



**BENEMÉRITA  
UNIVERSIDAD  
AUTÓNOMA DE PUEBLA**

*FACULTAD DE CIENCIAS FÍSICO  
MATEMÁTICAS*

*Desempeño de intervalos de confianza para  
una proporción y criterios para su  
aplicación.*

*T E S I S*

*para obtener el título de:*

**LICENCIADO EN  
MATEMÁTICAS**

*presenta:*

*MARCOS MORALES CORTÉS*

*directores de tesis:*

*DR. FÉLIX ALMENDRA ARAO  
DRA. HORTENSIA J. REYES CERVANTES*

*PUEBLA, PUE.*

*FEBRERO 2018*





# *Agradecimientos*

*Mis sueños son mentiras, que algún día dejarán de serlo.*  
*N.A.C.H.*

En primera instancia agradezco a los directores de esta tesis, Dr. Félix Almendra Arao y Dra. Hortensia J. Reyes Cervantes, por su dedicación, apoyo, paciencia y sabios consejos, que me brindaron a lo largo de este trabajo, pero sobre todo, gracias por tolerar mis errores.

Le agradezco a mis sinodales Dr. Hugo A. Cruz Suárez, Dr. Bulmaro Juárez Hernández y M. C. José J. Castro Alva, quienes aún con sus múltiples ocupaciones, aceptaron revisar la presente tesis y me ayudaron a mejorarla sustancialmente.

A mis padres, Vicente Morales Rosas y Modesta Cortés Sánchez por apoyarme en todo momento, por sus sabios consejos, palabras de aliento y por su gran amor, en ocasiones no merecido. Mejores padres, la vida no pudo haberme dado.

A mis hermanos, gracias por su apoyo, cariño, por estar en los momentos más importantes de mi vida, pero sobre todo, gracias por estar conmigo en los malos momentos. Éste logro también es de ustedes.

A mi institución, gracias por haberme permitido formarme en ella, gracias a todos los profesores por su gran aporte en mi educación, que el día de hoy se verá reflejado en la culminación de mis estudios.

A Mariela Meza Ríos, gracias por sus consejos, por no dejarme claudicar y por ayudarme a llegar hasta aquí. Eres lo que siempre estuve buscando.

Agradezco el apoyo de la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP) por el apoyo en la impresión de esta tesis con el proyecto titulado, Pruebas de no inferioridad suponiendo observaciones discretas y otras aplicaciones II.

**¡Muchas gracias a todos!**



# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Revisión de la literatura . . . . .	3
1.2. Objetivos . . . . .	6
<b>2. Preliminares</b>	<b>7</b>
2.1. Distribución binomial . . . . .	7
2.2. Algunos tipos de convergencia . . . . .	10
2.3. Estimación puntual . . . . .	20
2.4. Pruebas de hipótesis. . . . .	23
2.4.1. Métodos de evaluación de pruebas. . . . .	25
2.4.2. Pruebas para muestras grandes. . . . .	26
2.5. Intervalos de confianza . . . . .	28
2.5.1. Introducción a intervalos de confianza. . . . .	29
2.5.2. Métodos para encontrar estimadores por intervalo. . . . .	30
<b>3. Intervalos de confianza para una proporción</b>	<b>35</b>
3.1. El intervalo de Wald . . . . .	36
3.2. El intervalo de Wilson . . . . .	37
3.3. El intervalo de Agresti-Coull . . . . .	38
3.4. El intervalo arcoseno . . . . .	39
3.5. El intervalo de Clopper-Pearson . . . . .	41

<b>4. Comparación de los intervalos de confianza</b>	<b>43</b>
4.1. Probabilidad de cobertura del intervalo de Wald. . . . .	43
4.1.1. Desempeño del intervalo de Wald . . . . .	45
4.2. Factores de comparación de un intervalo de confianza . . . . .	53
4.3. Comportamiento bajo los criterios más usuales. . . . .	57
4.4. Comportamiento del intervalo de Wald en términos de los valo- res de los parámetros ( $n$ y $p$ ). . . . .	61
4.5. Desempeño de los intervalos alternativos . . . . .	64
4.5.1. Intervalo de Wilson . . . . .	64
4.5.2. Intervalo de Agresti-Coull . . . . .	72
4.5.3. Intervalo arcoseno . . . . .	79
4.6. Comparación de los intervalos . . . . .	85
<b>5. Conclusiones</b>	<b>101</b>
<b>A. Programas para la comparación de intervalos de confianza</b>	<b>103</b>

# Capítulo 1

## Introducción

La distribución binomial es muy conocida en estadística aplicada, está basada en la idea de un ensayo Bernoulli. Un ensayo Bernoulli es un experimento con solamente dos posibles resultados, llamados éxito y fracaso. Si se realizan  $n$  ensayos Bernoulli idénticos e independientes con una probabilidad de éxito conocida  $p$ , la variable aleatoria que cuenta el número de éxitos en los  $n$  ensayos tiene una distribución binomial. El uso de la distribución binomial es frecuente en diversas áreas como estudios de mercado, muestreo, industria, química, control de calidad, ingeniería, ensayos clínicos, en la industria farmacéutica, industrial y espacial entre otros. En la mayoría de los casos el parámetro  $p$  es desconocido y el investigador estará interesado en estimarlo ya sea por medio de un estimador puntual o por un estimador por intervalo (intervalo de confianza), en este trabajo se tratará el segundo caso.

En la presente tesis se revisa un problema importante en la práctica estadística, la estimación por intervalo de la probabilidad de éxito  $p$  en una distribución binomial. Anwer Khurshid en ([16]) presenta una amplia y detallada revisión de la literatura sobre el cálculo de intervalos de confianza de las distribuciones binomial y Poisson con sus variantes, que estuvo disponible hasta el año 2010. Sin embargo, el problema de la determinación de interva-

los de confianza para una proporción sigue abierto y por ello continúa bajo investigación.

En la mayoría de libros elementales de estadística, se presenta un intervalo de confianza que ha adquirido aceptación casi universal en la práctica. Dicho intervalo es conocido como el intervalo de Wald, ya que proviene de la prueba de Wald para muestras grandes para el caso binomial.

Sin embargo diversos autores entre los que destacan Agresti y Coull en [1], Agresti y Caffo en ([2]) y Brown et al. en ([8]) han demostrado que este intervalo tiene serios problemas sobre todo cuando  $p$  está cerca de los extremos del intervalo  $(0,1)$  o cuando  $n$  es pequeño. Incluso se ha observado también que la probabilidad de cobertura del intervalo de Wald puede estar muy por debajo del coeficiente de confianza incluso si  $p$  está alejado de los extremos del intervalo  $(0,1)$  y también cuando  $n$  es grande. Debido al comportamiento errático de las probabilidades de cobertura del intervalo de Wald, en diversos textos al presentarlo le incluyen una condición o sugerencia para su uso, ésto con la finalidad de mejorar su desempeño. En Brown et al. ([8]) los autores se dieron a la tarea de analizar 11 textos populares de estadística, obteniendo con ello una lista de condiciones sobre el uso del intervalo de Wald, tales condiciones posteriormente serán mencionadas. La obtención de un intervalo de confianza para un parámetro binomial es uno de los problemas estadísticos más comunes y básicos, se han realizado una gran cantidad de propuestas en más de 80 años desde el desarrollo original de los intervalos de confianza, que continúan incluso en el siglo XXI. En este trabajo se analizará el desempeño de algunas propuestas. Sin embargo, de acuerdo con Schilling y Doi en ([25]) después de todo éste tiempo y estudio, aún no ha surgido un intervalo de confianza óptimo.

En el capítulo 2 se hace un resumen de los conceptos básicos de la estimación puntual y de la estimación por intervalos, relacionados con observaciones Bernoulli que brindan las bases para la comprensión del presente trabajo.

En el capítulo 3 se presenta el intervalo de confianza más usado en estadística, es decir, el intervalo de Wald, después son presentados otros intervalos que destacan en la literatura como soluciones alternativas al intervalo de Wald, en todos los casos se realiza un breve resumen sobre su origen.

En el capítulo 4 se muestran las principales inconsistencias del intervalo de Wald, después se muestra que el comportamiento errático de su probabilidad de cobertura es más complejo de lo que se piensa, y que éste no mejora lo suficiente aún cuando se aplican las sugerencias para su uso ni cuando  $n$  es muy grande. El mal comportamiento del intervalo de Wald se presenta tanto para  $n$  fija con  $p$  variable como para  $p$  fija y  $n$  variable. Se analizará su comportamiento variando ambos parámetros haciendo uso de las cantidades obtenidas de la evaluación de un intervalo, el mismo análisis se realizará para comparar otros intervalos que son sugeridos en la literatura.

En el capítulo 5, se presentan las conclusiones sobre este análisis y al mismo tiempo se dan recomendaciones sobre la elección de un intervalo específico y de algún criterio para el uso en la práctica para diferentes valores de  $p$  y de  $n$ .

Para analizar el comportamiento de los intervalos de confianza bajo estudio en el presente trabajo, se escribieron programas en el lenguaje de programación R. Tales programas se encuentran en el apéndice A.

## 1.1. Revisión de la literatura

Agresti A. y Coull B. en ([1]) muestran que los intervalos exactos además de tener probabilidades de cobertura que son mayores al nivel de confianza nominal  $1 - \alpha$ , es decir, de ser conservadores, en ocasiones tales probabilidades de cobertura pueden ser bastante mayores a ese nivel de confianza. También muestran que los intervalos de Agresti-Coull y de Wilson en ocasiones pueden tener probabilidades de cobertura inferiores al nivel de confianza nominal  $1 - \alpha$ ,

pero la probabilidad de cobertura es cercana a ese nivel. De ésta manera, Agresti A. y Coull B. en ([1]) concluyen al igual que Santner en ([24]), que para la mayoría de las aplicaciones se deben preferir los intervalos de Wilson y de Agresti-Coull. También recomiendan que los profesores que imparten cursos básicos presenten el intervalo de Wilson en lugar del intervalo de Wald.

Las conclusiones obtenidas por Brown et al. en ([8]) son que el intervalo de Wilson o el intervalo de Jeffreys (intervalo bayesiano) sea usado para  $n$  pequeños ( $n \leq 40$ ), ya que estos dos intervalos son comparables tanto en su error de cobertura absoluto como en su longitud y sugieren que cualquiera de los dos podría ser usado. En el caso de  $n$  más grandes, los intervalos de Wilson, de Jeffreys y de Agresti-Coull son todos comparables y recomiendan para  $n \geq 40$  al intervalo de Agresti-Coull por ser el más simple de presentar.

Cepeda et al. en ([11]) y Reyes-Cervantes et al. en ([21]) concuerdan con las recomendaciones realizadas por Brown et al. en ([8]), sin embargo Cepeda et al en ([11]) afirma que si lo que se desea es que en promedio se tenga una probabilidad de cobertura de  $100(1-\alpha)\%$ , entonces la alternativa más recomendada está dada por el intervalo de Wilson, sin embargo también añaden que este procedimiento es muy complicado de presentar en niveles elementales de estadística.

En [8] Brown et al. también muestran, que el comportamiento de los intervalos de Wilson y de Jeffreys puede mejorarse para  $p$  cerca de 0 o de 1 modificando los puntos finales de tales intervalos. Posteriormente, Agresti y Coull en sus comentarios al artículo ([8]) de Brown et al. mencionan que una comparación más justa de las longitudes esperadas debería realizarse tomando en cuenta las versiones modificadas de los intervalos de Wilson y de Jeffreys; Agresti y Coull realizan un análisis tomando en cuenta las modificaciones mencionadas y concluyen que, en general, las recomendaciones de Brown et al. sobre el método de elección son razonables. También añaden que su preferencia es

utilizar el intervalo de Wilson para la práctica estadística y el intervalo de Agresti-Coull para la enseñanza en cursos de estadística elemental.

Pires y Amado en ([20]) concluyen que cuando el conservadurismo es obligatorio, es decir, cuando se exige que la probabilidad de cobertura debe ser mayor o igual al nivel de confianza nominal, se debe elegir el intervalo Clopper-Pearson, o, casi de manera equivalente, el intervalo arcoseno con la corrección de Anscombe. Si el conservadurismo no es obligatorio, entonces la recomendación es elegir, en el caso del 95 % de confianza, el método arcoseno con corrección por continuidad o el método de Agresti-Coull y para una confianza del 99 %, el método recomendado es el de Agresti-Coull. El intervalo de Wilson con corrección por continuidad sigue siendo una opción adecuada, sin embargo puede ser demasiado amplio si  $p$  está cerca de 0 o de 1.

En el reporte realizado por Reyes et al. en ([21]) se concluye que los criterios tienen buen desempeño para tamaños de muestra superiores a  $n = 1000$ , teniendo un desempeño similar, destacando la condición de que el intervalo de Wald sea usado solo si  $n \geq 50$  y  $0.2 \leq p \leq 0.8$ .

Måns Thulin en ([18]) muestra que el costo de usar un intervalo exacto (específicamente el intervalo de Clopper-Pearson) en lugar de un intervalo aproximado (Wald, Wilson y Agresti-Coull) está dado en términos de su longitud esperada, siendo ésta más grande para los intervalos exactos. Al mismo tiempo proporciona expresiones de cuánto más grande es la longitud esperada de los intervalos exactos y de cuánto aumenta el tamaño de la muestra para alcanzar una longitud fija esperada.

En ([25]), Schilling y Doi hacen mención de que después de todo este tiempo y estudio, la obtención de un intervalo de confianza óptimo para el parámetro  $p$  de la distribución binomial continúa sin resolverse y que una de las razones es que se han usado dos estándares distintos, uno requiere que el intervalo de confianza cumpla que el ínfimo de las probabilidades de cobertura sea mayor

o igual que el nivel de confianza  $(1 - \alpha)$  y el otro permite que eso se cumpla sólo aproximadamente. Al mismo tiempo presentan el nuevo método LCO (Longitud/Cobertura Optima) que es óptimo con respecto a la longitud y la cobertura cuando el procedimiento de confianza cumpla que el ínfimo de las probabilidades de cobertura sea mayor o igual que el nivel de confianza, y proporcionan una versión aproximada que supera los procedimientos existentes para el criterio de cobertura aproximado. En la cuestión de elegir entre intervalos exactos y aproximados, Schilling y Doi sugieren al menos para el caso binomial de una muestra, que el usuario aclare qué tipo de intervalo se está utilizando, dada su diferencia sustancial en el rendimiento (cobertura y longitud).

## 1.2. Objetivos

1. Revisión bibliográfica de estudios que se han realizado sobre los intervalos de confianza para una proporción.
2. Escribir programas en el ambiente de programación R con la finalidad de calcular las respectivas probabilidades de cobertura y longitudes de los intervalos de confianza para una proporción.
3. Comparar diversos intervalos de confianza para una proporción así como algunos criterios para su aplicación.
4. Establecer recomendaciones acerca del uso de los intervalos de confianza para una proporción así como de los criterios para su utilización.

# Capítulo 2

## Preliminares

Un problema estadístico muy común es la estimación de los parámetros que ayudan a caracterizar una variable aleatoria (v.a.). Un intervalo de confianza nos permite hacer inferencias sobre qué valores se pueden esperar para un parámetro. Los intervalos de confianza dependen de la muestra aleatoria, del tamaño muestral y del nivel de confianza seleccionado. A continuación se presentan algunos conceptos básicos relativos a la distribución binomial, prueba de hipótesis e intervalos de confianza (ver por ejemplo ([10])) así como otros conceptos.

### 2.1. Distribución binomial

En la mayoría de las disciplinas del conocimiento cuando se trabaja con individuos (objetos, plantas, animales o personas) es frecuente realizar experimentos en los cuales se tienen dos posibles resultados mutuamente excluyentes, llamados éxito (**E**) y fracaso (**F**) en donde  $P(\mathbf{E})=p$  y la  $P(\mathbf{F})=1 - p$  con  $p \in [0, 1]$ , a tales experimentos se les conoce como *experimentos Bernoulli* ( $B(p)$ ), si se llevan a cabo  $n$  repeticiones independientes de dicho experimento de tal forma que en cada una de las repeticiones  $P(\mathbf{E}) = p$  y  $P(\mathbf{F}) = 1 - p$ ,

la variable aleatoria  $X$  que cuenta el número de éxitos en los  $n$  ensayos es una variable aleatoria binomial ( $X \sim B(n, p)$ ) con función masa de probabilidad,

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \dots, n; \\ 0, & \text{de otra forma,} \end{cases} \quad 0 \leq p \leq 1, \quad (2.1)$$

y función de distribución acumulativa,

$$F_X(y) = \sum_{x=0}^y \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 \leq y \leq n. \quad (2.2)$$

**Teorema 2.1.** Si  $X \sim B(n, p)$  entonces (a)  $E(X) = np$  y (b)  $Var(X) = np(1-p)$ .

Otra variable aleatoria de gran importancia y que posteriormente será usada es una variable aleatoria normalmente distribuida, definida como.

**Definición 2.1.** Se dice que la variable aleatoria continua  $X$  tiene una distribución normal si su función de densidad está dada por la siguiente expresión

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty$$

donde  $\mu \in \mathbb{R}$  y  $\sigma > 0$  son los parámetros de la distribución y se denota como  $X \sim N(\mu, \sigma^2)$ .

Las áreas bajo la función de densidad normal correspondientes a  $P(a \leq X \leq b)$  requieren la evaluación de una integral, es decir,

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} dx.$$

**Definición 2.2** (Función generadora de momentos). Sea  $X$  una variable aleatoria con función de distribución acumulativa  $F_X$ . La función generadora de momentos de  $X$  denotada por  $M_X(t)$ , es

$$M_X(t) = E[e^{tX}],$$

con la condición de que la esperanza exista para  $t$  en alguna vecindad del 0. Esto es, existe  $h > 0$  tal que, para todo  $t$  con,  $-h < t < h$ ,  $E[e^{tX}]$  existe. Más explícitamente se puede escribir a la función generadora de momentos de  $X$  como

$$M_X(t) = E[e^{tX}] = \begin{cases} \sum e^{tx}p(x), & \text{si } X \text{ es discreta;} \\ \int_{-\infty}^{\infty} e^{tx}f(x)dx, & \text{si } X \text{ es continua.} \end{cases} \quad (2.3)$$

**Teorema 2.2.** Si  $X$  tiene función generadora de momentos  $M_X(t)$ , entonces

$$E[X^n] = M_X^{(n)}(0),$$

donde  $M_X^{(n)}(0) = \frac{d^n}{dt^n}M_X(t) |_{t=0}$ .

*Demostración.* La demostración se realizará para el caso cuando  $X$  es continua,

$$\frac{d}{dt}M_X(t) = \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx,$$

como  $e^{tx}f_X(x)$  es diferenciable con respecto a  $t$ , por la Regla de Leibnit'z ([10])

la expresión anterior es equivalente a,

$$\begin{aligned} & \int_{-\infty}^{\infty} \left( \frac{d}{dt} e^{tx} \right) f_X(x) dx \\ &= \int_{-\infty}^{\infty} (x e^{tx}) f_X(x) dx \\ &= E[X e^{tX}]. \end{aligned}$$

Así,

$$\frac{d}{dt}M_X(t) |_{t=0} = E[X e^{tX}] |_{t=0} = E[X].$$

Procediendo de una manera similar, se establece que

$$\frac{d^n}{dt^n}M_X(t) |_{t=0} = E[X^n e^{tX}] |_{t=0} = E[X^n]. \quad \square$$

**Teorema 2.3.** La función generadora de momentos de una variable aleatoria normalmente distribuida  $X$  está dada por

$$M_X(t) = e^{\left(\mu t + \frac{t^2 \sigma^2}{2}\right)}, \quad -\infty < t < \infty.$$

**Teorema 2.4.** Si  $X$  es una variable aleatoria normalmente distribuida con parámetros  $\mu$  y  $\sigma$ , entonces  $E(X) = \mu$  y  $Var(X) = \sigma^2$ .

Se dice que una variable aleatoria  $X$  tiene una distribución normal estándar si tiene una distribución normal con parámetros  $\mu = 0$  y  $\sigma^2 = 1$  ( $X \sim N(0, 1)$ ). Es posible transformar una variable aleatoria normal no estándar en una estándar mediante la siguiente operación.

**Teorema 2.5.** *Si  $X \sim N(\mu, \sigma^2)$  entonces la variable aleatoria  $Z = \frac{X - \mu}{\sigma}$  tiene una distribución  $N(0, 1)$ .*

En la presente tesis es mencionado en varias ocasiones el valor  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \alpha/2)$  donde  $\Phi(z)$  es la función de distribución de una normal estándar,  $\kappa$  es el  $(1 - \frac{\alpha}{2})$  cuantil de la misma distribución y es un valor tal que,

$$P(Z \leq \kappa) = P(Z \leq z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}; \quad (2.4)$$

en la mayoría de ejemplos presentados en esta tesis  $\alpha = 0.05$  entonces  $\kappa$  es el valor tal que

$$P(Z \leq \kappa) = 1 - \frac{\alpha}{2} = 0.975;$$

como

$$P(Z \leq \kappa) = 1 - P(Z \geq \kappa)$$

entonces  $\kappa$  es tal que  $P(Z \geq \kappa) = \frac{\alpha}{2} = 0.025$ , por lo tanto  $\kappa \approx 1.96$  ([27], pág. 848).

## 2.2. Algunos tipos de convergencia

### Convergencia en probabilidad

**Definición 2.3.** *Se dice que una sucesión de variables aleatorias  $X_1, X_2, \dots, X_n$  converge en probabilidad a la variable aleatoria  $X$  denotado por  $X_n \xrightarrow{P} X$  si, para cada  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0, \text{ o equivalentemente, } \lim_{n \rightarrow \infty} P(|X_n - X| \leq \epsilon) = 1,$$

**Teorema 2.6.** Si  $X_i$  y  $Y_i$  son dos sucesiones de variables aleatorias que satisfacen que  $X_i \xrightarrow{P} X$  y  $Y_i \xrightarrow{P} Y$ , entonces

1.  $X_i \xrightarrow{P} X \iff X_i - X \xrightarrow{P} 0$ .
2.  $X_i \pm Y_i \xrightarrow{P} X \pm Y$ .
3. Si  $k$  es constante  $\implies kX_i \xrightarrow{P} kX$ .
4. Si  $X_i \xrightarrow{P} k \implies X_i^2 \xrightarrow{P} k^2$ .
5. Si  $X_i \xrightarrow{P} a$ ,  $Y_i \xrightarrow{P} b$ ,  $a, b$  constantes  $\implies X_i Y_i \xrightarrow{P} ab$ .
6. Sea  $X_i \neq 0$ , si  $X_i \xrightarrow{P} 1 \implies \frac{1}{X_i} \xrightarrow{P} 1$ .
7. Sean  $a, b$  constantes con  $b \neq 0$ . Si  $X_i \xrightarrow{P} a$ ,  $Y_i \xrightarrow{P} b$ , con  $X_i \neq 0 \implies \frac{X_i}{Y_i} \xrightarrow{P} \frac{a}{b}$ .
8.  $X_i Y \xrightarrow{P} XY$ .
9.  $X_i Y_i \xrightarrow{P} XY$ .

*Demostración.* ■ (1) se sigue inmediatamente de la definición.

- La expresión (2) se sigue de,

$$P(|(X_i + Y_i) - (X + Y)| \geq \epsilon) \leq P(|X_i - X| \geq \epsilon/2) + P(|Y_i - Y| \geq \epsilon/2).$$

- (3) se basa en el hecho,

$$P(|kX_i - kX| > |k|\epsilon) = P(|X_i - X| > \epsilon) \text{ para cualquier } \epsilon > 0.$$

- (4) se sigue de manera inmediata debido a que,

$$\{|X_i^2 - k^2| > \delta\} \subset \{|X_i - k| > \epsilon^*\}, \text{ al tomar } \epsilon^* = \min\left\{1, \frac{\delta}{2|k| + 1}\right\}, \text{ dado cualquier } \delta > 0.$$

- Para (5), aplicando (2), (3) y (4) se tiene que,
 
$$X_i Y_i = \frac{(X_i + Y_i)^2 - (X_i - Y_i)^2}{4} \xrightarrow{P} \frac{(a+b)^2 - (a-b)^2}{4} = ab.$$
- Para (6), de
 
$$P\left(\left|\frac{1}{X_i} - 1\right| \geq \epsilon\right) = P\left(\frac{1}{X_i} \geq 1 + \epsilon\right) + P\left(\frac{1}{X_i} \leq 1 - \epsilon\right) = P\left(\frac{1}{X_i} \geq 1 + \epsilon\right) + P\left(\frac{1}{X_i} \leq 0\right) + P\left(0 < \frac{1}{X_i} \leq 1 - \epsilon\right),$$
 donde cada término derecho tiende a 1 cuando  $n \rightarrow \infty$ .
- (7) se cumple usando (6) y (5).
- Para (8), como  $Y$  es una variable aleatoria, entonces dado  $\delta > 0$ , existe un  $k > 0$ , tal que  $P(|Y| > k) < \frac{\delta}{2}$ . De manera que
 
$$P(|X_i Y - XY| > \epsilon) = P(|X_i - X||Y| \geq \epsilon, |Y| > k) + P(|X_i - X|, |Y| \leq k) < \frac{\delta}{2} + P(|X_i - X| > \frac{\epsilon}{k}).$$
- Para (9), de (1) y (5) se tiene que  $(X_i - X)(Y_i - Y) \xrightarrow{P} 0$ , desarrollando y simplificando se tiene que  $X_i Y_i - XY_i - X_i Y + XY \xrightarrow{P} 0$ , usando (8), se llega a que  $X_i Y_i + XY \xrightarrow{P} 2XY$  y finalmente  $X_i Y_i \xrightarrow{P} XY$ .

□

**Lema 2.1.** [Desigualdad de Chebyshev] Para toda variable aleatoria  $X$  y para toda  $a > 0$  y  $c \in \mathbb{R}$ , se tiene

$$E[X - c]^2 \geq a^2 P(|X - c| \geq a).$$

**Teorema 2.7.** [Ley débil de los grandes números] Si  $X_1, X_2, \dots$  es un conjunto de variables aleatorias independientes e idénticamente distribuidas con  $E[X_i] = \mu$  y  $V[X_i] = \sigma^2 < \infty$ . Se define  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ , entonces  $\bar{X}_n$  converge en probabilidad a  $\mu$ .

*Demostración.* Haciendo uso de la desigualdad de Chebychev, se tiene, que para cada  $\epsilon > 0$

$$P(|\bar{X}_n - \mu| \geq \epsilon) = P((\bar{X}_n - \mu)^2 \geq \epsilon^2) \leq \frac{E[(\bar{X}_n - \mu)^2]}{\epsilon^2} = \frac{V[\bar{X}_n]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Así,

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \epsilon) \leq \lim_{n \rightarrow \infty} \frac{\sigma^2}{n\epsilon^2} = 0,$$

por tanto  $\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \epsilon) = 0$ . □

**Ejemplo 2.1.** Si  $X_1, X_2, \dots$  es una muestra aleatoria de una población  $B(p)$ , probar que  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  converge en probabilidad a  $p$ .

**Solución.** La solución es inmediata usando el teorema 2.7.

**Teorema 2.8.** Sean  $X_1, X_2, \dots, X_n$  un conjunto de variables aleatorias independientes e idénticamente distribuidas. Si  $X_n \xrightarrow{P} X$  y  $g$  es una función continua definida sobre  $\mathbb{R}$ , entonces  $g(X_n) \xrightarrow{P} g(X)$  cuando  $n \rightarrow \infty$ .

*Demostración.* Puesto que  $X$  es una variable aleatoria, se puede, dado un  $\epsilon > 0$ , encontrar una constante  $k = k(\epsilon)$  tal que,

$$P\{|X| > k\} < \frac{\epsilon}{2}.$$

También, como  $g$  es continua sobre  $\mathbb{R}$ , entonces  $g$  es uniformemente continua sobre  $[-k, k]$ . De ello se deduce que existe un  $\delta = \delta(\epsilon, k)$  tal que

$$|g(x_n) - g(x)| < \epsilon,$$

siempre que  $|x| \leq k$  y  $|x_n - x| < \delta$ . Sean

$$A = \{|X| \leq k\}, \quad B = \{|X_n - X| < \delta\}, \quad C = \{|g(X_n) - g(X)| < \epsilon\}.$$

Entonces  $\omega \in A \cap B \implies \omega \in C$ , de modo que

$$A \cap B \subseteq C,$$

y por lo tanto que,

$$P\{C^c\} \leq P\{A^c\} + P\{B^c\},$$

Esto es,

$$P\{|g(X_n) - g(X)| \geq \epsilon\} \leq P\{|X_n - X| \geq \delta\} + P\{|X| > k\} < \epsilon$$

para  $n \geq N(\epsilon, \delta, k)$ , donde  $N(\epsilon, \delta, k)$  se elige de manera que

$$P\{|X_n - X| \geq \delta\} < \frac{\epsilon}{2} \quad \text{para } n \geq N(\epsilon, \delta, k).$$

□

**Corolario 2.1.** Si  $X_i$  es una sucesión de variables aleatorias tal que  $X_i \xrightarrow{P} c$  y  $f$  es una función continua en  $c$ , entonces

$$f(X_i) \xrightarrow{P} f(c).$$

**Ejemplo 2.2.** Si  $X_1, X_2, \dots$  es una muestra aleatoria de una población  $B(p)$ , probar que  $\sqrt{\overline{X}_n(1 - \overline{X}_n)/n}$  converge en probabilidad a  $\sqrt{p(1-p)/n}$ .

**Solución.** La prueba es inmediata usando el corolario 2.1 y el ejemplo 2.1.

**Ejemplo 2.3.** Si  $X_1, X_2, \dots$  es una muestra aleatoria de una población  $B(p)$ , probar que

$$\frac{\sqrt{p(1-p)/n}}{\sqrt{\overline{X}_n(1 - \overline{X}_n)/n}}$$

converge en probabilidad a 1.

**Solución.** Usando el ejemplo 2.2, se tiene que  $\sqrt{\overline{X}_n(1 - \overline{X}_n)/n} \xrightarrow{P} \sqrt{p(1-p)/n}$ .

Por el teorema 2.6 incisos (3) y (6) se obtiene que,  $\frac{\sqrt{p(1-p)/n}}{\sqrt{\overline{X}_n(1 - \overline{X}_n)/n}} \xrightarrow{P} 1$ .

**Teorema 2.9** (Ley fuerte de los grandes números). Sean  $X_1, X_2, \dots$  variables aleatorias independientes e idénticamente distribuidas con  $E[X_i] = \mu$  y  $\text{Var}(X_i) = \sigma^2 < \infty$  y se define  $\overline{X}_n = (1/n) \sum_{i=1}^n X_i$ . Entonces, para todo  $\epsilon > 0$ ,

$$P\left(\lim_{n \rightarrow \infty} |\overline{X}_n - \mu| < \epsilon\right) = 1.$$

*Demostración.* (Suponiendo cuarto momento finito). Dada la idéntica distribución de los elementos de la sucesión, cualquier elemento de ésta se denota simplemente por  $X$ . Suponga que  $E[X - \mu]^2 = \sigma^2$  y observe que  $E[X - \mu] = 0$ . Entonces por la independencia,

$$E|\sum_{i=1}^n (X_i - \mu)|^4 = nE|X - \mu|^4 + 3n(n-1)\sigma^4.$$

Por la desigualdad de Chebyshev (Lema 2.1) aplicada a la variable  $|\sum_{i=1}^n (X_i - \mu)|$  y la función  $g(x) = x^4$  se obtiene, para  $\epsilon > 0$ ,

$$\begin{aligned} P(|\sum_{i=1}^n (X_i - \mu)| > n\epsilon) &\leq E|\sum_{i=1}^n (X_i - \mu)|^4 / (n\epsilon)^4 \\ &= (nE|X - \mu|^4 + 3n(n-1)\sigma^4) / (n\epsilon)^4. \end{aligned}$$

Sea el evento  $A_n = (|\frac{1}{n} \sum_{i=1}^n X_i - \mu| > \epsilon)$ . Entonces  $\sum_{n=1}^{\infty} P(A_n) < \infty$ . Por el lema de Borel-Cantelli ([22], pág. 38) la probabilidad de que ocurra una infinidad de eventos  $A_n$  es cero, es decir, con probabilidad uno, sólo un número finito de estos eventos ocurre. Por lo tanto con probabilidad uno, existe un número natural  $n$  a partir del cual ningún evento  $A_n$  se verifica. Es decir,

$$P(\lim_{n \rightarrow \infty} |\frac{1}{n} \sum_{i=1}^n X_i - \mu| \leq \epsilon) = 1.$$

Como esta afirmación vale para cualquier  $\epsilon > 0$ , se cumple que

$$P(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i - \mu) = 1.$$

□

## Convergencia en distribución

**Definición 2.4.** Sea  $X_1, X_2, \dots$ , una sucesión de variables aleatorias con función de distribución de probabilidad acumulativa  $F_{X_i}$  y  $X$  es una variable aleatoria con función de distribución de probabilidad acumulativa  $F_X$ . Se dice que  $X_i$  converge en distribución a  $X$  (denotado como  $X_i \xrightarrow{D} X$ ) si

$$\lim_{i \rightarrow \infty} F_{X_i} = F_X,$$

en todos los puntos donde  $F_X$  es continua.

