

COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

COMPARACIÓN DE PRUEBAS EXACTAS Y ASINTÓTICAS DE NO INFERIORIDAD PARA DOS PROPORCIONES INDEPENDIENTES

CECILIA RAMÍREZ FIGUEROA

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA OBTENER
EL GRADO DE:

DOCTORA EN CIENCIAS

MONTECILLO, TEXCOCO, EDO DE MÉXICO

2008

La presente tesis, titulada: "Comparación de pruebas exactas y asintóticas de no inferioridad para dos proporciones independientes", realizada por la alumna: "Cecilia Ramírez Figueroa", bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

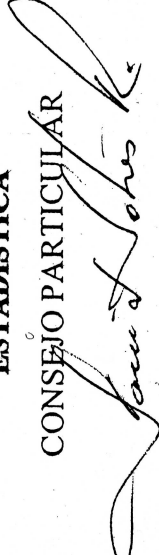
DOCTORA EN CIENCIAS

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA

ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO



DR. DAVID A. SOTRES RAMOS

ASESOR



DR. FÉLIX ALMENDRA ARAO

ASESOR



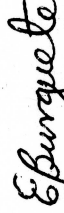
DR. JOSÉ A. VILLASEÑOR ALVA

ASESOR



DR. HUMBERTO VAQUERA HUERTA

ASESOR



DR. ESTEBAN BURGUETE HERNÁNDEZ

COMPARACIÓN DE PRUEBAS EXACTAS Y ASINTÓTICAS DE NO INFERIORIDAD PARA DOS PROPORCIONES INDEPENDIENTES

CECILIA RAMÍREZ FIGUEROA, DRA.

COLEGIO DE POSGRADUADOS, 2008

RESUMEN

Las pruebas estadísticas de no inferioridad son procedimientos de prueba contruidos con el objetivo de demostrar que un tratamiento nuevo con menores efectos secundarios o menor costo, no es substancialmente inferior que otro tratamiento control el cual se sabe que es efectivo.

Existen varias pruebas de no inferioridad que son usadas en la práctica; sin embargo, hasta ahora no se sabe cuáles de éstas tienen buen control del nivel de significancia y alta potencia cuando el tamaño de muestra es pequeño.

La presente investigación es para comparar los niveles de significancia reales y las potencias de las pruebas exactas y asintóticas de no inferioridad de Blackwelder, Farrington-Manning, Böhning-Viwatwongkasem, Hauck-Anderson, la prueba de razón de verosimilitudes y dos variantes de estas pruebas.

Las pruebas se comparan en un amplio escenario que incluye tamaños de muestra que van desde 5 hasta 100 para diseños balanceados (iguales tamaños de muestra), para cinco márgenes de no inferioridad y para tres niveles de significancia nominales. El cálculo del nivel de significancia real se efectuó enumerando todos los puntos muestrales posibles. Algunas pruebas que se utilizan en la práctica resultaron ser conservadoras en exceso: Blackwelder, Böhning-Viwatwongkasem y Hauck-Anderson. Se destaca la prueba de Farrington-Manning exacta como la que tuvo un mejor desempeño en cuanto a su nivel de significancia real y a su potencia.

Palabras clave: Pruebas estadísticas, no inferioridad, diferencia de proporciones independientes

COMPARACIÓN DE PRUEBAS EXACTAS Y ASINTÓTICAS DE NO INFERIORIDAD PARA DOS PROPORCIONES INDEPENDIENTES

CECILIA RAMÍREZ FIGUEROA, DRA.

COLEGIO DE POSGRADUADOS, 2008

ABSTRACT

Non-inferiority statistical tests are testing procedures built with the objective of demonstrating that a new treatment (with fewer secondary adverse reactions or less cost) is not substantially inferior in efficacy than the gold standard with proved efficacy.

There exist several non-inferiority tests that can be used in practice; however, so far it is not known which of these tests have a good significance level control and high power when the sample size is small.

The aim of this research is to compare the actual significance levels and powers of seven exact non-inferiority tests for two binomial proportions.

We compare the tests in a broad scenario which includes sample sizes from 5 to 100 for balanced designs, for five non-inferiority margins and for three levels of nominal significance. The calculation of the actual significance level was obtained by considering all possible sample points. Some tests that are used in practice were too conservative: Blackwelder, Böhning-Vivatwongkasem y Hauck-Anderson. The Farrington-Manning exact test was highlighted as the best in terms of its actual significance level and its power.

Key words: statistical test, non inferiority, difference of two independent proportions

AGRADECIMIENTOS

Al Colegio de Postgraduados, por todas las condiciones que me proporcionó para continuar mi formación académica.

Al Consejo Nacional de Ciencia y Tecnología, por brindarme los recursos necesarios para realizar los estudios de Doctorado.

A mis profesores, por la paciencia y confianza que me brindaron a lo largo de mi estancia en el colegio.

Al Dr David Sotres Ramos por su paciencia, sus valiosos comentarios y su acertada dirección.

Al Dr. Félix Almendra Arao por su infinito apoyo, por el tiempo y dedicación que me brindó y sin el cual me hubiera sido mas difícil la investigación.

Al Dr. José Villaseñor Alva, quien me ha ayudado a tener confianza en mis conocimientos y ha sido muy generoso con su tiempo, sus ideas y su gran amabilidad

Al Dr. Humberto Vaquera Huerta por sus observaciones y haber revisado el escrito.

Al Dr. Esteban Burguete Hernández por sus atinadas sugerencias e invaluable apoyo incondicional y continuo.

A todo el personal secretarial y administrativo, de quienes recibí mucho apoyo.

A todos mis amigos del Colegio de Postgraduados por su eterna disposición para ayudarme.

A mi querida familia, con la que nací y a la que yo elegí, por todo lo que me han dado siempre y por el apoyo incondicional eterno.

A mi Mamá

CONTENIDO

CAPITULO 1. INTRODUCCIÓN.....	1
1.1. Antecedentes.....	1
1.2. Revisión de literatura.....	2
1.3. Objetivos	4
1.4. Organización de la tesis.....	5
CAPITULO 2. PRUEBAS DE NO INFERIORIDAD.....	6
2.1. Generalidades	6
2.2. Modelo binomial	8
2.3. El problema de calcular el nivel de significancia exacto	11
CAPITULO 3. PRUEBAS SELECCIONADAS.....	14
3.1. Pruebas exactas.....	15
3.1.1. La estadística de blackwelder (T_1)	15
3.1.2. Prueba de Farrington y Manning (T_2)	18
3.1.3. Prueba de Böhning-Viwatwongkasem (T_3)	20
3.1.4. Prueba de Hauck y Anderson (T_4)	22
3.1.5. Una combinación de la Prueba de Farrington Manning con la Prueba de Hauck-Anderson (T_5)	24
3.1.6. Una combinación de la Prueba de Böhning-Viwatwongkasem con la Prueba de Hauck-Anderson (T_6)	24
3.1.7. La estadística de razón de verosimilitudes generalizada (T_7).....	25
3.2. Regiones críticas de las pruebas exactas.....	25
3.3. Pruebas asintóticas	28
3.4. Regiones críticas de las pruebas asintóticas.....	32
3.4.1. Correcciones por continuidad.....	34

CAPITULO 4. METODOLOGIA PARA EL NIVEL DE SIGNIFICANCIA ..35

- 4.1. Resultados que permiten simplificar el cálculo del nivel de significancia...36
- 4.2. Método de cálculo del nivel de significancia39

CAPITULO 5. RESULTADOS Y DISCUSION41

- 5.1. Nivel de significancia de las pruebas exactas41
 - 5.1.1.Prueba T_1 43
 - 5.1.2.Prueba T_2 44
 - 5.1.3.Prueba T_3 45
 - 5.1.4.Prueba T_4 46
 - 5.1.5.Prueba T_5 47
 - 5.1.6.Prueba T_6 48
 - 5.1.7.Prueba T_7 49
- 5.2. Potencia de las pruebas exactas.....51
- 5.3. Pruebas exactas versus pruebas asintóticas57
- 5.4. Tablas de la prueba T_2 exacta.....61

CAPITULO 6. CONCLUSIONES70

- 6.1. Recomendaciones72
- 6.2. Investigación futura.....72

BIBLIOGRAFIA73

APÉNDICE I. SIMBOLOGÍA UTILIZADA77

APÉNDICE II. PROGRAMAS DE CÓMPUTO79

ÍNDICE DE FIGURAS

Figura 1. Región de rechazo original (r original) y región de rechazo redefinida (r redefinida) de la estadística de la prueba t_1 con $n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$	17
Figura 2. Región de rechazo original (R original) y Región de rechazo redefinida (R redefinida) de la estadística T_4 con $n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$	23
Figura 3. Región de rechazo (RR) y región de rechazo restringida al espacio muestral lícito (RR eml) para T_7 , $n=20$ y $d_0=0.15$	33
Figura 4. Nivel de Significancia para la prueba T_1 según el tamaño de muestra, $\alpha=0.01$	43
Figura 5. Nivel de Significancia para la prueba T_2 según el tamaño de muestra, $\alpha=0.01$	44
Figura 6. Nivel de Significancia para la prueba T_3 según el tamaño de muestra, $\alpha=0.01$	45
Figura 7. Nivel de Significancia para la prueba T_4 según el tamaño de muestra, $\alpha=0.01$	46
Figura 8. Nivel de Significancia para la prueba T_5 según el tamaño de muestra, $\alpha=0.01$	48
Figura 9. Nivel de Significancia para la prueba T_6 según el tamaño de muestra, $\alpha=0.01$	49
Figura 10. Nivel de Significancia para la prueba T_7 según el tamaño de muestra, $\alpha=0.01$	50
Figura 11. Potencia exacta de las pruebas T_2 y T_5 como función del tamaño de muestra para $d_0=0.15$, $\alpha=0.05$, $p_1=p_2=0.10$	56
Figura 12. Niveles de significancia de la prueba T_2 en sus versiones exacta, sin corrección por continuidad (sin cc) y con corrección por	

continuidad (con cc) para el nivel nominal de $\alpha = 0.01$ y el margen de no inferioridad de $d_0 = 0.15$ 59

Figura 13. Niveles de significancia de la prueba T_2 en sus versiones exacta,

sin corrección por continuidad (sin cc) y con corrección por

continuidad (con cc) para el nivel nominal de $\alpha = 0.05$ y el margen

de no inferioridad $d_0 = 0.15$ 60

ÍNDICE DE TABLAS

Tabla 1. Valores redefinidos para $\hat{\sigma}_1(x, y)$	16
Tabla 2. Porcentaje(*) de los Niveles de significancia que pertenecen al intervalo $[0.80\alpha, \alpha]$	42
Tabla 3. Pruebas para las cuales las muestras colocan más del 90% de sus niveles de significancia reales en un intervalo de $[0.8\alpha, \alpha]$	51
Tabla 4. Niveles de significancia reales para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.10$	53
Tabla 5. Niveles de significancia reales para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.15$	54
Tabla 6. Niveles de significancia reales para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.20$	55
Tabla 7. Porcentaje de niveles de significancia reales que pertenecen al intervalo $[.04, .05]$	57
Tabla 8. Porcentajes de los niveles de significancia reales que pertenecen al intervalo especificado para la prueba T_2 en su versión exacta y asintótica sin y con corrección por continuidad para el nivel nominal $\alpha=0.01$	58
Tabla 9. Porcentajes de los niveles de significancia reales que pertenecen al intervalo especificado para la prueba T_2 en su versión exacta y asintótica sin y con corrección por continuidad para un nivel nominal de $\alpha=0.05$	60
Tabla 10. Constantes críticas (c_α) y Niveles de Significancia Reales (R.L.S.) Para T_2 exacta, con $\alpha = 0.01$	62
Tabla 11. Constantes críticas (c_α) y Niveles de Significancia Reales (R.L.S.) Para T_2 exacta, con $\alpha = 0.05$	66

CAPITULO 1. INTRODUCCIÓN

1.1. Antecedentes

En investigación clínica, frecuentemente es necesario comparar los efectos de dos tratamientos, y en los casos donde es conocida la eficacia de un tratamiento tradicional, los ensayos de placebo controlado están cayendo en desuso por cuestiones de ética. Esta situación se está volviendo muy común, especialmente en tratamiento de enfermedades cardiacas, cáncer y otras enfermedades crónicas, por lo que los investigadores están incrementando el uso de diseños de no inferioridad, en los cuales un tratamiento en prueba es comparado con un tratamiento tradicional (Gao y Ware, 2008).

Debido a que el tratamiento control se presume que es efectivo, no es necesario demostrar que el tratamiento en prueba es más efectivo (superior) que el tratamiento control. Más bien, el tratamiento en prueba debe ofrecer una formulación más conveniente, menores efectos secundarios u otras características que lo hagan una alternativa mejor para el tratamiento estándar. Por lo tanto, el objetivo del ensayo de no inferioridad es demostrar que el tratamiento nuevo que se está probando no es “sustancialmente inferior” en eficacia al tratamiento estándar, (Chen et al., 2000).

1.2. Revisión de literatura

Las pruebas de no inferioridad para comparar dos proporciones independientes se pueden estudiar desde la perspectiva de pruebas asintóticas (Chen et al., 2000; Li and Chuang-Stein, 2006. Almendra y Sotres, 2008) o desde el enfoque de métodos exactos (Römel y Mansmann, 1999; Chan, 1998; Chan, 1999; Chan, 2003; Skipka *et al.*, 2004).

Las pruebas de no inferioridad se pueden especificar para el riesgo relativo o para la razón de momios (Chan, 1998; Munk *et al.*, 2005), además del tema de la presente investigación que es la diferencia de proporciones.

Las comparaciones de pruebas de no inferioridad para dos proporciones independientes se han realizado con base en simulaciones o en cálculos sobre una porción arbitraria del

espacio paramétrico (Chen *et al.*, 2000; Chan, 2003; Skipka *et al.*, 2004) y no se han encontrado análisis comparativos de pruebas exactas de no inferioridad y en consecuencia tampoco de dichas pruebas exactas contra pruebas asintóticas donde se efectúe el cálculo de los niveles de significancia reales y de las potencias. La razón tal vez sea la gran cantidad de tiempo de cómputo necesario para el cálculo.

Almendra (2007) realiza una revisión de las omisiones de trabajos que analizan el nivel de significancia y enfatiza la importancia de la publicación del trabajo de Röhmel (2005) que realiza una demostración formal de los procedimientos para obtener el nivel de significancia exacto que se habían utilizado sin una justificación formal. La dificultad del cálculo de los niveles de significancia para pruebas de no inferioridad radica en la existencia de parámetros perturbadores, que de acuerdo con Basu (1977), “la eliminación de parámetros perturbadores es universalmente reconocido como un problema mayor en estadística”.

Röhmel y Mansmann (1999) retomaron la investigación que hizo Barnard (1947) en la eliminación de los parámetros perturbadores y la adaptaron al problema de pruebas de no inferioridad. Posteriormente Röhmel (2005) investigó la teoría de las pruebas de no inferioridad y obtuvo resultados importantes del procedimiento para la prueba exacta de no inferioridad de Farrington y Manning (1990). Estas contribuciones permiten reducir el tiempo de cómputo porque el nivel de significancia no se calcula como el supremo en todo el espacio nulo sino como el máximo en una parte de la frontera del espacio nulo.

Almendra (2007) generalizó el teorema demostrado por Röhmel y Mansmann haciéndolo extensivo a pruebas asintóticas y también en el sentido de facilitar su aplicación con condiciones menos restrictivas.

1.3. Objetivos

Comparar los niveles de significancia y las potencias de pruebas exactas de no inferioridad para dos proporciones independientes de Blackwelder, Farrington-Manning, Böhning-Viwatwongkasem, Hauck-Anderson, la prueba de razón de verosimilitudes generalizada y dos variantes de estas pruebas.

Comparar los niveles de significancia y las potencias de pruebas exactas contra pruebas asintóticas de no-inferioridad para dos proporciones independientes de Blackwelder, Farrington-Manning, Böhning-Viwatwongkasem, Hauck-Anderson, la prueba de razón de verosimilitudes generalizada y dos variantes de estas pruebas.

1.4. Organización de la tesis

La organización de esta tesis es la siguiente inicia con la introducción en el capítulo uno. Posteriormente, en el capítulo 2 se describirán generalidades de las pruebas de no inferioridad desde la aplicación de este tipo de pruebas en los diseños de ensayos clínicos hasta la especificación de las hipótesis. Posteriormente se precisa el modelo binomial y el nivel de significancia de una prueba con el problema que existe en su cálculo por la existencia de parámetros perturbadores y se describe una solución por el método de maximización sobre los parámetros perturbadores.

En el capítulo 3 se realiza la descripción de las pruebas estudiadas en su versión exacta y asintótica, así como las regiones críticas. En el capítulo 4 se explica detalladamente la forma en que se obtienen los niveles de significancia y se presentan los resultados que permiten simplificar los cálculos necesarios para comparar las pruebas. En el capítulo 5 se presentan los resultados de los niveles de significancia y de la potencia de las pruebas estudiadas. En el capítulo 6 se presentan las conclusiones.

CAPITULO 2. PRUEBAS DE NO INFERIORIDAD

2.1. Generalidades

La necesidad de comparar grupos en muchos campos de la ciencia es un tema de constante investigación. En los campos de la biología y de medicamentos el ensayo de placebo controlado es el ensayo estándar. En la evaluación de muchos tratamientos nuevos, sin embargo, el uso de ensayos de placebo controlado para la evaluación de nuevos tratamientos médicos provoca problemas éticos. Los ensayos de no inferioridad son la principal alternativa a este problema. En lugar del placebo, el grupo control es el control activo, es decir un tratamiento de eficacia demostrada. El objetivo de los ensayos de no inferioridad de control activo no es demostrar superioridad de un tratamiento nuevo experimental o de investigación sobre el control activo. Más bien es establecer que el efecto del tratamiento experimental cuando se compara con el control activo, no está por debajo de un margen de no inferioridad. Muchas cuestiones acompañan estos ensayos, también formas de diseñarlos, implementarlos, analizarlos e interpretarlos (Rodary et al. 1989).

Un ensayo de no inferioridad es un ensayo diseñado con el objetivo de mostrar que el tratamiento experimental no es sustancialmente inferior en eficacia al tratamiento estándar. El nuevo producto puede ofrecer ventajas de seguridad. Por ejemplo, un producto nuevo anti-infeccioso que no presente resistencia a bacterias, un producto para males respiratorios de infantes prematuros sintético como alternativa a un derivado animal que presente más riesgos, un nuevo inhalador para el tratamiento del asma puede evitar contener CFC's en contraste con los productos estándar (Ebutt y Frith, 1998) . En el caso de tratamientos para el VIH, los productos nuevos pueden tener un régimen más simple de promover la adherencia y potencialmente reducir la resistencia. También es posible que los costos, la mercadotecnia y ganancias potenciales sean las razones subyacentes. Por ejemplo los costos del nuevo producto pueden ser menores o el promotor puede mejorar el acceso al mercado.

El razonamiento estadístico para evaluar no inferioridad (y equivalencia) se encuentra en el artículo de Blackwelder (1982). A continuación se presenta un breve resumen. Sea p_2 la probabilidad de eficacia para tratamiento nuevo o experimental, similarmente sea p_1 la probabilidad de eficacia del control activo o tratamiento de referencia. Las hipótesis nula y alternativa para probar no inferioridad son:

$$H_0 : p_1 - p_2 \geq d_0 \text{ Versus } H_1 : p_1 - p_2 < d_0 \quad (2.1)$$

donde d_0 conocido es el margen de no inferioridad, esto es, por cuanto puede exceder p_1 a p_2 para seguir considerando no inferior a p_2 con respecto a p_1 ($d_0 > 0$). La hipótesis nula establece que el control activo p_1 excede al tratamiento de referencia por al menos d_0 ; si

esta hipótesis no es rechazada entonces el control activo es considerado superior al tratamiento experimental con respecto a la eficacia. La hipótesis alternativa establece que el control activo puede tener mejor eficacia que el tratamiento experimental, pero no en más que d_0 . En tal caso, se puede decir que el producto investigado es no inferior al control activo. Es decir, rechazar la hipótesis nula implica concluir no inferioridad del tratamiento nuevo sobre el tratamiento estándar o conocido.

2.2. Modelo binomial

Si para comparar la eficacia de un tratamiento nuevo contra un tratamiento estándar se diseña un ensayo clínico de no inferioridad y se requiere extraer dos muestras aleatorias e independientes (n_1, n_2) de dos poblaciones distintas y se observa el número de individuos que de cada muestra tienen cierta propiedad dicotómica (éxito o fracaso), entonces los resultados comúnmente se presentan en el formato de una tabla 2x2, o tabla de contingencia como la siguiente:

Respuesta	Muestra 1	Muestra 2
Éxito	x_1	x_2
Fracaso	$n_1 - x_1$	$n_2 - x_2$
Total	n_1	n_2

Si (X_1, X_2) son las variables aleatorias y (x_1, x_2) son sus valores experimentales o realizaciones, entonces este tipo de muestreo se ajusta a un modelo en que las variables X_i ($i=1,2$) se distribuyen como binomiales con parámetros n_i y p_i , donde p_1 y p_2 son las

proporciones de individuos con respuesta exitosa de la población 1 y en la población 2, respectivamente.

Por lo anterior, la probabilidad de observar una configuración como la de la tabla de arriba está dada por la función de densidad binomial conjunta de las dos variables aleatorias independientes (X_1, X_2)

$$P(x_1, x_2 | n_1, n_2, p_1, p_2) = \binom{n_1}{x_1} p_1^{x_1} (1-p_1)^{n_1-x_1} \binom{n_2}{x_2} p_2^{x_2} (1-p_2)^{n_2-x_2} \quad (2.2)$$

En los ensayos de no inferioridad, la hipótesis de interés que se desea probar (hipótesis de no inferioridad) es la hipótesis alternativa en el siguiente juego de hipótesis:

$$H_0 : p_1 - p_2 \geq d_0 \text{ versus } H_a : p_1 - p_2 < d_0 \quad (2.3)$$

donde d_0 es el margen de no inferioridad y es una constante positiva conocida. En el contexto de ensayos clínicos los valores usuales para d_0 son 0.10, 0.15 y 0.20, FDA (1992, 1998). Cuando $d_0=0$ el problema se reduce a la hipótesis clásicas de una cola.

Eentonces de acuerdo con el modelo binomial el espacio muestral es $\mathcal{X} = \{(x_1, x_2) \in \{0, \dots, n_1\} \times \{0, \dots, n_2\}\}$ y se define el espacio muestral lícito como

$$\mathcal{X}' = \{(x_1, x_2) \in \mathcal{X} : \hat{p}_1 - \hat{p}_2 < d_0\} \text{ donde } \hat{p}_i = X_i / n_i \text{ (i=1,2)}$$

es el estimador de máxima verosimilitud para p_i .

Se destaca la importancia del espacio muestral lícito observada por Martin Andrés y Herranz Tejedor (2004)a, donde los autores enfatizan que de no restringirse a esta condición del espacio muestral se pueden obtener conclusiones falsas, por ejemplo en los casos muestrales donde $\hat{p}_1 - \hat{p}_2 \geq d_0$ y si la prueba indica rechazar la hipótesis nula es ilógico concluir que $p_1 - p_2 < d_0$. El espacio de parámetros es $\Theta = \{(p_1, p_2) \in [0,1]^2\}$ y se encuentra representado gráficamente en la Figura 1.

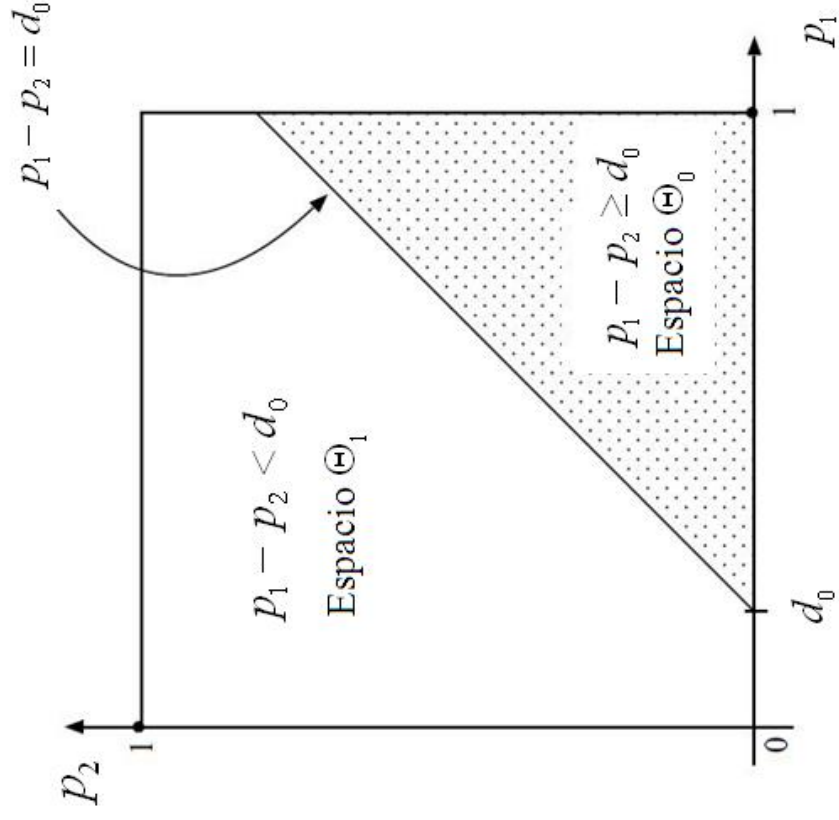


Figura 1. Espacio de parámetros de la región de Θ_0 y de Θ_1

2.3. El problema de calcular el nivel de significancia exacto

Bajo la hipótesis nula ($H_0 : p_1 - p_2 \geq d_0$) la probabilidad de observar un punto muestral particular $((X_1, X_2) = (x_1, x_2))$ está dada por la función de verosimilitud conjunta

$$L(p_1, p_2; x_1, x_2, n_1, n_2) = \binom{n_1}{x_1} p_1^{x_1} (1 - p_1)^{n_1 - x_1} \binom{n_2}{x_2} p_2^{x_2} (1 - p_2)^{n_2 - x_2} \quad (2.4)$$

y la función de potencia de una prueba T es

$$\beta_T(p_1, p_2) = \sum_{(x_1, x_2) \in R_T(\alpha)} L(p_1, p_2; x_1, x_2) \quad (2.5)$$

donde $R_T(\alpha)$ es la región crítica de una prueba T para un nivel de significancia nominal

(α) y el tamaño de una prueba T está definido por $\sup_{(p_1, p_2) \in \Theta_0} \beta_T(p_1, p_2)$ con

$$\Theta_0 = \{(p_1, p_2) \in \Theta : p_1 - p_2 \geq d_0\}$$
 como el espacio nulo.

