RINCON SOLIS

VOCAL

DR. JOSE ANDRES CHRISTEN GRACIA

SECRETARIO

DR. GABRIEL NUÑEZ ANTONIO

El presente documento cuenta con la firma –autógrafa, escaneada o digital, según corresponda- del funcionario universitario competente, que certifica que las firmas que aparecen en esta acta – Temporal, digital o dictamen- son auténticas y las mismas que usan los c.c. profesores mencionados en ella



Agradecimientos

Agradezco a la **Universidad Autónoma Metropolitana**, que abre las puertas de manera incondicional a todos aquellos(as) que luchamos por lograr nuestros sueños y metas.

Agradezco, al Consejo Nacional de Ciencia y Tecnología (CONACyT), por la beca otorgada para mis estudios de doctorado.

Agradezco muy ampliamente al Dr. Andrey Novikov por haberme compartido sus conocimientos, por sus asesorías y enseñanzas, por su tiempo y profesionalismo, así como por su paciencia y apoyo incondicional para llevar acabo este trabajo.

Agradezco a todos y cada uno de los sinodales por sus observaciones, comentarios y sugerencias en la revisión de la tesis, con lo cual lograron que el presente trabajo fuera más claro y conciso.

- Dr. José Raúl Montes de Oca Machorro.
- Dr. José Andrés Christen Gracia.
- Dr. Gabriel Núñez Antonio.
- Dr. Julio César García Corte.
- Dr. Luis Antonio Rincón Solís.

A cada uno de ustedes les quedo muy agradecido.

Agradezco a cada uno de mis amigos, con los cuales compartí momentos de frustración, de alegrías, de sueños, etc., y que finalmente eso es lo que se lleva muy dentro de sí.

Agradezco a mi familia que son el motor de motivación para seguir en busca de nuevas metas.

A todos muchas gracias.



Índice general

Antecedentes	3
Introducción	7
1. Pruebas de hipótesis estadísticas y análisis secuencial	9
1.1. Conceptos probabilísticos	9
1.2. Pruebas de hipótesis en el enfoque clásico	10
1.3. Análisis estadístico secuencial	12
1.4. Procesos estocásticos secuenciales	13
1.5. Planteamiento de problemas	14
1.6. Reglas generales de paro y de decisión	15
1.7. Planteamiento general del error tipo I , tipo II y costo	16
2. Teoría secuencial multietápica	19
2.1. La regla de decisión óptima	21
2.2. Paro óptimo	22
2.3. Minimización del costo	26
3. Planteamiento general multietápico del proceso de Wiener con deriva lineal	27
3.1. El proceso de Wiener con deriva lineal en etapas	28
3.2. Fórmulas multietápicas del proceso de Wiener con deriva lineal	31
3.3. El error tipo I	36
3.4. El error tipo II	37
3.5. Costo promedio bajo la hipótesis nula y la alternativa	37
3.6. Costo promedio bajo la hipótesis nula	37
3.7. Resultados de la evaluación del proceso de Wiener con deriva lineal	38
3.8. Análisis de los resultados numéricos	44
4. El problema modificado de Kiefer-Weiss con distribuciones uniformes	47
4.1. Decisión óptima para el problema modificado de Kiefer-Weiss	48
4.2. Determinación del valor del error tipo II	50
4.3. Paro óptimo cuando $0 \leq \theta \leq \theta_1$	50
4.3.1. Error tipo I y número promedio muestral	53
4.4. Paro óptimo cuando $\theta_1 < \theta < 1$	55
4.4.1. Error tipo I y número promedio muestral	59
4.5. Paro óptimo para $1 < \theta$	60

ÍNDICE GENERAL

4.6. Error tipo <i>I</i> y número promedio muestral	62
4.7. Resultados numéricos para el caso $\theta = \theta_0$	63
4.8. Errores y número promedio muestral con distribuciones uniformes usando el criterio de Wald	64
4.9. La razón de verosimilitud con distribuciones uniformes	65
4.10. El error tipo <i>I</i> y error tipo <i>II</i>	65
4.11. El costo promedio	66
5. Conclusiones sobre el trabajo desarrollado en la tesis	67
6. Apéndices	69
Apéndice A	71
6.1. Tablas de resultados numéricos en tres etapas	71
Apéndice B	73
6.2. Tablas de resultados numéricos en cuatro etapas	73
Apéndice C	75
6.3. Evaluación numérica del proceso de Wiener con deriva lineal en una y dos etapas	75
6.4. Minimización de la función de Lagrange	76
6.5. El error tipo <i>I</i>	77
6.6. El error tipo <i>II</i>	77
6.7. Número promedio muestral bajo las dos hipótesis	78
6.8. Número promedio muestral en caso asimétrico	79
6.9. Programa principal	79
6.10. Presentación de resultados	80
Apéndice D	81
6.11. Código para la evaluación numérica del proceso de Wiener con deriva lineal en tres etapas	81
6.12. Funciones para tres etapas	81
6.13. Funciones para dos etapas	82
6.14. Cálculo de la integral numérica	82
6.15. Funciones para una etapa	83
6.16. Minimización de la función de Lagrange	83
6.17. Programa principal	84
6.18. Error tipo <i>I</i> , tipo <i>II</i> y número promedio muestral	85
6.19. El error tipo <i>I</i>	86
6.20. El error tipo <i>II</i>	86
6.21. Número promedio muestral en el caso simétrico	87
6.22. Número promedio muestral en el caso asimétrico	88
Apéndice E	91
6.23. Código para la evaluación numérica del problema de Kiefer-Weiss modificado	91

Bibliografía

93

Antecedentes

El trabajo es enmarcado en el área del análisis estadístico secuencial, específicamente en pruebas de hipótesis secuenciales. El origen del análisis estadístico secuencial viene de los trabajos fundamentales de Abraham Wald (ver, por ejemplo, [Wald (1945)] and [Wald (1950)]). La idea de este enfoque es permitir para cualquier procedimiento estadístico, un muestreo secuencial, en el cual la muestra se forma agregando observaciones una por una, empezando desde una sola observación. Finalmente el número de observaciones se determina en curso del muestreo mismo, dependiendo de las condiciones en las que se generan los datos que se presentan durante el muestreo.

El enfoque original de Wald (ver, por ejemplo, [Wald (1950)]) supone que las observaciones se toman una por una. El conjunto de los resultados obtenidos en cada etapa se analiza con el objetivo de decidir si es necesario seguir con el muestreo y proceder a la siguiente etapa, o dar por terminado el experimento estadístico. Formalmente esta etapa de muestreo es controlada por una regla llamada *la regla de paro*. En la observación en que el experimento termina, se procede a la toma de decisión final: en esta fase se analizan los datos obtenidos en todas las etapas del experimento y se toma la decisión final, si se rechaza o no la hipótesis nula, para hacer esto se requiere de *la regla de decisión*.

El resultado más famoso del análisis secuencial es la optimalidad de la llamada prueba secuencial de razón de probabilidades (sequential probability ratio test, SPRT), en el problema de pruebas de dos hipótesis simples cuando las observaciones son independientes e idénticamente distribuidas (i.i.d.). La SPRT, en promedio, requiere de una tercera a una cuarta parte del número de las observaciones, en comparación con el método estadístico clásico (basado en una sola muestra), que son necesarias para lograr la misma calidad de la inferencia (probabilidades de error tipo I y tipo II).

Uno de los problemas de la estadística secuencial es que no se utiliza en la práctica, porque se requiere de mayor trabajo al realizar las n pruebas (una a una), que realizar sólo una en la que están contenidas todas las n observaciones [Schmitz (1993)]. Si se considera que cada una de las observaciones tiene un costo, entonces es de interés considerar el número promedio de observaciones ya sea con respecto a la hipótesis nula o la hipótesis alternativa. Esto es crucial para la optimalidad de la SPRT (ver [Cressie and Morgan(1993)] y [Schmitz (1993)]). Otro problema es que, aunque el *promedio* de observaciones que usa la SPRT es mínimo, este número teóricamente no queda acotado (ver prefacio [Schmitz (1993)]), y sólo se resuelve la cuestión del número de observaciones a la hora de realizar el experimento.

Parece no ser casualidad que a unos cuantos años del descubrimiento de Wald, empezaron intentos de “truncar” la prueba con tal de tener un experimento con un máximo de observaciones que no sea mayor a un número dado. El enfoque de pruebas secuencialmente planeadas de [Cressie and Morgan(1993)] y [Schmitz (1993)] venía dirigido esencialmente contra el primero de los dos problemas arriba mencionados. El trabajo [Cressie and Morgan(1993)] y algunos otros relacionados trata el caso de dos hipótesis simples para observaciones i.i.d.

En las notas de [Schmitz (1993)] se desarrolla una teoría de procedimientos secuencialmente planeados y se hace uso de ella para tratar los problemas de pruebas de hipótesis también para el caso de observaciones i.i.d., en un marco un poco más general que en [Cressie and Morgan(1993)], quedando abierta la posibilidad de tratar los casos de observaciones de una estructura más general que el caso i.i.d. Sin embargo, las pruebas desarrolladas en [Cressie and Morgan(1993)] y [Schmitz (1993)] y otros artículos relacionados no combaten el segundo de los “males” de los problemas secuenciales: la imposibilidad, generalmente hablando, de dar una cota para el número máximo de observaciones en el experimento, ya que el tamaño de la muestra que se toma en cada etapa varía en función de las observaciones anteriores y, aparte, el número de las etapas generalmente hablando queda indefinido.

En [Novikov (2004)] se propone una simple idea de cómo combinar las ventajas del análisis secuencial, con la ventaja de tener el control sobre el número total de observaciones que se tomarán en el experimento: usar diseño de dos etapas, cada una de una duración fija y definida antes de iniciar el experimento. La duración de las etapas se considera como parámetro (variable) del procedimiento que permite una optimización adicional, con tal de minimizar el número promedio muestral del experimento completo (de dos etapas). Este tipo de experimentos bietápicos se menciona para fines ilustrativos en [Ghosh et al. (1997)], sin darle más desarrollo a los mismos. Una ventaja de los diseños como en [Novikov (2004)] es que en ellos se pueden aplicar, sin complicaciones mayores a los procesos estocásticos mucho más generales que el caso i.i.d., incluso a procesos a tiempo continuo, lo que quedó ilustrado en [Novikov (2004)].

Los resultados numéricos obtenidos en [Novikov (2004)] dan evidencia que los diseños bietápicos tienen una eficiencia inesperadamente alta, y pueden competir aún contra los procedimientos óptimos, en los casos cuando éstos pueden ser calculados explícitamente. En casos de procesos con dependencia entre observaciones la estructura de los procedimientos óptimos se complica considerablemente y es difícil creer que en casos más generales que i.i.d. alguna vez se podrán encontrar explícitamente las pruebas óptimas, como es esto en el caso de la SPRT. Dada la eficiencia de la estadística secuencial en los diseños de dos etapas, ésta deja abiertas posibilidades para el cálculo de las pruebas óptimas: algunos ejemplos están dados en [Novikov (2004)] y [Novikov (2006)].

En [Novikov (2004)] se propone una simple idea de cómo combinar las ventajas del análisis secuencial, con la ventaja de tener el control sobre el número total de observaciones que se tomarán en el experimento: usar diseño de dos etapas, cada una de una

duración fija y definida antes de iniciar el experimento. La duración de las etapas se considera como parámetro (variable) del procedimiento que permite una optimización adicional, con tal de minimizar el número promedio muestral del experimento completo (de dos etapas). Este tipo de experimentos bietápicos se menciona para fines ilustrativos en [Ghosh et al. (1997)], sin darle más desarrollo a los mismos. Una ventaja de los diseños como en [Novikov (2004)] es que en ellos se pueden aplicar, sin complicaciones mayores a los procesos estocásticos mucho más generales que el caso i.i.d., incluso a procesos a tiempo continuo, lo que quedó ilustrado en [Novikov (2004)].

Los resultados numéricos obtenidos en [Novikov (2004)] dan evidencia que los diseños bietápicos tienen una eficiencia inesperadamente alta, y pueden competir aún contra los procedimientos óptimos, en los casos cuando éstos pueden ser calculados explícitamente. En casos de procesos con dependencia entre observaciones la estructura de los procedimientos óptimos se complica considerablemente y es difícil creer que en casos más generales que i.i.d. alguna vez se podrán encontrar explícitamente las pruebas óptimas, como es esto en el caso de la SPRT. Dada la eficiencia de la estadística secuencial en los diseños de dos etapas, ésta deja abiertas posibilidades para el cálculo de las pruebas óptimas: algunos ejemplos están dados en [Novikov (2004)] y [Novikov (2006)].

Introducción

El objetivo básico del presente trabajo es minimizar el tamaño de la muestra al realizar una prueba de hipótesis secuencialmente en k etapas, manteniendo la calidad de la inferencia igual al de la estadística clásica.

En el capítulo 1, se plantean los conceptos básicos sobre la estadística secuencial como son: plan de muestreo, etapas, los tamaños de las etapas, el modelo estadístico secuencial para observaciones i.i.d. También se dan las definiciones secuenciales de un proceso estocástico tanto para caso discreto como continuo. Se plantean los problemas que se van a resolver en la tesis. En forma general se presenta el error tipo I , el error tipo II y suponiendo que cada observación tiene un costo, entonces se define el costo por el uso de las observaciones. Debido a que los errores dependen del tiempo de paro y de la regla de decisión, estos se definen como $\alpha(\psi, \phi)$, $\beta(\psi, \phi)$ y el costo del experimento como $k(\psi, \phi)$.

En el capítulo 2, se construye la teoría secuencial para k etapas. Por medio del método de los multiplicadores de Lagrange, se crea una función a la cual se le llama *la función de Lagrange* en la cuál participan el error tipo I , el error tipo II y el costo de las observaciones. La función de Lagrange es la que se optimiza, considerando restricciones sobre los errores $\alpha(\psi, \phi)$, $\beta(\psi, \phi)$. Se concluye este capítulo con la optimización de la regla de paro, la regla de decisión y las duraciones óptimas de las etapas.

En el capítulo 3, partiendo de una muestra aleatoria con observaciones i.i.d. Se aplica la teoría desarrollada al proceso de Wiener con deriva lineal. Se obtiene la estructura en general y en forma particular de tal manera que se dan las fórmulas para la evaluación numérica del proceso, obteniendo explícitamente la forma para calcular los errores tipo I y tipo II . Con respecto al costo, se obtienen las fórmulas para el costo bajo de ambas hipótesis (H_0 y H_1) y bajo alguna de las dos. Con estas fórmulas se realiza una evaluación numérica y se presenta los resultados tanto en tablas como en gráficas de contorno. Todo este trabajo se hizo para 2, 3 y 4 etapas. En el caso de dos etapas, los resultados numéricos se pueden comparar con los que presentó A. Novikov en [Novikov (2004)].

En el capítulo 4, se resuelve el problema modificado de Kiefer-Weiss, en el cual participa una tercera hipótesis diferente de H_0 y H_1 . Al resolver este problema surgen diferentes casos y en cada uno de ellos se dan las fórmulas explícitas para calcular los errores tipo I y tipo II , además del número promedio muestral (ASN). Un caso particular es cuando la tercera hipótesis coincide con H_0 o H_1 , lo cual se reduce a la

Introducción

prueba de H_0 vs H_1 . Para este caso se realizan simulaciones por medio de MatLab y se presentan los resultados numéricos de los errores tipo I y tipo II además del número promedio de observaciones. Los cálculos se realizan para dos, tres y cuatro etapas.

Otro punto importante que se considera es que, ya con la teoría desarrollada, se hace la formulación para el cálculo de los errores y número promedio muestral con distribuciones uniformes utilizando el criterio de Wald. Como ya se mencionó, estos resultados son consecuencia del presente trabajo por lo cual no se dan referencias bibliográficas.

En el apéndice 6 se presentan los resultados numéricos del proceso de Wiener con deriva lineal para 3 etapas.

En el apéndice 6.1 están los resultados numéricos del proceso de Wiener con deriva lineal para 4 etapas.

En los apéndices 6.2 y 6.10 se presenta el código para la evaluación numérica del proceso de Wiener con deriva lineal en dos y tres etapas.

En el apéndice 6.22 está el código para la evaluación numérica para el problema modificado de Kiefer-Weiss, en diferentes etapas.

Capítulo 1

Pruebas de hipótesis estadísticas y análisis secuencial

Una de las bases importantes en la estadística inferencial, son las pruebas de hipótesis, en estas es de gran relevancia considerar el error tipo I , el error tipo II y el tamaño de la muestra. Con respecto a estos puntos, el presente trabajo está motivado por la eficiencia que se obtuvo al realizar pruebas de dos hipótesis simples tomando las observaciones en dos etapas, esto se puede ver en [Novikov (2004)].

Antes de iniciar con el trabajo secuencial, se hará un breve repaso sobre los conceptos básicos de los modelos probabilísticos, y las pruebas de hipótesis de desde el punto de vista clásico.

1.1. Conceptos probabilísticos

En este apartado se da un breve repaso de los modelos probabilísticos. Antes de esto, se da el siguiente resultado de teoría de la medida que es necesario para fundamentar debidamente lo que se propone.

Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Se denota como $M^+(\Omega, \mathcal{F})$ a la colección de todas las funciones no negativas \mathcal{F} -medibles de Ω a $\bar{\mathbb{R}}$.

Teorema 1.1. *Sean λ y μ dos medidas σ -finitas definidas sobre \mathcal{F} y supóngase que λ es absolutamente continua con respecto a μ . Entonces existe una función f en $M^+(\Omega, \mathcal{F})$, tal que*

$$\lambda(E) = \int_E f d\mu, \quad E \in \mathcal{F}. \quad (1.1)$$

Además, la función f es determinada de manera única μ -casi donde quiera.

La función f que participa en 1.1, se le llama *la derivada de Radon-Nikodym* de λ con respecto a μ y es denotada por $d\lambda/d\mu$.

La suposición general que se hace sobre el modelo estadístico es la siguiente.

1. Pruebas de hipótesis estadísticas y análisis secuencial

Sea (Ω, \mathcal{F}) un espacio medible, sea X_1, X_2, \dots , una sucesión de variables aleatorias independientes e idénticamente distribuidas (v.a. i.i.d.) en este espacio. Supóngase que P es la distribución de las variables aleatorias y que existe la densidad (la derivada de Radon-Nikodym)

$$f(x) = \frac{dP}{d\mu}(x) \quad (1.2)$$

con respecto a una medida σ -finita μ .

Debido a que las variables aleatorias son independientes, para cualquier $n \geq 1$, la función de densidad conjunta del vector aleatorio (X_1, \dots, X_n) , es

$$f^n = f^n(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) \quad (1.3)$$

con respecto a la medida

$$\mu^n = \mu \times \dots \times \mu.$$

Sea Ω el espacio muestral y \mathcal{F}_n con $n \geq 1$ es la σ -álgebra que contiene a (X_1, \dots, X_n) . En el caso en que $n = 0$, es decir, cuando no se tienen observaciones, la sigma álgebra es $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

1.2. Pruebas de hipótesis en el enfoque clásico

En esta parte, se enuncian los conceptos básicos para realizar pruebas de hipótesis con el enfoque de la estadística clásica.

Definición 1.1. *Una hipótesis estadística o simplemente hipótesis es una afirmación o conjetura que se hace acerca de la distribución de una variable aleatoria X .*

La hipótesis nula se denota como H_0 , y las hipótesis alternativas como H_i con $i = 1, \dots, k$.

Definición 1.2. *La hipótesis estadística es simple si esta específica de forma única a la distribución, en otro caso se dice que es compuesta.*

Definición 1.3. *Una prueba de hipótesis es una regla que indica si se rechaza o no se rechaza la hipótesis nula.*

Al tomar la decisión de rechazar o no rechazar la hipótesis nula, se puede cometer dos tipos de errores.

Definición 1.4. *Rechazar la hipótesis nula cuando ésta es verdadera se le conoce como el error de tipo I y a la probabilidad de cometer este error se le conoce como α , que es el tamaño de la región crítica. A esta probabilidad también se le conoce como el nivel de significancia.*

Definición 1.5. *No rechazar H_0 siendo que ésta es falsa se le conoce como el error de tipo II y la probabilidad de cometer este error se denota como β .*

1.2. Pruebas de hipótesis en el enfoque clásico

	H_0 cierta	H_1 cierta
Rechazar H_0	Error tipo I , (α)	Decisión correcta
No rechazar H_0	Decisión correcta	Error tipo II , (β)

Tabla 1.1. Tipos de error

Estas definiciones se pueden ver en la tabla 1.1. Para tomar la decisión sobre si se rechaza o no la hipótesis nula, se utiliza la información de la muestra aleatoria X_1, \dots, X_n , la cual se supone sigue cierta distribución. Es claro que si se cambia de muestra, la decisión de rechazar o no la hipótesis nula también podría cambiar.

Sea Ω el espacio muestral y sea C un subconjunto de Ω en el cual se rechaza H_0 , a este subconjunto se le llama “región crítica” o “región de rechazo”, la región de no rechazo de la hipótesis nula es $\Omega - C$.

Definición 1.6. La región crítica de una prueba de hipótesis, es un subconjunto C del espacio muestral Ω que corresponde al rechazo de la hipótesis nula.

Es claro que se espera escoger una prueba estadística y una región crítica de manera que las probabilidades de error tipo I y error tipo II , sean lo más pequeñas posibles. Notar que en una hipótesis simple H_0 , el nivel de significancia es también el tamaño de la prueba. Considerando las probabilidades de error al realizar una prueba de hipótesis, es muy conveniente tener la definición de la *función de potencia*

Definición 1.7. La función de potencia, $\pi(\theta)$ en la prueba de hipótesis de H_0 es la probabilidad de rechazo de H_0 , cuando el verdadero valor del parámetro es θ .

Supóngase que se tienen dos o más pruebas del mismo tamaño y el interés está en determinar aquella prueba que tenga las probabilidades de error más pequeñas.

Supóngase que las v.a. X_1, \dots, X_n tienen función de densidad de probabilidad $f(x_1, \dots, x_n)$ y considerar su región crítica C .

La notación de la función de potencia correspondiente a C es

$$\pi_C(\theta) = P[(X_1, \dots, X_n) \in C | \theta] \tag{1.4}$$

Definición 1.8. La prueba de hipótesis de $H_0 : \theta = \theta_0$ vs $H_0 : \theta = \theta_1$ basada sobre la región crítica C^* se dice que es una prueba más potente de tamaño α , si

1. $\pi_{C^*}(\theta_0) = \alpha$, y
2. $\pi_{C^*}(\theta_1) \geq \pi_C(\theta_1)$ para cualquier otra región crítica C de tamaño α , (es decir, $\pi_C(\theta_0) = \alpha$).

A la región crítica C^* se le llama *la región crítica más potente* de tamaño α .

El teorema de Neyman-Pearson es uno de los resultados más importantes en el que se sustenta la estadística inferencial clásica y que a continuación se enuncia.

1. Pruebas de hipótesis estadísticas y análisis secuencial

Teorema 1.2. (Lema de Neyman-Pearson). Sea $\mathbf{X} = X_1, X_2, \dots, X_n$, una muestra aleatoria de tamaño n de una función de densidad o de probabilidad, correspondiente a θ_i es $f(\mathbf{x}|\theta_i)$ con $i = 0, 1$. Para realizar la prueba de $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$, para la prueba sea usa la región de rechazo C que satisfice

$$f(\mathbf{x}|\theta_1) > kf(\mathbf{x}|\theta_0), \text{ si } x \in C \quad (1.5)$$

y

$$f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0), \text{ si } x \in C^c \quad (1.6)$$

para algún $k > 0$, y

$$\alpha = P_{\theta_0}(\mathbf{X} \in C). \quad (1.7)$$

Entonces

1. (Suficiencia) cualquier prueba que satisfice a (1.5) - (1.7) es una prueba UMP de nivel α .
2. (Necesario) Si existe una prueba que satisfice a (1.5) - (1.7) con $k > 0$, entonces cada prueba UMP de nivel α es una prueba de tamaño α (que satisfice (1.7)) y cualquier prueba UMP de nivel α satisfice a (1.5) y (1.6) exepto en un conjunto A tal que $P_{\theta_0}(X \in A) = P_{\theta_1}(X \in A) = 0$.

Con este teorema (Lema de Neyman-Pearson), concluye el breve repaso de los conceptos de la estadística clásica.

1.3. Análisis estadístico secuencial

En el análisis secuencial multietápico, en el que se trabaja, el total de las observaciones se dividen en k grupos cada uno tamaño fijo. A la toma secuencial de cada grupo de observaciones se le llama etapa, de esta manera se dice que los datos son tomados en k etapas con $0 < k < \infty$. Siguiendo el procedimiento de análisis secuencial como el que planteó Wald (ver [Wald (1947)]), se procede de la siguiente manera: Se da una regla para realizar una de las siguientes tres decisiones en cada etapa del experimento. 1) aceptar la hipótesis nula, 2) rechazar la hipótesis nula 3) continuar con el experimento tomando un grupo más de observaciones. Entonces, así descrito esto es una prueba de hipótesis secuencial. Sobre la base de la información del primer grupo de observaciones se pueden tomar cualquiera de las tres decisiones mencionadas. Si la primera o la segunda se toman, entonces el experimento termina. Si se toma la tercer decisión, otra vez, en base a las observaciones de los dos grupos de observaciones se toma una de las tres decisiones ya mencionadas. Si se opta por la tercer decisión, entonces se toman las observaciones del tercer grupo y así se continua. Este proceso continua hasta que se toma la primera o segunda decisión. El número n de grupos de observaciones que se requiere para un procedimiento de este tipo es una variable aleatoria, el valor de n depende de los resultados de las observaciones.

Es importante aclarar que el procedimiento descrito puede ser que nunca tome la tercer decisión y por lo tanto, el procedimiento no termine. En el caso que nos interesa se consideran k grupos de datos, y el número de observaciones de cada grupo es fijo.

De esta manera, el experimento secuencial descrito puede detenerse en cualquier etapa n con $1 \leq n \leq k$.

Sea X_t , $t \in T$, un proceso estocástico, con $T = \mathbb{N}$ (tiempo discreto), o $T = [0, \infty)$ (tiempo continuo). Supongamos que la distribución del proceso depende de algún “parámetro” desconocido θ y se quiere realizar el contraste de la hipótesis simple $H_0 : \theta = \theta_0$ contra una alternativa simple $H_1 : \theta = \theta_1$, con $\theta_1 \neq \theta_0$. El objetivo de este trabajo es mostrar el procedimiento para pruebas de hipótesis secuenciales con un máximo finito de k etapas, $k \geq 2$.

Se supondrá que las duraciones de las etapas son números fijos positivos t_1, \dots, t_k , dados antes de que el experimento empiece. A diferencia [Cressie and Morgan(1993)] y [Schmitz and Sueselbeck (1983)], se supondrá que cada uno de los t_1, \dots, t_k , son números naturales fijos y que son establecidos con anticipación. En el caso de tiempo discreto $t_i \in \mathbb{N}$, en tiempo continuo $t_i \in (0, \infty)$. El vector $t = (t_1, t_2, \dots, t_k)$ es llamado *plan de muestreo*.

En forma general, para un plan de muestreo t y para algún $n \leq k$ sea

$$X^{(t|n)} = (X_1, \dots, X_{t_1}, X_{t_1+1}, \dots, X_{t_1+t_2}, \dots, X_{t_1+t_2+\dots+t_n}).$$

El caso particular es cuando $n = k$, y entonces por definición $X^{(t)} = X^{(t|k)}$. Notemos que si el vector t es unidimensional, es decir $t = N$, entonces $X^{(N)} = (X_1, \dots, X_N)$.

Para cualquier $t \in \mathcal{P}^k$, el vector de datos $X^{(t|k)} = (X_1^{(t)}, X_2^{(t)}, \dots, X_k^{(t)})$, tiene que la función de densidad

$$f_{\theta}^{(t|k)}(x^{(t|k)}) = f_{\theta}^{(t|k)}(x_1^{(t)}, x_2^{(t)}, \dots, x_k^{(t)}) \tag{1.8}$$

es la derivada de Radon-Nikodym con respecto a la medida producto

$$\mu^{(t|k)} = \mu_{t_1} \otimes \mu_{t_2} \otimes \dots \otimes \mu_{t_k}$$

siendo μ_{t_1} medidas σ -finitas sobre los respectivos espacios.

La afirmación anterior es debido a que

$$f_{\theta}^{(t|k-1)}(x^{(t|k-1)}) = \int_{x_k^{(t)}} f_{\theta}^{(t|k)}(x_1^{(t)}, x_2^{(t)}, \dots, x_k^{(t)}) d\mu^{(t|k)}$$

si $E = x_k^{(t)}$, $f = f_{\theta}^{(t|k)}$ y $\mu = d\mu^{(t|k)}$, se tienen todas las condiciones del Teorema 1.1.

Se aclara que no se supone alguna estructura paramétrica de las distribuciones participantes, por lo que “ θ ” puede ser lo que sea; el único requisito es que identifique de forma única la distribución.

1.4. Procesos estocásticos secuenciales

Supóngase que las observaciones son generadas por un proceso estocástico y que son usadas en un procedimiento estadístico secuencial, con el objetivo de observar la diferencia entre dos hipótesis simples acerca de la distribución del proceso. A lo largo de la tesis, se supone que las observaciones son tomadas en un número finito de k etapas,

(siendo $k \geq 2$, porque si $k = 1$ se tiene una prueba de hipótesis de la manera clásica). En particular, el caso discreto queda cubierto cuando los datos son tomados en una etapa de duración unitaria, y cada valor del proceso X_1, X_2, \dots, X_k puede potencialmente ser observada. La prueba secuencial correspondiente a este esquema son usualmente llamadas en la literatura como pruebas “truncadas” o de “horizonte finito”.

En el caso de tiempo continuo, supóngase que las etapas respectivas están formadas por “partes” de un proceso $\{X_s, s_{i-1} < s \leq s_i\}$, $i = 1, \dots, k$, con duraciones de etapas positivas $t_i = s_i - s_{i-1}$, $i = 1, \dots, k$. Alternativamente, se puede observar una versión discretizada de un proceso de tiempo continuo, cuando las observaciones disponibles para el análisis estadístico son $X_{s_1}, X_{s_2}, \dots, X_{s_k}$, con $s_1 < s_2 < \dots < s_k$.