**Teorema 2.10.** Sean  $a$  y  $b$  números reales con  $b \neq 0$ . Si  $X_i \xrightarrow{D} X$ , entonces  $bX_i + a \xrightarrow{D} bX + a$ .

*Demostración.*

Sea  $b > 0$  y sea  $x$  un punto de continuidad de  $bX + a$ ,  $-\infty < x < \infty$ . Entonces  $\frac{x-a}{b}$  es un punto de continuidad de  $X$ , además

$$P(bX_i + a \leq x) = P\left(X_i \leq \frac{x-a}{b}\right) \longrightarrow P\left(X \leq \frac{x-a}{b}\right) = P(bX + a \leq x).$$

La demostración para  $b < 0$  es análoga. □

**Teorema 2.11.** Sean  $X_i$  y  $Y_i$  sucesiones de variables aleatorias independientes. Si  $|X_i - Y_i| \xrightarrow{P} 0$  y  $Y_i \xrightarrow{D} Y$ , entonces  $X_i \xrightarrow{D} Y$ .

*Demostración.* Si  $x$  es un punto de continuidad de la función de densidad de la variable aleatoria  $Y$  y  $\epsilon > 0$ . Entonces

$$P\{X_i \leq x\} = P\{Y_i \leq x + Y_i - X_i\} = P\{Y_i \leq x + Y_i - X_i, Y_i - X_i \leq \epsilon\} + P\{Y_i \leq x + Y_i - X_i, Y_i - X_i > \epsilon\} \leq P\{Y_i \leq x + \epsilon\} + P\{Y_i - X_i > \epsilon\}.$$

De donde se sigue que,

$$\overline{\lim}_{i \rightarrow \infty} P\{X_i \leq x\} \leq \lim_{i \rightarrow \infty} P\{Y_i \leq x + \epsilon\}.$$

De forma similar tenemos que,

$$\lim_{i \rightarrow \infty} P\{X_i \leq x\} \geq \overline{\lim}_{i \rightarrow \infty} P\{Y_i \leq x - \epsilon\}.$$

Puesto que  $\epsilon > 0$  es arbitrario y  $x$  es un punto de continuidad de  $P\{Y \leq x\}$ , el resultado se consigue haciendo que  $\epsilon \rightarrow 0$ . □

**Corolario 2.2.** Si  $X_i \xrightarrow{P} X$  entonces  $X_i \xrightarrow{D} X$ .

*Demostración.* La demostración es directa del teorema 2.11. □

**Teorema 2.12.** Sean  $X_1, X_2, \dots$ , una sucesión de variables aleatorias tal que existen las  $M_{X_i}(t)$ , siendo las funciones generadoras de momentos de las  $X_i$ , además suponga que

$$\lim_{i \rightarrow \infty} M_{X_i}(t) = M_X(t), \text{ para } t \text{ en alguna vecindad del } 0,$$

donde  $M_X(t)$  es una función generadora de momentos. Entonces hay una única función de distribución acumulativa  $F_X$  y cuyos momentos son determinados por  $M_X(t)$ , para todo  $x$  donde  $F_X(x)$  es continua, se tiene

$$\lim_{i \rightarrow \infty} F_{X_i}(x) = F_X(x). \quad (2.5)$$

Es decir, la convergencia, para  $|t| < h$  de funciones generadoras de momentos a una función generadora de momentos implica la convergencia de las funciones de distribución acumulativas.

*Demostración.* Ver en ([12]). □

**Teorema 2.13** (Teorema central del límite). Sea  $X_1, X_2, \dots, X_n$  variables aleatorias independientes e idénticamente distribuidas cuyas funciones generadoras de momentos existen en alguna vecindad alrededor 0 (esto es,  $M_{X_i}(t)$  existe para  $|t| < h$  para algún  $h$  positivo). Sea  $E[X_i] = \mu$  y  $\text{Var}(X_i) = \sigma^2 > 0$  ( $\mu$  y  $\sigma^2$  son finitos puesto que la función generatriz de momentos existe). Se define  $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$  y sea  $G_n(x)$  la función de distribución acumulativa de  $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ . Entonces para  $-\infty < x < \infty$ :

$$\lim_{n \rightarrow \infty} G_n(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy;$$

esto es,  $\sqrt{n}(\bar{X}_n - \mu)/\sigma$  tiene distribución límite normal estándar.

*Demostración.* Se prueba que para  $|t| < h$ , la función generadora de momentos de  $\sqrt{n}(\bar{X}_n - \mu)/\sigma$  converge hacia  $e^{t^2/2}$ , la función generatriz de momentos de

una variable aleatoria normal estándar ( $N(0, 1)$ ).

Se define  $Y_i = (X_i - \mu)/\sigma$  y sea  $M_Y(t)$  la función generatriz de momentos de las  $Y_i$ s, las cuales existen para  $|t| < \sigma h$ . Sea  $W$  tal que:

$$W = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i,$$

entonces:

$$M_W(t) = M_{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i}(t) = M_{\sum_{i=1}^n Y_i}(t/\sqrt{n}) = [M_Y(t/\sqrt{n})]^n.$$

Expandiendo en series de Taylor,

$$M_Y(t/\sqrt{n}) = \sum_{k=0}^{\infty} M_Y^{(k)}(0) \frac{(t/\sqrt{n})^k}{k!} \text{ donde } M_Y^{(k)}(0) = \left. \frac{d^k}{dt^k} M_Y(t) \right|_{t=0}.$$

Como la función generadora de momentos existe para  $|t| < \sigma h$ , la expresión en serie de potencias es válida si  $t < \sqrt{n}\sigma h$ . Usando el hecho de que  $M_Y^{(0)} = 1$ ,  $M_Y^{(1)} = 0$  y  $M_Y^{(2)} = 1$ , ya que por construcción la media y varianza de  $Y$  son 0 y 1 respectivamente. Entonces:

$$M_Y(t/\sqrt{n}) = 1 + \frac{(t/\sqrt{n})^2}{2!} + R_Y(t/\sqrt{n})$$

donde  $R_Y$  es el residuo en la expansión de Taylor.

$$R_Y(t/\sqrt{n}) = \sum_{k=3}^{\infty} M_Y^{(k)}(0) \frac{(t/\sqrt{n})^k}{k!},$$

para  $t \neq 0$  fijo, se tiene:

$$\lim_{n \rightarrow \infty} \frac{R_Y(t/\sqrt{n})}{(t/\sqrt{n})^2} = 0.$$

Como  $t$  es fija, se tiene:

$$\lim_{n \rightarrow \infty} \frac{R_Y(t/\sqrt{n})}{(1/\sqrt{n})^2} = \lim_{n \rightarrow \infty} n R_Y\left(\frac{t}{\sqrt{n}}\right) = 0$$

y lo anterior se cumple también en  $t=0$  ya que  $R_Y(0/\sqrt{n}) = 0$ , así que para  $t$  fijo, se tiene;

$$\begin{aligned}\lim_{n \rightarrow \infty} [M_Y(t/\sqrt{n})]^n &= \lim_{n \rightarrow \infty} \left[ 1 + \frac{(t/\sqrt{n})^2}{2!} + R_Y(t/\sqrt{n}) \right]^n \\ &= \lim_{n \rightarrow \infty} \left[ 1 + \frac{1}{n} \left( \frac{t^2}{2} + nR_Y(t/\sqrt{n}) \right) \right]^n\end{aligned}$$

y por último se tiene:

$$\lim_{n \rightarrow \infty} [M_Y(t/\sqrt{n})]^n = e^{t^2/2},$$

que es la función generatriz de momentos de la distribución  $N(0,1)$ .  $\square$

**Teorema 2.14** (Teorema de Slutsky). *Si  $X_i \xrightarrow{D} X$  y  $Y_i \xrightarrow{P} c$ , donde  $c$  es una constante, entonces:*

1.  $X_i + Y_i \xrightarrow{D} X + c$ ,
2.  $X_i Y_i \xrightarrow{P} 0$  si  $c = 0$  y  $X_i Y_i \xrightarrow{D} cX$  si  $c \neq 0$ .

*Demostración.* Para 1, como  $X_i \xrightarrow{D} X$ , entonces por el teorema 2.10  $X_i + c \xrightarrow{D} X + c$ . Además  $Y_i - c = (X_i + Y_i) - (X_i + c) \xrightarrow{P} 0$ . Así utilizando el teorema 2.11, se tiene que

$$X_i + Y_i \xrightarrow{D} X + c.$$

Para 2, si  $c = 0$ . Entonces, para cualquier número fijo  $k > 0$ ,

$$\begin{aligned}P\{|X_i Y_i| > \epsilon\} &= P\{|X_i Y_i| > \epsilon, |Y_i| \leq \frac{\epsilon}{k}\} + P\{|X_i Y_i| > \epsilon, |Y_i| > \frac{\epsilon}{k}\} \\ &\leq P\{|X_i| > k\} + P\{|Y_i| > \frac{\epsilon}{k}\}.\end{aligned}$$

Como  $Y_i \xrightarrow{P} 0$  y  $X_i \xrightarrow{D} X$ , entonces, para cualquier  $k > 0$  fijo se tiene que,

$$X_i Y_i \xrightarrow{P} 0.$$

Si  $c \neq 0$ , entonces

$$X_i Y_i - cX_i = X_i(Y_i - c),$$

y dado que  $X_i \xrightarrow{D} X$ ,  $Y_i \xrightarrow{P} c$ , se cumple que,  $X_i(Y_i - c) \xrightarrow{P} 0$ , y por el teorema 2.11 se cumple que,

$$X_i Y_i \xrightarrow{D} cX.$$

□

## 2.3. Estimación puntual

**Definición 2.5.** (*Muestra aleatoria*) A la sucesión de variables aleatorias  $X_1, X_2, \dots, X_n$  se le llama muestra aleatoria de tamaño  $n$  de la población  $f(x)$ . Si  $X_1, X_2, \dots, X_n$  son variables aleatorias mutuamente independientes y la función de densidad de probabilidad de cada  $X_i$  es la misma función  $f(x)$ . De manera alternativa se dice que  $X_1, X_2, \dots, X_n$  son variables aleatorias independientes e idénticamente distribuidas con función de densidad  $f(x)$ .

**Definición 2.6.** (*Espacio paramétrico*)

Sea  $X$  una variable aleatoria con función de densidad  $f(\cdot, \theta)$ , con  $\theta$  desconocido. Al conjunto de todos los posibles valores que toma  $\theta$  se le llama espacio paramétrico y se denotará por  $\Theta$ .

El propósito de la estadística frecuentemente es usar la información contenida en una muestra con el fin de realizar estimaciones acerca de una población. Para realizar esto, un investigador tiene que obtener datos muestrales de la población estudiada. Las conclusiones pueden entonces basarse en los valores calculados de varias cantidades muestrales. La estimación tiene varias aplicaciones prácticas. Por ejemplo, un fabricante de máquinas lavadoras podría estar interesado en estimar la proporción de lavadoras que esperaría que fallen antes de la expiración de la garantía de un año. También se podría estar interesado en estimar otros parámetros poblacionales tales como la media poblacional, la varianza y la desviación estándar. Por ejemplo se puede estimar la media del tiempo de espera en una caja registradora del supermercado o la desviación estándar del error de medición de un instrumento electrónico.

Existen dos tipos de estimación *la estimación puntual y la estimación por intervalo*, en esta sección se estudiará la *estimación puntual*.

**Definición 2.7.** *Una estadística es una función que sólo depende de las variables aleatorias que participan en la muestra y que no contiene cantidades desconocidas, es decir, una estadística es de la forma,*

$$T = t(X_1, \dots, X_n).$$

**Definición 2.8.** *Sea  $X$  una variable aleatoria con función de densidad de probabilidad  $f(\cdot, \theta)$ , con  $\theta$  desconocido. Sean  $X_1, \dots, X_n$ , una muestra aleatoria y sean  $x_1, \dots, x_n$ , los valores muestrales correspondientes, y  $\theta \in \Theta$ . Si  $g(X_1, \dots, X_n)$  es una estadística cuyo rango (o soporte) es un subconjunto del espacio paramétrico de la población, entonces diremos que  $g$  es un estimador puntual de  $\theta$ .*

**NOTACIÓN:**

$$\begin{aligned}\hat{\Theta} &= g(X_1, \dots, X_n), \text{ es el estimador de } \theta; \\ \hat{\theta} &= g(x_1, \dots, x_n), \text{ es una estimación de } \theta.\end{aligned}$$

La estimación puntual consiste en encontrar un número, con base en las observaciones realizadas de la variable aleatoria, que sirva como estimación del parámetro desconocido  $\theta$ .

Por ejemplo, en la mayoría de casos la probabilidad de éxito  $p$  asociada con cierto ensayo Bernoulli es desconocida y se desea estimar su valor. Una manera de lograr esto consiste en realizar  $n$  ensayos independientes y contabilizar el número de éxitos  $X$ . Para estimar  $p$  se calcula la proporción muestral  $\hat{p}$  definida como

$$\hat{p} = \frac{X}{n}$$

donde  $X \sim B(n, p)$ .

## Estimadores de máxima verosimilitud

**Definición 2.9.** Si  $f(\mathbf{x}|\theta)$  denota la función de probabilidad o densidad conjunta de la muestra  $X = (X_1, \dots, X_n)$ . Entonces, dado que  $X = \mathbf{x}$  es observado, la función de  $\theta$  definida por:

$$L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta),$$

es llamada función de verosimilitud.

Si  $X_1, \dots, X_n$  es una muestra independiente e idénticamente distribuida de una población con función de probabilidad o densidad  $f(x|\theta_1, \dots, \theta_k)$ , la función de verosimilitudes se define por:

$$L(\theta|\mathbf{x}) = L(\theta_1, \dots, \theta_k|x_1, \dots, x_n) = \prod_{i=1}^n f(x_i|\theta_1, \dots, \theta_k).$$

**Definición 2.10.** Para cada punto muestral  $\mathbf{x}$ , sea  $\hat{\theta}(\mathbf{x})$  el valor del parámetro en que  $L(\theta|\mathbf{x})$  toma su valor máximo como función de  $\theta$ , con  $\mathbf{x}$  fijo. Un estimador de máxima verosimilitud del parámetro  $\theta$  basado en la muestra  $X$  es  $\hat{\theta}(X)$ .

Si la función de verosimilitud es diferenciable en  $\theta_i$ , los posibles candidatos para estimadores de máxima verosimilitud son los valores de  $(\theta_1, \dots, \theta_k)$  que resuelven:

$$\frac{\partial}{\partial \theta_i} L(\theta|\mathbf{x}) = 0, \quad i = 1, 2, \dots, k. \quad (2.6)$$

Note que las soluciones en (2.6) son solo posibles candidatos para los estimadores de máxima verosimilitud ya que la primera derivada igualada a cero es sólo una condición necesaria para un máximo, no es una condición suficiente. Además, los ceros de la primera derivada ubican solo puntos extremos en el interior del dominio de una función. Si los extremos se producen en los límites la primera, derivada no puede ser cero. Por lo tanto, el límite debe verificarse por separado para los extremos.

Los puntos en los que las primeras derivadas son 0 pueden ser mínimos locales o globales, máximos globales o locales o puntos de inflexión. Nuestro trabajo es encontrar un máximo global

**Ejemplo 2.4.** Sea  $X_1, \dots, X_n$  variables aleatorias independientes e idénticamente distribuidas según  $B(p)$ .

$$L(p|\mathbf{x}) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i}.$$

Tomando logaritmo a ambos lados obtenemos:

$$\log L(p|\mathbf{x}) = \sum_{i=1}^n x_i \log p + (n - \sum_{i=1}^n x_i) \log(1-p).$$

Derivando e igualando a 0  $\left( \frac{d}{dp} \log L(p|\mathbf{x}) = 0 \right)$  se obtiene que  $\hat{p} = \frac{\sum_{i=1}^n x_i}{n}$  es un punto crítico de la función de verosimilitud. Luego, verificando que:

$$\left. \frac{d^2}{dp^2} \log L(p|\mathbf{x}) \right|_{p=\hat{p}} < 0$$

se concluye que  $\hat{p} = \frac{\sum_{i=1}^n X_i}{n} = \bar{X}$  es el estimador de máxima verosimilitud.

En muchas ocasiones, para encontrar el estimador de máxima verosimilitud no puede aplicarse el criterio de la segunda derivada y el punto donde la función de verosimilitud alcanza su máximo se tendría que obtener usando algún otro criterio. Además en algunos casos el estimador de máxima verosimilitud no existe y en otros no es único ([15]).

## 2.4. Pruebas de hipótesis.

**Definición 2.11.** [Espacio de probabilidad] Un espacio de probabilidad es una terna  $(\Omega, \mathcal{F}, P)$ , en donde  $\Omega$  es un conjunto arbitrario,  $\mathcal{F}$  es una  $\sigma$ -álgebra de subconjuntos de  $\Omega$  y  $P$  es una medida de probabilidad definida sobre  $\mathcal{F}$ .

Consideramos el problema estadístico que contiene un parámetro desconocido, digamos  $\theta$ . Se supone que  $\theta$  pertenece a un espacio paramétrico  $\Theta \subset R^n$ . Supongamos además que  $\Theta$  es una partición de dos subconjuntos que son disjuntos  $\Theta_0$  y  $\Theta_1$ , así el problema consiste en decidir si el parámetro desconocido  $\theta$  pertenece a  $\Theta_0$  o a  $\Theta_1$ .

En varios aspectos, el procedimiento formal para pruebas de hipótesis es semejante al método científico. Éste observa la naturaleza, fórmula una teoría y la confronta con lo que es observado. En nuestro contexto, el investigador plantea alguna hipótesis respecto a uno o más parámetros poblacionales: les da algunos valores específicos con base a su experiencia. Después toma una muestra de la población y compara sus observaciones con la hipótesis. Si las observaciones no concuerdan con la hipótesis, la rechaza usando un valor fijo de equivocación entre cero y uno; de lo contrario, concluye que no existe evidencia para rechazar la hipótesis o que la muestra no detecta diferencia entre los valores real e hipotético de los parámetros poblacionales.

**Definición 2.12.** *Dado un experimento aleatorio se define el espacio muestral de dicho experimento, denotado por  $\Omega$ , como el conjunto de todos los posibles resultados del experimento en cuestión.*

**Definición 2.13.** *Una hipótesis es una declaración acerca de un parámetro poblacional.*

**Definición 2.14.** *Las dos hipótesis complementarias en un problema de prueba de hipótesis son llamadas hipótesis nula e hipótesis alternativa y se denotan por  $H_0$  y  $H_1$ , respectivamente.*

El objetivo de una prueba de hipótesis es decidir, con base en una muestra de una población, cuál de las dos hipótesis complementarias es verdadera. Si  $\theta$  denota un parámetro poblacional, el formato general de las hipótesis nula es  $H_0 : \theta \in \Theta_0$  y de la alternativa es  $H_1 : \theta \in \Theta_0^c$ ; donde  $\Theta_0$  es algún subconjunto del espacio paramétrico y  $\Theta_0^c$  es su complemento.

**Definición 2.15.** *Un procedimiento de prueba de hipótesis es una regla que especifica:*

- i. *Para que valores muestrales no se rechaza  $H_0$  y se encuentra que es verdadera.*
- ii. *Para que valores muestrales la decisión es rechazar  $H_0$  y no se rechaza  $H_1$  encontrándose que es verdadera.*

*El conjunto del espacio muestral para el cual  $H_0$  sería rechazada es llamado la **región de rechazo (RR) o región crítica**. El complemento de la región de rechazo es llamado la región de **no rechazo**.*

Típicamente una prueba de hipótesis está especificada por un estadístico de prueba  $T(X_1, X_2, \dots, X_n) = T(\mathbf{X})$ , es decir, por una función que solo depende de la muestra.

### 2.4.1. Métodos de evaluación de pruebas.

En la decisión de aceptar o rechazar la hipótesis nula  $H_0$ , un experimentador podría cometer un error. Usualmente, las pruebas de hipótesis son evaluadas y comparadas con base en dichas posibilidades de error. En esta sección se discutirá como esas probabilidades pueden ser controladas.

#### Probabilidades de error y potencia de prueba

Una prueba de hipótesis para  $H_0 : \theta \in \Theta_0$  vs  $H_1 : \theta \in \Theta_0^c$  puede conducir a dos tipos de error.

	Decisión	
Hipótesis verdadera	No se rechaza $H_0$	Se rechaza $H_0$
$H_0$	Decisión correcta	Error tipo I
$H_1$	Error tipo II	Decisión correcta

**Definición 2.16.** *Suponga que  $R$  denota la región de rechazo para una prueba dada. Entonces para  $\theta \in \Theta_0$ , la prueba comete un error si  $X \in R$ , llamado error tipo I; la probabilidad de error tipo I se define como  $P_\theta(X \in R)$ . De manera similar, para  $\theta \in \Theta_0^c$  se comete un error si  $X \in R^c$  llamado error tipo II; la probabilidad de error tipo II se define como  $P_\theta(X \in R^c)$ .*

$$P_\theta(X \in R) = \begin{cases} \text{probabilidad del error tipo I,} & \text{si } \theta \in \Theta_0; \\ 1 - \text{probabilidad del error tipo II,} & \text{si } \theta \in \Theta_0^c. \end{cases}$$

La probabilidad de un error tipo I será denotada por  $\alpha(\theta)$ . La probabilidad de un error tipo II será denotada por  $\beta(\theta)$ .

### 2.4.2. Pruebas para muestras grandes.

Un método común para construir una prueba estadística está basado en un estimador que tiene una distribución normal asintótica. Suponga que se desea probar una hipótesis acerca de un parámetro  $\theta$  de valor real y  $W_n = W(X_1, X_2, \dots, X_n)$  es un estimador puntual de  $\theta$  basado en una muestra aleatoria de tamaño  $n$ , que ha sido encontrado por algún método de estimación. Por ejemplo,  $W_n$  puede ser el estimador de máxima verosimilitud de  $\theta$ . Una prueba aproximada basada en una aproximación normal, puede ser justificada de la siguiente manera.

Si  $\sigma_n^2$  denota la varianza de  $W_n$  y usando el teorema central del límite (sección teorema 2.13) o de alguna de sus variaciones para probar que, cuando  $n \rightarrow \infty$ ,  $\frac{W_n - \theta}{\sigma_n}$  converge en distribución a una variable aleatoria normal estándar ( $N(0, 1)$ ), entonces  $\frac{W_n - \theta}{\sigma_n}$  se puede comparar con una distribución normal estándar. Por lo tanto, se tiene la base para una prueba aproximada.

Hay muchos detalles a ser verificados en el argumento del párrafo anterior, pero esta idea aplica en muchas situaciones. Por ejemplo, si  $W_n$  es un estimador de máxima verosimilitud, el argumento de arriba es usualmente válido. Note

que la distribución de  $W_n$  y quizás el valor de  $\sigma_n$  dependen de  $\theta$ . La convergencia por lo tanto, más formalmente dice, que para cada valor fijo  $\theta \in \Theta$ , si usamos la correspondiente distribución de  $W_n$  y el valor correspondiente de  $\sigma_n$ ,  $\frac{W_n - \theta}{\sigma_n}$  converge a una normal estándar. Si para cada  $n$ ,  $\sigma_n$  es una constante calculable (que puede depender sólo de  $\theta$  y no de otro parámetro desconocido), entonces podría ser derivada una prueba basada en  $\frac{W_n - \theta}{\sigma_n}$ .

En algunos casos,  $\sigma_n$  puede depender de parámetros desconocidos. En tal caso buscaremos el estimador  $S_n$  de  $\sigma_n$  con la propiedad de que  $\frac{\sigma_n}{S_n}$  converja en probabilidad a 1. Entonces utilizando el teorema de Slutsky (2.14) podemos deducir que  $\frac{W_n - \theta}{S_n}$  converge en distribución a una distribución normal estándar. Una prueba para muestras grandes puede ser basada en este hecho.

Suponga que deseamos probar

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta \neq \theta_0.$$

Una prueba aproximada está basada en el estadístico  $Z_n = \frac{W_n - \theta_0}{S_n}$  y se rechazaría  $H_0$  si y sólo si  $Z_n < z_{\alpha/2}$  o  $Z_n > z_{\alpha/2}$ . Si  $H_0$  es verdadero, entonces  $\theta = \theta_0$  y  $Z_n$  converge en distribución a una distribución normal estándar. Así la probabilidad de error tipo 1,  $P_{\theta_0}(Z_n < z_{\alpha/2} \text{ o } Z_n > z_{\alpha/2}) \rightarrow P(Z < z_{\alpha/2} \text{ o } Z > z_{\alpha/2}) = \alpha$ .

Para ejemplificar este hecho veamos un ejemplo.

**Ejemplo 2.5.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de una población  $B(p)$ , considere probar  $H_0 : p \leq p_0$  vs  $H_1 : p > p_0$  donde  $0 < p_0 < 1$  es un valor específico. El estimador de máxima verosimilitud de  $p$  basado en una muestra de tamaño  $n$ , es  $\hat{p}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . Debido a que  $\hat{p}_n$  es solo la media muestral, el teorema central del límite se aplica y establece que para cualquier  $p$ ,  $0 < p < 1$ ,  $\frac{\hat{p}_n - p}{\sigma_n}$  converge a una variable aleatoria normal estándar. Aquí  $\sigma_n = \sqrt{p(1-p)/n}$  es un valor que depende del parámetro desconocido  $p$ . Un estimador razonable de  $\sigma_n$  es  $S_n = \sqrt{\hat{p}_n(1-\hat{p}_n)/n}$  y puede mostrarse que  $\frac{\sigma_n}{S_n}$

converge en probabilidad a 1 (ver ejemplo 2.3). Así para cualquier  $p$ , por el teorema de Slutsky se tiene que,

$$\frac{\hat{p}_n - p}{\sqrt{\frac{\hat{p}_n(1 - \hat{p}_n)}{n}}} \rightarrow N(0, 1). \quad (2.7)$$

El estadístico de prueba  $Z_n$  está definido mediante la sustitución de  $p$  por  $p_0$  y la prueba para muestras grandes rechaza  $H_0$  si  $Z_n > z_\alpha$ .

Si había interés en probar la hipótesis  $H_0 : p = p_0$  vs  $H_1 : p \neq p_0$  donde  $0 < p_0 < 1$  es un valor específico, la estrategia de arriba es otra vez aplicable. Sin embargo, en este caso, hay una prueba aproximada alternativa. Por el teorema central del límite, para cualquier  $p$ ,  $0 < p < 1$ ,

$$\frac{\hat{p}_n - p}{\sqrt{p(1 - p)/n}} \rightarrow N(0, 1). \quad (2.8)$$

Por lo tanto, se deduce que, si la hipótesis nula es verdadera, el estadístico

$$\acute{Z}_n = \frac{\hat{p}_n - p_0}{\sqrt{p_0(1 - p_0)/n}} \rightarrow N(0, 1) \quad (\text{aproximadamente}). \quad (2.9)$$

La prueba aproximada nivel  $\alpha$  rechaza  $H_0$  si  $|\acute{Z}_n| > z_{\alpha/2}$ .

## 2.5. Intervalos de confianza

Al tomar una muestra aleatoria (m.a.)  $X_1, X_2, \dots, X_n$  de alguna población con función de densidad  $f(\cdot, \theta)$  con  $\theta$  desconocido, podemos a partir de la muestra, estimar un valor puntual del parámetro  $\theta$ . Sin embargo al tomar otra muestra aleatoria  $Y_1, Y_2, \dots, Y_n$  de la misma población, obtendríamos una estimación del parámetro, en general diferente al de la primera muestra. Así cada muestra aleatoria proporcionará un valor diferente para el parámetro. ¿Cuál de estos valores será el más cercano a  $\theta$ ?. Es imposible precisar cuál de todos esos valores es el más cercano. Ya que en general la probabilidad de que el

valor estimado sea igual al parámetro es cero.

A continuación se presentan algunos conceptos básicos relativos a los intervalos de confianza (ver por ejemplo ([10])).

### 2.5.1. Introducción a intervalos de confianza.

**Definición 2.17.** Una estimación por intervalo de un parámetro  $\theta$  consiste en un par de funciones  $L(x_1, x_2, \dots, x_n)$  y  $U(x_1, x_2, \dots, x_n)$  de una muestra que satisface  $L(\mathbf{x}) \leq U(\mathbf{x})$  para todo  $\mathbf{x} \in \Omega$ , donde  $\Omega$  es el espacio muestral. Si  $X = \mathbf{x}$  es observado, la inferencia  $L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})$  está hecha. El intervalo aleatorio  $[L(X), U(X)]$  es llamado un estimador por intervalo.

El propósito de usar un estimador por intervalo en lugar de un estimador puntual es tener alguna garantía de que el estimador contenga al parámetro de interés. La certeza de esta garantía se encuentra determinada por las siguientes definiciones.

**Definición 2.18.** Para un estimador por intervalo  $[L(X), U(X)]$  de un parámetro  $\theta$ , la probabilidad de cobertura denotado por **PC** es la probabilidad de que el intervalo aleatorio contenga al verdadero parámetro,  $\theta$ . Esto se denota por  $P_\theta(\theta \in [L(X), U(X)])$  o por  $P(\theta \in [L(X), U(X)] | \theta)$ .

**Definición 2.19.** Para un estimador por intervalo  $[L(X), U(X)]$  de un parámetro  $\theta$ , el coeficiente de confianza de  $[L(X), U(X)]$  es el ínfimo de la probabilidades de cobertura, es decir  $\inf_\theta P_\theta(\theta \in [L(X), U(X)])$ .

**Observación 1.** Es importante aclarar dos puntos:

1. El parámetro no es la cantidad aleatoria. La cantidad aleatoria es el intervalo de confianza.
2. En  $P_\theta(\theta \in [L(X), U(X)])$ , la probabilidad se refiere a  $X$  no a  $\theta$ .

Los estimadores por intervalo junto con una medida de confianza (usualmente un coeficiente de confianza) son también conocidos como intervalos de confianza (**IC**). Aunque principalmente se trata con intervalos de confianza existen conjuntos más generales (conjuntos de confianza).

Un intervalo de confianza con un coeficiente de confianza igual a  $1 - \alpha$ , es simplemente llamado un intervalo de confianza  $1-\alpha$ , donde  $\alpha$  es llamado nivel de significancia del intervalo e indica el porcentaje de error que podemos cometer en la construcción del intervalo.

En el proceso de tomar muestras aleatorias de una población y calcular un intervalo de confianza con coeficiente de confianza  $1 - \alpha$  para un parámetro  $\theta$ , en estos intervalos a largo plazo la proporción de intervalos que cubren al parámetro  $\theta$  es equivalente a  $1 - \alpha$ . Una forma de ver esto es imaginarlo de la misma forma que Boomsma en ([7]), como el juego de lanzar dardos de diferentes longitudes a un blanco ( $\theta$ ), los intervalos en este caso son las longitudes de los palos de tal manera que el blanco no es impactado en  $100\alpha\%$  de los casos.

### **2.5.2. Métodos para encontrar estimadores por intervalo.**

Existen diversas formas de construir intervalos de confianza, algunas de ellas son: inversión de pruebas estadísticas, uso de cantidades pivotaes, el pivoteo de la función de distribución acumulativa e Intervalos Bayesianos (ver por ejemplo [10]).

#### **Inversión de pruebas estadísticas**

Existe una fuerte correspondencia entre prueba de hipótesis y estimadores por intervalo, podemos decir en general que a cada conjunto de confianza le corresponde una prueba y viceversa.

La correspondencia está descrita en el siguiente teorema.

**Teorema 2.15.** Para cada  $\theta_0 \in \Theta$ , si  $A(\theta_0)$  es la región de no rechazo de una prueba de nivel  $\alpha$  de  $H_0 : \theta = \theta_0$ . Para cada  $\mathbf{x} \in \Omega$ , se define un conjunto  $C(\mathbf{x})$ , en el espacio paramétrico como

$$C(\mathbf{x}) = \{\theta_0 : \mathbf{x} \in A(\theta_0)\}. \quad (2.10)$$

Entonces el conjunto aleatorio  $C(X)$  es un conjunto de confianza  $1 - \alpha$ . Inversamente, si  $C(X)$  es un conjunto de confianza  $1 - \alpha$ . Para cualquier  $\theta_0 \in \Theta$ , se define

$$A(\theta_0) = \{\mathbf{x} : \theta_0 \in C(\mathbf{x})\}.$$

Entonces  $A(\theta_0)$  es la región de no rechazo de una prueba de nivel  $\alpha$  de  $H_0 : \theta = \theta_0$ .