El problema de calcular el nivel de significancia exacto de la prueba se puede enfrentar de dos maneras. La primera tiene que ver con el enfoque de pruebas condicionadas, como por ejemplo la prueba exacta de Fisher que se utiliza en el caso particular en donde $d_0 = 0$. La segunda forma de enfrentar el problema es la aplicación de métodos no condicionados como lo es el principio del máximo, detallado más adelante.

Se define al nivel de significancia como:

$$\sup_{\theta \in \Theta_0} P_\theta(T \leq t_0) \quad (2.6)$$

(ver Mood, et al., 1974)

donde t_0 es un valor observado de la estadística de prueba T y Θ_0 es el espacio nulo.

Al usar $T(x_1, x_2; d_0)$ como una estadística de prueba seleccionada, valores pequeños proporcionan evidencia en contra de la hipótesis nula. Aplicando (2.6), se puede llevar a cabo una prueba no condicionada resultando el tamaño de la prueba expresado mediante:

$$\sup_{(p_1, p_2) \in \Theta_0} \beta_T(p_1, p_2) \quad (2.7)$$

Note que el nivel de significancia se determina al encontrar el supremo de la función de potencia sobre el espacio nulo, el cual es un espacio de dos dimensiones, dicho cálculo implica un alto consumo de tiempo al ser una búsqueda computacional intensiva. Sin embargo, bajo ciertas condiciones (que se especificarán más adelante) es posible probar que el supremo en (2.7) es un máximo y se alcanza en una parte de la frontera, ver por ejemplo Röhmel (2005). Esto es, el supremo ocurre sobre el conjunto $\{(p_1, p_2) : p_1 - p_2 = d_0\}$. Por lo tanto, se puede simplificar el cálculo del nivel de significancia en (2.7), obteniéndose en nivel de significancia exacto, α^* , que es el nivel de la prueba más cercano a α que nos produce el valor crítico.

$$\alpha^* = \max_{\{(p_1, p_2) : p_1 - p_2 = d_0\}} \left(\sum_{T(x_1, x_2; d_0) \leq T(x_1^*, x_2^*; d_0)} L(p_1, p_2; x_1, x_2, n_1, n_2,) \right) \quad (2.8)$$

Una prueba de nivel α es aquella que rechaza H_0 si y sólo si α^* es menor o igual que el error objetivo α . Esto garantiza que dicha prueba no es liberal, aunque si cabe la posibilidad de que sea una prueba muy conservadora. El algoritmo para el cálculo del nivel de significancia debe ser simplificado para reducir el espacio de la búsqueda, lo cual se logra con los resultados propuestos por Barnard (1947) y demostrados para pruebas de no

inferioridad por Röhmel y Mansmann (1999), Almendra (2008). Los conceptos involucrados en la simplificación de dicho cálculo son las condiciones de convexidad de Barnard (C) y de simetría en la misma cola (SMC), mismos que se estudiarán más adelante. Algunos trabajos de investigación han utilizado estos resultados sin realizar una demostración completa (Chan, 1999; Mehrotra et al. 2003).

CAPITULO 3. PRUEBAS SELECCIONADAS

Las pruebas estudiadas tienen como regla de decisión: rechazar la hipótesis nula cuando $T < C_\alpha$, donde T es la estadística de prueba y C_α es el valor crítico para un nivel α dado.

Seis de las estadísticas de prueba analizadas son de la forma siguiente:

$$T_i(x_1, x_2) = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2} - d_0}{\hat{\sigma}_i} \quad (3.1)$$

para $i = 1, 2, 3, 4, 5, 6$; donde $\hat{\sigma}_i$ es un estimador de la desviación estándar de $d = \hat{p}_1 - \hat{p}_2$ con

$p_i = \frac{x_i}{n_i}$ para $i = 1, 2$. La séptima estadística es aquella que corresponde a la prueba de razón de verosimilitudes generalizada.

Las primeras seis estadísticas de prueba consideradas en este trabajo (ver (3.1)) se diferencian únicamente en que utilizan diferentes estimadores de la desviación estándar $\hat{\sigma}_i$.

3.1. Pruebas exactas

Las pruebas exactas están dirigidas a controlar el error tipo I, es decir el nivel verdadero α^* de estas pruebas nunca excede el nivel nominal asignado α . En esta investigación se estudian las siete pruebas siguientes: Blackwelder, Farrington-Manning, Böhning-Viwatwongkasen, Hauck-Anderson, dos combinaciones de las anteriores y la de razones de verosimilitudes generalizada.

3.1.1. La estadística de blackwelder (T_1)

Esta prueba presenta la gran ventaja de simplicidad de cálculo y además que los estimadores de proporción (x_i / n_i) son los estimadores de máxima verosimilitud.

La estadística de prueba se define como:

$$T_1(x_1, x_2) = \frac{\hat{p}_1 - \hat{p}_2 - d_0}{\hat{\sigma}_1}$$

donde $\hat{\sigma}_1$ es el estimador consistente de la desviación estándar de $d = \hat{p}_1 - \hat{p}_2$ dado por

$$\hat{\sigma}_1 = \left(\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2} \right)^{1/2}$$

La estadística T_1 se construye usando la estimación clásica para la desviación estándar, nótese que $\hat{\sigma}_1 = 0$ para $\hat{p}_i = 0, 1$, en estos casos T_1 no está definida. Por ello Almendra (2008) propone redefinir a $\hat{\sigma}_1(x_1, x_2)$ en aquellos puntos del espacio muestral donde se anula, de la siguiente forma:

$$\text{Observación 1. Sea } \varphi(n, x) = \frac{x \left(1 - \frac{x}{n} \right)}{n} = \frac{x(n-x)}{n^2}$$

Como $\hat{\sigma}_1(x, y) = 0$ para $(x, y) \in \{0, n_1\} \times \{0, n_1\}$ entonces $T_1(x, y)$ no está definida para $(x, y) \in \{0, n_1\} \times \{0, n_1\}$ por lo tanto, la relación propuesta esta dada en la Tabla 1:

Tabla 1. Valores redefinidos para $\hat{\sigma}_1(x, y)$

Valores del espacio muestral	Valores redefinidos para $\hat{\sigma}_1(x, y)$
$(x = 0, y = 0)$	$(x = .01, y = .01) \quad \sqrt{\varphi(n_1, .01) + \varphi(n_2, .01)}$
$(x = 0, y = n_2)$	$(x = .01, y = n_2 - .01) \quad \sqrt{\varphi(n_1, .01) + \varphi(n_2, n_2 - .01)}$
$(x = n_1, y = 0)$	$(x = n_1 - .01, y = .01) \quad \sqrt{\varphi(n_1, n_1 - .01) + \varphi(n_2, .01)}$
$(x = n_1, y = n_2)$	$(x = n_1 - .01, y = n_2 - .01) \quad \sqrt{\varphi(n_1, n_1 - .01) + \varphi(n_2, n_2 - .01)}$

En la Figura 2 se puede apreciar gráficamente un ejemplo de la redefinición de la región crítica de la estadística T_1 . Se ha ilustrado la redefinición para la configuración de

$n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$. Se observa que sin la redefinición la región de rechazo original no incluirían los puntos $(0,0)$, $(0,29)$ y $(29,29)$ con lo que la región de rechazo no cumpliría la condición de convexidad de Barnard y no sería sencillo calcular el nivel de significancia exacto de la prueba. Al modificar la región de rechazo e incluirle los puntos $(0,0)$, $(0,29)$ y $(29,29)$ la región de rechazo cumple la condición de convexidad de Barnard, que posteriormente se definirá a detalle.

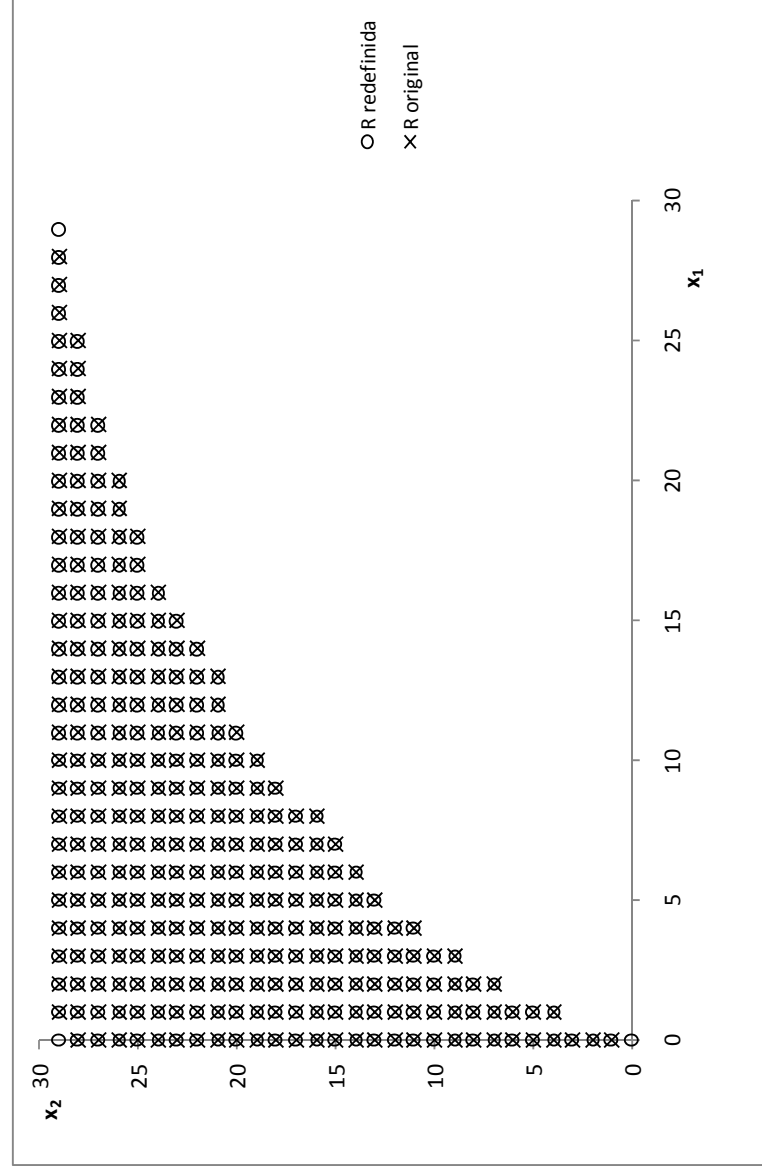


Figura 2. Región de rechazo original (R original) y Región de rechazo redefinida (R redefinida) de la estadística de la prueba T_1 con $n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$

3.1.2. Prueba de Farrington y Manning (T_2)

La estadística de prueba se define como:

$$T_2(x_1, x_2; d_0) = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2} - d_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \quad (3.2)$$

donde \hat{p}_1 y \hat{p}_2 son los estimadores de máxima verosimilitud restringidos a la hipótesis nula para p_1 y p_2 respectivamente, que se describen a continuación.

Dada una realización, $X_1 = x_1$ y $X_2 = x_2$, el logaritmo de la función de verosimilitud conjunta en (2.4) resulta en:

$$\begin{aligned} \log L(p_1, p_2; x_1, x_2) &= \log \binom{n_1}{x_1} + x_1 \log p_1 + (n_1 - x_1) \log(1 - p_1) \\ &\quad + \log \binom{n_2}{x_2} + x_2 \log p_2 + (n_2 - x_2) \log(1 - p_2) \end{aligned}$$

Tomando las derivadas parciales de la log verosimilitud con respecto a p_1 y con la restricción definida,

$$\begin{aligned} \frac{\partial}{\partial p_1} \log L(p_1, p_2; x_1, x_2) &= \frac{x_1}{p_1} - \frac{n_1 - x_1}{1 - p_1} + \frac{x_2}{p_2} - \frac{n_2 - x_2}{1 - p_2} \\ &= \frac{x_1}{p_1} - \frac{n_1 - x_1}{1 - p_1} + \frac{x_2}{p_1 - d_0} - \frac{n_2 - x_2}{1 - p_1 - d_0} \end{aligned}$$

Iguando a cero la derivada

$$(n_1 + n_2) p_1^3 - (n_1 + n_2 + x_1 + x_2 + 2n_1 d_0 + n_2 d_0) p_1^2 + (n_1 d_0^2 + 2x_1 d_0 + n_2 d_0 + n_1 d_0 + x_1 + x_2) p_1 - x_1 d_0 (1 + d_0) = 0$$

Dividiendo ambos lados de la ecuación por n_1

$$ap_1^3 + bp_1^2 + cp_1 + d = 0 \quad (3.3)$$

donde

$$a = 1 + \frac{n_2}{n_1}$$

$$b = - \left(1 + \frac{n_2 + x_1 + x_2}{n_1} + d_0 \left(\frac{n_2 + 2}{n_1} \right) \right)$$

$$c = d_0^2 + d_0 \left(\frac{2x_1 + n_2}{n_1} + 1 \right) + \frac{x_1 + x_2}{n_1}$$

$$d = -d_0(1 + d_0) \frac{x_1}{n_1}$$

Miettinen y Nurminen (1985) y Farrington y Manning (1990) demostraron que el estimador de máxima verosimilitud restringido para p_1 es la solución única para (3.3) donde

$p_1 \in (d_0, 1)$, es decir:

$$\hat{p}_1 = \frac{2}{3} \sqrt{r^2 - 3s} \cos \left(\frac{1}{3} \arccos \left(- \left(\frac{2r^2}{27} - \frac{rs}{3} + t \right) / 2 \left(\frac{\sqrt{r^2 - 3s}}{3} \right)^3 \right) + \frac{4}{3} \pi \right) - \frac{r}{3} \quad (3.4)$$

$$\hat{p}_2 = \hat{p}_1 - d_0$$

donde

$$r = \frac{x_1 + x_2 + n_1(1 + 2d_0) + n_2(1 + d_0)}{n_1 + n_2}$$

$$s = \frac{x_2 + x_1(1 + 2d_0) + d_0(n_2 + n_1(1 + d_0))}{n_1 + n_2}$$

$$t = \frac{-x_1 d_0(1 + d_0)}{n_1 + n_2}$$

La prueba T_2 estima la desviación estándar tomando como estimadores de las proporciones a los estimadores de máxima verosimilitud restringidos a la hipótesis nula

3.1.3. Prueba de Böhning-Viwatwongkasem (T_3)

Aunque los estimadores de proporciones se utilizan en muchas áreas de aplicación Böhning y Viwatwongkasem (2005) insisten que no se les ha estudiado en forma satisfactoria. Así que el estimador de proporción convencional ($p_i = x_i / n_i$), esto es, el de máxima verosimilitud presenta muchos problemas para estimar la varianza en tamaños de muestra pequeños. Por ejemplo, la varianza de \hat{p} cuando se estima con $\frac{(x/n)(1-(x/n))}{n}$ se anula para $x = 0$ y $x = n$. Böhning y Viwatwongkasem sugieren usar una clase de estimadores de

la forma $\tilde{p}_c = \frac{X+b}{n+2b}$ $b > 0$. Existen varias propuestas para el valor de b (1/6, 1/2, 1 ó 2),

aunque son resultado de la mera intuición al añadir observaciones al número de éxitos o al número de fracasos. Otra argumentación llega de la línea bayesiana, cuando se usa $b=1$, porque este estimador corresponde al estimador de Bayes cuando se toma como distribución a priori una distribución Beta (1,1).

Entonces la estadística de la prueba de Böhning-Viwatwongkasen (T_3) es

$$T_3(x_1, x_2; d_0) = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2} - d_0}{\sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2}}}$$

con $\tilde{p}_i = \frac{x_i+1}{n_i+2}$ para $i = 1, 2$. Con esta modificación se logra que la varianza

$\sigma_3 = \sqrt{\frac{\tilde{p}_{1,b}(1-\tilde{p}_{1,b})}{n_1} + \frac{\tilde{p}_{2,b}(1-\tilde{p}_{2,b})}{n_2}}$ sea diferente de 0 para todo $b > 0$, y se evita el

problema de indeterminación en $T_3(x_1, x_2; d_0)$.

Algunas propiedades de la propuesta de Böhning-Viwatwongkasen (para $b = 1$).

- a) $\tilde{p}_{i,b} \in (0,1)$ para toda $b > 0$
- b) $\lim_{b \rightarrow 0} \tilde{p}_{i,b} = \hat{p}_i$
- c) $\lim_{b \rightarrow 0} E(\tilde{p}_{i,b}) = \lim_{b \rightarrow 0} \frac{n_i p_i + b}{n_i + 2b} = p_i = E(\hat{p}_i)$

$$d) \lim_{n_i \rightarrow \infty} E(\tilde{p}_{i,b}) = \lim_{n_i \rightarrow \infty} \frac{n_i p_i + b}{n_i + 2b} = p_i = E(\hat{p}_i)$$

$$e) \text{var}(\tilde{p}_{i,b}) = \frac{n_i p_i (1-p_i)}{(n_i + 2b)^2} \leq \frac{p_i (1-p_i)}{n_i} = \text{var}(\hat{p}_i)$$

$$f) \lim_{b \rightarrow 0} \text{var}(\tilde{p}_{i,b}) = \text{var}(\hat{p}_i)$$

$$g) \lim_{n_i \rightarrow \infty} \text{var}(\tilde{p}_{i,b}) = 0 = \lim_{n_i \rightarrow \infty} \text{var}(\hat{p}_i)$$

3.1.4. Prueba de Hauck y Anderson (T₄)

Hauck y Anderson (1986) estudiaron siete intervalos de confianza basados en la aproximación normal para la diferencia de dos proporciones binomiales de donde resulta la recomendación de usar $(n_i - 1)$ en lugar de n_i en la estimación del error estándar.

Entonces la estadística de prueba para la diferencia de dos proporciones para el problema de probar no inferioridad resulta ser

$$T_4(x_1, x_2; d_0) = \frac{\hat{p}_1 - \hat{p}_2 - d_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1 - 1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2 - 1}}}$$

donde $\hat{p}_i = x_i / n_i$ y $n_i > 1$ para $i = 1, 2$.

Nótese que al igual que la estadística T_1 , la presente estadística no está definida para algunos valores del espacio muestral es decir $\hat{\sigma}_4 = 0$ para $\hat{p}_i = 0, 1$. Entonces se propone redefinir $\hat{\sigma}_4$ de la misma forma que se hizo con $\hat{\sigma}_1$.

En la Figura 3 se pueden observar la región de rechazo original y la región de rechazo redefinida para T_4 con $n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$.

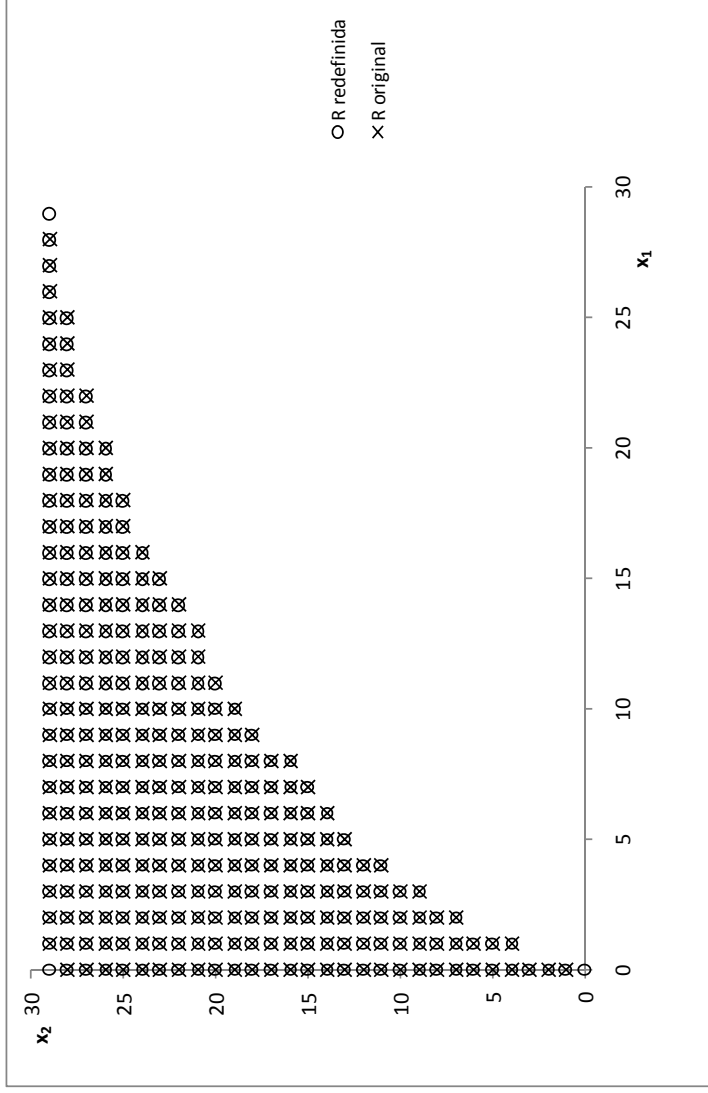


Figura 3. Región de rechazo original (R original) y Región de rechazo redefinida (R redefinida) de la estadística de la prueba T_4 con $n_1=n_2=29$, $d_0=0.15$ y $\alpha=0.05$

3.1.5. Una combinación de la Prueba de Farrington Manning con la Prueba de

Hauck-Anderson (T_5)

La estadística que resulta de una combinación de la estadística de Farrington Manning con la estadística de Hauck-Anderson y que se analizará en la presente investigación es como sigue.

$$T_5(x_1, x_2; d_0) = \frac{\hat{p}_1 - \hat{p}_2 - d_0}{\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1 - 1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2 - 1}}}$$

Con $\hat{p}_i = x_i / n_i$ y \hat{p}_i es el estimador de máxima verosimilitud restringido bajo la hipótesis nula para $i = 1, 2$.

3.1.6. Una combinación de la Prueba de Böhning-Vivattongkasen con la

Prueba de Hauck-Anderson (T_6)

La estadística resultante de esta combinación es como sigue

$$T_6(x_1, x_2; d_0) = \frac{\hat{p}_1 - \hat{p}_2 - d_0}{\sqrt{\frac{\tilde{p}_1(1 - \tilde{p}_1)}{n_1 - 1} + \frac{\tilde{p}_2(1 - \tilde{p}_2)}{n_2 - 1}}}$$

Con $\hat{p}_i = x_i / n_i$, para $i = 1, 2$ y $\tilde{p}_i = \frac{x_i + 1}{n_i + 2}$

3.1.7. La estadística de razón de verosimilitudes generalizada (T₇)

La razón de verosimilitudes para el problema del contraste en (2.3) esta dada por:

$$T_7(X_1, X_2) = \frac{\sup_{H_0} L(p_1, p_2; X_1, X_2)}{\sup_{H_0 \cup H_1} L(p_1, p_2; X_1, X_2)}$$

$$= \begin{cases} 1 & \text{si } \hat{p}_1 \geq \hat{p}_2 + d_0 \\ \frac{\binom{v}{p_1}^{x_1} \binom{v}{1-p_1}^{n_1-x_1} \binom{v}{p_2}^{x_2} \binom{v}{1-p_2}^{n_1-x_1}}{\binom{\hat{p}_1}{p_1}^{x_1} \binom{\hat{p}_1}{1-\hat{p}_1}^{n_1-x_1} \binom{\hat{p}_2}{p_2}^{x_2} \binom{\hat{p}_2}{1-\hat{p}_2}^{n_2-x_2}} & \text{si } \hat{p}_1 < \hat{p}_2 + d_0 \end{cases}$$

con \hat{p}_1 y \hat{p}_2 como en (3.4).

3.2. Regiones críticas de las pruebas exactas

La prueba correspondiente a la estadística T_i para un nivel de significancia nominal especificado α es: rechazar $H_0 : p_1 - p_2 \geq d_0$ en favor de $H_a : p_1 - p_2 < d_0$ si y sólo si se cumple que $T_i(x_1, x_2) \leq T_i^*(x_1^*, x_2^*)$ donde el punto (x_1^*, x_2^*) es tal que

$$T(x_1^*, x_2^*) = \max_{(p_1, p_2) \in \Theta_0} \left\{ T(a, b) : \sup_{T(x_1, x_2) \leq T(a, b)} \sum L(p_1, p_2; x_1, x_2) \right\} \leq \alpha$$

por lo tanto la prueba correspondiente a la estadística T_i tiene región de rechazo

$$R_{T_i}(\alpha) = \left\{ (x_1, x_2) \in \{0, \dots, n_1\} \times \{0, \dots, n_2\} : T(x_1, x_2) \leq T(x_1^*, x_2^*) \right\}$$

Para simplificar la escritura se utilizará la misma notación T_i para referirnos tanto a la prueba de hipótesis como a la estadística de prueba T_i que la genera. De la región crítica de la prueba se calcula una aproximación para α^* que debe cumplir con $\alpha^* \leq \alpha$. Una aproximación para α^* se puede obtener usando un programa de cómputo escrito en el paquete S-PLUS® (los detalles se encuentran en el apéndice II). Para cada tamaño de muestra $n = n_1 = n_2$ se efectúan los siguientes pasos. Para cada prueba T_i , con $i=1, 2, 3, 4, 5, 6, 7$;

- a) Se calcula T_i para todos los puntos del espacio muestral.
- b) El conjunto de todos los resultados posibles de T_i se ordenan en forma creciente.
- c) Se comienza a formar la región crítica (RC) con el punto del espacio muestral que produjo el valor de T_i más pequeño, en caso de que haya empates del valor de T_i se incluyen todos los puntos que produjeron el mismo valor. En este punto también se deben considerar las condiciones específicas de cada una de las pruebas, por ejemplo el espacio muestral lícito.
- d) Se calcula la aproximación del nivel de significancia real.
- e) Si el valor resultante del nivel de significancia de la prueba es menor que el nivel nominal α se le añaden a la RC todos los puntos del espacio muestral que produjeron el siguiente valor de T_i más pequeño.
- f) Se sigue este proceso iterativo hasta obtener una región crítica RC cuyo tamaño α^* se acerque lo más posible, al error objetivo α , sin sobrepasarlo.