Sea X_t , $t \in T$, un proceso estocástico, con $T = \mathbb{N}$ (tiempo discreto), o $T = [0, \infty)$ (tiempo continuo). Supongamos que la distribución del proceso depende de algún “parámetro” desconocido θ y se quiere realizar la prueba de hipótesis simple $H_0 : \theta = \theta_0$ contra una alternativa simple $H_1 : \theta = \theta_1$, con $\theta_1 \neq \theta_0$.

Para cualquier plan de muestreo t , se define la suma del tamaño de las etapas, como

$$s_n(t) = t_1 + \dots + t_n, \quad n = 1, 2, \dots, k \quad (1.9)$$

siendo $s_0(t) = 0$.

Se denotará como

$$X_n^{(t)} = \{X_s\}_{s_{n-1}(t) < s \leq s_n(t)}$$

a los datos observados en la etapa n , y

$$X^{(t;n)} = (X_1^{(t)}, X_2^{(t)}, \dots, X_n^{(t)})$$

todos los datos observados hasta la etapa n , con $n = 1, 2, \dots, k$.

1.5. Planteamiento de problemas

A continuación se enuncian, en forma general, los problemas que se resuelven en la tesis.

Considerando un plan de muestreo t con $t_i > 0$ para todo $i = 2, 3, \dots, k$ fijo y determinado antes que empiece el experimento, el objetivo de este trabajo es mostrar el procedimiento para pruebas de hipótesis secuenciales con un máximo finito de k etapas, $k \geq 2$.

Los elementos que participan en el proceso secuencial multietápico son la regla de paro y la regla de decisión, es claro que estas deben ser determinadas en su forma más general.

Si además se considera que las observaciones tienen un costo, entonces, es de gran interés considerar la minimización del costo promedio de observaciones.

Para mostrar la eficiencia de la estadística secuencial multietápica se realizan evaluaciones de la teoría desarrollada en casos particulares, uno de ellos es para el proceso de Wiener con deriva lineal y el otro es para un proceso con distribuciones uniformes. En cada caso se presentarán las fórmulas para obtener los valores de los errores, los valores promedios de los costos y, además, se presentan los resultados numéricos.

Otro punto de interés es fijar un valor θ diferente a θ_0 y θ_1 , y buscar la minimización del costo promedio, cuando las observaciones son i.i.d. y con horizonte finito. A este se le llama el problema modificado de Kiefer-Weiss, (ver [Kiefer, J., Weiss, L., (1957)] y [Weiss (1962)]).

Un tema que también se considera en la tesis es determinar la estructura de un proceso estocástico utilizando el criterio de Wald. En este se determinan las fórmulas para el error tipo I , el error tipo II y el número promedio de observaciones.

1.6. Reglas generales de paro y de decisión

Para probar H_0 vs H_1 dado un plan de muestreo t , se define una prueba de hipótesis multietápica como el par (ψ_t, ϕ_t) , con ψ_t regla de paro y ϕ_t regla de decisión (terminal). La regla de paro está dada por

$$\psi_t = (\psi_{t|1}, \psi_{t|2}, \dots, \psi_{t|k-1}),$$

donde

$$\psi_{t|n} = \psi_{t|n}(x^{(t|n)}) = \psi_{t|n}(x_1, \dots, x_{t_1+\dots+t_n}), \quad n = 1, 2, \dots, k-1,$$

es una función medible y toma valores en $[0, 1]$.

El valor de $\psi_{t|n}(x^{(t|n)})$ es interpretado como la probabilidad condicional *de parar y proceder a tomar una decisión*, dado que se está en la etapa n del experimento, y las observaciones tomadas son $x^{(t|n)}$. Entonces, $1 - \psi_{t|n}(x^{(t|n)})$ es la esperanza condicional de no parar y tomar las observaciones de la etapa $n + 1$ del experimento sin tomar alguna decisión. Si el experimento no se detiene en alguna etapa $n < k$, entonces la toma de decisión se debe hacer en la última etapa k , con base en el vector completo de observaciones $x^{(t)}$ y el experimento debe terminar en este momento.

Para cumplir con la última condición de paro, ψ_t debe satisfacer

$$(1 - \psi_{t|1}(x^{(t|1)}))(1 - \psi_{t|2}(x^{(t|2)})) \dots (1 - \psi_{t|k}(x^{(t|k)})) = 0 \quad (1.10)$$

esta restricción es distintiva de las pruebas multietápicas.

Una vez que el experimento se detiene, se debe tomar la decisión *rechazar o no rechazar H_0* , para esto se define la función de decisión

$$\phi_t = (\phi_{t|1}, \phi_{t|2}, \dots, \phi_{t|k})$$

donde

$$\phi_{t|n} = \phi_{t|n}(x^{(t|n)}) = \phi_{t|n}(x_1, \dots, x_{t_1+\dots+t_n}), \quad n = 1, 2, \dots, k.$$

Se supone que estas funciones son medibles y que toman valores en $[0, 1]$.

La función $\phi_{t|n}(x^{(t|n)})$ se interpreta como la probabilidad condicional de *rechazar* la hipótesis nula H_0 , dado que el experimento se detiene en la etapa n , siendo $x^{(t|n)}$ el vector de datos observados. Entonces $1 - \phi_{t|n}(x^{(t|n)})$ es la probabilidad de *aceptar* H_0 bajo estas condiciones.

El procedimiento descrito anteriormente genera un tiempo de paro τ_ψ (que toma el valor de n cuando el algoritmo indica detener el experimento en esta etapa) y una función de decisión δ_ϕ (que toma el valor de 1 cuando H_0 se rechaza y 0 cuando se acepta, de acuerdo a la decisión aleatorizada en base de $\phi_{t|n}$ cuando el procedimiento de prueba se detiene en la etapa n).

1.7. Planteamiento general del error tipo I, tipo II y costo

Para alguna prueba (ψ_t, ϕ_t) se define la *función de potencia* $P(\theta; \psi_t, \phi_t)$ como la probabilidad (total) de rechazo de H_0 , cuando el verdadero valor del parámetro es θ .

A continuación se probará que la función de potencia en el caso k -etápico es de la forma

$$P(\theta; \psi_t, \phi_t) = \sum_{n=1}^k P_\theta(\tau_\psi = n, \delta_\phi = 1) = \sum_{n=1}^k E_\theta[r_{t|n}\psi_{t|n}\phi_{t|n}] \quad (1.11)$$

siendo

$$r_{t|1} = (1 - \psi_1)(1 - \psi_2) \dots (1 - \psi_{t|n-1}) \quad (1.12)$$

para $n = 2, 3, \dots, k$ y $r_{t|1} = 1$.

Demostración. Debido a que el experimento debe parar a lo más en la etapa k , entonces $\psi_k = 1$.

$$\begin{aligned} P(\theta; \psi_t, \phi_t) &= P_\theta(\text{rechazar } H_0) \\ &= P_\theta(\psi_{t|1} = 1, \phi_{t|1} = 1) + P_\theta((1 - \psi_{t|1})\psi_{t|2} = 1, \phi_{t|2} = 1) + \dots + \\ &\quad P_\theta((1 - \psi_{t|1}), \dots, (1 - \psi_{t|k-1})\psi_{t|k} = 1, \phi_{t|k} = 1) \\ &= E_\theta[\psi_{t|1}\phi_{t|1} + (1 - \psi_{t|1})\psi_{t|2}\phi_{t|2} + \dots + (1 - \psi_{t|1}) \dots (1 - \psi_{t|k-2})\psi_{t|k-1}\phi_{t|k-1} \\ &\quad + (1 - \psi_1) \dots (1 - \psi_{t|k-1})\psi_{t|k}\phi_{t|k}] \\ &= \sum_{n=1}^k E_\theta[r_{t|n}\psi_{t|n}\phi_{t|n}]. \end{aligned} \quad (1.13)$$

Si en (1.11) se calcula la esperanza bajo θ_0 , se obtiene la expresión para el error tipo I y este es

$$\alpha(\psi_t, \phi_t) = \sum_{n=1}^k E_{\theta_0}(1 - \psi_{t|1}) \dots (1 - \psi_{t|n-1})\psi_{t|n}\phi_{t|n} \quad (1.14)$$

y si $\theta = \theta_1$ y se calcula la probabilidad de no rechazar H_0 , se obtiene la expresión para el cálculo del error tipo II

$$\beta(\psi_t, \phi_t) = \sum_{n=1}^k E_1(1 - \psi_{t|1}) \dots (1 - \psi_{t|n-1})\psi_{t|n}(1 - \phi_{t|n}). \quad (1.15)$$

La probabilidad $\alpha(\psi_t, \phi_t) = P(\theta_0; \psi_t, \phi_t)$ es llamada la probabilidad de error tipo I, y $\beta(\psi_t, \phi_t) = 1 - P(\theta_1; \psi_t, \phi_t)$ es la probabilidad de error tipo II.

Normalmente se busca que las probabilidades de error estén acotadas de la siguiente manera

$$\alpha(\psi_t, \phi_t) \leq \alpha, \quad \text{y} \quad \beta(\psi_t, \phi_t) \leq \beta \quad (1.16)$$

con $\alpha, \beta \in (0, 1)$, (ver [Wald, A., Wolfowitz, J., (1948)]).

Suponiendo que cada observación tiene un costo, entonces el costo por el grupo de observaciones de la n -ésima etapa es $c_n(t) > 0$, y el costo total de todas las observaciones tomadas hasta la etapa n . es $C_n(t) = c_1(t) + c_2(t) + \dots + c_n(t)$.

Entonces el costo promedio de todo el experimento secuencial es

$$\begin{aligned} K(\theta; \psi_t) &= \sum_{n=1}^k C_n(t) P_\theta(\tau_\psi = n) \\ &= \sum_{n=1}^k C_n(t) E_\theta(1 - \psi_{t|1}) \dots (1 - \psi_{t|n-1})\psi_{t|n} \end{aligned} \quad (1.17)$$

La forma en que generalmente se plantean los costos del experimento es $c_n(t) = ct_n$ (proporcional al número de observaciones), (ver [Wald, A., Wolfowitz, J., (1948)]). Para ejemplos más generales (ver [Schmitz and Sueselbeck (1983)], entre muchos otros), son del tipo $c_n(t) = ct_n + d$ (proporcional al número de observaciones en el grupo, más el costo por el grupo). En el presente trabajo siempre se supondrá que $c_n(t) > 0$, para todo $n = 1, 2, \dots, k$.

Sea \mathcal{P}^k el conjunto de todos los planes de muestreo, a continuación se enuncian algunos casos de nuestro interés. En el caso de tiempo discreto, sea $\mathcal{P}^k = \{(1, 1, \dots, 1)\}$ el cual corresponde a todas las pruebas secuenciales truncadas, o $\mathcal{P}^k = \{(m, m, \dots, m)\}$ que corresponde a pruebas secuenciales grupales con el mismo grupo de observaciones en cada grupo, etc. En el caso de tiempo continuo, se puede pensar en $\mathcal{P}^k = \{(\Delta t, \Delta t, \dots, \Delta t)\}$ que corresponde a procesos truncados a tiempos discretizados, o $\mathcal{P}^k = (0, \infty)^k$ que corresponde a todas las pruebas multietápicas, etc.

Para cualquier plan de muestreo $t \in \mathcal{P}^k$, sea \mathcal{F}_t la clase de todas las reglas de paro que satisfacen (1.10), y \mathcal{T}_t la clase de todas las clases multietápicas (ψ_t, ϕ_t) basadas sobre este plan de muestreo.

Sea

$$\mathfrak{F}^k = \bigcup_{t \in \mathcal{P}^k} \mathcal{F}_t,$$

y

$$\mathfrak{T}^k = \bigcup_{t \in \mathcal{D}^k} \mathcal{T}_t.$$

El principal problema a resolver es: la minimización del costo promedio $K(\theta; \psi_t)$ en la clase de las pruebas $(\psi_t, \phi_t) \in \mathfrak{T}^k$ sujetas a (1.16).

Dado el valor del parámetro θ , lo que se quiere minimizar es $K(\theta, \psi_t)$ y esto puede variar dependiendo de las necesidades. Por ejemplo, si se quiere minimizar $K(\theta_0, \psi_t)$ o $K(\theta_1, \psi_t)$ (como se plantea en [Wald, A., Wolfowitz, J., (1948)]) o se minimizan ambos costos a la vez. Esto parece ser una posibilidad afortunada únicamente para observaciones i.i.d. y horizonte infinito). Otra posibilidad para θ en la minimización de $K(\theta, \psi_t)$ es fijar un valor entre θ_0 y θ_1 , este problema es conocido en el caso de observaciones i.i.d. y horizonte infinito como “el problema modificado de Kiefer-Weiss”, y es algunas veces útil para resolver el problema original de Kiefer-Weiss, es decir, la minimización de $\sup_{\theta} K(\theta, \psi_t)$ (ver [Kiefer, J., Weiss, L., (1957)], [Weiss (1962)], ver también [Lorden, G., (1980)] y [Schmitz and Sueselbeck (1983)]).

Capítulo 2

Teoría secuencial multietápica

El objetivo de este capítulo es desarrollar la teoría necesaria para realizar pruebas de hipótesis simples, tomando las observaciones en grupos de manera secuencial, para eso se crea una función en la que participa el error tipo I , el error tipo II y el costo del experimento. El objetivo es encontrar una regla de paro ψ_t que minimice el costo dado en (1.17), sobre todas las pruebas sujetas a (1.16).

Se inicia definiendo la función de Lagrange como

$$L(\psi_t, \phi_t; \lambda_0, \lambda_1) = K(\theta; \psi_t) + \lambda_0 \alpha(\psi_t, \phi_t) + \lambda_1 \beta(\psi_t, \phi_t) \quad (2.1)$$

donde $\lambda_0 \geq 0$ y $\lambda_1 \geq 0$, son constantes llamadas multiplicadores.

El siguiente teorema es una aplicación directa de la idea del método de multiplicadores de Lagrange.

Teorema 2.1. *Sea $\mathcal{T} \subset \mathfrak{T}^k$ una clase secuencial k -etápica. Existe una prueba $(\psi, \phi) \in \mathcal{T}$, tal que para todas las pruebas $(\psi'_t, \phi'_t) \in \mathcal{T}$, se cumple*

$$L(\psi_t, \phi_t; \lambda_0, \lambda_1) \leq L(\psi'_t, \phi'_t; \lambda_0, \lambda_1) \quad (2.2)$$

con

$$\alpha(\psi_t, \phi_t) = \alpha \quad y \quad \beta(\psi_t, \phi_t) = \beta. \quad (2.3)$$

Entonces, para cada prueba $(\psi'_t, \phi'_t) \in \mathcal{T}$, tal que

$$\alpha(\psi'_t, \phi'_t) \leq \alpha \quad y \quad \beta(\psi'_t, \phi'_t) \leq \beta \quad (2.4)$$

se tiene

$$K(\theta; \psi_t) \leq K(\theta; \psi'_t). \quad (2.5)$$

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

Demostración. Sea $(\psi'_t, \phi'_t) \in \mathcal{T}$ que satisface (2.2), debido a (2.3)

$$\begin{aligned} K(\theta; \psi'_t) + \lambda_0 \alpha + \lambda_1 \beta &\geq K(\theta; \psi'_t) + \lambda_0 \alpha(\psi'_t, \phi'_t) + \lambda_1 \beta(\psi'_t, \phi'_t) \\ &= L(\psi'_t, \phi'_t) \\ &\geq K(\theta; \psi_t) + \lambda_0 \alpha(\psi_t, \phi_t) + \lambda_1 \beta(\psi_t, \phi_t) \\ &= K(\theta; \psi_t) + \lambda_0 \alpha + \lambda_1 \beta \end{aligned} \quad (2.6)$$

2. Teoría secuencial multietápica

de la cadena de desigualdades se sigue que

$$K(\theta; \psi) \leq K(\theta; \psi')$$

esto demuestra la primera parte del teorema. Para que $K(\theta; \psi) = K(\theta; \psi')$, se requiere que haya igualdades en (2.6). \square

El Teorema 2.1 muestra que las pruebas que minimizan el costo promedio bajo las restricciones en las probabilidades de error, se pueden encontrar a través de la minimización de la función de Lagrange (por ejemplo (2.2)). Respectivamente, se busca una prueba (ψ, ϕ) tal que

$$\min_{t, \psi_t, \phi_t} L(\psi_t, \phi_t) = \min_t \min_{\psi_t} \min_{\phi_t} L(\psi_t, \phi_t) \quad (2.7)$$

en la clase \mathfrak{T}^k .

El lado derecho de (2.7) indica la manera para hacer la minimización: primero se minimiza sobre ϕ_t una t dada, luego a ψ_t con t dada y finalmente sobre $t \in \mathcal{P}^k$. Este será el plan de trabajo a seguir para el resto de esta sección.

Para simplificar la notación se hacen las siguientes aclaraciones: Para alguna regla de paro $\psi_t \in \mathfrak{F}^k$, se definen las funciones

$$s_{t|n}^\psi = (1 - \psi_{t|1}) \cdots (1 - \psi_{t|n-1}) \psi_{t|n} \quad \text{y} \quad r_{t|n}^\psi = (1 - \psi_{t|1}) \cdots (1 - \psi_{t|n-1}),$$

con $n = 1, 2, \dots, k$.

Además se define I_A , como la función indicadora del evento A .

El siguiente lema (Lemma 5.1 dado en [Novikov (2009)]), es un resultado de gran ayuda para los objetivos que se persiguen.

Lema 2.1. *Sea ϕ , F_1 y F_2 algunas funciones medibles sobre un espacio medible con una medida μ , tal que*

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

y

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

Entonces

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

con igualdad si y sólo si

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

μ -casi donde quiera.

2.1. La regla de decisión óptima

Regresando a la búsqueda de la optimización de la función de Lagrange. El siguiente teorema da la mejor regla de decisión final, para realizar la prueba de H_0 vs H_1 para cualquier plan de muestreo dado.

Teorema 2.2. *Sea $t \in \mathcal{P}^k$ un plan de muestreo fijo, y $\psi_t \in \mathcal{F}_t$ una regla de paro dada. Entonces*

$$\lambda_0 \alpha(\psi_t, \phi_t) + \lambda_1 \beta(\psi_t, \phi_t) \geq \sum_{n=1}^k \int r_{t|n}^\psi \psi_{t|n} \min\{\lambda_0 f_{\theta_0}^{(t|n)}, \lambda_1 f_{\theta_1}^{(t|n)}\} d\mu^{(t|n)} \quad (2.8)$$

cualquiera que sea la regla de decisión ϕ_t . Hay igualdad en (2.8) si y sólo si

$$I_{\{\lambda_0 f_{\theta_0}^{(t|n)} < \lambda_1 f_{\theta_1}^{(t|n)}\}} \leq \phi_t \leq I_{\{\lambda_0 f_{\theta_0}^{(t|n)} \leq \lambda_1 f_{\theta_1}^{(t|n)}\}} \quad (2.9)$$

$\mu^{(t|n)}$ -casi donde quiera, sobre $\mathcal{S}_{t|n}^\psi = \{x^{(t|n)} : r_{t|n}^\psi(x^{(t|n)}) \psi_{t|n}(x^{(t|n)}) > 0\}$, para todo $n = 1, 2, \dots, k$.

Demostración. Debido a (1.14) y (1.15)

$$\lambda_0 \alpha(\psi_t, \phi_t) + \lambda_1 \beta(\psi_t, \phi_t) = \sum_{n=1}^k \int s_{t|n}^\psi [\phi_{t|n} \lambda_0 f_0^{(t|n)} + (1 - \phi_{t|n}) \lambda_1 f_1^{(t|n)}] d(\mu^{(t|n)}) \quad (2.10)$$

aplicando el lema 2.1, a cada uno de los sumandos del lado derecho de (2.10), se tiene

$$\lambda_0 \alpha(\psi_t, \phi_t) + \lambda_1 \beta(\psi_t, \phi_t) \geq \sum_{n=1}^k \int s_{t|n}^\psi \min\{\lambda_0 f_0^{(t|n)}, \lambda_1 f_1^{(t|n)}\} d(\mu^{(t|n)})$$

con igualdad, si y sólo si, ϕ_n satisface a (2.9), $\mu^{(t|n)}$ -casi donde quiera, sobre $\mathcal{S}_{t|n}^\psi$, para algún $n = 1, 2, \dots, k$. \square

Dado que la función de Lagrange ahora sólo depende del tiempo de paro, ésta puede redefinirse como

$$L_t(\psi_t) = L_t(\psi_t; \lambda_0, \lambda_1) = \inf_{\phi} L_t(\psi_t, \phi; \lambda_0, \lambda_1). \quad (2.11)$$

Corolario 1. *Para cualquier plan de muestro $t \in \mathcal{P}^k$ y para cualquier regla de paro $\psi_t \in \mathcal{F}_t$*

$$L_t(\psi_t) = \sum_{n=1}^k \int s_{t|n}^\psi (C_n(t) f_{\theta}^{(t|n)} + l^{(t|n)}) d\mu^{(t|n)} \quad (2.12)$$

donde por definición

$$l^{(t|n)} \equiv \min\{\lambda_0 f_{\theta_0}^{(t|n)}, \lambda_1 f_{\theta_1}^{(t|n)}\}. \quad (2.13)$$

Demostración. Esta se obtiene sustituyendo a (1.17) y (2.10) en (2.11), posteriormente se aplica el teorema 2.2. La igualdad es debido a que se satisface la regla de decisión (2.9). \square

2.2. Paro óptimo

Dado que se tiene una regla de decisión óptima para cualquier tiempo de paro. El interés ahora está en minimizar a $L(\psi_t)$ (la función de Lagrange) para un plan de diseño $t \in \mathcal{P}^k$. Lo cual es un problema de paro óptimo y se resuelve utilizando la “inducción hacia atrás”.

Se define el operador B_t como sigue: para cualquier función $v = v(x^{(t|n)})$ se tiene que

$$B_n^{(t)}v(x^{(t|n)}) = \int v(x^{(t|n)}, x_t^{(t)})d\mu_t(x_n^{(t)}) \quad (2.14)$$

(por el teorema de Fubini, $B_n^{(t)}$ está bien definida para cada $x^{(t|n)}$ e integrable con respecto a $\mu^{t|n-1}$).

La estructura de las reglas de paro óptimo y que minimiza parte de (2.7), es determinada por las funciones $V^{(t|n)}$ para $n \leq k$, las cuales se calculan de la siguiente manera: $V^{(t|k)} \equiv l^{(t|k)}$ (ver (2.13)) y recursivamente para $n = k - 1, k - 2, \dots, 1$,

$$V^{(t|n)} = \text{mín}\{l^{(t|n)}, c_{n+1}(t)f_\theta^{(t|n)} + B_{t_{n+1}}^{(t)}V^{(t|n+1)}\}. \quad (2.15)$$

De manera natural, las condiciones de paro solo serán relevantes cuando se trata de datos con probabilidad condicional distinta de cero, es decir, sobre el conjunto respectivo

$$R_{t|n}^\psi = \{x^{(t|n)} : r_{t|n}^\psi(x^{(t|n)}) > 0\}$$

para todo $n = 1, 2, \dots, k$. El siguiente teorema describe todas las pruebas óptimas k -etápicas para un plan de muestreo fijo t .

Teorema 2.3. *Sea $t \in \mathcal{P}^k$ y cualquier plan de muestreo fijo. Para toda $\psi_t \in \mathcal{F}_t$*

$$L_1(\psi_t) \geq c_1(t) + B_1^{(t)}V^{(t|1)}. \quad (2.16)$$

Hay igualdad en (2.16), si y sólo si, se cumplen las siguientes condiciones para todo $n = 1, 2, \dots, k - 1$

$$I_{\{l^{(t|n)} < c_{n+1}(t)f_\theta^{(t|n)} + B_{t_{n+1}}^{(t)}V^{(t|n+1)}\}} \leq \psi_{t|n} \leq I_{\{l^{(t|n)} \leq c_{n+1}(t)f_\theta^{(t|n)} + B_{t_{n+1}}^{(t)}V^{(t|n+1)}\}} \quad (2.17)$$

$\mu^{(t|n)}$ -casi donde quiera sobre $R_{t|n}^\psi$, y $\psi_{t|k} = 1$ casi donde quiera sobre $R_{t|k}^\psi$.

El siguiente lema proporciona ayuda para la demostración del teorema 2.3.

Lema 2.2. *Sea $t = (t_1, \dots, t_k)$, $k \geq 2$ y sea $v = v(x^{(t|k)})$ alguna función no-negativa integrable (con respecto a $\mu^{t|k}$). Entonces*

$$\begin{aligned} & \int s_{t|k-1}^\psi (C_{k-1}(t)f_\theta^{(t|k-1)} + l^{(t|k-1)})d\mu^{(t|k-1)} + \int r_{t|k}^\psi (C_k(t)f_\theta^{(t|k)} + v)d\mu^{(t|k)} \\ & \geq \int r_{t|k-1}^\psi (C_{k-1}(t)f_\theta^{(t|k-1)} + Q_{t_k}v)d\mu^{(t|k-1)} \end{aligned} \quad (2.18)$$

con

$$Q_{t_k}v = \text{mín}\{l^{(t|k-1)}, c_k(t)f_\theta^{(t|k-1)} + B_k^{(t)}v\} \quad (2.19)$$

siendo $B_k^{(t)}$ el operador definido en (2.14).

La cota inferior en (2.18) se puede obtener, si y sólo si

$$I_{\{l^{(t|k-1)} < Q_{t_k} v\}} \leq \psi_{t|k-1} \leq I_{\{l^{(t|k-1)} \leq Q_{t_k} v\}}$$

$\mu^{t|k-1}$ -casi donde quiera, sobre

$$R_{t|k-1}^\psi = \{x^{(t|k-1)} : r_{t|k-1}^\psi(x^{(t|k-1)}) > 0\}.$$

Demostración. Por el Teorema de Fubini, la parte izquierda de (2.18) se puede ver como

$$\begin{aligned} & \int s_{t|k-1}^\psi (C_{k-1}(t) f_\theta^{(t|k-1)} + l^{(t|k-1)}) d\mu^{(t|k-1)} + \\ & \int r_{t|k}^\psi \int (C_k(t) f_\theta^{(t|k)} + v) d\mu^{t_k} d\mu^{(t|k-1)} \\ &= \int r_{t|k-1}^\psi [\psi_{t|k-1} (C_{k-1}(t) f_\theta^{(t|k-1)} + l^{(t|k-1)}) + \\ & (1 - \psi_{t|k-1}) \int (C_k(t) f_\theta^{(t|k)} + v) d\mu^{t_k}] d\mu^{(t|k-1)} \end{aligned} \quad (2.20)$$

pero $f_\theta^{(t|k)}(x^{(t|k)})$ es una función de densidad conjunta de $(X_1, \dots, X_{t_1+\dots+t_k})$, entonces

$$\int f_\theta^{(t|k)}(x_1, \dots, x_{t_1+\dots+t_k}) d\mu^{t_k} = f_\theta^{(t|k-1)}(x_1, \dots, x_{t_1+\dots+t_{k-1}})$$

así que la parte derecha de (2.20) es

$$\begin{aligned} &= \int r_{t|k-1}^\psi [\psi_{t|k-1} C_{k-1}(t) f_\theta^{(t|k-1)} + \psi_{t|k-1} l^{(t|k-1)} \\ & + (1 - \psi_{t|k-1}) (C_k(t) f_\theta^{(t|k-1)} + \int v d\mu^{t_k})] d\mu^{(t|k-1)} \\ &= \int r_{t|k-1}^\psi [\psi_{t|k-1} C_{k-1}(t) f_\theta^{(t|k-1)} + \psi_{t|k-1} l^{(t|k-1)} \\ & + (1 - \psi_{t|k-1}) (C_{k-1}(t) + c_k(t)) f_\theta^{(t|k-1)} + \int v d\mu^{t_k}] d\mu^{(t|k-1)}, \end{aligned}$$

por la definición del operador $B_n^{(t)}$ y simplificando, esto es

$$\begin{aligned} &= \int r_{t|k-1}^\psi [C_{k-1}(t) f_\theta^{(t|k-1)} + \psi_{t|k-1} l^{(t|k-1)} + \\ & (1 - \psi_{t|k-1}) (c_k(t) f_\theta^{(t|k-1)} + B_k^{(t)} v)] d\mu^{(t|k-1)}. \end{aligned} \quad (2.21)$$

Aplicando el Lema 2.1, se sigue que (2.21) es mayor o igual a

$$\begin{aligned} & \int r_{t|k-1}^\psi [C_{k-1}(t) f_\theta^{(t|k-1)} + \min\{l^{(t|k-1)}, c_k f_\theta^{(t|k-1)} + B_{t_k} v\}] d\mu^{(t|k-1)} \\ &= \int r_{t|k-1}^\psi (C_{k-1}(t) f_\theta^{(t|k-1)} + Q_{t_k} v) d\mu^{(t|k-1)}. \end{aligned} \quad (2.22)$$

Debido a la definición de $Q_{t_k} v$ en (2.19), hay igualdad entre (2.21) y (2.22) si y sólo si se cumple (2.17), μ -casi donde quiera sobre $R_{t|k}^\psi$. \square

2. Teoría secuencial multietápica

Ahora ya se tienen las condiciones para hacer la prueba pendiente.