*Demostración.* Para la primera parte, ya que  $A(\theta_0)$  es la región de no rechazo de una prueba de nivel  $\alpha$ ,

$$P_{\theta_0}(X \notin A(\theta_0)) \leq \alpha \text{ y por lo tanto } P_{\theta_0}(X \in A(\theta_0)) \geq 1 - \alpha.$$

Debido a que  $\theta_0$  es arbitrario, escribimos  $\theta$  en lugar de  $\theta_0$ . La inecuación de arriba junto con (2.10), muestran que la probabilidad de cobertura del conjunto  $C(X)$  está dada por

$$P_{\theta}(\theta \in C(X)) = P_{\theta}(X \in A(\theta)) \geq 1 - \alpha.$$

Mostrando que  $C(X)$  es un conjunto de confianza  $1 - \alpha$ .

Para la segunda parte, el error de probabilidad tipo 1 para la prueba de  $H_0 : \theta = \theta_0$  con región de no rechazo  $A(\theta_0)$  es

$$P_{\theta_0}(X \notin A(\theta_0)) = P_{\theta_0}(\theta_0 \notin C(X)) \leq \alpha.$$

Así que se trata de una prueba de nivel  $\alpha$

□

## Uso de cantidades pivotaes.

**Definición 2.20.** Una variable aleatoria  $Q(X, \theta) = Q(X_1, X_2, \dots, X_n, \theta)$  es una cantidad pivotal, o pivote, si la distribución de  $Q(X, \theta)$  es independiente del parámetro. Es decir, si  $X \sim F(\mathbf{x}, \theta)$  entonces  $Q(X, \theta)$  tiene la misma distribución para todos los valores de  $\theta$ .

### Método de la Cantidad Pivotal

Si  $Q(X, \theta) = Q(X_1, X_2, \dots, X_n, \theta)$  es una cantidad pivotal para el parámetro  $\theta$ , entonces, para cualquier  $0 < 1 - \alpha < 1$  fijo, existen  $q_1$  y  $q_2$  dependientes de  $\alpha$  tales que,

$$P[q_1 < Q < q_2] = 1 - \alpha,$$

si para cada valor muestral  $x_1, x_2, \dots, x_n$  tenemos que:

$$q_1 < q(x_1, x_2, \dots, x_n; \theta) < q_2,$$

si y sólo si,  $t_1(x_1, x_2, \dots, x_n) < \theta < t_2(x_1, x_2, \dots, x_n)$  para  $t_1$  y  $t_2$  funciones que no dependen de  $\theta$ , entonces  $(t_1(x_1, x_2, \dots, x_n), t_2(x_1, x_2, \dots, x_n))$  es un intervalo de confianza  $1 - \alpha$ .

### Pivoteo de la función de distribución acumulativa

En la sección anterior se vio que un pivote  $Q$ , conduce a un conjunto de confianza de la forma,

$$C(\mathbf{x}) = \{\theta_0 : a \leq Q(\mathbf{x}, \theta_0) \leq b\}.$$

Si para cualquier  $\mathbf{x}$ , la función  $Q(\mathbf{x}, \theta)$  es una función monótona de  $\theta$ , entonces  $C(\mathbf{x})$  está garantizado a ser un intervalo. Los pivotes que son principalmente construidos usando transformaciones de forma y escala, resultan una función  $Q$  monótona y por tanto conducen a intervalos de confianza. En esta sección se trabajará con otro pivote, que es totalmente general y que con menos suposiciones garantiza un intervalo.

Basamos nuestra construcción de un intervalo de confianza para un parámetro  $\theta$  en una estadística  $T$  con función de distribución acumulativa  $F_T(t)$ . En la práctica usualmente se toma  $T$  una estadística suficiente para  $\theta$ , más no necesariamente debe ser así, consideraremos el caso discreto en el teorema 2.16.

**Definición 2.21.** Una función de distribución acumulativa  $F_X$  es estocásticamente mayor que una función de distribución acumulativa  $F_Y$  si  $F_X(t) \leq F_Y(t)$  para todo  $t$  y  $F_X(t) < F_Y(t)$  para algún  $t$

**Teorema 2.16.** Si  $T$  es una estadística discreta con función de distribución acumulativa  $F_T(t | \theta) = P(T \leq t | \theta)$ . Si  $0 < \alpha < 1$  es un valor fijo. Suponga que para todo  $t \in \mathcal{T}$ , el espacio muestral de  $T$ ,  $\theta_L(t)$  y  $\theta_U(t)$  pueden definirse como sigue:

1. Si  $F_T(t | \theta)$  es una función decreciente de  $\theta$  para cada  $t$ , se define  $\theta_L(t)$  y  $\theta_U(t)$  por:

$$P(T \leq t | \theta_U(t)) = \frac{\alpha}{2}, P(T \geq t | \theta_L(t)) = \frac{\alpha}{2}.$$

2. Si  $F_T(t | \theta)$  es una función creciente de  $\theta$  para cada  $t$ , se define  $\theta_L(t)$  y  $\theta_U(t)$  por:

$$P(T \geq t | \theta_U(t)) = \frac{\alpha}{2}, P(T \leq t | \theta_L(t)) = \frac{\alpha}{2}.$$

Entonces, el intervalo aleatorio  $[\theta_L(T), \theta_U(T)]$  es un intervalo de confianza  $1 - \alpha$  para  $\theta$ .

*Demostración.* Se hará un bosquejo de la demostración para la parte (1), para una demostración más exhaustiva consulte en ([10]).

Sea  $\bar{F}_T(t|\theta)$  la función definida por

$$\bar{F}_T(t|\theta) = P(T \geq t|\theta).$$

Debido a que  $F_T(t|\theta)$  es una función decreciente de  $\theta$  para cada  $t$ . Se puede mostrar que  $\bar{F}_T(t|\theta)$  es una función no decreciente de  $\theta$  para cada  $t$ . Por lo tanto, se deduce que

$$\begin{aligned}\theta > \theta_U(t) &\implies F_T(t|\theta) < \frac{\alpha}{2}, \\ \theta < \theta_L(t) &\implies \bar{F}_T(t|\theta) < \frac{\alpha}{2}.\end{aligned}$$

Para el intervalo  $[\theta_L(T), \theta_U(T)]$ ,

$$\begin{aligned}P_\theta(\theta_L(T) \leq \theta \leq \theta_U(T)) &= P_\theta(\theta \leq \theta_U(T)) - P_\theta(\theta < \theta_L(T)) \\ &= 1 - P_\theta\left(F_T(T|\theta) < \frac{\alpha}{2}\right) - P_\theta\left(\bar{F}_T(T|\theta) < \frac{\alpha}{2}\right).\end{aligned}$$

Ahora,  $F_T(T|\theta)$  es estocásticamente mayor que una variable aleatoria uniforme ([10], pág. 78), esto es,  $P_\theta(F_T(T|\theta) \leq x) \leq x$ . Además, éste hecho también implica que  $\bar{F}_T(T|\theta)$  es estocásticamente más grande que una variable aleatoria uniforme. Así, se tiene que

$$P_\theta\left(F_T(T|\theta) < \frac{\alpha}{2}\right) \leq \frac{\alpha}{2} \text{ y } P_\theta\left(\bar{F}_T(T|\theta) < \frac{\alpha}{2}\right) \leq \frac{\alpha}{2}.$$

Mostrando que  $[\theta_L(T), \theta_U(T)]$  es un intervalo de confianza  $1 - \alpha$ . □

# Capítulo 3

## Intervalos de confianza para una proporción

Al construir un intervalo de confianza usualmente se desea que su probabilidad de cobertura esté cercana al nivel de confianza nominal  $(1 - \alpha)$ . Debido a que la distribución binomial es una variable aleatoria discreta no es posible construir intervalos que alcancen de manera exacta un nivel de confianza específico, a no ser que se aleatorice, sin embargo las pruebas aleatorizadas no son de utilidad práctica y en este trabajo no son de nuestro interés. Algunos métodos para construir intervalos de confianza para una proporción trabajan con métodos aproximados, en especial usando propiedades de muestras grandes. Otros trabajan con estadística bayesiana y algunos más con procedimientos exactos, éstos son la inversión de la prueba binomial de colas iguales, sin embargo estos intervalos cumplen el hecho de que para cualquier  $p$  fijo, su probabilidad de cobertura es siempre mayor o igual que el nivel de confianza nominal  $(1 - \alpha)$  y en ocasiones éstas pueden estar muy cercanas a 1. En esta tesis se tratarán algunos intervalos para una proporción construidos por métodos aproximados.

### 3.1. El intervalo de Wald

Suponga que  $X_1, X_2, \dots, X_n$  es una muestra aleatoria de una población  $B(p)$ , considere probar

$$H_0 : p = p_0 \text{ vs } H_1 : p \neq p_0,$$

donde  $0 < p_0 < 1$  es un valor específico.

En el ejemplo 2.5 se dedujo, que cuando  $n \rightarrow \infty$   $\frac{\hat{p} - p_0}{\sqrt{\hat{p}(1 - \hat{p})/n}} \rightarrow N(0, 1)$ . Entonces una prueba aproximada está basada en éste estadístico por lo cual la región de rechazo es  $\left\{ \mathbf{x} : \left| \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1 - \hat{p})/n}} \right| > z_{\frac{\alpha}{2}} \right\}$ . Por lo cual  $H_0$  no se rechaza para puntos muestrales que satisfacen  $\left| \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1 - \hat{p})/n}} \right| \leq z_{\frac{\alpha}{2}}$ . Por el teorema 2.15, al invertir la región de no rechazo se obtendrá un intervalo de confianza para  $p_0$ , el cual cumple que,

$$\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1 - \hat{p})/n} \leq p_0 \leq \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1 - \hat{p})/n}. \quad (3.1)$$

En textos de introducción a la estadística, para el caso de una proporción, se presenta el intervalo de confianza mostrado en (3.1) el cual ha adquirido aceptación casi universal. Sin embargo de acuerdo con la literatura revisada, su probabilidad de cobertura comúnmente es menor de lo deseado. El procedimiento está basado en la aproximación normal y garantiza que para cualquier  $p \in (0, 1)$  su probabilidad de cobertura converge a  $1 - \alpha$  cuando  $n \rightarrow \infty$ . Este intervalo es conocido como el intervalo de Wald, ya que proviene de invertir la región de aceptación de la conocida prueba de Wald para muestras grandes.

En la expresión (3.1)  $X$  = número de éxitos en  $n$  realizaciones independientes,  $\hat{p} = \frac{X}{n}$ ,  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$  y  $\hat{q} = 1 - \hat{p}$ ,  $\Phi(z)$  es la función de distribución de una normal estándar. En ocasiones el intervalo de Wald es también llamado intervalo estándar y será denotado por  $I_E$ .

A continuación son presentados otros intervalos para una proporción que resaltan en la literatura; éstos son el intervalo de Wilson, el Intervalo de Agresti-

Coull, el intervalo arcoseno y el intervalo de Clopper-Pearson.

## 3.2. El intervalo de Wilson

Suponga que  $X_1, X_2, \dots, X_n$  es una muestra aleatoria de una población  $B(p)$ , considere probar

$$H_0 : p = p_0 \text{ vs } H_1 : p \neq p_0,$$

donde  $p_0$  es un valor fijo,  $0 < p_0 < 1$ .

Del ejemplo 2.5 se dedujo que, cuando  $n \rightarrow \infty$   $\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \rightarrow N(0, 1)$ . Entonces una prueba aproximada está basada en éste estadístico por lo cual la región de rechazo es  $\left\{ \mathbf{x} : \left| \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \right| > z_{\frac{\alpha}{2}} \right\}$ . De donde,  $H_0$  no se rechaza para puntos muestrales que satisfacen  $\left| \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \right| \leq z_{\frac{\alpha}{2}}$ , Por el teorema 2.15, al invertir la región de no rechazo se obtendrá un intervalo de confianza para  $p_0$ , manipulando ésta region, se obtiene que,

$$\begin{aligned} |X - p_0 n| &\leq \kappa \sqrt{np_0(1-p_0)}, & \text{donde } z_{\frac{\alpha}{2}} &= \kappa, \\ X^2 - 2p_0 n X &\leq -n(p_0^2(n + \kappa^2) - p_0 \kappa^2) \\ X^2 - 2p_0 n X &\leq -n \frac{p_0^2(n + \kappa^2)^2 - 2p_0^2(n + \kappa^2) \frac{\kappa^2}{2p_0} + \frac{\kappa^4}{4} - \frac{\kappa^4}{4}}{n + \kappa^2} \\ X^2 - 2p_0 n X &\leq -n \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2 - \frac{\kappa^4}{4}}{n + \kappa^2} \\ X^2 - \frac{2p_0(n + \kappa^2)}{\kappa^2 n} \left( \frac{\kappa^2 n^2}{n + \kappa^2} \right) X &\leq \frac{\kappa^2 n^2}{n + \kappa^2} \left( - \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2 - \frac{\kappa^4}{4}}{\kappa^2 n} \right) \\ X^2 \left( \frac{n + \kappa^2}{\kappa^2 n^2} \right) - \frac{2X p_0(n + \kappa^2)}{\kappa^2 n} &\leq - \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2 - \frac{\kappa^4}{4}}{\kappa^2 n} \\ \frac{X^2}{\kappa^2 n} + \frac{X^2}{n^2} - \frac{2X p_0(n + \kappa^2)}{\kappa^2 n} &\leq - \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2 - \frac{\kappa^4}{4}}{\kappa^2 n} \\ \frac{X^2 + X \kappa^2 - 2X p_0(n + \kappa^2)}{\kappa^2 n} &\leq \frac{X}{n} - \frac{X^2}{n^2} + \frac{\kappa^2}{4n} - \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2}{\kappa^2 n} \\ \frac{X^2 - 2X p_0 \left( n + \kappa^2 - \frac{\kappa^2}{2p_0} \right)}{\kappa^2 n} &\leq \frac{X}{n} - \frac{X^2}{n^2} + \frac{\kappa^2}{4n} - \frac{p_0^2(n + \kappa^2 - \frac{\kappa^2}{2p_0})^2}{\kappa^2 n} \\ \frac{\left( p_0 \left( n + \kappa^2 - \frac{\kappa^2}{2p_0} \right) - X \right)^2}{\kappa^2 n} &\leq \hat{p} \hat{q} + \frac{\kappa^2}{4n} \end{aligned}$$

$$\begin{aligned} \frac{|p_0 n + p_0 \kappa^2 - X - \frac{\kappa^2}{2}|}{\kappa n^{1/2}} &\leq (\hat{p}\hat{q} + \frac{\kappa^2}{4n})^{1/2} \\ \left| p_0 - \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} \right| &\leq \frac{\kappa n^{1/2}}{n + \kappa^2} (\hat{p}\hat{q} + \frac{\kappa^2}{4n})^{1/2} \\ \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} - \frac{\kappa n^{1/2}}{n + \kappa^2} (\hat{p}\hat{q} + \frac{\kappa^2}{4n})^{1/2} &\leq p_0 \leq \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} + \frac{\kappa n^{1/2}}{n + \kappa^2} (\hat{p}\hat{q} + \frac{\kappa^2}{4n})^{1/2}. \end{aligned} \quad (3.2)$$

El intervalo de Wilson mostrado en (3.2), se obtiene al invertir la región de aceptación de la prueba de Wald para muestras grandes sólo que en lugar de utilizar el error estimado estándar  $(\hat{p}\hat{q})^{1/2}n^{-1/2}$  se utiliza el error estándar  $(pq)^{1/2}n^{-1/2}$ . El intervalo de Wilson fue aparentemente introducido por Wilson en ([28]), donde  $X$  = número de éxitos en  $n$  realizaciones independientes,  $\hat{p} = \frac{X}{n}$ ,  $\hat{q} = 1 - \hat{p}$  y  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$  donde  $\Phi(z)$  tiene una distribución normal estándar.

El intervalo de Wilson tiene un atractivo teórico, es la inversión de la aproximación del teorema central del límite (ver teorema 2.13) a la familia de las pruebas de colas iguales bajo la hipótesis nula  $H_0 : p = p_0$ . Por lo tanto, no se rechaza  $H_0$  basado en la aproximación del teorema central del límite si y sólo si  $p_0$  está en este intervalo. Como Wilson mostró, el argumento consiste en encontrar la solución de una ecuación cuadrática; ver por ejemplo Tamhane y Dunlop en ([26], ejercicio 9.39).

### 3.3. El intervalo de Agresti-Coull

Para la presentación en los cursos básicos de estadística, sería bueno tener un intervalo alternativo que tenga una expresión familiar al intervalo de Wald (3.1), el presente intervalo usa un nuevo estimador  $\tilde{p}$  en lugar de  $\hat{p}$ . Lo anterior se logra mediante el uso del centro de la región de Wilson. Sea  $\tilde{X} = X + \frac{\kappa^2}{2}$ ,  $\tilde{n} = n + \kappa^2$ ,  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$  donde  $\Phi(z)$  tiene una distribución normal estándar. Tomando en cuenta las expresiones anteriores, si  $\tilde{p} = \frac{\tilde{X}}{\tilde{n}}$  y  $\tilde{q} = 1 - \tilde{p}$ ,

el intervalo de Agresti-Coull es

$$\left[ \tilde{p} - \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2}, \tilde{p} + \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2} \right]. \quad (3.3)$$

Para el caso de un intervalo de confianza del 95 %, se tendría que  $\kappa = 1.96$ . Si en la expresión (3.3) en lugar de utilizar  $\kappa = 1.96$  se utiliza 2, el intervalo resultante es el intervalo en Agresti y Coull en ([4]) que añade 2 éxitos y 2 fracasos al intervalo de Wald. Por esta razón el intervalo expresado en (3.3) se conoce como el intervalo Agresti-Coull. Buscando antecedentes históricos, Brown et al. en ([8]) mencionan que el libro *Statistics for the Life Sciences* ([23]), es el primer texto de introducción a la estadística que recomienda el uso de este intervalo.

### 3.4. El intervalo arco seno

Si  $X_n$  (o  $X$  por simplicidad) es una variable aleatoria binomial con parámetros  $n$  y  $p$  ( $b(n, p)$ ). De esta manera  $\frac{X}{n}$  es aproximadamente una  $N(p, p(1-p)/n)$ . Expertos en estadística a menudo buscan funciones de estadísticas cuya varianza no dependa de los parámetros. En este caso la varianza de  $\frac{X}{n}$  depende de  $p$ . ¿Se puede encontrar una función  $u$  que dependa de  $\frac{X}{n}$ , cuya varianza sea libre de  $p$ ? Como  $\frac{X}{n}$  converge en probabilidad a  $p$ , se puede aproximar  $u(\frac{X}{n})$  por los dos primeros términos de la expansión en series de Taylor alrededor de  $p$ , es decir, por

$$u\left(\frac{X}{n}\right) \doteq v\left(\frac{X}{n}\right) = u(p) + \left(\frac{X}{n} - p\right)u'(p).$$

Por supuesto,  $v(\frac{X}{n})$  es una función lineal de  $\frac{X}{n}$  y así también tiene una distribución aproximadamente normal; ésta tiene media  $u(p)$  y varianza  $[u'(p)]^2 \frac{p(1-p)}{n}$ . Sin embargo, lo último que queremos es que sea esencialmente libre de  $p$ ; así establecemos esto igual a una constante, obteniendo la ecuación diferencial

$$u'(p) = \frac{c}{\sqrt{p(1-p)}}$$

Una solución es  $u(p) = (2c)\text{arcoseno}\sqrt{p}$ .

Si tomamos  $c = \frac{1}{2}$ , se tiene, ya que  $u(\frac{X}{n})$  es aproximadamente igual a  $v(\frac{X}{n})$ , que

$$u(\frac{X}{n}) = \text{arcoseno}\sqrt{\frac{X}{n}},$$

por el método Delta ([14]), tiene una distribución aproximadamente normal con media  $\text{arcoseno}\sqrt{p}$  y varianza  $\frac{1}{4n}$ , la cual es libre de  $p$ .

Suponga que  $X_1, X_2, \dots, X_n$  es una muestra aleatoria de una población  $B(p)$ , considere probar

$$H_0 : p = p_0 \text{ vs } H_1 : p \neq p_0,$$

donde  $0 < p_0 < 1$  es un valor específico.

Anteriormente se concluyó que

$$2n^{1/2}(\text{arcoseno}\sqrt{\frac{X}{n}} - \text{arcoseno}\sqrt{p_0}) \rightarrow N(0, 1).$$

Sin embargo, Anscombe en ([5]) mostró que reemplazando  $\hat{p} = \frac{X}{n}$  por  $\check{p} = (X + 3/8)/(n + 3/4)$  da una mejor estabilización en la varianza, además

$$2n^{1/2}[\text{arcoseno}(\check{p}^{1/2}) - \text{arcoseno}(p_0^{1/2})] \rightarrow N(0, 1) \quad \text{cuando } n \rightarrow \infty.$$

Entonces una prueba aproximada se basa en éste estadístico por lo cual la región de rechazo es  $\left\{ \mathbf{x} : |2n^{1/2}(\text{arcoseno}(\check{p}^{1/2}) - \text{arcoseno}(p_0^{1/2}))| > z_{\frac{\alpha}{2}} \right\}$ . Entonces  $H_0$  no se rechaza para puntos muestrales que satisfacen

$$|2n^{1/2}(\text{arcoseno}(\check{p}^{1/2}) - \text{arcoseno}(p_0^{1/2}))| \leq z_{\frac{\alpha}{2}},$$

Por el teorema 2.15, al invertir la región de no rechazo se obtendrá un intervalo de confianza para  $p_0$ , manipulando ésta region, se obtiene que,

$$-z_{\frac{\alpha}{2}} \leq 2n^{1/2}(\text{arcoseno}(\check{p}^{1/2}) - \text{arcoseno}(p_0^{1/2})) \leq z_{\frac{\alpha}{2}}$$

$$\begin{aligned}
-\frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}} &\leq \arccoseno(\check{p}_0^{1/2}) - \arccoseno(p^{1/2}) \leq \frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}} \\
\arccoseno(\check{p}^{1/2}) - \frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}} &\leq \arccoseno(p_0^{1/2}) \leq \arccoseno(\check{p}^{1/2}) + \frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}} \\
\text{seno}^2(\arccoseno(\check{p}^{1/2}) - \frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}}) &\leq p_0 \leq \text{seno}^2(\arccoseno(\check{p}^{1/2}) + \frac{1}{2}n^{-1/2}z_{\frac{\alpha}{2}}).
\end{aligned} \tag{3.4}$$

El intervalo mostrado en (3.4) es un intervalo de confianza para  $p$  del  $100(1 - \alpha)\%$ , conocido como el intervalo arccoseno, En algunos artículos el intervalo presentado en (3.4) es llamado el intervalo arccoseno con la corrección de Anscombe ([20]).

### 3.5. El intervalo de Clopper-Pearson

Con la finalidad de evitar la aproximación muchos textos de estadística avanzada recomiendan, al intervalo de Clopper-Pearson ([13]), el cual es la inversión de la prueba binomial de colas iguales en lugar de su aproximación normal. Algunos autores se refieren a éste como un procedimiento *exacto* a causa de su derivación de la distribución binomial.

Si  $X = x$  se observa, entonces el intervalo Clopper-Pearson se define por:  $CI_{CP} = [L_{CP}(x), U_{CP}(x)]$  donde  $L_{CP}(x)$  y  $U_{CP}(x)$  son respectivamente las soluciones en  $p$  a las ecuaciones

$$P_p(X \geq x) = \frac{\alpha}{2} \text{ y } P_p(X \leq x) = \frac{\alpha}{2},$$

excepto que el límite inferior es 0 cuando  $x = 0$  y el límite superior es 1 cuando  $x = n$ . De acuerdo con Agresti y Coull ([1]), cuando  $x = 1, \dots, n-1$ , el intervalo de confianza de Clopper-Pearson equivale a

$$\left[1 + \frac{n-x+1}{xF_{2x, 2(n-x+1), 1-\frac{\alpha}{2}}}\right]^{-1} < p < \left[1 + \frac{n-x}{(x+1)F_{2(x+1), 2(n-x), \frac{\alpha}{2}}}\right],$$

donde  $F_{a,b,c}$  denota el  $1 - c$  de la distribución  $F$  con grados de libertad  $a$  y  $b$ .

Otra forma de presentar al intervalo de Clopper-Pearson se obtiene al mostrar que el punto final inferior es el  $\frac{\alpha}{2}$  cuantil de una distribución beta  $Beta(x, n - x + 1)$ , y el punto final superior es el  $1 - \frac{\alpha}{2}$  cuantil de una distribución beta  $Beta(x + 1, n - x)$ .

El intervalo de Clopper-Pearson garantiza que la probabilidad de cobertura es siempre mayor o igual que el nivel de confianza nominal por lo cual es conservador. Sin embargo, para cualquier  $p$  fijo, la probabilidad de cobertura puede ser mucho más grande que  $(1 - \alpha)$  a menos que  $n$  sea suficientemente grande, la figura 3.1 en ([8]) muestra el desempeño del intervalo de Clopper-Pearson para  $n = 50$  y  $p \in (0, 1)$ .

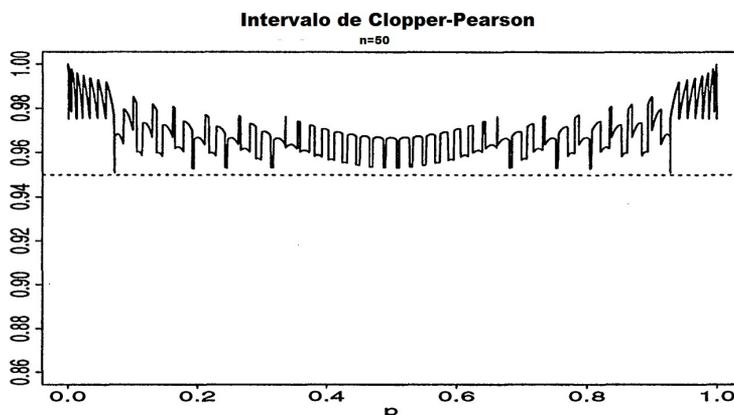


Figura 3.1: Intervalo de Clopper-Pearson para  $n = 50$  y  $p \in (0, 1)$ .

El intervalo de Clopper-Pearson es bastante conservador y no es una buena opción para el uso práctico, a menos que se requiera que la probabilidad de cobertura sea mayor que el nivel de confianza nominal, además de acuerdo con Måns Thulin en ([18]) el intervalo de Clopper-Pearson tiene una longitud esperada mayor que los intervalos de Wald, Wilson y Agresti-Coull. Por lo cual el intervalo de Clopper-Pearson no se considerará en el estudio realizado en la presente tesis.

# Capítulo 4

## Comparación de los intervalos de confianza

En el capítulo 3 se presentaron algunos intervalos de confianza para una proporción, ahora se procederá a analizar el desempeño de tales intervalos haciendo énfasis principalmente en el intervalo de Wald debido a que es el intervalo de confianza para una proporción más usado en estadística.

### 4.1. Probabilidad de cobertura del intervalo de Wald.

Debido a la naturaleza discreta de la variable aleatoria es imposible encontrar intervalos de confianza tal que su probabilidad de cobertura sea igual al coeficiente de confianza  $1 - \alpha$ , sin embargo se esperaría que la probabilidad de cobertura se encuentre cercana a  $1 - \alpha$ , en otras palabras que el valor  $|\mathbf{PC} - (1 - \alpha)|$  sea pequeño.

La probabilidad de cobertura para  $n$  y  $p$  será denotada por  $\mathbf{PC}(n, p)$ . Por la definición (2.18) la probabilidad de cobertura de un intervalo de confianza

para  $p$  con  $p \in (0, 1)$  se define como;

$$P_p(p \in [L(X), U(X)])$$

de la expresión (3.1) se tiene que la probabilidad de cobertura del intervalo de Wald es

$$P_p\left(\hat{p} - \kappa n^{-1/2}(\hat{p}\hat{q})^{1/2} \leq p \leq \hat{p} + \kappa n^{-1/2}(\hat{p}\hat{q})^{1/2}\right), \quad (4.1)$$

observar que la probabilidad de cobertura depende de  $\alpha$ ,  $n$  y  $X$ .

La proposición 4.1 evalúa el desempeño en términos de la probabilidad de cobertura del intervalo de Wald para  $\alpha$ ,  $n$  y  $p$ .

**Proposición 4.1.** *Para  $\alpha$ ,  $n$  y  $p$ , la probabilidad de cobertura del intervalo de Wald puede ser expresada como,*

$$P_p\left[\frac{n}{1+c}\left(p+\frac{c}{2}-\sqrt{-cp^2+cp+\frac{c^2}{4}}\right) \leq X \leq \frac{n}{1+c}\left(p+\frac{c}{2}+\sqrt{-cp^2+cp+\frac{c^2}{4}}\right)\right] \quad (4.2)$$

donde  $c = \frac{k^2}{n}$  y  $X \sim Bin(n, p)$ .

*Demostración.*

$$\begin{aligned} & P_p(\hat{p} - \kappa n^{-1/2}(\hat{p}\hat{q})^{1/2} \leq p \leq \hat{p} + \kappa n^{-1/2}(\hat{p}\hat{q})^{1/2}) \\ &= P_p(|\hat{p} - p| \leq \kappa n^{-1/2}(\hat{p}\hat{q})^{1/2}) = P_p((\hat{p} - p)^2 \leq \kappa^2 n^{-1}(\hat{p}\hat{q})). \end{aligned}$$

Si  $c = \kappa^2 n^{-1} = \frac{\kappa^2}{n}$ , entonces la última expresión se puede escribir como,

$$\begin{aligned} & P_p(\hat{p}^2 - 2\hat{p}p + p^2 \leq c\hat{p} - c\hat{p}^2) = P_p((1+c)\hat{p}^2 - (2p+c)\hat{p} \leq -p^2) \\ &= P_p\left((1+c)[\hat{p}^2 - \frac{2p+c}{1+c}\hat{p}] \leq -p^2\right) \\ &= P_p\left(\hat{p}^2 - \frac{2p+c}{1+c}\hat{p} \leq \frac{-p^2}{(1+c)}\right) = P_p\left((\hat{p} - \frac{p+\frac{c}{2}}{1+c})^2 \leq \frac{-p^2}{(1+c)} + (\frac{p+\frac{c}{2}}{1+c})^2\right) \\ &= P_p\left(|\hat{p} - \frac{p+\frac{c}{2}}{1+c}| \leq \sqrt{\frac{-p^2}{(1+c)} + (\frac{p+\frac{c}{2}}{1+c})^2}\right) = P_p\left(|\hat{p} - \frac{p+\frac{c}{2}}{1+c}| \leq \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) \\ &= P_p\left(-\frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c} \leq \hat{p} - \frac{p+\frac{c}{2}}{1+c} \leq \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) \\ &= P_p\left(\frac{p+\frac{c}{2}}{1+c} - \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c} \leq \hat{p} \leq \frac{p+\frac{c}{2}}{1+c} + \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) \end{aligned}$$

$$\begin{aligned}
&= P_p \left( \frac{p + \frac{c}{2}}{1+c} - \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \leq \frac{X}{n} \leq \frac{p + \frac{c}{2}}{1+c} + \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \right) \\
&= P_p \left( n \left( \frac{p + \frac{c}{2}}{1+c} - \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \right) \leq X \leq n \left( \frac{p + \frac{c}{2}}{1+c} + \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \right) \right). \quad \square
\end{aligned}$$

La probabilidad de cobertura del intervalo de Wald para  $n$  y  $p$ , se puede escribir de forma compacta como sigue,

$$\mathbf{PC}(n, p) = \sum_{x=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} \binom{n}{x} p^x (1-p)^{n-x}, \quad (4.3)$$

donde  $x_1 = \frac{n}{1+c} \left( p + \frac{c}{2} - \sqrt{-cp^2 + cp + \frac{c^2}{4}} \right)$ ,  $x_2 = \frac{n}{1+c} \left( p + \frac{c}{2} + \sqrt{-cp^2 + cp + \frac{c^2}{4}} \right)$ ,  $c = \frac{k^2}{n}$  y  $\lceil x_1 \rceil$ ,  $\lfloor x_2 \rfloor$  son las funciones techo y piso, respectivamente.

Por ejemplo si  $X$  se distribuye de forma binomial con parámetros  $p = 0.4$  y  $n = 100$  y consideramos a  $\alpha = 0.05$ , entonces  $\lfloor x_2 \rfloor = 49$  y  $\lceil x_1 \rceil = 31$  y usando (4.3) se tiene que  $\mathbf{PC}(100, 0.4) = .9481$ .

#### 4.1.1. Desempeño del intervalo de Wald

Las figuras y cuadros que se presentan en este capítulo son el resultado de cálculos realizados mediante un programa escrito por el autor en el lenguaje de programación R.