Insistiendo en el proceso anterior, si denominamos RC_i a la región crítica RC obtenida en un paso i cualquiera, la siguiente región crítica RC_{i+1} se obtienen de la siguiente forma: Sea C_i el conjunto de puntos candidatos a ser añadidos a la RC_i . Sea C_{ij} cada uno de los subconjuntos de C_i tales que todos sus puntos empatan por la estadística T_i considerada.

Entonces si t_{i_1}, \dots, t_{i_r} son los r valores posibles de $T(x_1, x_2)$ en los puntos de C_i , entonces

$$\text{para toda } j \in \{1, \dots, r\}; C_{ij} = \{(x_1, x_2) \in C_i \mid T(x_1, x_2) = t_{ij}\}$$

Las RC_{ij} serán las r posibles regiones críticas a considerar en el paso $i+1$, pues RC_{i+1} se formará añadiendo a la RC_i aquel punto (o puntos, si es que hay empates) de C_i que tenga o tengan el menor valor de T_i . De esta forma, si $t_{ik} = \min\{t_{ij}, j = 1, \dots, r\}$ se tendrá que

$$RC_{i+1} = RC_{ik}$$

Por lo tanto como lo que se pretende es obtener una región crítica $RC(\alpha)$, para un nivel de significancia objetivo α , el proceso anterior se repite hasta obtener una RC cuyo tamaño α^* se acerque lo más posible, al error objetivo α sin rebasarlo.

Para determinar que se ha finalizado el proceso iterativo y que ya no se deben añadir más puntos a la $RC(\alpha)$, se observa que

$$\text{si } \alpha^*_{i+1} > \alpha \Rightarrow RC(\alpha) = RC_i$$

$$\text{si } \alpha^*_{i+1} = \alpha \Rightarrow RC(\alpha) = RC_{i+1}$$

$$\text{si } \alpha^*_{i+1} < \alpha \Rightarrow RC_i \subseteq RC(\alpha)$$

En las dos primeras situaciones el proceso concluye, y el error real de la prueba es α_i^* en el primer caso y α^*_{i+1} en el segundo. Pero en la última situación el proceso debe continuar (Martin Andrés, A., and A. Silva Mato, 1994).

3.3. Pruebas asintóticas

Las pruebas asintóticas que se estudian utilizan las mismas estadísticas de prueba descritas en el apartado 3.1 que corresponden a las pruebas exactas. Por lo tanto se procederá a describir en detalle las propiedades asintóticas de las estadísticas de prueba.

Sea $N=n_1+n_2$, en relación con $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \zeta \in (0,1)$, un caso muy simple en el que se cumple es

para $n_2 = kn_1$ con $k \in \mathbb{N}$ entonces $\frac{n_1}{N} = \frac{n_1}{n_1 + kn_1} = \frac{1}{1+k}$ por lo tanto $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \frac{1}{1+k} \in (0,1)$.

En los enunciados siguientes se usará la notación $\hat{\sigma}_N = \sqrt{\frac{\hat{\sigma}_{n_1}^2}{n_1} + \frac{\hat{\tau}_{n_2}^2}{n_2}}$

Las demostraciones de los resultados de esta sección se encuentran desarrolladas en Almendra (2007).

Proposición. Sean V_1, \dots, V_{n_1} y W_1, \dots, W_{n_2} variables aleatorias independientes con función de distribución F_1 y F_2 ; medias μ_1 y μ_2 ; varianzas finitas $\sigma^2 > 0$ y $\tau^2 > 0$ respectivamente Sean

$N = n_1 + n_2$; $\sigma_N = \sqrt{\frac{\sigma^2}{n_1} + \frac{\tau^2}{n_2}}$. Supóngase que $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \zeta \in (0,1)$ sean $\{\hat{\sigma}_{n_1}\}$ una sucesión

consistente de estimadores de σ y $\{\hat{\tau}_{n_2}\}$ una sucesión consistente de estimadores de τ y

$\hat{\sigma}_N = \sqrt{\frac{\hat{\sigma}_{n_1}^2}{n_1} + \frac{\hat{\tau}_{n_2}^2}{n_2}}$. Entonces

$$\frac{\bar{V} - \bar{W} - (\mu_1 - \mu_2)}{\hat{\sigma}_N} \xrightarrow{L} N(0,1)$$

Proposición. Sean las variables aleatorias independientes $V_1, \dots, V_{n_1} \sim \text{Bernoulli}(p_1)$ y

$W_1, \dots, W_{n_2} \sim \text{Bernoulli}(p_2)$; $X_1 = \sum_{i=1}^{n_1} V_i$ y $X_2 = \sum_{i=1}^{n_2} W_i$ y supongamos que $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \zeta \in (0, 1)$.

Entonces

$$\frac{\frac{X_1}{n_1} - \frac{X_2}{n_2} - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \xrightarrow{L} N(0, 1)$$

es decir,

$$\frac{\hat{d} - d}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \xrightarrow{L} N(0, 1)$$

Proposición. Sean las variables aleatorias independientes $V_1, \dots, V_{n_1} \sim \text{Bernoulli}(p_1)$;

$W_1, \dots, W_{n_2} \sim \text{Bernoulli}(p_2)$; $X_1 = \sum_{i=1}^{n_1} V_i$ y $X_2 = \sum_{i=1}^{n_2} W_i$ y supongamos que $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \zeta \in (0, 1)$.

Entonces

1. $\{\hat{\sigma}_{n_1} = \hat{p}_1(1 - \hat{p}_1)\}$, $\{\check{\sigma}_{n_1} = \check{p}_1(1 - \check{p}_1)\}$ y $\{\tilde{\sigma}_{n_1} = \tilde{p}_1(1 - \tilde{p}_1)\}$ son sucesiones consistentes

de estimadores de $\sigma = p_1(1 - p_1)$; donde $\hat{p}_1 = \frac{X_1}{n_1}$ es el estimador de máxima

verosimilitud de \check{p}_1 es el estimador de máxima verosimilitud de p_1 restringida bajo la

hipótesis nula y $\tilde{p}_1 = \frac{X_1 + c}{n_1 + 2c}$ con c un número real positivo.

2. $\{\hat{\sigma}_{n_2} = \hat{p}_2(1 - \hat{p}_2)\}$, $\{\check{\sigma}_{n_2} = \check{p}_2(1 - \check{p}_2)\}$ y $\{\tilde{\sigma}_{n_2} = \tilde{p}_2(1 - \tilde{p}_2)\}$ son sucesiones consistentes de estimadores de $\tau = p_2(1 - p_2)$; donde $\hat{p}_2 = \frac{X_2}{n_2}$ es el estimador de

máxima verosimilitud de p_2 , \check{p}_2 es el estimador de máxima verosimilitud de p_2 restringida bajo la hipótesis nula y $\tilde{p}_2 = \frac{X_2 + c}{n_2 + 2c}$ con c un número real positivo.

Proposición. Sean las variables aleatorias independientes $V_1, \dots, V_{n_1} \sim \text{Bernoulli}(p_1)$;

$W_1, \dots, W_{n_2} \sim \text{Bernoulli}(p_2)$, $X_1 = \sum_{i=1}^{n_1} V_i$ y $X_2 = \sum_{i=1}^{n_2} W_i$ y supongamos que $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{N} = \zeta \in (0, 1)$.

Entonces

$$\frac{X_1 + \frac{X_2}{n_2} - (p_1 - p_2)}{\hat{\sigma}_i} \xrightarrow{L} N(0, 1)$$

para $i = 1, 2, 3, 4, 5, 6$.

Puesto que para que se cumpla la convergencia asintótica, esencialmente se ha impuesto en términos generales la condición $\lim_{n_1, n_2 \rightarrow \infty} \frac{n_1}{n_1 + n_2} = \zeta \in (0, 1)$, de aquí en adelante siempre que se trate con resultados asintóticos en relación a las estadísticas mencionadas, se supondrá que se cumple esa condición, nótese también que tal condición es equivalente a

$$\lim_{\min(n_1, n_2) \rightarrow \infty} \frac{n_1}{n_2} = k \in (0, 1) \text{ con } k = \frac{\zeta}{1 - \zeta}.$$

Para el caso particular de esta investigación las estadísticas de prueba comparadas, esto es T_i para $i = 1, \dots, 6$ convergen en distribución a la distribución normal estándar de acuerdo con las proposiciones anteriores. Por otra parte la estadística de prueba T_7 o de razón de verosimilitudes generalizada tiene el comportamiento asintótico siguiente: $-2 \ln T_7(X_1, X_2)$ se distribuye como $\chi^2_{1-2\alpha}(1)$, una chi cuadrada con un grado de libertad, Munk et al. (2005).

3.4. Regiones críticas de las pruebas asintóticas

Las seis pruebas asintóticas para las correspondientes estadísticas del tipo (3.1) para un nivel de significancia nominal α se describen a continuación.

Sea $d = p_1 - p_2$, entonces la prueba T_1 consiste en rechazar $[H_0 : d \geq d_0]$ en favor de $[H_a : d < d_0]$ (ver 2.1) si y sólo si

$$T_i(x_1, x_2) < z_\alpha$$

donde:

$$d = p_1 - p_2$$

z_α es el percentil superior α de la distribución normal estándar, es decir, $P(Z > z_\alpha) = \alpha$;

para $i=1, 2, 3, 4, 5, 6$ donde Z es una variable aleatoria con distribución normal estándar.

Entonces la región de rechazo es de la forma $R_{T_i}(\alpha) = \{(x_1, x_2) \in \mathcal{X} : T_i(x_1, x_2) < -z_\alpha\}$

Nótese que estas pruebas son consistentes para el caso $n_1=n_2$ ya que si se denota a \hat{d} mediante \hat{d}_n , a $\hat{\sigma}$ mediante $\hat{\sigma}_n$ y a cualquiera de estas seis pruebas por W_n , entonces

$$W_n = \frac{\hat{d}_n - d_0}{\hat{\sigma}_n} = \frac{\hat{d}_n - d}{\hat{\sigma}_n} + \frac{d - d_0}{\hat{\sigma}_n} \xrightarrow[n \rightarrow \infty]{} -\infty$$

ya que recuérdese que $\frac{\hat{d}_n - d}{\hat{\sigma}_n}$ tiene distribución normal

estándar; por lo tanto $P(W_n < -z_\alpha) \xrightarrow[n \rightarrow \infty]{} 1$.

La prueba asintótica para la estadística del tipo (2.3) es:

Prueba T_7 : rechazar $H_0 : d \geq d_0$ en favor de $H_a : d < d_0$ (ver 2.1) si y sólo si $-2 \ln T_7(x_1, x_2) > \chi_{1-2\alpha}^2(1)$ y en consecuencia tiene región de rechazo de la forma

$R_T(\alpha) = \{(x_1, x_2) \in \mathcal{X} : -2 \ln T_T(x_1, x_2) > \chi_{1-2\alpha}^2(1)\}$ donde $\chi_{1-2\alpha}^2(1)$ es el percentil superior $1 - 2\alpha$ de la distribución ji-cuadrada con un grado de libertad, es decir, $P(\chi_1^2 > \chi_{1-2\alpha}^2(1)) = 1 - 2\alpha$.

Almendra (2007) demuestra lo siguiente.

Proposición. Para estadísticas tipo (3.1), las regiones críticas correspondientes al espacio muestral y al espacio muestral lícito coinciden, es decir, si T es una estadística del tipo (2.2) sea $t_0 < 0$ si $R_T = \{(x_1, x_2) \in \mathcal{X} : T(x_1, x_2) \leq t_0\}$ y $R'_T = \{(x_1, x_2) \in \mathcal{X}' : T(x_1, x_2) \leq t_0\}$ son la región de rechazo y la región de rechazo restringida al espacio muestral lícito, respectivamente, entonces $R_T = R'_T$.

No obstante, para la estadística de razón de verosimilitudes generalizada los espacios muestral y muestral lícito no coinciden en general tanto para las pruebas asintóticas como exactas, como se ilustra en la Figura 4 para el caso de pruebas exactas.

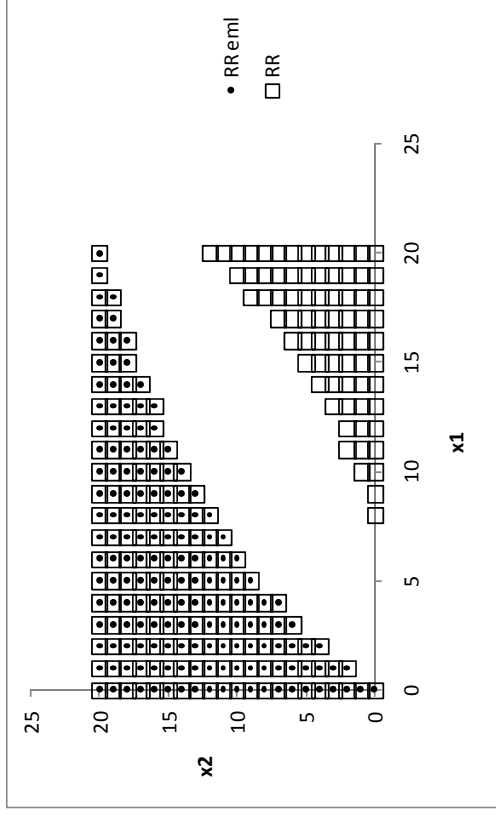


Figura 4. Región de rechazo (RR) y región de rechazo restringida al espacio muestral lícito (RR eml) para T_7 exacta, $n=20$ y $d_0=0.15$.

3.4.1. Correcciones por continuidad

Las pruebas estadísticas asintóticas que se van a comparar en este trabajo muestran un comportamiento liberal es decir generalmente exceden el nivel nominal para el que fueron construidas (Almendra, 2007) por lo que se hace necesario el uso de una corrección por continuidad que pueda ayudar a revertir ese comportamiento indeseado de las pruebas. Derivado de los resultados del estudio realizado por Almendra (2007) de pruebas asintóticas se ha determinado usar la corrección por continuidad siguiente

$$cc = \frac{1}{4n}$$

Para modificar las estadísticas de prueba de la siguiente manera

$$T_{7c}(X_1, X_2) = \frac{\hat{p}_1 - \hat{p}_1 - d_0 + cc}{\hat{\sigma}_i}; \text{ para } i=1,2,3,4,5,6$$

En cuanto a la estadística de prueba T_7 con la corrección por continuidad incluida es

$$T_{7c}(X_1, X_2) = \frac{\sup_{H_0} L(p_1, p_2; X_1, X_2)}{\sup_{H_0 \cup H_1} L(p_1, p_2; X_1, X_2)} + cc$$

CAPITULO 4. METODOLOGIA PARA EL NIVEL DE SIGNIFICANCIA

El cálculo de los niveles de significancia presenta la dificultad que para calcular la probabilidad de cometer el error tipo I se debe enfrentar la presencia de parámetros en dos dimensiones.

Recordando la función de verosimilitud conjunta de (X_1, X_2) es:

$$L(X_1 = x_1, X_2 = x_2 | n_1, n_2, p_1, p_2) = \binom{n_1}{x_1} p_1^{x_1} (1 - p_1)^{n_1 - x_1} \binom{n_2}{x_2} p_2^{x_2} (1 - p_2)^{n_2 - x_2}$$

esta función depende de los valores de (p_1, p_2) . Si se representan los pares de valores (x_1, x_2) en una cuadrícula de $(n_1 + 1) \times (n_2 + 1)$ puntos, entonces la región crítica de cualquier prueba para las hipótesis de estudio se define como un subconjunto de esta cuadrícula. Por lo tanto la probabilidad de cometer el error tipo I para $(p_1, p_2) \in \Theta_0$ es:

$$P((X_1, X_2) \in R_{T(\alpha)} | (p_1, p_2)) = \sum_{(x_1, x_2) \in R_{T(\alpha)}} L(p_1, p_2; x_1, x_2)$$

Para una región crítica dada (R_T) no está dada sino que se construye a partir de varias cosas, entre ellas el nivel nominal. El nivel de significancia se determina por la siguiente expresión:

$$\alpha^* = \max_{\{(p_1, p_2); p_1 - p_2 \geq d_0\}} \left\{ \sum_{(x_1, x_2) \in R_T(\alpha)} L(p_1, p_2; x_1, x_2) \right\} \quad (4.1)$$

4.1. Resultados que permiten simplificar el cálculo del nivel de significancia

El proceso de maximizar en todo el espacio nulo se presenta como una tarea difícil por lo que es conveniente simplificar el cálculo del nivel de significancia, reduciendo el espacio de la búsqueda, lo cual se logra con los resultados propuestos por Barnard y demostrados para pruebas de no inferioridad por Rhömel y Mansmann (1999), y por Almendra (2007). Los conceptos involucrados en la simplificación de dicho cálculo son las condiciones de convexidad de Barnard (C) y de simetría en la misma cola (SMC), los cuales se definen en la presente sección.

4.1.1. Condición de Convexidad de Barnard

Definición de Convexidad de Barnard

Sea T una estadística de prueba para las hipótesis en (2.1) con región crítica R_T . Se dice que R_T satisface la condición de convexidad de Barnard (C) si se cumplen las dos condiciones siguientes

- a) $(x, y) \in R_T \Rightarrow (x-1, y) \in R_T$ para toda $1 \leq x \leq n_1$, $0 \leq y \leq n_2$
 b) $(x, y) \in R_T \Rightarrow (x, y+1) \in R_T$ para toda $0 \leq x \leq n_1$, $0 \leq y \leq n_2 - 1$

Röhmel y Mansmann (1999) demostraron el siguiente resultado para pruebas exactas

Teorema. Sea T una estadística de prueba para el problema de prueba de hipótesis

$$H_0 : \theta_2 \leq g(\theta_1) \text{ versus } H_a : \theta_2 > g(\theta_1)$$

con región de rechazo dada por $R = \{(x, y) : T(x, y) \leq T(x_0, y_0)\}$ y sean

$$\Psi_0 = \{\theta = (\theta_1, \theta_2) : \theta_2 \leq g(\theta_1)\}; \quad a \leq \theta_1 \leq 1, \quad 0 \leq a \text{ y } g(a) = 0; \text{ si } a > 0 \text{ con } g \text{ creciente}$$

y no interseca a $\theta_2 = \theta_1$. Si la región R_T satisface la condición de convexidad de

Barnard, entonces el supremo $\sup_{(p_1, p_2) \in \Psi_0} \beta_T(p_1, p_2)$ es un máximo y se alcanza en un

punto de la frontera sobre la curva $\theta_2 = g(\theta_1)$. Entonces

$$\Psi^* = \{\theta^* = (\theta_1, \theta_2) : \theta_2 = g(\theta_1)\} \text{ y el nivel de significancia es}$$

$$\sup_{\theta \in \Psi_0} \beta_T(p_1, p_2) = \max_{\theta \in \Psi^*} \beta_T(p_1, p_2)$$

Almendra (2007) generalizó el teorema de Röhmel y Mansmann para pruebas tanto exactas como asintóticas y además volviendo menos restrictivo el principio de convexidad. Con esto, un mayor número de pruebas estadísticas cumplen con las nuevas condiciones definidas en dicho trabajo, llamadas condiciones de semiconvexidad, logrando con ello simplificar los cálculos para una gama más amplia de regiones críticas.

Mostrar que una región de rechazo cumple la condición de convexidad de Barnard no es un asunto simple, de hecho solamente se ha demostrado para la prueba T_2 exacta (Röhmel, 2005) y para T_2 asintótica (Almendra 2007), por lo cual para las demás pruebas la verificación se hará numéricamente mediante un programa de cómputo escrito en S-PLUS [®].

4.1.2. Simetría en la misma cola

Definición de Simetría en la Misma Cola

Se dice que una región crítica R satisface *la propiedad de simetría en la misma cola (SMC)* si $(x, y) \in R \Rightarrow (n - y, n - x) \in R$.

Almendra (2008) demostró las siguientes proposiciones.

Proposición. Si $n = n_1 = n_2$, entonces R_{IC} cumple la propiedad de simetría en la misma cola para $i = 1, 3, 4, 6$

Proposición. Sean $n = n_1 = n_2$ y $R(\alpha)$ una región crítica para el problema de prueba de hipótesis.

$$[H_0 : d \geq d_0] \text{ versus } [H_a : d < d_0]$$

para el nivel de significancia nominal α , si $R(\alpha)$ cumple la Condición (C) y la condición (SMC), entonces el nivel de significancia está dado por

$$\alpha^* = \max_{\substack{p_2 = p_1 - d_0 \\ p_1 \in [d_0, \frac{1+d_0}{2}]} \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} \binom{n_1}{x_1} p_1^{x_1} (1-p_1)^{n_1-x_1} \binom{n_2}{x_2} p_2^{x_2} (1-p_2)^{n_2-x_2} I_{[(x_1, x_2) \in R_T(\alpha)]} \quad (4.2)$$

4.2. Método de cálculo del nivel de significancia

Se usó la fórmula en (4.2) para calcular una aproximación del nivel de significancia exacto de todas las pruebas consideradas. La aproximación del nivel de significancia exacto α^* , en la fórmula (4.2), se hizo reemplazando el intervalo $[d_0, (1+d_0)/2]$ por el conjunto de puntos: $\{d_0 + (.001)^i : i = 0, 1, 2, \dots, 500(1 - d_0)\}$, y al valor obtenido de esta aproximación de α^* usando tal conjunto discreto se le ha llamado *nivel de significancia real* de la prueba, el cual será denotado por α_R .

Es importante subrayar que la fórmula (4.2), para el cálculo del nivel de significancia (α^*) de cualquiera de las pruebas consideradas aquí, es una fórmula exacta. Así que la única aproximación que se hace para calcular el nivel de significancia real (α^*_R) es al reemplazar al intervalo $[d_0, (1+d_0)/2]$ por el conjunto de puntos $\{d_0 + (.001)^i : i = 0, 1, 2, \dots, 500(1 - d_0)\}$. Almendra y Sotres (2008) verificaron que para las pruebas asintóticas el grado de error en que incurre esta aproximación se estimó de la siguiente manera. Para todas las pruebas consideradas y para cada una de las doce combinaciones de parámetros ($n_1=n_2=30, 50$ y 100), ($\alpha=0.05$ y 0.10), ($d_0 = 0.10$ y 0.15) se analizó el error incurrido al estimar α^* por α_R y en todos los casos se verificó que en los valores adyacentes al máximo estimado (α_R) la variación en la función de potencia fue menor o igual a 0.00001 , es decir, nuestra estimación del error $|\alpha^* - \alpha_R|$ es que es menor o igual 0.00001 . Por ejemplo para $n_1=n_2=50$, $\alpha=0.05$ y $d_0 = 0.10$ los valores adyacentes al máximo estimado (α_R) de la función de potencia son los siguientes:

$$\beta_{\Gamma}(0.549,0.449)=0.0441077$$

$$\beta_{\Gamma}(0.550,0.450)=0.0441080 = \alpha_R \text{ (m\u00e1ximo aproximado)}$$

$$\beta_{\Gamma}(0.551,0.451)=0.0441077,$$

en este caso $\beta_{\Gamma}(0.550, .0450)-\beta_{\Gamma}(0.549, .449)=0.0000003$, el cual es menor a 0.00001. Con base en este an\u00e1lisis es razonable concluir que el error incurrido al estimar α^* por α_R es aceptable desde el punto de vista pr\u00e1ctico, ya que el nivel de significancia real (α_R) difiere del nivel de sinificancia exacto (α^*) cuando mucho en la quinta cifra decimal.

De tal modo que para comparar el nivel de significancia real de las 7 pruebas cada prueba se consider\u00f3 la configuraci\u00f3n siguiente:

- Nivel de significancia nominal $\alpha = 0.01, 0.05$ y 0.10
- M\u00e1rgenes de no inferioridad $d_0 \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$
- Tama\u00f1o de muestra. Se han elegido tama\u00f1os de muestra balanceados, para

$$n_1 = n_2 \in \{5, 6, 7, 8, 9, \dots, 100\}.$$

Adem\u00e1s, el intervalo de $p_1 \in [d_0, (1+d_0)/2]$ se reemplaz\u00f3 por el conjunto de los puntos

$$\{d_0 + (.001)^i : i = 0, 1, 2, \dots, 500(1 - d_0)\} \text{ y } p_2 = p_1 - d_0.$$

CAPITULO 5. RESULTADOS Y DISCUSION

5.1. Nivel de significancia de las pruebas exactas

Para comparar el nivel de significancia real de las 7 pruebas exactas se calculó el nivel de significancia real para un amplio escenario de tamaños de muestra desde 5 hasta 100 con incrementos de 1 y para cinco valores del margen de no inferioridad d_0 (ver sección anterior). Los resultados se presentan en la Tabla 2.

Las pruebas exactas siempre mantienen el nivel nominal, es decir $\alpha^* \leq \alpha$. Sin embargo, tienen el inconveniente de que pueden ser demasiado conservadoras, lo deseable es que la distancia del nivel de significancia real al nivel nominal, $\alpha - \alpha^*$, sea lo más pequeña posible. Por esta razón es que se calculó el porcentaje de niveles de significancia reales que pertenecen al intervalo de $[0.80\alpha, \alpha]$, es decir que estén bastante cercanos al nivel nominal α elegido para la prueba. En la Tabla 2 se presenta el comportamiento de los niveles de significancia reales de las siete pruebas consideradas.

