

**4 Nuevos métodos para mejorar su
Estimación de una sola población
Proporción en Minitab**



Un problema común en estadística básica es la estimación de la proporción de individuos con una determinada característica de interés en una población. Por ejemplo, un ingeniero de calidad puede querer estimar la proporción de defectos en un gran lote de unidades producidas en masa en un día determinado; Un científico médico puede querer investigar la proporción de individuos en alguna comunidad que fueron vacunados contra un patógeno específico pero experimentaron la enfermedad relacionada de todos modos; Un director de campaña puede estar interesado en la proporción de votantes registrados que tienen la intención de votar por su candidato.

Los métodos de estimación de intervalos más conocidos para este problema son el método de aproximación normal de libro de texto denominado intervalo de confianza (IC) de Wald y el IC exacto de Clopper-Pearson (1934). Por un lado, el IC de Wald es extremadamente liberal en el sentido de que el nivel de confianza real (o probabilidad de cobertura) del IC está muy por debajo del nivel nominal objetivo, especialmente cuando la proporción real es cercana a 0 o 1 (véase la Figura 1). Por otro lado, el IC exacto de

Clopper-Pearson es excesivamente conservador en el sentido de que el nivel de confianza real (o probabilidad de cobertura) del IC está muy por encima del nivel nominal objetivo. Ambos métodos ya no deberían utilizarse para ninguna aplicación práctica (véase Agresti-Coull, 1998; Brown et al., 2001).

En los últimos años, sin embargo, han desempeñado un papel importante en el desarrollo de mejores métodos de IC con mejores probabilidades de cobertura intermedia. Por ejemplo, el IC aproximado de Agresti-Coull es un ajuste en el IC de Wald; el IC exacto de Blaker (2000, 2001) utiliza los límites de confianza de Clopper-Pearson como estimaciones iniciales en un algoritmo numérico iterativo. Teniendo en cuenta estos métodos recientemente mejorados, Minitab ha actualizado la herramienta estadística para estimar una sola proporción de población para incluir los siguientes 4 métodos: el IC de Blaker ajustado y los métodos de prueba, el IC de Wilson/puntuación y los métodos de prueba (con y sin corrección de continuidad), y el IC de Agresti-Coull y los métodos de prueba. Además, para cada uno de estos métodos, Minitab se asegura de que el IC y la prueba produzcan resultados consistentes.

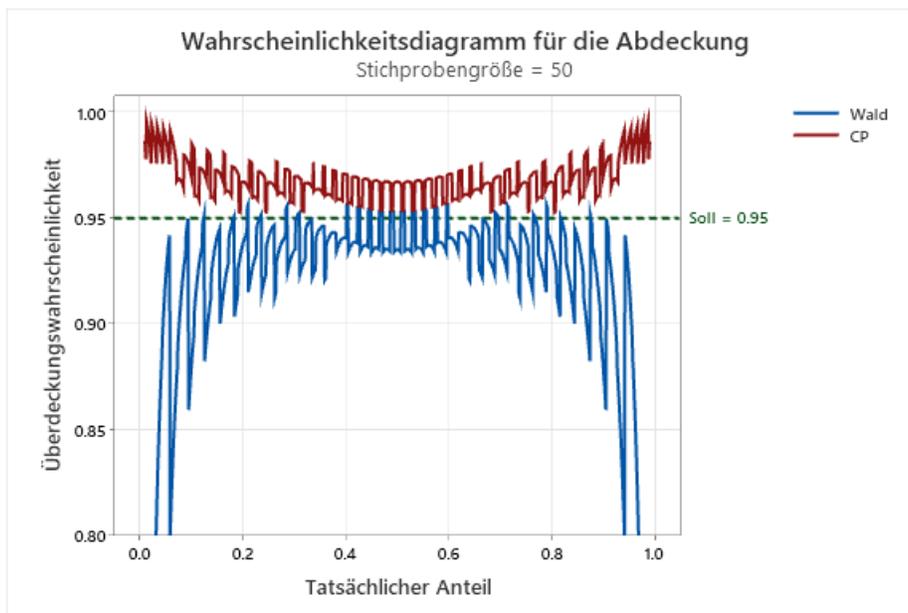


Figura 1: Comparación de las probabilidades de cobertura de los IC de Wald y los IC de Clopper-Pearson (CP) en función de la proporción real cuando el tamaño de la muestra es de 50. El gráfico ilustra que los IC de Wald y los IC de Clopper Pearson son excesivamente liberales y conservadores, respectivamente, particularmente cuando la proporción real está cerca de 0 o 1. Suponiendo que las proporciones verdaderas se distribuyen uniformemente en el intervalo (0, 1), las probabilidades medias de cobertura basadas en una muestra de tamaño 50 son de 0,901 y 0,969 para el IC de Wald y el IC de Clopper-Pearson, respectivamente.