Con la finalidad de analizar el comportamiento del intervalo de Wald se revisaron algunos artículos relacionados con el tema, por ejemplo ([4], [8] y [9]), a continuación, se presentan los casos en los cuales el intervalo de Wald es demasiado inconsistente en términos de su probabilidad de cobertura.

Usualmente en los problemas sobre control de calidad se desea que el porcentaje de artículos defectuosos de un producto sea pequeño, así que no es extraña la necesidad del cálculo de intervalos de confianza para tales  $p$ , y por tanto es importante analizar el comportamiento de la probabilidad de cobertura del intervalo de Wald para tales  $p$ .

**Ejemplo 4.1.** [Comportamiento del intervalo de Wald para  $p$  cercano a 0]  
 La figura 4.1 presenta las probabilidades de cobertura para el intervalo de Wald del 95 % de confianza nominal para  $p = 0.001$  y  $n \in \llbracket 1, 10000 \rrbracket$ .

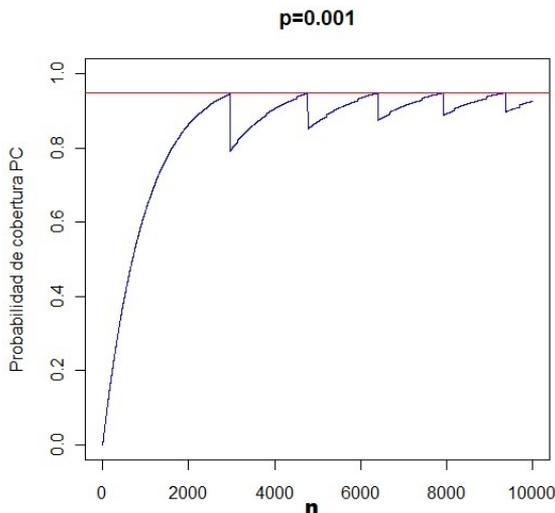


Figura 4.1: Probabilidad de cobertura,  $n \in \llbracket 1, 10000 \rrbracket$  y  $p = 0.001$ .

Se puede observar de la gráfica que la oscilación es considerable y que la probabilidad de cobertura se acerca al coeficiente de confianza nominal (0.95) pero no de manera monótona, también se observa que para que la  $PC \geq 0.92$  se tendría que tomar un tamaño de muestra bastante grande ya que por ejemplo para  $n=2000$  la  $PC(2000, 0.001) = 0.8637$ . La gráfica que representa la probabilidad de cobertura comienza en 0 y conforme aumenta el valor de  $n$  también aumenta la probabilidad de cobertura, este proceso continúa hasta llegar a  $n = 2959$  donde la probabilidad de cobertura es 0.9447 pero cae de inmediato hasta 0.7915 en  $n = 2960$ , nuevamente la probabilidad de cobertura aumenta de manera monótona hasta  $n = 4771$  donde su probabilidad de cobertura es 0.9474 pero cae de inmediato hasta 0.8511 en  $n = 4772$ . Se puede apreciar de lo anterior que la probabilidad de cobertura del intervalo de Wald es deficiente para  $p = 0.001$ , aún con tamaños de muestra bastante grandes. Otros

puntos  $(n, 0.001)$  donde su probabilidad de cobertura está muy por debajo de 0.95 son  $n = 100$  donde  $\mathbf{PC}(100, 0.001)$  es solamente 0.0952, en  $n = 1000$  apenas alcanza 0.6317 y como se ha señalado para  $n = 2000$  es 0.8637, un comportamiento similar se muestra en ([8], pág. 105).

De manera análoga, el comportamiento del intervalo de Wald es errático en términos de probabilidad de cobertura para tamaños de muestra pequeños ([8]), un primer acercamiento a esta cuestión se observa en la figura 4.1, sin embargo esta situación se mantiene aún para  $p$  no cercanos a 0 o 1. Un ejemplo para evidenciar esta situación se muestra en la figura 4 de Brown et al. en ([8]), que se reproduce en el siguiente ejemplo.

**Ejemplo 4.2** (Comportamiento del intervalo de Wald para tamaños de muestra pequeños).

Se presenta la probabilidad de cobertura del intervalo de Wald del 99 % de confianza nominal para  $n = 20$  y  $p \in \{0.001, 0.002, \dots, 0.999\}$ .

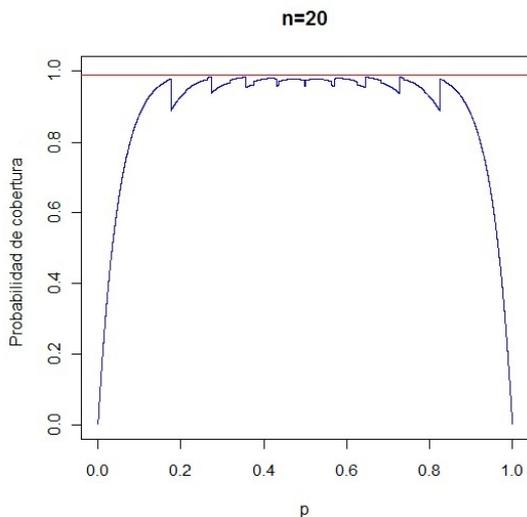


Figura 4.2: Probabilidad de cobertura del intervalo de Wald,  $n = 20$ .

Se observa que no hay un solo valor de  $p$  ( $p \in \{0.001, 0.002, \dots, 0.999\}$ ) para

el cual  $PC(n,p) \geq 0.99$ , la probabilidad de cobertura más cercana a 0.99 se alcanza en  $p = 0.272$  (por simetría también en  $p = 0.728$ ) y ésta es 0.9831, el valor promedio de las probabilidades de cobertura es 0.8835.

Debido a que los principales problemas del intervalo de Wald surgen cuando  $p$  está cerca de 0 o de 1 y cuando el tamaño muestral  $n$  es pequeño ([1], [2], [4], [8], [9], [17] y [19]), un usuario desprevenido pensaría que el intervalo de Wald tendría un buen desempeño en términos de su probabilidad de cobertura si  $p$  no está cerca de 0 o de 1 o si  $n$  no es pequeño, estos casos se analizarán en los siguientes ejemplos.

### Ejemplo 4.3.

La figura 4.3 muestra las probabilidades de cobertura del intervalo de Wald del 95 % de confianza nominal para  $n = 100$  y  $p \in \{0.001, 0.002, \dots, 0.999\}$ .

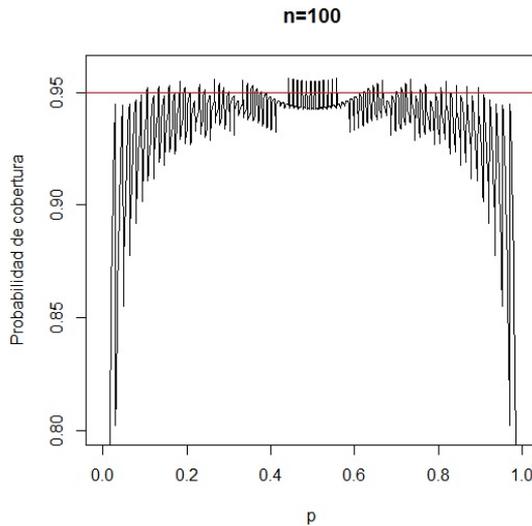


Figura 4.3: Probabilidad de cobertura,  $p \in \{0.001, 0.002, \dots, 0.999\}$  y  $n = 100$ .

Se observa nuevamente que las probabilidades de cobertura para  $n = 100$  y  $p \in \{0.001, 0.002, 0.999\}$  en su mayoría son menores que 0.95 y su comportamiento es bastante inestable, dicho de otra forma, existen saltos en las

probabilidades de cobertura para valores de  $p$  cercanos y consecutivos, estos saltos son más evidentes cuando  $p$  está cercano a 0 o bien a 1, más aún, se siguen presentando probabilidades de cobertura pequeñas para  $p$  cercanos a 0 o bien a 1, por ejemplo en  $p = 0.001$  la probabilidad de cobertura es solamente 0.0952. La probabilidad de cobertura solo es razonable para valores de  $p$  cercanos a 0.5.

En el ejemplo anterior se observó que el intervalo de Wald tiene un desempeño aceptable en términos de su probabilidad de cobertura cuando  $p$  está cerca de 0.5, sin embargo esto no es del todo cierto ya que las inconsistencias y pobres probabilidades de cobertura siguen presentes para tales valores de  $p$ , para ver esto consideremos el siguiente ejemplo presentado en ([9], pág. 163).

**Ejemplo 4.4.** La figura 4.4 muestra el comportamiento de la probabilidad de cobertura del intervalo de Wald del 95 % de confianza nominal para  $p = 0.5$  y  $n \in \llbracket 10, 100 \rrbracket$ .

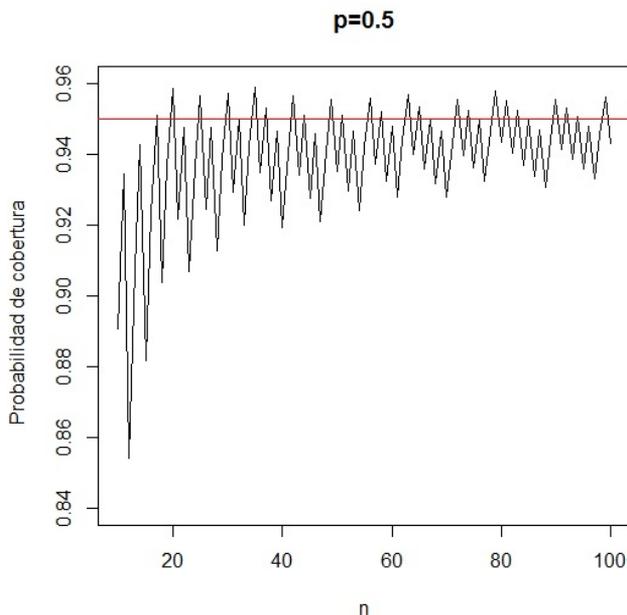


Figura 4.4: Probabilidad de cobertura,  $p = 0.5$  y  $n \in \llbracket 10, 100 \rrbracket$ .

Se observa que el intervalo de Wald mejora notablemente cuando se considera  $p = 0.5$  y que el grado de oscilación disminuye aunque sigue presente, la probabilidad de cobertura se acerca al nivel de confianza pero no de manera monótona, por ejemplo en  $n = 17$ ,  $\mathbf{PC}(17, 0.5) = 0.951$  y en  $n = 40$ ,  $\mathbf{PC}(40, 0.5) = 0.9193$ . Para  $10 \leq n \leq 100$  se cumple que  $\mathbf{PC}(n, 0.5) \geq 0.854$ , para  $p = 0.5$  aún continúan presentándose probabilidades de cobertura pequeñas para valores de  $n$  pequeños, por ejemplo  $\mathbf{PC}(12, 0.5) = 0.854$  y  $\mathbf{PC}(15, 0.5) = 0.8815$ . De acuerdo con Brown et al. en ([9]) para que  $\mathbf{PC}(n, 0.5) \geq 0.94$  se tendría que tomar  $n \geq 194$ .

Como se pudo observar de los ejemplos 4.3 y 4.4 la probabilidad de cobertura del intervalo de Wald continúa siendo inconsistente para valores de  $p$  no cercanos a 0 o a 1 así también cuando el tamaño de muestra  $n$  no es pequeño, lo anterior ha sido señalado en diversos artículos, por ejemplo ([1]), ([8]) y ([9]).

Además de la probabilidad de cobertura otro aspecto importante en la evaluación del desempeño de un intervalo de confianza es su longitud esperada ( $\mathcal{LE}$ ), un buen intervalo de confianza deberá ser estrecho. La longitud esperada para  $n$  y  $p$  está definida de la siguiente manera

$$\mathcal{LE}(n, p) = \sum_{x=0}^n (U(x, n) - L(x, n)) \binom{n}{x} p^x (1-p)^{n-x}. \quad (4.4)$$

Donde  $U(x, n)$  y  $L(x, n)$  son los límites superior e inferior del intervalo de confianza, respectivamente. La *longitud media esperada* ( $\mathcal{LME}$ ) del intervalo de confianza es

$$\mathcal{LME} = \int_0^1 \left( \sum_{x=0}^n (U(x, n) - L(x, n)) \binom{n}{x} p^x (1-p)^{n-x} \right) dp. \quad (4.5)$$

**Ejemplo 4.5** (Longitud esperada del intervalo de Wald del 95% para  $n = 30$  y  $p \in \{0.001, 0.002, \dots, 0.999\}$ ).

Se puede observar de la figura 4.5 que cuando  $p$  está alejado de los extremos

del intervalo  $(0,1)$  la longitud esperada es grande, también se observa que ésta converge a cero cuando  $p$  tiende a 0 o 1. Se podría pensar que por lo anterior el desempeño del intervalo de Wald es adecuado para  $p$  cercanos a 0 o 1, sin embargo no es así, debido a su pobre desempeño con respecto a la probabilidad de cobertura para éstos valores de  $p$  (ejemplo 4.1). La longitud media esperada para  $p \in \{0.001, 0.002, \dots, 0.999\}$  es 0.2705.

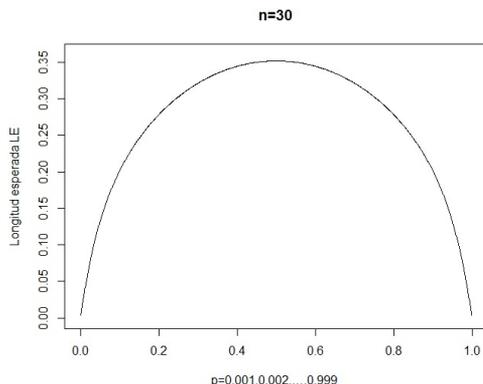


Figura 4.5: Longitud esperada cuando  $n = 30$  y  $p \in \{0.001, 0.002, \dots, 0.999\}$ .

Debido a que el intervalo de Wald presenta un mal desempeño en términos de su probabilidad de cobertura en  $p$  cercanos a 0 o a 1 y de manera similar cuando el tamaño de muestra  $n$  es pequeño y éste no desaparece cuando se consideran valores cercanos a  $p = 0.5$  ni cuando  $n$  es "suficientemente grande", diversos textos populares en la literatura acompañan al intervalo de Wald con condiciones para su uso. De una muestra de 11 textos populares sobre estadística, Brown et al. en ([8]) obtuvieron una lista de tales condiciones, otros investigadores como Leemis y Trivedy en ([17]) realizaron algo similar. En éste análisis se considerarán las condiciones que enlistan Brown et al. en ([8]).

El intervalo de confianza de Wald se puede utilizar si:

1.  $np \geq 5$  y  $n(1 - p) \geq 5$  o  $np \geq 10$  y  $n(1 - p) \geq 10$ ;

2.  $np(1 - p) \geq 5$  o  $np(1 - p) \geq 10$ ;
3.  $n\hat{p} \geq 5$  y  $n(1 - \hat{p}) \geq 5$  o  $n\hat{p} \geq 10$  y  $n(1 - \hat{p}) \geq 10$ ;
4.  $\hat{p} \pm 3\sqrt{\hat{p}(1 - \hat{p})/n}$  no contiene al 0 o al 1;
5.  $n$  bastante grande;
6.  $n \geq 50$  a menos que  $p$  sea muy pequeño.

Se observará más adelante que estos criterios no son adecuados para el intervalo de Wald en varios aspectos ([8] proposición 1 pág. 106 y [9] tabla 1 pág. 164), sin embargo dado que el intervalo de Wald aún sigue siendo el intervalo de confianza para el parámetro binomial más usado en niveles elementales se estudiará su comportamiento con base en ellos. Para esto sea  $\mathcal{P} = \{0.001, 0.002, \dots, 0.99\}$ ,  $p \in \mathcal{P}$  y  $n$  tomará los valores que se indique, se calculará y graficará la respectiva probabilidad de cobertura para los puntos de la forma  $(n, p)$  cuando se cumplan los criterios siguientes;

- **Criterio 1**  $np \geq 5$  y  $n(1 - p) \geq 5$  ;
- **Criterio 2**  $np \geq 10$  y  $n(1 - p) \geq 10$ ;
- **Criterio 3**  $np(1 - p) \geq 5$ ;
- **Criterio 4**  $np(1 - p) \geq 10$ ;
- **Criterio 5**  $n \geq 100$ ;
- **Criterio 6**  $n \geq 50$  y  $0.2 < p < 0.8$ .

Con la finalidad de analizar las condiciones que enlistan Brown et al. en ([8]), se desglosó la condición 1 en el criterio 1 y 2, la condición 2 en el criterio 3 y 4, para la condición 5 consideramos que  $n = 100$  es "bastante grande" dando origen al criterio 5 y para el criterio 6 se consideran los valores de  $p$

cuando  $0.2 < p < 0.8$  dando origen al criterio 6, es importante notar que las condiciones 3 y 4 son verificables, pero son inútiles porque en el contexto de las probabilidades de cobertura frecuentista, una condición a base de datos no tiene un significado.

## 4.2. Factores de comparación de un intervalo de confianza

En esta sección se tomarán en cuenta los 6 criterios que fueron presentados al final de la sección anterior y con base en ellos se definen los siguientes factores, que nos servirán para analizar el desempeño de un intervalo de confianza.

### Porcentaje de puntos adecuados.

Sea  $\mathbb{N}$  el conjunto de los números naturales,  $\mathcal{N} = \llbracket k, k + m \rrbracket$  con  $k, m \in \mathbb{N}$  y  $\mathcal{P} = \{0.001, 0.002, \dots, 0.999\}$ . Para  $n \in \mathcal{N}$  y  $p \in \mathcal{P}$ , se definen los siguientes conjuntos.

$$\mathcal{M}_i = \{(n, p) \in \mathcal{N} \times \mathcal{P} \mid (n, p) \text{ cumple el criterio } i\}$$

$$\mathcal{A}_i = \{(n, p) \in \mathcal{M}_i \mid \mathbf{PC}(n, p) \geq 1 - \alpha\}, \quad (4.6)$$

donde  $\alpha$  es un número real entre 0 y 1.

Entonces el porcentaje de puntos  $(n, p) \in \mathcal{N} \times \mathcal{P}$  tal que cumplen el criterio  $i$  y la  $\mathbf{PC}(n, p) \geq 0.95$  será llamado el *porcentaje de puntos adecuados* y será denotado como  $\mathcal{PA}_i$ , así

$$\mathcal{PA}_i = 100 * \frac{\text{Card}(\mathcal{A}_i)}{\text{Card}(\mathcal{M}_i)}. \quad (4.7)$$

Al momento de calcular la probabilidad de cobertura para  $n$  y  $p$ , se observa que, en general:

$$|(1 - \alpha) - \mathbf{PC}(n, p)| \neq 0,$$

por esta razón se consideran dos tipos de errores en la cobertura, el primero está definido de la siguiente manera.

### **Error de cobertura por defecto.**

Sea

$$\mathcal{I}_i = \{(n, p) \in \mathcal{M}_i \mid \mathbf{PC}(n, p) < 1 - \alpha\}. \quad (4.8)$$

Si  $(n, p) \in \mathcal{I}_i$ , el *error de cobertura por defecto* de  $(n, p)$  cuando  $(n, p)$  cumple el criterio  $i$  es

$$\mathcal{ECD}_i(n, p) = (1 - \alpha) - \mathbf{PC}(n, p). \quad (4.9)$$

Generalmente, al fijar  $n$  y variar  $p$  surgen una gran cantidad de puntos  $(n, p)$  tal que la  $\mathbf{PC}(n, p) < (1 - \alpha)$  lo mismo ocurre si se fija  $p$  y se varía  $n$ , entonces en ambos casos se obtienen varios *errores de cobertura por defecto* (uno para cada  $(n, p) \in \mathcal{I}_i$ ). Con esta información podemos calcular el valor promedio de los *errores de cobertura por defecto*, esto con la intención de tener otra medida de la efectividad de un intervalo de confianza. En caso de que éste valor sea pequeño implicaría que, el intervalo de confianza tiene un buen comportamiento ya que aunque existan puntos  $(n, p)$  tal que  $\mathbf{PC}(n, p) < (1 - \alpha)$  las probabilidades de cobertura estarían en promedio cercanas a  $1 - \alpha$ , lo cual es deseable.

### **Error de cobertura promedio por defecto.**

Si  $(n, p) \in \mathcal{I}_i$ , entonces el *error de cobertura promedio por defecto* para el criterio  $i$  es,

$$\mathcal{ECPD}_i = \frac{\sum_{(n,p) \in \mathcal{I}_i} ((1 - \alpha) - \mathbf{PC}(n, p))}{\text{Card}(\mathcal{I}_i)}. \quad (4.10)$$

En algunas ocasiones para hacer más sencilla la comparación en lugar del *error de probabilidad de cobertura promedio por defecto*, se hablará de la *probabilidad de cobertura promedio por defecto* y ésta se calcula de la siguiente forma.

### **Probabilidad de cobertura promedio por defecto**

Si  $(n, p) \in \mathcal{I}_i$ , entonces la *probabilidad de cobertura promedio por defecto* para el criterio  $i$  es:

$$\mathcal{PCPD}_i = (1 - \alpha) - \mathcal{ECPD}_i. \quad (4.11)$$

El segundo tipo de error en la cobertura, está definido de la siguiente manera.

### **Error de cobertura por exceso ( $\mathcal{ECE}$ ).**

Recordar que  $\mathcal{M}_i = \{(n, p) \in \mathcal{N} \times \mathcal{P} \mid (n, p) \text{ cumple el criterio } i\}$  y  $\mathcal{A}_i = \{(n, p) \in \mathcal{M}_i \mid \mathbf{PC}(n, p) > 1 - \alpha\}$  entonces.

Si  $(n, p) \in \mathcal{A}_i$ , el *error de cobertura por exceso* de  $(n, p)$  cuando  $(n, p)$  cumple el criterio  $i$  es

$$\mathcal{ECE}_i(n, p) = \mathbf{PC}(n, p) - (1 - \alpha). \quad (4.12)$$

De manera similar, al fijar  $n$  y variar  $p$  surgen una gran cantidad de puntos  $(n, p)$  tal que la  $\mathbf{PC}(n, p) > (1 - \alpha)$  lo mismo ocurre si se fija  $p$  y se varía  $n$ , entonces tenemos varios *errores de cobertura por exceso* (uno para cada  $(n, p) \in \mathcal{A}_i$ ). Con esta información podemos calcular el valor promedio de los *errores de cobertura por exceso*, esto con la intención de tener una medida más de la efectividad de un intervalo de confianza. En caso de que este valor sea pequeño implicaría que, el intervalo de confianza tiene un buen comportamiento ya que aunque existan puntos  $(n, p)$  tal que  $\mathbf{PC}(n, p) > (1 - \alpha)$  éstas estarían en promedio cercanas a  $1 - \alpha$ , lo cual también es deseable.

### **Error de cobertura promedio por exceso.**

Si  $(n, p) \in \mathcal{A}_i$ , entonces el *error de cobertura promedio por exceso* para el criterio  $i$  es

$$\mathcal{ECP\mathcal{E}}_i = \frac{\sum_{(n,p) \in \mathcal{A}_i} (\mathbf{PC}(n, p)) - (1 - \alpha)}{\text{Card}(\mathcal{A}_i)}. \quad (4.13)$$

En algunas ocasiones para hacer más sencilla la comparación en lugar del *error de probabilidad de cobertura promedio por exceso*, se hablará de la *probabilidad de cobertura promedio por exceso* y ésta se calcula de la siguiente forma.

### **Probabilidad de cobertura promedio por exceso.**

Si  $(n, p) \in \mathcal{A}_i$  entonces la *probabilidad de cobertura promedio por exceso* para el criterio  $i$  es:

$$\mathcal{PCPE}_i = \mathcal{ECPPE}_i + (1 - \alpha). \quad (4.14)$$

Otro aspecto por medio del cual se analizará el desempeño de un intervalo de confianza será su longitud media esperada (denotada por  $\mathcal{LME}$ ), definida anteriormente, sin embargo en tal definición se considera a  $p \in (0, 1)$  mientras que en este análisis solo consideramos a  $p \in \mathcal{P}$ .

### Longitud media esperada

Si  $(n, p) \in \mathcal{M}_i$ , la *longitud media esperada* para el criterio  $i$  es,

$$\mathcal{LME}_i = \frac{\sum_{(n,p) \in \mathcal{M}_i} (\mathcal{LE}(n, p))}{\text{Card}(\mathcal{M}_i)}. \quad (4.15)$$

Es importante observar que Agresti y Min en ([3]) concluyen que al usar variables discretas se obtienen comportamientos inesperados en los intervalos de confianza y esto es independiente del método usado para construirlos. Por esta razón en la tesis se considerará que un intervalo de confianza tiene un buen desempeño cuando se cumplan las dos condiciones siguientes:

- Cuando el intervalo conceda entre todos los intervalos en consideración, los  $\mathcal{ECPPE}$  y  $\mathcal{ECPD}$  menores, de manera similar, cuando otorgue las  $\mathcal{PCPE}$  y  $\mathcal{PCPD}$  mas cercanas al nivel de confianza nominal.
- Cuando el intervalo conceda entre todos los intervalos en consideración, la  $\mathcal{LME}$  menor.

### 4.3. Comportamiento bajo los criterios más usuales.

Los criterios han sido sugeridos con la finalidad de que el intervalo de Wald mejore su desempeño de una forma notable, nos interesaría saber cuál es el grado de esa mejoría y por esta razón en la presente sección se analizará el desempeño del intervalo de Wald sujeto a tales criterios, haciendo uso de los factores de comparación definidas en la sección anterior.

Como un primer acercamiento al análisis del desempeño del intervalo de Wald con base en los criterios, veamos el caso mostrado en el ejemplo 4.4. En tal ejemplo  $p = 0.5$  por lo que el criterio 1 se satisface a partir de  $n = 10$  pero  $\mathbf{PC}(12, 0.5) = 0.8540$  y  $\mathbf{PC}(15, 0.5) = 0.8815$ , el criterio 2 se satisface a partir de  $n = 20$  pero  $\mathbf{PC}(23, 0.5) = 0.9068$  y  $\mathbf{PC}(28, 0.5) = 0.9128$ , el criterio 3 se satisface a partir de  $n = 11$  pero los valores  $n = 12$  y  $n = 15$  dan probabilidades de cobertura deficientes, el criterio 4 se satisface a partir de  $n = 22$  pero nuevamente para  $n = 23$  y  $n = 28$  se obtienen probabilidades de cobertura deficientes. Para el criterio 5 encontramos que  $\mathbf{PC}(106, 0.5) = 0.9355$  y  $\mathbf{PC}(117, 0.5) = 0.936$ , para el criterio 6 se satisface que  $0.2 < p < 0.8$  y encontramos que  $\mathbf{PC}(54, 0.5) = 0.9241$  y  $\mathbf{PC}(61, 0.5) = 0.9278$ .

Las siguientes figuras así como los cuadros fueron realizados usando la expresión (4.3) y el coeficiente de confianza nominal 0.95, es decir, el valor de  $\alpha$  es 0.05. La figura 4.6 presenta el comportamiento de la probabilidad de cobertura del intervalo de Wald con base en los criterios definidos, para  $n = 100$  y  $p \in \mathcal{P}$ . Los puntos en negro representan las probabilidades de cobertura cuando  $(100, p)$  cumple el criterio establecido  $i$  con  $i = 1, 2, \dots, 6$  y los puntos en azul cuando no lo cumple.

Como  $n = 100$  el criterio 1 se cumpliría para  $0.05 \leq p \leq 0.95$  y como se observa de la figura 4.6 las probabilidades de cobertura en su mayoría están

muy por debajo de 0.95, el criterio 2 se cumple para  $0.1 \leq p \leq 0.9$  y su desempeño en este criterio es similar al del criterio 1. El criterio 3 se cumpliría para  $0.06 \leq p \leq 0.94$ , el criterio 4 para  $0.12 \leq p \leq 0.88$  y se observa que el intervalo de Wald continúa teniendo un pobre desempeño. El criterio 5 se cumple para todo  $p$  y el criterio 6 para  $0.2 < p < 0.8$ .

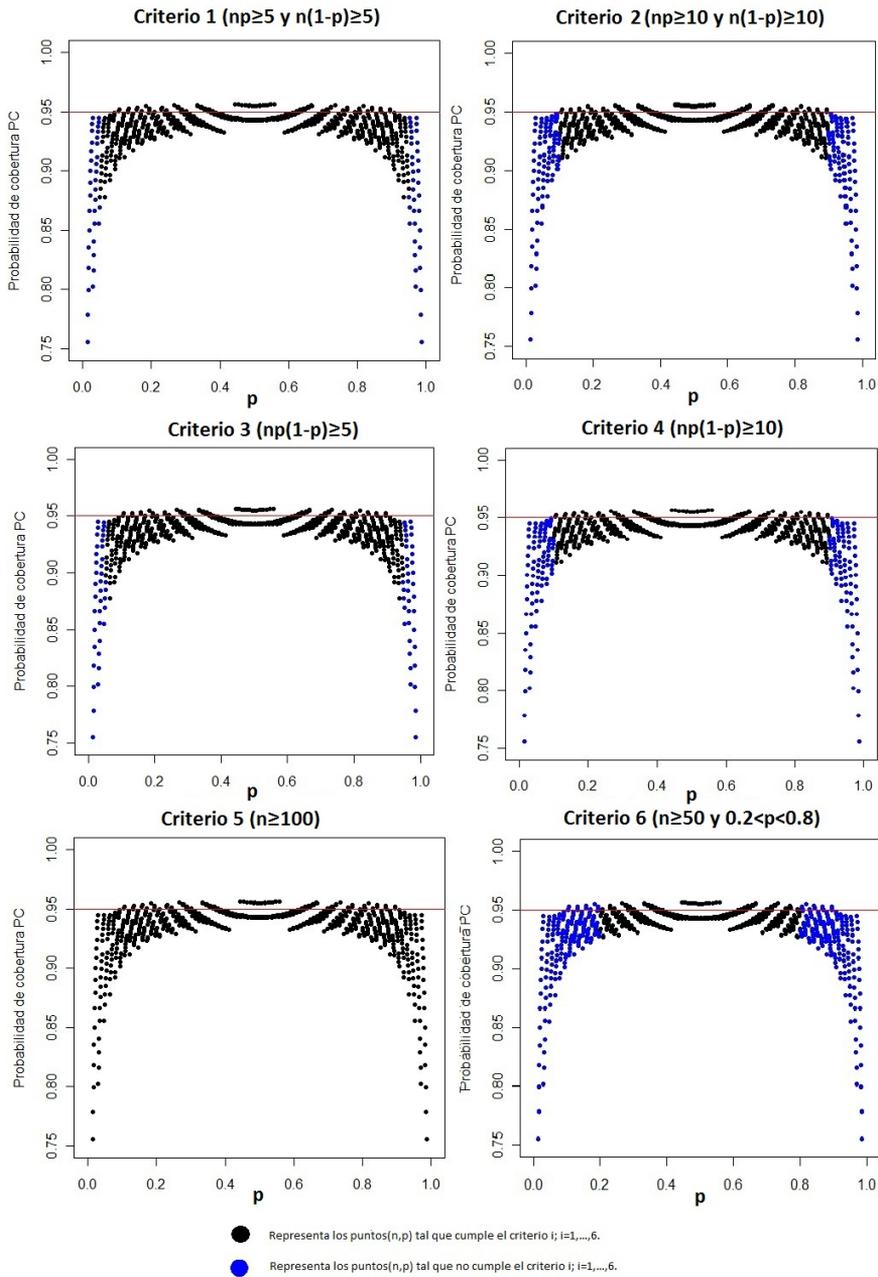


Figura 4.6: Probabilidades de cobertura del intervalo de Wald con  $n = 100$  y  $p \in \mathcal{P}$ , con base en los criterios.

También se observa que los criterios 1, 2, 3, 4 y 6 eliminan algunos puntos  $(100, p)$  con  $p \in \mathcal{P}$  en los cuales la probabilidad de cobertura es extremadamente irregular y en algunos casos cercana a 0. Se observa también que las probabilidades de cobertura cuando  $(100, p)$  cumple algún criterio continúan situándose en su mayoría por debajo del coeficiente de confianza (0.95). Esto ejemplifica que sí en efecto, los criterios mejoran el desempeño del intervalo de Wald pero no de la manera esperada.

Los factores definidas en el sección 4.2 nos ayudarán a tener una mejor apreciación y estos son mostrados en el cuadro 4.1.

Factor \ Criterio (i)	1	2	3	4	5	6
$\mathcal{PA}$	13.11 %	14.75 %	13.18 %	14.71 %	11.81 %	16.03 %
$\mathcal{ECPD}$	0.0121	0.0095	0.0118	0.0091	0.0308	0.0075
$\mathcal{ECPE}$	0.0028	0.0028	0.0028	0.0029	0.0028	0.0030
$\mathcal{LME}$	0.1634	0.1712	0.1638	0.1729	0.1524	0.1825

Cuadro 4.1: Factores de comparación del intervalo de Wald para  $p \in \mathcal{P}$  y  $n = 100$ .