Tabla 2. Porcentaje (*) de los Niveles de significancia que pertenecen al intervalo $[0.80\alpha, \alpha]$.

alfa (α)	Intervalo $[0.8\alpha, \alpha]$	d_0	Porcentaje de NS que pertenecen al intervalo						
			T_1	T_2	T_3	T_4	T_5	T_6	T_7
0.01	[0.008, 0.01]	0.050	6.25	94.79	67.71	7.29	94.79	65.63	46.32
		0.100	15.63	96.88	48.96	15.63	95.83	48.96	49.47
		0.150	25.00	97.92	47.92	25.00	97.92	47.92	59.15
		0.200	29.17	95.83	43.75	29.17	95.83	43.75	66.32
		0.250	34.38	93.75	50.00	34.38	95.83	50.00	72.63
0.05	[0.04, 0.05]	0.050	15.63	96.88	78.13	17.71	96.88	76.04	46.32
		0.100	28.13	98.96	67.71	28.13	98.96	67.71	63.16
		0.150	39.58	95.83	66.67	39.58	94.79	66.67	47.37
		0.200	46.88	93.75	64.58	46.88	94.79	64.58	77.89
		0.250	54.17	95.83	71.88	54.17	95.83	71.88	83.16
0.1	[0.08, 0.1]	0.050	23.96	88.54	79.17	23.96	88.54	79.17	52.63
		0.100	41.67	90.63	73.96	41.67	91.67	73.96	65.26
		0.150	48.96	96.88	79.17	48.96	96.88	79.17	70.42
		0.200	61.46	94.79	76.04	60.42	94.79	76.04	80.00
		0.250	64.58	97.92	84.38	64.58	97.92	84.38	87.37

(*) Nota: Para cada prueba T_i , el porcentaje reportado se calculó dividiendo el número de niveles de significancia reales de T_i que pertenecen al intervalo $[0.80\alpha, \alpha]$ entre los 96 tamaños de muestra considerados, pues $\#\{5,6,\dots,100\} = 96$.

5.1.1. Prueba T_1

Se puede observar en la Tabla 2 que la prueba T_1 para el nivel nominal de $\alpha=0.01$ tiene un comportamiento demasiado conservador, acentuándose en los niveles más bajos del margen de no inferioridad d_0 ya que para $d_0=0.05$ solamente el 6.25% de las muestras sus niveles de significancia pertenecen al intervalo de [0.008, 0.01]. En la Figura 5 se ilustra el comportamiento del nivel de significancia real para la prueba T_1 para $\alpha=0.01$, y para tres márgenes de no inferioridad. El comportamiento del nivel de significancia real de T_1 para $\alpha = 0.05$ y 0.10 es también muy conservador y similar al caso $\alpha = 0.01$.

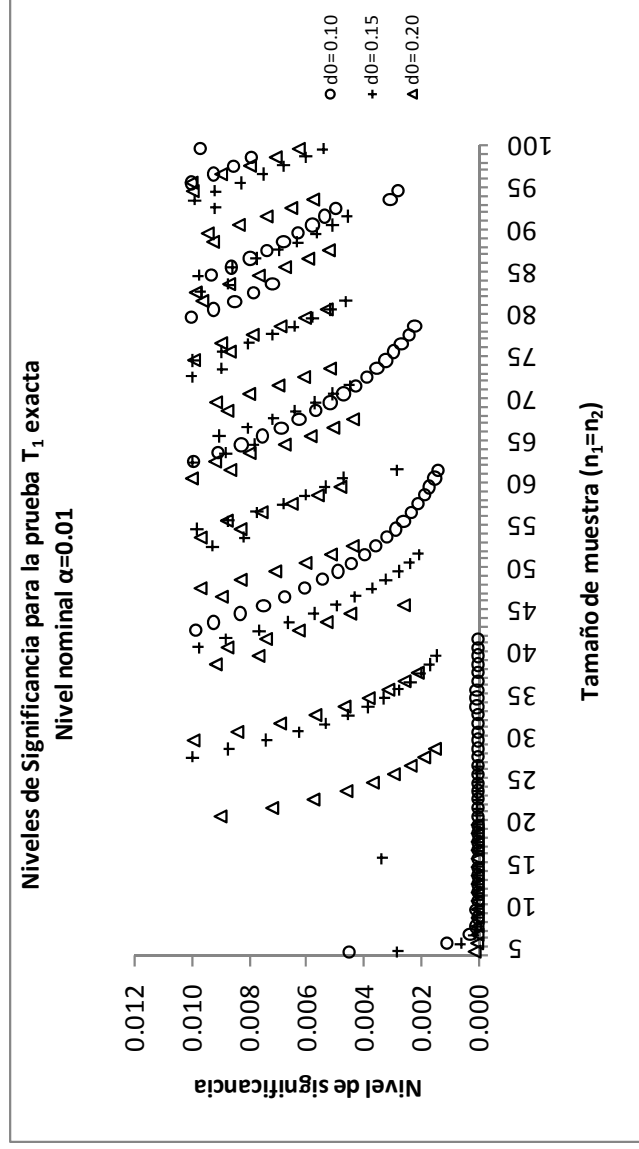


Figura 5. Nivel de Significancia para la prueba T_1 según el tamaño de muestra, $\alpha=0.01$

5.1.2. Prueba T_2

En la Tabla 2 y en la Figura 6 se puede observar que el nivel de significancia de la prueba T_2 para el nivel nominal de $\alpha=0.01$ por lo menos el 93.75 % de las muestras estudiadas sus niveles de significancia pertenecen al intervalo $[0.008, 0.01]$ lo que nos habla de que es una prueba que mantiene muy bien el nivel sin ser demasiado conservadora.

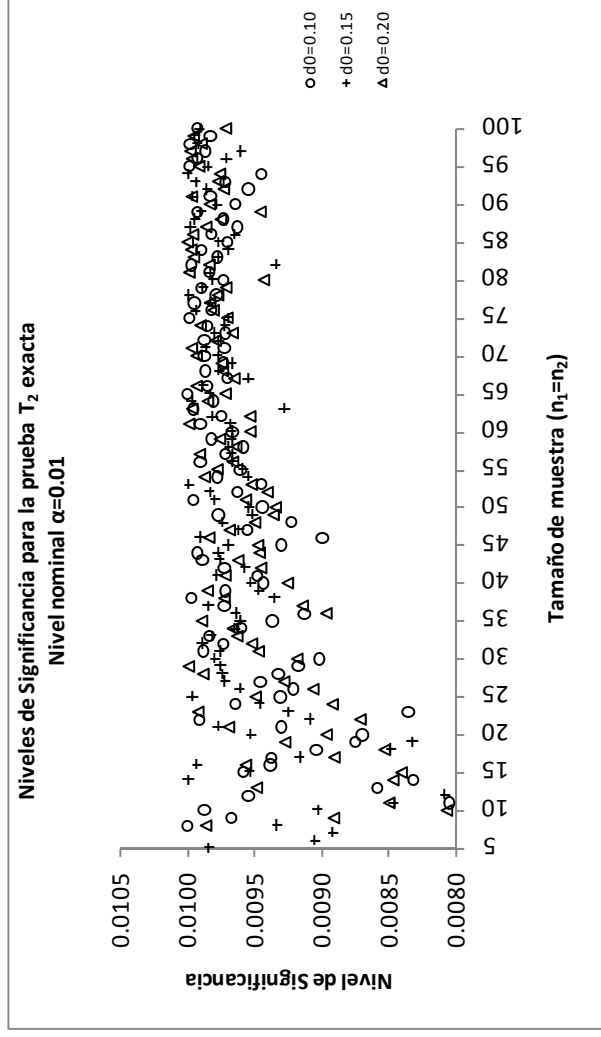


Figura 6. Nivel de Significancia para la prueba T_2 según el tamaño de muestra, $\alpha=0.01$

Para un nivel nominal de $\alpha=0.05$ la prueba T_2 mantiene el nivel de significancia en el intervalo impuesto también para al menos el 93.75 % de las muestras, y para el nivel nominal de $\alpha=0.10$ logra colocar al menos el 88.54 % de las muestras en el intervalo especificado para el nivel nominal. No se pudo observar una mejoría en el nivel de significancia a través de los tres valores de nivel nominal estudiados.

5.1.3. Prueba T_3

Para la prueba T_3 y el nivel de significancia nominal $\alpha=0.01$ aproximadamente para la mitad de las muestras el valor del nivel de significancia cae en el intervalo especificado. Con lo cual muestra un desempeño bueno considerando que la forma de la estadística es muy sencilla.

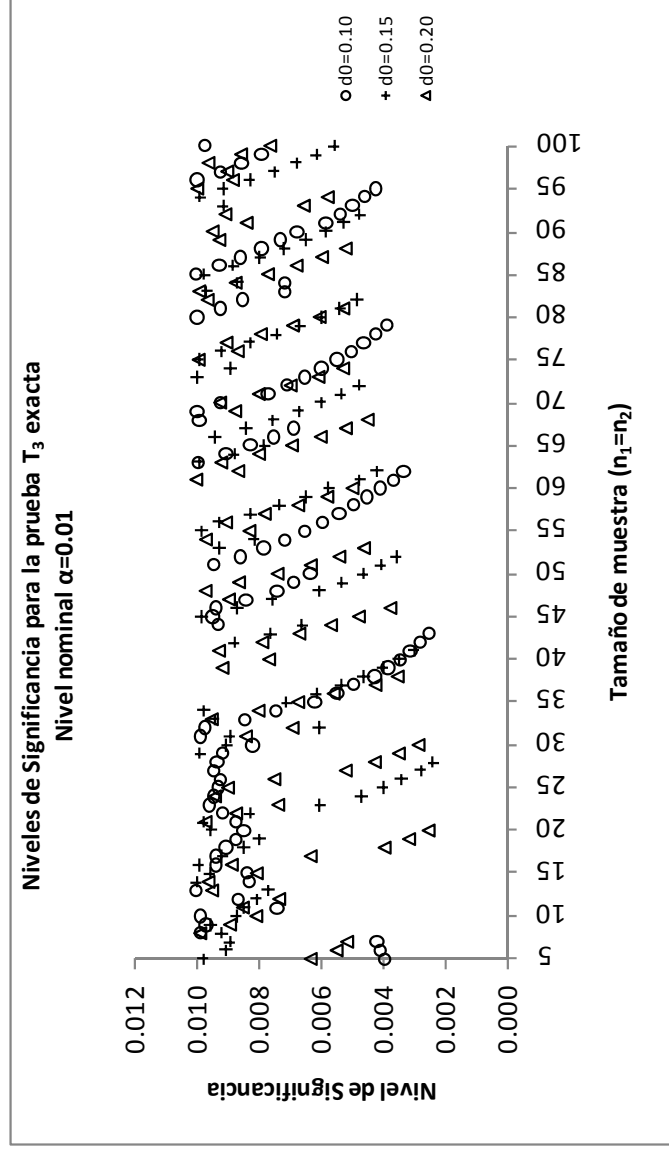


Figura 7. Nivel de Significancia para la prueba T_3 según el tamaño de muestra, $\alpha=0.01$

La prueba T_3 para el nivel nominal de $\alpha=0.05$ muestra un mejor comportamiento de los niveles de significancia exactos con respecto al nivel nominal $\alpha=0.01$, porque mientras que allá lo mas que se alcanza es que el 67.71% de las muestras entren en el intervalo especificado, en el nivel de $\alpha=0.05$ se puede lograr que el 78.13% de las muestras se coloquen en el intervalo [0.04, 0.05]. Lo que es más para el nivel nominal de $\alpha=0.10$ la prueba T_3 muestra su mejor comportamiento de los tres niveles estudiados, por lo que se

puede decir que esta prueba se desempeña mejor para niveles grandes de $\alpha=0.10$ y se puede detectar una mejoría conforme el nivel nominal aumenta. Este comportamiento se muestra en la Figura 7.

5.1.4. Prueba T_4

El comportamiento del nivel de significancia exacto de la prueba T_4 (que también puede verse como una modificación de T_1) para el nivel nominal de $\alpha=0.01$ muestra un comportamiento bastante conservador porque en su mejor valor solo el 34.38% de las muestras pertenecen al intervalo elegido, [0.008, 0.01]. Los valores de los niveles de significancia son muy parecidos a los obtenidos en la prueba T_1 . El desempeño del nivel de significancia se mejora conforme crece el margen de no inferioridad d_0 . En la Figura 8 se muestran los niveles de significancia exactos.

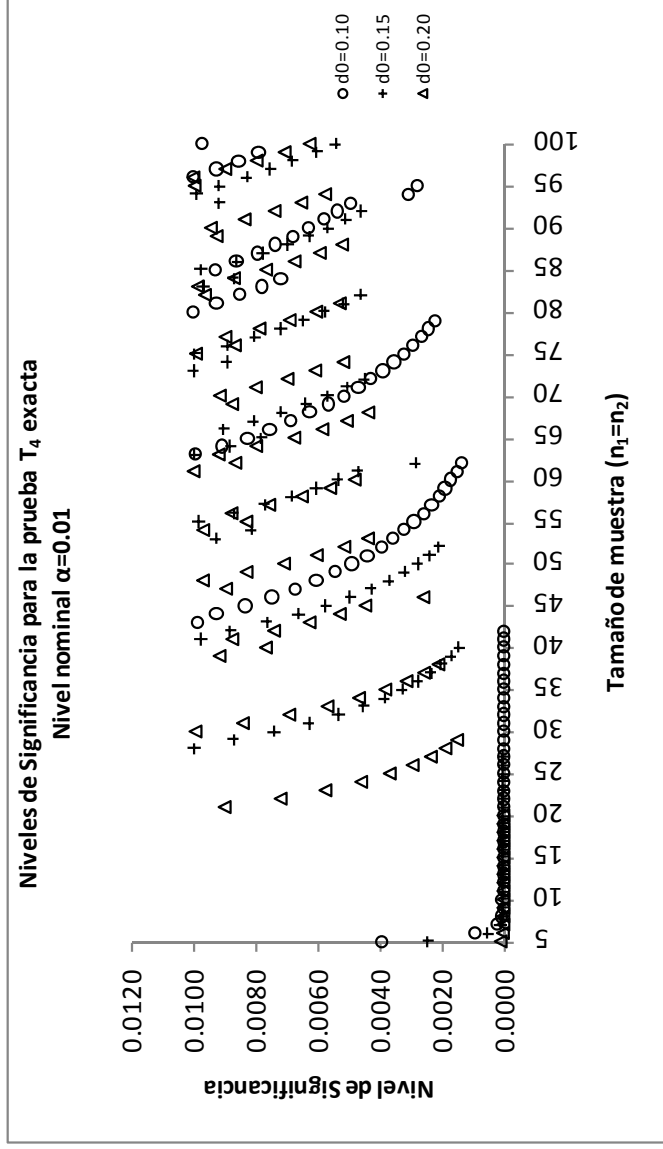


Figura 8. Nivel de Significancia para la prueba T_4 según el tamaño de muestra, $\alpha=0.01$

Para el nivel nominal de $\alpha=0.05$ la prueba T_4 es muy conservadora porque los porcentajes de los tamaños de muestra para las cuales el nivel de significancia real pertenece al intervalo $[0.04, 0.05]$ van desde solamente el 17.71% hasta el 54.17% en su mejor valor. La prueba T_4 muestra mejoría conforme crece el margen de no inferioridad d_0 , de lo que se puede decir que el nivel de significancia exacto depende del valor d_0 .

La prueba T_4 para el nivel nominal de $\alpha=0.10$ muestra el mejor comportamiento de todos los niveles nominales estudiados aunque en su mejor valor solamente el 64.58% de las muestras estudiadas colocan el valor del nivel de significancia exacto en el intervalo $[0.08, 0.10]$ y eso es para el valor de $d_0=0.25$ porque para $d_0=0.05$ solamente el 23.96 % de las muestras no son conservadoras. Esta prueba no tiene un comportamiento deseable porque es muy conservadora en el nivel de significancia.

La prueba T_4 surgió de una modificación de la prueba T_1 , pero ambas pruebas muestran un comportamiento casi idéntico por lo que no se puede decir que dicha modificación mejora en algún modo a la prueba T_1 y ambas tienen un comportamiento demasiado conservador.

5.1.5. Prueba T_5

La prueba T_5 muestra un comportamiento muy semejante a la prueba T_2 , y recordando que surge como una modificación a la prueba T_2 entonces la modificación no surtió mucho efecto, como se puede ver en las Figuras a continuación de los niveles de significancia.

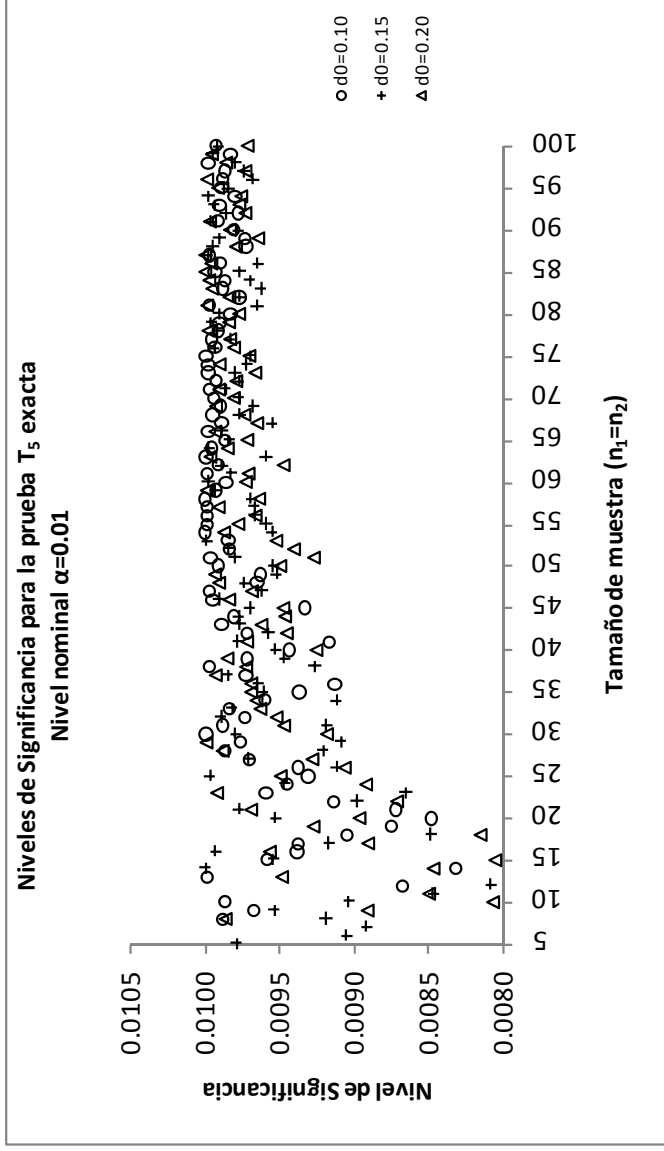


Figura 9. Nivel de Significancia para la prueba T_5 según el tamaño de muestra, $\alpha=0.01$

El comportamiento de los niveles de significancia exactos de la prueba T_5 es muy bueno, basta con observar en la Figura 9 como están concentrados los puntos del nivel de significancia exacto en torno al nivel de significancia nominal y aunque para muestras pequeñas los valores de los son más bajos no dejan de caer en el intervalo [0.008, 0.01]

La prueba T_5 no muestra una dependencia con el tamaño de muestra y volviendo a observar la Tabla 2 no se puede observar un patrón conforme crece el margen de no inferioridad

5.1.6. Prueba T_6

La prueba T_6 resulta de una modificación de la prueba T_3 y en cuanto al nivel de significancia ambas pruebas tienen un comportamiento casi idéntico como se puede observar en la Tabla 2.

La prueba T_6 no tiene una estadística muy compleja de calcular pero es muy conservadora para algunos tamaños de muestra y exhibe un comportamiento muy semejante para cualquier nivel del margen de no inferioridad d_0 , en la Figura 10 se ilustra para tres niveles de d_0 .

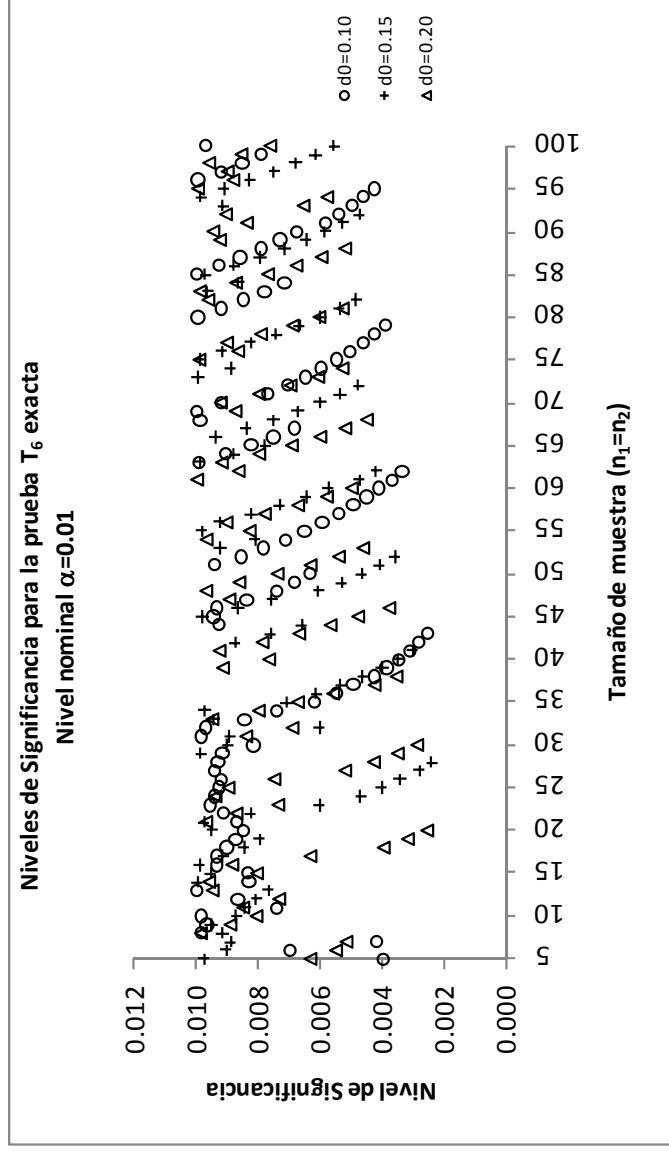


Figura 10. Nivel de Significancia para la prueba T_6 según el tamaño de muestra, $\alpha=0.01$

5.1.7. Prueba T_7

El comportamiento de la prueba T_7 o de razón de verosimilitudes generalizada no es muy bueno y se ve afectado por los valores más pequeños del nivel nominal, $\alpha=0.01$, y del margen de no inferioridad, $d_0=0.05$.

Para el nivel nominal de $\alpha=0.01$ la prueba T_7 logra que el 72% de los niveles de significancia de las muestras analizadas pertenezcan al intervalo de $[0.008, 0.01]$ para un margen de $d_0=0.25$ mientras que para el margen de $d_0=0.05$ lo logra para el 46 % de las

muestras. Lo anterior es observando los extremos estudiados de los márgenes de no inferioridad en la Figura 11 se pueden observar en comportamiento de $d_0=0.10, 0.15, 0.20$ y se les ve muy concentrados cercanas al nivel nominal, lo que indica que tiene un buen comportamiento.

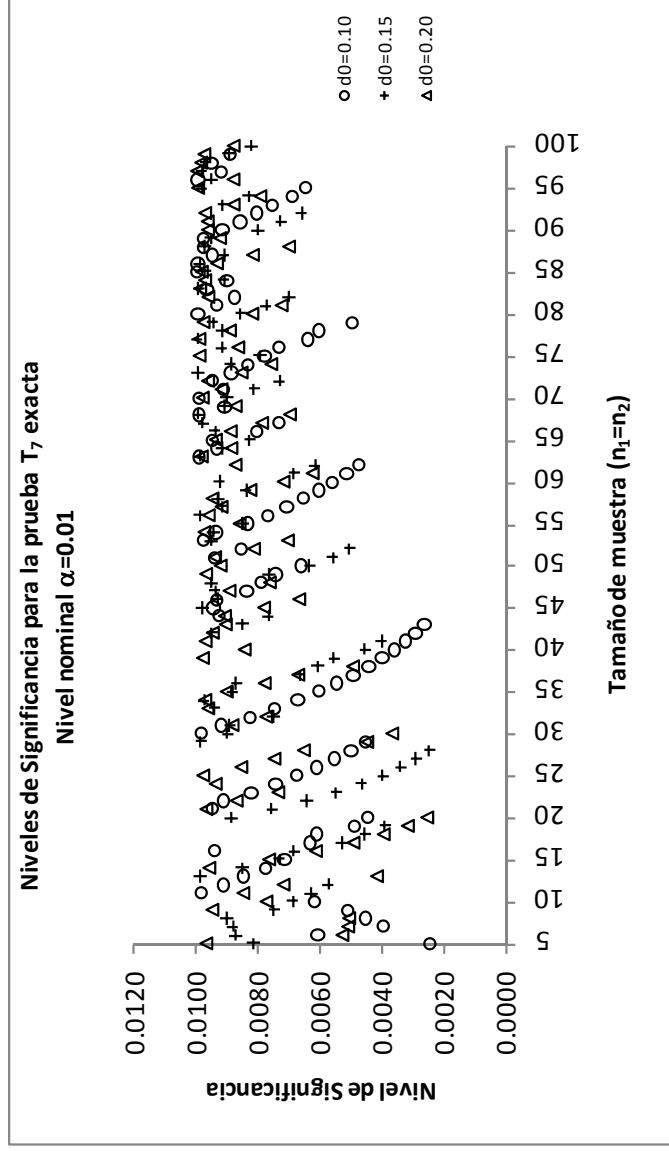


Figura 11. Nivel de Significancia para la prueba T_7 según el tamaño de muestra, $\alpha=0.01$

Para el nivel nominal de $\alpha=0.05$ la prueba T_7 logra colocar el nivel de significancia dentro del intervalo $[0.04, .05]$ para el 83% de sus muestras con $d_0=0.25$, sin embargo, para los demás márgenes el comportamiento de la prueba es pobre. En la Tabla 2 también se puede observar que el desempeño de la prueba mejora conforme aumenta el valor de margen de no inferioridad.

Para el nivel nominal de $\alpha=0.10$ los niveles de significancia se mejoran porque logra colocar el nivel de significancia dentro del intervalo $[0.08, 1.00]$ para el 80% o más de sus muestras en dos márgenes de no inferioridad $d_0=0.20$ y 0.25 .