Conoce los 4 nuevos métodos de estimación

Los 4 nuevos métodos consisten en 1 IC exacto y un método de prueba llamado método de Blaker ajustado y 3 IC y métodos de prueba aproximados, incluido el método Wilson/puntuación (Wilson), el método Wilson/puntuación con corrección de continuidad de Yates (Wilson CC) y el método Agresti-Coull (AC). Un método exacto en este contexto significa que no se utiliza ninguna aproximación en la derivación del método, a diferencia de los métodos aproximados que se obtienen utilizando algunas formas de procedimientos de aproximación normales.

1. El método de Blaker ajustado

El método de Blaker ajustado, debido a Klaschka y Reiczigel (2021), es una modificación del IC exacto y los métodos de prueba de Blaker (2000, 2001). La modificación aborda la naturaleza intensiva en computación del algoritmo original de Blaker y las inconsistencias ocasionales entre su CI y los resultados de las pruebas. Al igual que el IC de Blaker original, el IC ajustado resultante es exacto, anidado y está contenido en el IC de Clopper-Pearson. Como resultado, el IC de Blaker ajustado es menos conservador que el IC de Clopper-Pearson. El IC está anidado en el sentido de que un IC con un nivel de confianza más alto siempre contiene un IC con un nivel de confianza más bajo. Por ejemplo, un IC de Blaker bilateral del 95% (ajustado) siempre contiene el correspondiente IC bilateral del 90%. La anidación es una propiedad atractiva de los métodos de IC exactos derivados de una distribución discreta como el binomio. Por ejemplo, el CI de Clopper-Pearson está anidado. Sin embargo, hay métodos de IC exactos disponibles que no están necesariamente anidados. Por ejemplo, el llamado IC Blyth-Still-Casella (Blyth y Still, 1983; Casella, 1986) está garantizado como el IC exacto más corto, pero no está anidado. Crow (1956) CI tampoco está anidado. Los cálculos de IC basados en Blaker o en el método de Blaker ajustado son más complejos que los métodos clásicos de IC antes mencionados porque requieren algoritmos numéricos. Sin embargo, con las innovaciones actuales en la tecnología informática, ya no debemos rehuir la implementación de algoritmos complejos que produzcan mejores resultados. La Figura 2 ilustra las mejoras del IC de Blaker ajustado sobre el IC de Clopper-Pearson.

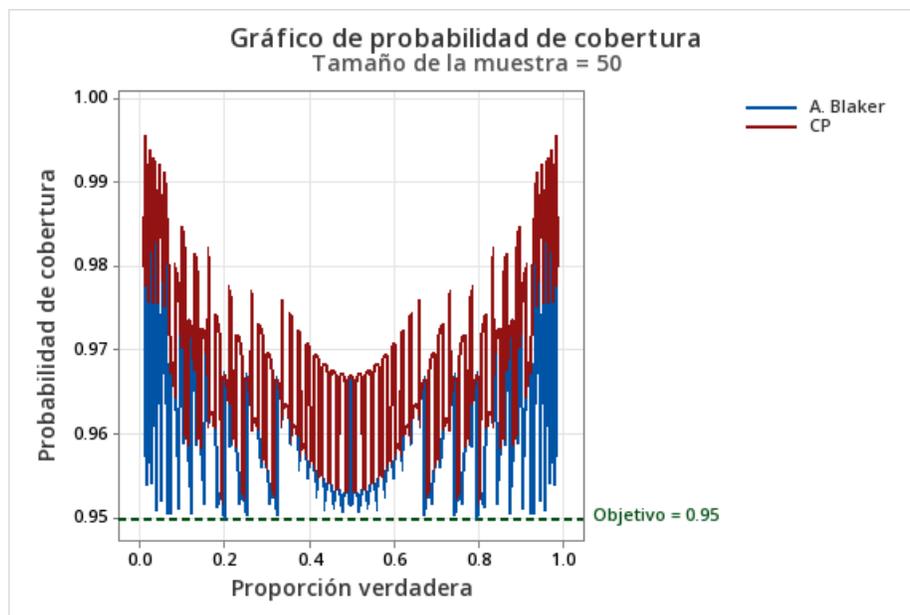


Figura 2: Comparación de las probabilidades de cobertura de los IC ajustados de Blaker (A. Blaker) y los IC de Clopper-Pearson (CP) en función de la proporción real cuando el tamaño de la muestra es de 50. El gráfico indica que la probabilidad de cobertura del IC de Clopper-Pearson es al menos la del IC de Blaker ajustado. Esto es consistente con el hecho de que los IC de Blaker ajustados están contenidos en los IC de Clopper-Pearson. Para cualquier muestra dada de tamaño 50, las probabilidades medias de cobertura (suponiendo que la proporción real se distribuya uniformemente en el intervalo de unidades) son de 0,960 y 0,969 para el IC de Blaker ajustado y el IC de Clopper-Pearson, respectivamente.