Del cuadro 4.1 se observa que para los criterios bajo análisis, el intervalo de Wald presenta porcentajes de puntos adecuados ( $\mathcal{PA}$ ) pequeños, siendo mayor el obtenido por el criterio 6 (no es extraño esto, ya que en los  $p$  no cercanos a 0 o 1 es donde el intervalo de Wald tiene un mejor desempeño) pero solo alcanza el 16.03 %. Los errores de cobertura promedio por defecto ( $\mathcal{ECPD}$ ) no son tan pequeños cómo se esperaría, siendo el más pequeño el obtenido también por el criterio 6 el cual es 0.0075. Los errores de cobertura promedio por defecto al ser grandes proporcionan probabilidades de cobertura promedio por defecto ( $\mathcal{PCPD}_i = 0.95 - \mathcal{ECPD}_i$ ) inferiores al nivel de confianza (0.95), para este caso el valor más cercano a 0.95 se obtiene con el criterio 6, el cual logra una probabilidad de cobertura promedio por defecto de 0.9425. Los errores

de cobertura promedio por exceso ( $\mathcal{ECP}\mathcal{E}$ ) son pequeños y esto implica que la probabilidad de cobertura para  $(n, p)$  cuando  $\mathbf{PC}(n, p) > 0.95$  están muy próxima a 0.95. Las longitudes medias esperadas ( $\mathcal{LM}\mathcal{E}$ ) del intervalo de Wald son grandes, la menor se alcanza con el criterio 5.

#### 4.4. Comportamiento del intervalo de Wald en términos de los valores de los parámetros $(n$ y $p)$ .

Brown et al. ([8]), en sus ejemplos 1, 3 y 4 exhiben las inconsistencias en términos de la probabilidad de cobertura del intervalo de Wald, también muestran que puntos  $(n, p)$  con probabilidad de cobertura cercana al nivel de confianza nominal surgen de una manera impredecible, de manera similar surgen puntos  $(n, p)$  con probabilidad de cobertura alejada del nivel de confianza nominal. Además, muestran que la probabilidad de cobertura puede variar significativamente para tamaños de muestra  $n$  próximos con  $p$  fija, esto también es cierto para valores de  $p$  cercanos entre sí con  $n$  fija. Tales situaciones pueden observarse también en las figuras 4.1 y 4.3 de la sección 4.1.1. Por éstas razones se realizará un análisis detallado del comportamiento del intervalo de Wald variando los parámetros  $n$  y  $p$ .

Para esto sea  $i \in \llbracket 1, 10 \rrbracket$ ,  $p \in \mathcal{P} = \{0.001, 0.002, \dots, 0.999\}$  y se consideran los siguientes conjuntos  $A_i$  donde,  $A_i = \llbracket 100(i-1) + 1, 100i \rrbracket$ , así  $A_1 = \llbracket 1, 100 \rrbracket$ ,  $A_2 = \llbracket 101, 200 \rrbracket, \dots, A_{10} = \llbracket 901, 1000 \rrbracket$ .

Considerando los conjuntos  $A_i$  con  $i \in \llbracket 1, 10 \rrbracket$ ,  $\mathcal{P}$  y usando las expresiones (4.10), (4.11), (4.13), (4.14) y (4.15) se procederá a calcular los factores de comparación del intervalo de Wald para  $(n, p) \in A_i \times \mathcal{P}$  con base en los criterios bajo estudio. Los resultados se muestran en el cuadro 4.2.

	criterio 1 ( $np \geq 5$ y $n(1-p) \geq 5$ )				criterio 2 ( $np \geq 10$ y $n(1-p) \geq 10$ )				criterio 3 ( $np(1-p) \geq 5$ )			
	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$
$A_1$	10.69	0.9326	0.9527	0.2383	13.58	0.9377	0.9527	0.228	11.49	0.9346	0.9527	0.2252
$A_2$	17.71	0.9399	0.9522	0.1329	19.04	0.942	0.9522	0.1377	17.77	0.9401	0.9522	0.1331
$A_3$	21.66	0.9424	0.9519	0.1005	22.53	0.9438	0.9519	0.1031	21.68	0.9425	0.9519	0.1006
$A_4$	24.31	0.9437	0.9517	0.0842	25.03	0.9448	0.9517	0.0858	24.32	0.9438	0.9517	0.0842
$A_5$	25.9	0.9446	0.9516	0.0739	26.5	0.9454	0.9516	0.075	25.91	0.9446	0.9516	0.0739
$A_6$	27.64	0.9451	0.9515	0.0666	28.16	0.9459	0.9515	0.0675	27.65	0.9452	0.9515	0.0666
$A_7$	28.69	0.9456	0.9514	0.0611	29.14	0.9462	0.9514	0.0618	28.69	0.9456	0.9514	0.0611
$A_8$	29.73	0.946	0.9513	0.0568	30.14	0.9465	0.9513	0.0574	29.74	0.946	0.9513	0.0568
$A_9$	30.6	0.9462	0.9513	0.0533	30.98	0.9467	0.9513	0.0538	30.6	0.9462	0.9513	0.0533
$A_{10}$	31.35	0.9465	0.9512	0.0504	31.67	0.9469	0.9512	0.0508	31.35	0.9465	0.9512	0.0504

(a)

	criterio 4 ( $np(1-p) \geq 10$ )				criterio 5 ( $n \geq 100$ )				criterio 6 ( $n \geq 50$ y $0.2 < p < 0.8$ )			
	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$	$\mathcal{P}A$	$\mathcal{P}CPD$	$\mathcal{P}CPE$	$\mathcal{L}ME$
$A_1$	15.13	0.9394	0.9527	0.2122	11.81	0.9192	0.9528	0.1524	15.89	0.94	0.9527	0.2134
$A_2$	19.23	0.9421	0.9522	0.1384	16.51	0.9264	0.9522	0.1265	22.87	0.9442	0.9523	0.1512
$A_3$	22.62	0.9438	0.9519	0.1033	20.8	0.9344	0.9519	0.0974	27.51	0.9461	0.9519	0.1163
$A_4$	25.07	0.9448	0.9517	0.0859	23.63	0.9381	0.9517	0.0823	30.39	0.9469	0.9517	0.0981
$A_5$	26.53	0.9454	0.9516	0.0751	25.35	0.9403	0.9516	0.0725	31.83	0.9474	0.9516	0.0865
$A_6$	28.19	0.9459	0.9515	0.0675	27.17	0.9418	0.9515	0.0656	33.84	0.9477	0.9514	0.0782
$A_7$	29.15	0.9462	0.9514	0.0619	28.27	0.9428	0.9514	0.0604	34.76	0.9479	0.9514	0.072
$A_8$	30.15	0.9465	0.9513	0.0574	29.37	0.9437	0.9513	0.0562	35.74	0.9481	0.9513	0.067
$A_9$	30.98	0.9467	0.9513	0.0538	30.28	0.9443	0.9513	0.0528	36.5	0.9482	0.9512	0.0629
$A_{10}$	31.68	0.9469	0.9512	0.0508	31.03	0.9449	0.9512	0.0499	37.22	0.9483	0.9512	0.0595

(b)

Cuadro 4.2: Factores de comparación del intervalo de Wald, con  $(n, p) \in A_i \times \mathcal{P}$

Se observa del cuadro 4.2 (a) y (b) que el intervalo de Wald continúa teniendo probabilidades de cobertura para  $n$  y  $p$  en su mayoría inferiores al coeficiente de confianza nominal (0.95) aún cuando se ha hecho uso de los criterios bajo análisis. Los porcentajes de puntos adecuados para el criterio  $i$  ( $\mathcal{PA}_i$ ) continúan siendo pequeños. Para  $n \in A_1$  estos son bastante pequeños siendo mayor el obtenido cuando  $n \geq 50$  y  $0.2 < p < 0.8$  (criterio 6) aunque solo alcanza el 15.89 %. Conforme  $n$  aumenta de  $A_1$  a  $A_{10}$  los porcentajes de puntos adecuados aumentan para cada criterio siendo los obtenidos por el criterio 6 siempre los mayores, aunque solo alcanza el valor máximo de 37.22 % para  $n \in A_{10}$ . Las probabilidades de cobertura promedio por defecto para el criterio  $i$  ( $\mathcal{PCPD}_i$ ) también aumentan conforme  $n$  aumenta de  $A_1$  a  $A_{10}$ , pero cuando  $n \in A_1$  éstas se encuentran alejadas de 0.95 para todos los criterios sobre todo con el criterio 5. También se observa que las probabilidades de cobertura promedio por defecto más cercanas al nivel nominal son obtenidas con el criterio 6 y estas alcanzan su valor máximo de 0.9483 cuando  $n \in A_{10}$ . Las probabilidades de cobertura promedio por exceso para el criterio  $i$  ( $\mathcal{PCPE}_i$ ) decrecen conforme  $n$  aumenta de  $A_1$  a  $A_{10}$ , sin embargo éstas son bastante similares para todos los criterios y están muy cercanas al nivel de confianza 0.95, aún para  $n \in A_1$  estas son 0.9527 para los criterios 1, 2, 3, 4 y 6 y 0.9528 para el criterio 5, sin embargo para  $n \in A_{10}$  éstas disminuyen hasta 0.9512 en todos los criterios, excepto en el criterio 6 ahí se se tiene una probabilidad de cobertura promedio por exceso de 0.9512 desde  $A_9$ . Las longitudes medias esperadas para el criterio  $i$  ( $\mathcal{LME}_i$ ) también mantienen un decrecimiento en todos los criterios cuando  $n$  aumenta de  $A_1$  a  $A_{10}$ , sin embargo las longitudes medias esperadas menores son obtenidos siempre con el criterio 5.

## 4.5. Desempeño de los intervalos alternativos

### 4.5.1. Intervalo de Wilson

Recordemos que el **intervalo de Wilson** es

$$\left[ \frac{X + \kappa^2/2}{n + \kappa^2} - \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2}, \frac{X + \kappa^2/2}{n + \kappa^2} + \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \right]$$

donde  $\hat{p} = \frac{X}{n}$ ,  $\hat{q} = 1 - \hat{p}$  y  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$  siendo  $\Phi(z)$  la función de distribución de una variable aleatoria normal estándar.

Su probabilidad de cobertura se encuentra indicada de la siguiente manera:

$$P_p \left( \frac{X + \kappa^2/2}{n + \kappa^2} - \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \leq p \leq \frac{X + \kappa^2/2}{n + \kappa^2} + \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \right), \quad (4.16)$$

observar que la probabilidad de cobertura depende de  $\alpha$ ,  $n$  y  $X$ .

La proposición 4.2 evalúa el desempeño en términos de la probabilidad de cobertura del intervalo de Wilson para  $\alpha$ ,  $n$  y  $p$ .

**Proposición 4.2.** *Para  $\alpha$ ,  $n$  y  $p$ , la probabilidad de cobertura del intervalo de Wilson puede ser expresada como,*

$$P_p \left( pn - \kappa \sqrt{np(1-p)} \leq X \leq pn + \kappa \sqrt{np(1-p)} \right) \quad (4.17)$$

donde  $X \sim \text{Bin}(n, p)$ .

*Demostración.*

$$\begin{aligned} & P_p \left( \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} - \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \leq p \leq \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} + \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \right) \\ &= P_p \left( \left| p - \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} \right| \leq \frac{\kappa n^{1/2}}{n + \kappa^2} \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \right) \\ &= P_p \left( \frac{|pn + p\kappa^2 - X - \frac{\kappa^2}{2}|}{\kappa n^{1/2}} \leq \left( \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right)^{1/2} \right) \end{aligned}$$

$$\begin{aligned}
&= P_p \left( \frac{(p(n+\kappa^2 - \frac{\kappa^2}{2p}) - X)^2}{\kappa^2 n} \leq \hat{p}\hat{q} + \frac{\kappa^2}{4n} \right) \\
&= P_p \left( \frac{X^2 + X\kappa^2 - 2Xp(n+\kappa^2)}{\kappa^2 n} \leq \frac{X}{n} - \frac{X^2}{n^2} + \frac{\kappa^2}{4n} - \frac{p^2(n+\kappa^2 - \frac{\kappa^2}{2p})^2}{\kappa^2 n} \right) \\
&= P_p \left( X^2 \left( \frac{n + \kappa^2}{\kappa^2 n^2} \right) - \frac{2Xp(n+\kappa^2)}{\kappa^2 n} \leq -\frac{p^2(n+\kappa^2 - \frac{\kappa^2}{2p})^2 - \frac{\kappa^4}{4}}{\kappa^2 n} \right) \\
&= P_p \left( X^2 - \frac{2p(n+\kappa^2)}{\kappa^2 n} \left( \frac{\kappa^2 n^2}{n+\kappa^2} \right) X \leq \frac{\kappa^2 n^2}{n+\kappa^2} \left( -\frac{p^2(n+\kappa^2 - \frac{\kappa^2}{2p})^2 - \frac{\kappa^4}{4}}{\kappa^2 n} \right) \right) \\
&= P_p \left( X^2 - 2pnX \leq -n \frac{p^2(n+\kappa^2 - \frac{\kappa^2}{2p})^2 - \frac{\kappa^4}{4}}{n+\kappa^2} \right) \\
&= P_p \left( X^2 - 2pnX \leq -n \frac{p^2(n+\kappa^2)^2 - 2p^2(n+\kappa^2)\frac{\kappa^2}{2p} + \frac{\kappa^4}{4} - \frac{\kappa^4}{4}}{n+\kappa^2} \right) \\
&= P_p \left( X^2 - 2pnX \leq -n(p^2(n + \kappa^2) - p\kappa^2) \right) \\
&= P_p \left( (X - pn)^2 \leq -n^2 p^2 - np^2 \kappa^2 + np\kappa^2 + p^2 n^2 \right) \\
&= P_p \left( |X - pn| \leq \kappa \sqrt{np(1-p)} \right) \\
&= P_p \left( pn - \kappa \sqrt{np(1-p)} \leq X \leq pn + \kappa \sqrt{np(1-p)} \right). \quad \square
\end{aligned}$$

La probabilidad de cobertura del intervalo de Wilson para  $n$  y  $p$ , se puede escribir de forma compacta como sigue

$$\mathbf{PC}(n, p) = \sum_{x=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} \binom{n}{x} p^x (1-p)^{n-x} \quad (4.18)$$

donde  $x_1 = pn - k\sqrt{np(1-p)}$  y  $x_2 = pn + k\sqrt{np(1-p)}$ ,  $\lceil x_1 \rceil$  y  $\lfloor x_2 \rfloor$  son las funciones techo y piso, respectivamente.

Con la intención de comparar el desempeño en términos de la probabilidad de cobertura del intervalo de Wilson con respecto al intervalo de Wald se presenta la figura 4.7 en la cual se muestran las respectivas probabilidades de cobertura de tales intervalos para  $n = 30$  y  $n = 60$ . Otras figuras con el mismo objetivo pueden encontrarse en ([8]).

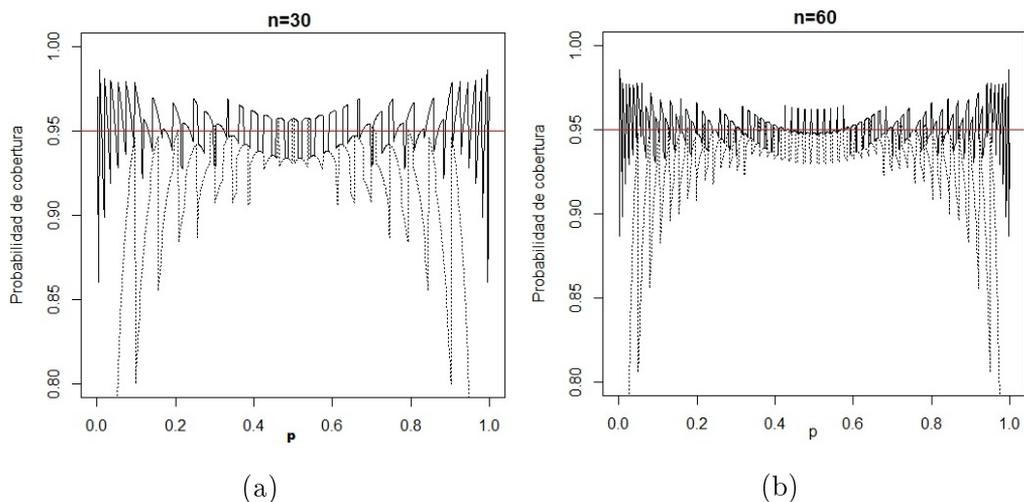


Figura 4.7: Probabilidad de cobertura para  $p \in \mathcal{P}$ , (a)  $n = 30$  y (b)  $n = 60$ . Intervalo de Wald (línea punteada), intervalo de Wilson (línea continua).

Se observa de la figura 4.7 que el desempeño en términos de la probabilidad de cobertura del intervalo de Wilson es preferible al del intervalo de Wald, ya que las probabilidades de cobertura para  $n = 30$  y  $n = 60$  y  $p \in \mathcal{P}$  se encuentran más cercanas a 0.95 y esto es más evidente para valores de  $p$  cercanos a 0 o a 1, en los cuales la probabilidad de cobertura del intervalo de Wald es cercana a 0.

La figura 4.8 muestra las longitudes esperadas del intervalo de Wald y del intervalo de Wilson para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$ .

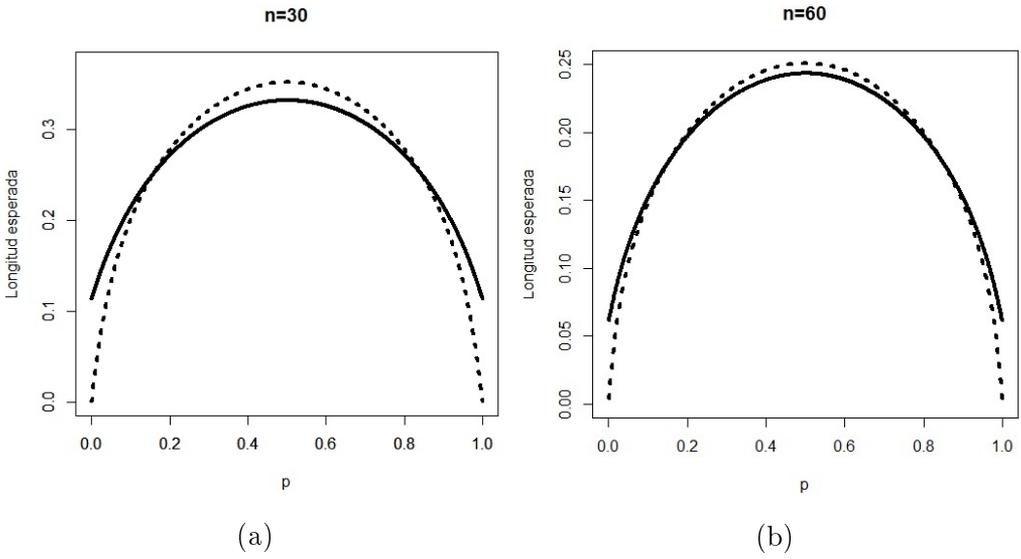


Figura 4.8: Longitud esperada para  $p \in \mathcal{P}$ , (a)  $n = 30$  y (b)  $n = 60$ . Intervalo de Wilson (—), intervalo de Wald (....).

De la figura anterior se observa, en (a) que la longitud esperada del intervalo de Wald es mayor que la del intervalo de Wilson cuando  $0.157 \leq p \leq 0.842$  y menor cuando  $p < 0.157$  o  $p > 0.842$ , en (b) la longitud esperada del intervalo de Wald es mayor que la del intervalo de Wilson cuando  $0.151 \leq p \leq 0.848$  y menor cuando  $p < 0.151$  o  $p > 0.848$ . De lo anterior si sólo tomamos en cuenta la longitud esperada se podría pensar que el desempeño del intervalo de Wald es mejor al del intervalo de Wilson para valores de  $p$  cercanos a 0 o bien a 1, sin embargo como se observa de la figura 4.7 para tales  $p$ , la probabilidad de cobertura dada por el intervalo de Wald está muy alejada del nivel de confianza.

En la figura 4.7 se observa que el intervalo de Wilson presenta (en un grado menor) un mal desempeño en términos de la probabilidad de cobertura cuando  $p$  está cerca de 0 o de 1. En la figura 4.9 analizamos más de cerca esta cuestión para  $p = 0.999$  y  $n \in \llbracket 1, 1000 \rrbracket$ .

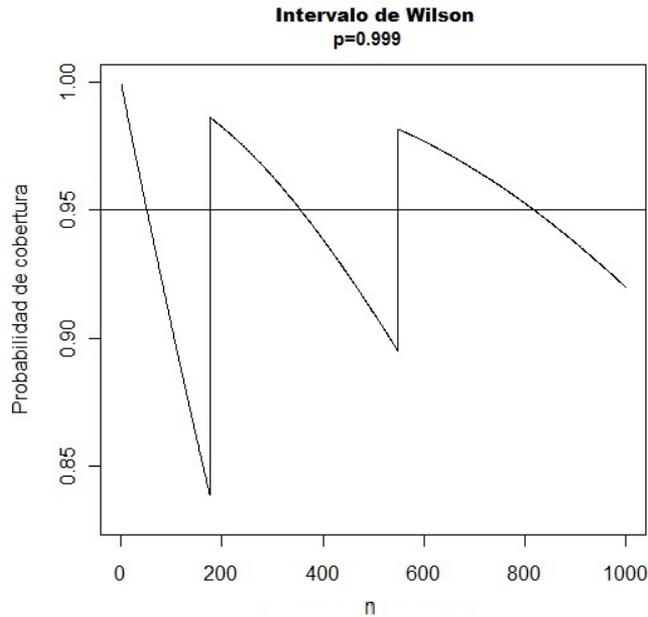


Figura 4.9: Probabilidad de cobertura del intervalo de Wilson, para  $p = 0.999$  y  $n \in \llbracket 1, 1000 \rrbracket$ .

Se observa que la probabilidad de cobertura es deficiente para  $p = 0.999$ . La probabilidad de cobertura para  $n$  pequeños está cerca de 1 y conforme  $n$  aumenta ésta decrece hasta caer a 0.8385 en  $n = 176$ , pero inmediatamente aumenta a 0.9861 en  $n = 177$ . Nuevamente la probabilidad de cobertura decrece hasta llegar a 0.895 en  $n = 548$  e inmediatamente aumenta hasta 0.9817 en  $n = 549$ , otros valores en los cuales se presentan cambios drásticos en la probabilidad de cobertura son en  $n = 1020$  y  $n = 1556$ . Como puede observarse el intervalo de Wilson mantiene un comportamiento similar pero es mejor al del intervalo de Wald para  $p$  cerca del 0 o del 1 (figura 4.1).

Como se mostró anteriormente la probabilidad de cobertura del intervalo de Wilson presenta un pobre desempeño para valores de  $p$  cercanos a 0 o a 1. Lo anterior es ampliamente conocido en la literatura, además éste comportamiento se mantiene para todo  $n$  y toda  $\alpha$  ([4] y [8]). Sin embargo cuando  $p$  no está

cerca de 0 o de 1 el intervalo de Wilson mantiene probabilidades de cobertura cercanas al nivel nominal ([8] y [19]).

Debido a que el intervalo de Wilson continúa presentando irregularidades en su probabilidad de cobertura cuando el parámetro  $p$  está cercano a 0 o bien a 1, se analizará su comportamiento con base en los criterios anteriormente mencionados. La Figura 4.10 presenta la misma situación pero ahora  $n = 100$  y  $p \in \mathcal{P}$ . Los puntos en negro representan las probabilidades de cobertura cuando  $(100, p)$  con  $p \in \mathcal{P}$  cumple el criterio establecido  $i$  con  $i = 1, 2, \dots, 6$  y los puntos en azul cuando no lo cumple.

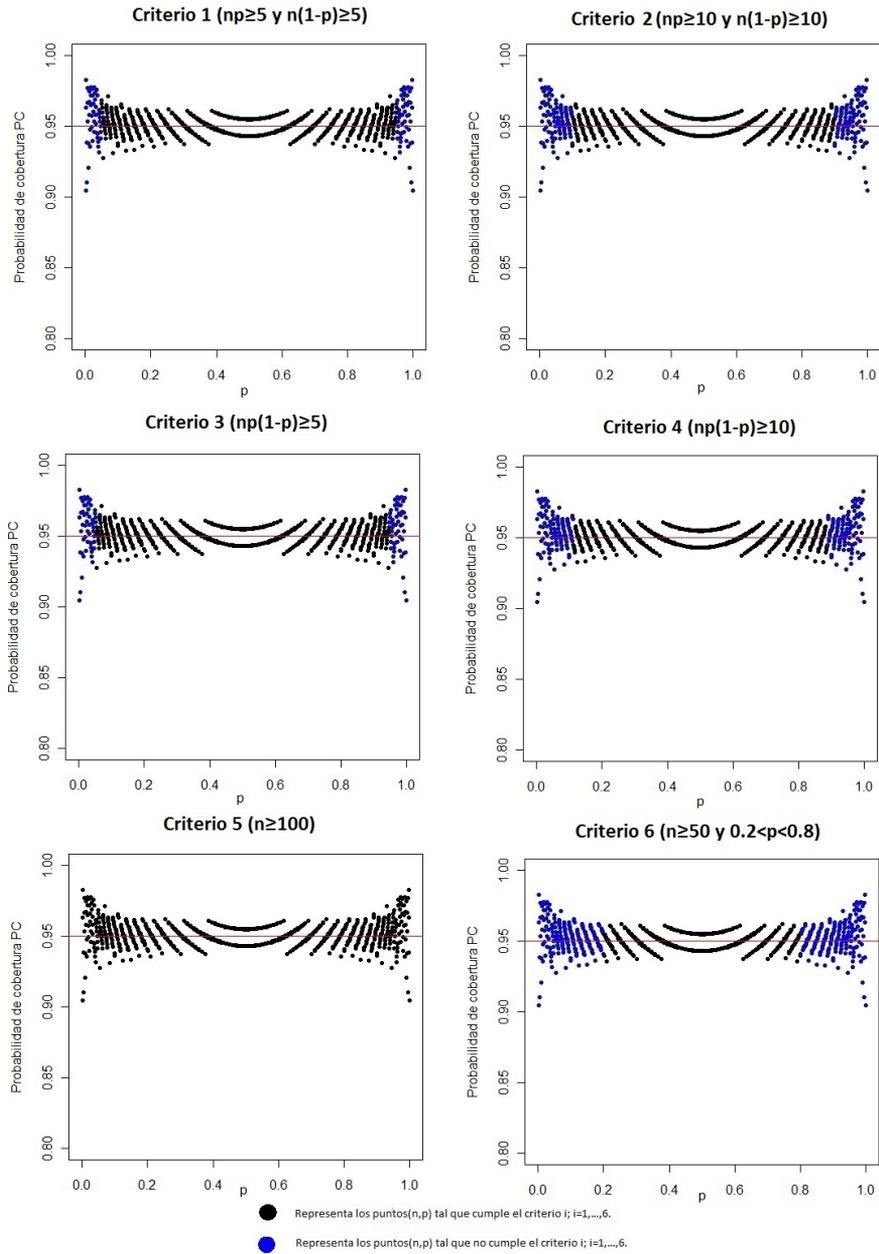


Figura 4.10: Probabilidades de cobertura del intervalo de Wilson cuando  $n = 100$  y  $p \in \mathcal{P}$ , para los criterios 1-6.

De la figura 4.10 se puede constatar que el desempeño del intervalo de Wilson es mejor al del intervalo de Wald (figura 4.6), las probabilidades de cobertura para  $n = 100$  y  $p \in \mathcal{P}$  que se obtienen con el intervalo de Wilson están más próximas al nivel nominal (0.95), en valores de  $p$  cercanos a 0 o a 1 es más clara esta mejoría. También se observa que los criterios bajo análisis eliminan puntos  $(100, p)$  en los cuales la probabilidad de cobertura obtenida por el intervalo de Wilson está alejada del nivel de confianza nominal.

El cuadro 4.3 presenta los valores de los factores de comparación (sección 4.2) que provee el intervalo de Wilson cuando  $n = 100$  y  $p \in \mathcal{P}$ .

Factor \ Criterio	1	2	3	4	5	6
$\mathcal{PA}$	54.78	53.50	54.53	52.90	55.46	52.09
$\mathcal{ECPD}$	0.0050	0.0049	0.0050	0.0049	0.0057	0.0046
$\mathcal{ECP\mathcal{E}}$	0.0052	0.0049	0.0052	0.0048	0.0064	0.0046
$\mathcal{LME}$	0.1619	0.1691	0.1623	0.1707	0.1524	0.1796

Cuadro 4.3: Factores de comparación del intervalo de Wilson con base en los criterios  $i$  para  $i \in \llbracket 1, 6 \rrbracket$ , para  $n = 100$  y  $p \in \mathcal{P}$ .

Del cuadro 4.3 puede observarse que para los criterios bajo análisis, el intervalo de Wilson presenta porcentajes de puntos adecuados ( $\mathcal{PA}$ ) mayores a los obtenidos por el intervalo de Wald (cuadro 4.1) siendo el mayor obtenido por el criterio 5. Los errores de cobertura promedio por defecto ( $\mathcal{ECPD}$ ) son inferiores a los obtenidos por el intervalo de Wald y los errores de cobertura promedio por exceso ( $\mathcal{ECP\mathcal{E}}$ ) son mayores a los obtenidos por el intervalo de Wald pero tales valores son pequeños, los errores de cobertura promedio por defecto y por exceso más pequeños para el intervalo de Wilson son obtenidos por el criterio 6. Las longitudes medias esperadas ( $\mathcal{LME}$ ) son también menores a los obtenidos por el intervalo de Wald, la menor es obtenida por el criterio 5.

## 4.5.2. Intervalo de Agresti-Coull

Recordemos que el **intervalo de Agresti-Coull** es

$$\left[ \tilde{p} - \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2}, \tilde{p} + \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2} \right]$$

donde  $\tilde{p} = \frac{\tilde{X}}{\tilde{n}}$ ,  $\tilde{X} = X + \frac{\kappa^2}{2}$ ,  $\tilde{n} = n + \kappa^2$ ,  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$ ;  $\Phi(z)$  es la función de distribución de una normal estándar y  $\tilde{q} = 1 - \tilde{p}$ .

Su probabilidad de cobertura se encuentra indicada de la siguiente manera:

$$P_p \left( \tilde{p} - \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2} \leq p \leq \tilde{p} + \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2} \right), \quad (4.19)$$

observar que la probabilidad de cobertura depende de  $\alpha$ ,  $n$  y  $X$ .

La proposición 4.3 evalúa el desempeño en términos de la probabilidad de cobertura del intervalo de Agresti-Coull para  $\alpha$ ,  $n$  y  $p$ .

**Proposición 4.3.** *Para  $\alpha$ ,  $n$  y  $p$ , la probabilidad de cobertura del intervalo de Agresti-Coull puede ser expresada como,*

$$P_p \left( \frac{n + \kappa^2}{1 + c} \left[ p + \frac{c}{2} - \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] - \frac{\kappa^2}{2} \leq X \leq \frac{n + \kappa^2}{1 + c} \left[ p + \frac{c}{2} + \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] - \frac{\kappa^2}{2} \right). \quad (4.20)$$

donde  $c = \frac{\kappa^2}{\tilde{n}}$  y  $X \sim \text{Bin}(n, p)$ .