5.2. Potencia de las pruebas exactas

Las pruebas que son candidatas para que se les calcule la potencia deben cumplir primeramente que su nivel de significancia real esté muy cercano al nivel nominal, es decir que para un determinado margen de no inferioridad d_0 , aproximadamente el 90% o más de los niveles de significancia de las muestras estudiadas pertenezcan al intervalo de $[0.8\alpha, \alpha]$. De esta restricción y para el nivel nominal de $\alpha=0.05$ quedan como finalistas las pruebas T_2 y T_5 según la Tabla 3.

Tabla 3. Pruebas para las cuales las muestras colocan más del 90% de sus niveles de significancia reales (α^*) en un intervalo de $[0.8\alpha, \alpha]$

alfa (α)	Intervalo $[0.8\alpha, \alpha]$	Porcentaje de NS que pertenecen al Intervalo	
		T_2	T_5
0.05	[0.04, 0.05]	98.96	98.96
		95.83	94.79
		93.75	94.79
0.1	[0.08, 0.1]	90.63	91.67
		96.88	96.88
		94.79	94.79
	0.25	97.92	97.92

Posteriormente se revisó en cada margen de no inferioridad aquellos tamaños de muestra donde la diferencia máxima entre los niveles de significancia de las dos pruebas T_2 y T_5 resultó menor o igual que 0.0001. En las Tablas 4, 5 y 6 se muestran los tamaños de muestra que presentan esta característica.

En la Tabla 4 se presentan los niveles de significancia reales de las pruebas T_2 y T_5 para los tamaños de muestra en los cuales la diferencia fue menor o igual que 0.0001 cuando el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad es $d_0=0.10$. Del total de muestras analizadas de $5 \leq n \leq 100$ en 82 (85%) de ellas hubo una diferencia menor o igual de 0.0001.

En la Tabla 5 se presentan los niveles de significancia reales de las pruebas T_2 y T_5 para los tamaños de muestra en los cuales la diferencia fue menor o igual que 0.0001 cuando el margen de no inferioridad es $d_0=0.15$. Del total de muestras analizadas de $5 \leq n \leq 100$ en 73 (76%) de ellas hubo una diferencia menor o igual de 0.0001

Tabla 4. Niveles de significancia para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0 = 0.10$

n	T2	T5	Diferenc	n	T2	T5	Diferenc
5	0.04807	0.04807	0.0000	50	0.04994	0.04994	0.0000
6	0.04646	0.04646	0.0000	51	0.04901	0.04901	0.0000
7	0.04475	0.04475	0.0000	52	0.04974	0.04974	0.0000
9	0.04263	0.04263	0.0000	53	0.04941	0.04941	0.0000
10	0.04121	0.04121	0.0000	54	0.04945	0.04945	0.0000
11	0.03993	0.03993	0.0000	55	0.04852	0.04852	0.0000
12	0.04792	0.04792	0.0000	56	0.04911	0.04911	0.0000
13	0.04661	0.04661	0.0000	57	0.04960	0.04960	0.0000
14	0.04525	0.04525	0.0000	58	0.04835	0.04835	0.0000
15	0.04933	0.04933	0.0000	59	0.04905	0.04905	0.0000
16	0.04656	0.04656	0.0000	60	0.04407	0.04407	0.0000
17	0.04764	0.04764	0.0000	61	0.04993	0.04993	0.0000
18	0.04856	0.04856	0.0000	62	0.04878	0.04878	0.0000
19	0.04670	0.04670	0.0000	63	0.04919	0.04919	0.0000
20	0.04485	0.04485	0.0000	64	0.04939	0.04939	0.0000
21	0.04965	0.04965	0.0000	65	0.04963	0.04963	0.0000
22	0.04675	0.04675	0.0000	66	0.04973	0.04973	0.0000
23	0.04803	0.04803	0.0000	67	0.04983	0.04983	0.0000
24	0.04603	0.04603	0.0000	68	0.04984	0.04984	0.0000
25	0.04649	0.04649	0.0000	69	0.04981	0.04981	0.0000
26	0.04846	0.04846	0.0000	70	0.04977	0.04972	0.0001
29	0.04938	0.04938	0.0000	75	0.04897	0.04893	0.0000
30	0.04755	0.04755	0.0000	77	0.04961	0.04961	0.0000
31	0.04875	0.04875	0.0000	78	0.04928	0.04928	0.0000
32	0.04986	0.04986	0.0000	79	0.04974	0.04974	0.0000
33	0.04802	0.04802	0.0000	80	0.04949	0.04949	0.0000
34	0.04734	0.04734	0.0000	81	0.04999	0.04999	0.0000
35	0.04847	0.04847	0.0000	83	0.04889	0.04889	0.0000
36	0.04819	0.04819	0.0000	84	0.04978	0.04978	0.0000
37	0.04730	0.04730	0.0000	85	0.04952	0.04952	0.0000
38	0.04852	0.04852	0.0000	86	0.04955	0.04955	0.0000
39	0.04825	0.04825	0.0000	87	0.04843	0.04843	0.0000
40	0.04950	0.04950	0.0000	88	0.04612	0.04612	0.0000
41	0.04930	0.04930	0.0000	89	0.04977	0.04977	0.0000
42	0.04908	0.04908	0.0000	90	0.04972	0.04972	0.0000
43	0.04867	0.04867	0.0000	91	0.04969	0.04969	0.0000
44	0.04709	0.04709	0.0000	92	0.04962	0.04962	0.0000
46	0.04839	0.04839	0.0000	93	0.04948	0.04948	0.0000
47	0.04882	0.04882	0.0000	94	0.04934	0.04934	0.0000
48	0.04850	0.04850	0.0000	95	0.04988	0.04988	0.0000
49	0.04896	0.04896	0.0000	96	0.04964	0.04964	0.0000

Tabla 5. Niveles de significancia para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.15$

n	T_2	T_5	Diferencia
5	0.03334	0.03334	0.0000
6	0.03076	0.03076	0.0000
8	0.04716	0.04716	0.0000
9	0.04312	0.04312	0.0000
10	0.03935	0.03935	0.0000
11	0.04385	0.04385	0.0000
12	0.04880	0.04880	0.0000
13	0.04449	0.04449	0.0000
14	0.04055	0.04055	0.0000
15	0.04846	0.04846	0.0000
16	0.04776	0.04776	0.0000
17	0.04331	0.04331	0.0000
20	0.04966	0.04966	0.0000
21	0.04601	0.04601	0.0000
22	0.04882	0.04882	0.0000
23	0.04804	0.04804	0.0000
24	0.04870	0.04870	0.0000
25	0.04560	0.04560	0.0000
26	0.04862	0.04862	0.0000
27	0.04913	0.04913	0.0000
28	0.04634	0.04634	0.0000
29	0.04492	0.04492	0.0000
30	0.04824	0.04824	0.0000
31	0.04885	0.04885	0.0000
32	0.04894	0.04894	0.0000

n	T_2	T_5	Diferencia
33	0.04910	0.04910	0.0000
40	0.04915	0.04915	0.0000
41	0.04957	0.04957	0.0000
42	0.04749	0.04749	0.0000
43	0.04884	0.04884	0.0000
44	0.04809	0.04809	0.0000
45	0.04846	0.04846	0.0000
46	0.04809	0.04809	0.0000
47	0.04980	0.04980	0.0000
48	0.04927	0.04927	0.0000
49	0.04954	0.04954	0.0000
51	0.04942	0.04943	0.0000
54	0.04913	0.04913	0.0000
55	0.04949	0.04949	0.0000
56	0.04911	0.04911	0.0000
57	0.04910	0.04910	0.0000
58	0.04991	0.04991	0.0000
59	0.04843	0.04843	0.0000
60	0.04958	0.04958	0.0000
61	0.04902	0.04902	0.0000
62	0.04950	0.04950	0.0000
63	0.04973	0.04973	0.0000
67	0.04908	0.04908	0.0000
68	0.04877	0.04877	0.0000

n	T_2	T_5	Diferencia
69	0.04885	0.04885	0.0000
70	0.04974	0.04974	0.0000
71	0.04903	0.04903	0.0000
72	0.04976	0.04976	0.0000
73	0.04908	0.04908	0.0000
74	0.04933	0.04933	0.0000
75	0.04939	0.04939	0.0000
79	0.04915	0.04915	0.0000
80	0.04881	0.04881	0.0000
81	0.04973	0.04973	0.0000
82	0.04918	0.04918	0.0000
83	0.04970	0.04970	0.0000
84	0.04999	0.04999	0.0000
85	0.04911	0.04911	0.0000
86	0.04994	0.04994	0.0000
87	0.04966	0.04966	0.0000
90	0.04913	0.04910	0.0000
91	0.04985	0.04985	0.0000
92	0.04686	0.04686	0.0000
93	0.04974	0.04974	0.0000
94	0.04994	0.04994	0.0000
95	0.04914	0.04914	0.0000
96	0.04919	0.04919	0.0000
97	0.04980	0.04980	0.0000

En la Tabla 6 se presenta los niveles de significancia de las pruebas T_2 y T_5 para los tamaños de muestra en los cuales la diferencia fue menor o igual que 0.0001 cuando el margen de no inferioridad es $d_0=0.20$. Del total de muestras analizadas de $5 \leq n \leq 100$ en 78 (81%) de ellas hubo una diferencia menor o igual de 0.0001.

Tabla 6. Niveles de significancia para los tamaños de muestra $5 \leq n \leq 100$ para las pruebas T_2 , y T_5 para el nivel de significancia nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.20$

n	T_2	T_5	Diferencia	n	T_2	T_5	Diferencia	n	T_2	T_5	Diferencia
5	0.02256	0.02256	0.0000	40	0.04828	0.04828	0.0000	71	0.04922	0.04922	0.0000
6	0.03922	0.03922	0.0000	41	0.04814	0.04814	0.0000	72	0.04974	0.04974	0.0000
7	0.03334	0.03334	0.0000	42	0.04947	0.04947	0.0000	73	0.04994	0.04994	0.0000
8	0.04528	0.04528	0.0000	43	0.04905	0.04905	0.0000	74	0.04983	0.04983	0.0000
9	0.03943	0.03943	0.0000	44	0.04909	0.04909	0.0000	75	0.04944	0.04944	0.0000
10	0.04511	0.04511	0.0000	45	0.04913	0.04913	0.0000	76	0.04973	0.04973	0.0000
11	0.03939	0.03939	0.0000	47	0.04998	0.04998	0.0000	79	0.04887	0.04887	0.0000
12	0.04454	0.04454	0.0000	48	0.04883	0.04883	0.0000	80	0.04905	0.04905	0.0000
16	0.04856	0.04856	0.0000	49	0.04518	0.04518	0.0000	81	0.04898	0.04898	0.0000
18	0.04613	0.04613	0.0000	50	0.04956	0.04956	0.0000	82	0.04981	0.04981	0.0000
19	0.04570	0.04570	0.0000	51	0.04844	0.04844	0.0000	83	0.04998	0.04998	0.0000
20	0.04482	0.04482	0.0000	52	0.04883	0.04883	0.0000	84	0.04968	0.04968	0.0000
21	0.04647	0.04647	0.0000	53	0.04998	0.04998	0.0000	86	0.04928	0.04928	0.0000
22	0.04809	0.04809	0.0000	56	0.04927	0.04927	0.0000	87	0.04947	0.04947	0.0000
23	0.04927	0.04927	0.0000	57	0.04993	0.04993	0.0000	88	0.04923	0.04923	0.0000
24	0.04850	0.04850	0.0000	58	0.04882	0.04882	0.0000	89	0.04982	0.04982	0.0000
25	0.04794	0.04794	0.0000	59	0.04921	0.04921	0.0000	90	0.04994	0.04994	0.0000
26	0.04749	0.04749	0.0000	60	0.04919	0.04919	0.0000	91	0.04962	0.04962	0.0000
30	0.04975	0.04975	0.0000	61	0.04986	0.04986	0.0000	93	0.04955	0.04955	0.0000
31	0.04943	0.04943	0.0000	63	0.04997	0.04997	0.0000	94	0.04942	0.04942	0.0000
32	0.04950	0.04950	0.0000	64	0.04976	0.04976	0.0000	95	0.04912	0.04912	0.0000
33	0.04935	0.04935	0.0000	65	0.04868	0.04868	0.0000	96	0.04965	0.04965	0.0000
34	0.04814	0.04814	0.0000	66	0.04904	0.04904	0.0000	97	0.04969	0.04969	0.0000
35	0.04959	0.04959	0.0000	67	0.04906	0.04906	0.0000	98	0.04995	0.04995	0.0000
36	0.04605	0.04605	0.0000	68	0.04876	0.04876	0.0000	99	0.04938	0.04942	0.0000
39	0.04934	0.04934	0.0000	69	0.04911	0.04911	0.0000	100	0.04958	0.04958	0.0000

La comparación de las potencias se realizó en los puntos $(p_1, p_2) \in \Theta_{AD}$, donde:

$$\Theta_{AD} = \left\{ (p_1, p_2) : p_1 = 0.05i, p_2 = \max(p_1 - d_0, 0) + 0.05_j \right. \\ \left. i = 0, 1, \dots, 20; j \text{ tal que } p_2 \in [\max(p_1 - d_0, 0), 1] \right\} \\ = \left\{ (p_1, p_2) : p_1 = 0.05i, p_2 = \max(p_1 - d_0, 0) + 0.05_j \right. \\ \left. i = 0, 1, \dots, 20; j = 0, 1, \dots, 20(1 - \max(p_1 - d_0, 0)) \right\}$$

Como resultado de la comparación de potencias para el margen de no inferioridad de $d_0=0.10$ tenemos que para todos los tamaños de muestra de la Tabla 4 las pruebas T_2 y T_5 resultaron iguales en la potencia en el 100% de los casos.

En el caso del margen de no inferioridad de $d_0=0.15$ para todos los tamaños de muestra de la Tabla 5 las pruebas T_2 y T_5 también resultaron iguales en la potencia en el 100% de los casos.

Para el margen de no inferioridad de $d_0=0.20$ para todos los tamaños de muestra de la Tabla 6 se analizaron las potencias de las pruebas T_2 y T_5 y resultaron iguales para los 78 tamaños de muestra.

Con fines ilustrativos se presenta gráficamente en la Figura 12 la potencia de las pruebas T_2 y T_5 según el tamaño de muestra para la configuración $d_0=0.15$, $\alpha=0.05$, $p_1=p_2=0.10$.

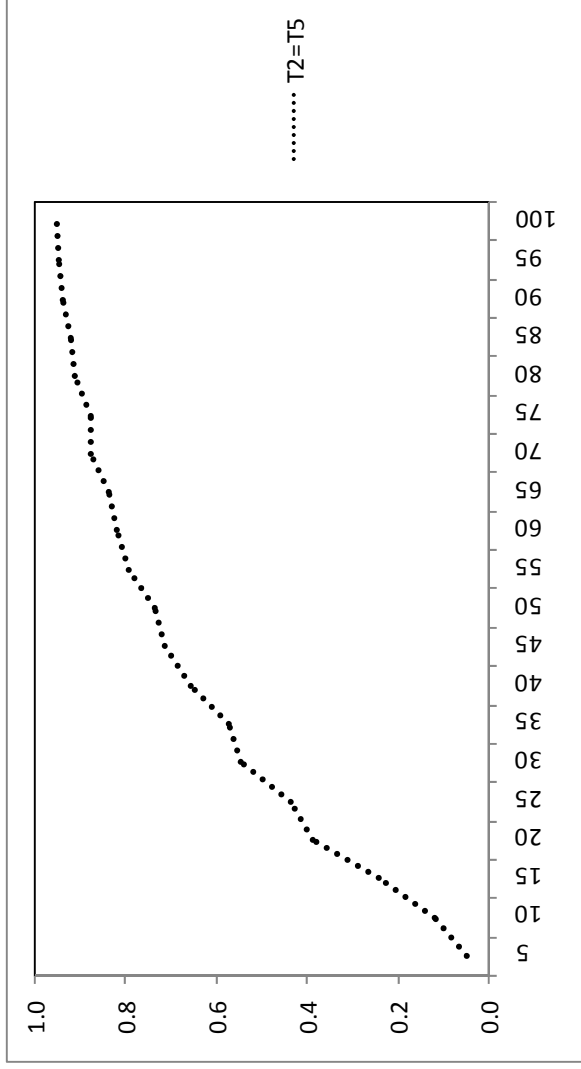
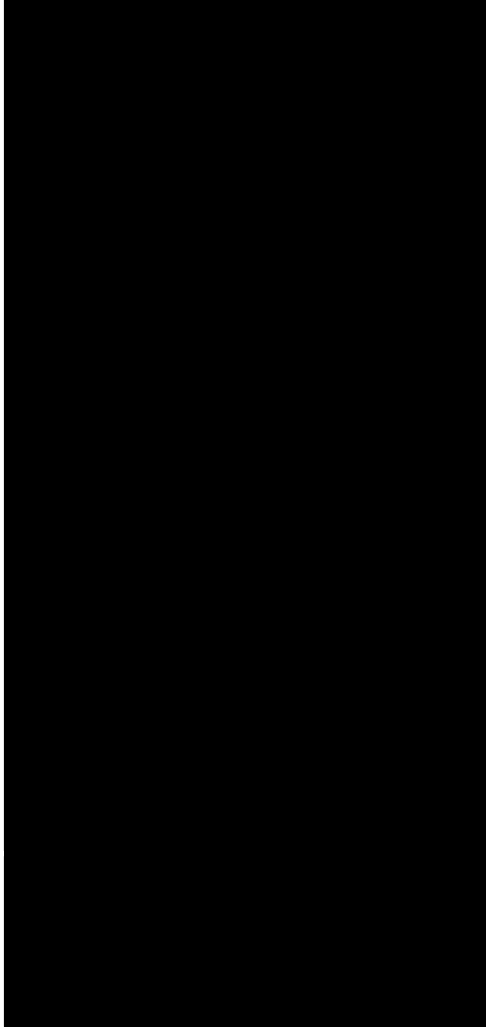


Figura 12. Potencia exacta de las pruebas T_2 y T_5 como función del tamaño de muestra para $d_0=0.15$, $\alpha=0.05$, $p_1=p_2=0.10$

5.3. Pruebas exactas versus pruebas asintóticas

Para realizar la comparación de las pruebas exactas contra las pruebas asintóticas hay que remitirse a la investigación de Almendra y Sotres (2008) en donde se compararon las siete pruebas asintóticas presentadas. En la investigación mencionada se concluye que las pruebas que mantienen los niveles de significancia a un nivel adecuado; es decir que para un nivel nominal de $\alpha=0.05$ un alto porcentaje están en el intervalo $[0.04, 0.05]$, son las pruebas T_2 y T_5 con corrección por continuidad según se puede observar en la Tabla 7.

Tabla 7. Porcentaje de niveles de significancia reales que pertenecen al intervalo $[.04, .05]$



De la Tabla 7 podemos observar que las mejores pruebas asintóticas son las pruebas T_2 y la prueba T_5 con corrección por continuidad para los tres márgenes de no inferioridad estudiados.

Recordando que las pruebas exactas que resultaron mejores en el comportamiento del nivel de significancia fueron las pruebas T_2 y T_5 . Por lo se hace una comparación entre las pruebas asintóticas y exactas considerando el tamaño de muestra. En la Tabla 8 y en la Figura 12 se presenta una comparación de la prueba T_2 considerando el tamaño de muestra para el nivel nominal de $\alpha=0.01$.

Tabla 8. Porcentajes de los niveles de significancia que pertenecen al intervalo especificado para la prueba T_2 en su versión exacta y asintótica sin y con corrección por continuidad para el nivel nominal $\alpha=0.01$.

intervalo	d_0	$5 \leq n \leq 30$						$31 \leq n \leq 100$			
		Exacta		Asintótica		Asintótica		Exacta		Asintótica	Asintótica
		sin cc	con cc	sin cc	con cc	sin cc	con cc	sin cc	con cc	sin cc	con cc
[0,.008)	0.10	12	4	77							4
	0.15	8		50							1
	0.20	15		65		1					6
[.008,.01]	0.10	88	4	23	100						89
	0.15	92	12	50	100						93
	0.20	85	4	35	99						81
(.01,.012)	0.10		50						99		7
	0.15		58						84		6
	0.20		46						74		13
≥ 0.012	0.10		42						1		
	0.15		31						16		
	0.20		50						26		

En la Tabla 8 y en la Figura 13 se puede observar que la prueba T_2 en su versión asintótica necesita utilizar la corrección por continuidad (cc) ya que sin ella es demasiado liberal porque la mayoría de los niveles de significancia están muy por encima del valor nominal. Ahora, considerando los tamaños de muestra de $5 \leq n \leq 30$, la prueba T_2 exacta tiene un comportamiento mejor del nivel de significancia respecto a la prueba asintótica con cc porque esta se vuelve demasiado conservadora. En el tamaño de muestra de $31 \leq n \leq 100$ aunque la versión exacta sigue siendo superior, el comportamiento de los niveles de significancia de la prueba asintótica es bueno porque la mayoría de los tamaños de muestra

tienen un nivel de significancia que pertenece al intervalo de interés que para este caso es $[0.008, 0.01]$, mostrando su mejor comportamiento para el margen de no inferioridad de $d_0=0.15$.

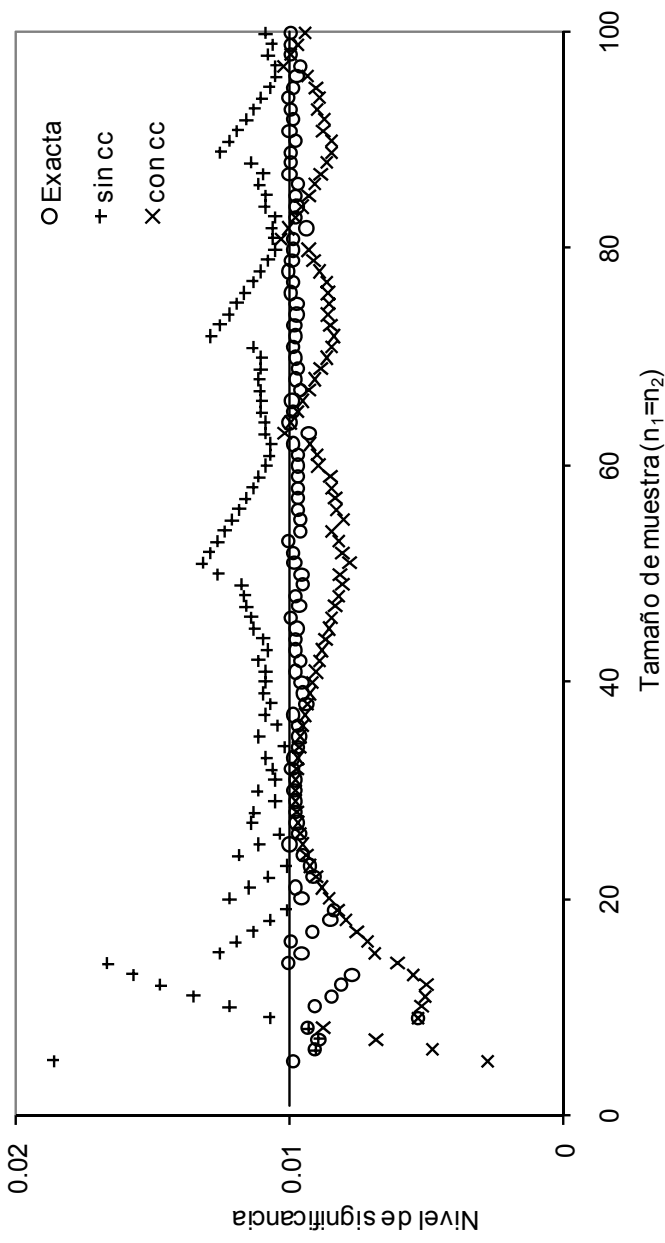


Figura 13. Niveles de significancia de la prueba T_2 en sus versiones exacta, sin corrección por continuidad (sin cc) y con corrección por continuidad (con cc) para el nivel nominal de $\alpha = 0.01$ y el margen de no inferioridad de $d_0=0.15$

En la Tabla 9 y en la Figura 14 se presenta el comportamiento de la prueba T_2 para el nivel de significancia nominal $\alpha=0.05$.

Tabla 9. Porcentajes de los niveles de significancia reales que pertenecen al intervalo especificado para la prueba T_2 en su versión exacta y asintótica sin y con corrección por continuidad para un nivel nominal de $\alpha=0.05$.