Demostración. Del Teorema 2.3. Debido al Corolario 1, la función de Lagrange es

$$\sum_{n=1}^{k-1} \int s_{t|n}^{\psi} (C_n(t) f_{\theta}^{(t|n)} + l^{(t|n)}) d\mu^{(t|n)} + \int s_{t|k}^{\psi} (C_k(t) f_{\theta}^{(t|k)} + V^{(t|k)}) d\mu^{(t|k)} \quad (2.23)$$

y se quiere probar que esto es mayor o igual que la siguiente expresión

$$\sum_{n=1}^{k-2} \int s_{t|n}^{\psi} (C_n(t) f_{\theta}^{(t|n)} + l^{(t|n)}) d\mu^{(t|n)} + \int s_{t|k-1}^{\psi} (C_{k-1}(t) f_{\theta}^{(t|k-1)} + V^{(t|k-1)}) d\mu^{(t|k-1)} \quad (2.24)$$

para probar esto, es suficiente ver que la diferencia entre (2.23) y (2.24) es mayor o igual a cero. Desarrollando a (2.23) para $n = k - 1$ y posteriormente a esto se le resta la expresión (2.24), se obtiene

$$\int s_{t|k-1}^{\psi} (C_{k-1}(t) f_{\theta}^{(t|k-1)} + l^{(t|k-1)}) d\mu^{(t|k-1)} + \int s_{t|k}^{\psi} (C_k(t) f_{\theta}^{(t|k)} + V^{(t|k)}) d\mu^{(t|k)} \geq 0 \quad (2.25)$$

porque por definición cada una de las funciones que participan en (2.25) son no negativas. Por otra parte, en particular se tiene que (2.25)

$$\geq \int s_{t|k-1}^{\psi} (C_{k-1}(t) f_{\theta}^{(t|k-1)} + V^{(t|k-1)}) d\mu^{(t|k-1)} \quad (2.26)$$

y esto se cumple debido al Lema 2.2, con igualdad en (2.26) si y sólo si

$$I_{\{l^{(t|k-1)} < Q_{t|k-1} V\}} \leq \psi_{t|k-1} \leq I_{\{l^{(t|k-1)} \leq Q_{t|k-1} V\}}$$

donde

$$Q_{t|k-1} = c_k(t) f_{\theta}^{(t|k-1)} + B_k^{(t)} V^{(t|k)}$$

siendo $B_k^{(t)}$ es el operador definido en (2.14).

Análogamente, se demuestra que la parte derecha de (2.26) es

$$\geq \sum_{n=1}^{k-3} \int s_{t|n}^{\psi} (C_n(t) f_{\theta}^{(t|n)} + l^{(t|n)}) d\mu^{t|n} + \int s_{t|k-2}^{\psi} (C_{k-2}(t) f_{\theta}^{(t|k-2)} + V^{(t|k-2)}) d\mu^{t|k-2}$$

con igualdad si y sólo si

$$I_{\{l^{(t|k-2)} < Q_{t|k-2} V\}} \leq \psi_{t|k-2} \leq I_{\{l^{(t|k-2)} \leq Q_{t|k-2} V\}}$$

donde

$$Q_{t|k-2} = c_{k-1}(t) f_{\theta}^{(t|k-2)} + B_{k-1}^{(t)} V^{(t|k-1)}.$$

Aplicando consecutivamente el Lema 2.2 finalmente se obtiene (2.16), con igualdad si y sólo si la regla de paro satisface (2.17). \square

Del Teorema 2.3, se tiene que la parte media del mínimo en (1.16) es igual a $c_1(t) + B_k^{(t)} V^{(t|1)}$, para completar la minimización de (2.7) es necesario buscar planes de muestreo que alcancen

$$\min_{t \in \mathcal{P}^k} \{c_1(t) + B_1^{(t)} V^{(t|1)}\}. \quad (2.27)$$

Nota 2.1. Tener en cuenta que en el esquema actual se requiere que la primera etapa esté siempre presente (esto explica el término $c_1(t)$ en (2.16)). Formalmente, las reglas de paro óptimas del teorema 2.3 se minimizan bajo esta condición, sin embargo, en ocasiones hay algo mejor que hacer antes de comenzar la primera etapa: no hacer nada. Por ejemplo, si $\min\{\lambda_0, \lambda_1\} \leq c_1(t) + B_{t_1} V^{(t|1)}$, entonces la decisión trivial $I_{\{\lambda_0 \leq \lambda_1\}}$ indica que no se requiere ninguna observación, ésto funciona mejor que las reglas óptimas del teorema 2.3 (por la misma razón que se explica en la Nota 2.3 en [Novikov (2009)]).

Si hay más de un plan de muestreo en \mathcal{P}^k , parece que la única forma de minimizar a (2.27) es de forma numérica, para la mayoría de las aplicaciones. Para pruebas secuenciales de grupo del mismo tamaño ($\mathcal{P}^k = \{(m, m, \dots, m), m \in \mathbb{N}\}$), la optimización sobre m puede ser una tarea importante para las aplicaciones, cuando el costo incluye una parte pagada por grupo (por ejemplo, $c_n(t) = ct_n + d, d > 0$).

Si sólo hay un elemento en \mathcal{P}^k (por ejemplo, para todas las pruebas secuenciales truncadas, $\mathcal{P}^k = \{(1, 1, \dots, 1)\}$), no se necesita la minimización de (2.27). En particular, se obtiene inmediatamente de los Teoremas 2.1 a 2.3, las pruebas truncadas óptimas de modelos para observaciones i.i.d. A continuación se mencionan algunos usos importantes de estos resultados para este caso particular.

Nota 2.2. *Se sabe que la prueba secuencial óptima en el problema modificado de Kiefer-Weiss (ver [Lorden, G., (1980)]), para las familias Koopman-Darmois está truncada, por lo que estas se pueden obtener de los Teoremas 2.1 a 2.3. La prueba óptima (3) en [Lorden, G., (1980)] tiene que ser un caso particular de las pruebas óptimas de los Teoremas 2.1 a 2.3 para este problema, porque ahí se describen todas las pruebas aleatorizadas que minimizan la función de Lagrange y esta es la forma en que la prueba (3) en [Lorden, G., (1980)] se obtiene.*

Nota 2.3. Alternativas bilaterales. Hay una manera de incluir en el esquema actual el caso importante para aplicaciones de alternativas “bilaterales” (ver [Jennison (1999)]). Se puede tomar formalmente

$$f_{\theta_1}^{(t|n)} = \pi' f_{\theta'}^{(t|n)} + \pi'' f_{\theta''}^{(t|n)}$$

con $\theta' < \theta_0 < \theta''$, $\pi' > 0$, $\pi'' > 0$, $\pi' + \pi'' = 1$, como la “densidad alternativa” y modificar las fórmulas en los teoremas 2.1-2.3 respectivamente. El método *por sí mismo* proporciona pruebas para las cuales

$$\pi' P_{\theta'}(\delta = 0) + \pi'' P_{\theta''}(\delta = 0) = \beta \tag{2.28}$$

dejando la posibilidad de usar π' y π'' como “ajustes” adicionales para hacer que las probabilidades (de error tipo II) en (2.28) sean iguales.

Nota 2.4. Pruebas bayesianas multi-etápicas. De manera similar, se pueden tratar las pruebas Bayesianas. Estas se obtienen de los teoremas 2.2 y 2.3 tomando formalmente para la evaluación de $K(\theta; \psi_t)$ en (1.17) a

$$f_{\theta}^{(t|n)} = \pi_0 f_{\theta_0}^{(t|n)} + \pi_1 f_{\theta_1}^{(t|n)}$$

donde π_0, π_1 son algunos números positivos tales que $\pi_0 + \pi_1 = 1$ (cf. [Wald, A., Wolfowitz, J., (1948)]). Usando esta opción en el teorema 2.3, se obtiene la estructura de todas las pruebas secuenciales k -etápicas Bayesianas.

2.3. Minimización del costo

En esta sección se resolverá el problema de minimización del costo del experimento, es decir, de $K(\theta; \psi_t)$ sobre todas las k etapas con probabilidades de error sujetas a

$$\alpha(\psi_t, \phi_t) \leq \alpha, \quad y \quad \beta(\psi_t, \phi_t) \leq \beta.$$

Combinando los resultados de los teoremas 2.2 y 2.3, se tiene el siguiente resultado.

Teorema 2.4. *Sea k un número natural, supóngase que existe un plan de muestreo $t^* = (t_1^*, \dots, t_k^*) > 0$, tal que para alguna $t = (t_1, t_2, \dots, t_k) > 0$,*

$$V^{(t|0)}(t^*) \leq V^{(t|0)}(t) \tag{2.29}$$

(ver (2.15)) y sea $(\psi(t^*), \phi(t^*))$ alguna prueba para la cual $\psi(t^*)$ cumple (2.17) con $n = 1$, $t = t^*$ y $\phi(t^*)$ satisface (2.9). Dado

$$\alpha(\psi(t^*), \phi(t^*)) = \alpha \quad y \quad \beta(\psi(t^*), \phi(t^*)) = \beta$$

entonces para alguna prueba k etápica $(\psi_t, \phi_t) \in \Delta^k$, tal que

$$\alpha(\psi_t, \phi_t) \leq \alpha \quad y \quad \beta(\psi_t, \phi_t) \leq \beta \tag{2.30}$$

se tiene que

$$K(\theta; \psi(t)) \geq K(\theta; \psi(t^*)). \tag{2.31}$$

La desigualdad en (2.31) es estricta, si una de las desigualdades en (2.30) es estricta.

Por el teorema 2.4, lo que queda del problema es encontrar

$$\min_{t=(t_1, \dots, t_k)} V^{(t|0)}. \tag{2.32}$$

Como esta prueba está construida con componentes de los Teoremas 2.2 y 2.3, entonces (2.32) es óptima. La única forma de dar solución a (2.32) es por medio de la minimización numérica. En [Novikov (2004)] y [Novikov (2006)] se realizó este trabajo para la prueba de dos hipótesis simples para un proceso de Wiener con deriva lineal en dos etapas ($k = 2$).

Los Teoremas y las demostraciones en la teoría expuesta son una generalización de los Teoremas que se presentan en [Novikov (2004)] y [Novikov (2006)]. Es decir en forma general son resultados que se obtuvieron en el desarrollo del presente trabajo y que se han obtenido partiendo del caso de dos etapas.

Capítulo 3

Planteamiento general multietápico del proceso de Wiener con deriva lineal

Los resultados teóricos de las secciones anteriores son muy generales. A continuación se dan dos aplicaciones igualmente muy generales y después se harán más específicas de tal manera que se obtendrán las fórmulas explícitas para realizar evaluaciones de las mismas y mostrar resultados numéricos.

Para el tiempo continuo, la suposición de (1.8) también es bastante general, por lo que es necesario realizar un ejemplo para un caso en concreto.

Supóngase que el proceso de Wiener con deriva lineal está disponible para observaciones en etapas

$$X_s = \sigma W(s) + \theta s, \quad s \geq 0 \quad (3.1)$$

donde $W(s)$ es un proceso de Wiener estándar, así que $EW(s) = 0$ y $Var(W(s)) = s$, $s \geq 0$. Sea μ_t la distribución de $\{W(s)\}_{s \in [0,t]}$, se sabe que la distribución de $\{X_s\}_{s \in [0,t]}$ es absolutamente continuo con respecto a μ_t por lo que la derivada respectiva de Radon-Nikodym es

$$f_\theta(\{x_s\}_{s \in [0,t]}) = \exp\left(x_t \frac{\theta}{\sigma^2} - \frac{t \theta^2}{2 \sigma^2}\right).$$

Sea $t \in \mathcal{P}^k$ algún plan de muestreo. Usando la independencia de los incrementos de W y $s_n(t)$ (definido en (1.9)), se tiene que

$$f_\theta^{(t|n)}(x^{(t|n)}) = \exp\left(x_{s_n(t)} \frac{\theta}{\sigma^2} - \frac{s_n(t) \theta^2}{2 \sigma^2}\right) \quad (3.2)$$

es la derivada de Radon-Nikodym con respecto a la medida producto $\mu^n = \mu \otimes \mu \otimes \dots \otimes \mu$.

3.1. El proceso de Wiener con deriva lineal en etapas

Supóngase que un proceso de Wiener con deriva lineal (3.1) es observado en k etapas. Sea $\mathcal{P}^k = (0, 1)^k$, con $k \geq 2$.

El interés está en probar $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ con $\theta_0 \neq \theta_1$, suponiendo que σ en (3.1) es conocida. Sin pérdida de generalidad se puede suponer que $\sigma = 1$ y $\theta_0 = 0$. De esta manera, el valor de θ_1 es el único parámetro que permanece en el modelo y para simplificar la notación se considera $\theta_1 = \theta$.

Con respecto al costo, sea $c_n(t) = t_n$ con $t \in \mathcal{P}^k$ y $n \leq k$. Se pretende hacer el cálculo del costo ponderado bajo H_0 y H_1

$$c_0 K(\theta_0; \psi_t) + c_1 K(\theta_1; \psi_t)$$

siendo c_0 y c_1 constantes no negativas.

Una pequeña diferencia con el criterio Bayesiano al final de la nota 2.4, es que aquí una de las constantes pueden ser cero. Se supondrá que $c_0 > 0$ y bajo esta suposición considerar que $c_n(t) = t_n$, entonces $K(\theta_i, \psi)$ se convierte en la duración promedio de la prueba bajo H_0 y H_1 .

Se tienen dos alternativas para los valores de c_0 y c_1 : $c_0 = 1, c_1 = 1$ y $c_0 = 1, c_1 = 0$.

- La primera opción corresponde al criterio Bayesiano que otorga ponderaciones iguales a la duración media tanto en H_0 como en H_1 , a esto le llamamos *el criterio equilibrado Bayesiano*.
- El segundo caso (cuando $c_0 = 1, c_1 = 0$) y se obtiene el costo promedio bajo H_0 , sin considerar el valor bajo H_1 , a esto se le llamará *el criterio bajo H_0* .

Es de esperarse que en el segundo caso, la duración promedio de la prueba tendrá valores menores en comparación con el criterio del equilibrio Bayesiano (probablemente a expensas de la duración promedio elevada de la prueba bajo H_1).

Sea $t \in \mathcal{P}^k$ cualquier plan de muestreo. De acuerdo a las especificaciones del modelo, se toma (3.2) con $\sigma = 1$

$$f_{\theta}^{(t|n)}(x^{(t|n)}) = \exp\left(x_{s_n(t)}\theta - \frac{s_n(t)}{2}\theta^2\right) = z_{\theta}(x_{s_n(t)}, s_n(t)) \quad (3.3)$$

donde, por definición

$$z_{\theta}(x, t) = \exp(x\theta - t\theta^2/2). \quad (3.4)$$

La regla de paro óptimo del teorema 2.3, se puede expresar de una manera más compacta en términos de las siguientes funciones $\rho^{(t|n)}(z)$ con $z \geq 0$ y $n = 1, \dots, k$.

Empezando con $n = k$

$$\rho^{(t|n)}(z) = g(z) = \min\{\lambda_0, \lambda_1 z\} \quad (3.5)$$

y $z \geq 0$, recursivamente para $n = k - 1, k - 2, \dots, 1$,

$$\rho^{(t|n)}(z) = \min \left\{ g(z), (c_0 + c_1 z)t_{n+1} + \int_{-\infty}^{\infty} \rho^{(t|n+1)}(z z_{\theta}(x, t_{n+1})) d\Phi \left(\frac{x}{\sqrt{t_{n+1}}} \right) \right\} \quad (3.6)$$

donde Φ es la función de distribución acumulada de la normal estándar.

Proposición 3.1. *Para las funciones $V^{(t|n)}$ y $\rho^{(t|n)}(z)$ definidas en (2.15) y (3.6) respectivamente, se tiene que*

$$V^{(t|n)}(x^{(t|n)}) = \rho^{(t|n)}(z_{\theta}(x_{s_n(t)}, s_n(t))) \quad (3.7)$$

para todo $n = k, k - 1, \dots, 1$.

Demostración. Esta es por inducción hacia atrás. Para $n = k$

$$\begin{aligned} V^{(t|k)}(x^{(t|k)}) &= \min \left\{ \lambda_0, \lambda_1 \frac{f_{\theta_1}^{(t|k)}(x^{(t|k)})}{f_{\theta_0}^{(t|k)}(x^{(t|k)})} \right\} f_{\theta_0}^{(t|k)}(x^{(t|k)}) \\ &= \min \{ \lambda_0, \lambda_1 z \} f_{\theta_0}^{(t|k)}(x^{(t|k)}) \end{aligned} \quad (3.8)$$

donde

$$z = \frac{f_{\theta_1}^{(t|k)}}{f_{\theta_0}^{(t|k)}} \quad (3.9)$$

y por otra parte

$$\rho^{(t|k)}(z) = \min \{ \lambda_0, \lambda_1 z \}. \quad (3.10)$$

Supóngase que (3.7) se cumple para $n = m < k$, es decir

$$\begin{aligned} V^{(t|m)}(x^{(t|m)}) &= \min \left\{ g(z), (c_0 + c_1 z)t_{m+1} + \int_{-\infty}^{\infty} \rho^{(t|m+1)}(z z_{\theta}(x, t_{m+1})) d\Phi \left(\frac{x}{\sqrt{t_{m+1}}} \right) \right\} \\ &= \rho^{(t|m)}(z^{(t|m)}). \end{aligned} \quad (3.11)$$

Por demostrar que (3.7) se cumple para $n = m - 1$, es decir

$$V^{(t|m-1)}(x^{(t|m-1)}) = \min \left\{ g(z), (c_0 + c_1 z)t_m + \int_{-\infty}^{\infty} \rho^{(t|m)}(z z_{\theta}(x, t_m)) d\Phi \left(\frac{x}{\sqrt{t_m}} \right) \right\}.$$

Demostración, para $n = m - 1$, se tiene

$$V^{(t|m-1)}(x^{(t|m-1)}) = \min \left\{ l(z), (c_0 + c_1 z)t_m + \int_{-\infty}^{\infty} V^{(t|m)}(x^{(t|m)}, x_k^{(t)}) d\mu_t(x_k^{(t)}) \right\}$$

por definición $l(z) = g(z)$ y debido a la hipótesis de inducción (ver (3.11)), se sigue

$$\begin{aligned} &= \min \left\{ g(z), (c_0 + c_1 z)t_m + \int_{-\infty}^{\infty} \rho^{(t|m)}(z_{\theta}(x_{s_m(t)}, s_m(t))) d\Phi \left(\frac{x}{\sqrt{t_m}} \right) \right\} \\ &= \rho^{(t|m-1)}(z). \end{aligned}$$

Con esto concluye la demostración de (3.7). □

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

En particular, el valor mínimo en el lado derecho de (2.16) del teorema 2.3 adquiere la siguiente forma

$$(c_0 + c_1)t_1 + \int_{-\infty}^{\infty} \rho^{(t_1)}(z_\theta(x, t_1)) d\Phi\left(\frac{x}{\sqrt{t_1}}\right). \quad (3.12)$$

Esta es la expresión que se minimizará numéricamente para encontrar las duraciones óptimas de la etapa de acuerdo con (2.7).

Otros elementos importantes relacionados con la optimización en la duración de la etapa dada en (2.7), son las respectivas reglas de paro y decisión, las cuales también tienen una forma compacta en términos de $\rho^{(t^k)}$.

Se analizará la desigualdad en el lado derecho de (2.17), con la finalidad de definir una regla de paro óptimo. Es fácil ver que (2.17) se puede reescribir como

$$g(z) \leq (c_0 + c_1 z)t_{n+1} + \int_{-\infty}^{\infty} \rho^{(t_{n+1})}(z z_\theta(x, t_{n+1})) d\Phi\left(\frac{x}{\sqrt{t_{n+1}}}\right) \quad (3.13)$$

suponiendo que

$$z = z_\theta(x_{s_n(t)}, s_n(t)). \quad (3.14)$$

Para simplificar la notación, se denotará la integral en el lado derecho de (3.13) como $I^{(t^n)}(z)$. No es difícil ver (como en [Novikov, A., (2008)]) que $I^{(t^n)}(z)$ es una función cóncava de z en $[0, \infty)$. Entonces,

$$F(z) = \lambda_1 z - (c_0 + c_1 z)t_{n+1} - I^{(t^n)}(z)$$

es una función convexa de $z \geq 0$ y $F(0) = -c_0 t_{n+1} < 0$.

De manera similar

$$G(z) = \lambda_0 - (c_0 + c_1 z)t_{n+1} - I^{(t^n)}(z)$$

es convexa y $\lim_{z \rightarrow \infty} G(z) < 0$

De estas propiedades se sigue que

- Si $F(\lambda_0/\lambda_1) (= G(\lambda_0/\lambda_1)) \geq 0$, entonces existe un único $A_n \leq \lambda_0/\lambda_1$ tal que $F(A_n) = 0$ y un único $B_n \geq \lambda_0/\lambda_1$ tal que $G(B_n) = 0$. Además, (3.13) sólo se cumple si $z < A_n$ o $z > B_n$.
- Si $F(\lambda_0/\lambda_1) < 0$, entonces (3.13) se cumple para todo $z \geq 0$.

Como consecuencia, se tiene que la condición de no parar (lo contrario a (3.13)) sólo ocurre cuando $z \in (A_n, B_n)$, siendo A_n y B_n las raíces de la ecuación

$$\min\{\lambda_0, \lambda_1 z\} = (c_0 + c_1 z)t_{n+1} + \int_{-\infty}^{\infty} \rho^{(t_{n+1})}(z z_\theta(x, t_{n+1})) d\Phi\left(\frac{x}{\sqrt{t_{n+1}}}\right). \quad (3.15)$$

El intervalo (A_n, B_n) definido de la forma descrita anteriormente, es llamado intervalo de continuación. Si A_n y B_n no existen (o existen, pero $A_n = B_n$), la prueba debe

detenerse en la etapa n , independientemente de los datos recibidos. Todas las demás reglas de paro óptimas de (2.17) serán equivalentes a ésta, porque difieren sólo en los puntos de datos que tienen probabilidad cero.

En cuanto a la regla de decisión que se aplicará una vez que se ha detenido el proceso, se sigue del teorema 2.2 que H_0 debe rechazarse al momento de paro si $z \geq \lambda_0/\lambda_1$, siendo z definida en (3.14). Si el proceso se detiene debido a que $z \notin (A_n, B_n)$, esto implica que H_0 se rechazará cuando $z \geq B_n$, y se aceptará cuando $z \leq A_n$.

Para obtener (2.27) ahora se debe minimizar numéricamente a (3.12) sobre todos los positivos t_1, t_2, \dots, t_k . Como era de esperar, solo se necesita una rutina para minimizar (3.12) cuando $\theta = 1$: las duraciones de etapa óptimas para otros θ se pueden obtener de aquellas calculadas para $\theta = 1$, dividiendo cada valor de t_1, \dots, t_k por θ^2 . En ese caso, los intervalos de continuación siguen siendo los mismos. Por esta razón, solo se analiza el caso $\theta = 1$ en la siguiente sección. Para otros valores de θ , las pruebas óptimas se modifican como se describió anteriormente y todos los valores relativos como los de eficiencia relativa, no dependen de θ .

3.2. Fórmulas multietápicas del proceso de Wiener con deriva lineal

En este momento se tiene todo lo necesario para dar solución al problema de minimización de la función de lagrange dado en (2.1), esto se hará con el proceso de Wiener con deriva lineal.

Para $n = k$, por (3.8)

$$V^{(t|k)}(x^{(t|k)}) = \min \{ \lambda_0, \lambda_1 z_k \} f_{\theta_0}^{(t|k)}(x^{(t|k)}) \quad (3.16)$$

donde

$$z_k = z_\theta(x_{s_k(t)}, s_k(t)) = e^{\theta x_k - \frac{\theta^2 t_k}{2}} \quad (3.17)$$

ver (3.4), a partir de (3.16) se obtiene

$$U^{(t|k)}(z z_k) = \frac{V^{(t|k)}}{f_{\theta_0}^{(t|k)}(t_k; x_k)} = \min \{ \lambda_0, \lambda_1 z z_k \}. \quad (3.18)$$

Para $n = k - 1$

$$U^{(t|k-1)}(z z_{k-1}) = \min \left\{ g(z), (c_0 + c_1 z) t_k + \int f_{\theta_0}(t_k; x_k) U^{(t|k)}(z z_k) dx_k \right\} \quad (3.19)$$

siendo g la función que está definida en (3.5).

Haciendo $\lambda_0 = \lambda_1 z z_k$ (ver (3.18)), sustituyendo a z_k (ver (3.17)) y despejando a x_k , se obtiene

$$x_k = \frac{1}{\theta} \ln \left(\frac{\lambda_0}{\lambda_1 z} \right) + \frac{\theta t_k}{2}$$

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

y entonces

$$U^{(t|k)}(zz_k) = \begin{cases} \lambda_1 z z_k & \text{if } z \leq x_k \\ \lambda_0 & \text{if } x_k \leq z. \end{cases} \quad (3.20)$$

Así, la integral en (3.19) es

$$= \int_{-\infty}^{x_k} \lambda_1 z f_{\theta_0}(t_k; x_k) z_k dx_k + \int_{x_k}^{\infty} \lambda_0 f_{\theta_0}(t_k; x_k) dx_k$$

sustituyendo a z_k (ver (3.9)) en la primer integral y en la segunda debido a que f_{θ_0} es función de densidad

$$= \lambda_1 z \int_{-\infty}^{x_k} f_{\theta_1}(t_k; x_k) dx_k + \lambda_0 \left[1 - \int_{-\infty}^{x_k} f_{\theta_0}(t_k; x_k) dx_k \right] \quad (3.21)$$

integrando y agregando el costo, la parte derecha del mínimo de (3.19) es equivalente a

$$F_k(t_k; z) = (c_0 + c_1 z)t_k + \lambda_1 z \Phi(l_k) + \lambda_0 (1 - \Phi(m_k)) \quad (3.22)$$

donde

$$l_k = \frac{1}{\theta \sqrt{t_k}} \ln\left(\frac{\lambda_0}{\lambda_1 z}\right) - \frac{\theta \sqrt{t_k}}{2} \quad (3.23)$$

y

$$m_k = \frac{1}{\theta \sqrt{t_k}} \ln\left(\frac{\lambda_0}{\lambda_1 z}\right) + \frac{\theta \sqrt{t_k}}{2} \quad (3.24)$$

debido a (3.22), se tiene que (3.19) es equivalente a

$$U^{(t|k-1)}(zz_{k-1}) = \min \{g(z), F_k(t_k; z)\}. \quad (3.25)$$

La función U para $n = k - 2$, es

$$U^{(t|k-2)}(zz_{k-2}) = \min \left\{ g(z), (c_0 + c_1 z)t_{k-2} + \int f_{\theta_0}(t_{k-1}; x_{k-1}) U^{(t|k-1)}(zz_{k-1}) dx_{k-1} \right\} \quad (3.26)$$

para la integral de (3.26) se determinan los límites de integración haciendo $F_k(t_k; z) = g(z)$ (ver (3.25)), se obtienen las siguientes igualdades

$$F_k(t_k; z) = \lambda_1 z \quad (3.27)$$

y

$$F_k(t_k; z) = \lambda_0 \quad (3.28)$$

resolviendo con respecto a z se obtienen A_{k-1} y B_{k-1} respectivamente.

Las constantes A_{k-1} y B_{k-1} particionan el eje Z y la función $U^{(t|k-1)}$ dada en (3.25) queda definida de la siguiente manera

$$U^{(t|k-1)}(zz_{k-1}) = \begin{cases} \lambda_1 z z_{k-1} & \text{si } z z_{k-1} \leq A_{k-1} \\ F_k(t_k; z) & \text{si } A_{k-1} \leq z z_{k-1} \leq B_{k-1} \\ \lambda_0 & \text{si } B_{k-1} \leq z z_{k-1}. \end{cases} \quad (3.29)$$

Sea z_0 un número cualquiera en el eje z , haciendo $z z_{k-1} = z_0$ se tiene que $z_{k-1} = \frac{z_0}{z}$, ahora sustituyendo a z_{k-1} (ver (3.17)) y despejando a x_{k-1} , se obtiene

$$x_{k-1} = \frac{\ln(z_0) - \ln(z)}{\theta} + \frac{\theta t_{k-1}}{2}, \quad (3.30)$$

3.2. Fórmulas multietápicas del proceso de Wiener con deriva lineal

en particular si $z_0 = A_{k-1}$ ó $z_0 = B_{k-1}$ en (3.30), entonces se pueden definir

$$A_{k-1}^* = \frac{\ln(A_{k-1}) - \ln(z)}{\theta} + \frac{\theta t_{k-1}}{2} \quad (3.31)$$

y

$$B_{k-1}^* = \frac{\ln(B_{k-1}) - \ln(z)}{\theta} + \frac{\theta t_{k-1}}{2}. \quad (3.32)$$

Debido a estas aclaraciones, (3.29) es equivalente a

$$U^{(t|k-1)}(z z_{k-1}) = \begin{cases} \lambda_1 z z_{k-1} & \text{si } x_{k-1} \leq A_{k-1}^* \\ F_k(t_k; z) & \text{si } A_{k-1}^* < x_{k-1} < B_{k-1}^* \\ \lambda_0 & \text{si } A_{k-1}^* \leq x_{k-1} \end{cases} \quad (3.33)$$

así por (3.33) la integral dada en (3.26) es igual a

$$\begin{aligned} &= \lambda_1 z \int_{-\infty}^{A_{k-1}^*} f_{\theta_0}(t_{k-1}; x_{k-1}) z_{k-1} dx_{k-1} + \int_{A_{k-1}^*}^{B_{k-1}^*} f_{\theta_0}(t_{k-1}; x_{k-1}) F_k(t_k; z) dx_{k-1} \\ &+ \lambda_0 \int_{B_{k-1}^*}^{\infty} f_{\theta_0}(t_{k-1}; x_{k-1}) dx_{k-1}. \end{aligned} \quad (3.34)$$

A continuación se realiza el cálculo de estas integrales. Sustituyendo z_{k-1} (ver (3.9)) en la primer integral de (3.34), se obtiene

$$\lambda_1 z \int_{-\infty}^{A_{k-1}^*} f_{\theta_1}(t_{k-1}; x_{k-1}) dx_{k-1} = \lambda_1 z \Phi(l_{k-1}) \quad (3.35)$$

con

$$l_{k-1} = \frac{\ln(A_{k-1}) - \ln(z)}{\theta \sqrt{t_{k-1}}} - \frac{\theta \sqrt{t_{k-1}}}{2}. \quad (3.36)$$

Para realizar la segunda integral de (3.34), observar que en ésta participan dos funciones con diferentes variables, con esto en mente se particiona el eje z en el intervalo $[A_{k-1}, B_{k-1}]$

$$\Delta z = \frac{B_{k-1} - A_{k-1}}{n}, \quad y \quad z_i = A_{k-1} + i \Delta z, \quad i = 0, 1, 2, \dots, n \quad (3.37)$$

para alguna $n \in \mathbb{N}$ fija.