*Demostración.*

$$\begin{aligned} & P_p(\tilde{p} - \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2} \leq p \leq \tilde{p} + \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2}) \\ &= P_p(|\tilde{p} - p| \leq \kappa \tilde{n}^{-1/2} (\tilde{p}\tilde{q})^{1/2}) = P_p((\tilde{p} - p)^2 \leq \kappa^2 \tilde{n}^{-1} (\tilde{p}\tilde{q})), \quad \text{si } c = \kappa^2 \tilde{n}^{-1} = \frac{\kappa^2}{\tilde{n}} \\ &= P_p(\tilde{p}^2 - 2\tilde{p}p + p^2 \leq c\tilde{p} - c\tilde{p}^2) = P_p((1+c)\tilde{p}^2 - (2p+c)\tilde{p} \leq -p^2) \\ &= P_p\left((1+c)[\tilde{p}^2 - \frac{2p+c}{1+c}\tilde{p}] \leq -p^2\right) = P_p\left(\tilde{p}^2 - \frac{2p+c}{1+c}\tilde{p} \leq -\frac{p^2}{1+c}\right) \\ &= P_p\left(\left(\tilde{p} - \frac{p+\frac{c}{2}}{1+c}\right)^2 \leq -\frac{p^2}{1+c} + \left(\frac{p+\frac{c}{2}}{1+c}\right)^2\right) = P_p\left(\left|\tilde{p} - \frac{p+\frac{c}{2}}{1+c}\right| \leq \sqrt{-\frac{p^2}{1+c} + \left(\frac{p+\frac{c}{2}}{1+c}\right)^2}\right) \\ &= P_p\left(\left|\tilde{p} - \frac{p+\frac{c}{2}}{1+c}\right| \leq \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) = P_p\left(-\frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c} \leq \tilde{p} - \frac{p+\frac{c}{2}}{1+c} \leq \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) \\ &= P_p\left(\frac{p+\frac{c}{2}}{1+c} - \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c} \leq \frac{\tilde{X}}{\tilde{n}} \leq \frac{p+\frac{c}{2}}{1+c} + \frac{\sqrt{-p^2c+pc+\frac{c^2}{4}}}{1+c}\right) \end{aligned}$$

$$\begin{aligned}
&= P_p \left( \frac{p + \frac{c}{2}}{1+c} - \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \leq \frac{X + \frac{\kappa^2}{2}}{n + \kappa^2} \leq \frac{p + \frac{c}{2}}{1+c} + \frac{\sqrt{-p^2c + pc + \frac{c^2}{4}}}{1+c} \right) \\
&= P_p \left( \frac{n + \kappa^2}{1+c} \left[ p + \frac{c}{2} - \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] - \frac{\kappa^2}{2} \leq X \leq \frac{n + \kappa^2}{1+c} \left[ p + \frac{c}{2} + \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] - \frac{\kappa^2}{2} \right). \quad \square
\end{aligned}$$

La probabilidad de cobertura del intervalo de Agresti-Coull para  $n$  y  $p$ , se puede escribir de forma compacta como sigue

$$\mathbf{PC}(n, p) = \sum_{x=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} \binom{n}{x} p^x (1-p)^{n-x}, \quad (4.21)$$

donde  $c = \frac{\kappa^2}{n}$ ,  $x_1 = \frac{n + \kappa^2}{1+c} \left[ p + \frac{c}{2} - \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] - \frac{\kappa^2}{2}$ ,

$x_2 = \frac{n + \kappa^2}{1+c} \left[ p + \frac{c}{2} + \sqrt{-p^2c + pc + \frac{c^2}{4}} \right] + \frac{\kappa^2}{2}$ ,  $\lceil x_1 \rceil$  y  $\lfloor x_2 \rfloor$  son las funciones techo y piso, respectivamente.

A continuación se muestra el desempeño en términos de probabilidad de cobertura del intervalo de Agresti-Coull con respecto al intervalo de Wald para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$ , figuras con el mismo objetivo se encuentran en ([4]).

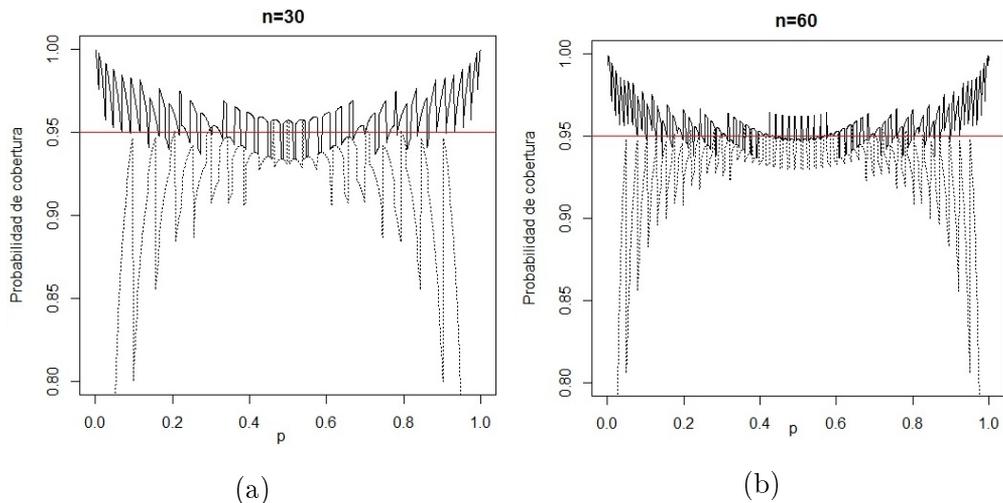


Figura 4.11: Probabilidad de cobertura para  $p \in \mathcal{P}$ , (a)  $n = 30$  y (b)  $n = 60$ . Intervalo de Wald (línea punteada), intervalo de Agresti-Coull (línea continua).

Se observa de la figura 4.11 que las probabilidades de cobertura obtenidas por el intervalo de Agresti-Coull para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$  se encuentran más cerca de 0.95 que las obtenidas por medio del intervalo de Wald. Lo anterior es más notable para valores de  $p$  cercanos a 0 o a 1, sin embargo se observa también que para tales valores de  $p$  las probabilidades de cobertura están muy cercanas a 1.

La figura 4.12 muestra las longitudes esperadas del intervalo de Wald y del intervalo de Agresti-Coull para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$ .

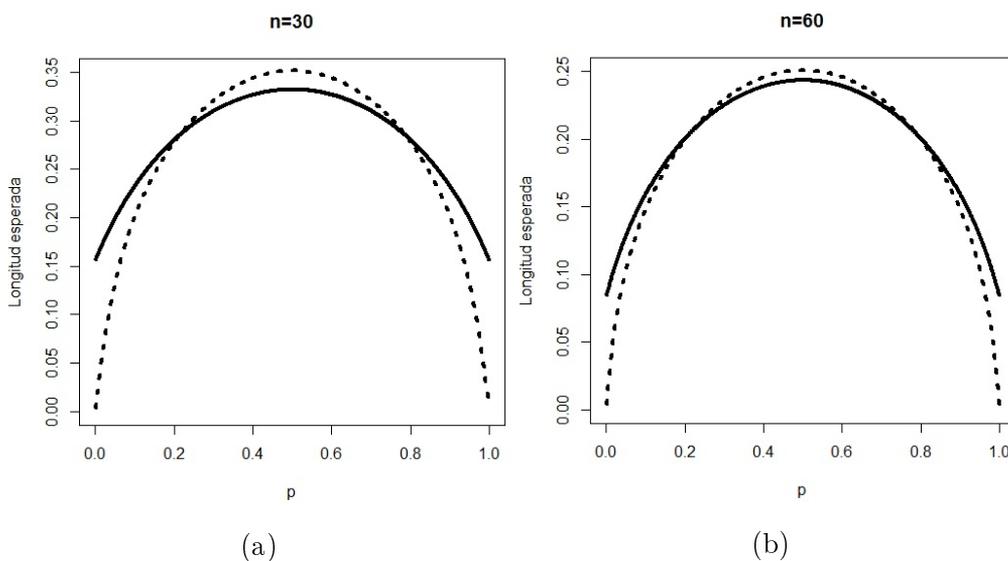


Figura 4.12: Longitud esperada para  $p \in \mathcal{P}$ , (a)  $n = 30$  y (b)  $n = 60$ . Intervalo de Agresti-Coull (—), intervalo de Wald (.....).

De la figura anterior se observa que, en (a) la longitud esperada del intervalo de Wald es mayor que la del intervalo de Agresti-Coull cuando  $0.214 \leq p \leq 0.785$  y menor cuando  $p < 0.214$  o  $p > 0.785$ , en (b) la longitud esperada del intervalo de Wald es mayor que la del intervalo de Agresti-Coull cuando  $0.213 \leq p \leq 0.786$  y menor cuando  $p < 0.213$  o  $p > 0.786$ . Nuevamente, en términos de la longitud esperada el intervalo de Wald parecería ser mejor que

el intervalo de Agresti-Coull para valores de  $p$  cercanos a 0 o bien de 1, sin embargo para tales  $p$  la probabilidad de cobertura dada por el intervalo de Wald está muy alejada del nivel de confianza (figura 4.11).

Con la finalidad de examinar el comportamiento del intervalo de Agresti-Coull cuando  $p$  está cerca de 0 o de 1 se presenta la figura 4.13, la cual muestra el comportamiento de la probabilidad de cobertura del intervalo de Agresti-Coull para  $p = 0.999$  y  $n \in \llbracket 1, 10000 \rrbracket$ .

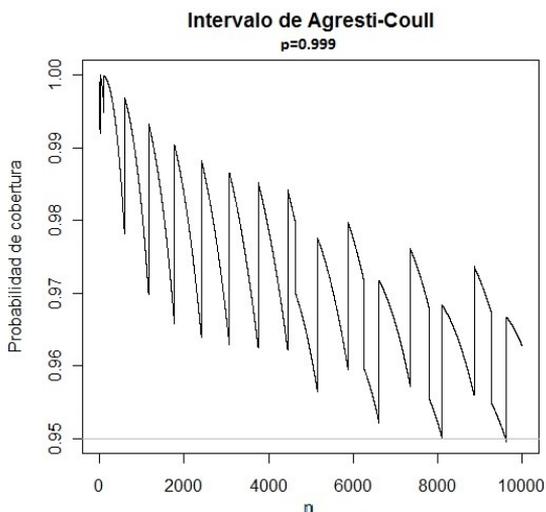


Figura 4.13: Probabilidad de cobertura del intervalo de Agresti-Coull, cuando  $p = 0.999$  y  $n \in \llbracket 1, 10000 \rrbracket$ .

Se observa de la figura 4.13 que el comportamiento del intervalo de Agresti-Coull para  $p$  cercanos a 0 o a 1 es contrario al comportamiento del intervalo de Wald (figura 4.1) para los mismos casos. En el intervalo de Agresti-Coull las probabilidades de cobertura para  $n$  pequeños son cercanas a 1, por ejemplo para  $n = 10$  la  $\mathbf{PC}(10, 0.999) = 0.9999$ , las probabilidades de cobertura comienzan cerca de 1 y de ahí decrecen aunque no de manera monótona al nivel nominal 0.95. Los valores de  $n$  para los cuales se aseguraría una probabilidad de cobertura cercana a 0.95 son demasiado grandes como puede verse en la figu-

ra 4.1, por ejemplo en  $n = 2000$  la  $\mathbf{PC}(2000, 0.999) = 0.9835$ . Finalmente se observa que hay saltos drásticos en la probabilidad de cobertura, por ejemplo en  $n = 587$  la  $\mathbf{PC}(587, 0.999) = 0.9782$  pero aumenta a 0.9969 en  $n = 588$ , en  $n = 1770$  la  $\mathbf{PC}(1770, 0.999) = 0.9658$  pero inmediatamente aumenta a 0.9904 en  $n = 1771$ . Agresti y Min muestran en su artículo ([4]) que el intervalo de Agresti-Coull puede ser bastante conservador para valores de  $p$  cercano a 0 o bien a 1, es decir, en tales  $p$  presenta probabilidades de cobertura cercanas a 1, de acuerdo con Agresti y Min, ello es preferible a tener probabilidades de cobertura muy bajas, las cuales son obtenidas por el intervalo de Wald en esas regiones.

Debido a que el intervalo de Agresti-Coull es bastante conservador cuando  $p$  está cerca de 0 o bien a 1, se observará su desempeño con base en los criterios bajo estudio. Para  $n = 100$  y  $p \in \mathcal{P}$  la figura 4.14 muestra el desempeño del intervalo de Agresti-Coull con base en tales criterios, los puntos en negro representan las probabilidades de cobertura para  $(100, p)$  con  $p \in \mathcal{P}$  cuando  $(100, p)$  cumple el criterio establecido  $i$  y los puntos en azul cuando  $(100, p)$  no lo cumple.

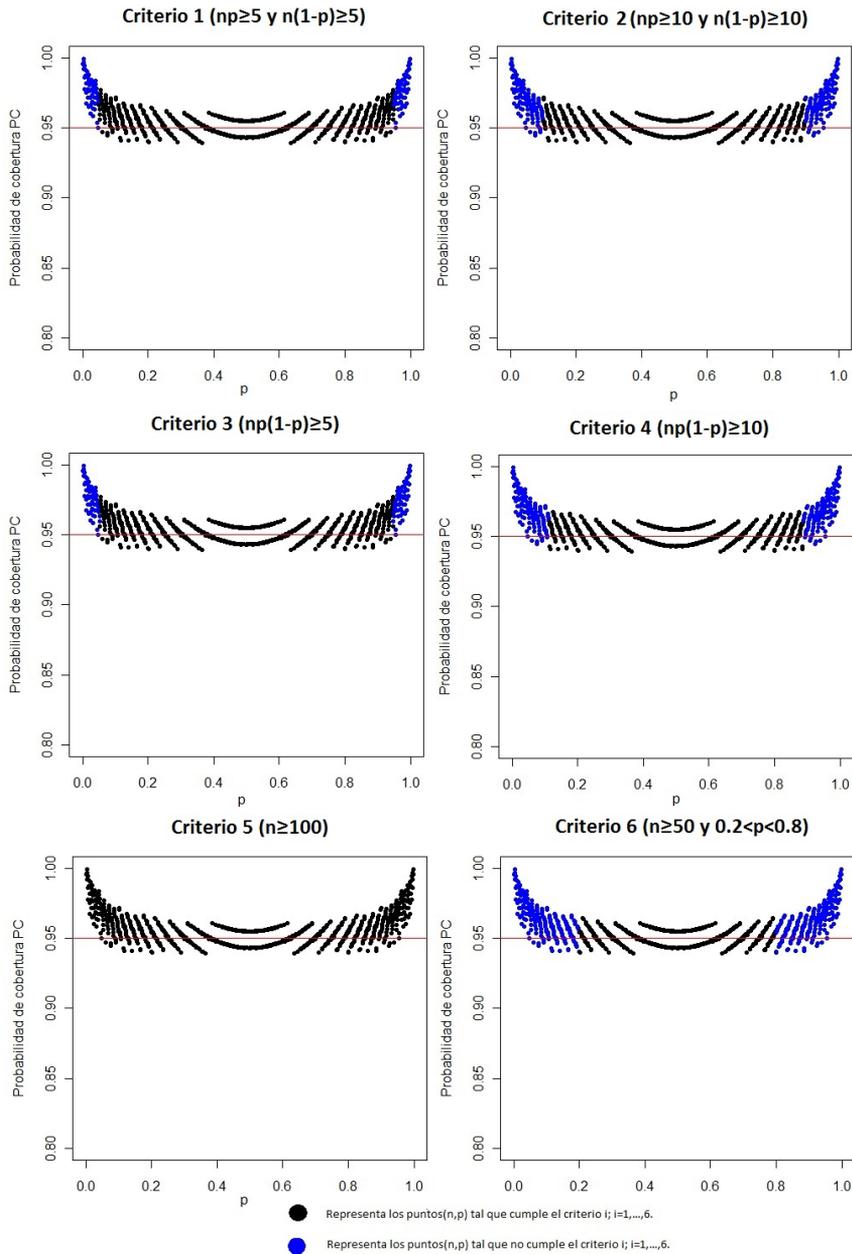


Figura 4.14: Probabilidades de cobertura del intervalo de Agresti-Coull cuando  $n=100$  y  $p \in \mathcal{P}$ , con base en los criterios.

De la figura 4.14 puede observarse que las probabilidades de cobertura en su mayoría son mayores que el nivel de confianza nominal (0.95), también que los criterios 1, 2, 3, 4 y 6 excluyen puntos  $(100, p)$  con  $p \in \mathcal{P}$  tal que  $\mathbf{PC}(n, p) \geq 0.95$  sin embargo en la mayoría de estos puntos se cumple que  $|\mathbf{PC}(100, p) - 0.95|$  es demasiado grande, es decir, tales puntos tienen una probabilidad de cobertura que sobrepasa por mucho a 0.95, tales puntos están cercanos a los límites 0 y 1 que es donde el intervalo de Agresti-Coull es bastante conservador.

El cuadro 4.4 presenta los factores de comparación que se obtienen para el intervalo de Agresti-Coull cuando  $n = 100$  y  $p \in \mathcal{P}$ .

Factor \ Criterio	1	2	3	4	5	6
<i>PA</i>	65.22	62.38	65.03	61.42	68.47	57.76
<i>ECPD</i>	0.0041	0.0042	0.0041	0.0042	0.0041	0.0042
<i>ECPE</i>	0.0071	0.0059	0.0070	0.0057	0.0099	0.0049
<i>LME</i>	0.1633	0.1701	0.1637	0.1716	0.1546	0.1801

Cuadro 4.4: Factores obtenidas por el intervalo de Agresti-Coull con base a los criterios 1, 2, ..., 6, cuando  $n = 100$  y  $p \in \mathcal{P}$ .

Del cuadro 4.4 puede observarse que para los criterios bajo análisis, los porcentajes de puntos adecuados son mayores a los obtenidos por los intervalos de Wilson (cuadro 4.3) y por supuesto a los obtenidos por el intervalo de Wald (cuadro 4.1), los errores de cobertura promedio por defecto (*ECPD*) son inferiores a los obtenidos por el intervalo de Wilson, los errores de cobertura promedio por exceso (*ECPE*) son superiores a los obtenidos por el intervalo de Wilson y las longitudes medias esperadas (*LME*) son ligeramente mayores a las obtenidos por el intervalo de Wilson.

### 4.5.3. Intervalo arcoseno

Recordemos que el **intervalo arcoseno** es

$$\left[ \text{seno}^2\left(\text{arcoseno}(\check{p}^{1/2}) - \frac{1}{2}\kappa n^{-1/2}\right), \text{seno}^2\left(\text{arcoseno}(\check{p}^{1/2}) + \frac{1}{2}\kappa n^{-1/2}\right) \right]$$

donde  $\check{p} = \frac{X + \frac{3}{8}}{n + \frac{3}{4}}$ , y  $\kappa = z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$  donde  $\Phi(z)$  es la función de distribución de una normal estándar.

Su probabilidad de cobertura se indica de la siguiente manera:

$$P_p\left(\text{seno}^2(\text{arcoseno}(\check{p}^{1/2}) - \frac{1}{2}\kappa n^{-1/2}) \leq p \leq \text{seno}^2(\text{arcoseno}(\check{p}^{1/2}) + \frac{1}{2}\kappa n^{-1/2})\right), \quad (4.22)$$

observar que la probabilidad de cobertura depende de  $\alpha$ ,  $n$  y  $X$ .

La proposición 4.4 evalúa el desempeño en términos de la probabilidad de cobertura del intervalo arcoseno para  $\alpha$ ,  $n$  y  $p$ .

**Proposición 4.4.** *Para  $\alpha$ ,  $n$  y  $p$ , la probabilidad de cobertura del intervalo arcoseno puede ser expresada como,*

$$P_p\left(\left(n + \frac{3}{4}\right)\text{seno}^2(\text{arcoseno}(\sqrt{p}) - \frac{1}{2}\kappa n^{-1/2}) - \frac{3}{8} \leq X \leq \left(n + \frac{3}{4}\right)\text{seno}^2(\text{arcoseno}(\sqrt{p}) + \frac{1}{2}\kappa n^{-1/2}) - \frac{3}{8}\right) \quad (4.23)$$

donde  $X \sim \text{Bin}(n, p)$ .

*Demostración.*

$$\begin{aligned} & P_p\left(\text{seno}^2(\text{arcoseno}(\check{p}^{1/2}) - \frac{1}{2}\kappa n^{-1/2}) \leq p \leq \text{seno}^2(\text{arcoseno}(\check{p}^{1/2}) + \frac{1}{2}\kappa n^{-1/2})\right) \\ &= P_p\left(\text{arcoseno}(\check{p}^{1/2}) - \frac{1}{2}\kappa n^{-1/2} \leq \text{arcoseno}(\sqrt{p}) \leq \text{arcoseno}(\check{p}^{1/2}) + \frac{1}{2}\kappa n^{-1/2}\right) \\ &= P_p\left(\text{arcoseno}(\sqrt{p}) - \frac{1}{2}\kappa n^{-1/2} \leq \text{arcoseno}(\check{p}^{1/2}) \leq \text{arcoseno}(\sqrt{p}) + \frac{1}{2}\kappa n^{-1/2}\right) \\ &= P_p\left(\text{seno}^2(\text{arcoseno}(\sqrt{p}) - \frac{1}{2}\kappa n^{-1/2}) \leq \check{p} \leq \text{seno}^2(\text{arcoseno}(\sqrt{p}) + \frac{1}{2}\kappa n^{-1/2})\right) \\ &= P_p\left(\text{seno}^2(\text{arcoseno}(\sqrt{p}) - \frac{1}{2}\kappa n^{-1/2}) \leq \frac{X + \frac{3}{8}}{n + \frac{3}{4}} \leq \text{seno}^2(\text{arcoseno}(\sqrt{p}) + \frac{1}{2}\kappa n^{-1/2})\right) \end{aligned}$$

$$P_p \left( \left( n + \frac{3}{4} \right) \text{seno}^2(\text{arcoseno}(\sqrt{p}) - \frac{1}{2} \kappa n^{-\frac{1}{2}}) - \frac{3}{8} \leq X \right. \\ \left. \leq \left( n + \frac{3}{4} \right) \text{seno}^2(\text{arcoseno}(\sqrt{p}) + \frac{1}{2} \kappa n^{-\frac{1}{2}}) - \frac{3}{8} \right). \quad \square$$

La probabilidad de cobertura del intervalo arcoseno para  $n$  y  $p$ , puede escribirse de forma compacta como sigue

$$\mathbf{PC}(n, p) = \sum_{x=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} \binom{n}{x} p^x (1-p)^{n-x}, \quad (4.24)$$

donde  $x_1 = \left( n + \frac{3}{4} \right) \text{seno}^2(\text{arcoseno}(\sqrt{p}) - \frac{1}{2} \kappa n^{-1/2}) - \frac{3}{8}$ ,  $x_2 = \left( n + \frac{3}{4} \right) \text{seno}^2(\text{arcoseno}(\sqrt{p}) + \frac{1}{2} \kappa n^{-1/2}) - \frac{3}{8}$ ,  $\lceil x_1 \rceil$  y  $\lfloor x_2 \rfloor$  son las funciones techo y piso, respectivamente.

A continuación se presentan gráficas que muestran el comportamiento de la probabilidad de cobertura del intervalo arcoseno y del intervalo de Wald para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$ . Otras gráficas que muestran el comportamiento del intervalo arcoseno pueden encontrarse en ([8]).

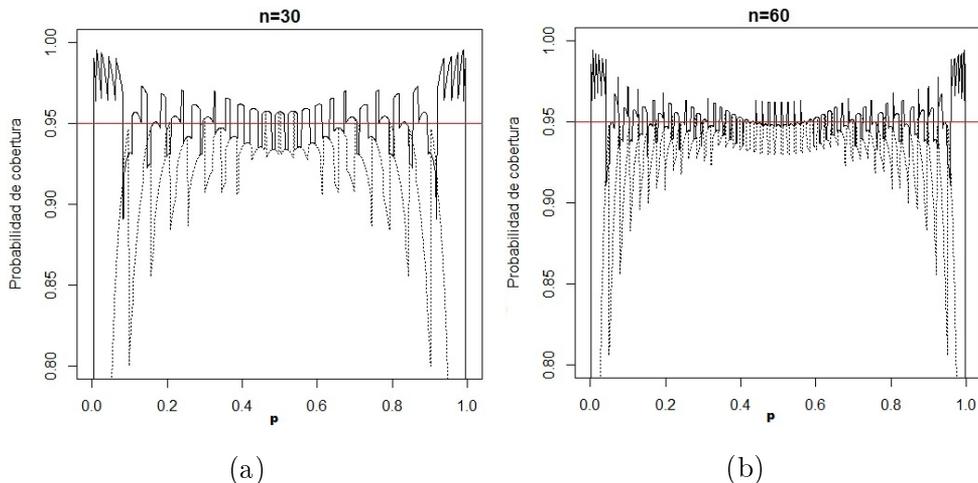


Figura 4.15: Probabilidad de cobertura para  $p \in \mathcal{P}$ , (a)  $n = 30$  y (b)  $n = 60$ . Intervalo de Wald (línea punteada), intervalo arcoseno (línea continua).

Es claro de la figura 4.15 que aunque el intervalo arcoseno no produzca una cobertura aceptable ésta es mejor a la obtenida por medio del intervalo de Wald ya que las probabilidades de cobertura para  $n = 30$ ,  $n = 60$  y  $p \in \mathcal{P}$  se encuentran más cerca del nivel de confianza nominal, sin embargo para  $p$  bastante cerca de 0 o de 1 la probabilidad de cobertura del intervalo arcoseno es menor a la del intervalo de Wald (figura 4.16).

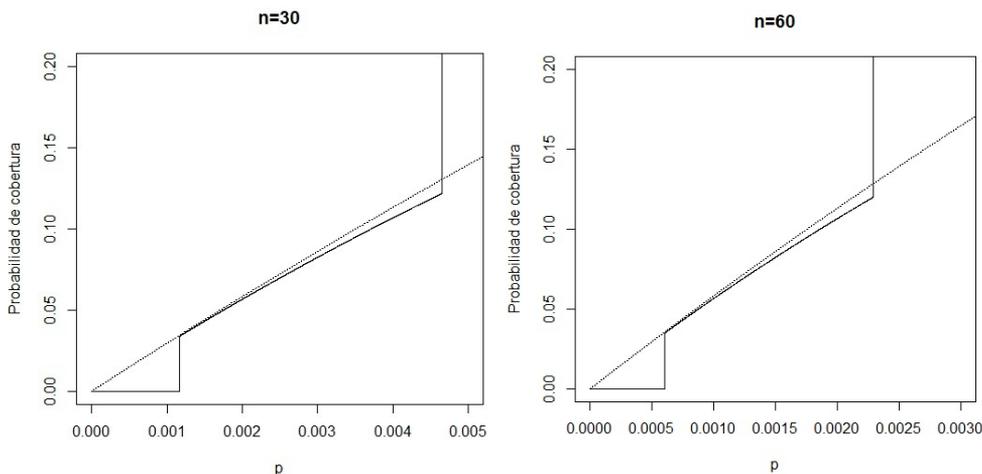


Figura 4.16: Probabilidad de cobertura para  $n = 30$  (a),  $n = 60$  (b). Intervalo de Wald (línea punteada), intervalo de arcoseno (línea continua).

Con la finalidad de analizar el desempeño del intervalo arcoseno en puntos cercanos a 0 o a 1 se presenta la figura 4.17 que muestra su probabilidad de cobertura para  $p = 0.999$  y  $n \in \llbracket 1, 10000 \rrbracket$ .

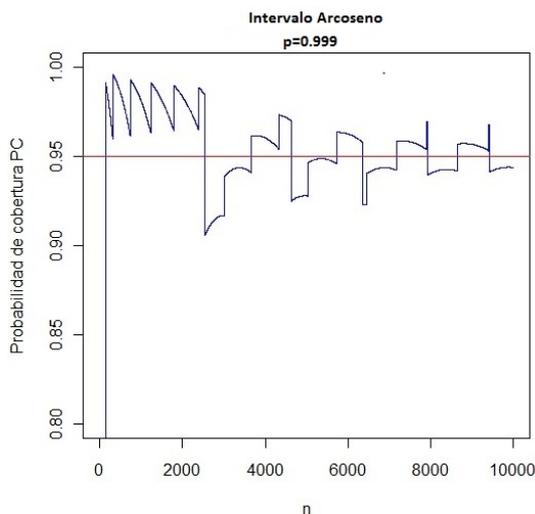


Figura 4.17: Probabilidad de cobertura del intervalo arcoseno, cuando  $p = 0.001$  y  $n \in \llbracket 1, 10000 \rrbracket$ .

De la figura 4.17 se observa que probabilidades de cobertura obtenidas por el intervalo arcoseno para  $p = 0.999$  están demasiado alejadas del nivel nominal 0.95 e incluso para  $n \leq 35$  éstas son 0, en  $n = 36$  la probabilidad de cobertura es solamente 0.0348 y aumenta conforme aumenta  $n$  hasta llegar a 0.1188 en  $n = 136$ , inmediatamente alcanza 0.9915 en  $n = 137$ . Al igual que en los intervalos anteriores los saltos en la probabilidad de cobertura surgen de manera imprevista, sin embargo existen saltos que son demasiado notables tales como el que ocurre en  $n = 136$ , situaciones similares ocurren en  $n = 2534$  donde la probabilidad de cobertura es 0.9887 y cae hasta 0.9057 en  $n = 2535$ . Otros saltos drásticos en la cobertura surgen en  $n = 4631$  y en  $n = 6353$ .

Veamos el comportamiento del intervalo arcoseno con respecto a los criterios anteriormente mencionados para  $n = 100$  y  $p \in \mathcal{P}$ . Los puntos en negro representan las probabilidades de cobertura cuando  $(100, p)$  con  $p \in \mathcal{P}$  cumple el criterio establecido  $i$  y los puntos en azul cuando  $(n, p)$  no lo cumple.

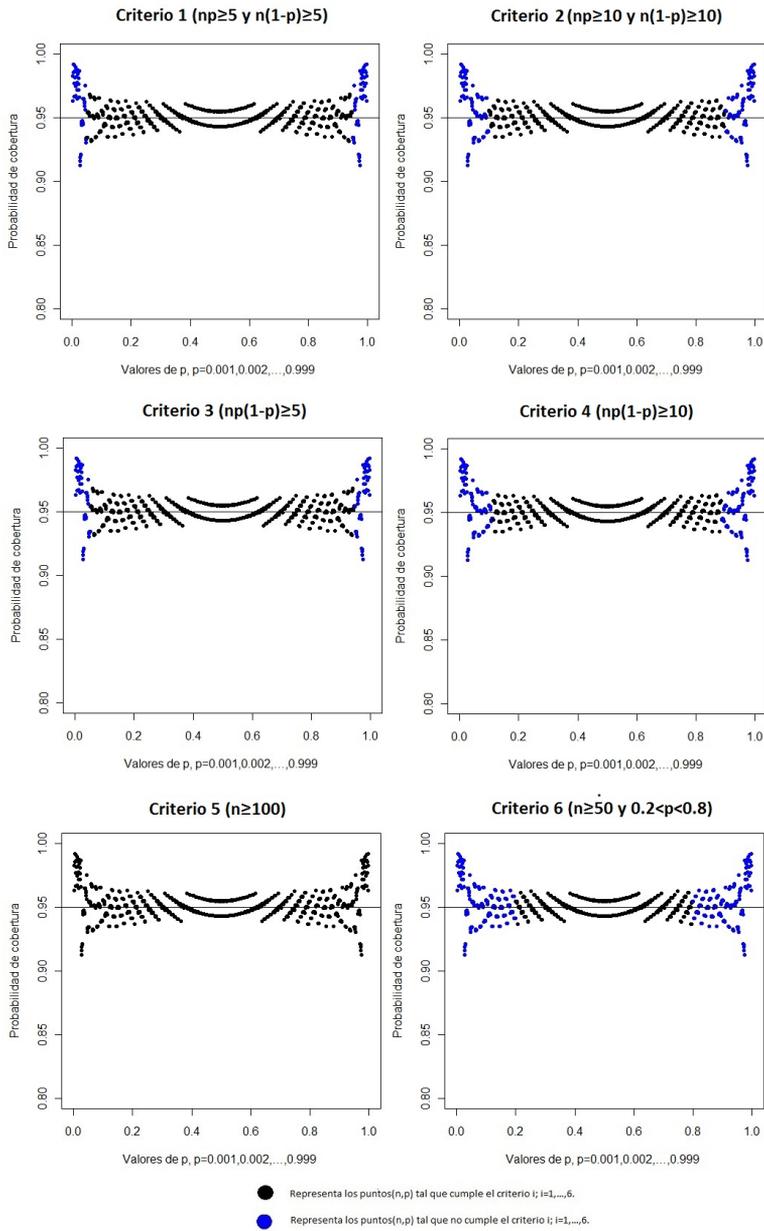


Figura 4.18: Probabilidades de cobertura del intervalo arco seno para  $n=100$  y  $p \in \mathcal{P}$ , con base en los criterios.

Puede observarse de la figura 4.18 que los criterios eliminan puntos cercanos a 0 o a 1 en los cuales el intervalo arco seno produce malas probabilidades de cobertura y también se observa que la cantidad de puntos  $(100, p)$  con probabilidades de cobertura mayor que 0.95 es similar a la cantidad de puntos  $(100, p)$  con probabilidades de cobertura menor que 0.95. Los factores de comparación para este caso pueden observarse en el cuadro 4.5.

Factor \ Criterio	1	2	3	4	5	6
$\mathcal{PA}$	55.56	54.38	55.64	54.71	56.46	54.09
$\mathcal{ECPD}$	0.0056	0.0049	0.0055	0.0049	0.0103	0.0044
$\mathcal{ECPE}$	0.0052	0.0049	0.0052	0.0050	0.0075	0.0047
$\mathcal{LME}$	0.196	0.196	0.196	0.196	0.196	0.196

Cuadro 4.5: Factores obtenidas por el intervalo arco seno considerando los seis criterios bajo análisis, para  $n = 100$  y  $p \in \mathcal{P}$ .