Nivel nominal	intervalo	d_0	$5 \leq n \leq 30$			$31 \leq n \leq 100$		
			Asintótica		Exacta	Asintótica		Exacta
			sin cc	con cc		sin cc	con cc	
$\alpha = 0.05$	[0,.04)	0.10		46	4			
		0.15		69	15			3
		0.20		50	23	1		
	[.04,.05]	0.10		54	96	100		99
		0.15		31	85	100		91
		0.20		46	77	99		99
	(.05,.06)	0.10			58		91	1
		0.15			50		91	6
		0.20			65		87	1
	≥ 0.06	0.10			27		9	
		0.15			50		9	
		0.20			35		13	

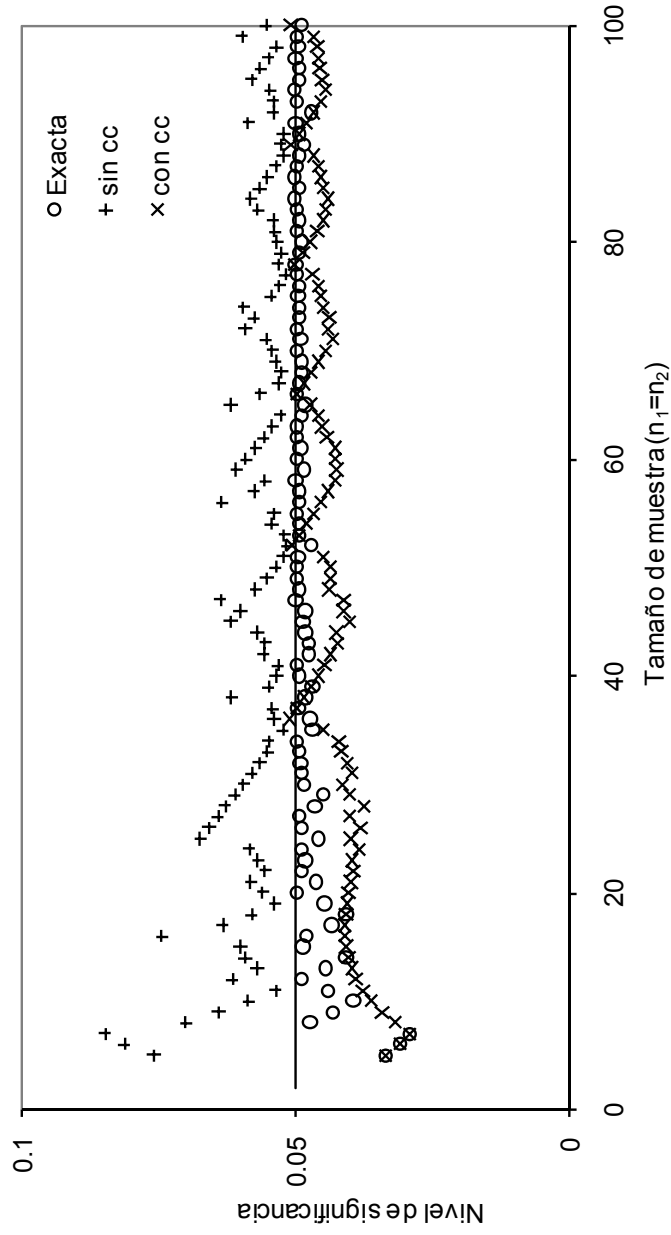


Figura 14. Niveles de significancia de la prueba T_2 en sus versiones exacta, sin corrección por continuidad (sin cc) y con corrección por continuidad (con cc) para el nivel nominal de $\alpha = 0.05$ y el margen de no inferioridad $d_0=0.15$

Para el nivel nominal de $\alpha=0.05$ en la Tabla 9 podemos apreciar que la prueba T_2 requiere utilizar una corrección por continuidad (cc) ya que sin ella se vuelve demasiado liberal al colocar el N.S. para la mayoría de los tamaños de muestra por encima del nivel nominal. Comparando los N.S. para los tamaños de muestra $5 \leq n \leq 30$ para la prueba T_2 exacta y asíntótica con cc se puede ver que esta última es muy conservadora y que se ve afectada por el margen de no inferioridad especificado. La prueba exacta tiene los mayores porcentajes del N.S. en el intervalo de interés que es [0.04, 0.05] lo que no la caracteriza como conservadora.

La prueba T_2 exacta es mejor que la prueba asíntótica porque sus niveles de significancia reales son uniformemente mayores, para tamaños de muestra de $5 \leq n_1=n_2 \leq 100$, es decir es más eficiente en el sentido de que esta por debajo del nivel nominal pero por encima de la prueba asíntótica.

Para los tamaños de muestra $31 \leq n \leq 100$ las dos versiones de la prueba T_2 , la exacta y la asíntótica, no resultan conservadoras ya que las dos concentran más del 90% de los niveles de significancia en el intervalo de interés [0.04, 0.05].

5.4. Tablas de la prueba T_2 exacta

Se han calculado en la Tabla 10 y en la Tabla 11 la prueba T_2 exacta que contienen la constante crítica y el nivel de significancia para tamaños de muestra de $5 \leq n \leq 200$, con lo cual se puede realizar la prueba exacta.

Tabla 10. Constantes críticas (c_α) y Niveles de Significancia (R.L.S.) para T_2 exacta, $\alpha = 0.01$

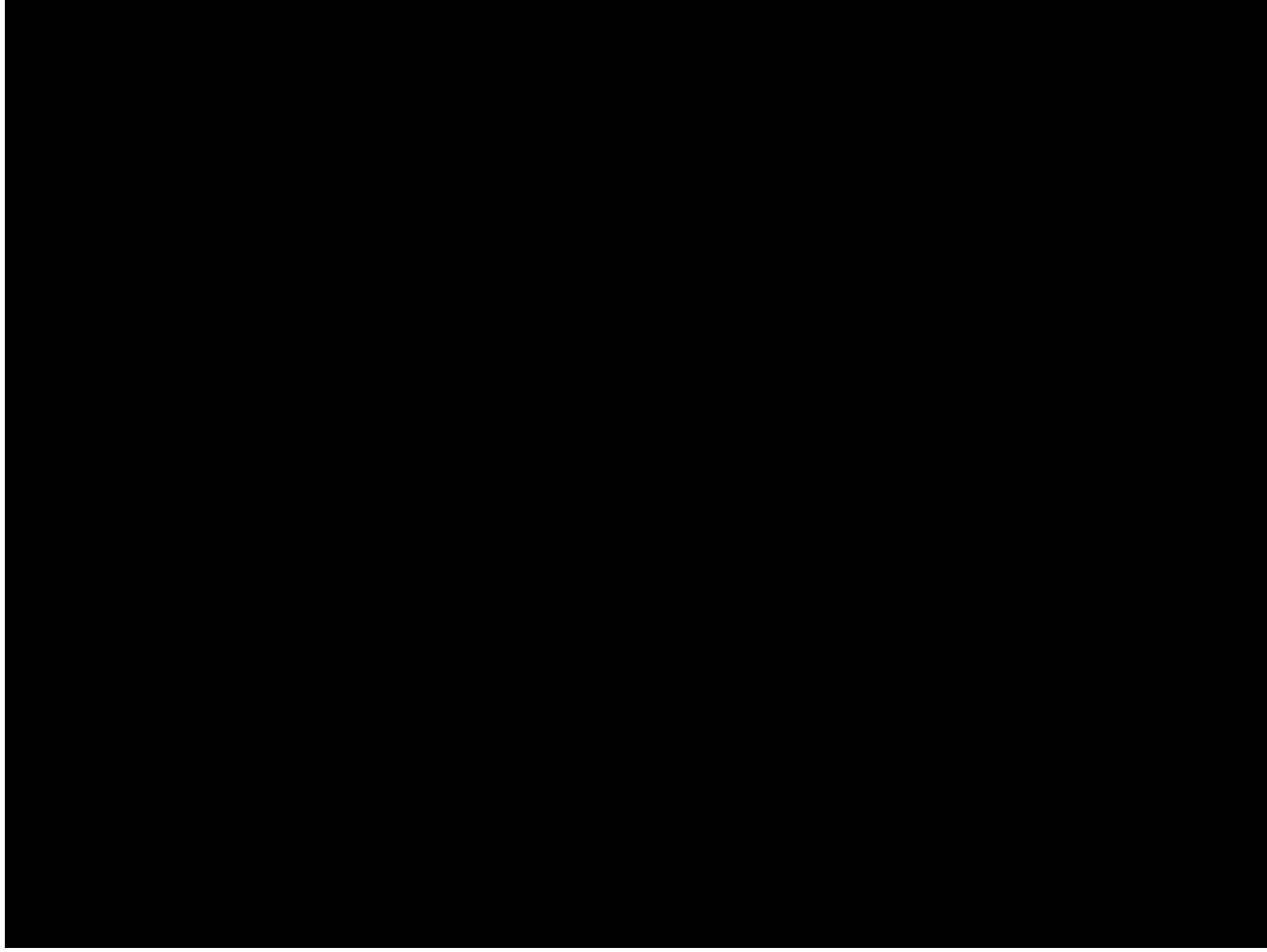


Tabla 10. Continuación

n	Nivel de Significancia Nominal $\alpha=0.01$									
	$d_0 = 0.05$		$d_0 = 0.1$		$d_0 = 0.15$		$d_0 = 0.2$		$d_0 = 0.25$	
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real
51	-2.3773	0.0095	-2.4091	0.0100	-2.3766	0.0098	-2.3890	0.0096	-2.4109	0.0097
52	-2.3741	0.0095	-2.4084	0.0096	-2.3721	0.0098	-2.3814	0.0094	-2.4340	0.0096
53	-2.3716	0.0095	-2.4077	0.0094	-2.3686	0.0100	-2.3799	0.0095	-2.3783	0.0099
54	-2.3695	0.0095	-2.4018	0.0098	-2.3805	0.0095	-2.3783	0.0099	-2.3816	0.0094
55	-2.3680	0.0097	-2.4015	0.0096	-2.3866	0.0096	-2.3769	0.0098	-2.3902	0.0097
56	-2.3454	0.0100	-2.3964	0.0099	-2.3921	0.0097	-2.3926	0.0097	-2.3894	0.0099
57	-2.4132	0.0093	-2.3964	0.0097	-2.3940	0.0097	-2.3860	0.0099	-2.3872	0.0099
58	-2.3998	0.0099	-2.3971	0.0096	-2.3915	0.0097	-2.4034	0.0096	-2.4084	0.0098
59	-2.3902	0.0099	-2.3921	0.0098	-2.4007	0.0097	-2.4101	0.0098	-2.4266	0.0097
60	-2.3815	0.0099	-2.3931	0.0097	-2.4054	0.0097	-2.4234	0.0095	-2.3766	0.0099
61	-2.3736	0.0100	-2.3871	0.0099	-2.4110	0.0097	-2.3693	0.0100	-2.3782	0.0099
62	-2.3681	0.0098	-2.3900	0.0097	-2.4164	0.0098	-2.3703	0.0095	-2.3674	0.0100
63	-2.3680	0.0098	-2.3861	0.0100	-2.4263	0.0093	-2.3683	0.0100	-2.3874	0.0096
64	-2.3586	0.0099	-2.3875	0.0098	-2.3526	0.0100	-2.3756	0.0098	-2.3940	0.0099
65	-2.3568	0.0100	-2.3842	0.0100	-2.3447	0.0098	-2.3766	0.0097	-2.4093	0.0097
66	-2.3559	0.0100	-2.3820	0.0099	-2.3613	0.0099	-2.3770	0.0099	-2.4293	0.0096
67	-2.3672	0.0096	-2.3865	0.0097	-2.3732	0.0095	-2.3884	0.0096	-2.3681	0.0098
68	-2.3648	0.0096	-2.3841	0.0099	-2.3596	0.0098	-2.3952	0.0097	-2.3634	0.0100
69	-2.3628	0.0096	-2.3828	0.0097	-2.3731	0.0097	-2.4035	0.0097	-2.3793	0.0100
70	-2.3480	0.0099	-2.3840	0.0099	-2.3585	0.0098	-2.4161	0.0099	-2.3783	0.0099
71	-2.3505	0.0100	-2.3864	0.0097	-2.3594	0.0099	-2.3579	0.0100	-2.3982	0.0097
72	-2.3442	0.0099	-2.3822	0.0099	-2.3736	0.0098	-2.3695	0.0098	-2.4129	0.0098
73	-2.3524	0.0098	-2.3864	0.0097	-2.3674	0.0098	-2.3672	0.0097	-2.4378	0.0093
74	-2.3577	0.0098	-2.3844	0.0098	-2.3782	0.0097	-2.3719	0.0099	-2.3774	0.0098
75	-2.3575	0.0099	-2.3829	0.0100	-2.3798	0.0097	-2.3792	0.0097	-2.3733	0.0098
76	-2.4046	0.0094	-2.3852	0.0098	-2.3723	0.0099	-2.3792	0.0098	-2.3842	0.0098
77	-2.3971	0.0098	-2.3820	0.0099	-2.3836	0.0098	-2.3816	0.0098	-2.3897	0.0099
78	-2.3906	0.0097	-2.3864	0.0098	-2.3845	0.0100	-2.3947	0.0098	-2.4024	0.0099
79	-2.3846	0.0098	-2.3856	0.0099	-2.3926	0.0099	-2.4062	0.0097	-2.4245	0.0098
80	-2.3769	0.0098	-2.3482	0.0097	-2.4004	0.0098	-2.4216	0.0094	-2.3693	0.0096
81	-2.3741	0.0099	-2.3474	0.0098	-2.4085	0.0098	-2.3594	0.0100	-2.3638	0.0100
82	-2.3695	0.0099	-2.3471	0.0100	-2.4183	0.0093	-2.3598	0.0098	-2.3749	0.0099
83	-2.3618	0.0100	-2.3896	0.0098	-2.4586	0.0098	-2.3599	0.0099	-2.3868	0.0097
84	-2.3656	0.0098	-2.3892	0.0098	-2.3570	0.0097	-2.3698	0.0100	-2.3959	0.0100
85	-2.3590	0.0098	-2.3921	0.0097	-2.3583	0.0098	-2.3682	0.0100	-2.4173	0.0097
86	-2.3624	0.0098	-2.3923	0.0098	-2.3704	0.0096	-2.3769	0.0100	-2.3619	0.0099
87	-2.3600	0.0098	-2.3944	0.0096	-2.3602	0.0100	-2.3901	0.0099	-2.3725	0.0098
88	-2.3579	0.0099	-2.3952	0.0097	-2.3547	0.0099	-2.4025	0.0097	-2.3682	0.0100
89	-2.3477	0.0099	-2.3961	0.0099	-2.3592	0.0099	-2.4186	0.0095	-2.3875	0.0098
90	-2.3465	0.0099	-2.3984	0.0096	-2.3718	0.0098	-2.3660	0.0098	-2.3929	0.0099
91	-2.3507	0.0099	-2.3994	0.0098	-2.3645	0.0100	-2.3538	0.0100	-2.4120	0.0100
92	-2.3520	0.0099	-2.4018	0.0095	-2.3697	0.0099	-2.3645	0.0097	-2.3619	0.0099
93	-2.3449	0.0099	-2.4029	0.0097	-2.3761	0.0099	-2.3692	0.0098	-2.3703	0.0098
94	-2.3482	0.0099	-2.4052	0.0094	-2.3798	0.0100	-2.3735	0.0098	-2.3729	0.0100
95	-2.3420	0.0100	-2.3431	0.0100	-2.3867	0.0098	-2.3814	0.0099	-2.3803	0.0100
96	-2.3406	0.0100	-2.3499	0.0099	-2.3961	0.0097	-2.3878	0.0100	-2.3912	0.0099
97	-2.3463	0.0099	-2.3414	0.0099	-2.4047	0.0096	-2.4036	0.0100	-2.4109	0.0098
98	-2.3463	0.0099	-2.3478	0.0100	-2.3524	0.0099	-2.3546	0.0099	-2.3610	0.0099
99	-2.3497	0.0099	-2.3453	0.0098	-2.3548	0.0099	-2.3520	0.0100	-2.3733	0.0098
100	-2.3391	0.0098	-2.3504	0.0099	-2.3522	0.0099	-2.3695	0.0097	-2.3746	0.0099

Tabla 10. Continuación

n	Nivel de Significancia Nominal $\alpha=0.01$									
	$d_0 = 0.05$		$d_0 = 0.1$		$d_0 = 0.15$		$d_0 = 0.2$		$d_0 = 0.25$	
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real
101	-2.3500	0.0099	-2.3460	0.0098	-2.3611	0.0099	-2.3710	0.0097	-2.3805	0.0099
102	-2.3417	0.0100	-2.3441	0.0100	-2.3587	0.0098	-2.3706	0.0100	-2.3913	0.0099
103	-2.3938	0.0096	-2.3503	0.0097	-2.3539	0.0100	-2.3780	0.0099	-2.3635	0.0099
104	-2.3886	0.0099	-2.3516	0.0098	-2.3617	0.0099	-2.3799	0.0100	-2.3694	0.0098
105	-2.3847	0.0098	-2.3526	0.0099	-2.3638	0.0100	-2.3444	0.0100	-2.3624	0.0100
106	-2.3801	0.0098	-2.3494	0.0097	-2.3688	0.0099	-2.3610	0.0098	-2.3744	0.0098
107	-2.3775	0.0098	-2.3529	0.0099	-2.3695	0.0099	-2.3547	0.0099	-2.3782	0.0099
108	-2.3724	0.0100	-2.3456	0.0099	-2.3704	0.0100	-2.3642	0.0099	-2.3896	0.0100
109	-2.3711	0.0098	-2.3530	0.0098	-2.3527	0.0099	-2.3672	0.0099	-2.3575	0.0099
110	-2.3679	0.0099	-2.3576	0.0098	-2.3484	0.0098	-2.3711	0.0098	-2.3675	0.0097
111	-2.3395	0.0099	-2.3490	0.0099	-2.3567	0.0098	-2.3695	0.0099	-2.3654	0.0099
112	-2.3503	0.0100	-2.3459	0.0100	-2.3536	0.0098	-2.3793	0.0099	-2.3775	0.0098
113	-2.3555	0.0100	-2.3522	0.0098	-2.3507	0.0100	-2.3493	0.0099	-2.3799	0.0099
114	-2.3533	0.0100	-2.3571	0.0098	-2.3496	0.0099	-2.3479	0.0099	-2.3619	0.0099
115	-2.3545	0.0099	-2.3560	0.0099	-2.3623	0.0098	-2.3580	0.0100	-2.3607	0.0100
116	-2.3548	0.0100	-2.3521	0.0099	-2.3559	0.0099	-2.3602	0.0100	-2.3651	0.0098
117	-2.3556	0.0100	-2.3555	0.0100	-2.3638	0.0098	-2.3593	0.0099	-2.3658	0.0099
118	-2.3544	0.0100	-2.3614	0.0098	-2.3626	0.0099	-2.3671	0.0099	-2.3776	0.0098
119	-2.3534	0.0100	-2.3600	0.0099	-2.3632	0.0100	-2.3719	0.0099	-2.3812	0.0100
120	-2.3487	0.0100	-2.3579	0.0099	-2.3655	0.0100	-2.3746	0.0100	-2.3541	0.0100
121	-2.3508	0.0100	-2.3566	0.0099	-2.3666	0.0100	-2.3548	0.0100	-2.3619	0.0098
122	-2.3493	0.0100	-2.3614	0.0099	-2.3464	0.0100	-2.3539	0.0100	-2.3664	0.0099
123	-2.3509	0.0099	-2.3612	0.0099	-2.3507	0.0098	-2.3554	0.0100	-2.3686	0.0099
124	-2.3469	0.0099	-2.3596	0.0099	-2.3498	0.0099	-2.3533	0.0100	-2.3733	0.0100
125	-2.3438	0.0099	-2.3609	0.0099	-2.3480	0.0099	-2.3644	0.0099	-2.3617	0.0100
126	-2.3459	0.0099	-2.3623	0.0100	-2.3548	0.0098	-2.3653	0.0100	-2.3606	0.0098
127	-2.3491	0.0099	-2.3486	0.0100	-2.3477	0.0099	-2.3741	0.0098	-2.3581	0.0099
128	-2.3442	0.0099	-2.3515	0.0099	-2.3583	0.0098	-2.3746	0.0100	-2.3664	0.0099
129	-2.3472	0.0098	-2.3416	0.0100	-2.3583	0.0098	-2.3525	0.0100	-2.3732	0.0098
130	-2.3512	0.0098	-2.3388	0.0099	-2.3615	0.0098	-2.3543	0.0100	-2.3792	0.0100
131	-2.3468	0.0098	-2.3443	0.0098	-2.3582	0.0099	-2.3549	0.0100	-2.3559	0.0100
132	-2.3458	0.0098	-2.3436	0.0098	-2.3611	0.0099	-2.3543	0.0099	-2.3636	0.0098
133	-2.3388	0.0100	-2.3464	0.0099	-2.3643	0.0099	-2.6094	0.0050	-2.3627	0.0099
134	-2.3454	0.0099	-2.3408	0.0099	-2.3676	0.0100	-2.3673	0.0099	-2.3673	0.0098
135	-2.3401	0.0099	-2.3487	0.0098	-2.3406	0.0099	-2.3739	0.0099	-2.3766	0.0099
136	-2.3410	0.0099	-2.3450	0.0098	-2.3503	0.0099	-2.3441	0.0100	-2.3802	0.0100
137	-2.3431	0.0099	-2.3494	0.0099	-2.3502	0.0098	-2.3536	0.0099	-2.3579	0.0100
138	-2.3405	0.0098	-2.3397	0.0100	-2.3505	0.0099	-2.3531	0.0099	-2.3638	0.0098
139	-2.3461	0.0098	-2.3418	0.0100	-2.3512	0.0098	-2.3532	0.0099	-2.3618	0.0100
140	-2.3379	0.0100	-2.3482	0.0098	-2.3523	0.0099	-2.3541	0.0100	-2.3682	0.0099
141	-2.3352	0.0100	-2.3495	0.0098	-2.3534	0.0100	-2.3630	0.0099	-2.3749	0.0100
142	-2.3373	0.0099	-2.3474	0.0099	-2.3556	0.0098	-2.3687	0.0099	-2.3573	0.0099
143	-2.3441	0.0099	-2.3466	0.0099	-2.3577	0.0099	-2.3693	0.0100	-2.3525	0.0099
144	-2.3394	0.0098	-2.3497	0.0099	-2.3601	0.0099	-2.3506	0.0099	-2.3590	0.0099
145	-2.3450	0.0098	-2.3446	0.0099	-2.3623	0.0099	-2.3517	0.0099	-2.3677	0.0098
146	-2.3401	0.0100	-2.3462	0.0100	-2.3649	0.0099	-2.3537	0.0098	-2.3712	0.0100
147	-2.3348	0.0100	-2.3467	0.0100	-2.3417	0.0099	-2.3545	0.0099	-2.3538	0.0100
148	-2.3378	0.0099	-2.3507	0.0100	-2.3481	0.0098	-2.3571	0.0100	-2.3616	0.0098
149	-2.3360	0.0099	-2.3467	0.0100	-2.3422	0.0100	-2.3633	0.0100	-2.3647	0.0098
150	-2.3394	0.0098	-2.3508	0.0100	-2.3521	0.0098	-2.3696	0.0100	-2.3606	0.0100

Tabla 10. Continuación

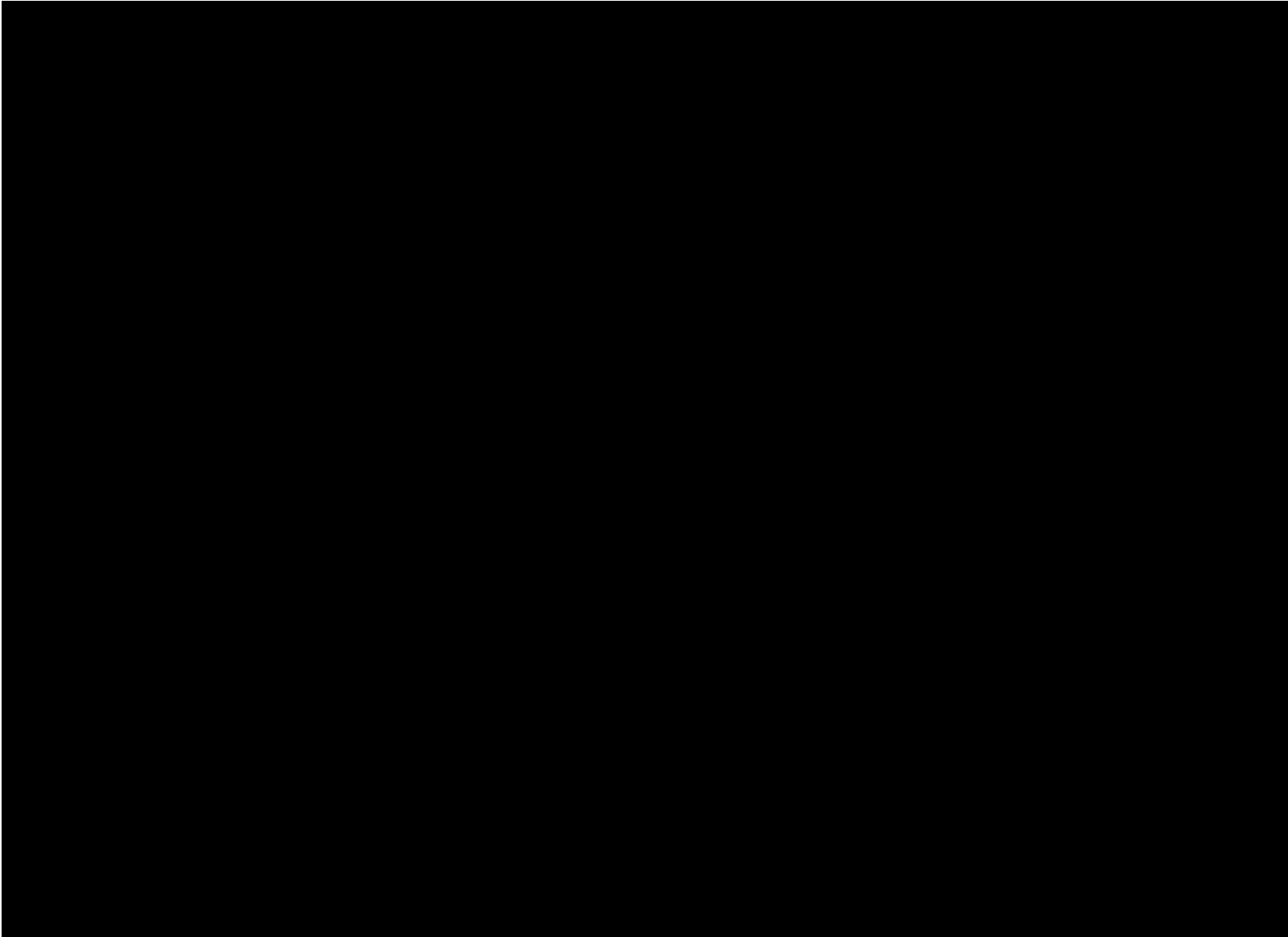


Tabla 11. Constantes críticas (c_α) y Niveles de Significancia (R.L.S.) Para T2 exacta, $\alpha = 0.05$