Evaluando a F_k en la partición (3.37) se obtiene la matriz

$$U = (z_i, F_k(t_k; z_i)) \quad (3.38)$$

pero las z_i están en función de x_i tal y como se mostró en (3.30), entonces la partición dada en (3.37) en término de x_i queda de la siguiente forma

$$x_i \equiv \frac{\ln(z_i) - \ln(z)}{\theta} + \frac{\theta t_{k-1}}{2} \quad \text{para } i = 0, 1, \dots, n \quad (3.39)$$

en estos puntos se evalúan las funciones que participan en la segunda integral de (3.34), así se obtiene la matriz

$$I = [x_i, f_0(t_{k-1}; \bar{x}_i) F_k(t_k; z_i)] \quad (3.40)$$

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

y esta integral debido a (3.37) y (3.39), se calcula a través de

$$\sum_{i=0}^{n-1} \frac{f_{\theta_0}(t_{k-1}; x_i)F_k(t_k; z_i) + f_{\theta_0}(t_{k-1}; x_{i+1})F_k(t_k; z_i)}{2} (x_{i+1} - x_i) = S_{k-1}(t_k, t_{k-1}; z). \quad (3.41)$$

La tercer integral de (3.34) es

$$= \lambda_0 \left(1 - \int_{-\infty}^{B_{k-1}^*} f_{\theta_0}(t_{k-1}; x_{k-1}) dx_{k-1} \right) = \lambda_0 (1 - \Phi(m_{k-1})) \quad (3.42)$$

siendo

$$m_{k-1} = \frac{\ln(B_{k-1}) - \ln(z)}{\theta\sqrt{t_{k-1}}} + \frac{\theta\sqrt{t_{k-1}}}{2}. \quad (3.43)$$

Entonces, la integral dada en (3.34) más el costo debido a (3.35), (3.41) y (3.42), es

$$F_{k-1}(t_k, t_{k-1}; z) = (c_0 + c_1 z)t_{k-1} + \lambda_1 z \Phi(l_{k-1}) + S_{k-1}(t_k, t_{k-1}; z) + \lambda_0 (1 - \Phi(m_{k-1})). \quad (3.44)$$

Debido a (3.44), se tiene que (3.26) es equivalente a

$$U^{(t|k-2)}(z) = \text{mín} \{g(z), F_{k-1}(t_k, t_{k-1}; z)\}. \quad (3.45)$$

Para obtener $U^{(t|n)}$ para $n = k-3, \dots, 1$ se procede de forma análoga a como se obtuvo $U^{(t|k-2)}$, suponiendo que en cada caso se cumple

$$g(z) > (c_0 + c_1 z)t_{n+1} + \int f_{\theta_0}(t_{n+1}; x_{n+1}) U^{(t|n+1)}(z z_{n+1}) dx_{n+1}.$$

Siguiendo este proceso recursivo se llega a

$$U^{(t|1)}(z) = \text{mín} \left\{ g(z), (c_0 + c_1 z)t_2 + \int f_{\theta_0}(t_2; x_2) U^{(t|2)}(z z_2) dx_2 \right\}$$

y haciendo los cálculos análogos a los realizados en el caso $n = k-2$, se obtiene

$$U^{(t|1)}(z) = \text{mín} \{g(z), F_2(t_k, \dots, t_2; z)\}, \quad (3.46)$$

donde

$$F_2(t_k, \dots, t_2; z) = (c_0 + c_1 z)t_2 + \lambda_1 z \Phi(l_1) + S_2(t_k, \dots, t_2; z) + \lambda_0 (1 - \Phi(m_1)) \quad (3.47)$$

con l_1 y m_1 definidos en (3.36) y (3.43) con $k = 2$, respectivamente.

Para $n = 0$, se tiene que $f_{\theta_0}^{(t|0)}(x) = 1 = f_{\theta_1}^{(t|0)}(x) = 1$, por lo que

$$g(z) = \text{mín} \{\lambda_0, \lambda_1\} = \lambda$$

entonces

$$U^{(t|0)}(z) = \text{mín} \{ \lambda, (c_0 + c_1)(t_1) + \int f_{\theta_0}(t_1; x_1) U^{(t|1)}(z_1) dx_1 \}$$

y suponiendo que el mínimo no es λ , entonces se obtiene el mínimo de la función de Lagrange

$$(c_0 + c_1)t_1 + \int f_{\theta_0}(t_1; x_1) U^{(t|1)}(z_1) dx_1. \quad (3.48)$$

Para realizar la integral dada en (3.48), se debe determinar z tal que $F_2(t_k, \dots, t_2; z) = g(z)$. Esto nos lleva a los siguientes dos casos:

- Haciendo $F_2(t_k, \dots, t_2; z) = \lambda_1 z$, se obtiene A_1 .
- Y haciendo $F_2(t_k, \dots, t_2; z) = \lambda_0$, se obtiene B_1 .

Debido a esto, la función $U^{(t_1)}$ es equivalente a

$$U^{(t_1)}(z_1) = \begin{cases} \lambda_1 z_1 & \text{si } z_1 \leq A_1 \\ F_k(t_k; z) & \text{si } A_1 \leq z_1 \leq B_1 \\ \lambda_0 & \text{si } B_1 \leq z_1. \end{cases} \quad (3.49)$$

Si z_0 es un número cualquiera en el eje z , haciendo $z_1 = z_0$, sustituyendo a z_1 (ver (3.17)) y despejando a x_1 se tiene

$$x_1 = \frac{\ln(z_0)}{\theta} + \frac{\theta t_1}{2} \quad (3.50)$$

en particular si $z_0 = A_1$ o $z_0 = B_1$, se pueden definir

$$A_1^* = \frac{\ln(A)}{\theta} + \frac{\theta t_1}{2} \quad (3.51)$$

y

$$B_1^* = \frac{\ln(B)}{\theta} + \frac{\theta t_1}{2}. \quad (3.52)$$

Considerando estas nuevas constantes, entonces (3.49) es equivalente a

$$U^{(t_1)}(z_1) = \begin{cases} \lambda_1 & \text{if } x_1 \leq A_1^* \\ F_k(t_k, \dots, t_2; z) & \text{if } A_1^* \leq z_1 \leq B_1^* \\ \lambda_0 & \text{if } B_1^* \leq z_1 \end{cases} \quad (3.53)$$

debido a esto la integral de (3.48) se divide en tres, las cuales se realizan a continuación.

- La primera

$$\int_{-\infty}^{A_1^*} \lambda_1 z_1 f_{\theta_0}(t_1; x_1) dx_1 = \lambda_1 \int_{-\infty}^{A_1^*} f_{\theta_1}(t_1; x_1) dx_1 = \lambda_1 \Phi(l_1) \quad (3.54)$$

siendo

$$l_1 = \frac{\ln(A_1)}{\theta\sqrt{t_1}} - \frac{\theta\sqrt{t_1}}{2}. \quad (3.55)$$

- La segunda

$$\int_{A_1^*}^{B_1^*} f_{\theta_0}(t_1; x_1) F_2(t_k, \dots, t_2; z) dx_1 \quad (3.56)$$

como $F_1(t_k, \dots, t_2; z)$ está en función de z , se hace una partición del intervalo (A_1, B_1) análogo a (3.37) y evaluando la función F_1 en los puntos de esta partición se obtiene la matriz

$$U_1 = (z_i, F_1(t_k, \dots, t_2; z_i)) \quad (3.57)$$

a partir de (3.50), la partición anterior sobre el eje x es

$$x_i \equiv \frac{\ln(z_i)}{\theta} + \frac{\theta t_1}{2}, \quad i = 0, 1, \dots, n \quad (3.58)$$

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

evaluando a f_{θ_0} que participa en (3.56) se genera la matriz

$$I = [x_i, f_{\theta_0}(t_1; x_i)F_2(t_k, \dots, t_2; z_i)] \quad (3.59)$$

debido a estos cálculos, la integral dada en (3.56) por (3.57) y (3.59), es

$$S_1(t_k, \dots, t_1) = \sum_{i=0}^{n-1} \frac{f_{\theta_0}(t_1; x_i)F_2(t_k, \dots, t_2; z_i) + f_{\theta_0}(t_1; x_{i+1})F_2(t_k, \dots, t_2; z_{i+1})}{2} (x_{i+1} - x_i). \quad (3.60)$$

▪ La tercera

$$\lambda_0 \int_{B_1^*}^{\infty} f_{\theta_0}(t_1; x_1) dx_1 = \lambda_0(1 - \Phi(m_1)) \quad (3.61)$$

con

$$m_1 = \frac{\ln(B_1)}{\theta\sqrt{t_1}} + \frac{\theta\sqrt{t_1}}{2}. \quad (3.62)$$

De (3.54), (3.60) y (3.61) se obtiene la integral de (3.48). Por lo tanto, el mínimo de la función de Lagrange (ver (3.48)), es

$$(c_0 + c_1)t_1 + \lambda_1\Phi(l_1) + S_1(t_k, \dots, t_1) + \lambda_0(1 - \Phi(m_1)). \quad (3.63)$$

Una vez obtenido el limite inferior de la función de Lagrange en (3.63), el siguiente objetivo es determinar la fórmulas para calcular el error tipo I , el error tipo II y el número promedio muestral.

3.3. El error tipo I

A partir de los resultados de la sección anterior, se determina la expresión para obtener el error tipo I . Para la k -ésima etapa, la función correspondiente a F_k (ver (3.22)) es

$$A_k(t_k; z) = 1 - \Phi(m_k)$$

con m_k definida en (3.24). Para $k - 1$, la función que corresponde a F_{k-1} (ver (3.44)) es

$$A_{k-1}(t_k, t_{k-1}; z) = S_{k-1}(t_k, t_{k-1}; z) + (1 - \Phi(m_{k-1}))$$

S_{k-1} y m_{k-1} definidas en (3.41) y (3.43) respectivamente, los límites de S_{k-1} están dados en (3.31) y (3.32).

Procediendo de esta manera, la función correspondiente a (3.63) para el error tipo I es

$$\alpha = S_1(t_k, \dots, t_1) + (1 - \Phi(m_1)), \quad (3.64)$$

con S_1 y m_1 definidas en (3.60) y (3.62), los límites para calcular S_1 son A_1^* y B_1^* los cuales están dados en (3.51) y (3.52) respectivamente.

3.4. El error tipo II

Obtención del error tipo II. Para la k -ésima etapa, la función correspondiente a F_k (ver (3.22)) es

$$B_k(t_k; z) = z\Phi(l_k)$$

con l_k definida en (3.23).

Para $k - 1$, la función que corresponde a F_{k-1} (ver (3.44)) es

$$B_{k-1}(t_k, t_{k-1}; z) = z\Phi(l_{k-1}) + S_{k-1}(t_k, t_{k-1}; z)$$

l_{k-1} y S_{k-1} están definidas en (3.36) y (3.41) respectivamente, los límites de S_{k-1} se definieron en (3.31) y (3.32).

Siguiendo este camino, la función correspondiente a (3.63) para el error tipo II es

$$\beta = z\Phi(l_1) + S_1(t_k, \dots, t_1) \quad (3.65)$$

l_1 y S_1 se definieron respectivamente en (3.55) y (3.60), los límites para calcular S_1 son A_1^* y B_1^* están dados en (3.51) y (3.52).

3.5. Costo promedio bajo la hipótesis nula y la alternativa

El costo promedio, en la k -ésima etapa la función correspondiente a F_k (ver (3.22)) es

$$C_k(t_k; z) = (c_0 + c_1 z)t_k.$$

La función que corresponde a F_{k-1} (ver (3.44)) es

$$C_{k-1}(t_k, t_{k-1}; z) = (c_0 + c_1 z)t_{k-1} + S_{k-1}(t_k, t_{k-1}; z).$$

con S_{k-1} calculada en (3.41), los límites de S_{k-1} son A_{k-1}^* y B_{k-1}^* definidos en (3.31) y (3.32).

Finalmente, el costo promedio se obtiene con la función correspondiente a (3.63), es

$$\text{Costo} = (c_0 + c_1)t_1 + S_1(t_k, \dots, t_2) \quad (3.66)$$

S_1 está dada en (3.41), con límites A_1^* y B_1^* definidos en (3.51) y (3.52).

3.6. Costo promedio bajo la hipótesis nula

Para el cálculo del número promedio muestral bajo H_0 se hace $c_1 = 0$ en las funciones $U^{(t|n)}$ con $n = 1, \dots, k$, entonces el costo de la k -ésima etapa es

$$H_k(t_k; z) = c_0(t_k).$$

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

El costo para la etapa $k - 1$ debido a (3.34), es

$$H_{k-1}(t_{k-1}; z) = c_0(t_{k-1}) + \int_{A_{k-1}^*}^{B_{k-1}^*} c_0 t_k f_{\theta_0}^{(t|k)}(x^{(t|k)})$$

con A_{k-1}^* y B_{k-1}^* definidos en (3.31) y (3.32), integrando numéricamente se obtiene

$$H_{k-1}(t_{k-1}; z) = c_0(t_{k-1}) + S_{k-1}(t_k, t_{k-1}; z).$$

con S_{k-1} dada en (3.41)) y sus límites de integración están dados en (3.31) y (3.32). Finalmente la función que corresponde a (3.63), con respecto al costo promedio, es

$$\text{Costo}_0 = c_0(t_1) + S_1(t_k, \dots, t_2) \quad (3.67)$$

S_1 es la integral que se obtiene en (3.41) con límites A_1^* y B_1^* dados en (3.51) y (3.52).

3.7. Resultados de la evaluación del proceso de Wiener con deriva lineal

Los resultados numéricos que se presentan a continuación son para la prueba de las hipótesis $H_0 : \theta = 0$ vs $H_1 : \theta = 1$ sobre el coeficiente de deriva θ de un proceso Wiener con deriva lineal (ver sección 3.1 para más detalles). La optimización de (2.27) se realiza sobre los planes de etapa $(0, \infty)^k$ con $k = 2, 3, 4$. Para la evaluación numérica se utilizó el lenguaje de programación R ([R Core Team, 2013.]). Todos los resultados numéricos fueron verificados independientemente utilizando el sistema MATLAB ([MATLAB, 2010. version 7.10.0 (R2010a.)]).

Las Tablas 3.1, 3.2 y 3.3 muestran los valores de los parámetros de las pruebas óptimas de 2, 3 y 4 etapas para el criterio equilibrado bayesiano. Es decir, en el caso $c_0 = c_1 = 1$, para esto se hizo $\lambda_0 = \lambda_1$ para que se obtuviera $\alpha = \beta$. Las duraciones de prueba promedio respectivas bajo H_0 y H_1 las cuales también salen iguales y se denotan como N_2, N_3 y N_4 , respectivamente. Por otra parte, t_1, t_2, \dots , son las duraciones de las etapas óptimas obtenidas mediante la minimización de (3.12). $\log(B_1), \log(B_2)$, etc., son los logaritmos en base 10 de los puntos finales superiores de los intervalos de continuación en las etapas 1, 2, etc., definidos en (3.15). Los intervalos de continuación también salen simétricos, en el sentido de que $\log(A_1) = -\log(B_1)$, $\log(A_2) = -\log(B_2)$, etc.

Para evaluar la eficiencia, las pruebas óptimas de 2, 3 y 4 etapas, se comparan con la prueba óptima completamente secuencial (SPRT) y con la prueba no secuencial óptima de Neyman-Pearson.

Para cualquier par de errores α y β alcanzadas por la prueba k -etápica, se encontró la SPRT con los mismos valores de α y β y se compararon ambas duraciones promedio. Las fórmulas para la duración promedio de la prueba de la SPRT se encuentran en la sección 4.2 en [Shiryayev]. Para el promedio bajo de H_0 y H_1 (los denotamos como NW_0 y NW_1 , respectivamente), se tiene

$$NW_0 = \frac{w(\alpha, \beta)}{\theta^2/2}, \quad NW_1 = \frac{w(\beta, \alpha)}{\theta^2/2}$$

3.7. Resultados de la evaluación del proceso de Wiener con deriva lineal

donde

$$w(x, y) = x \ln \frac{x}{1-y} + (1-x) \ln \frac{1-x}{y}.$$

Para la prueba de Neyman-Pearson, la duración de la prueba respectiva es

$$NP = \frac{(\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta))^2}{\theta^2}.$$

La tabla 4.1 contiene los resultados de la comparación de las pruebas k -etápicas con la SPRT de Wald. Los resultados que se presentan son los valores de la razón entre la duración promedio de las pruebas k -etápicas y la duración promedio de la SPRT, para todos los $\alpha = \beta$ involucrados. Para 2 etapas, el valor medio observado es 1.5 ± 0.028 , para 3 etapas 1.3 ± 0.013 , y para 4 etapas 1.23 ± 0.006 , con una disminución perceptible al incrementarse el número de etapas. De hecho, es sorprendente que con solo *cuatro* etapas, el exceso de la duración promedio sea tan pequeña como un 24%, por arriba de la prueba *óptima*, la cual controla el proceso de prueba en *cualquier* momento, sobre el horizonte *infinito*.

Debido a la teoría, parece bastante razonable que cuando $k \rightarrow \infty$, la proporción respectiva se aproxime a 1, sea cual sea el valor de $\alpha = \beta$.

La tabla 4.2 contiene los resultados respectivos de la comparación de las pruebas multi-etápicas con la prueba de Neyman-Pearson. Para 2 etapas, el el valor medio de la proporción observada es 0.68 ± 0.024 , para 3 etapas 0.6 ± 0.027 , y para 4 etapas 0.56 ± 0.029 .

λ_0, λ_1	α, β	t_1	t_2	$\log(B_1)$	N2
40	0.1	3.46	4.29	0.55	4.96
80	0.05	5.59	7.23	0.77	7.80
190	0.02	8.40	11.45	1.06	11.37
350	0.01	10.39	14.63	1.26	13.80
700	0.005	12.62	18.33	1.50	16.43

Tabla 3.1. Pruebas óptimas de 2 etapas.

λ_0, λ_1	α, β	t_1	t_2	t_3	$\log(B_1)$	$\log(B_2)$	N3
35	0.1	2.55	2.44	3.79	0.64	0.51	4.46
67	0.05	4.06	3.84	6.43	0.88	0.71	6.86
154	0.02	6.09	5.64	10.39	1.20	0.98	9.84
300	0.01	7.71	6.98	13.8	1.46	1.21	12.06
570	0.005	9.22	8.17	17.20	1.71	1.43	14.06

Tabla 3.2. Pruebas óptimas de 3 etapas.

En las Figuras 3.1- 3.3 se preseta la eficiencia relativa de las pruebas para 2, 3 y 4 etapas, respectivamente. En cada gráfica, se uso λ_0 y λ_1 con logaritmos naturales de 3 a 8, con paso 1. Para cada par (de λ_0 y λ_1), se calcularon las duraciones óptimas de las

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

λ_0, λ_1	α, β	t_1	t_2	t_3	t_4	$\log(B_1)$	$\log(B_2)$	$\log(B_3)$	N4
33	0.1	2.12	1.72	2.32	3.54	0.69	0.62	0.49	4.26
63	0.05	3.37	2.64	3.68	6.16	0.95	0.86	0.69	6.50
144	0.02	5.05	3.77	5.43	10.04	1.28	1.18	0.96	9.24
270	0.01	6.31	4.57	6.87	13.23	1.54	1.42	1.17	11.16
520	0.005	7.63	5.24	7.87	16.72	1.81	1.68	1.40	13.01

Tabla 3.3. Pruebas óptimas de 4 etapas.

α, β	NW	$N2/NW$	$N3/NW$	$N4/NW$
0.1	3.52	1.41	1.27	1.21
0.05	5.30	1.47	1.29	1.23
0.02	7.47	1.52	1.32	1.24
0.01	9.01	1.53	1.34	1.24
0.005	10.48	1.57	1.34	1.24

Tabla 3.4. Eficiencia con respecto a la SPRT.

α, β	NP	$N2/NP$	$N3/NP$	$N4/NP$
0.1	6.57	0.75	0.68	0.65
0.05	10.82	0.72	0.63	0.60
0.02	16.87	0.67	0.58	0.55
0.01	21.65	0.64	0.56	0.52
0.005	26.54	0.62	0.53	0.49

Tabla 3.5. Eficiencia con respecto a la prueba de Neyman-Pearson.

3.7. Resultados de la evaluación del proceso de Wiener con deriva lineal

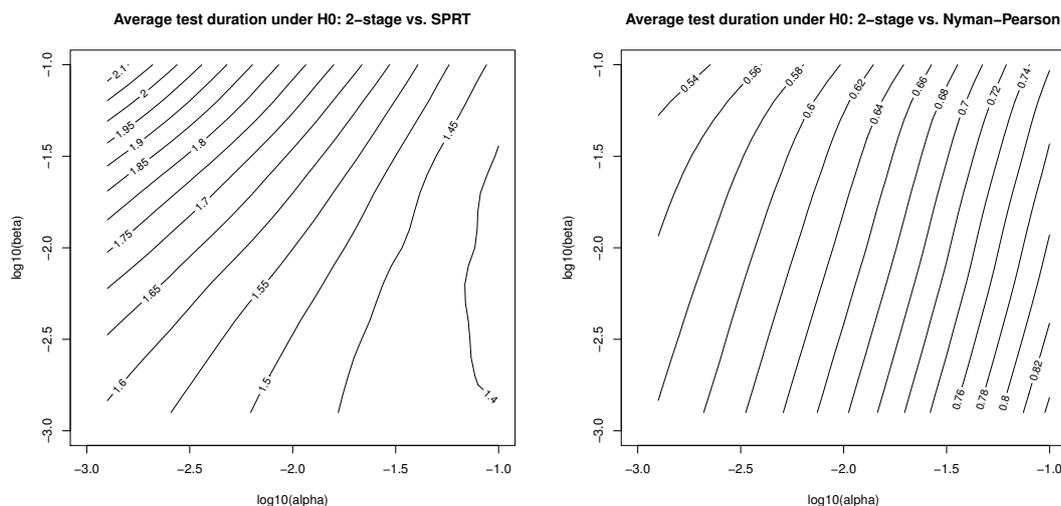


Figura 3.1. Eficiencia de pruebas de 2 etapas bajo H_0 . Criterio equilibrado Bayesiano.

etapas y los correspondientes α y β , además la duración promedio de la prueba bajo H_0 y H_1 , así como la duración promedio del SPRT para estos α y β , bajo de H_0 y H_1 , y finalmente la duración de la prueba de una etapa (Neyman-Pearson) con los mismos α y β . Es importante anotar que algunos pares de λ_0 y λ_1 no proporcionaron pruebas multietápicas que alcancen el valor mínimo de (3.12), siendo este igual a $\min\{\lambda_0, \lambda_1\}$ (ver Nota 2.1). Para cada pareja α y β se calculó la eficiencia relativa de la prueba de multietápicas, en la forma de la relación de su duración promedio de la SPRT respectiva, y a la duración de la prueba de Neyman-Pearson.

Para una mejor visualización los datos obtenidos de esta manera, se ajustó el modelo no paramétrico LOESS ¹ a cada uno de los conjuntos establecidos de eficiencia. Las Figuras 3.1- 3.3 son gráficos de contorno correspondientes a los modelos ajustados. Para esto los datos se obtuvieron utilizando el criterio equilibrado bayesiano.

Para evaluar el efecto de una posible disminución de la duración promedio usando el criterio de H_0 (únicamente), se presentan las figuras 3.4 y 3.5 respectivamente, realizadas para 4 etapas; los resultados para 2 y 3 etapas muestran en gran medida el mismo patrón. La Figura 3.4 presenta la eficiencia con respecto a la SPRT y la prueba de Neyman-Pearson bajo H_0 , y la Figura 3.5 muestra la eficiencia bajo H_1 . Es notable cómo la disminución relativamente pequeña en la duración promedio bajo H_0 se “compensa” con un aumento muy notable bajo H_1 .

¹ tt <https://www.rdocumentation.org/packages/stats/versions/3.6.1/topics/loess>

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

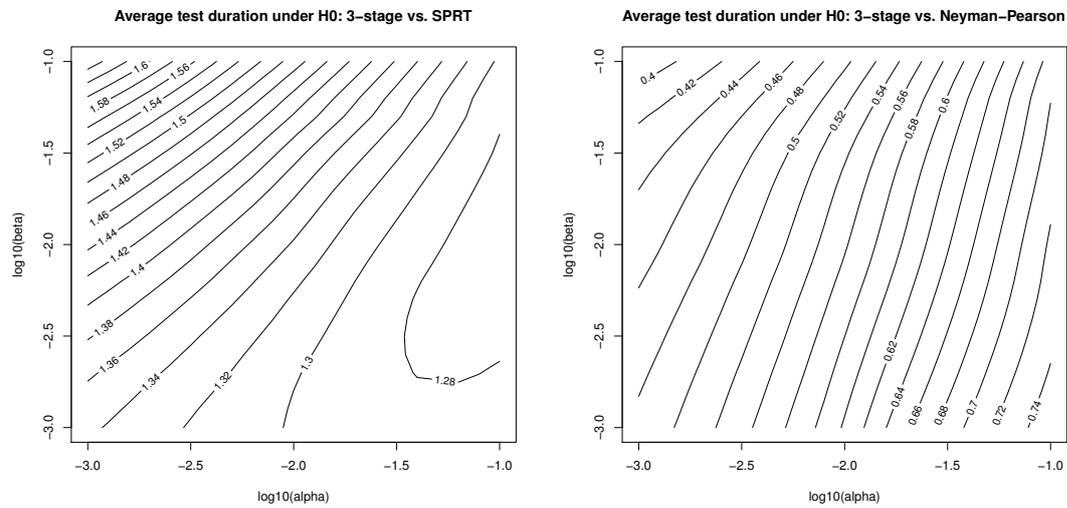


Figura 3.2. Eficiencia de pruebas de 3 etapas bajo H_0 . Criterio equilibrado Bayesiano.

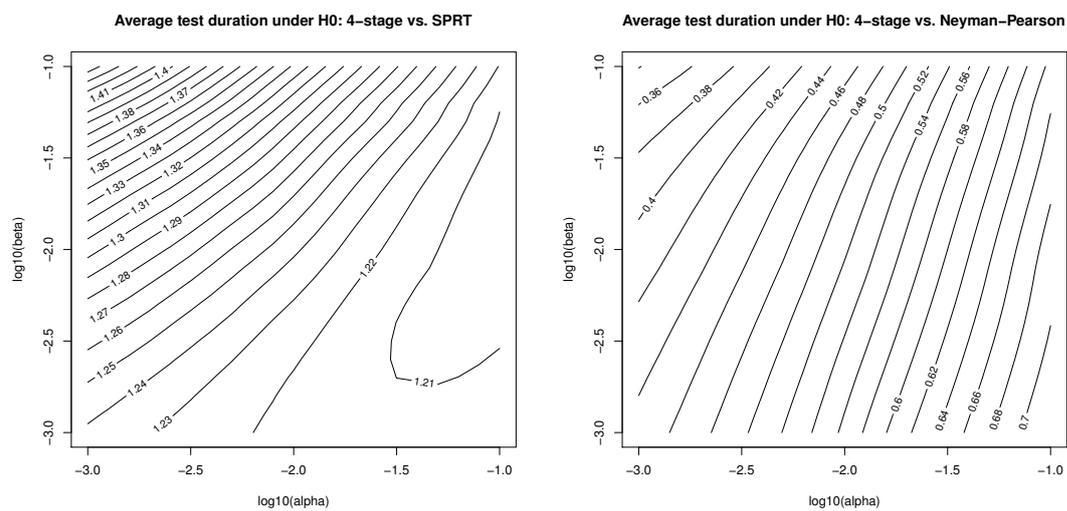


Figura 3.3. Eficiencia de pruebas de 4 etapas bajo H_0 . Criterio equilibrado Bayesiano.

3.7. Resultados de la evaluación del proceso de Wiener con deriva lineal

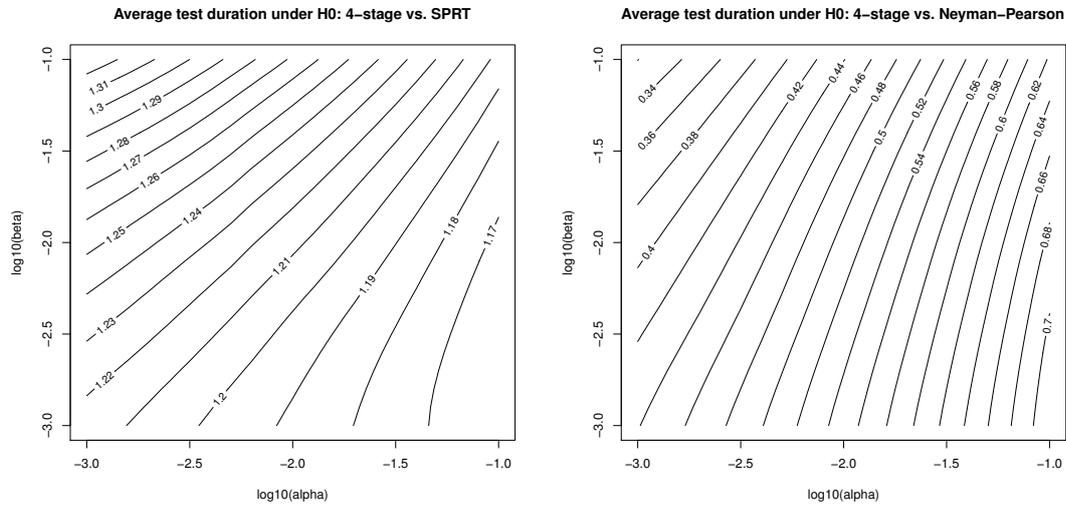


Figura 3.4. Eficiencia de pruebas de 4 etapas bajo H_0 . Criterio de H_0 únicamente.