Del cuadro 4.5 se observa que los porcentajes de puntos adecuados ( $\mathcal{PA}$ ) obtenidos por el intervalo arco seno son mayores a los obtenidos por los intervalos de Wald y de Wilson (cuadros 4.1 y 4.3) pero menores a los obtenidos por el intervalo de Agresti-Coull (cuadro 4.4). Los errores de cobertura promedio por defecto  $\mathcal{ECPD}$  son menores a los obtenidos por los intervalos de Wald, mayores o iguales a los obtenidos por medio del intervalo de Wilson excepto para el criterio 6, ahí los errores de cobertura promedio por defecto del intervalo arco seno son menores, los errores de cobertura promedio por defecto del intervalo arco seno son mayores a los obtenidos por el Intervalo de Agresti-Coull. Los errores de cobertura promedio por exceso ( $\mathcal{ECPE}$ ) son mayores a los obtenidos por los intervalos de Wald y de Wilson y menores a los del Intervalo de Agresti-Coull. Las longitudes medias esperadas ( $\mathcal{LME}$ ) son mucho mayores a las obtenidas por otros tres intervalos.

## 4.6. Comparación de los intervalos

Con el objetivo de observar de una manera clara las diferencias entre los intervalos, a continuación se presenta un caso particular para el intervalo de Wald, cuando el tamaño de muestra  $n$  es pequeño. La amplia literatura nos asevera que para estos casos el intervalo de Wald presenta probabilidades de cobertura muy alejadas del nivel nominal. Sea  $n = 20$  y  $p \in \mathcal{P}$ . Se denotará al Intervalo de Wald por  $I_E$ , al Intervalo de Wilson  $I_W$ , al intervalo de Agresti-Coull  $I_{AC}$  y al intervalo arcoseno  $I_A$ . La figura 4.19 muestra lo anteriormente dicho.

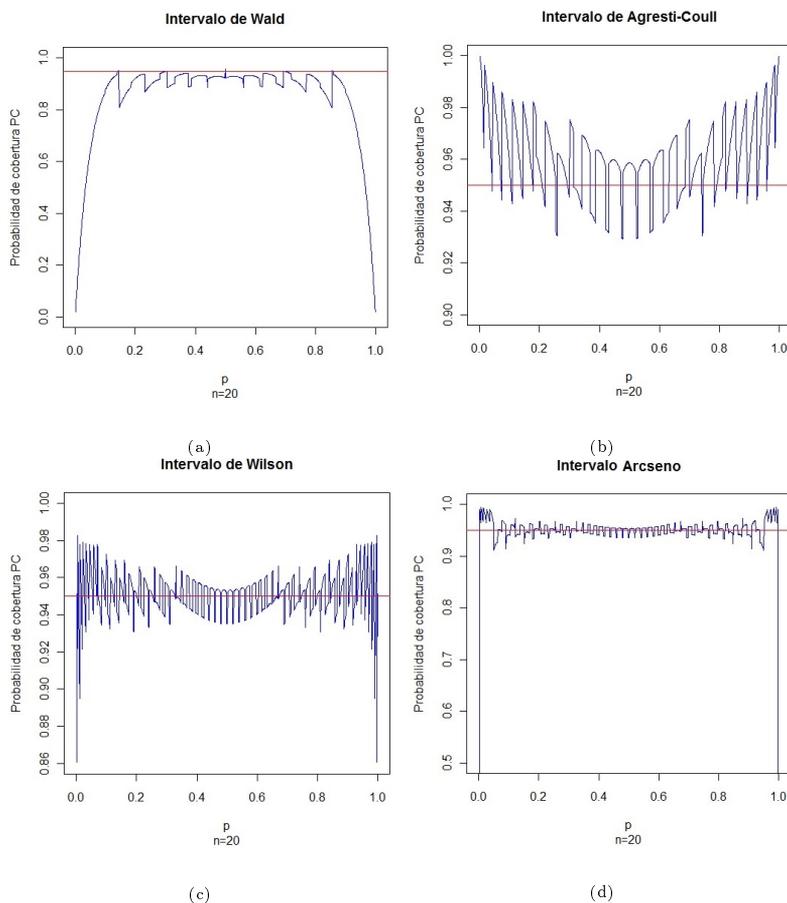


Figura 4.19: Probabilidad de cobertura,  $n = 20$  y  $p \in \mathcal{P}$ .

De la figura 4.19 se observa que los intervalos Agresti-Coull, Wilson y arcoseno muestran una probabilidad de cobertura preferible a la del intervalo de Wald, en especial, el comportamiento del intervalo de Wilson parece ser el mejor sobre todo para  $p$  cercanos a 0 o a 1, ya que para tales  $p$  el intervalo de Agresti-Coull otorga probabilidades de cobertura cercanas a 1 y el intervalo arcoseno a 0.

La figura 4.20 representa las longitudes esperadas de los 4 intervalos para  $n = 20$  y  $p \in \mathcal{P}$ .

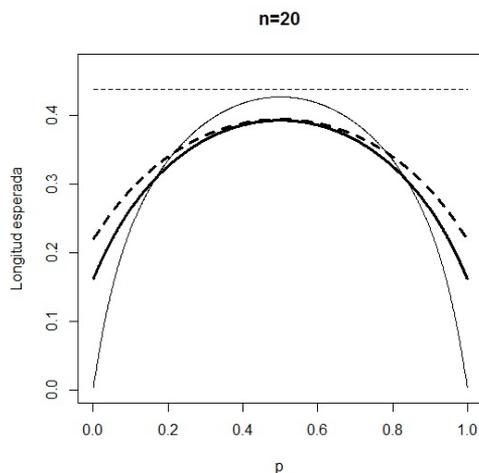


Figura 4.20: Longitud esperada,  $n = 20$  y  $p \in \mathcal{P}$ . (—) intervalo de Wald, (-.-.-) intervalo de Agresti-Coull, (—) intervalo de Wilson (-.-) intervalo arcoseno.

Como puede observarse de la figura 4.20, la longitud esperada del intervalo de Wald es la más corta cuando  $p \leq 0.166$  y cuando  $p \geq 0.834$ , mientras que para  $0.166 < p < 0.834$  el intervalo de Wilson posee la longitud esperada más corta, la longitud esperada del intervalo de Agresti-Coull es mayor a la de Wilson sin embargo para  $p$  cercanos a  $1/2$  son bastante similares. La longitud esperada del intervalo arcoseno es mayor a los otros intervalos y ésta es similar para todo  $p$ .

El cuadro 4.6 muestra los factores de comparación para los 4 intervalos para  $n = 20$  y  $p \in \mathcal{P}$ .

Factor \ Intervalo	$I_E$	$I_{AC}$	$I_W$	$I_A$
$\mathcal{PA}$	0.5	80.58	59.56	62.76
$\mathcal{ECPD}$	0.1039	0.0083	0.0129	0.0409
$\mathcal{ECPE}$	0.0027	0.0166	0.0141	0.0190
$\mathcal{LME}$	0.3243	0.3409	0.3255	0.4383

Cuadro 4.6: Factores de comparación de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , para  $n = 20$ ,  $p \in \mathcal{P}$ .

El intervalo que genera un porcentaje de puntos adecuados ( $\mathcal{PA}$ ) mayor es el intervalo de Agresti-Coull, también posee el error de cobertura promedio por defecto ( $\mathcal{ECPD}$ ) más pequeño, sin embargo su error de cobertura promedio por exceso ( $\mathcal{ECPE}$ ) y su longitud media esperada ( $\mathcal{LME}$ ) son superiores a los obtenidos por el intervalo de Wilson. Considerando que el 80.58 % de puntos  $(20, p)$  cumplen que su  $\mathbf{PC} \geq 0.95$  y que su longitud media esperada es mayor que la del intervalo de Wilson, sería una mala elección considerar al intervalo de Agresti-Coull como el idóneo para este caso.

Otro de los comportamientos drásticos del intervalo de Wald y de los demás intervalos es cuando el parámetro  $p$  es pequeño consideremos el caso  $p = 0.005$  y  $n \in [1, 1000]$ , la figura 4.21 muestra las respectivas probabilidades de cobertura para los cuatro intervalos.

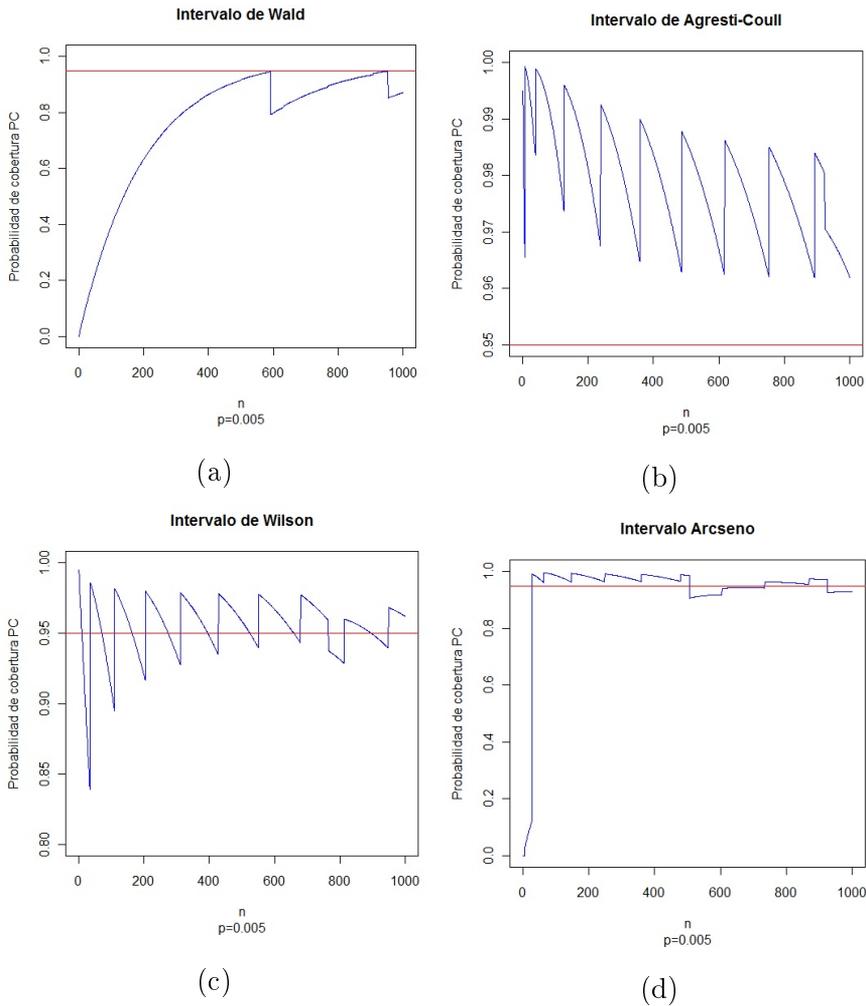


Figura 4.21: Probabilidad de cobertura para los intervalos Wald, Agresti-Coull, Wilson y arcoseno con  $p=0.005$  y  $n \in \{1, 2, \dots, 1000\}$ .

De la figura 4.21 se observa que el intervalo con un mejor desempeño en términos de la probabilidad de cobertura es el intervalo de Wilson ya que el intervalo de Agresti-Coull cumple que para todo  $n \in \llbracket 1, 1000 \rrbracket$  la  $\mathbf{PC}(n, 0.005) \geq 0.9619$  las cuales están muy por encima del nivel nominal. El intervalo arcoseno provee probabilidades de cobertura pobres, para  $n \leq 27$  estas son cercanas a 0 y para  $28 \leq n \leq 506$  su  $\mathbf{PC}(n, p) \geq 0.9612$ .

El cuadro 4.7 muestra los valores de los factores de comparacion de los 4 intervalos para  $p = 0.005$  y  $n \in \llbracket 1, 1000 \rrbracket$ .

Factor \ Intervalo	$I_E$	$I_{AC}$	$I_W$	$I_A$
$\mathcal{PA}$	0	100	67.80	67.50
$\mathcal{ECPD}$	0.1863	0	0.0159	0.0781
$\mathcal{ECP\mathcal{E}}$	0	0.0289	0.0152	0.0259
$\mathcal{LM\mathcal{E}}$	0.0122	0.0335	0.0273	0.1211

Cuadro 4.7: Factores de comparación de los intervalos (sin criterio),  $p = 0.005$ ,  $n \in \llbracket 1, 1000 \rrbracket$ .

Del cuadro 4.7 se observa, que el intervalo con un mejor desempeño en términos de probabilidad de cobertura para  $p = 0.005$  es el intervalo de Wilson ya que en el intervalo de Agresti-Coull cumple siempre que  $\mathbf{PC}(n, p) \geq 0.9612$  y por tanto tiene el error de cobertura promedio por exceso  $\mathcal{ECP\mathcal{E}}$  mayor, el cual es superior al obtenido por el intervalo de Wilson. El intervalo arco seno provee un porcentaje de puntos adecuados  $\mathcal{PA}$  similar al obtenido por el intervalo de Wilson pero el error de cobertura promedio por exceso y el error de cobertura promedio por defecto así como la longitud media esperada son grandes comparados con los obtenidos por el intervalo de Wilson.

Debido a que el intervalo de Wald es bastante inestable y también a que los intervalos alternativos continúan presentando inconsistencia. A continuación se analizará cuál de los intervalos alternativos debería ser la opción más recomendada para cualquier tamaño de muestra y observar si alguno de los criterios bajo análisis podría mejorar su desempeño. Para lograr esto, a continuación se compara el comportamiento de los 4 intervalos (Wald, Agresti-Coull, Wilson y arco seno) de la misma manera en la que se comparó únicamente al intervalo de Wald (cuadro 4.2), es decir, se realizará la comparación del desempeño de los

intervalos y criterios al mismo tiempo. Usando las definiciones de los factores de comparación  $\mathcal{PA}_i$  (4.7),  $\mathcal{PCPD}_i$  (4.11),  $\mathcal{PCPE}_i$  (4.14) y  $\mathcal{LME}_i$  (4.15).

Sean  $A_i$  con  $i = 0, 1, 2, \dots, 9$  los conjuntos definidos de la siguiente manera;

$$A_i = \llbracket 100i + 1, 100(i + 1) \rrbracket.$$

Se procederá a calcular los factores de comparación con base en los criterios bajo estudio para  $n \in A_i$  y  $p \in \mathcal{P}$ , los resultados se muestran a continuación.

Algunos conceptos que es conveniente recordar para entender los siguientes cuadros son.

$A_i$  representa el conjunto en el cual  $n$  toma valores,

$\mathcal{PA}$  Porcentaje de puntos  $(n, p)$  tal que  $(n, p)$  cumple el criterio  $i$  y su  $\mathbf{PC}(n, p) \geq 0.95$ ,

$\mathcal{PCPD}$  Probabilidad de cobertura promedio para puntos  $(n, p)$  tal que  $(n, p)$  cumple el criterio  $i$  y su  $\mathbf{PC}(n, p) < 0.95$ ,

$\mathcal{PCPE}$  Probabilidad de cobertura promedio para puntos  $(n, p)$  tal que  $(n, p)$  cumple el criterio  $i$  y su  $\mathbf{PC}(n, p) > 0.95$ ,

$\mathcal{LME}$  Longitud media esperada para puntos  $(n, p)$  tal que  $(n, p)$  cumple el criterio  $i$ .

**Criterio 1 ( $np \geq 5$  y  $n(1-p) \geq 5$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$A_1$	10.69	0.9326	0.9527	0.2383	54.83	0.9439	0.9567	0.2307	67.22	0.9451	0.9582	0.2332	55.46	0.9436	0.9568	0.2731
$A_2$	17.71	0.9399	0.9522	0.1329	53.41	0.9462	0.9543	0.1322	64.9	0.9468	0.9558	0.1331	52.82	0.9458	0.9543	0.1618
$A_3$	21.66	0.9424	0.9519	0.1005	52.88	0.9469	0.9535	0.1003	63.06	0.9474	0.9547	0.1008	52.16	0.9466	0.9534	0.1244
$A_4$	24.31	0.9437	0.9517	0.0842	52.48	0.9473	0.953	0.084	62.05	0.9478	0.954	0.0843	51.15	0.9471	0.953	0.1049
$A_5$	25.9	0.9446	0.9516	0.0739	52.18	0.9476	0.9527	0.0738	60.85	0.948	0.9536	0.074	51.22	0.9474	0.9527	0.0925
$A_6$	27.64	0.9451	0.9515	0.0666	52.4	0.9478	0.9525	0.0665	60.45	0.9482	0.9532	0.0667	51.35	0.9476	0.9524	0.0836
$A_7$	28.69	0.9456	0.9514	0.0611	51.88	0.9479	0.9523	0.0611	59.6	0.9483	0.953	0.0612	51.08	0.9478	0.9522	0.0769
$A_8$	29.73	0.946	0.9513	0.0568	51.96	0.9481	0.9522	0.0568	59.18	0.9484	0.9528	0.0569	50.8	0.9479	0.9521	0.0716
$A_9$	30.6	0.9462	0.9513	0.0533	51.69	0.9482	0.952	0.0533	58.81	0.9485	0.9527	0.0534	50.95	0.9481	0.952	0.0672
$A_{10}$	31.35	0.9465	0.9512	0.0504	51.64	0.9482	0.9519	0.0504	58.45	0.9485	0.9525	0.0504	50.99	0.9481	0.9518	0.0636

(a)

**Criterio 2 ( $np \geq 10$  y  $n(1-p) \geq 10$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$A_1$	13.58	0.9377	0.9527	0.228	53.33	0.9447	0.9557	0.2218	62.29	0.9453	0.9564	0.223	54.26	0.9446	0.9558	0.2501
$A_2$	19.04	0.942	0.9522	0.1377	52.9	0.9463	0.954	0.1366	62.41	0.9468	0.9549	0.1373	52.69	0.9461	0.954	0.1616
$A_3$	22.53	0.9438	0.9519	0.1031	52.63	0.947	0.9532	0.1027	61.43	0.9474	0.954	0.1031	52.54	0.9468	0.9532	0.1244
$A_4$	25.03	0.9448	0.9517	0.0858	52.27	0.9474	0.9528	0.0856	60.89	0.9478	0.9535	0.0859	51.44	0.9473	0.9528	0.1049
$A_5$	26.5	0.9454	0.9516	0.075	52.01	0.9476	0.9526	0.0749	59.94	0.948	0.9532	0.0751	51.48	0.9475	0.9525	0.0925
$A_6$	28.16	0.9459	0.9515	0.0675	52.22	0.9478	0.9523	0.0674	59.7	0.9482	0.9529	0.0676	51.59	0.9477	0.9523	0.0836
$A_7$	29.14	0.9462	0.9514	0.0618	51.77	0.948	0.9522	0.0618	58.96	0.9483	0.9527	0.0619	51.23	0.9479	0.9521	0.0769
$A_8$	30.14	0.9465	0.9513	0.0574	51.72	0.9481	0.952	0.0574	58.61	0.9484	0.9525	0.0575	50.95	0.948	0.952	0.0716
$A_9$	30.98	0.9467	0.9513	0.0538	51.62	0.9482	0.9519	0.0538	58.31	0.9485	0.9524	0.0539	51.06	0.9481	0.9519	0.0672
$A_{10}$	31.67	0.9469	0.9512	0.0508	51.58	0.9483	0.9518	0.0508	58.02	0.9485	0.9523	0.0508	51.09	0.9482	0.9518	0.0636

(b)

Cuadro 4.8: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , Criterio 1 (a), Criterio 2 (b).

Para el criterio 1 (a) tenemos que para todo  $A_i$  los porcentajes de puntos adecuados ( $\mathcal{PA}_i$ ) mayores son obtenidos usando el intervalo de Agresti-Coull así como también las probabilidades de cobertura promedio por defecto ( $\mathcal{PCPD}_i$ ) más cercanas al nivel nominal pero al mismo tiempo provee las probabilidades de cobertura promedio por exceso ( $\mathcal{PCPE}_i$ ) más alejadas de 0.95. Las probabilidades de cobertura promedio por exceso más cercanas a 0.95 son logradas por los intervalos de Wilson y arcoseno. Sin embargo las longitudes medias esperas  $\mathcal{LME}_i$  del intervalo arcoseno son demasiado grandes comparadas con las  $\mathcal{LME}_i$  del intervalo de Wilson, las cuales son las mínimas. La diferencia entre las probabilidades de cobertura por exceso del intervalo de Agresti-Coull y el intervalo de Wilson es 0.0015 en  $A_1$  y decrecen hasta 0.0006 en  $A_{10}$ .

Para el criterio 2 (b) tenemos las mismas observaciones y en este criterio las diferencias entre las probabilidades de cobertura promedio por exceso del intervalo de Agresti-Coull y el intervalo de Wilson son 0.0009 en  $A_1$  y decrecen hasta 0.0005 en  $A_{10}$ .

**Criterio 3 ( $np(1-p) \geq 5$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$
$A_1$	11.49	0.9346	0.9527	0.2252	54.73	0.9443	0.9562	0.2194	66.18	0.9453	0.9576	0.2215	55.01	0.9439	0.9563	0.2578
$A_2$	17.77	0.9401	0.9522	0.1331	53.28	0.9462	0.9543	0.1323	64.79	0.9468	0.9557	0.1333	52.83	0.9458	0.9543	0.1618
$A_3$	21.68	0.9425	0.9519	0.1006	52.83	0.9469	0.9535	0.1003	63.02	0.9474	0.9547	0.1008	52.2	0.9466	0.9534	0.1244
$A_4$	24.32	0.9438	0.9517	0.0842	52.46	0.9473	0.953	0.0841	62.04	0.9478	0.954	0.0844	51.17	0.9471	0.953	0.1049
$A_5$	25.91	0.9446	0.9516	0.0739	52.16	0.9476	0.9527	0.0738	60.84	0.948	0.9536	0.074	51.23	0.9474	0.9527	0.0925
$A_6$	27.65	0.9452	0.9515	0.0666	52.39	0.9478	0.9525	0.0665	60.44	0.9482	0.9532	0.0667	51.36	0.9476	0.9524	0.0836
$A_7$	28.69	0.9456	0.9514	0.0611	51.87	0.9479	0.9523	0.0611	59.59	0.9483	0.953	0.0612	51.08	0.9478	0.9522	0.0769
$A_8$	29.74	0.946	0.9513	0.0568	51.96	0.9481	0.9522	0.0568	59.17	0.9484	0.9528	0.0569	50.8	0.948	0.9521	0.0716
$A_9$	30.6	0.9462	0.9513	0.0533	51.69	0.9482	0.952	0.0533	58.81	0.9485	0.9527	0.0534	50.95	0.9481	0.952	0.0672
$A_{10}$	31.35	0.9465	0.9512	0.0504	51.64	0.9482	0.9519	0.0504	58.45	0.9485	0.9525	0.0504	50.99	0.9481	0.9518	0.0636

(c)

**Criterio 4 ( $np(1-p) \geq 10$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$	$PA$	$PCPD$	$PCPE$	$\mathcal{LME}$
$A_1$	15.13	0.9394	0.9527	0.2122	53.59	0.9449	0.9552	0.2075	61.41	0.9455	0.9557	0.2085	54.57	0.9449	0.9553	0.2316
$A_2$	19.23	0.9421	0.9522	0.1384	52.95	0.9463	0.954	0.1373	62.26	0.9468	0.9547	0.1379	52.54	0.9461	0.954	0.1615
$A_3$	22.62	0.9438	0.9519	0.1033	52.75	0.947	0.9532	0.1029	61.3	0.9474	0.954	0.1033	52.62	0.9468	0.9532	0.1244
$A_4$	25.07	0.9448	0.9517	0.0859	52.27	0.9474	0.9528	0.0857	60.82	0.9478	0.9535	0.086	51.5	0.9473	0.9528	0.1049
$A_5$	26.53	0.9454	0.9516	0.0751	51.98	0.9476	0.9525	0.075	59.89	0.948	0.9532	0.0752	51.54	0.9475	0.9525	0.0925
$A_6$	28.19	0.9459	0.9515	0.0675	52.19	0.9478	0.9523	0.0674	59.67	0.9482	0.9529	0.0676	51.63	0.9477	0.9523	0.0836
$A_7$	29.15	0.9462	0.9514	0.0619	51.74	0.948	0.9522	0.0618	58.94	0.9483	0.9527	0.0619	51.25	0.9479	0.9521	0.0769
$A_8$	30.15	0.9465	0.9513	0.0574	51.7	0.9481	0.952	0.0574	58.6	0.9484	0.9525	0.0575	50.97	0.948	0.952	0.0716
$A_9$	30.98	0.9467	0.9513	0.0538	51.61	0.9482	0.9519	0.0538	58.3	0.9485	0.9524	0.0539	51.07	0.9481	0.9519	0.0672
$A_{10}$	31.68	0.9469	0.9512	0.0508	51.57	0.9483	0.9518	0.0508	58.01	0.9485	0.9523	0.0508	51.1	0.9482	0.9518	0.0636

(d)

Cuadro 4.9: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , Criterio 3 (c), Criterio 4 (d).

Para el criterio 3 (c) tenemos nuevamente que para todo  $A_i$  los porcentajes de puntos adecuados  $\mathcal{PA}_i$  mayores son obtenidos por el intervalo de Agresti-Coull de manera similar, con él se obtienen las probabilidades de cobertura por defecto  $\mathcal{PCPD}_i$  más cercanas al nivel nominal pero provee las probabilidades de cobertura por exceso  $\mathcal{PCPE}_i$  más alejadas del nivel nominal, éstas son logradas por los intervalos de Wilson y arcoseno, sin embargo las longitudes medias esperadas  $\mathcal{LME}_i$  del intervalo arcoseno son demasiado grandes comparadas a las obtenidas por medio del intervalo de Wilson. Las diferencias entre las probabilidades de cobertura por exceso del intervalo de Agresti-Coull y el intervalo de Wilson se encuentran entre 0.0014 y 0.0006.

Para el criterio 4 (d) tenemos las mismas observaciones que para el criterio 6, las diferencias entre las probabilidades de cobertura por defecto del intervalo de Agresti-Coull y el intervalo de Wilson se encuentran entre 0.0008 y 0.0005.

**Criterio 5 ( $n \geq 100$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$A_1$	11.81	0.9192	0.9528	0.1524	55.46	0.9443	0.9564	0.1524	68.47	0.9459	0.9599	0.1546	56.46	0.9397	0.9575	0.196
$A_2$	16.51	0.9264	0.9522	0.1265	54.26	0.9456	0.9553	0.1264	67.28	0.9468	0.958	0.1278	53.72	0.9438	0.9562	0.1621
$A_3$	20.8	0.9344	0.9519	0.0974	53.5	0.9467	0.9541	0.0974	64.52	0.9474	0.9561	0.098	52.77	0.9462	0.9545	0.1245
$A_4$	23.63	0.9381	0.9517	0.0823	52.98	0.9472	0.9535	0.0823	63.1	0.9478	0.9551	0.0826	51.61	0.9468	0.9538	0.105
$A_5$	25.35	0.9403	0.9516	0.0725	52.34	0.9474	0.953	0.0725	61.68	0.948	0.9545	0.0728	51.59	0.9472	0.9534	0.0925
$A_6$	27.17	0.9418	0.9515	0.0656	52.69	0.9476	0.9528	0.0656	61.13	0.9482	0.9539	0.0658	51.61	0.9474	0.9528	0.0836
$A_7$	28.27	0.9428	0.9514	0.0604	52.19	0.9479	0.9525	0.0604	60.18	0.9483	0.9537	0.0605	51.3	0.9476	0.9527	0.0769
$A_8$	29.37	0.9437	0.9513	0.0562	52.21	0.948	0.9524	0.0562	59.68	0.9484	0.9533	0.0563	51.05	0.9478	0.9525	0.0716
$A_9$	30.28	0.9443	0.9513	0.0528	51.9	0.9481	0.9522	0.0528	59.25	0.9485	0.9531	0.0529	51.09	0.9479	0.9523	0.0672
$A_{10}$	31.03	0.9449	0.9512	0.0499	51.08	0.9481	0.9521	0.0499	58.86	0.9485	0.9529	0.05	51.1	0.9479	0.9522	0.0636

(e)

**Criterio 6 ( $n \geq 50$  y  $0.2 < p < 0.8$ )**

	Intervalo de Wald				Intervalo de Wilson				Intervalo de Agresti-Coull				Intervalo arcoseno			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$A_1$	15.89	0.94	0.9527	0.2134	52.83	0.9452	0.9551	0.2087	59.61	0.9456	0.9555	0.2096	53.91	0.9452	0.9551	0.2298
$A_2$	22.87	0.9442	0.9523	0.1512	51.88	0.9466	0.9535	0.1495	56.38	0.9468	0.9537	0.1498	52.74	0.9466	0.9536	0.1621
$A_3$	27.51	0.9461	0.9519	0.1163	51.56	0.9474	0.9527	0.1155	54.72	0.9475	0.9528	0.1157	52.44	0.9474	0.9527	0.1245
$A_4$	30.39	0.9469	0.9517	0.0981	51.27	0.9478	0.9523	0.0977	54.05	0.9479	0.9523	0.0977	51.88	0.9478	0.9523	0.105
$A_5$	31.83	0.9474	0.9516	0.0865	50.63	0.948	0.952	0.0862	53.16	0.9481	0.9521	0.0862	51.35	0.9481	0.952	0.0925
$A_6$	33.84	0.9477	0.9514	0.0782	51.1	0.9482	0.9518	0.078	53.29	0.9483	0.9518	0.078	51.62	0.9482	0.9518	0.0836
$A_7$	34.76	0.9479	0.9514	0.072	50.61	0.9484	0.9517	0.0718	52.72	0.9484	0.9517	0.0718	51.14	0.9484	0.9517	0.0769
$A_8$	35.74	0.9481	0.9513	0.067	50.77	0.9485	0.9515	0.0668	52.65	0.9485	0.9516	0.0669	51.22	0.9485	0.9515	0.0716
$A_9$	36.5	0.9482	0.9512	0.0629	50.74	0.9486	0.9515	0.0628	52.49	0.9486	0.9515	0.0628	51.19	0.9486	0.9515	0.0672
$A_{10}$	37.22	0.9483	0.9512	0.0595	50.55	0.9486	0.9514	0.0594	52.19	0.9487	0.9514	0.0594	51.32	0.9486	0.9514	0.0636

(f)

Cuadro 4.10: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , Criterio 5 (e), Criterio 6 (f).

Para el criterio 5 (e) tenemos que para todo  $A_i$  los porcentajes de puntos adecuados  $\mathcal{PA}_i$  mayores son obtenidos por medio del intervalo de Agresti-Coull de manera similar con él se obtienen las probabilidades de cobertura promedio por defecto  $\mathcal{PCPD}_i$  más cercanas al nivel nominal pero provee las probabilidades de cobertura promedio por exceso  $\mathcal{PCPE}_i$  más alejadas del nivel nominal, estas son logradas por el intervalo de Wilson y arcoseno, sin embargo las longitudes medias esperadas  $\mathcal{LME}_i$  del intervalo arcoseno son demasiado grandes comparadas a las obtenidas por medio del intervalo de Wilson. Las diferencias entre las probabilidades de cobertura promedio por exceso del intervalo de Agresti-Coull y del intervalo de Wilson se encuentran entre 0.0035 y 0.0008.

Para el criterio 6 (f) tenemos las mismas observaciones que para el criterio 5, sin embargo, las diferencias entre las las probabilidades de cobertura promedio por exceso del intervalo de Agresti-Coull y el intervalo de Wilson se encuentran entre 0.0004 y 0.

En este apartado se presentan estudios similares a los anteriores para tamaños de muestra grandes  $25 \leq n \leq 10000$ . Esto con el objetivo de corroborar que las inconsistencias del intervalo de Wald no desaparecen para tales tamaños de muestra y de analizar el comportamiento de los intervalos alternativos para tamaños de muestra muy grandes.