n	Nivel de Significancia Nominal $\alpha=0.05$									
	$d_0 = 0.05$		$d_0 = 0.10$		$d_0 = 0.15$		$d_0 = 0.20$		$d_0 = 0.25$	
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real
5	-2.0581	0.0392	-1.8724	0.0481	-2.0258	0.0333	-2.1865	0.0226	-2.1480	0.0384
6	-1.9319	0.0397	-1.8696	0.0465	-2.0362	0.0308	-1.9608	0.0392	-1.9180	0.0402
7	-1.8631	0.0443	-1.8776	0.0448	-2.0573	0.0291	-1.9842	0.0333	-1.9970	0.0389
8	-1.8229	0.0445	-1.8916	0.0492	-1.8011	0.0472	-1.8683	0.0453	-1.7990	0.0483
9	-1.7990	0.0455	-1.8877	0.0426	-1.8307	0.0431	-1.9077	0.0394	-1.8816	0.0391
10	-1.7852	0.0462	-1.8712	0.0412	-1.8644	0.0393	-1.8541	0.0451	-1.9625	0.0346
11	-1.7780	0.0466	-1.8661	0.0399	-1.8105	0.0439	-1.8820	0.0394	-1.8577	0.0422
12	-1.7755	0.0470	-1.7724	0.0479	-1.7578	0.0488	-1.8403	0.0445	-1.8059	0.0446
13	-1.8732	0.0376	-1.7702	0.0466	-1.7924	0.0445	-1.8811	0.0394	-1.7868	0.0474
14	-1.7797	0.0477	-1.7739	0.0453	-1.8285	0.0406	-1.8516	0.0500	-1.7885	0.0478
15	-1.7702	0.0480	-1.7472	0.0493	-1.7355	0.0485	-1.8339	0.0481	-1.8038	0.0485
16	-1.7915	0.0448	-1.7747	0.0466	-1.7599	0.0478	-1.7461	0.0486	-1.8512	0.0416
17	-1.7380	0.0500	-1.7595	0.0476	-1.8034	0.0433	-1.8069	0.0445	-1.8598	0.0480
18	-1.7285	0.0490	-1.7294	0.0486	-1.8386	0.0406	-1.7526	0.0461	-1.7752	0.0445
19	-1.7223	0.0489	-1.7421	0.0467	-1.7693	0.0456	-1.7480	0.0457	-1.7547	0.0465
20	-1.7185	0.0488	-1.7564	0.0448	-1.7225	0.0497	-1.7678	0.0448	-1.7252	0.0484
21	-1.7168	0.0485	-1.7129	0.0497	-1.7522	0.0460	-1.7470	0.0465	-1.7543	0.0470
22	-1.7167	0.0482	-1.7398	0.0467	-1.7183	0.0488	-1.7375	0.0481	-1.7539	0.0466
23	-1.7181	0.0478	-1.7254	0.0480	-1.7262	0.0480	-1.7149	0.0493	-1.7637	0.0492
24	-1.7206	0.0497	-1.7410	0.0460	-1.7212	0.0487	-1.7414	0.0485	-1.7936	0.0472
25	-1.7754	0.0467	-1.7193	0.0465	-1.7473	0.0456	-1.7514	0.0479	-1.7037	0.0498
26	-1.7607	0.0463	-1.7066	0.0485	-1.7119	0.0486	-1.7652	0.0475	-1.7077	0.0498
27	-1.7489	0.0465	-1.7809	0.0453	-1.7175	0.0491	-1.7799	0.0473	-1.7177	0.0494
28	-1.7393	0.0472	-1.8107	0.0457	-1.7403	0.0463	-1.7095	0.0492	-1.7742	0.0435
29	-1.7256	0.0483	-1.6963	0.0494	-1.7423	0.0449	-1.7663	0.0468	-1.7269	0.0495
30	-1.7260	0.0475	-1.7115	0.0475	-1.7287	0.0482	-1.7036	0.0498	-1.7481	0.0474
31	-1.7216	0.0474	-1.7022	0.0487	-1.7287	0.0489	-1.7034	0.0494	-1.7739	0.0481
32	-1.7185	0.0474	-1.6920	0.0499	-1.7317	0.0489	-1.7076	0.0495	-1.6968	0.0486
33	-1.7165	0.0472	-1.7069	0.0480	-1.7317	0.0491	-1.6993	0.0494	-1.7495	0.0447
34	-1.7058	0.0485	-1.6946	0.0473	-1.7444	0.0497	-1.7258	0.0481	-1.7230	0.0478
35	-1.7153	0.0468	-1.6972	0.0485	-1.7588	0.0468	-1.7181	0.0496	-1.7164	0.0492
36	-1.7060	0.0495	-1.7068	0.0482	-1.7666	0.0473	-1.7537	0.0461	-1.7250	0.0494
37	-1.7139	0.0475	-1.7118	0.0473	-1.6892	0.0494	-1.7573	0.0469	-1.7459	0.0489
38	-1.7188	0.0470	-1.6955	0.0485	-1.7170	0.0481	-1.6944	0.0493	-1.6937	0.0500
39	-1.7212	0.0465	-1.7015	0.0482	-1.7449	0.0468	-1.6763	0.0493	-1.7085	0.0498
40	-1.6855	0.0493	-1.6804	0.0495	-1.6950	0.0491	-1.7076	0.0483	-1.6971	0.0499
41	-1.6877	0.0489	-1.6908	0.0493	-1.6895	0.0496	-1.6999	0.0481	-1.7197	0.0480
42	-1.6904	0.0485	-1.6914	0.0491	-1.7053	0.0475	-1.7012	0.0495	-1.7256	0.0481
43	-1.6834	0.0499	-1.6959	0.0487	-1.6867	0.0488	-1.6996	0.0491	-1.7328	0.0492
44	-1.7518	0.0473	-1.7094	0.0471	-1.7005	0.0481	-1.7169	0.0491	-1.7652	0.0465
45	-1.7438	0.0468	-1.7526	0.0454	-1.6950	0.0485	-1.7275	0.0491	-1.6887	0.0499
46	-1.7344	0.0490	-1.6952	0.0484	-1.7116	0.0481	-1.7484	0.0476	-1.6942	0.0481
47	-1.7280	0.0484	-1.6911	0.0488	-1.7025	0.0498	-1.6820	0.0500	-1.7104	0.0485
48	-1.7226	0.0483	-1.6876	0.0485	-1.7072	0.0493	-1.6942	0.0488	-1.7079	0.0500
49	-1.7121	0.0500	-1.6705	0.0490	-1.7182	0.0495	-1.7294	0.0452	-1.7330	0.0473
50	-1.7074	0.0499	-1.6809	0.0499	-1.7223	0.0496	-1.6943	0.0496	-1.7570	0.0480

Tabla 11. Continuación

n	Nivel de Significancia Nominal $\alpha=0.05$									
	$d_0 = 0.05$	$d_0 = 0.10$	$d_0 = 0.15$	$d_0 = 0.20$	$d_0 = 0.25$					
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real	C_α					
51	-1.7032	0.0498	-1.6911	0.0490	-1.7338	0.0494	-1.7020	0.0484	-1.6949	0.0477
52	-1.6997	0.0498	-1.6755	0.0497	-1.7483	0.0469	-1.7008	0.0488	-1.6849	0.0491
53	-1.6967	0.0498	-1.6800	0.0494	-1.6889	0.0491	-1.7120	0.0500	-1.6959	0.0485
54	-1.6946	0.0498	-1.6874	0.0495	-1.6871	0.0491	-1.7328	0.0482	-1.7353	0.0460
55	-1.7040	0.0481	-1.6865	0.0485	-1.6823	0.0495	-1.7027	0.0499	-1.7274	0.0491
56	-1.7180	0.0461	-1.6887	0.0491	-1.6749	0.0491	-1.6937	0.0493	-1.6861	0.0497
57	-1.7364	0.0439	-1.6868	0.0496	-1.6912	0.0491	-1.6845	0.0499	-1.6911	0.0482
58	-1.6837	0.0494	-1.6970	0.0484	-1.6840	0.0499	-1.6908	0.0488	-1.7036	0.0481
59	-1.6879	0.0500	-1.6927	0.0490	-1.6958	0.0484	-1.6920	0.0492	-1.7301	0.0458
60	-1.6905	0.0490	-1.7358	0.0441	-1.6931	0.0496	-1.6952	0.0492	-1.7110	0.0494
61	-1.6912	0.0487	-1.6874	0.0499	-1.6936	0.0490	-1.7078	0.0499	-1.7318	0.0490
62	-1.6921	0.0485	-1.6901	0.0488	-1.7027	0.0495	-1.7250	0.0492	-1.6856	0.0494
63	-1.6793	0.0499	-1.6922	0.0492	-1.7056	0.0497	-1.6783	0.0500	-1.6769	0.0496
64	-1.6802	0.0498	-1.6940	0.0494	-1.7196	0.0489	-1.6691	0.0498	-1.6907	0.0494
65	-1.6719	0.0495	-1.6914	0.0496	-1.7316	0.0481	-1.6862	0.0487	-1.7014	0.0494
66	-1.6798	0.0498	-1.6891	0.0497	-1.7053	0.0495	-1.6891	0.0490	-1.7125	0.0499
67	-1.6847	0.0488	-1.6919	0.0498	-1.6768	0.0491	-1.6915	0.0491	-1.7405	0.0486
68	-1.6866	0.0485	-1.6977	0.0498	-1.6865	0.0488	-1.6994	0.0488	-1.6886	0.0487
69	-1.6790	0.0492	-1.7001	0.0498	-1.6789	0.0489	-1.7111	0.0491	-1.6863	0.0499
70	-1.6732	0.0497	-1.7030	0.0498	-1.6732	0.0497	-1.7258	0.0496	-1.6999	0.0489
71	-1.6710	0.0495	-1.7063	0.0497	-1.6821	0.0490	-1.6835	0.0492	-1.7028	0.0499
72	-1.6791	0.0490	-1.7096	0.0498	-1.6862	0.0498	-1.6804	0.0497	-1.7227	0.0489
73	-1.6700	0.0498	-1.7159	0.0487	-1.6903	0.0491	-1.6815	0.0499	-1.7002	0.0488
74	-1.6757	0.0489	-1.7196	0.0487	-1.6972	0.0493	-1.6863	0.0498	-1.6840	0.0491
75	-1.6726	0.0498	-1.7233	0.0490	-1.7021	0.0494	-1.6944	0.0494	-1.6804	0.0496
76	-1.6646	0.0494	-1.7295	0.0477	-1.7087	0.0492	-1.6999	0.0497	-1.6933	0.0495
77	-1.6779	0.0490	-1.6717	0.0496	-1.7203	0.0495	-1.7128	0.0495	-1.7100	0.0495
78	-1.6684	0.0498	-1.6691	0.0493	-1.6683	0.0498	-1.7336	0.0478	-1.7377	0.0478
79	-1.6742	0.0491	-1.6650	0.0497	-1.6801	0.0491	-1.6857	0.0489	-1.6700	0.0498
80	-1.6712	0.0497	-1.6302	0.0495	-1.6805	0.0488	-1.6860	0.0491	-1.6866	0.0489
81	-1.6745	0.0493	-1.6195	0.0500	-1.6672	0.0497	-1.6840	0.0490	-1.6907	0.0500
82	-1.6776	0.0489	-1.6291	0.0487	-1.6785	0.0492	-1.6837	0.0498	-1.7055	0.0492
83	-1.6671	0.0498	-1.6816	0.0489	-1.6775	0.0497	-1.6934	0.0500	-1.7245	0.0489
84	-1.6582	0.0499	-1.6694	0.0498	-1.6776	0.0500	-1.7045	0.0497	-1.6843	0.0487
85	-1.6692	0.0494	-1.6720	0.0495	-1.6930	0.0491	-1.7231	0.0487	-1.6777	0.0492
86	-1.6753	0.0497	-1.6781	0.0496	-1.6903	0.0499	-1.6794	0.0493	-1.6885	0.0492
87	-1.6737	0.0498	-1.6819	0.0484	-1.6998	0.0497	-1.6768	0.0495	-1.6986	0.0496
88	-1.7217	0.0479	-1.7094	0.0461	-1.7095	0.0493	-1.6856	0.0492	-1.7130	0.0499
89	-1.7190	0.0488	-1.6793	0.0498	-1.7218	0.0484	-1.6851	0.0498	-1.6740	0.0491
90	-1.7174	0.0483	-1.6806	0.0497	-1.6729	0.0491	-1.6903	0.0499	-1.6743	0.0494
91	-1.7148	0.0491	-1.6781	0.0497	-1.6709	0.0498	-1.7008	0.0496	-1.6867	0.0494
92	-1.7134	0.0487	-1.6794	0.0496	-1.6938	0.0469	-1.7164	0.0492	-1.6936	0.0498
93	-1.7110	0.0495	-1.6867	0.0495	-1.6716	0.0497	-1.6764	0.0495	-1.7061	0.0499
94	-1.7096	0.0490	-1.6858	0.0493	-1.6779	0.0499	-1.6717	0.0494	-1.6818	0.0497
95	-1.7074	0.0498	-1.6879	0.0499	-1.6820	0.0491	-1.6848	0.0491	-1.6778	0.0496
96	-1.7061	0.0493	-1.6885	0.0496	-1.6842	0.0492	-1.6848	0.0496	-1.6811	0.0495
97	-1.7051	0.0491	-1.6927	0.0500	-1.6898	0.0498	-1.6899	0.0497	-1.7016	0.0481
98	-1.7029	0.0496	-1.6973	0.0497	-1.6962	0.0494	-1.6964	0.0500	-1.7006	0.0500
99	-1.7020	0.0493	-1.7020	0.0494	-1.7041	0.0496	-1.7125	0.0494	-1.7252	0.0489
100	-1.6999	0.0499	-1.7069	0.0491	-1.7172	0.0486	-1.6697	0.0496	-1.6778	0.0496

Tabla 11. Continuación

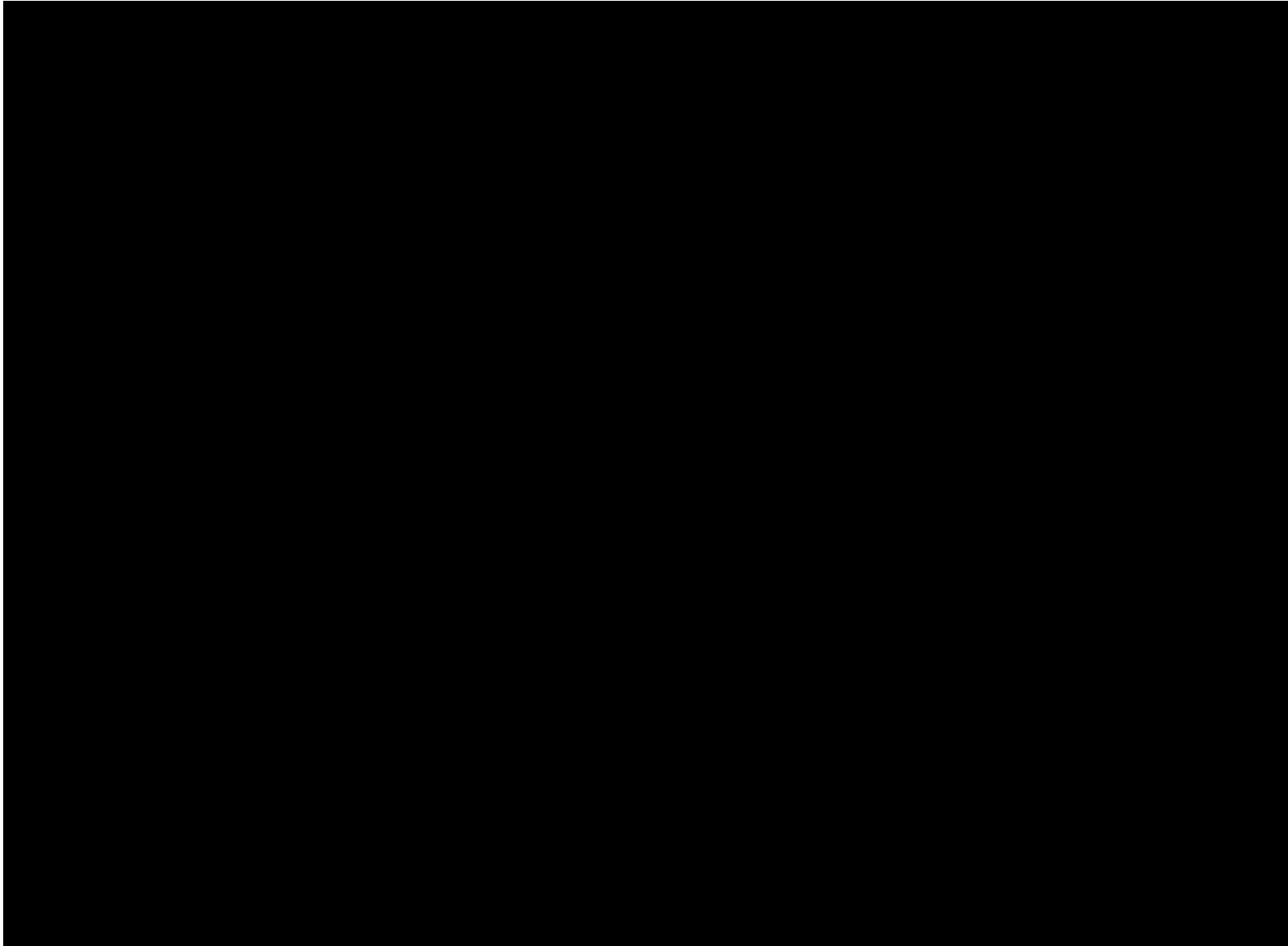


Tabla 11. Continuación

n	Nivel de Significancia Nominal $\alpha=0.05$									
	$d_0 = 0.05$	$d_0 = 0.10$	$d_0 = 0.15$	$d_0 = 0.20$						
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real						
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real						
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real						
	C_α	Nivel de Significancia Real	C_α	Nivel de Significancia Real						
151	-1.6579	0.0500	-1.6712	0.0495	-1.6651	0.0499	-1.6816	0.0496	-1.6767	0.0495
152	-1.6587	0.0499	-1.6800	0.0486	-1.6640	0.0499	-1.6866	0.0500	-1.6817	0.0500
153	-1.6633	0.0496	-1.6718	0.0495	-1.6664	0.0498	-1.6623	0.0498	-1.6886	0.0500
154	-1.6553	0.0500	-1.6714	0.0497	-1.6614	0.0496	-1.6636	0.0499	-1.6669	0.0495
155	-1.6578	0.0499	-1.6719	0.0499	-1.6700	0.0499	-1.6697	0.0498	-1.6721	0.0497
156	-1.6635	0.0495	-1.6763	0.0500	-1.6762	0.0495	-1.6759	0.0495	-1.6785	0.0495
157	-1.6549	0.0499	-1.6783	0.0500	-1.6801	0.0496	-1.6787	0.0495	-1.6802	0.0495
158	-1.6586	0.0498	-1.6809	0.0498	-1.6822	0.0498	-1.6849	0.0497	-1.6893	0.0498
159	-1.6589	0.0500	-1.6603	0.0497	-1.6560	0.0500	-1.6676	0.0493	-1.6712	0.0496
160	-1.6556	0.0499	-1.6606	0.0500	-1.6579	0.0500	-1.6694	0.0494	-1.6705	0.0497
161	-1.6601	0.0494	-1.6572	0.0500	-1.6601	0.0500	-1.6702	0.0494	-1.6754	0.0500
162	-1.6576	0.0497	-1.6586	0.0497	-1.6626	0.0498	-1.6721	0.0498	-1.6798	0.0498
163	-1.6607	0.0494	-1.6666	0.0495	-1.6675	0.0496	-1.6735	0.0499	-1.6899	0.0499
164	-1.6553	0.0499	-1.6659	0.0493	-1.6679	0.0499	-1.6812	0.0496	-1.6645	0.0497
165	-1.6605	0.0495	-1.6605	0.0498	-1.6721	0.0495	-1.6657	0.0499	-1.6693	0.0497
166	-1.6559	0.0499	-1.6646	0.0495	-1.6758	0.0496	-1.6628	0.0497	-1.6762	0.0499
167	-1.6566	0.0496	-1.6644	0.0499	-1.6785	0.0499	-1.6668	0.0496	-1.6817	0.0496
168	-1.6572	0.0499	-1.6660	0.0496	-1.6816	0.0500	-1.6699	0.0500	-1.6644	0.0498
169	-1.6589	0.0497	-1.6647	0.0499	-1.6589	0.0499	-1.6735	0.0496	-1.6697	0.0497
170	-1.6579	0.0500	-1.6721	0.0495	-1.6648	0.0498	-1.6756	0.0499	-1.6747	0.0498
171	-1.6561	0.0499	-1.6714	0.0497	-1.6645	0.0497	-1.6849	0.0498	-1.6753	0.0497
172	-1.6593	0.0495	-1.6706	0.0498	-1.6662	0.0495	-1.6625	0.0498	-1.6840	0.0497
173	-1.6619	0.0497	-1.6730	0.0499	-1.6670	0.0498	-1.6639	0.0498	-1.6646	0.0496
174	-1.6603	0.0493	-1.6743	0.0499	-1.6669	0.0499	-1.6678	0.0496	-1.6662	0.0497
175	-1.6575	0.0498	-1.6757	0.0499	-1.6702	0.0500	-1.6697	0.0498	-1.6738	0.0495
176	-1.6613	0.0494	-1.6781	0.0499	-1.6762	0.0498	-1.6775	0.0497	-1.6748	0.0499
177	-1.6727	0.0483	-1.6613	0.0496	-1.6803	0.0499	-1.6823	0.0497	-1.6831	0.0498
178	-1.6630	0.0495	-1.6563	0.0499	-1.6585	0.0498	-1.6599	0.0500	-1.6678	0.0496
179	-1.6572	0.0500	-1.6594	0.0497	-1.6634	0.0497	-1.6616	0.0499	-1.6700	0.0496
180	-1.6586	0.0496	-1.6618	0.0494	-1.6738	0.0489	-1.6698	0.0499	-1.6709	0.0498
181	-1.6569	0.0498	-1.6599	0.0498	-1.6661	0.0499	-1.6678	0.0499	-1.6778	0.0496
182	-1.6593	0.0495	-1.6638	0.0495	-1.6671	0.0496	-1.6721	0.0498	-1.6854	0.0497
183	-1.6583	0.0498	-1.6604	0.0499	-1.6681	0.0497	-1.6772	0.0499	-1.6687	0.0495
184	-1.6592	0.0496	-1.6634	0.0496	-1.6720	0.0498	-1.6627	0.0500	-1.6751	0.0488
185	-1.6576	0.0498	-1.6634	0.0498	-1.6742	0.0497	-1.6594	0.0500	-1.6730	0.0495
186	-1.6645	0.0495	-1.6699	0.0495	-1.6785	0.0498	-1.6682	0.0498	-1.6798	0.0497
187	-1.6570	0.0499	-1.6695	0.0497	-1.6604	0.0497	-1.6663	0.0500	-1.6869	0.0498
188	-1.6603	0.0495	-1.6695	0.0498	-1.6657	0.0498	-1.6708	0.0499	-1.6709	0.0494
189	-1.6573	0.0498	-1.6685	0.0499	-1.6634	0.0499	-1.6759	0.0497	-1.6677	0.0498
190	-1.6621	0.0494	-1.6692	0.0499	-1.6673	0.0496	-1.6816	0.0498	-1.6755	0.0497
191	-1.6603	0.0497	-1.6724	0.0498	-1.6642	0.0498	-1.6675	0.0494	-1.6791	0.0497
192	-1.6571	0.0500	-1.6757	0.0498	-1.6667	0.0499	-1.6649	0.0498	-1.6712	0.0498
193	-1.6618	0.0497	-1.6765	0.0499	-1.6685	0.0499	-1.6718	0.0495	-1.6789	0.0486
194	-1.6569	0.0499	-1.6571	0.0497	-1.6731	0.0498	-1.6705	0.0500	-1.6713	0.0495
195	-1.6616	0.0496	-1.6594	0.0494	-1.6769	0.0499	-1.6746	0.0498	-1.6741	0.0498
196	-1.6582	0.0499	-1.6561	0.0499	-1.6566	0.0499	-1.6782	0.0500	-1.6790	0.0500
197	-1.6621	0.0496	-1.6643	0.0496	-1.6665	0.0491	-1.6630	0.0500	-1.6668	0.0498
198	-1.6582	0.0498	-1.6590	0.0499	-1.6660	0.0495	-1.6631	0.0500	-1.6666	0.0496
199	-1.6603	0.0495	-1.6617	0.0496	-1.6614	0.0498	-1.6680	0.0495	-1.6714	0.0497
200	-1.6597	0.0497	-1.6607	0.0499	-1.6621	0.0500	-1.6692	0.0496	-1.6747	0.0498

CAPITULO 6. CONCLUSIONES

De la comparación de las siete pruebas de no inferioridad para la diferencia de dos proporciones independientes T_1 :Blackwelder; T_2 :Farrington-Manning; T_3 :Böhning-Viwatwongkasem; T_4 :Hauck-Anderson; T_5 : una variante de Farrington-Manning; T_6 : una variante de Böhning-Viwatwongkasem y T_7 : la prueba de razón de verosimilitudes generalizada, en sus versiones exactas algunas de ellas resultaron ser comparables entre si en cuanto al nivel de significancia real y a la potencia. Se encontró que las pruebas T_1 y T_4 y T_3 y T_6 mostraron un comportamiento muy similar en los valores del nivel de significancia y que son excesivamente conservadoras, a tal grado que para los tamaños de muestra pequeños (menores de $n_1=n_2=30$) el nivel de significancia está muy cercano a cero. La prueba de razón de verosimilitudes generalizada (T_7) tiene un comportamiento muy conservador para los valores más pequeños del margen de no inferioridad y aunque es un poco más difícil que las pruebas T_2 y T_5 y requiere de un mayor tiempo de cómputo para el cálculo de sus niveles de significancia, no tiene un comportamiento mejor del nivel de significancia ni de la potencia que T_2 y T_5 .

Las pruebas T_2 y T_5 exactas son iguales en los valores del nivel de significancia, y como T_5 es resultado de una modificación de la prueba T_2 entonces dicha modificación no surtió ningún efecto en tal sentido. Se realizaron comparaciones de las potencias de las pruebas T_2 y T_5 para el nivel nominal de $\alpha=0.05$ para un gran número de tamaños de muestra, y la potencia resultó igual para ambas pruebas comparadas.