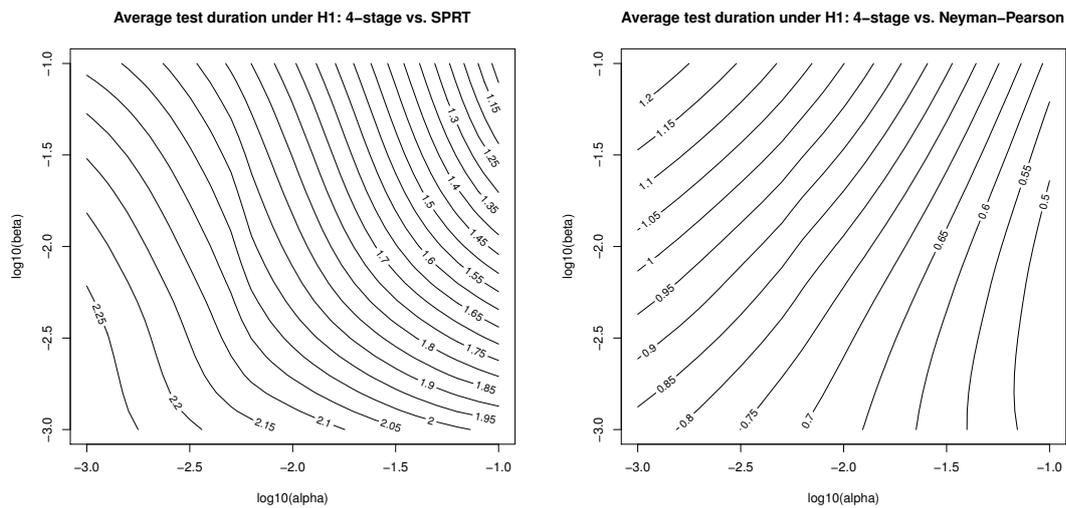


Figura 3.5. Eficiencia de pruebas de 4 etapas bajo H_1 . Criterio de H_1 únicamente.

3.8. Análisis de los resultados numéricos

Los resultados numéricos muestran una alta eficiencia en procedimientos de prueba secuencial multietápico, para una amplia gama de α y β , cuyos valores son representativos de los que se usan habitualmente en aplicaciones prácticas.

Para $\alpha = \beta$ se observa que la proporción promedio para pruebas en 2 etapas la media muestral de la SPRT es de 1.5 ± 0.028 ; para 3 etapas es 1.3 ± 0.013 , y para 4 etapas 1.23 ± 0.006 , con una disminución notable con el aumento del número de etapas. Para más de 4 etapas seguramente se obtendrán proporciones con niveles más bajos. El hecho es que, con una planificación óptima de las etapas, es posible obtener niveles de eficiencia cercanos a 1 incluso para un número relativamente bajo de etapas.

Para un amplio rango de parejas α, β , la relación se mantiene prácticamente igual, excepto en casos muy asimétricos: la media muestral observada para 2 etapas es de 1.58 ± 0.036 (más de $n = 31$ de parejas α, β), para 3 etapas 1.38 ± 0.023 ($n = 32$), para 4 etapas 1.26 ± 0.013 ($n = 32$) esto se ilustra en las Figuras 3.1 a 3.3). Los resultados anteriores se obtuvieron utilizando el criterio del equilibrio bayesiano.

Al hacer el cálculo promedio únicamente bajo H_0 no se toma en cuenta la duración promedio bajo H_1 , por lo tanto, se obtienen valores de proporción más bajos de la duración promedio bajo H_0 . Sin embargo, la disminución no es muy impresionante, el análisis corresponde a 4 etapas y se ilustra en la Figura 3.4). En la media sobre de 36 pares de α, β , la proporción de la duración promedio de pruebas multietápicas a la SPRT, con los mismos α, β , es de 1.26 ± 0.01 y parece que no hay disminución, en comparación con el caso equilibrado bayesiano (1.26 ± 0.013). Pero no hay contradicción, porque estos resultados son para otros pares α, β diferentes a los observados para el criterio equilibrado bayesiano. Más impresionante es que para H_1 la relación entre la duración media de las pruebas de 4 etapas y la SPRT fue de 2.08 ± 0.042 (este es un resumen de los resultados correspondientes a la Figura 3.5), lo que significa que la duración promedio de las pruebas de 4 etapas es aproximadamente el doble que la del SPRT, según la alternativa. En comparación con la prueba de Neyman-Pearson, la media muestral de la relación bajo H_1 fue de 0.97 ± 0.043 , lo que significa que la duración promedio de la prueba de 4 etapas bajo H_1 en la media es casi tan grande como la duración de la prueba de Neyman-Pearson, de hecho, en ocasiones es notablemente más grande bajo H_1 (ver Figura 3.5).

La conclusión es que el criterio equilibrado bayesiano es definitivamente más recomendado para optimizar las pruebas de etapas múltiples en el caso del proceso de Wiener que el criterio de H_0 solamente. Esto puede parecer sorprendente, porque, para observaciones i.i.d. y horizonte infinito, el criterio de H_0 produce la misma SPRT óptima (ver [Lorden, G., (1980)]) que el criterio bayesiano (ver [Wald, A., Wolfowitz, J., (1948)]).

Como se mencionó anteriormente (ver Nota 2.3), se puede aplicar la construcción de esta sección a pruebas bilaterales, las cuales que son muy importantes para las aplicaciones (ver [Jennison (1999)]). Para lograr esto, será necesario utilizar la siguiente expresión

$$“f_{\theta_1}^{(t|n)}(x^{(t|n)})” = \pi' z_{-1}(x_{s_n(t)}, s_n(t)) + \pi'' z_1(x_{s_n(t)}, s_n(t))$$

(por ejemplo (3.3)). Todas las expresiones que definen el paro óptimo se modifican respectivamente, y ya no serían tan simples como en (3.6), pero siguen siendo bastante tratables numéricamente. La gran diferencia con el caso unilateral que se vió anteriormente en esta sección, es que no existe una forma conocida de la prueba secuencial óptima en este caso bilateral y ni siquiera se sabe si dicha prueba existe. Esto sería tema de otro trabajo.

Un uso muy inmediato de los resultados que se obtuvieron en esta sección, es la construcción de pruebas de etapas múltiples asintóticamente óptimas para familias localmente asintóticamente normales (LAN) para observaciones i.i.d. Para las familias LAN (ver [Le Cam, L., Yang, G.L., 2000.]), la razón de probabilidad converge en distribución a la razón de probabilidad del proceso Wiener, lo que permite construir pruebas multi-etápicas discretas asintóticamente óptimas sobre la base de las pruebas multi-etápicas óptimas para el proceso Wiener (vea una construcción análoga para pruebas de dos etapas en [Novikov, A., 2005.]). La gran ventaja de las pruebas multi-etápicas es que para esta aproximación, solo se necesita la convergencia de distribuciones de dimensiones finitas, que se deriva directamente de la condición LAN. Es bien sabido que la condición LAN puede cumplirse no solo para observaciones i.i.d. (ver por ejemplo, [Roussas, G.G., 1972] para el caso de los procesos ergódicos de Markov), por lo que las aproximaciones del proceso Wiener tienen una naturaleza mucho más general que la del caso i.i.d. Esto brinda la oportunidad de aplicar las mismas pruebas multi-etápicas en tiempo discreto, basadas en las aproximaciones de Wiener para procesos mucho más generales que para observaciones i.i.d. En particular, es fácil imaginar que para algunos de ellos no habrá una prueba secuencial óptima conocida (como en el caso general de los procesos ergódicos de Markov, por ejemplo). Por supuesto, siempre existe la posibilidad de utilizar la SPRT de Wald, pero para establecer su óptimalidad asintótica se necesitan los resultados como el “principio de invariancia estadística” (ver [Greenwood, P.E., Shiryaev, A.N., 1985.] para el caso i.i.d) que no siempre está disponible para observaciones dependientes. Por lo tanto, la alternativa para tales casos será: “pruebas multi-etápicas asintóticamente óptimas” o “SPRT con propiedades óptimas cuestionables, incluso asintóticas”.

Otro aspecto de la metodología desarrollada está relacionado con procesos limitantes diferentes del proceso de Wiener. Se sabe (véase [Stadje, W., 1984.]) que procesos distintos del proceso de Wiener pueden aparecer en la distribución limitante de la razón de probabilidad, bajo la condición de contigüidad. Si este es el caso, ¿cómo debería ser la prueba secuencial óptima para el proceso limitante? ¿Cómo calcular sus características?, no siempre hay respuestas explícitas. Se puede encontrar un ejemplo en [Novikov, A., 1988.] donde se podría caracterizar la prueba secuencial óptima, para el proceso limitante, solo para un caso muy particular del modelo con saltos. Para tratar casos más generales, en su lugar se podrían usar pruebas de varias etapas para el proceso limitante. A partir de estos, se pueden derivar pruebas multi-etápicas asintóticamente óptimas para el problema original en tiempo discreto.

Concluyendo, se espera haber mostrado que las pruebas secuenciales multi-etápicas pueden ser una herramienta útil para muchos problemas de análisis secuencial.

3. Planteamiento general multietápico del proceso de Wiener con deriva lineal

Capítulo 4

El problema modificado de Kiefer-Weiss con distribuciones uniformes

Una vez que se tiene la teoría desarrollada para pruebas de hipótesis multietápicas, la cual fue expuesta en los capítulos 1 y 2 en el presente trabajo. Ahora se realizará una aplicación de esta teoría al problema de Kiefer-Weiss modificado, con distribuciones uniformes, por lo tanto los resultados que a continuación se desarrollan son de autoría propia y del asesor.

Como ya se ha dicho, en el presente capítulo se da una aplicación de la teoría general de paro óptimo con distribuciones uniformes. Aprovechando la simplicidad de estas distribuciones, se da solución al problema de Kiefer-Weiss modificado en el cual participan tres distribuciones de prueba cuyos parámetros son: θ , θ_0 y θ_1 . En el presente trabajo se considera a $\theta_0 = 0$, $0 < \theta_1 < 1$ y se permite que θ tome cualquier valor en \mathbb{R} . Debido a la libertad de θ , al aplicar la teoría de paro se generan diferentes resultados, cada uno viene plasmado en un Teorema, (que determina la estructura a partir de las funciones recursivas $V^{(t|n)}$), de éste se desprende un corolario que indica la forma de la regla de paro óptimo, y por último se hacen las formulaciones de los errores de tipo I , de tipo II y el costo promedio muestral.

Se inicia el trabajo con las definiciones de las distribuciones uniformes

$$f_{\theta_0}^{(t|n)}(x^{(t|n)}) = I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)}), \quad f_{\theta_1}^{(t|n)}(x^{(t|n)}) = I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)}) \quad (4.1)$$

y

$$f_{\theta}^{(t|n)}(x^{(t|n)}) = I_{A_{\theta}^{(t|n)}}(x^{(t|n)}) \quad (4.2)$$

siendo

$$\begin{aligned} A_{\theta_0}^{(t|n)} &= \{(x_1, \dots, x_{t_1}, x_{t_1+1}, \dots, x_{t_1+\dots+t_n}) : x_i \in A_{\theta_0}, i = 1, \dots, |t^{(n)}|\} \\ A_{\theta_1}^{(t|n)} &= \{(x_1, \dots, x_{t_1}, x_{t_1+1}, \dots, x_{t_1+\dots+t_n}) : x_i \in A_{\theta_1}, i = 1, \dots, |t^{(n)}|\} \\ A_{\theta}^{(t|n)} &= \{(x_1, \dots, x_{t_1}, x_{t_1+1}, \dots, x_{t_1+\dots+t_n}) : x_i \in A_{\theta}, i = 1, \dots, |t^{(n)}|\} \end{aligned} \quad (4.3)$$

con

$$A_{\theta_0} = [0, 1], \quad A_{\theta_1} = [\theta_1, \theta_1 + 1], \quad A_\theta = [\theta, 1 + \theta]. \quad (4.4)$$

Para simplificar la notación, se define

$$I_{A_3^{(t|n)}}(x^{(t|n)}) = I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)}) I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)}) \quad (4.5)$$

donde

$$A_3^{(t|n)} = \{(x_1, \dots, x_{t_1}, x_{t_1+1}, \dots, x_{t_1+\dots+t_n}) : x_i \in A_3, i = 1, \dots, |t^{(n)}|\} \quad (4.6)$$

y $A_3 = A_{\theta_0} \cap A_{\theta_1}$ con $0 < \theta_1 < 1$, (ver (4.4)).

A continuación se muestra una lista de los valores que θ puede tomar, siguiendo ese orden se analizarán los resultados que generan las funciones recursivas, $V^{(t|n)}$.

- $0 \leq \theta \leq \theta_1$.
- $\theta_1 < \theta < 1$.
- $\theta < 0$ or $1 < \theta$.
- Si $\theta = \theta_0$ or $\theta = \theta_1$, se tienen únicamente las dos distribuciones dadas en (4.1).

4.1. Decisión óptima para el problema modificado de Kiefer-Weiss

Observar que en el teorema de la regla de decisión (Teorema 2.2, sección 2.1), solo participan las indicadoras $f_{\theta_0}^{(t|n)}$ y $f_{\theta_1}^{(t|n)}$ dadas en (4.1), por lo cual, se puede calcular la regla de decisión sin alguna otra condición.

Para lo que sigue, se realiza la siguiente aclaración. Sean $x, y \in \mathbb{R}$, entonces

$$I_{\{x < y\}}^{(\leq)} = \begin{cases} 1 & \text{si y sólo si } x < y \\ & (\leq) \\ 0 & \text{c.o.c.} \end{cases} \quad (4.7)$$

Las reglas de decisión óptimas para distribuciones uniformes dadas en (4.1), son

$$I_{\{\lambda_0 I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)}) < \lambda_1 I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)})\}} \quad (4.8)$$

$$\leq \phi_t \leq$$

$$I_{\{\lambda_0 I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)}) \leq \lambda_1 I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)})\}} \quad (4.9)$$

con $n = 1, 2, \dots, k$.

Proposición 4.1. *La indicadora dada en (4.8) es igual a*

$$\phi_{t|n} = I_{A_{\theta_1}^{(t|n)}}, \quad \text{con } n = 1, 2, \dots, k. \quad (4.10)$$

Demostración. Se debe probar que (4.8) y (4.10) son iguales.

- Si $I_{A_{\theta_1}^{(t|n)}} = 1 = \phi_{t|n}$; para determinar el valor de la indicadora dada en (4.8) se consideran los siguientes dos casos.
 - Si $I_{A_{\theta_1}^{(t|n)} \setminus A_{\theta_0}^{(t|n)}} = 1$, se tiene que $I_{\{0 < \lambda_1\}}$ y toma el valor de 1 debido a que se cumple la desigualdad estricta de (4.7).
 - Si $I_{A_3^{(t|n)}} = 1$, se tiene que $I_{\{\lambda_0 < \lambda_1\}}$ toma el valor de 1 debido a que se cumple la desigualdad estricta de (4.7).
- Si $I_{A_{\theta_1}^{(t|n)}} = 0 = \phi_{t|n}$, y para (4.8)
 - Si $I_{A_{\theta_0}^{(t|n)} \setminus A_{\theta_1}^{(t|n)}} = 1$, entonces $I_{\{\lambda_0 < 0\}} = 0$ toma el valor de 0 porque no se cumple la desigualdad estricta de (4.7).
 - Si $I_{A_{\theta_0}^{(t|n)}} = 0 = I_{A_{\theta_1}^{(t|n)}}$ evaluando esto en la indicadora dada en (4.8) se obtiene $I_{\{\lambda_0 < 0\}}$ y esta vale cero porque se supone que $\lambda_0 > 0$, y es claro que la desigualdad estricta de (4.7) no se cumple.

Por lo tanto (4.8) y (4.10) son iguales. □

Proposición 4.2. *La indicadora dada en (4.9) es igual a*

$$\phi_{t|n} = I_{(A_{\theta_0}^{(t|n)} \setminus A_{\theta_1}^{(t|n)})^c}, \text{ con } n = 1, 2, \dots, k. \quad (4.11)$$

Demostración. Se debe probar que (4.9) y (4.11) son iguales.

- Si $I_{(A_{\theta_0}^{(t|n)} \setminus A_{\theta_1}^{(t|n)})^c} = 1 = \phi_{t|n}$, analizando por partes esta indicadora para determinar el valor de (4.9)
 - Si $I_{A_3^{(t|n)}} = 1$ y la indicadora dada en (4.9) es $I_{\{\lambda_0 \leq \lambda_1\}}$ la cual vale 1 sólo si se cumple (4.7).
 - Si $I_{(A_{\theta_1}^{(t|n)} \setminus A_{\theta_0}^{(t|n)})} = 1$ la indicadora dada en (4.9) es $I_{\{0 \leq \lambda_1\}}$ esta vale 1 porque se satisface (4.7).
 - Si $I_{(A_{\theta_0}^{(t|n)} \cup A_{\theta_1}^{(t|n)})^c} = 1$ esto implica que $I_{A_{\theta_0}^{(t|n)}} = 0 = I_{A_{\theta_1}^{(t|n)}}$, la indicadora dada en (4.9) es $I_{\{0 \leq 0\}}$ y como en particular se cumple la igualdad en (4.7) entonces el valor de (4.9) es 1.
- Si $I_{(A_{\theta_0}^{(t|n)} \setminus A_{\theta_1}^{(t|n)})^c} = 0 = \phi_{t|n}$, entonces $I_{\{\lambda_0 \leq 0\}}$ y esto contradice el supuesto de $\lambda_0 > 0$, por lo que el valor de (4.9) es 0.

Por lo tanto (4.9) y (4.11) son iguales. □

Con respecto a los errores, viendo a (1.15) y a los resultados de las proposiciones 4.1 y 4.2, se puede calcular la fórmula del error tipo II , aún sin tener las expresiones explícitas para los tiempos de paro óptimo.

4.2. Determinación del valor del error tipo II

Proposición 4.3. *El error tipo II para un proceso con las distribuciones uniformes dadas en (4.1) es cero.*

Demostración. Para esto basta evaluar (1.15) en (4.10) y (4.11), respectivamente.

- Sustituyendo (4.10) en (1.15),

$$\beta(\psi_{t|n}, \phi_{t|n}) = \sum_{n=1}^k E_1 s_{t|n}^\psi (1 - \phi_{t|k}) = \sum_{n=1}^k E_1 s_{t|n}^\psi I_{A_{\theta_1}^c}^{(t|k)} = 0.$$

- Para el otro caso, se sustituye (4.11) en (1.15)

$$\beta(\psi_{t|n}, \phi_{t|n}) = \sum_{n=1}^k E_1 s_{t|n}^\psi (1 - \phi_{t|k}) = \sum_{n=1}^k E_1 s_{t|n}^\psi I_{A_{\theta_0}^{(t|k)} \setminus A_{\theta_1}^{(t|k)}} = 0.$$

Observar que en cualquiera de los dos casos analizados, el tiempo de paro no influye en el cálculo del valor del error tipo II, es decir, $\beta = 0$ para cualquiera que sea la “forma” del tiempo de paro. □

4.3. Paro óptimo cuando $0 \leq \theta \leq \theta_1$

Como ya se ha visto al inicio del presente capítulo, el valor de θ (ver (4.2)), puede tomar diferentes valores en la recta real. Para lo que sigue se determinarán las funciones de paro a partir del teorema 2.3 (Sección 2.2).

Teorema 4.1. *Considerar $0 \leq \theta \leq \theta_1$. Las funciones $V^{t|n}$ definidas en (2.15) evaluadas en las distribuciones uniformes dadas en (4.2) y (4.1), son descritas por*

$$V^{t|n}(x^{(t|n)}) = v_{t|n} I_{A_3}^{(t|n)}(x^{(t|n)}) \quad (4.12)$$

con $v_{t|k} = \lambda$, y $v_{t|n}$ definidas recursivamente por $n = k - 1, \dots, 1$, como

$$v_{t|n} = \min\{\lambda, c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}\}. \quad (4.13)$$

Demostración. La prueba es por inducción hacia atrás. Verifiquemos que (4.12) se cumple para $n = k$, de (2.15) se tiene que $V^{(t|k)} = I^{(t|k)}$ para $n = k$, por lo cual

$$\begin{aligned} V^{(t|n)}(x^{(t|n)}) &= \min\{\lambda_0 I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)}), \lambda_1 I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)})\} \\ &= \min\{\lambda_0, \lambda_1\} I_{A_3}^{(t|n)}(x^{(t|n)}) \\ &= \lambda I_{A_3}^{(t|n)}(x^{(t|n)}) \\ &= v_{t|k} I_{A_3}^{(t|n)}(x^{(t|n)}) \end{aligned} \quad (4.14)$$

siendo

$$\min\{\lambda_0, \lambda_1\} = \lambda = v_{t|k}. \quad (4.15)$$

En la hipótesis de inducción se supone que (4.12) se cumple para $n = m < k$

$$V^{t|m}(x^{(t|m)}) = v_{t|m} I_{A_3^{(t|m)}}(x^{(t|m)}). \quad (4.16)$$

Ahora se probará que (4.12) se cumple para $n = m - 1 < k$, es decir, por demostrar que

$$V^{t|m-1}(x^{(t|m-1)}) = v_{t|m-1} I_{A_3^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.17)$$

Para probar (4.17) se sustituye (4.16) en (2.15), así que

$$\begin{aligned} V^{(t|m-1)}(x^{(t|m-1)}) &= \text{mín}\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ &\quad + \int \cdots \int v_{t|m} I_{A_3^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m}\} \end{aligned}$$

integrando

$$= \text{mín}\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) + v_{t|m}(1 - \theta_1)^{t_m} I_{A_3^{(t|m-1)}}(x^{(t|m-1)})\}. \quad (4.18)$$

Análisis de (4.18).

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 0$, se sigue directamente que

$$V^{(t|m-1)}(x^{(t|m-1)}) = 0. \quad (4.19)$$

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 1$ y como $A_3^{(t|m-1)} \subset A_\theta^{(t|m-1)}$ entonces $I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) = 1$ por lo cual

$$V^{(t|m-1)}(x^{(t|m-1)}) = \text{mín}\{\lambda, c_m(t) + v_{t|m}(1 - \theta_1)^{t_m}\} I_{A_3^{(t|m-1)}}(x^{(t|m-1)})$$

pero la constante en esta igualdad es $v_{t|m-1}$ definida en (4.13), por lo cual

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1} I_{A_3^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.20)$$

A partir de (4.19) y (4.20) se tiene (4.17). Por lo tanto (4.12) se cumple y con esto concluye la demostración. \square

A continuación se determinan las funciones de paro.

Sea $0 \leq \theta \leq \theta_1$, entonces las reglas de paro óptimas bajo las condiciones del Teorema 4.1 son

$$\begin{aligned} I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) < c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}} I_{A_3^{(t|n)}}(x^{(t|n)})\}} & \quad (4.21) \\ & \leq \psi_{t|n} \leq \end{aligned}$$

$$I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) \leq c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}} I_{A_3^{(t|n)}}(x^{(t|n)})\}} \quad (4.22)$$

para todo $n = 1, \dots, k - 1$, y $v_{t|n}$ están definidas en (4.13).

Proposición 4.4. *La función indicadora dada en (4.21), es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} & \text{si } n < k \text{ y } n \notin N_0 \\ I_{A_\theta^{(t|n)}} & \text{si } n \in N_0 \\ 1 & \text{si } n = k \end{cases} \quad (4.23)$$

con $n = 1, \dots, k$, y siendo

$$N_0 = \{0 < n < k : \lambda < c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}\}. \quad (4.24)$$

Demostración. Se debe verificar que (4.21) y (4.23) son iguales.

- Si $I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} = 1 = \psi_{t|n}$, para (4.21) se tiene $I_{\{0 < c_{n+1}(t)\}}$ y como se satisface (4.7), entonces el valor de la indicadora es 1.
- Si $I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} = 0 = \psi_{t|n}$, y el valor de $I_{\{0 < 0\}}$ es cero porque no se cumple (4.7).
- El caso $I_{A_\theta^{(t|n)}} = 0$ y $I_{A_3^{(t|n)}} = 1$, no es posible porque $A_3^{(t|n)} \subset A_\theta^{(t|n)}$.
- Si $I_{A_\theta^{(t|n)}} = 1 = \psi_{t|n}$, y si $n \in N_0$, se tiene que $I_{\{\lambda < c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}\}}$ y su valor es 1 porque se cumple (4.7).

Por lo tanto, (4.21) y (4.23) son iguales. \square

Proposición 4.5. *La indicadora dada en (4.22) es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_3^c(t|n)} & \text{si } n < k \text{ y } n \notin N_1 \\ 1 & \text{si } n \in N_1 \cup \{k\} \end{cases} \quad (4.25)$$

con $n = 1, \dots, k$, siendo

$$N_1 = \{0 < n < k : \lambda \leq c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}\}. \quad (4.26)$$

Demostración. Se necesita probar que (4.22) y (4.25) son iguales.

- Si $I_{A_\theta^{(t|n)}} = 1$ y $I_{A_3^{(t|n)}} = 0$ entonces $I_{A_3^c(t|n)} = 1 = \psi_{t|n}$, y para (4.22) se tiene $I_{\{0 \leq c_{n+1}(t)\}}$ cuyo valor es 1 porque se satisface (4.7).
- Si $I_{A_\theta^{(t|n)}} = 0$ y $I_{A_3^{(t|n)}} = 0$ entonces $I_{A_3^c(t|n)} = 1 = \psi_{t|n}$, y la indicadora (4.22) es $I_{\{0 \leq 0\}}$ cuyo valor es 1, porque en particular se cumple la igualdad en (4.7).
- Si $I_{A_\theta^{(t|n)}} = 0$ y $I_{A_3^{(t|n)}} = 1$, este caso no es posible porque $A_3^{(t|n)} \subset A_\theta^{(t|n)}$.
- Si $I_{A_\theta^{(t|n)}} = 1$ y $I_{A_3^{(t|n)}} = 1$ dependiendo sobre la localización de n se tienen los siguientes dos casos para (4.22).
 - Si $n \notin N_1$ entonces $I_{A_3^c(t|n)} = 0 = \psi_{t|n}$, para la indicadora (4.22) se tiene $I_{\{\lambda \leq c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}\}}$ la cual toma el valor de cero porque $n \notin N_1$ y (4.7) no se cumple.
 - Si $n \in N_1$ entonces $\psi_{t|n} = 1$, y el valor de (4.22) es 0, porque si $n \in N_1$ entonces $\lambda > c_{n+1}(t) + v_{t|n+1}(1 - \theta_1)^{t_{n+1}}$ y se satisface a (4.7).

Por lo tanto, (4.22) y (4.25) son iguales. \square

4.3.1. Error tipo I y número promedio muestral

Antes de iniciar los cálculos, se hacen las siguientes definiciones. A partir de (4.24), se tiene que

$$n_0 = \begin{cases} \min(N_0) & \text{si } N_0 \neq \emptyset \\ k & \text{si } N_0 = \emptyset \end{cases} \quad (4.27)$$

y de (4.26)

$$n_1 = \begin{cases} \min(N_1) & \text{si } N_1 \neq \emptyset \\ k & \text{si } N_1 = \emptyset. \end{cases} \quad (4.28)$$

Cálculo del error tipo I con (4.10) y (4.23). Para $n < k$ y $n \notin N_0$, se tiene $\psi_{t|n} = I_{(A_\theta^{(t|n)} \setminus A_3^{(t|n)})}$ y $\psi_{t|n_0} = I_{A_\theta^{(t|n_0)}}$, sustituyendo esto en (1.14)

$$\begin{aligned} \alpha(\psi_{t|n}, \phi_{t|n}) &= E_0 [I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})} I_{A_{\theta_1}^{(t|1)}} + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} I_{(A_\theta^{(t|2)} \setminus A_3^{(t|2)})} I_{A_{\theta_1}^{(t|2)}} \\ &\quad + \cdots + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_3^{(t|n_0-1)})^c} I_{A_\theta^{(t|n_0)}} I_{A_{\theta_1}^{(t|n_0)}} \\ &\quad + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_3^{(t|n_0-1)})^c} I_{A_\theta^{(t|n_0)}} I_{(A_\theta^{(t|n_0+1)} \setminus A_3^{(t|n_0+1)})} I_{A_{\theta_1}^{(t|n_0+1)}} \\ &\quad + \cdots + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_3^{(t|n_0-1)})^c} \\ &\quad I_{A_\theta^{(t|n_0)}} I_{(A_\theta^{(t|n_0+1)} \setminus A_3^{(t|n_0+1)})^c} \cdots I_{(A_\theta^{(t|k-1)} \setminus A_3^{(t|k-1)})^c} 1 I_{A_{\theta_1}^{(t|k)}}] \end{aligned}$$

como $E_0 [I_{(A_\theta^{(t|n)} \setminus A_3^{(t|n)})} I_{A_{\theta_1}^{(t|n)}}] = 0$ y haciendo los cálculos, esto es

$$= (1 - \theta_1)^{t_1} (1 - \theta_1)^{t_2} \cdots (1 - \theta_1)^{t_{n_0-1}} (1 - \theta_1)^{t_{n_0}}$$

por lo tanto, el error tipo I es

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{\sum_{i=1}^{n_0} t_i}. \quad (4.29)$$

Cálculo del número promedio muestral. A partir de (4.23), $\psi_{t|n} = I_{(A_\theta^{(t|n)} \setminus A_3^{(t|n)})}$ y $\psi_{t|n_0} = I_{A_\theta^{(t|n_0)}}$ sustituyendo esto en (1.17)

$$\begin{aligned} K(\theta; \psi_{t|n}) &= E_\theta [C_1(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})} + C_2(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} I_{(A_\theta^{(t|2)} \setminus A_3^{(t|2)})} \\ &\quad + \cdots + C_{n_0}(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_3^{(t|n_0-1)})^c} I_{A_\theta^{(t|n_0)}} \\ &\quad + C_{n_0+1}(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_3^{(t|n_0-1)})^c} I_{A_\theta^{(t|n_0)}} \psi_{t|n_0+1} \\ &\quad + \cdots + C_k(t) I_{(A_\theta^{(t|1)} \setminus A_{\theta_1}^{(t|1)})^c} \cdots I_{(A_\theta^{(t|n_0-1)} \setminus A_{\theta_1}^{(t|n_0-1)})^c} I_{A_\theta^{(t|n_0)}} (1 - \psi_{t|n_0+1}) \\ &\quad \cdots (1 - \psi_{t|n_0+1}) \psi_{t|k} \end{aligned}$$

haciendo los cálculos

$$\begin{aligned} &= C_1(t) (1 - (1 - \theta_1)^{t_1}) + C_2(t) (1 - \theta_1)^{t_1} (1 - (1 - \theta_1)^{t_2}) + \cdots + \\ &\quad C_{n_0-1}(t) (1 - \theta_1)^{t_1 + \cdots + t_{n_0-2}} (1 - (1 - \theta_1)^{t_{n_0-1}}) + C_{n_0}(t) (1 - \theta_1)^{\sum_{j=1}^{n_0-1} t_j} \end{aligned}$$

como $C_n(t) = c_1(t) + \cdots + c_n(t)$ sustituyendo y haciendo aritmética, se concluye que

$$K(\theta; \psi_{t|n}) = c_1(t) + \sum_{i=2}^{n_0} c_i(t) (1 - \theta_1)^{\sum_{j=1}^{i-1} t_j}. \quad (4.30)$$

4. El problema modificado de Kiefer-Weiss con distribuciones uniformes

Nota 4.1. Debido a (4.29), (4.30) y a la proposición 4.3, el mínimo de (4.25) es

$$c_1(t) + \sum_{i=2}^{n_0} c_i(t)(1 - \theta_1)^{\sum_{j=1}^{i-1} t_j} + \lambda_0(1 - \theta_1)^{\sum_{i=1}^{n_0} t_i}. \quad (4.31)$$

A continuación se determina el error tipo I y el número promedio con las funciones de decisión y de paro dadas en (4.11) y (4.25), respectivamente.