CRITERIO 1 ( $np \geq 5$ y $n(1-p) \geq 5$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25,1000 \rrbracket$	25.47	0.9435	0.9516	0.0847	61.19	0.9479	0.9539	0.0846	52.44	0.9474	0.953	0.0842	51.62	0.9472	0.9529	0.1043
$\llbracket 1001,2000 \rrbracket$	33.9	0.9472	0.951	0.0405	57.15	0.9488	0.952	0.0406	51.39	0.9486	0.9516	0.0418	50.62	0.9485	0.9515	0.0513
$\llbracket 2001,3000 \rrbracket$	36.83	0.948	0.9508	0.031	55.8	0.949	0.9515	0.0311	51.13	0.9489	0.9512	0.031	50.55	0.9488	0.9512	0.0394
$\llbracket 3001,4000 \rrbracket$	38.48	0.9484	0.9507	0.0262	55.02	0.9492	0.9513	0.0262	50.98	0.949	0.951	0.0261	50.44	0.949	0.951	0.0332
$\llbracket 4001,5000 \rrbracket$	39.65	0.9486	0.9507	0.023	54.41	0.9493	0.9511	0.023	50.79	0.9492	0.9509	0.023	50.31	0.9491	0.9509	0.0293
$\llbracket 5001,6000 \rrbracket$	40.4	0.9487	0.9506	0.0208	54.2	0.9493	0.951	0.0208	50.82	0.9492	0.9508	0.0208	50.35	0.9492	0.9492	0.0265
$\llbracket 6001,7000 \rrbracket$	41.00	0.9488	0.9506	0.0191	53.9	0.9494	0.9509	0.0191	50.77	0.9493	0.9508	0.0191	50.36	0.9492	0.9507	0.0243
$\llbracket 7001,8000 \rrbracket$	41.56	0.949	0.9505	0.0178	53.6	0.9494	0.9509	0.0178	50.73	0.9493	0.9507	0.0178	50.42	0.9493	0.9507	0.0226
$\llbracket 8001,9000 \rrbracket$	41.97	0.949	0.9505	0.0167	53.44	0.9494	0.9508	0.0167	50.64	0.9494	0.9507	0.0167	50.28	0.9493	0.9506	0.0213
$\llbracket 9001,10000 \rrbracket$	42.46	0.9491	0.9505	0.0158	53.31	0.9495	0.9507	0.0158	50.7	0.9494	0.9506	0.0158	50.29	0.9494	0.9506	0.0206

(a)

CRITERIO 2 ( $np \geq 10$ y $n(1-p) \geq 10$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25,1000 \rrbracket$	26.45	0.9449	0.9516	0.084	59.9	0.9479	0.9533	0.0837	52.14	0.9475	0.9527	0.0835	51.69	0.9474	0.9527	0.1012
$\llbracket 1001,2000 \rrbracket$	34.14	0.9476	0.951	0.0408	56.84	0.9488	0.9518	0.0408	51.31	0.9486	0.9515	0.0421	50.68	0.9485	0.9515	0.0513
$\llbracket 2001,3000 \rrbracket$	36.98	0.9482	0.9508	0.0312	55.62	0.949	0.9514	0.0312	51.09	0.9489	0.9512	0.0311	50.59	0.9489	0.9512	0.0394
$\llbracket 3001,4000 \rrbracket$	38.58	0.9485	0.9507	0.0262	54.9	0.9492	0.9512	0.0262	50.93	0.9491	0.951	0.0261	50.44	0.949	0.951	0.0332
$\llbracket 4001,5000 \rrbracket$	39.73	0.9487	0.9507	0.0231	54.32	0.9493	0.951	0.0231	50.76	0.9492	0.9509	0.0231	50.33	0.9491	0.9509	0.0293
$\llbracket 5001,6000 \rrbracket$	40.48	0.9488	0.9506	0.0208	54.11	0.9493	0.951	0.0208	50.82	0.9492	0.9508	0.0208	50.4	0.9492	0.9492	0.0265
$\llbracket 6001,7000 \rrbracket$	41.08	0.9489	0.9506	0.0192	53.8	0.9494	0.9509	0.0192	50.74	0.9493	0.9507	0.0192	50.39	0.9493	0.9507	0.0243
$\llbracket 7001,8000 \rrbracket$	41.64	0.949	0.9505	0.0178	53.51	0.9494	0.9508	0.0178	50.69	0.9493	0.9507	0.0178	50.37	0.9493	0.9507	0.0226
$\llbracket 8001,9000 \rrbracket$	42.05	0.9491	0.9505	0.0167	53.34	0.9494	0.9507	0.0168	50.63	0.9494	0.9506	0.0167	50.31	0.9494	0.9506	0.0213
$\llbracket 9001,10000 \rrbracket$	42.54	0.9492	0.9505	0.0158	53.22	0.9495	0.9507	0.0158	50.67	0.9494	0.9506	0.0158	50.31	0.9494	0.9506	0.0206

(b)

Cuadro 4.11: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , criterio 1 (a), criterio 2 (b).

CRITERIO 3 ( $np(1-p) \geq 5$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25,1000 \rrbracket$	25.54	0.9437	0.9516	0.0844	61.11	0.9479	0.9539	0.0843	52.4	0.9474	0.9529	0.0839	51.62	0.9472	0.9529	0.1038
$\llbracket 1001,2000 \rrbracket$	33.9	0.9472	0.951	0.0405	57.15	0.9488	0.952	0.0406	51.39	0.9486	0.9516	0.0418	50.62	0.9485	0.9515	0.0513
$\llbracket 2001,3000 \rrbracket$	36.83	0.948	0.9508	0.031	55.8	0.949	0.9515	0.0311	51.13	0.9489	0.9512	0.031	50.55	0.9488	0.9512	0.0394
$\llbracket 3001,4000 \rrbracket$	38.48	0.9484	0.9507	0.0262	55.02	0.9492	0.9513	0.0262	50.98	0.949	0.951	0.0262	50.44	0.949	0.951	0.0332
$\llbracket 4001,5000 \rrbracket$	39.65	0.9486	0.9507	0.023	54.41	0.9493	0.9511	0.023	50.79	0.9492	0.9509	0.023	50.31	0.9491	0.9509	0.0293
$\llbracket 5001,6000 \rrbracket$	40.4	0.9487	0.9506	0.0208	54.2	0.9493	0.951	0.0208	50.82	0.9492	0.9508	0.0208	50.35	0.9492	0.9492	0.0265
$\llbracket 6001,7000 \rrbracket$	41.00	0.9488	0.9506	0.0191	53.9	0.9494	0.9509	0.0191	50.77	0.9493	0.9508	0.0191	50.36	0.9492	0.9507	0.0243
$\llbracket 7001,8000 \rrbracket$	41.56	0.949	0.9505	0.0178	53.6	0.9494	0.9509	0.0178	50.73	0.9493	0.9507	0.0178	50.42	0.9493	0.9507	0.0226
$\llbracket 8001,9000 \rrbracket$	41.97	0.949	0.9505	0.0167	53.44	0.9494	0.9508	0.0167	50.64	0.9494	0.9507	0.0167	50.28	0.9493	0.9506	0.0213
$\llbracket 9001,10000 \rrbracket$	42.46	0.9491	0.9505	0.0158	53.31	0.9495	0.9507	0.0158	50.7	0.9494	0.9506	0.0158	50.29	0.9494	0.9506	0.0206

(c)

CRITERIO 4 ( $np(1-p) \geq 10$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25,1000 \rrbracket$	26.72	0.9451	0.9515	0.0816	59.77	0.9479	0.9532	0.0815	52.13	0.9476	0.9526	0.0812	51.68	0.9474	0.9526	0.0985
$\llbracket 1001,2000 \rrbracket$	34.14	0.9476	0.951	0.0408	56.84	0.9488	0.9518	0.0408	51.3	0.9486	0.9515	0.0421	50.68	0.9485	0.9515	0.0513
$\llbracket 2001,3000 \rrbracket$	36.98	0.9482	0.9508	0.0312	55.62	0.949	0.9514	0.0312	51.09	0.9489	0.9512	0.0311	50.59	0.9489	0.9512	0.0394
$\llbracket 3001,4000 \rrbracket$	38.58	0.9485	0.9507	0.0262	54.9	0.9492	0.9512	0.0262	50.93	0.9491	0.951	0.0261	50.44	0.949	0.951	0.0332
$\llbracket 4001,5000 \rrbracket$	39.73	0.9487	0.9507	0.0231	54.32	0.9493	0.951	0.0231	50.76	0.9492	0.9509	0.0231	50.33	0.9491	0.9509	0.0293
$\llbracket 5001,6000 \rrbracket$	40.48	0.9488	0.9506	0.0208	54.11	0.9493	0.951	0.0208	50.81	0.9492	0.9508	0.0208	50.4	0.9492	0.9492	0.0265
$\llbracket 6001,7000 \rrbracket$	41.08	0.9489	0.9506	0.0192	53.8	0.9494	0.9509	0.0192	50.74	0.9493	0.9507	0.0192	50.39	0.9493	0.9507	0.0243
$\llbracket 7001,8000 \rrbracket$	41.64	0.949	0.9505	0.0178	53.51	0.9494	0.9508	0.0178	50.69	0.9493	0.9507	0.0178	50.37	0.9493	0.9507	0.0226
$\llbracket 8001,9000 \rrbracket$	42.05	0.9491	0.9505	0.0167	53.34	0.9494	0.9507	0.0168	50.63	0.9494	0.9506	0.0167	50.31	0.9494	0.9506	0.0213
$\llbracket 9001,10000 \rrbracket$	42.54	0.9492	0.9505	0.0158	53.22	0.9495	0.9507	0.0158	50.67	0.9494	0.9506	0.0158	50.31	0.9494	0.9506	0.0206

(d)

Cuadro 4.12: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , criterio 3 (c), criterio 4 (d).

CRITERIO 5 ( $n \geq 100$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25, 1000 \rrbracket$	25.81	0.9392	0.9515	0.0738	61.75	0.948	0.9546	0.0742	52.64	0.9474	0.9531	0.0738	51.77	0.947	0.9534	0.0942
$\llbracket 1001, 2000 \rrbracket$	33.7	0.9463	0.951	0.0403	57.4	0.9488	0.9523	0.0404	51.51	0.9485	0.9517	0.0416	50.71	0.9484	0.9517	0.0513
$\llbracket 2001, 3000 \rrbracket$	36.72	0.9478	0.9508	0.031	55.93	0.949	0.9516	0.031	51.19	0.9489	0.9513	0.031	50.56	0.9487	0.9513	0.0394
$\llbracket 3001, 4000 \rrbracket$	38.4	0.9481	0.9507	0.0261	55.11	0.9492	0.9514	0.0261	51.02	0.949	0.9511	0.0261	50.41	0.949	0.951	0.0332
$\llbracket 4001, 5000 \rrbracket$	39.58	0.9485	0.9507	0.023	54.5	0.9493	0.9512	0.023	50.82	0.9491	0.9509	0.023	50.33	0.9491	0.9509	0.0293
$\llbracket 5001, 6000 \rrbracket$	40.4	0.9487	0.9506	0.0208	54.2	0.9493	0.951	0.0208	50.82	0.9492	0.9508	0.0208	50.35	0.9492	0.9492	0.0265
$\llbracket 6001, 7000 \rrbracket$	41.00	0.9488	0.9506	0.0191	53.9	0.9494	0.9509	0.0191	50.77	0.9493	0.9508	0.0191	50.36	0.9492	0.9507	0.0243
$\llbracket 7001, 8000 \rrbracket$	41.56	0.949	0.9505	0.0178	53.6	0.9494	0.9509	0.0178	50.73	0.9493	0.9507	0.0178	50.42	0.9493	0.9507	0.0226
$\llbracket 8001, 9000 \rrbracket$	41.97	0.949	0.9505	0.0167	53.44	0.9494	0.9508	0.0167	50.64	0.9494	0.9507	0.0167	50.28	0.9493	0.9506	0.0213
$\llbracket 9001, 10000 \rrbracket$	42.46	0.9491	0.9505	0.0158	53.31	0.9495	0.9507	0.0158	50.7	0.9494	0.9506	0.0158	50.29	0.9494	0.9506	0.0206

(e)

CRITERIO 6 ( $n \geq 50$ y $0.2 < p < 0.8$ )																
	WALD				AGRESTI-COULL				WILSON				ARCOSENO			
	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$	$PA$	$PCPD$	$PCPE$	$LM\mathcal{E}$
$\llbracket 25, 1000 \rrbracket$	31.41	0.9467	0.9515	0.0947	53.84	0.948	0.9523	0.0941	51.11	0.9479	0.9522	0.094	51.78	0.9479	0.9522	0.1014
$\llbracket 1001, 2000 \rrbracket$	39.26	0.9487	0.951	0.0481	51.78	0.9489	0.9511	0.048	50.48	0.9489	0.9511	0.0495	50.79	0.9489	0.9511	0.0513
$\llbracket 2001, 3000 \rrbracket$	41.62	0.9491	0.9508	0.0369	51.4	0.9492	0.9509	0.0369	50.37	0.9492	0.9508	0.0369	50.69	0.9492	0.9508	0.0394
$\llbracket 3001, 4000 \rrbracket$	42.84	0.9492	0.9507	0.0311	51.16	0.9493	0.9507	0.0311	50.3	0.9493	0.9507	0.0311	50.51	0.9493	0.9507	0.0332
$\llbracket 4001, 5000 \rrbracket$	43.66	0.9493	0.9506	0.0274	51.02	0.9494	0.9506	0.0274	50.26	0.9494	0.9506	0.0274	50.5	0.9494	0.9506	0.0293
$\llbracket 5001, 6000 \rrbracket$	44.26	0.9494	0.9505	0.0248	50.98	0.9494	0.9506	0.0248	50.31	0.9494	0.9506	0.0248	50.46	0.9494	0.9496	0.0265
$\llbracket 6001, 7000 \rrbracket$	44.65	0.9494	0.9505	0.0228	50.85	0.9495	0.9505	0.0228	50.21	0.9495	0.9505	0.0228	50.46	0.9495	0.9505	0.0243
$\llbracket 7001, 8000 \rrbracket$	45.00	0.9495	0.9505	0.0212	50.82	0.9495	0.9505	0.0212	50.23	0.9495	0.9505	0.0212	50.35	0.9497	0.9505	0.0226
$\llbracket 8001, 9000 \rrbracket$	45.33	0.9495	0.9504	0.0199	50.76	0.9495	0.9505	0.0199	50.2	0.9495	0.9505	0.0199	50.3	0.9495	0.9505	0.0213
$\llbracket 9001, 10000 \rrbracket$	45.67	0.9495	0.9504	0.0188	50.8	0.9496	0.9504	0.0188	50.27	0.9496	0.9504	0.0188	50.38	0.9496	0.9504	0.0206

(f)

Cuadro 4.13: Factores de  $I_E$ ,  $I_{AC}$ ,  $I_W$  y  $I_A$ , criterio 5 (e), criterio 6 (f).

Como puede observarse, el desempeño inconsistente del intervalo de Wald continúa presente aún para tamaños de muestra bastante grandes  $1000 < n < 10,000$ ; además, los criterios sugeridos para el intervalo de Wald aplicado a estos tamaños de muestra continúan siendo inadecuados. Las probabilidades de cobertura de los intervalos de Wilson y Agresti-Coull para estos valores son bastante similares y cercanas al nivel de confianza sobre todo para  $p \in (0.2, 0.8)$ . Las probabilidades de cobertura del intervalo arco seno se encuentran también cercanas al nivel de confianza pero su longitud media esperada es bastante grande comparada con los demás intervalos.

# Capítulo 5

## Conclusiones

A continuación se exhiben las principales conclusiones obtenidas en el presente trabajo.

1. Los resultados de este trabajo refuerzan las conclusiones de diversos artículos que tratan este tema (Khurshid & Ageel [16]), el intervalo de Wald no tiene un buen desempeño en términos de probabilidades de cobertura (ya que las probabilidades de cobertura se comportan de forma errática y por debajo del nivel nominal en muchos casos). Adicionalmente, se concluye que los criterios que se sugieren para su aplicación no son del todo adecuados ya que las inconsistencias continúan presentándose aún cuando estos criterios son aplicados.
2. Los intervalos alternativos producen probabilidades de cobertura mejores que el intervalo de Wald y esto es más notable cuando  $p$  está cerca del 0 o del 1 o cuando el tamaño muestral  $n$  es pequeño. Los criterios sugeridos para el intervalo de Wald mejoran aún más el desempeño de los intervalos alternativos, sobre todo los criterios 4 y 6 en los cuales las probabilidades de cobertura están bastante cerca del nivel de confianza y además, para el criterio 6 las probabilidades de cobertura obtenidas por los intervalos

de Wilson y Agresti-Coull son muy similares y podría hacerse uso de ambos indistintamente.

3. Para  $p \leq 0.2$  o  $p \geq 0.8$  en los cuales sea difícil o costoso cumplir con alguno de los criterios 4 o 6, se recomienda usar el intervalo de Wilson para  $n < 200$  y el intervalo de Agresti-Coull  $n \geq 200$  y si se aplica el criterio 4, el resultado será aún mejor.
4. En caso de que no sea posible cumplir con alguno de los criterios 4 o 6, se recomienda usar el intervalo de Wilson para  $n < 200$  y el intervalo de Agresti-Coull para  $n \geq 200$ , esto es debido a la dificultad del manejo teórico de los conceptos básicos en cualquier curso de licenciatura que aplica la metodología a un conjunto de datos. Es importante notar que, para  $n \geq 200$  la probabilidad de cobertura promedio por exceso del intervalo de Agresti-Coull es mayor que la probabilidad de cobertura promedio por exceso del intervalo de Wilson solo por 0.002. Brown et al. en ([8]) recomiendan usar el intervalo de Wilson para  $n < 30$  y el intervalo de Agresti-Coull para  $n \geq 30$ , sin embargo para valores de  $p$  cercanos a 0 o a 1 la probabilidad de cobertura del intervalo de Agresti-Coull está muy cercana a 1, aún para valores de  $n$  mayores que 30.
5. Para  $p$  con  $0.2 < p < 0.8$  y  $n \geq 100$  también se podría recomendar el intervalo arcoseno, esté intervalo provee probabilidades de cobertura cercanas al nivel de confianza, sin embargo su longitud esperada es demasiado superior a la longitud esperada de los intervalos de Wald, Wilson y Agresti-Coull.

# Apéndice A

## Programas para la comparación de intervalos de confianza

# FUNCIÓN PARA CALCULAR LAS PROBABILIDADES DE COBERTURA. ASI COMO  $\mathcal{PA}$ ,  $\mathcal{ECPD}$ ,  $\mathcal{PCPD}$ ,  $\mathcal{ECP\mathcal{E}}$  y  $\mathcal{PCPE}$  ESTO PARA EL INTERVALO DE WALD. (PARA CALCULAR LOS FACTORES DE COMPARACION DE LOS OTROS INTERVALOS USAR LAS EXPRESIONES (4.18),(4.21) y (4.24).

```
CovPro <- function(n,p,Alpha){
z = qnorm(1 - Alpha/2, 0, 1)
c =  $\frac{z^2}{n}$ 
x1e <- ceiling(n * ((p + c/2)/(1 + c)) - n * (sqrt(-c * p^2 * c + c^2/4)/(1 + c)))
x2e <- floor(n * ((p + c/2)/(1 + c)) + n * (sqrt(-c * p^2 + p * c + c^2/4)/(1 + c)))
print(sum(dbinom(x1e : x2e, n, p)))
}
# n1: Tamaño de prueba inicial
n1 = 100
# n2: Tamaño de prueba final
n2 = 100
# Alpha: Valor de Alpha
Alpha = 0.05
# Delta: Incremento para p
Delta = 0.001
# principio y fin de p
pf = 1/Delta - 1
Vcc <- rep(0,6) #Puntos de la forma (n,p) tal que cumplen el criterio i
Lcp <- rep(0,6) #Puntos de la forma (n,p) tal que cumplen el criterio i y su PC  $\geq$  (1 - Alpha)
VecLcp <- rep(0,6) #Porcentajes de puntos de la forma (n,p) tal que cumplen el criterio i y su PC  $\geq$ 
```

```

(1 - Alpha)
Vcn <- rep(0,6) #Puntos de la forma (n,p) tal que cumplen el criterio i y su PC<(1-Alpha)
s<- rep(0,6) #Error de cobertura para puntos de la forma (n,p) tal que cumplen el criterio i y su PC<(1-Alpha)
VecLcpn <- rep(0,6) #Error de cobertura promedio para puntos de la forma (n,p) tal que cumplen el
criterio i y su PC<(1-Alpha)
VecPcp <- rep(0,6) #Cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio i y su
PC<(1-Alpha)
Vcnn <- rep(0,7) #Puntos de la forma (n,p) tal que cumplen el criterio i y su PC>(1-Alpha)
sn<- rep(0,7) #Error de cobertura para puntos de la forma (n,p) tal que cumplen el criterio i y su PC>(1-Alpha)
VecLcpnn <- rep(0,7) #Error de cobertura promedio para puntos de la forma (n,p) tal que cumplen el
criterio i y su PC>(1-Alpha)
VecPcpn <- rep(0,7) #Cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio i y
su PC>(1-Alpha)
for (n in n1:n2) {
for (pc in 1:pf) {
p=pc*Delta
# Criterion 1
if ( n*p>=5 && n*(1-p)>=5 )
{
Vcc[1]=Vcc[1]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[1]=Lcp[1]+1
}
}
else{
Vcn[1]=Vcn[1]+1
s[1]=s[1]+((1-Alpha)-CovPro(n,p,Alpha))
}
if ( CovPro(n,p,Alpha)>1-Alpha) {
Vcnn[1]=Vcnn[1]+1
sn[1]=sn[1]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
# Criterion 2
if ( n*p>=10 && n*(1-p)>=10 ) {
Vcc[2]=Vcc[2]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[2]=Lcp[2]+1
}
}
else{
Vcn[2]=Vcn[2]+1
s[2]=s[2]+((1-Alpha)-CovPro(n,p,Alpha))
}
}

```

```

if ( CovPro(n,p,Alpha)>1-Alpha) {
Vcnn[2]=Vcnn[2]+1
sn[2]=sn[2]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
# Criterion 3
if ( n*p*(1-p)>=5) {
Vcc[3]=Vcc[3]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[3]=Lcp[3]+1
}
else{
Vcn[3]=Vcn[3]+1
s[3]=s[3]+((1-Alpha)-CovPro(n,p,Alpha))
}
if ( CovPro(n,p,Alpha)>1-Alpha) {
Vcnn[3]=Vcnn[3]+1
sn[3]=sn[3]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
# Criterion 4
if ( n*p*(1-p)>=10 ) {
Vcc[4]=Vcc[4]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[4]=Lcp[4]+1
}
else{
Vcn[4]=Vcn[4]+1
s[4]=s[4]+((1-Alpha)-CovPro(n,p,Alpha))
}
if ( CovPro(n,p,Alpha)>1-Alpha) {
Vcnn[4]=Vcnn[4]+1
sn[4]=sn[4]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
# Criterion 5. We are considering n big when n>=100
if ( n >= 100 ) {
Vcc[5]=Vcc[5]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[5]=Lcp[5]+1
}
else{
Vcn[5]=Vcn[5]+1
s[5]=s[5]+((1-Alpha)-CovPro(n,p,Alpha))
}
if ( CovPro(n,p,Alpha)>1-Alpha) {

```

```

Vcnn[5]=Vcnn[5]+1
sn[5]=sn[5]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
# Criterion 6. We are considering  $p$  small when  $p \leq .2$ ,
# By Symmetry we include the dual situation  $p \geq .8$ 
if ( n >= 50 && p > .2 && p < .8 ) {
Vcc[6]=Vcc[6]+1
if ( CovPro(n,p,Alpha) >= 1-Alpha) {
Lcp[6]=Lcp[6]+1
}
else{
Vcn[6]=Vcn[6]+1
s[6]=s[6]+((1-Alpha)-CovPro(n,p,Alpha))
}
if ( CovPro(n,p,Alpha) > 1-Alpha) {
Vcnn[6]=Vcnn[6]+1
sn[6]=sn[6]+(CovPro(n,p,Alpha)-(1-Alpha)) }
}
}
}
for ( i in 1:6 ) {
VecLcp[i] <- Lcp[i]*100/Vcc[i]
VecLcpn[i] <- s[i]/Vcn[i]
VecPcp[i] <- ((1-Alpha)-VecLcpn[i])
VecLcpnn [i] <- sn[i]/Vcnn[i]
VecPcpn[i] <- ((1-Alpha)+VecLcpnn[i])
}
round(VecLcp,2) #Porcentajes
round(VecLcpn,4) #Error de cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio
i y su  $PC < (1-\text{Alpha})$ 
round(VecPcp,4) #Cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio i y su
 $PC < (1-\text{Alpha})$ 
round(VecLcpnn,4) #Error de cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio
i y su  $PC > (1-\text{Alpha})$ 
round(VecPcpn,4) #Cobertura promedio para puntos de la forma (n,p) tal que cumplen el criterio i y su
 $PC > (1-\text{Alpha})$ 

```

Para calcular otros tamaños de muestra solo modificar  $n_1$  y  $n_2$ .

# FUNCIÓN PARA CALCULAR LA  $\mathcal{LM}\mathcal{E}$  PARA EL INTERVALO DE WALD. (PARA CALCULAR LAS  $\mathcal{LM}\mathcal{E}$  DE LOS OTROS INTERVALOS USAR LAS EXPRESIONES (3.2), (3.3) y (3.4).

```

longcov <-function(n,p,Alpha){
s=0
for(i in 0:n){
k = qnorm(1 - Alpha/2, 0, 1)
s = s + (2 * k * n-1/2 * (i/n * (1 - i/n))1/2 * dbinom(i, n, p))}
print(s)}
# n1: Tamaño de prueba inicial
n1 = 100
# n2: Tamaño de prueba final
n2 = 100
# Alpha: Valor de Alpha
Alpha = 0.05
# Delta: Incremento para p
Delta = 0.001
# principio y fin de p
pf = 1/Delta - 1
Vcc < -rep(0, 6) #Puntos de la forma (n,p) tal que cumplen el criterio i
Lm < -rep(0, 6) #Puntos de la forma (n,p) tal que cumplen el criterio i y su PC ≥ (1 - Alpha)
VecLme < -rep(0, 6) #Porcentaje
for (n in n1:n2)
{
for (pc in 1:pf)
{
p = pc * Delta
# Criterion 1
if ( n*p>=5 && n*(1-p)>=5 ) {
Vcc[1]=Vcc[1]+1
Lm[1]=Lm[1]+longcov(n,p,Alpha)}
}
}
}

```

```

# Criterion 2
if ( n*p>=10 && n*(1-p)>=10 ) {
Vcc[2]=Vcc[2]+1
Lm[2]=Lm[2]+longcov(n,p,Alpha)}
# Criterion 3
if ( n*p*(1-p)>=5 ) {
Vcc[3]=Vcc[3]+1
Lm[3]=Lm[3]+longcov(n,p,Alpha)}
# Criterion 4
if ( n*p*(1-p)>=10 ){
Vcc[4]=Vcc[4]+1
Lm[4]=Lm[4]+longcov(n,p,Alpha)}
# Criterion 5
if ( n>=100 ) {
Vcc[5]=Vcc[5]+1
Lm[5]=Lm[5]+longcov(n,p,Alpha)}
# Criterion 6. We are considering  $p$  small when  $p \leq .2$ ,
# By Symmetry we include the dual situation  $p \geq .8$  if (  $n \geq 50$  &&  $p > .2$  &&  $p < .8$  ) {
Vcc[6]=Vcc[6]+1
Lm[6]=Lm[6]+longcov(n,p,Alpha)}
}
}
for (i in 1:6) {
VecLme[i] <- Lm[i]/Vcc[i]}
round(VecLme,4) #longitud media esperada. Para calcular otros tamaños de muestra solo modificar n1 y
n2.

```

# Bibliografía

- [1] Agresti A. and Coull B., Approximate is Better than Exact for Interval Estimation of Binomial Proportions, *The American Statistician*, Vol. 52, No. 2, 119-126, 1998.
  
- [2] Agresti A. and Caffo B., Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures, *The American Statistician*, Vol. 54, No. 4, 280-288, American Statistical Association, 2000.
  
- [3] Agresti, A. and Min, Y., On Small-Sample Confidence Intervals for Parameters in Discrete Distribution, *Biometrics* 57, 963-971, 2001.
  
- [4] Agresti A. and Minon Y., On sample Size Guidelines for Teaching Inference about the Binomial Parameter in Introductory Statistics, Department of Statistics, University of Florida, Gainesville, Florida, 32611-8545, 2002.
  
- [5] Anscombe F. J., The transformation of Poisson, binomial and negative binomial data, *Biometrika* 35, 246-254, 1948.
  
- [6] Bickel P. y Doksum K., *Mathematical Statistics*, Prentice-Hall, Englewood Cliffs, NJ, 1997.

- [7] Boomsma A., Confidence Intervals for a Binomial Proportion, Department of Statistics & Measurement Theory University of Groningen, 2005.
- [8] Brown L. D., Cai T. T. and DasGupta A., Interval Estimation for a Binomial Proportion, *Statistical Science* 16, 101-133, 2001.
- [9] Brown L. D., Cai T. T. , DasGupta A., Confidence Intervals for a Binomial Proportion and Asymptotic Expansions, *The Annals of Statistics*, 160-201, 2002.
- [10] Casella, G. & Berger, R., *Statistical Inference*, Wadsworth & Brooks, California, 1990.
- [11] Cepeda Cuervo E., Aguilar W., Cervantes V., Corrales M., Díaz I., Rodríguez D., *Intervalos de confianza e intervalos de credibilidad para una proporción*, Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia, 2008.
- [12] Chung K. L., *A course in probability theory*, third edition, Academic Press, New York, 1974.
- [13] Clopper C. J., Pearson, E. S., The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika* 26, 404-413, 1934.
- [14] Hogg R. V., McKean J. W., Craig A. T., *Introduction to Mathematical Statistics*, Pearson Prentice Hall, United States of America, 2005.
- [15] Juárez Hernandez B., *Notas de clase, Introducción a la inferencia estadística, Obtención de estimadores, MM y MV*, 2013.
- [16] Khurshid, A. & Ageel, M., 'Binomial and Poisson Confidence Intervals and its Variants: A Bibliography', *Pakistan Journal Statistical and Operation Research* VI(1), 75-100, 2010.

- [17] Leemis L. M. and Trivedi K S., A Comparison of Approximate Interval Estimators for the Bernoulli Parameter, *stat*, 50(1), 1994.
- [18] Måns Thulin, The cost of using exact confidence intervals for a binomial proportion, Department of Mathematics, Uppsala University, 2013.
- [19] Newcombe R. G., Confidence Intervals for Proportions and Related Measures of Effect Size, Chapman & Hall/CRC Biostatistics Series, 2013.
- [20] Pires M., A., Amado C., Interval estimators for a binomial proportion: comparison of twenty methods, Department of Mathematics and CEMAT, Portugal, 2008.
- [21] Reyes Cervantes. H., Almendra Arao. F., Morales Cortes. M., Tajonar S F., Problemas al usar la aproximación normal en intervalos de confianza suponiendo datos Bernoulli, Memorias Simposio Internacional de Estadística, XXV Version, Armenia 5 al 8 de agosto, 2015. <http://simposioestadistica.unal.edu.co/historico-de-memorias/>
- [22] Rincón L., Curso intermedio de probabilidad, Departamento de matemáticas, Facultad de Ciencias UNAM, 2008.
- [23] Samuels, M. L. and Witmer, J. W. *Statistics for the Life Sciences*, 2nd ed., Prentice Hall, 1999.
- [24] Santner T. J., A note on teaching binomial confidence intervals, *Teaching Statistics*, 1998.
- [25] Schilling M. F. & Doi J. A., A Coverage Probability Approach to Finding an Optimal Binomial Confidence Procedure, *The American Statistician*, 2014.
- [26] Tamhane, A. C. and Dunlop, D. D., *Statistics and Data Analysis from Elementary to Intermediate*. Prentice Hall, Englewood Cliffs, NJ, 2000.

- [27] Wackerly D. D., Mendenhall W. & Scheaffer R. L., Estadística matemática con aplicaciones, séptima edición, 2008.
- [28] Wilson E. B., Probable Inference the Law of Succession and Statistical Inference, Journal of the American Statistical Association, 209-212, 1927.

# Índice alfabético

- Conjunto de valores de  $p$  bajo análisis  $\mathcal{P} = \{0.001, 0.002, \dots, 0.99\}$ , 50
- Error de cobertura por exceso de  $n$  y  $p$   $\mathcal{ECE}(n, p)$ , 53
- Error de cobertura por defecto de  $n$  y  $p$   $\mathcal{ECD}(n, p)$ , 52
- Error de cobertura promedio por defecto  $\mathcal{ECPD}$ , 52
- Error de cobertura promedio por exceso  $\mathcal{ECPÉ}$ , 53
- Intervalo Arcoseno  $I_A$ , 38
- Intervalo de Agresti-Coull  $I_{AC}$ , 37
- Intervalo de Clopper-Pearson  $CI_{CP}$ , 38
- Intervalo de confianza  $\mathbf{IC}$ , 28
- Intervalo de Wald  $I_E$ , 34
- Intervalo de Wilson  $I_W$ , 36
- Longitud esperada de  $n$  y  $p$   $\mathcal{LE}(n, p)$ , 48
- Longitud media esperada  $\mathcal{LME}$ , 48, 54
- Porcentaje de puntos adecuados  $\mathcal{PA}_i$ , 51
- Probabilidad de cobertura para  $n$  y  $p$   $\mathbf{PC}(n, p)$ , 41
- Probabilidad de cobertura promedio por defecto  $\mathcal{PCPD}$ , 53
- Probabilidad de cobertura promedio por exceso  $\mathcal{PCPE}$ , 54
- Probabilidad de cobertura  $\mathbf{PC}$ , 27