La comparación detallada de las pruebas exactas y asintóticas resultó en que la prueba T_2 exacta es la mejor para los tamaños de muestra $5 \leq n_1=n_2 \leq 30$. Para los tamaños de muestra $31 \leq n_1=n_2 \leq 100$ la prueba T_2 exacta también resultó mejor que la asintótica con corrección por continuidad, sin embargo para estos tamaños de muestra la versión asintótica con corrección por continuidad no es muy inferior a la exacta.

La prueba T_2 exacta es mejor que la prueba asintótica porque sus niveles de significancia reales son uniformemente mayores, para tamaños de muestra de $5 \leq n_1=n_2 \leq 100$, es decir es más eficiente en el sentido de que esta por debajo del nivel nominal pero por encima de la prueba asintótica. Debido a la simplicidad de cálculo de la prueba T_2 asintótica, podría preferirse con respecto a la prueba exacta.

Debido a que la versión exacta es mejor tanto para tamaños de muestra pequeños ($5 \leq n \leq 30$) como para tamaños de muestra moderados ($31 \leq n \leq 100$) se elaboró una tabla ($5 \leq n \leq 200$) con las constantes críticas y los niveles de significancia para la prueba T_2 exacta, contribuyendo con ello a facilitar el uso de esta prueba.

Los programas de cómputo escritos en S-PLUS® se pueden modificar fácilmente para obtener el valor aproximado del nivel de significancia real para diseños desbalanceados.

6.1. Recomendaciones

De las pruebas comparadas, para un tamaño de muestra $5 \leq n_1 = n_2 \leq 30$, se recomienda utilizar la prueba T_2 exacta para probar la no inferioridad de un tratamiento nuevo contra un tratamiento estándar.

Para un tamaño de muestra de $31 \leq n \leq 100$ se pueden usar para probar no inferioridad la prueba T_2 exacta o la prueba T_2 asintótica con corrección por continuidad porque ambas resultaron buenas, pero la ventaja de la prueba asintótica es su facilidad para calcularla.

6.2. Investigación futura

Los tópicos con los que puede continuar la presente investigación es:

- Comparación de pruebas asintóticas y exactas para otras medidas de discrepancia entre dos proporciones como por ejemplo el riesgo relativo p_1/p_2 y la razón de

$$\text{mismos } \frac{p_1(1-p_2)}{p_2(1-p_1)}.$$

- Extender la comparación de pruebas exactas y asintóticas para la diferencia de dos proporciones a diseños desbalanceados.
- Comparación de pruebas de equivalencia.

BIBLIOGRAFIA

- Almendra A., F. 2007. Comparación de algunas pruebas estadísticas asintóticas de no inferioridad para contrastar dos proporciones independientes. Tesis Doctoral, Especialidad en Estadística, ISEI, Colegio de Postgraduados, México. 87 p.
- Almendra A., F. 2008. A Study on the Classical Asymptotic Non-inferiority Test for Two Binomial Proportions. Drug Information Journal. Aceptado.
- Almendra A., F. y D. Sotres R. 2008 Comparación de algunas pruebas estadísticas asintóticas de no-inferioridad para dos proporciones independientes. Agrociencia. Aceptado.
- Barnard G., A. 1947. Significance tests for 2x2 tables. *Biometrika* 34:123-138.
- Basu, D. 1977. On the Elimination of Nuisance Parameters. *Journal of the American Statistical Association* 72: 355-366
- Blackwelder, W. C. (1982) “Proving the null hypothesis” in clinical trials. *Controlled Clinical Trials*. 3, 345-353.
- Böhning, D. and C. Vivotwongkasem (2005). Revisiting proportion estimators. *Statistical methods in medical research*. 14, 147-169
- Chan, I. S. F. (1998). Exact tests of equivalence and efficacy with a non zero lower bound for comparative studies. *Statistics in Medicine*. 17, 1403-1413.
- Chan, I. S. F. (1999). Author’s reply on Exact tests of equivalence and efficacy with a non zero lower bound for comparative studies. *Statistics in Medicine*. 18, 1735-1737.

- Chan, I. S. F. (2003). Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods. *Statistical methods in medical research*. 12, 37-58.
- Chen, J., Y. Tsong, and S. Kang. 2000. Test for equivalence or noninferiority between two proportions. *Drug Information Journal* 34: 569-578.
- Ebutt, A.F. and L. Frith. 1998. Practical issues in equivalence trials. *Statistics in Medicine* 17:1691-1701.
- Farrington, C., and G. Manning. 1990. Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Statistics in Medicine* 9: 1447-1454.
- Gao P., and J. Ware. 2008. Assessing non-inferiority: A combination approach. *Statistics in Medicine* 27: 392-406.
- Hauck, W., and S. Anderson. 1986. Comparison of large sample confidence interval methods for the difference of two binomial probabilities. *The American Statistician* 40: 318-322
- Li, Z., and C. Chuang-Stein. (2006). A note on comparing two binomial proportions in confirmatory non inferiority trials. *Drug Information Journal*. 40, 203-208.
- Martin Andrés, A., and I. Herranz Tejedor. 2002. Equivalence testing for binomial random variables: which test to use? *The American Statistician* 56(3): 253-254, Letter to the editor.
- Martin Andrés, A., and I. Herranz Tejedor. 2004a. Asymptotical test on the equivalence, substantial difference and non-inferiority problems with two proportions. *Biometrical Journal* 46: 305-319.

- Martin Andrés, A., and I. Herranz Tejedor. 2004b. Exact unconditional non-classics tests on the difference of two proportions. *Computational Statistics & Data Analysis* 45: 373-388.
- Martin Andrés, A., and A. Silva Mato. 1994. Choosing the optimal unconditioned test for comparing two independent proportions. *Computational Statistics & Data Analysis* 17: 555-574.
- Mehrotra, D. V., I. S. F.Chan, and R. L. Berger. 2003. A cautionary note on exact unconditional inference for the difference between two independent binomial proportions. *Biometrics* 59: 441-450.
- Miettinen, O., and M. Nurminen. 1985. Comparative analysis of two rates. *Statistics in Medicine* 4: 213-226.
- Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. *Introduction to the theory of Statistics*. 3a edition. New York: McGraw-Hill.
- Munk, A., G. Skipka, and B. Stratmann. 2005. Testing general hypotheses under binomial sampling: The two sample case: asymptotic theory and exact procedures. *Computational Statistics and Data Analysis*. 49(3): 723-739.
- Rodary C., C. Com-Nougue, and M. Tournade. 1989. How to establish equivalence between treatments: a one-sided clinical trial in pediatric oncology. *Statistics in Medicine* 8: 593-598.
- Röhmel, J., and U. Mansmann. 1999. Unconditional non asymptotic one sided tests for independent binomial proportions when the interest lies in showing non inferiority and or superiority. *Biometrical Journal* 2: 149-170.

- Röhmel, J. 2005. Problems with existing procedures to calculate exact unconditional p-values for non inferiority/superiority and confidence intervals for two binomials and how to resolve them. *Biometrical Journal* 47: 37-47.
- Skipka, G., A. Munk, and G. Freitag. 2004. Unconditional exact tests for the difference of binomial probabilities - contrasted and compared. *Computational Statistics and Data Analysis* 47: 757-773.
- Tu, D. 1997. A comparative study of some statistical procedures in establishing therapeutic equivalence of non systemic drugs with binary endpoints. *Drug Information Journal* 31: 1291-1300.

Apéndice I. Simbología utilizada

$B(n, p)$	Distribución Binomial con parámetros n y p .
$f(x, n, p) = \binom{n}{x} p^x (1-p)^{n-x}$	Función de probabilidad de la distribución binomial con parámetros n y p
$H_0 : p_1 - p_2 \geq d_0$	Hipótesis nula.
$d = p_1 - p_2$	Diferencia de dos proporciones
d_0	Margen de no inferioridad
EMV	Estimador de máxima verosimilitud
$EMVR$	Estimador de máxima verosimilitud restringido
$\hat{d} = \hat{p}_1 - \hat{p}_2$	Estimador de la diferencia de proporciones
$\hat{\sigma}_1$	Estimador de Blackwelder para $\sqrt{Var(\hat{d})}$
$\hat{\sigma}_2$	Estimador de Farrington-Manning (FM) para $\sqrt{Var(\hat{d})}$
$\hat{\sigma}_3$	Estimador de Böhning-Viwatwongkasen (BV) de $\sqrt{Var(\hat{d})}$
$\hat{\sigma}_4$	Estimador de Hauck-Anderson (HA) para $\sqrt{Var(\hat{d})}$

$\hat{\sigma}_5$	Estimador de FM-HA para $\sqrt{\widehat{Var}(\hat{d})}$
$\hat{\sigma}_6$	Estimador de BV-HA para $\sqrt{\widehat{Var}(\hat{d})}$
$\hat{p}_i = x_i / n_i$	EMV de p_i
\check{p}_1	EMVR de p_1 bajo la hipótesis $p_1 - p_2 \geq d_0$
$\check{p}_2 = \check{p}_1 - d_0$	EMVR de p_2 bajo la hipótesis $p_1 - p_2 \geq d_0$
$\tilde{p}_{i,b} = \frac{x_i + b}{n_i + 2b}$	Estimador de BV de p_i

Apéndice II. Programas de Cómputo

Se presentan los programas para el cálculo del nivel de significancia para las pruebas T_1 , T_2 , T_3 y T_7 . Para las pruebas restantes, T_4 , T_5 y T_6 , se omitió incluir el programa porque son una modificación de las primeras.

Prueba T_1

```
## En la Parte 1 se calculan y se ordenan los valores de las estadísticas de
prueba para el tamaño de muestra indicado con n##
##PARTE 1 PARTE 1 PARTE 1#####
n_82
cont_0
d0 0.1
iter_(1-d0)*900
pl_d0
numpares (n+1)*(n+1)
tes_matrix(0, n+1,n+1)
cxy_0
contot_matrix(0, numpares)
infor_matrix(NA,numpares,4)
contl_0
# inicia ciclo para x = 0,1,2...n
for (x in 0:n)
{
# inicia ciclo para y= 0,1,2...n
for (y in 0:n)
{
#*****NUEVA SD #Cálculo del estimador de la desviación estándar##
if(x==0 & y==0)
{xx_0.01
YY_0.01}
else if (x==n & y==0)
{xx_x-.01
YY_0.01}
else if (x==0 & y==n)
{xx_.01
YY_y-.01}
else if (x==n & y==n)
{xx_x-.01
YY_y-.01}
else
{xx_x
YY_y}
cat(xx, " , ",yy, fill=T)
```

```

sd_sqrt(((xx)/n)*(1-((xx)/n))/(n)) + (((yy)/n)*(1-((yy)/n))/(n))
txy_(((x)/n)-((y)/n)-d0)/sd

cat(x, " , ", y, " , t=", txy, fill=T)
cont1_cont1+1
infor[cont1,2]_x
infor[cont1,3]_y
infor[cont1,1]_txy
tes[x+1,y+1]_txy

txy_(((x)/n)-((y)/n)-d0+c1)/sd

#cat(x, " , ", y, " , t=", txy, fill=T)
cont1_cont1+1
infor[cont1,2]_x
infor[cont1,3]_y
infor[cont1,1]_txy
tes[x+1,y+1]_txy
}

}

infnw_sort.col(infor, "<ALL>", ascending=T)
##OK_OK OK OK (Fin de parte 1)
#####

##En la parte 2 se realiza la maximización sobre los parámetros y se obtiene el
nivel de significancia exacto##

##PARTE 2 PARTE 2 PARTE 2 PARTE 2 PARTE 2 PARTE 2###

mrbig_matrix(0, iter, 1)
conteo_0
repeat
{
cat("procesando ", conteo, " de ", numpares, fill=T)
conteo_conteo+1
cat (conteo, fill=T)

g_infnw[conteo,2]
h_infnw[conteo,3]

# inicia ciclo para iteraciones, i= 0,1,2...iter
for (i in 0:iter)
{
p1_d0+i*(0.001)
p2_p1-d0
mrbig[i]_mrbig[i]+ dbinom(g, n, p1) * dbinom(h, n, p2)
}

maxbig_max(mrbig)
infnw[conteo,4]_maxbig
if(maxbig > 0.1) break
}
#####

## En Infnew se encuentra el nivel de significancia calculado##
Infnew

xx_infnw[,2][which(infnw[,4] != "NA")]
yy_infnw[,3][which(infnw[,4] != "NA")]
plot(xx,yy)

```

#En CONVEXIDAD Y SIMETRIA se verifica numéricamente la condición de convexidad de
Bernard y de simetría en la misma cola##

```
#####CONVEXIDAD Y SIMETRIA#####
comb_matrix(0, n+1,n+1)
cont_0
repeat
{
  cont_cont+1
  x_infnew[cont,2]
  Y_infnew[cont,3]
  comb[x+1,y+1]_1
  pval_infnew[cont,4]
  if (pval > 0.05) break
}

#a
evalx_1
for (p in 1:n+1)
{
  for (q in 2:n+1)
  if (comb[p,q]==1 & comb[p-1,q]==1) evalx_evalx*1 else if (comb[p,q]==1 &
  comb[p-1,q]==0) evalx_evalx*0
}
evalx

#b
evaly_1
for (p in 1:n+1)
{
  for (q in 1:n)
  {if (comb[p,q]==1 & comb[p,q+1]==1) evaly_evaly*1 else if (comb[p,q]==1 &
  comb[p-1,q]==0) evaly_evaly*0
}
}
evaly

if (evalx==1 & evaly==1) cat("Si Existe Convexidad")
else cat (" No Existe Convexidad")
```

Prueba T2

En la Parte 1 se calculan y se ordenan los valores de las estadísticas de prueba para el tamaño de muestra indicado con n##
##PARTE 1 PARTE 1 PARTE 1####

```
n_70
#options(object.size=5e6)
cont_0
d0_0_1
iter_(1-d0)*900
pl_d0
numpares_(n+1)*(n+1)
tes_matrix(0, n+1,n+1)
cxy_0
contot_matrix(0, numpares)
infor_matrix(NA,numpares,4)
contl_0

# inicia ciclo para x = 0,1,2...n
for (x in 0:n)
{
# inicia ciclo para y= 0,1,2...n
  for (y in 0:n)
  {
#*****NUEVA SD
r_ -(x+y+n*(1+2*d0)+n*(1+d0))/(n+n)
s_ (y+x*(1+2*d0)+ d0*(n+n*(1+d0)))/(n+n)
t_ -x*d0*(1+d0)/(n+n)
pgorro1_ 2/3* sqrt((r^2)-(3*s)) *
cos( 1/3* acos
( - ( ( 2*r^3/27) - r*s/3+t) / (2*(sqrt(r^2-3*s)/3)^3) ) )
+4/3*pi )-(r/3)
pgorro2_pgorro1-d0
sd_sqrt((pgorro1*(1-pgorro1))/n+(pgorro2*(1-pgorro2))/n)

##### En esta parte se calcula en valor de las estadísticas###
txy_(((x)/n)-((y)/n)-d0)/sd

#cat(x, " , ",y, " , t=",txy,fill=T)
contl_contl+1
infor[contl,2]_x
infor[contl,3]_y
infor[contl,1]_txy
tes[x+1,y+1]_txy
}
}

infnew_sort.col(infor, "<ALL>", ascending=T)
##OK OK OK OK (Fin de parte 1)

#####

##PARTE 2 PARTE 2 PARTE 2 PARTE 2 PARTE 2####

mrbig_matrix(0,iter,1)
conteo_0
repeat
{
cat("procesando ", conteo, " de ",numpares, fill=T)
conteo_conteo+1
cat (conteo, fill=T)

g_infnw[conteo,2]
```

```

h_infnew[conteo,3]
# inicia ciclo para iteraciones, i= 0,1,2...iter
for (i in 0:iter)
{
  p1_d0+i*(0.001)
  p2_p1-d0
  mrbig[i]_mrbig[i]+ dbinom(g, n, p1) * dbinom(h, n, p2)
}
maxbig_max(mrbig)
infnew[conteo,4]_maxbig
if(maxbig > 0.1) break
}
#####
infnew

xx_infnew[,2][which(infnew[,4] != "NA")]
yy_infnew[,3][which(infnew[,4] != "NA")]
plot(xx,yy)

#En CONVEXIDAD Y SIMETRIA se verifica numéricamente la condición de convexidad de
Barnard y de simetría en la misma cola##

#####CONVEXIDAD Y SIMETRIA#####
comb_matrix(0, n+1,n+1)
cont_0
repeat
{
  cont_cont+1
  x_infnew[cont,2]
  y_infnew[cont,3]
  comb[x+1,y+1]_1
  pval_infnew[cont,4]
  if(pval > 0.05) break
}

#a
evalx_1
for (p in 1:n+1)
{
  for (q in 2:n+1)
  if (comb[p,q]==1 & comb[p-1,q]==1) evalx_evalx*1 else if (comb[p,q]==1 &
  comb[p-1,q]==0) evalx_evalx*0
}
evalx
#b
evaly_1
for (p in 1:n+1)
{
  for (q in 1:n)
  {if (comb[p,q]==1 & comb[p,q+1]==1) evaly_evaly*1 else if (comb[p,q]==1 &
  comb[p-1,q]==0) evaly_evaly*0
}
}
evaly
if (evalx==1 & evaly==1) cat("Si Existe Convexidad")
else cat (" No Existe Convexidad")

```

Prueba T3

```
## En la Parte 1 se calculan y se ordenan los valores de las estadísticas de
prueba para el tamaño de muestra indicado con n##
##PARTE 1 PARTE 1 PARTE 1#####
n_80
cat("procesando n=",n,fill=T)
#options(object.size=5e6)
cont_0
d0 0.25
iter_(1-d0)*400
p1_d0
numpares (n+1)*(n+1)
tes_matrix(0, n+1,n+1)
cxy_0
contot_matrix(0, numpares)
infor_matrix(NA,numpares,4)
contl_0

# inicia ciclo para x = 0,1,2...n
for (x in 0:n)
{
# inicia ciclo para y= 0,1,2...n
for (y in 0:n)
{
#
#cat(xx, " , ",yy, fill=T)

sd_sqrt(((x+1)/(n+2))*(1-((x+1)/(n+2)))/(n)) + ((y+1)/(n+2))*(1-
((y+1)/(n+2)))/(n))
txy_(((x)/n)-((y)/n)-d0)/sd
#cat(x, " , ",y, " , t=",txy,fill=T)
contl_contl+1
infor[contl,2]_x
infor[contl,3]_y
infor[contl,1]_txy
tes[x+1,y+1]_txy

}
}
infnw_sort.col(infor, "<ALL>" ,ascending=T)

#####
##OK OK OK OK (Fin de parte 1)

#####
##PARTE 2 PARTE 2 PARTE 2 PARTE 2 PARTE 2####
mrbig_matrix(0,iter,1)
conteo_0
repeat
{
#cat("procesando ", conteo, " de ",numpares, fill=T)
conteo_conteo+1
#cat (conteo, fill=T)

g_infnw[conteo,2]
h_infnw[conteo,3]

# inicia ciclo para iteraciones, i= 0,1,2...iter
for (i in 0:iter)
{
```

```

p1_d0+i*(0.001)
p2_p1-d0
mrbig[i]_mrbig[i]+ dbinom(g, n, p1) * dbinom(h, n, p2)
}

maxbig_max(mrbig)
infnew[conteo,4]_maxbig
if(maxbig > 0.1) break
}
#####
infnew

#En CONVEXIDAD Y SIMETRIA se verifica numéricamente la condición de convexidad de
Barnard y de simetría en la misma cola##

#####CONVEXIDAD Y SIMETRIA#####
comb_matrix(0, n+1,n+1)
repeat
{
  cont_cont+1
  x_infnew[cont,2]
  Y_infnew[cont,3]
  comb[x+1,y+1]_1
  pval_infnew[cont,4]
  if(pval > 0.05) break
}

#a
evalx_1
for (p in 1:n+1)
{
  for (q in 2:n+1)
    if (comb[p,q]==1 & comb[p-1,q]==1) evalx_evalx*1 else if (comb[p,q]==1 &
      comb[p-1,q]==0) evalx_evalx*0
  }
  evalx
  #b
  evaly_1
  for (p in 1:n+1)
  {
    for (q in 1:n)
      {if (comb[p,q]==1 & comb[p,q+1]==1) evaly_evaly*1 else if (comb[p,q]==1 &
        comb[p-1,q]==0) evaly_evaly*0
      }
    }
  evaly
  if (evalx==1 & evaly==1) cat("Si Existe Convexidad")
  else cat (" No Existe Convexidad")
}

```


Prueba T7

En la Parte 1 se calculan y se ordenan los valores de las estadísticas de prueba para el tamaño de muestra indicado con n##
##PARTE 1 PARTE 1 PARTE 1####

```
n_80
cont_0
d0_0.10
iter_(1-d0)*400
p1_d0
numpares_(n+1)*(n+1)
tes_matrix(0, n+1,n+1)
cxy_0
contot_matrix(0, numpares)
infor_matrix(NA,numpares,4)
contl_0

# inicia ciclo para x = 0,1,2...n
for (x in 0:n)
{
# inicia ciclo para y= 0,1,2...n
  for (y in 0:n)
  {
#*****NUEVA SD
pg1_x/n
pg2_y/n
ld_dbinom(x,n,pg1)*dbinom(y,n,pg2)

r_ -(x+y+n*(1+2*d0)+n*(1+d0))/(n+n)
s_ (y+x*(1+2*d0)+ d0*(n+n*(1+d0)))/(n+n)
t_ -x*d0*(1+d0)/(n+n)

pgorro1_2/3*sqrt((r^2)-(3*s)) *
cos(1/3*acos
(-((2*r^3/27) - r*s/3+t) / (2*(sqrt(r^2-3*s)/3)^3)
+4/3*pi)-(r/3)

pgorro2_pgorro1-d0

if (pgorro1<0) pgorro1_0
if (pgorro1>1) pgorro1_1
if (pgorro2<0) pgorro2_0
if (pgorro2>1) pgorro2_1

lu_dbinom(x,n,pgorro1)*dbinom(y,n,pgorro2)

#*****
txy7_(lu/ld)
txy_(txy7)
if (pg1-pg2<d0)
  contl_contl+1
  infor[contl,2]_x
  infor[contl,3]_y
  infor[contl,1]_txy
  tes[x+1,y+1]_txy
}
}

infnw_sort.col(infor, "<ALL>" , ascending=T)

#####
##OK OK OK (Fin de parte 1)
#####
```

```

##PARTE 2 PARTE 2 PARTE 2 PARTE 2 PARTE 2###
mrbig_matrix(0,iter,1)
conteo_0
repeat
{
cat("procesando ", conteo, " de ",numpares, fill=T)
conteo_conteo+1
cat (conteo, fill=T)

g_infnew[conteo,2]
h_infnew[conteo,3]
# inicia ciclo para iteraciones, i= 0,1,2...iter
for (i in 0:iter)
{
p1_d0+i*(0.001)
p2_p1-d0
mrbig[i]_mrbig[i]+ dbinom(g, n, p1) * dbinom(h, n, p2)
}
maxbig_max(mrbig)
infnew[Conteo,4]_maxbig
if(maxbig > 0.1) break
}
#####
Infnew

#####**CONVEXIDAD Y SIMETRIA*****
comb_matrix(0, n+1,n+1)
cont_0
repeat
{
cont_cont+1
x_infnew[cont,2]
Y_infnew[cont,3]
comb[x+1,y+1]_1
pval_infnew[cont,4]
if(pval > 0.05) break
}

#a
evalx_1
for (p in 1:n+1)
{
for (q in 2:n+1)

if (comb[p,q]==1 & comb[p-1,q]==1) evalx_evalx*1 else if (comb[p,q]==1 &
comb[p-1,q]==0) evalx_evalx*0
}
evalx
#b
evaly_1
for (p in 1:n+1)
{
for (q in 1:n)

{if (comb[p,q]==1 & comb[p,q+1]==1) evaly_evaly*1 else if (comb[p,q]==1 &
comb[p-1,q]==0) evaly_evaly*0
}
}
evaly
if (evalx==1 & evaly==1) cat("Si Existe Convexidad")
else cat (" No Existe Convexidad")

```

Cálculo de Potencia.

En este programa se calcula la potencia para un tamaño de muestra sobre todo el espacio de parámetros.

```
tabla1_read.table("salidaT2d020n90", sep =",", skip=1)
tabla1[3]
n_tabla1[1,3]
filas_length(tabla1[,1])
matbeta_matrix(nrow=filas, ncol=1)
d0_0.20
vec_matrix(NA,21,21)
for_(k in 0:20)
{
  p1_0.05*k
  jmax_20*(1-max(0,p1-d0))
  #cat("p1= ",p1,fill=T)
  {
    for_(j in 0:jmax)
    {
      p2a_max(0,p1-d0)+j*.05
      if (p2a>=1) p2_p2a-.0000000000001 else (p2_p2a)
      cat ("p1= ",p1,"", p2=" ",p2,fill=T)
      beta_0
      betat_0
      i_1
      #cat("procesando ",j, fill=T)
      #cat("p2 =",p2, fill=T)
      repeat
      {
        x1_tabla1[i,2]
        x2_tabla1[i,3]
        betat_dbinom(x1,n,p1)*dbinom(x2,n,p2)
        beta_beta+betat
        # cat("procesando ",i,"beta=", beta," pval=", tabla1[i,4], fill=T,)
      }
      i_i+1
      if(tabla1[i,4] > 0.05) break
    }
    vec[j+1,k+1]_beta
  }
}
vec
write.table(vec, file="NEWpotT2d020a05n90", append=T)
```

Comparación de Potencia. Con este programa se realizan comparaciones por pares de la potencia.

```
tabla1_read.table("potT2d010a05n41", sep = ",")
tabla2_read.table("potT5d010a05n41", sep = ",")
n_41
dif_abs(tabla1-tabla2)
d1_c(dif)
d2_na.exclude(d1)
tam_length(d2)
count_0
for(x in 1:tam)
{
    if (d2[x]<0.001) count_count+1
}
percent_count/tam
percent
datos_c(n,percent)
write(datos, file="comparaT2T5d10", ncol=2, append=T)
```