Cálculo del error tipo I . Debido a (4.25) $\psi_{t|n} = I_{A_3^c(t|n)}$ para $n < n_1 < k$, sustituyendo esto y además (4.11) en (1.14), se tiene

$$\begin{aligned} \alpha(\psi_{t|n}, \phi_{t|n}) &= E_0[I_{A_3^c(t|1)} I_{(A_{\theta_0}^{(t|1)} \setminus A_{\theta_1}^{(t|1)})^c} + I_{A_3^{(t|1)}} I_{A_3^c(t|2)} I_{(A_{\theta_0}^{(t|2)} \setminus A_{\theta_1}^{(t|2)})^c} \\ &\quad + \cdots + I_{A_3^{(t|1)}} \cdots I_{(A_3^{(t|n_1-1)})} 1 I_{(A_{\theta_0}^{(t|n_1)} \setminus A_{\theta_1}^{(t|n_1)})^c} \\ &\quad + I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|n_1-1)}} I_{(A_{\theta_1}^{(t|n_1-1)} \setminus A_{\theta_1}^{(t|n_1-1)})^c} 0 \psi_{t|n_1+1} \phi_{t|n_1+1} \\ &\quad + \cdots + I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|n_1-1)}} 0 (1 - \psi_{t|n_1+1}) \cdots (1 - \psi_{t|k-1}) \psi_{t|k} \phi_{t|k}] \\ &= (1 - \theta_1)^{t_1} \cdots + (1 - \theta_1)^{t_k} \end{aligned}$$

por lo tanto

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{t_1 + \cdots + t_{n_1}}. \quad (4.32)$$

Para el número promedio se sustituye (4.25) en (1.17), por lo cual

$$\begin{aligned} K(\theta; \psi_{t|n}) &= E_\theta[C_1(t) I_{A_3^c(t|1)} + C_2(t) I_{A_3^{(t|1)}} I_{A_3^c(t|2)} + \cdots + C_{n_1}(t) I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|n_1-1)}} I_{A_3^c(t|n_1)} \\ &\quad + C_{n_1+1}(t) I_{A_3^{(t|1)}} \cdots I_{(A_3^{(t|n_1-1)})} 0 \psi_{t|n_1+1} \\ &\quad + \cdots + C_k(t) I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|n_1-1)}} 0 (1 - \psi_{t|n_1+1}) \cdots (1 - \psi_{t|k-1}) \psi_{t|k} \end{aligned}$$

haciendo los cálculos

$$\begin{aligned} &= C_1(t)(1 - (1 - \theta_1)^{t_1}) + C_2(t)(1 - \theta_1)^{t_1}(1 - (1 - \theta_1)^{t_2}) + \cdots + \\ &\quad C_{n_1-1}(t)(1 - \theta_1)^{t_1 + \cdots + t_{n_1-2}}(1 - (1 - \theta_1)^{t_{n_1-1}}) + C_{n_1}(t)(1 - \theta_1)^{\sum_{j=1}^{n_1-1} t_j} \end{aligned}$$

sustituyendo $C_n(t)$ y simplificando, se llega a que

$$K(\theta; \psi_{t|n}) = c_1(t) + \sum_{i=2}^{n_1} c_i(t)(1 - \theta_1)^{\sum_{j=1}^{i-1} t_j}. \quad (4.33)$$

Concluyendo, debido a los resultados obtenidos en (4.32), (4.33) y además de la proposición 4.3, se tiene el valor mínimo de (4.25) es

$$c_1(t) + \sum_{i=2}^{n_1} c_i(t)(1 - \theta_1)^{\sum_{j=1}^{i-1} t_j} + \lambda_0(1 - \theta_1)^{\sum_{i=1}^{n_1} t_i}. \quad (4.34)$$

Una vez concluido el análisis, es necesario aclarar que hay dos casos particulares del teorema 4.1.

- Cuando $\theta = 0$, se tiene que $I_{A_\theta^{(t|n)}}(x^{(t|n)}) = I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)})$.
- Por otra parte, si $\theta = \theta_0$, entonces $I_{A_\theta^{(t|n)}}(x^{(t|n)}) = I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)})$.

En cualquiera de los dos casos mencionados las indicadoras que son: $I_{A_{\theta_1}^{(t|n)}}(x^{(t|n)})$ y $I_{A_{\theta_0}^{(t|n)}}(x^{(t|n)})$.

4.4. Paro óptimo cuando $\theta_1 < \theta < 1$

Teorema 4.2. *Sea $\theta_1 < \theta < 1$, entonces las funciones recursivas $V^{(t|n)}$ con las distribuciones uniformes dadas en (4.2) y (4.1) para $n = k, \dots, 1$, son*

$$V^{(t|n)}(x^{(t|n)}) = v_{t|n}^* I_{A_3^{(t|n)} \setminus A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n} I_{A_3^{(t|n)} \cap A_\theta^{(t|n)}}(x^{(t|n)}) \quad (4.35)$$

si $n = k$ se tiene que $v_{t|k} = v_{t|k}^* = \lambda$; y para $n = 1, \dots, k-1$

$$v_{t|n}^* = \lambda(1 - \theta_1)^{\sum_{i=n}^{k-1} t_{i+1}} \quad (4.36)$$

recursivamente para $n = k-1, \dots, 1$,

$$v_{t|n} = \min\{\lambda, c_{n+1}(t) + v_{t|n+1}^*(1 - \theta_1)^{t_{n+1}} + [v_{t|n+1} - v_{t|n+1}^*](1 - \theta)^{t_{n+1}}\}. \quad (4.37)$$

Demostración. Esta se hace por inducción hacia atrás.

A continuación se comprueba (4.35) para $n = k$

$$\begin{aligned} V^{(t|k)}(x^{(t|k)}) &= \min\{\lambda_0 I_{A_{\theta_0}^{(t|k)}}(x^{(t|k)}), \lambda_1 I_{A_{\theta_1}^{(t|k)}}(x^{(t|k)})\} \\ &= \lambda I_{A_3^{(t|k)}}(x^{(t|k)}) \\ &= \lambda \left(I_{A_3^{(t|k)} \setminus A_\theta^{(t|k)}}(x^{(t|k)}) + I_{A_3^{(t|k)} \cap A_\theta^{(t|k)}}(x^{(t|k)}) \right) \\ &= v_{t|k}^* I_{A_3^{(t|k)} \setminus A_\theta^{(t|k)}}(x^{(t|k)}) + v_{t|k} I_{A_3^{(t|k)} \cap A_\theta^{(t|k)}}(x^{(t|k)}). \end{aligned}$$

Supóngase que (4.35) es cierta para algún $n = m \leq k$, es decir

$$V^{(t|m)}(x^{(t|m)}) = v_{t|m}^* I_{A_3^{(t|m)} \setminus A_\theta^{(t|m)}}(x^{(t|m)}) + v_{t|m} I_{A_3^{(t|m)} \cap A_\theta^{(t|m)}}(x^{(t|m)}). \quad (4.38)$$

Por demostrar que (4.35) se cumple para $n = m-1$, es decir

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1}^* I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) + v_{t|m-1} I_{A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.39)$$

Por (2.15) y la hipótesis de inducción (4.38)

$$\begin{aligned} V^{(t|m-1)}(x^{(t|m-1)}) &= \min\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ &\quad + \int \cdots \int v_{t|m}^* I_{A_3^{(t|m)} \setminus A_\theta^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m} \\ &\quad + \int \cdots \int v_{t|m} I_{A_3^{(t|m)} \cap A_\theta^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m}\}. \end{aligned} \quad (4.40)$$

Análisis de la indicadora de la primera integral de (4.40), observar que

$$A_3^{(t|m)} \setminus A_\theta^{(t|m)} = A_3^{(t|m-1)} \times A_3^{(t|m)} \setminus A_\theta^{(t|m-1)} \times A_\theta^{(t|m)} \quad (4.41)$$

aplicando el teorema (Producto cartesiano de la diferencia) (ver [Clark (1971)]), se tiene que (4.41) es igual a

$$= \left(A_3^{(t|m-1)} \times (A_3^{(t|m)} \setminus A_\theta^{(t|m)}) \right) \cup \left((A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}) \times A_3^{(t|m)} \right) \quad (4.42)$$

4. El problema modificado de Kiefer-Weiss con distribuciones uniformes

y

$$\begin{aligned} & \left(A_3^{(t|m-1)} \times (A_3^{(t|m)} \setminus A_\theta^{(t|m)}) \right) \cap \left((A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}) \times A_3^{(t|m)} \right) \\ & = (A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}) \times (A_3^{(t|m)} \setminus A_\theta^{(t|m)}). \end{aligned} \quad (4.43)$$

Debido a (4.42) y (4.43) la primera integral de (4.40) es

$$\begin{aligned} & = V_{t|m}^* \left[\int I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) I_{A_3^{(t|m)} \setminus A_\theta^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m} \right. \\ & \quad + \int I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) I_{A_3^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m} \\ & \quad \left. - \int I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) I_{A_3^{(t|m)} \setminus A_\theta^{(t|m)}}(x^{(t|m)}) dx_{t_1+\dots+t_{m-1}+1} \cdots dx_{t_1+\dots+t_m} \right] \end{aligned}$$

integrando

$$\begin{aligned} & = v_{t|m}^* \left[\left[(1 - \theta_1)^{t_m} - (1 - \theta)^{t_m} \right] I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) + (1 - \theta_1)^{t_m} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right. \\ & \quad \left. - \left[(1 - \theta_1)^{t_m} - (1 - \theta)^{t_m} \right] I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right] \end{aligned}$$

simplificando

$$= v_{t|m}^* \left[\left[(1 - \theta_1)^{t_m} - (1 - \theta)^{t_m} \right] I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) + (1 - \theta)^{t_m} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right]. \quad (4.44)$$

En la primera integral de (4.40) se sustituye (4.44) y calculando la segunda integral

$$\begin{aligned} & V^{(t|m-1)}(x^{(t|m-1)}) = \min \left\{ \lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right. \\ & \quad + v_{t|m}^* \left[\left[(1 - \theta_1)^{t_m} - (1 - \theta)^{t_m} \right] I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) + (1 - \theta)^{t_m} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right] \\ & \quad \left. + v_{t|m} (1 - \theta)^{t_m} I_{A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \right\}. \end{aligned} \quad (4.45)$$

Análisis de (4.45). Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 0$, se sigue que

$$V^{(t|m-1)}(x^{(t|m-1)}) = 0. \quad (4.46)$$

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 1$ y $I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) = 0$, entonces

$$\begin{aligned} & V^{(t|m-1)}(x^{(t|m-1)}) = \min \left\{ \lambda, v_{t|m}^* (1 - \theta_1)^{t_m} \right\} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ & = \min \left\{ \lambda, \lambda (1 - \theta_1)^{\sum_{i=m}^k t_{i+1}} (1 - \theta_1)^{t_m} \right\} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ & = \lambda (1 - \theta_1)^{\sum_{i=m-1}^k t_{i+1}} I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \end{aligned}$$

pero $v_{t|m-1}^* = \lambda (1 - \theta_1)^{\sum_{i=m-1}^k t_{i+1}}$ (ver (4.36)).

Si $x^{(t|m-1)} \in A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}$, se tiene que

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1}^* I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.47)$$

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 1$ y $I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) = 1$, entonces

$$V^{(t|m-1)}(x^{(t|m-1)}) = \min \left\{ \lambda, c_m(t) + v_{t|m}^* \left[(1 - \theta_1)^{t_m} - (1 - \theta)^{t_m} \right] + v_{t|m}(1 - \theta)^{t_m} \right\} I_{A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}}(x^{(t|m-1)})$$

la constante en la última igualdad es $v_{t|m-1}$ (ver (4.37)). Por lo tanto, si $x^{(t|m-1)} \in A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}$, entonces

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1} I_{A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.48)$$

Concluyendo, debido a (4.46), (4.47) y (4.48)

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1}^* I_{A_3^{(t|m-1)} \setminus A_\theta^{(t|m-1)}}(x^{(t|m-1)}) + v_{t|m-1} I_{A_3^{(t|m-1)} \cap A_\theta^{(t|m-1)}}(x^{(t|m-1)})$$

y esto es precisamente (4.39). \square

Determinación de la regla de paro del teorema 4.2, para esto se analizan los valores de $v_{t|n}^*$ y $v_{t|n}$, por (4.36), se tiene

$$v_{t|n}^* < \lambda, \text{ para todo } n = 1, \dots, k-1. \quad (4.49)$$

Por otra parte, el valor mínimo de $v_{t|n}$ (ver (4.37)) puede no ser determinado porque en éste participa $v_{t|n+1}$, el cual también contiene otro valor mínimo. Esto se muestra a continuación

$$v_{t|n} = \min \left\{ \lambda, c_{n+1}(t) + \lambda(1 - \theta_1)^{\sum_{i=n+1}^{k-1} t_{i+1}} (1 - \theta_1)^{t_{n+1}} + [\min \{ \lambda, c_{n+2}(t) + v_{t|n+2}^* (1 - \theta_1)^{t_{n+2}} + [v_{t|n+2} - v_{t|n+2}^*] (1 - \theta)^{t_{n+2}} \} - \lambda(1 - \theta_1)^{\sum_{i=n+1}^{k-1} t_{i+1}}] (1 - \theta)^{t_{n+1}} \right\}.$$

Debido a esto, el valor de (4.35) no se puede determinar explícitamente y la regla de paro (2.17) no se puede obtener en general.

Una forma para determinar el mínimo de $v_{t|n}$ es considerar cierta configuración de $c_{n+1}(t)$, esto es, a partir de (4.37) hacer

$$v_{t|n} = c_{n+1}(t) + v_{t|n+1}^* (1 - \theta_1)^{t_{n+1}} + [v_{t|n+1} - v_{t|n+1}^*] (1 - \theta)^{t_{n+1}} \quad (4.50)$$

con $n = k-1, \dots, 1$.

Bajo las condiciones sobre el costo, se da el siguiente resultado.

Sea $\theta_1 < \theta < 1$, entonces la regla de paro está dada por

$$I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) < c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1}^* I_{A_3^{(t|n)} \setminus A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1} I_{A_3^{(t|n)} \cap A_\theta^{(t|n)}}(x^{(t|n)})\}} \quad (4.51)$$

$$\leq \psi_{t|n} \leq$$

$$I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) \leq c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1}^* I_{A_3^{(t|n)} \setminus A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1} I_{A_3^{(t|n)} \cap A_\theta^{(t|n)}}(x^{(t|n)})\}} \quad (4.52)$$

con $v_{t|n+1}^*$ y $v_{t|n+1}$ definidas en (4.36) y (4.50) respectivamente.

4. El problema modificado de Kiefer-Weiss con distribuciones uniformes

Proposición 4.6. *La indicadora dada en (4.51) es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} & \text{si } n < k \\ 1 & \text{si } n = k \end{cases} \quad (4.53)$$

para todo $n = 1, \dots, k$.

Demostración. Se debe probar que (4.51) y (4.53) son iguales.

- Si $I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} = 1 = \psi_{t|n}$, para (4.51) si $I_{A_3^{(t|n)}} = 0$ y $I_{A_\theta^{(t|n)}} = 1$, y la indicadora (4.51) vale 1, porque $0 < c_{n+1}(t)$ y se cumple (4.7).
- Si $I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}} = 0 = \psi_{t|n}$, para (4.51)
 - Si $I_{A_3^{(t|n)}} = 1$ y $I_{A_\theta^{(t|n)}} = 0$ evaluando esto en la indicadora (4.51) se tiene $\lambda < v_{t|n}^*$ pero esto contradice (4.49), por lo tanto el valor de (4.51) es cero.
 - Si $I_{A_3^{(t|n)}} = 0$ y $I_{A_\theta^{(t|n)}} = 0$, y el valor de (4.51) es cero debido a que no se cumple la desigualdad estricta de (4.7).
 - Si $I_{A_3^{(t|n)}} = 1$ y $I_{A_\theta^{(t|n)}} = 1$; se tiene que el valor de $I_{\{\lambda < c_{n+1}(t) + v_{t|n}\}}$ es 0, porque en (4.50) se supone que $\lambda > c_{n+1}(t) + v_{t|n}$ y no se cumple la desigualdad estricta de (4.7).

Por lo tanto, (4.51) y (4.53) son iguales. □

Proposición 4.7. *La indicadora dada en (4.52) es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_3^c(t|n)} & \text{si } n < k \\ 1 & \text{si } n = k \end{cases} \quad (4.54)$$

para todo $n = 1, \dots, k$.

Demostración. Se debe probar que (4.52) y (4.54) son iguales.

- Si $I_{A_3^c(t|n)} = 1 = \psi_{t|n}$, para (4.52)
 - Si $I_{A_3^{(t|n)}} = 0$ y $I_{A_\theta^{(t|n)}} = 1$ se tiene $I_{\{0 \leq c_{n+1}(t)\}}$ y esto vale 1 porque se cumple la desigualdad de (4.7).
 - Si $I_{A_3^{(t|n)}} = 0 = I_{A_\theta^{(t|n)}}$ y el valor de (4.52) es 1 porque ambos términos de la indicadora son cero y en particular se cumple la igualdad de (4.7).
- $I_{A_3^c(t|n)} = 0 = \psi_{t|n}$, para (4.52)
 - Si $I_{A_3^{(t|n)}} = 1$ y $I_{A_\theta^{(t|n)}} = 0$ evaluando esto en la indicadora (4.51) se tiene $\lambda < v_{t|n}^*$ pero esto contradice (4.49), por lo tanto, el valor de (4.51) es cero.
 - Si $I_{A_3^{(t|n)}} = 1$ y $I_{A_\theta^{(t|n)}} = 1$ se obtiene $I_{\{\lambda \leq c_{n+1}(t) + v_{t|n}\}}$ esto vale 0 debido a que se contradice (4.50) y no se cumple (4.7).

Por lo tanto, las expresiones (4.52) y (4.54) son iguales. □

4.4.1. Error tipo I y número promedio muestral

Cálculo del error tipo I con (4.10) y (4.53). Para $n < k$ se tiene que $\psi_{t|n} = I_{(A_\theta^{(t|n)} \setminus A_3^{(t|n)})}$ sustituyendo esto en (1.14)

$$\begin{aligned} \alpha(\psi_{t|n}, \phi_{t|n}) &= E_0[I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})} I_{A_{\theta_1}^{(t|1)}} + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} I_{(A_\theta^{(t|2)} \setminus A_3^{(t|2)})} I_{A_{\theta_1}^{(t|2)}} \\ &\quad + \cdots + I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|k-1)} \setminus A_3^{(t|k-1)})^c} \mathbf{1}_{A_{\theta_1}^{(t|k)}}] \\ &= (1 - \theta_1)^{t_1 + \cdots + t_k} \end{aligned}$$

por lo tanto

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{\sum_{i=1}^k t_i}. \quad (4.55)$$

Para el número promedio muestral, debido a (4.53) para $n < k$ $\psi_{t|n} = I_{A_\theta^{(t|n)} \setminus A_3^{(t|n)}}$ sustituyendo esto en (1.17)

$$\begin{aligned} K(\theta; \psi_{t|n}) &= E_\theta[C_1(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})} + C_2(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} I_{(A_\theta^{(t|2)} \setminus A_3^{(t|2)})} \\ &= + \cdots + C_k(t) I_{(A_\theta^{(t|1)} \setminus A_3^{(t|1)})^c} \cdots I_{(A_\theta^{(t|k-1)} \setminus A_3^{(t|k-1)})^c} \mathbf{1} \end{aligned}$$

haciendo los cálculos

$$\begin{aligned} &= C_1(t)(1 - (1 - \theta)^{t_1}) + C_2(t)(1 - \theta)^{t_1}(1 - (1 - \theta)^{t_2}) + \cdots + \\ &\quad C_{k-1}(t)(1 - \theta)^{t_1 + \cdots + t_{k-2}}(1 - (1 - \theta)^{t_{k-1}}) + C_k(t)(1 - \theta)^{\sum_{j=1}^{k-1} t_j} \end{aligned}$$

sustituyendo $C_n(t) = c_1(t) + \cdots + c_n(t)$ y haciendo las cuentas, se llega a que

$$K(\theta; \psi_{t|n}) = c_1(t) + \sum_{i=2}^k c_i(t)(1 - \theta_1)^{\sum_{j=1}^{i-1} t_j}. \quad (4.56)$$

Nota 4.2. Debido a (4.55), (4.56) y a la Proposición 4.3, el mínimo de (4.54) está dado por

$$c_1(t) + \sum_{i=2}^k c_i(t)(1 - \theta_1)^{\sum_{j=1}^{i-1} t_j} + \lambda_0(1 - \theta_1)^{\sum_{i=1}^k t_i}. \quad (4.57)$$

Cálculo del error tipo I con (4.11) y (4.54). Para $n < k$ $\psi_{t|n} = I_{A_3^c(t|n)}$ sustituyendo esto en (1.14)

$$\begin{aligned} \alpha(\psi_{t|n}, \phi_{t|n}) &= E_0[I_{A_3^c(t|1)} I_{(A_{\theta_0}^{(t|1)} \setminus A_{\theta_1}^{(t|1)})^c} + I_{A_3^c(t|1)} I_{A_3^c(t|2)} I_{(A_{\theta_0}^{(t|2)} \setminus A_{\theta_1}^{(t|2)})^c} \\ &\quad + \cdots + I_{A_3^c(t|1)} \cdots I_{A_3^c(t|k-1)} I_{(A_{\theta_0}^{(t|k)} \setminus A_{\theta_1}^{(t|k)})^c}] \\ &= (1 - \theta_1)^{t_1 + \cdots + t_k} \end{aligned}$$

por lo tanto

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{\sum_{i=1}^k t_i}. \quad (4.58)$$

Para el número promedio muestral se sustituye (4.54) en (1.17), así

$$K(\theta; \psi_{t|n}) = E_\theta[C_1(t) I_{A_3^c(t|1)} + C_2(t) I_{A_3^c(t|1)} I_{A_3^c(t|2)} + \cdots + C_k(t) I_{A_3^c(t|1)} \cdots I_{A_3^c(t|k-1)} \mathbf{1},$$

haciendo los cálculos

$$= C_1(t)(1 - (1 - \theta)^{t_1}) + C_2(t)(1 - \theta)^{t_1}(1 - (1 - \theta)^{t_2}) + \dots + \\ C_{k-1}(t)(1 - \theta)^{t_1 + \dots + t_{k-2}}(1 - (1 - \theta)^{t_{k-1}}) + C_k(t)(1 - \theta)^{\sum_{j=1}^{k-1} t_j},$$

sustituyendo $C_n(t)$ y simplificando

$$K(\theta; \psi_{t|n}) = c_1(t) + \sum_{i=2}^k c_i(t)(1 - \theta)^{\sum_{j=1}^{i-1} t_j}. \quad (4.59)$$

Nota 4.3. Debido a que el error tipo II es cero, (ver proposición 4.3, y a los valores de del error tipo I y al número promedio de observaciones dados en (4.58) y a (4.59) respectivamente, se tiene que el mínimo de (4.54) es el que está dado en (4.57).

4.5. Paro óptimo para $1 < \theta$

El análisis para el caso $\theta < 0$ es el mismo al que se realiza a continuación.

Teorema 4.3. Sea $1 < \theta$, entonces las funciones recursivas $V^{(t|n)}$ con las distribuciones (4.2) y (4.1) están dadas por

$$V^{t|n}(x^{(t|n)}) = v_{t|n} I_{A_3^{(t|n)}}(x^{(t|n)}) \quad (4.60)$$

con

$$v_{t|n} = \lambda(1 - \theta_1)^{\sum_{i=n}^{k-1} t_{i+1}} \quad (4.61)$$

con $n = 1, \dots, k - 1$ y $v_{t|k} = \lambda$.

Demostración. Esta se hará por inducción hacia atrás.

La comprobación para $n = k$ es completamente análoga a la realizada en (4.14).

En la hipótesis de inducción se supone que (4.60) es cierta para $n = m < k$

$$V^{(t|m)}(x^{(t|m)}) = v_{t|m} I_{A_3^{(t|m)}}(x^{(t|m)}). \quad (4.62)$$

Por demostrar que (4.60) es cierta para $n = m - 1$, es decir

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1} I_{A_3^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.63)$$

Para probar (4.63) se sustituye (4.62) en (2.15)

$$V^{(t|m-1)}(x^{(t|m-1)}) = \text{mín}\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ + \int \dots \int \lambda(1 - \theta_1)^{\sum_{i=m}^{k-1} t_{i+1}} I_{A_3^{(t|m)}}(x^{t|m}) dx_{t_1 + \dots + t_{m-1} + 1} \dots dx_{t_1 + \dots + t_m}\},$$

integrando

$$= \text{mín}\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ + \lambda(1 - \theta_1)^{\sum_{i=m}^{k-1} t_{i+1}} (1 - \theta_1)^{t_m} I_{A_3^{(t|m-1)}}(x^{(t|m-1)})\} \\ = \text{mín}\{\lambda I_{A_3^{(t|m-1)}}(x^{(t|m-1)}), c_m(t) I_{A_\theta^{(t|m-1)}}(x^{(t|m-1)}) \\ + \lambda(1 - \theta_1)^{\sum_{i=m-1}^{k-1} t_{i+1}} I_{A_3^{(t|m-1)}}(x^{(t|m-1)})\}. \quad (4.64)$$

Análisis de (4.64).

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 0$, se sigue directamente que

$$V^{(t|m-1)}(x^{(t|m-1)}) = 0. \quad (4.65)$$

Si $I_{A_3^{(t|m-1)}}(x^{(t|m-1)}) = 1$, entonces

$$V^{(t|m-1)}(x^{(t|m-1)}) = \lambda(1 - \theta_1) \sum_{i=m-1}^{k-1} \theta_1^{i+1} I_{A_3^{(t|m-1)}}(x^{(t|m-1)}),$$

la constante en la última igualdad es $v_{t|m-1}$ (ver (4.61)), entonces

$$V^{(t|m-1)}(x^{(t|m-1)}) = v_{t|m-1} I_{A_3^{(t|m-1)}}(x^{(t|m-1)}). \quad (4.66)$$

De (4.65) y (4.66) se obtiene (4.63) y con esto termina la demostración. \square

Sea $1 < \theta$, la regla de paro está dada por

$$I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) < c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1} I_{A_3^{(t|n)}}(x^{(t|n)})\}} \quad (4.67)$$

$$\leq \psi_{t|n} \leq$$

$$I_{\{\lambda I_{A_3^{(t|n)}}(x^{(t|n)}) \leq c_{n+1}(t) I_{A_\theta^{(t|n)}}(x^{(t|n)}) + v_{t|n+1} I_{A_3^{(t|n)}}(x^{(t|n)})\}}, \quad (4.68)$$

para todo $n = 1, \dots, k-1$, y $v_{t|n}$ están definidas en (4.61).

Proposición 4.8. *La indicadora dada en (4.67) es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_\theta^{(t|n)}} & \text{si } n < k \\ 1 & \text{si } n = k \end{cases} \quad (4.69)$$

para todo $n = 1, \dots, k$.

Demostración. Se necesita probar que (4.67) y (4.69) son iguales.

- Si $I_{A_\theta^{(t|n)}} = 1$ y $I_{A_3^{(t|n)}} = 0$, entonces $\psi_{t|n} = 1$, para (4.67) se tiene $I_{\{0 < c_{n+1}(t)\}}$ y esta indicadora vale 1 porque se cumplen las condiciones de (4.7).
- Si $I_{A_\theta^{(t|n)}} = 0$ y $I_{A_3^{(t|n)}} = 1$ entonces $\psi_{t|n} = 0$, para (4.67) se tiene $I_{\{\lambda < v_{t|n}\}}$ y el valor de la indicadora es 0 porque $\lambda > v_{t|n}$ (ver (4.61)) y no se cumple la condición de (4.7).
- Si $I_{A_\theta^c(t|n)} = 1$, entonces $\psi_{t|n} = 0$; y el valor de (4.67) es cero debido a que los términos de la indicadora son cero y no se cumple (4.7).

Por lo tanto (4.67) y (4.69) son iguales. \square

Proposición 4.9. *La indicadora dada en (4.68) es igual a*

$$\psi_{t|n} = \begin{cases} I_{A_3^c(t|n)} & \text{si } n < k \\ 1 & \text{si } n = k, \end{cases} \quad (4.70)$$

para todo $n = 1, \dots, k$.

4. El problema modificado de Kiefer-Weiss con distribuciones uniformes

Demostración. Se necesita probar que (4.68) y (4.70), son iguales.

- Si $I_{A_3^c(t|n)} = 1 = \psi_{t|n}$, y para (4.68) se tienen los siguientes casos
 - Si $I_{A_\theta(t|n)} = 1$ y $I_{A_3(t|n)} = 0$ y la indicadora dada en (4.68) es $I_{\{0 \leq c_{n+1}(t)\}}$ y el valor es 1 porque se cumple la condición de (4.7).
 - Si $I_{A_\theta(t|n)} = 0 = I_{A_3(t|n)}$ y como los términos de la indicadora dada en (4.68) son cero y en particular se cumple la igualdad de (4.7).
- Si $I_{A_\theta(t|n)} = 0$ y $I_{A_3(t|n)} = 1$ entonces $I_{A_3^c(t|n)} = 0 = \psi_{t|n}$, para (4.68) se tiene que $I_{\{\lambda \leq v_{t|n}\}}$ cuyo valor es cero porque $\lambda > v_{t|n}$ (ver (4.61)) y no se cumple la condición de (4.7).

Por lo tanto (4.68) y (4.70) son iguales. □

4.6. Error tipo I y número promedio muestral

A partir de los resultados del Teorema 4.3, se hacen los cálculos de los errores de tipo I y tipo II, además del número promedio muestral.

Para el error tipo I se sustituyen (4.10) y (4.69) en (1.14)

$$\alpha(\psi_{t|n}, \phi_{t|n}) = E_0[I_{A_\theta(t|1)}I_{A_{\theta_1}(t|1)} + I_{A_\theta^c(t|1)}I_{A_\theta(t|2)}I_{A_{\theta_1}(t|2)} + \cdots + I_{A_\theta^c(t|1)} \cdots I_{A_\theta^c(t|k-1)}1I_{A_{\theta_1}(t|k)}]$$

como $E_0[I_{A_\theta(t|n)}I_{A_{\theta_1}(t|n)}] = 0$ para todo $n = 1, \dots, k-1$, haciendo los cálculos

$$= (1 - \theta_1)^{t_1} \cdots (1 - \theta_1)^{t_k}.$$

y simplificando

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{\sum_{i=1}^k t_i}. \quad (4.71)$$

Para el número promedio de observaciones se sustituye (4.69) en (1.17)

$$\begin{aligned} K(\theta; \psi_{t|n}) &= E_\theta[C_1(t)I_{A_\theta(t|1)} + C_2(t)I_{A_\theta^c(t|1)}I_{A_\theta(t|2)} + C_3(t)I_{A_\theta^c(t|1)}I_{A_\theta^c(t|2)}I_{A_\theta(t|3)} \\ &\quad + \cdots + C_k(t)I_{A_\theta^c(t|1)} \cdots I_{A_\theta^c(t|k-1)}1] \\ &= C_1(t) \end{aligned}$$

como $C_1(t) = c_1(t)$

$$K(\theta; \psi_{t|n}) = c_1(t). \quad (4.72)$$

Nota 4.4. En el caso en que $1 < \theta$, debido a: (4.71), (4.72) y a la Proposición 4.3, el valor mínimo de (4.60), es

$$c_1(t) + \lambda_0(1 - \theta_1)^{\sum_{i=1}^k t_i}. \quad (4.73)$$

A continuación se deducen las fórmulas para el error tipo I y número promedio muestral, con las reglas de decisión y de paro dadas en (4.11) y (4.70), respectivamente.

Para el error tipo I se sustituye (4.11) y (4.70) en (1.14), así que

$$\begin{aligned} \alpha(\psi_{t|n}, \phi_{t|n}) &= E_0[I_{A_3^c(t|1)} I_{(A_{\theta_0}^{(t|1)} \setminus A_{\theta_1}^{(t|1)})^c} + I_{A_3^{(t|1)}} I_{A_3^c(t|2)} I_{(A_{\theta_0}^{(t|2)} \setminus A_{\theta_1}^{(t|2)})^c} \\ &\quad + \cdots + I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|k-1)}} \mathbb{1}_{I_{(A_{\theta_0}^{(t|k)} \setminus A_{\theta_1}^{(t|k)})^c}}] \\ &= (1 - \theta_1)^{t_1} \cdots (1 - \theta_1)^{t_k} \end{aligned}$$

entonces

$$\alpha(\psi_{t|n}, \phi_{t|n}) = (1 - \theta_1)^{t_1 + \cdots + t_k}. \quad (4.74)$$

Para el promedio muestral se sustituye (4.70) en (1.17), obteniendo

$$\begin{aligned} K(\theta; \psi_t) &= E_\theta[C_1(t) I_{A_3^c(t|1)} + C_2(t) I_{A_3^{(t|1)}} I_{A_3^c(t|2)} + C_3(t) I_{A_3^{(t|1)}} I_{A_3^{(t|2)}} I_{A_3^c(t|3)} \\ &\quad + \cdots + C_k(t) I_{A_3^{(t|1)}} \cdots I_{A_3^{(t|k-1)}} \mathbb{1}] \end{aligned}$$

como $E_\theta[I_{A_3^{(t|n)}}] = 0$ para todo $n = 1, \dots, k - 1$, entonces

$$K(\theta; \psi_t) = C_1(t) = c_1(t). \quad (4.75)$$

Nota 4.5. Considerando el caso $1 < \theta$. Debido a que (4.71) y (4.72) coinciden con (4.74) y (4.75), respectivamente, y además el error tipo II es cero. La cota mínima que se obtiene con el tiempo de paro dado en (4.9), coincide con la que está dada en (4.73).

4.7. Resultados numéricos para el caso $\theta = \theta_0$

En esta sección, presentamos los resultados numéricos de las simulaciones realizadas con (4.29) y (4.30). La tabla 4.1 contiene el error tipo I , el número promedio de observaciones (N), y el tamaño de cada etapa (T). Las estimaciones se realizaron para un proceso con dos etapas.

Para los cálculos el valor de N se fue modificando hasta obtener cierto valor deseado para α . Otro punto de interés para los cálculos fue fijar el error de tipo I para obtener el número promedio muestral. Estos resultados se presentan en la Tabla 4.2.

$\theta_1=0.1$		$\theta_1=0.15$		N	T
α	AC	α	NPM		
0.109	13.59	0.032	11.54	21	8, 13
0.052	16.27	0.010	13.54	28	10, 18
0.020	19.06	0.002	15.55	37	12, 25
0.010	20.65	0.0009	16.62	43	13, 30
0.005	22.20	0.0002	18.05	50	15, 35

Tabla 4.1. Diferentes tamaños muestrales.

Stages	T	NPM
1	28	28
2	10, 18	16.27
3	6, 8, 14	13.45
4	4, 5, 7, 12	12.21
5	3, 4, 5, 6, 10	11.50
6	3, 3, 4, 5, 6, 7	11.05

Tabla 4.2. Números promedio para diferentes etapas, el valor de $\alpha=0.052$, y $\theta_1=0.1$.

En la Tabla 4.3, están los datos de dos etapas. Se fija el valor de α , se presenta el tamaño de la muestra N el cual se divide en dos etapas T , y se da el número promedio muestral.

α	N	T	NPM
0.109	21	8, 13	13.59
0.052	28	10, 18	16.27
0.020	37	12, 25	19.06
0.010	43	13, 30	20.65
0.005	50	15, 35	22.20

Tabla 4.3. Diferentes tamaños de muestra.

Otro punto de interés es fijar el valor de α y el tamaño de la muestra N , después considerar los datos en diferentes etapas y observar el número promedio muestral. Los resultados se presentan en la Tabla 4.4.

Etapas	T	NPM
1	28	28
2	10, 18	16.27
3	6, 8, 14	13.45
4	4, 5, 7, 12	12.21
5	3, 4, 5, 6, 10	11.50
6	3, 3, 4, 5, 6, 7	11.05

Tabla 4.4. Promedio muestral, $N=28$ y $\alpha=0.052$.

4.8. Errores y número promedio muestral con distribuciones uniformes usando el criterio de Wald

Considerando el trabajo realizado con respecto a las pruebas óptimas multietápicas en el presente trabajo, ahora surge el interés de calcular el error tipo I , el error tipo

II además del número promedio de observaciones utilizando el criterio de la razón de verosimilitud de Wald. Los cálculos se harán con distribuciones uniformes.

Para lo que sigue, se hacen las siguientes definiciones

$$A_{\theta_0}^i = \{(x_1, \dots, x_i) | x_1, \dots, x_i \in A_{\theta_0}\}, \quad i = 1, 2, \dots$$

y

$$A_{\theta_1}^i = \{(x_1, \dots, x_i) | x_1, \dots, x_i \in A_{\theta_1}\}, \quad i = 1, 2, \dots$$

los conjuntos A_{θ_0} y A_{θ_1} definidos en (4.4) son a los que se hará referencia en el presente análisis. Además de estos conjuntos, se define

$$A_{\theta}^i = \{(x_1, \dots, x_i) | x_1, \dots, x_i \in A_{\theta_0} \cap A_{\theta_1}\}, \quad i = 1, 2, \dots$$

4.9. La razón de verosimilitud con distribuciones uniformes

La razón de verosimilitud con distribuciones uniformes es $Z^i = \frac{I_{A_{\theta_1}^i}}{I_{A_{\theta_0}^i}}$ y los valores que toma son

$$Z^i = \begin{cases} 1 & \text{si } I_{A_{\theta}^i} = 1 \\ 0 & \text{si } I_{A_{\theta_0}^i \setminus A_{\theta_1}^i} = 1 \\ \infty & \text{si } I_{A_{\theta_1}^i \setminus A_{\theta_0}^i} = 1 \end{cases} \quad (4.76)$$

entonces $Z^i = 1$ if $I_{A_{\theta}^i} = 1$ (continuar), por lo cual

$$1 - \psi_i = I_{A_{\theta}^i}, \text{ and } \psi_i = I_{(A_{\theta}^i)^c} \quad \text{con } i = 1, 2, \dots \quad (4.77)$$

El tiempo de paro es $t = \inf\{i \geq 0 : Z^i \neq 1\}$, con regla de decisión

$$\phi_i = I_{A_{\theta_1}^i \setminus A_{\theta_0}^i} \quad \text{con } i = 1, 2, \dots \quad (4.78)$$

A partir de (4.77) y (4.78)

$$\psi_i \phi_i = I_{(A_{\theta}^i)^c} I_{A_{\theta_1}^i \setminus A_{\theta_0}^i} = I_{A_{\theta_1}^i \setminus A_{\theta_0}^i} \quad \text{con } i = 1, 2, \dots \quad (4.79)$$

por otra parte, debido a (4.77) y además al tomar el complemento de (4.78)

$$\psi_i (1 - \phi_i) = I_{(A_{\theta}^i)^c} I_{(A_{\theta_1}^i \setminus A_{\theta_0}^i)^c} = I_{A_{\theta_0}^i \setminus A_{\theta_1}^i} \quad \text{con } i = 1, 2, \dots \quad (4.80)$$

4.10. El error tipo *I* y error tipo *II*

Para el error tipo *I* se sustituye (4.79) y la parte izquierda de (4.77) en (1.14)

$$\alpha(\psi, \phi) = E_0[I_{A_{\theta_1}^1 \setminus A_{\theta_0}^1} + I_{A_{\theta}^1} I_{A_{\theta_1}^2 \setminus A_{\theta_0}^2} + \dots] = 0. \quad (4.81)$$

Para el error tipo *II* se debe sustituir (4.80) y la parte izquierda de (4.77) en (1.15), obteniendose

$$\beta(\psi, \phi) = E_1[I_{A_{\theta_0}^1 \setminus A_{\theta_1}^1} + I_{A_{\theta}^1} I_{A_{\theta_0}^2 \setminus A_{\theta_1}^2} + \dots] = 0. \quad (4.82)$$

Como los valores del error tipo *I* y el del error tipo *II* valen cero, eso quiere decir que no hay error al tomar la decisión al aceptar o rechazar la hipótesis nula, según sea el caso.

4.11. El costo promedio

Para calcular el costo hay que sustituir (4.77) en (1.17), así que

$$\begin{aligned} K(\theta_0; \psi) &= C_1(t)E_0I_{(A_\theta^1)^c} + C_2(t)E_0I_{A_\theta^1}I_{(A_\theta^2)^c} + C_3(t)E_0I_{A_\theta^1}I_{A_\theta^2}I_{(A_\theta^3)^c} + \cdots + \\ &= C_1(t)\theta_1 + C_2(t)(1 - \theta_1)\theta_1 + C_3(t)(1 - \theta_1)(1 - \theta_1)\theta_1 + \cdots + \end{aligned}$$

como $C_n(t) = c_1(t) + \cdots + c_n(t)$

$$= c_1(t)\theta_1 + (c_1(t) + c_2(t))(1 - \theta_1)\theta_1 + (c_1(t) + c_2(t) + c_3(t))(1 - \theta_1)^2\theta_1 + \cdots +$$

desarrollando y agrupando

$$\begin{aligned} &= c_1(t)\theta_1[1 + (1 - \theta_1) + (1 - \theta_1)^2 + \cdots +] + c_2(t)\theta_1[(1 - \theta_1) + (1 - \theta_1)^2 + \cdots +] \\ &\quad + c_3(t)\theta_1[(1 - \theta_1)^2 + (1 - \theta_1)^3 + \cdots +] + \cdots + \end{aligned}$$

pero cada paréntesis contiene una serie geométrica, calculando el límite de cada serie se obtiene

$$= c_1(t)\theta_1 \frac{1}{\theta_1} + c_2(t)\theta_1 \frac{1 - \theta_1}{\theta_1} + c_3(t)\theta_1 \frac{(1 - \theta_1)^2}{\theta_1} + \cdots$$

simplificando

$$K(\theta_0; \psi) = c_1(t) + \sum_{i=2}^{\infty} c_i(t)(1 - \theta_1)^{\sum_{j=i}^{i-1} t_j}. \quad (4.83)$$

Finalmente se obtuvieron las fórmulas utilizando el criterio de Wald [Wald (1947)], para calcular: el error tipo I , el error tipo II y el costo del experimento.

Se clara que los resultados (4.81)- (4.83) fueron obtenidos como parte de la investigación en la presente tesis, por lo cual no se hace referencia de las fórmulas obtenidas a algún texto en particular.

Capítulo 5

Conclusiones sobre el trabajo desarrollado en la tesis

En el presente trabajo se lograron los siguientes resultados.
En el capítulo 1.

- Se presentó un breve repaso de la estadística clásica, y algunos resultados necesarios para el para la teoría que se desarrollará posteriormente.
- Se dedujo la forma de un proceso estocástico a tiempo continuo en etapas.
- Se obtuvieron las reglas generales del tiempo de paro y de la regla de decisión.
- Se planteó la forma en general el error tipo I , el error tipo II y el costo del experimento, todo esto de forma secuencial en k etapas.

En el capítulo 2, se desarrolló en la forma más general la teoría necesaria y suficiente para realizar pruebas de hipótesis cuando los datos se toman en grupos o etapas.

- Se determinó la forma de la regla de decisión óptima independientemente del tiempo de paro.
- Se obtuvo la regla de paro óptimo para un proceso k -etápico.

En la primera sección del capítulo 3, se presenta una aplicación muy general de la teoría desarrollada para un proceso con variables aleatorias i.i.d.

- En la sección 3.2, se da una aplicación más específica de la teoría, y la solución al proceso de Wiener con deriva lineal en etapas. Para esto se desarrollaron las fórmulas explícitas para k -etapas, a partir de éstas se presentan resultados de las evaluaciones numéricas para dos, tres y cuatro etapas.
- En los cálculos del costo promedio, se consideró cuando la prueba se hace para las hipótesis H_0 y H_1 o sólo una de ellas.
- Con respecto al punto anterior se creó un software en MatLab, para la evaluación numérica del proceso de Wiener para dos y tres etapas, el cual se puede consultar en los Apéndices 6.2 y 6.10, respectivamente.

5. Conclusiones sobre el trabajo desarrollado en la tesis

- Otro punto que se logró fue dar solución al problema de Kieffer-Weiss modificado con distribuciones uniformes. En el desarrollo del problema surgieron diferentes vertientes, en cada caso se determinaron las fórmulas para obtener el error de tipo I y de tipo II , además del número promedio muestral.
- Para un caso particular del problema de Kieffer-Weiss modificado se generó un software en MatLab (para obtener resultados numéricos del error de tipo I , del error de tipo II y el número promedio de observaciones), éste se puede ver en el apéndice 6.22.
- Al final del capítulo 3, se dió un resumen muy detallado sobre los resultados obtenidos de la evaluación numérica realizada.

En el capítulo 4, se dió solución al problema modificado de Kieffer-Weiss con distribuciones uniformes. En este problema participa una tercer hipótesis diferente a H_0 y H_1 .

- Al hacer el análisis se dieron diferentes casos y en cada uno de ellos se obtuvo.
 - La fórmula para el error tipo I .
 - La fórmula para el error tipo II .
 - La fórmula para el costo promedio de observaciones.
- Para un caso en particular, se hizo la evaluación numérica y se obtuvo el valor del error tipo I , el valor del error tipo II y el número promedio de observaciones.
- Otro de los resultados obtenidos fueron las fórmulas para el cálculo de errores y número promedio muestral con dos distribuciones uniformes usando el criterio de Wald (ver Sección 4.8). En este caso se pudo ver que el valor del error tipo I y del error tipo II es cero.

Capítulo 6

Apéndices

En esta sección se darán los resultados numéricos de las evaluaciones del proceso de Weiner con deriva lineal, para dos, tres y cuatro etapas. También se dan las tablas de resultados del problema de Kiefer-Weiss modificado, con distribuciones uniformes.

Otro punto de relevancia del presente trabajo es el código desarrollado en MatLab, para realizar las evaluaciones numéricas tanto para el proceso de Weiner con deriva lineal y para el problema de Kiefer-Weiss modificado.

Apéndice A

En este Apéndice se presentan las tablas con los resultados de la evaluación numérica del proceso de Wiener con deriva lineal en tres etapas.

6.1. Tablas de resultados numéricos en tres etapas

	λ_0	λ_1				
		30	65	150	300	570
α, β	30	0.118, 0.118	0.139, 0.046	0.161, 0.018	0.183, 0.008	0.213, 0.004
t_1, t_2, t_3		2.2, 2.1, 3.2	2.8, 2.8, 4.4	3.2, 3.7, 5.6	3.3, 4.4, 6.5	3.1, 5.2, 7.2
A_1, B_1		0.26, 3.83	0.11, 2.94	0.04, 2.58	0.02, 1.86	0.01, 1.53
A_2, B_2		0.34, 2.91	0.15, 2.09	0.06, 1.52	0.03, 1.18	0.01, 0.93
C		1	0.46	0.2	0.1	0.05
α, β		65	0.046, 0.139	0.052, 0.052	0.055, 0.020	0.057, 0.009
t_1, t_2, t_3	2.8, 2.8, 4.4		3.9, 3.7, 6.2	4.8, 4.6, 8.0	5.4, 5.3, 9.2	5.9, 6.0, 10.2
A_1, B_1	0.33, 8.99		0.13, 7.39	0.05, 6.21	0.02, 5.47	0.01, 4.89
A_2, B_2	0.47, 6.66		0.19, 5.03	0.08, 3.94	0.04, 3.31	0.02, 2.86
C	2.16		1	0.43	0.21	0.11
α, β	150		0.018, 0.161	0.020, 0.055	0.021, 0.021	0.021, 0.010
t_1, t_2, t_3		3.2, 3.7, 5.6	4.8, 4.6, 8.0	6.0, 5.5, 10.2	6.7, 6.2, 11.8	7.4, 6.8, 13.2
A_1, B_1		0.43, 20.79	0.16, 17.72	0.06, 15.50	0.03, 14.12	0.01, 17.07
A_2, B_2		0.65, 15.29	0.25, 11.75	0.10, 9.46	0.05, 8.17	0.02, 7.28
C		3	2.30	1	0.5	0.26
α, β		300	0.008, 0.185	0.010, 0.057	0.010, 0.021	0.010, 0.010
t_1, t_2, t_3	3.2, 4.4, 6.5		5.4, 5.3, 9.2	6.7, 6.2, 11.8	7.6, 6.9, 13.8	8.3, 7.4, 15.4
A_1, B_1	0.53, 40.13		0.18, 34.99	0.07, 31.19	0.03, 28.89	0.01, 27.19
A_2, B_2	0.84, 29.82		0.30, 22.96	0.12, 18.61	0.06, 16.21	0.03, 14.55
C	10		4.61	2	1	0.52
α, β	570		0.004, 0.217	0.005, 0.059	0.005, 0.021	0.005, 0.009
t_1, t_2, t_3		3.0, 5.1, 7.2	5.8, 5.9, 10.2	7.3, 6.8, 13.2	8.3, 7.4, 15.3	9.0, 7.9, 17.2
A_1, B_1		0.64, 72.22	0.20, 64.51	0.07, 58.28	0.03, 54.27	0.01, 51.93
A_2, B_2		1.06, 54.98	0.34, 42.27	0.13, 34.29	0.06, 29.93	0.03, 26.95
C		19	8.77	3.8	1.9	1

Tabla 6.1. Tres etapas y dos hipótesis. Probabilidades de error, tiempos de paro y constantes que determinan las regiones de decisión.

6. Apéndices

	λ_0	λ_1					
		20	40	80	100	150	200
α, β	20	0.033, 0.119	0.044, 0.057	0.056, 0.028	0.060, 0.022	0.068, 0.014	0.075, 0.010
t_1, t_2, t_3		2.5, 3.2, 5.8	3.6, 3.6, 6.0	4.6, 3.9, 6.1	4.9, 4.0, 6.1	5.5, 4.2, 6.2	5.9, 4.3, 6.2
A_1, B_1		0.30, 74.03	0.16, 40.06	0.08, 21.24	0.06, 17.26	0.04, 11.80	0.03, 8.99
A_2, B_2		0.43, 28.92	0.22, 15.0	0.11, 7.77	0.08, 6.26	0.05, 4.23	0.04, 3.20
α, β	80	0.002, 0.127	0.001, 0.062	0.005, 0.033	0.0006, 0.027	0.002, 0.15	0.002, 0.011
t_1, t_2, t_3		4.0, 6.1, 16.8	5.3, 8.0, 19.4	5.6, 7.4, 18.6	5.2, 8.0, 19.4	8.0, 9.7, 21.5	8.5, 10.0, 21.9
A_1, B_1		0.56, 45.00	0.31, 78.00	0.15, 26.99	0.12, 16613	0.09, 63.5	0.06, 56.9
A_2, B_2		0.91, 200524	0.51, 285500	0.24, 106435	0.20, 117424	0.14, 177203	0.11, 150759
α, β	100	0.002, 0.149	0.002, 0.062	0.004, 0.029	0.007, 0.033	0.002, 0.014	0.002, 0.011
t_1, t_2, t_3		3.1, 6.1, 15.5	5.2, 7.9, 17.9	6.6, 8.0, 17.9	2.8, 7.3, 17.3	8.4, 9.8, 20.0	8.5, 10.4, 20.7
A_1, B_1		0.56, 40.7	0.31, 50.3	0.15, 34.5	0.12, 18	0.09, 57.1	0.07, 63.5
A_2, B_2		0.87, 267398	0.48, 394451	0.24, 198735	0.18, 108117	0.14, 258550	0.10, 261207
α, β	150	0.001, 0.134	0.002, 0.073	0.0002, 0.035	0.0002, 0.029	0.0008, 0.015	0.001, 0.11
t_1, t_2, t_3		3.9, 7.1, 15.0	4.2, 7.6, 15.6	5.6, 13.3, 22.3	5.4, 13.2, 22.4	8.2, 11.9, 20.0	8.7, 11.2, 19.3
A_1, B_1		0.60, 69.1	0.33, 0.45.2	0.20, 781.5	0.16, 619.5	0.10, 172	0.07, 82.7
A_2, B_2		0.87, 842990	0.44, 553112	0.29, 3961555	0.23, 3030239	0.14, 1332448	0.1, 173695
α, β	200	0.0002, 0.129	0.0004, 0.063	0.0009, 0.039	0.001, 0.031	0.0003, 0.017	0.001, 0.011
t_1, t_2, t_3		4.3, 10.0, 17.1	5.5, 10.6, 17.7	4.6, 10.5, 17.5	4.6, 10.6, 17.7	2.9, 10.9, 17.9	8.8, 11.2, 18.2
A_1, B_1		0.69, 381.6	0.35, 271.5	0.17, 125	0.14, 108.2	0.09, 15594.7	0.07, 78
A_2, B_2		0.9, 6359605	0.14, 4341332	0.24, 2001047	0.19, 1728377	0.13, 1310556	0.10, 1169040

Tabla 6.2. Tres etapas y dos hipótesis. Probabilidades de error, tiempos de paro y constantes que determinan las regiones de decisión en cada etapa.

	λ_0	λ_1					
		20	40	80	100	150	200
Wald H_0, H_1	20	3.82, 5.26	5.10, 5.44	6.31, 5.34	6.69, 5.27	7.36, 5.12	7.82, 5.00
Neyman Pearson		9.07	10.75	12.23	12.66	13.42	13.92
ANS H_0, H_1		5.04, 8.28	6.53, 8.47	7.92, 8.28	8.36, 8.22	9.12, 8.14	9.64, 8.13
Wald H_0, H_1	80	4.07, 9.61	5.51, 11.48	6.67, 9.61	7.16, 13.91	8.28, 11.68	8.86, 11.57
Neyman Pearson		15.45	19.91	18.89	26.19	24.74	25.57
ANS H_0, H_1		6.20, 17.53	8.13, 23.54	9.36, 19.14	10.17, 26.03	11.65, 23.77	12.35, 23.15
Wald H_0, H_1	100	3.75, 9.46	5.48, 10.60	6.93, 10.18	6.63, 9.23	8.38, 11.52	8.88, 11.78
Neyman Pearson		14.95	18.59	20.15	18.25	24.64	25.93
ANS H_0, H_1		5.67, 16.17	8.06, 20.44	9.79, 18.98	9.80, 20.00	1.78, 22.15	12.38, 23.27
Wald H_0, H_1	150	3.98, 10.56	5.16, 10.43	6.65, 16.07	7.01, 15.84	8.32, 13.65	8.90, 12.33
Neyman Pearson		16.73	17.89	28.48	28.79	27, 88	26.84
ANS H_0, H_1		6.11, 17.22	7.50, 18.50	10.20, 37.81	10.62, 37.95	11.92, 28.31	12.48, 23.94
Wald H_0, H_1	200	4.08, 13.54	5.50, 14.00	6.46, 13.07	6.88, 12.90	8.04, 15.29	8.91, 22.92
Neyman Pearson		21.07	23.54	23.73	24.18	29.29	26.75
ANS H_0, H_1		6.61, 23.55	8.40, 26.91	9.53, 26.42	10.10, 26.80	10.61, 14.84	12.51, 22.92

Tabla 6.3. Tres etapas y una hipótesis. Números promedios.

Apéndice B

En esta sección se presenta el código para los resultados numéricos del proceso de Wiener en cuatro etapas.

6.2. Tablas de resultados numéricos en cuatro etapas

Tablas de resultados de la evaluación numérica del proceso de Wiener con deriva lineal en cuatro etapas.

	λ_0	λ_1				
		30	60	140	250	550
α, β	30	0.110, 0.110	0.127, 0.047	0.146, 0.018	0.161, 0.009	0.188, 0.004
t_1, t_2		1.94, 1.57	2.42, 2.06	2.70, 2.69	2.72, 3.17	2.58, 3.95
t_3, t_4		2.12, 3.25	2.77, 4.37	3.50, 5.56	3.97, 6.25	4.58, 7.06
A_1, B_1		0.22, 4.47	0.10, 3.64	0.04, 2.88	0.02, 2.47	0.01, 1.98
A_2, B_2		0.26, 3.83	0.12, 3.03	0.05, 2.35	0.02, 1.99	0.01, 1.59
A_3, B_3		0.34, 2.91	0.16, 2.15	0.06, 1.55	0.03, 1.26	0.01, 0.94
C		1	0.5	0.21	0.12	0.05
α, β		60	0.047, 0.127	0.052, 0.052	0.055, 0.020	0.056, 0.010
t_1, t_2	2.41, 2.05		3.28, 2.57	4.01, 3.19	4.39, 3.62	4.84, 4.27
t_3, t_4	2.77, 4.37		3.58, 5.94	4.44, 7.64	4.97, 8.65	5.63, 9.84
A_1, B_1	0.27, 9.77		0.11, 8.43	0.04, 7.23	0.02, 6.56	0.01, 5.76
A_2, B_2	0.32, 8.29		0.14, 6.90	0.05, 5.75	0.03, 5.16	0.01, 4.51
A_3, B_3	0.46, 6.13		0.21, 4.74	0.08, 3.67	0.05, 3.15	0.02, 2.61
C	2		1	0.42	0.24	0.11
α, β	140		0.018, 0.146	0.020, 0.055	0.020, 0.020	0.020, 0.010
t_1, t_2		2.67, 2.67	3.98, 3.16	5.02, 3.73	5.59, 4.08	6.26, 4.53
t_3, t_4		3.50, 5.56	4.44, 7.64	5.40, 9.90	5.97, 11.27	6.67, 12.90
A_1, B_1		0.34, 23.18	0.13, 20.80	0.05, 18.69	0.03, 17.54	0.01, 16.24
A_2, B_2		0.42, 19.68	0.17, 16.85	0.06, 14.61	0.03, 13.49	0.01, 12.27
A_3, B_3		0.64, 14.32	0.27, 11.25	0.11, 8.97	0.06, 7.90	0.02, 6.82
C		4.67	2.34	1	0.56	0.25
α, β		250	0.009, 0.163	0.010, 0.056	0.011, 0.020	0.011, 0.010
t_1, t_2	2.65, 3.12		4.33, 3.56	5.55, 4.08	6.17, 4.33	6.92, 4.72
t_3, t_4	3.97, 6.25		4.97, 8.65	5.98, 11.26	6.57, 12.85	7.29, 14.76
A_1, B_1	0.40, 40.47		0.15, 37.08	0.05, 33.96	0.03, 32.50	0.01, 30.60
A_2, B_2	0.50, 34.65		0.19, 29.98	0.07, 26.35	0.04, 24.54	0.01, 22.61
A_3, B_3	0.79, 25.13		0.31, 19.77	0.12, 15.85	0.07, 14.05	0.03, 12.25
C	8.34		4.16	1.78	1	0.45
α, β	550		0.004, 0.194	0.004, 0.059	0.005, 0.020	0.005, 0.010
t_1, t_2		2.36, 3.75	4.59, 4.10	6.06, 4.47	6.82, 4.72	7.45, 4.88
t_3, t_4		4.60, 7.06	5.65, 9.83	6.71, 12.86	7.32, 14.73	8.14, 17.02
A_1, B_1		0.49, 84.66	0.17, 79.33	0.06, 74.60	0.03, 71.86	0.01, 69.57
A_2, B_2		0.62, 73.44	0.22, 64.23	0.08, 56.99	0.04, 53.47	0.02, 49.47
A_3, B_3		1.05, 53.57	0.38, 42.00	0.14, 33.71	0.08, 29.93	0.03, 26.19
C		18.33	9.16	3.93	2.2	1

Tabla 6.4. Cuatro etapas y dos hipótesis. Probabilidades de error, tiempos de paro y constantes que determinan las regiones de decisión.

λ_0	Wald		Neyman- Pearson	ASN H_0	ASN H_1
30	3.25	3.25	5.99	3.93	3.93
60	5.17	5.17	10.51	6.34	6.34
140	7.40	7.40	16.67	9.16	9.16
240	8.72	8.71	20.72	10.80	10.82
520	10.49	10.47	26.54	12.98	13.04

Tabla 6.5. Cuatro etapas y dos hipótesis. Número promedios muestrales.

Apéndice C

6.3. Evaluación numérica del proceso de Wiener con deriva lineal en una y dos etapas

En esta parte se presenta el código, para la evaluación del proceso de Wiener en dos etapas. **Funciones para dos etapas**

```
% Las funciones  $\Phi(l_k)$  y  $1 - \Phi(m_k)$  (ver (3.22))
function[w]= f1t2( lam0,lam1,t2,the1
w=@(x) lam1*x*normcdf(Lim2H1(lam0,lam1,x,t2,the1),0,1);
end
```

```
function[w] = f0t2( lam0,lam1,t2,the1 )
w=@(x) lam0*(1-normcdf(Lim2H0(lam0,lam1,x,t2,the1),0,1));
end
```

```
% La función (3.23)
function[w] = Lim2H1(lam0,lam1,z,t2,the1)
w=(log(lam0)-log(lam1)-log(z))/(the1*sqrt(t2)) - the1*sqrt(t2)/2;
end
```

```
% La función (3.24)
function[w] = Lim2H0(lam0,lam1,z,t2,the1)
w= (log(lam0)-log(lam1)-log(z))/(the1*sqrt(t2)) + the1*sqrt(t2)/2;
end
```

Funciones para una etapa

```
% La función (3.55).
function[w] =Lim1H1(a1,t1,the1)
log(a1)/(the1*sqrt(t1))-the1*sqrt(t1)/2;
end
```

```
% La función (3.62).
function[w] = Lim1H0(b1,t1,the1)
w=log(b1)/(the1*sqrt(t1))+the1*sqrt(t1)/2;
end
```

6.4. Minimización de la función de Lagrange

Minimización de la función de Lagrange dada en la Sección 3.2.

```
function[FL,A] = Lagrange( lam0,lam1,t,the1,c0,c1,n)
% Cálculo de las funciones que participan en (3.22).

    z0=lam0/lam1;
    f=F(lam0,lam1,t(2),the1,c0,c1);
    if f(z0)>lam0
        A(1,1)=z0;
        A(1,2)=z0;
    else H1=@(z) f(z)-lam1*z;
        H0=@(z) f(z)-lam0;
        % Obtención de  $A_1$  (ver (3.27)).
        A(1,1)=fsolve(H1,z0);
        % Obtención de  $B_1$  (ver (3.28)).
        A(1,2)=fsolve(H0,z0);

        % La función (3.35).
        f1t1=lam1*normcdf(Lim1H1(A(1,1),t(1),the1),0,1);
        Obtención de (3.41).
        deltaz=(A(1,2)-A(1,1))/n;
        V1=[];
        for k=0:n
            V1=[V1;A(1,1)+deltaz*k,f(A(1,1)+deltaz*k)];
        end Z1=[];
        for k=1:n+1
            Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
        end
        J=[Z1(:,1),V1(:,2).*Z1(:,2)];
        clear V1 Z1;
        W1=[];
        for p=1:n
            W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
        end Y1=[];
        for m=1:n
            Y1=[Y1;W1(m,1)*W1(m,2)];
        end
        S1=sum(Y1);
        clear J1 W1 Y1;
        % La función (3.43).
        f0t1=lam0*(1-normcdf(Lim1H0(A(1,2),t(1),the1),0,1));
        % La función (3.44)
        Suma=t(1)*(c0+c1)+ S1 + f1t1 + f0t1 ;
        if Suma<min(lam0,lam1)
            FL=Suma;
```

```

else
FL=min(lam0,lam1);
end
end

```

6.5. El error tipo *I*

Las fórmulas para el error tipo *I* están en la sección 3.3.

```

function[w] = Alpha( lam0,lam1,t,A,the1,n)
syms x;
clear z;
% la función (3.42).
f=@(z)1-normcdf(Lim2H0(lam0,lam1,x,t(2),the1),0,1);
% La integral (3.41).
deltaz=(A(1,2)-A(1,1))/n;
V1=[];
for k=0:n
V1=[V1;A(1,1)+deltaz*k,f(A(1,1)+deltaz*k)];
end
Z1=[];
for k=1:n+1
Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear V1 Z1;
W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
clear J1 W1 Y1;
% La función (3.43).
alf=1-normcdf(Lim1H0(A(1,2),t(1),the1),0,1);
% La función (3.44).
w=alf+S1;
end

```

6.6. El error tipo *II*

Las fórmulas para el error tipo *II* están en la sección 3.4.

```

function[w] = Beta(lam0,lam1,t,A,the1,n)
f3=@(z)x*normcdf(Lim2H1(lam0,lam1,x,t(2),the1),0,1);
deltaz=(A(1,2)-A(1,1))/n;
V1=[];
for k=0:n
V1=[V1;A(1,1)+deltaz*k,f(A(1,1)+deltaz*k)];
end
Z1=[];
for k=1:n+1
Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear V1 Z1;
W1=[];
for p=1:n
W1=[W1;
(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
clear J1 W1 Y1;
bet=cdf('Normal',Lim1H1(A(1,1),t(1),the1),0,1);
w=bet+S1;

```

6.7. Número promedio muestral bajo las dos hipótesis

```

function[w] = CostoH0H1(t,A,c,c1,n,the1) La función (3.44).
deltaz=(A(1,2)-A(1,1))/n;
V1=[];
for k=0:n
V1=[V1;A(1,1)+deltaz*k,F(A(1,1)+deltaz*k)];
end
Z1=[];
for k=1:n+1
Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear Z1 V1;
W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end

```

```

end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
ww=(t(1)*(c+c1))+S1;
clear J W1 Y1;
end

```

6.8. Número promedio muestral en caso asimétrico

Cálculo del número promedio de H_0 (ver sección 3.6).

```

function[w] = CostoH0(t,A,the1)
p=normcdf([Lim1H0(A(1,1),t(1),the1),Lim1H0(A(1,2),t(1),the1)]);
w=t(1)+t(2)*(p(2)-p(1));
end

```

6.9. Programa principal

```

% Minimización de la función dada en (3.63).
clear all;
clc;
format long
% Se obtiene el caso simétrico si  $c_0 = 1 = c_1$ . Si  $c = 0$  ó  $c_1 = 0$ , entonces se tiene el
caso asimétrico.
c0=1; c1=1;
% Valores de  $\lambda_0$  y  $\lambda_1$ .
lambda0=exp(7);
lambda1=exp(7);
% Tamaño de la partición para la integral numérica.
n=400;
% valor de theta1.
theta=1;
% plan de muestreo.
VI=[10,10];
% Condiciones de nivel de error computacional.
options = optimset('Display','iter','TolFun',1e-8);
% Optimización de la función de lagrange y plan de muestreo  $T$ .
tt=fminsearch(@(t)Lagrange(lambda0,lambda1,t,theta,c0,c1,n),VI);
% Obtención el valor de la función de Lagrange y de las constantes  $A_1$  y  $B_1$ .
[FL,A]=Lagrange( lambda0,lambda1,tt,theta,c0,c1,n);

```

6.10. Presentación de resultados

```

. % Cálculo de error tipo I.
alfa=Alpha(lambda0,lambda1,tt,M,theta,n);
% Cálculo del error tipo II.
beta=Beta(lambda0,lambda1,tt,M,theta,n);
% Número promedio para H0.
PromedioH0=CostoH0(tt,M,theta);
% Número promedio para H1.
PromedioH1=CostoH1(tt,M,theta,n);
% Cálculo de la prueba de Wald para H0.
WaldH0=2*((1-alfa)*log((1-alfa)/beta)+alfa*log(alfa/(1-beta)));
% Número promedio de la prueba de Wald para H1.
WaldH1=2*((1-beta)*log((1-beta)/alfa)+beta*log(beta/(1-alfa))) ;
% Número promedio de la prueba de Neyman Pearson.
NP = (norminv((1 - alfa), 0, 1) + norminv((1 - beta), 0, 1))^2;
A=[];
A=[A;log(lambda0), log(lambda1), alfa, beta, PromedioH0, PromedioH1, WaldH0, WaldH1,
NP ];
A
B=[];
B=[ tt ];
B
C=[];
C=[C; M ];
log10(C)

```

Apéndice D

6.11. Código para la evaluación numérica del proceso de Wiener con deriva lineal en tres etapas

En esta parte se presenta el código en MatLab, para la evaluación numérica del proceso de Wiener en tres etapas. El primer objetivo es generar la función de Lagrange dada en (3.63).

6.12. Funciones para tres etapas

```
% Las funciones dadas en (3.22).
function[w]= f1t3(lam0,lam1,t3,the1)
w=@(x) lam1*x*normcdf(Lim3H1( lam0,lam1,x,t3,the1),0,1);
end

function[w] = f0t3(lam0,lam1,t3,the1)
w = @(x)lam0*(1-normcdf(Lim3H0(lam0,lam1,x,t3,the1),0,1));
end

% La función (3.23).
function[w] = Lim3H1(lam0,lam1,z,t3,the1)
w=((log(lam0)-log(lam1)-log(z))/(the1*sqrt(t3)))-the1*sqrt(t3)/2;
end

% La función (3.24).
function[w] = Lim3H0(lam0,lam1,z,t3,the1)
w=((log(lam0)-log(lam1)-log(z))/(the1*sqrt(t3)))+the1*sqrt(t3)/2;
end

% La función (3.30). %
function[w] = lim2(d2,z,the1,t2)
w=((log(d2)-log(z))/the1) + the1*t2/2;
end

% La función (3.50).
```

```
function[w]=lim1(d1,the1,t1)
w=(log(d1)/the1)+the1*t1/2;
end
```

6.13. Funciones para dos etapas

```
% La función (3.31).
function[w]=Lim2H1(a2,z,t2,the1)
w=((log(a2)-log(z))/(the1*sqrt(t2)))-the1*sqrt(t2)/2;
end
```

```
% La función (3.32).
function[w]=Lim2H0(b2,z,t2,the1)
w=((log(b2)-log(z))/(the1*sqrt(t2)))+the1*sqrt(t2)/2;
end
```

```
% La función (3.35).
function[w]=f1t2(lam1,a2,the1,t2)
w=@(x) lam1*x*normcdf(Lim2H1(a2,x,t2,the1),0,1);
end
```

```
% La función (3.42).
function[w]=f0t2(lam0,b2,the1,t2)
w=@(x) lam0*(1-normcdf(Lim2H0(b2,x,t2,the1),0,1));
end
```

6.14. Cálculo de la integral numérica

Cálculo de la segunda integral dada en (3.34), esta se obtiene de (3.37)-(3.41).

```
function w=DIntegral(F,a2,b2,t2,the1,n)
syms x;
dz=(b2-a2)/n;
U=[];
for k=0:n
U=[U;a2+dz*k,F(a2+dz*k)];
end
V=[];
clear k;
for k=1:n+1
V=[V;lim2(U(k,1),x,the1,t2), normpdf(lim2(U(k,1),x,the1,t2),0,sqrt(t2))];
end
I=[V(:,1),U(:,2).*V(:,2)];
clear U V k;
W=[];
```

```

for k=1:n
W=[W;(I(k+1,1)-I(k,1)),(I(k,2)+I(k+1,2))/2];
end Y=[];
clear k;
for k=1:n
Y=[Y;W(k,1)*W(k,2)];
end
w=inline(sum(Y));
clear I W Y;
end

```

6.15. Funciones para una etapa

```

% La función (3.54).
function w= f1t1(lam1,a1,the1,t1)
w=lam1*normcdf(Lim1H1(a1,t1,the1),0,1);
end

```

```

% La función (3.55).
function[w] =Lim1H1(a1,t1,the1)
w=(log(a1)/(the1*sqrt(t1)))-the1*sqrt(t1)/2;
end

```

```

% La función (3.61).
function[w] = f0t1(lam0,b1,the1,t1)
w=lam0*(1-normcdf(Lim1H0( b1,t1,the1),0,1));
end

```

```

% La función (3.62).
function[w] = Lim1H0(b1,t1,the1)
w=(log(b1)/(the1*sqrt(t1))) + the1*sqrt(t1)/2;
end

```

```

% La función (3.58).
function[w] = lim1(d1,the1,t1)
w=(log(d1)/the1) + the1*t1/2;
end

```

6.16. Minimización de la función de Lagrange

La función de Lagrange dada en la Sección 3.2, con $k = 3$.

```

function[FL,A] = fLagrangeFin(t,lam0,lam1,c0,c1,the1,n)
m=f1t3(lam0,lam1,t(3),the1);
p=f0t3(lam0,lam1,t(3),the1);

```

```

f3=@(z)t(3)*(c0+c1*z)+m(z)+p(z);
options = optimset('Display','iter','TolFun',1e-8);
h31=@(z)lam1*z-f3(z);
A(1,1)=fsolve(h31,z0,options);
h30=@(z)f3(z)-lam0;
A(1,2)=fsolve(h30,z0,options);
S2=DIIntegral(f3,A(1,1),A(1,2),t(2),the1,n);
h1=f1t2(lam1,A(1,1),the1,t(2));
h0=f0t2(lam0,A(1,2),the1,t(2));
f2=@(z)t(2)*(c0+c1*z)+h1(z)+S2(z)+h0(z);
H21=@(z)lam1*z-f2(z);
A(2,1)=fsolve(H21,z0,options);
H23=@(z)lam0-f2(z);
A(2,2)=fsolve(H23,z0,options);
dz=(A(2,2)-A(2,1))/n;
V1=[]; for i=0:n
V1=[V1;A(2,1)+dz*i,f2(A(2,1)+dz*i)];
end
clear i;
Z1=[];
for i=1:n+1
Z1=[Z1;lim1(V1(i,1),the1,t(1)), normpdf(lim1(V1(i,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear V1 Z1;
W1=[]; for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
clear J W1 Y1;
Suma=t(1)*(c0+c1)+f1t1(lam1,A(2,1),the1,t(1))+ S1+f0t1(lam0,A(2,2),the1,t(1));
if
Suma<min(lam0,lam1)
FL=Suma;
else
FL=min(lam0,lam1);
end
end
end

```

6.17. Programa principal

```

% La función de Lagrange dada en (3.63).
clear all;

```

```

clc;
format long
c0=1; c1=1;
% Si c0 = 1 = c1 los cálculos son del caso simétrico. Si c = 0 ó c1 = 0 se tiene el caso
asimétrico.
% Valores de  $\lambda_0$  y  $\lambda_1$ .
lambda0=30;
lambda1=30;
% Tamaño de la partición para las integrales numéricas.
n=100;
% Valor de theta1.
theta=1;
% Plan de muestreo.
VI=[3, 7, 7];
% Nivel de error computacional.
options = optimset('Display','iter','TolFun',1e-8);
% Optimización de la función de lagrange y plan de muestreo.
tt=fminsearch(@(t)fLagrangeFin(t,lambda0,lambda1,c,c1,theta,n),VI,options);
% Obtención de la función de Lagrange y de  $A_i$  y  $B_i$ .
[FL,A]=fLagrangeFin(tt,lambda0,lambda1,c,c1,theta,n);

```

6.18. Error tipo I, tipo II y número promedio muestral

```

% Cálculo de error tipo I (sección 3.3).
alfa=Alpha(tt,lambda0,lambda1,M,n,theta);
% El error tipo II.
beta=Beta(tt,lambda0,lambda1,M,n,theta);
% Número promedio para H0.
PromedioH0=CostoH0(tt,M,theta,n);
% Número promedio para H1.
PromedioH1=CostoH1(tt,M,theta,n);
% Número promedio de la prueba de Wald para H0.
WaldH0=2*((1-alfa)*log((1-alfa)/beta)+alfa*log(alfa/(1-beta)));
% Número promedio de la prueba de Wald para H1.
WaldH1=2*((1-beta)*log((1-beta)/alfa)+beta*log(beta/(1-alfa)));
% Número promedio de la prueba de Neyman Pearson.
NP = (norminv((1 - alfa), 0, 1) + norminv((1 - beta), 0, 1))^2;
Presentación de resultados.
A=[];
A=[A;log(lambda0), log(lambda1), alfa, beta, PromedioH0, PromedioH1, WaldH0, WaldH1,
NP];
B=[]; B=[tt];
C=[]; C=[C; log(M) ];
A

```

B
C

6.19. El error tipo *I*

Las fórmulas para el error de tipo *II* están en la sección 3.3.

```
function[w] = Alpha(t,lam0,lam1,A,n,the1)
syms x;
clear z;
% La función (3.22).
f3=@(z)1-normcdf(Lim3H0(lam0,lam1,z,t(3),the1),0,1);
% Ver la sección (6.14).
S2=DIntegral(f3,A(1,1),A(1,2),t(2),the1,n);
f2=@(z)S2(z)+1-normcdf(Lim2H0(A(1,2),z,t(2),the1),0,1);
% La función (3.42).
dz=(A(2,2)-A(2,1))/n;
V1=[]; clear k;
for k=0:n
V1=[V1;A(2,1)+dz*k,f2(A(2,1)+dz*k)];
end
Z1=[]; clear k; for k=1:n+1
Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear Z1 V1; W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
w=S1+1-normcdf(Lim1H0(A(2,2),t(1),the1));
% La función (3.61).
clear J W1 Y1;
end
```

6.20. El error tipo *II*

Las fórmulas para el error de tipo *II*, están en la sección 3.4.

```
function[w] = Beta(t,lam0,lam1,A,n,the1)
```

```

% La función (3.22).
f3=@(z)z*normcdf(Lim3H1(lam0,lam1,z,t(3),the1),0,1);
% Ver sección (6.14).
S2=DIntegral(f3,A(1,1),A(1,2),t(2),the1,n);
% La función (3.35).
f2=@(z)z*normcdf(Lim2H1(A(1,1),z,t(2),the1),0,1)+S2(z);
dz=(A(2,2)-A(2,1))/n;
V1=[];
for k=0:n
V1=[V1;A(2,1)+dz*k, f2(A(2,1)+dz*k)];
end
Z1=[];
for k=1:n+1
Z1=[Z1;lim1(V1(k,1),the1,t(1)), normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear V1 Z1; W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
w=normcdf(Lim1H1(A(2,1),t(1),the1))+S1;
clear J W1 Y1;
end

```

6.21. Número promedio muestral en el caso simétrico

Las fórmulas para número promedio muestral considerando el equilibrio Bayesiano, están en la sección 3.5.

```

function[w] = Costo(t,A,c,c1,n,the1)
f3=@(x)t(3)*(c+c1*x);
% DIntegral se encuentra en la sección (6.14).
S2=DIntegral(f3,A(1,1),A(1,2),t(2),the1,n);
% La función (3.44).
f2=@(x)t(2)*(c+c1*x)+S2(x);
deltaz=(A(2,2)-A(2,1))/n;
V1=[];
for k=0:n
Zb=A(2,1)+deltaz*k;

```

```

FZb=f2(Zb);
V1=[V1;Zb,FZb];
end
Z1=[];
for k=1:n+1
f0xF=normpdf(lim1(V1(k,1),the1,t(1)),0,sqrt(t(1)));
Z1=[Z1;lim1(V1(k,1),the1,t(1)),f0xF];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear Z1 V1;
W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1))/2,J(p,2)+J(p+1,2)];
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
w=t(1)*c+t(1)*c1+S1;
clear J W1 Y1;
end

```

6.22. Número promedio muestral en el caso asimétrico

% Cálculo del número promedio de H_0 , las fórmulas se están en la sección 3.6.

```

function[w] = CostoH0(t,A,the1,n)
y=@(x)cdf('Normal',lim2(A(1,2),x,the1,t(2)),0,sqrt(t(2)));
j=@(x)cdf('Normal',lim2(A(1,1),x,the1,t(2)),0,sqrt(t(2)));
f2=@(x)t(2)+t(3)*(y(x)-j(x));
dz=(A(2,2)-A(2,1))/n;
V1=[];
for i=0:n
V1=[V1;A(2,1)+dz*i,f2(A(2,1)+dz*i)];
end
Z1=[]; clear i;
for i=1:n+1
Z1=[Z1;lim1(V1(i,1),the1,t(1)),normpdf(lim1(V1(i,1), the1,t(1)),0,sqrt(t(1)))];
end
J=[Z1(:,1),V1(:,2).*Z1(:,2)];
clear Z1 V1; W1=[];
for p=1:n
W1=[W1;(J(p+1,1)-J(p,1)),(J(p,2)+J(p+1,2))/2];
end

```

```
end
Y1=[];
for m=1:n
Y1=[Y1;W1(m,1)*W1(m,2)];
end
S1=sum(Y1);
w=t(1)S1; % % Se obtiene la función (3.63).
clear J W1 Y1;
end
```


Apéndice E

6.23. Código para la evaluación numérica del problema de Kiefer-Weiss modificado

Aquí se presenta el código del capítulo 4.3, donde se resuelve el problema de Kiefer-Weiss modificado considerando el caso $0 \leq \theta \leq \theta_1$ y solo participan dos hipótesis. Las distribuciones que se utilizan son uniformes. La función que se optimiza es la que está dada en (4.31).

Código del programa principal

```
clear all; clc; format long;
N=28; % Tamaño de la muestra. %
L=[30,50]; % Valores de  $\lambda_0$  y  $\lambda_1$ . %
tr=[10]; % El vector de diseño  $t$ . %
b=1; c=1;
theta=0.1; % El valor de  $\theta_1$ . %
I=(b-theta);
FLag=KUni(tt,L,b,theta,c) % La función de Lagrange. %
h=round(fminsearch(@(t)KUni( t,N,L,I,c ),tr)); % Minimización de la función de La-
grange con respecto a  $t$ . %
tk=N-sum(h);
T=[h,tk] % Vector de diseño  $t$  óptimo. %
Alpha=KUniH0(T,b,theta) % Cálculo del error de tipo  $I$ . %
Costo=CostoKUnif(T,theta,b,c) % Cálculo del Costo. %
```

```
function[w] = KUni(T,N,L,I,c)
tk=N-sum(T);
t=[T,tk];
k=length(t);
v(k+1)=min(L(1),L(2));
for i=k:-1:1
v(i)=c*t(i)+v(i+1)*(b - theta)^t(i);
end
if v(1)<min(L(1),L(2))
w=v(1);
else
```

6. Apéndices

```
w=min(L(1),L(2));  
end  
end
```

Código para el cálculo del número promedio muestral dado en (4.30)

```
function[w] = CostoKUnif(t,theta,b,c)  
k=length(t);  
m=c*t(1);  
for i=2:k  
suma=0;  
for j=1:i-1  
suma=suma+t(j);  
end  
m=m+c*t(i)*(b - theta)^suma;  
end  
w=m;
```

Código para el cálculo del error de tipo I dado en (4.29)

```
function[w] = KUniH0(t,b,theta) k=length(t);  
V(k)=(b - theta)^t(k);  
for i=k-1:-1:1  
V(i)=V(i + 1) * (b - theta)^t(i);  
end  
w=V(1);
```

Bibliografía

- [Aivazian (1959)] Aivazian S. A. 1959. A comparison of the optimal properties of the Neyman-Pearson and the Wald sequential probability ratio test. *Theory Probability Apply.* 4, 86-93.
- [Clark (1971)] Elements of Abstract Algebra, Dover.
- [Cochlar and Vrana (1978)] Cochlar J., Vrana I. 1993. On the Optimum Sequential Test Of Two Hypotheses for Statistical Dependent Observations
- [Cressie and Morgan(1993)] Cressie, N., Morgan, P.B., 1993. The VRPT: A sequential testing procedure dominating the SPRT. *Econometric Theory* 9, 431-450.
- [Ghosh et al. (1997)] Ghosh, M., Mukhopadhyay, N., Sen, P.K., 1997. Sequential Estimation. Wiley, New York.
- [Greenwood, P.E., Shiryayev, A.N., 1985.] Contiguity and the statistical invariance principle. Gordon and Breach Science Publishers S.A., New York.
- [Ferguson (1967)] Ferguson, T. S., 1967. Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.
- [Jennison (1999)] Jennison, C., Turnbull, B. W., 1999. Group Sequential Methods with Applications to Clinical Trails. Chapman and Hall/CRC.
- [Kiefer and Weiss (1957)] Kiefer, J. and Weiss, L. (1957). Some Properties of Generalized Sequential Probability Ratio Tests. *Ann. Math. Statistics* 28, 57-75.
- [Kiefer, J., Weiss, L., (1957)] Some properties of generalized sequential probability ratio tests. *Annals of Mathematical Statistics* 28, 57-75.
- [Le Cam, L., Yang, G.L., 2000.] Asymptotics in Statistics: Some Basic Concepts. Springer Science & Business Media.
- [Lorden, G., (1980)] Structure of sequential tests minimizing an expected sample size. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 51, 291-302.
- [MATLAB, 2010. version 7.10.0 (R2010a).] The MathWorks Inc., Natick, Massachusetts.
- [Novikov, A., 1988.] A functional limit theorem for likelihood ratio with application to sequential hypothesis distinguishing, in: Proceedings of the First world congress of Bernoulli society, Sections 6-8, Moscow -Tula. pp. 246-252. In Russian.

BIBLIOGRAFÍA

- [Novikov (2002)] Novikov A. 2002. Efficiency of sequential hypotheses testing. *Aportaciones matemáticas. Serie Comunicaciones* 30 71-79.
- [Novikov (2004)] Novikov, A., 2004. Optimality of two-stage hypothesis tests. In: *COMPSTAT 2004 Proceedings in Computational Statistics*, ed. J. Antoch, Physica-Verlag, Heidelberg, pp. 1601-1608.
- [Novikov, A., 2005.] Asymptotic optimality of two-stage hypotheses tests. *Aportaciones Matemáticas, Serie Comunicaciones* 35, 37 - 43.
- [Novikov (2006)] Novikov, A., 2006. Locally most powerful two-stage tests. In: *PRAQUE STOCHASTICS 2006. Proc. of the joint session of 7th Prague Symp. on Asymp. Statist. and 15th Prague Conf. on Inform. Theory, Statist. Decision Functions and Random Proc.*, Prague, August 21 to 25, 2006, Prague: MATFYZ-PRESS, Charles University in Prague, pp. 554-567.
- [Novikov, A., (2008)] Optimal sequential tests for two simple hypotheses based on independent observations. *International Journal of Pure and Applied Mathematics* 45, 291-314.
- [Novikov (2009)] Novikov, A., 2009. Optimal sequential tests for two simple hypotheses. *Sequential Analysis* 28 (2), 188-217.
- [Schmitz(1993)] Schmitz, N., 1993. *Optimal Sequentially Planned Decision Procedures. Lecture Notes in Statistics* 79, Springer-Verlag, New York.
- [R Core Team, 2013.] R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- [Roussas, G.G., 1972] *Contiguity of probability measures: some applications in statistics.* Cambridge university press.
- [Schmitz (1993)] Schmitz, N., 1993. *Optimal Sequentially Planned Decision Procedures. Lecture Notes in Statistics; 79, Springer-Verlag, New York.*
- [Stadje, W., 1984.] Asymptotic efficiency of sequential testing using error sums. *Sankhya: The Indian Journal of Statistics. Series A, Part 1* 46, 117-134.
- [Schmitz and Sueselbeck (1983)] Schmitz, N., Süselbeck, B., 1983. Sequential probability ratio tests for homogeneous Markov chains. *Lecture Notes in Statistics* 20, 181-202.
- [Shiryaev] Shiryaev, A.N., 1978. *Optimal stopping rules.* Springer, Berlin.
- [Wald, A., Wolfowitz, J., (1948)] Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics* 19, 326 - 339.
- [Wald (1945)] Wald, A., 1945. *Sequential Tests of Statistical Hypotheses.* *Annals of Mathematical Statistics* 16, 117-186.

- [Wald (1947)] Sequential Analysis. New York: John Wiley and Sons. 1947.
- [Wald (1950)] Wald, A., 1950. Statistical Decision Functions, Wiley, New York.
- [Weiss (1962)] Weiss, L. On sequential tests which minimize the maximum expected sample size. *Journal Amer. Statist. Assoc.* 57(1962), 551-566.