2. Los métodos Wilson y Wilson CC

El método de IC de Wilson (1927) se deriva como el IC que corresponde a la prueba de puntuación, la prueba que utiliza el error estándar nulo, $\sqrt{p_0(1-p_0)/n}$, en contraposición al error estándar clásico $\sqrt{p(1-p)/n}$, en el denominador del estadístico de prueba. Por esta razón, también se le conoce como IC Wilson/score. Su probabilidad real de cobertura puede ser menor o mayor que el nivel nominal objetivo, pero permanecer cerca de él, excepto cuando la proporción real es cercana a 0 o 1 (véase la Figura 3). Se puede hacer un ajuste para que el IC de Wilson/puntuación sea conservador utilizando la corrección de continuidad de Yates. Minitab proporciona tanto versiones de IC (con y sin la corrección de continuidad de Yates) como sus pruebas de hipótesis de emparejamiento.

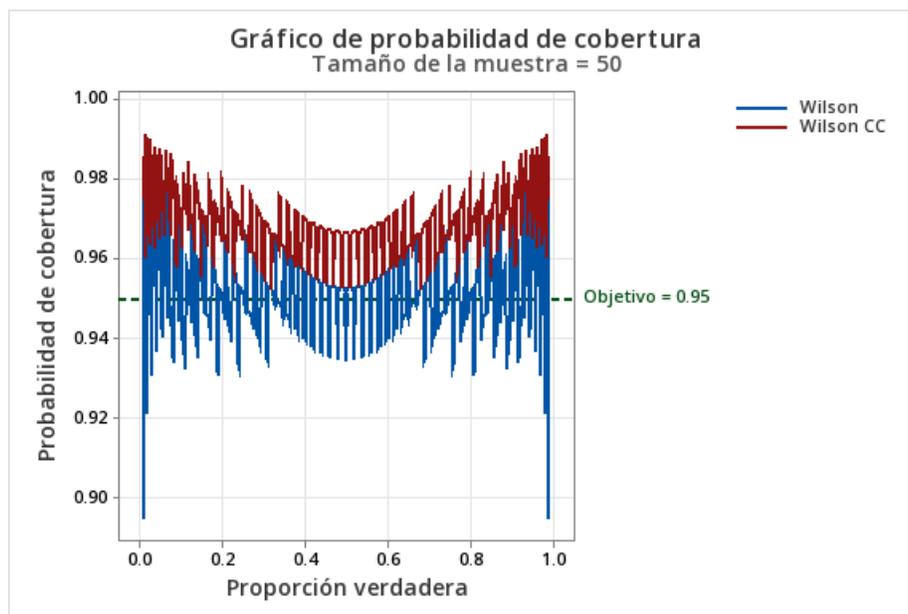


Figura 3: Comparación de las probabilidades de cobertura de los IC Wilson/score (Wilson) y los Wilson/score con la corrección de continuidad de Yates (Wilson CC) en función de la proporción verdadera cuando el tamaño de la muestra es 50. El gráfico muestra que los IC CC de Wilson son siempre conservadores, mientras que los IC de Wilson son conservadores y liberales, dependiendo de la magnitud de la proporción real. En particular, los IC de Wilson tienden a ser demasiado liberales cuando la proporción real está muy cerca de 0 o 1. Para cualquier muestra de tamaño 50, las probabilidades medias de cobertura son de 0,952 y 0,969 para el IC de Wilson y el IC CC de Wilson, respectivamente.

3. El método Agresti-Coull

El IC de Agresti-Coull se obtiene a partir de un ajuste del IC de Wald clásico excesivamente liberal. El IC resultante tiene propiedades de cobertura similares a las de los IC de Wilson, pero un poco más conservador en general. Además, los dos tipos de IC tienen el mismo punto medio, pero los IC de Wilson siempre están contenidos en los IC de Agresti-Coull. Como se ilustra en la Figura 4, tienen esencialmente las mismas probabilidades de cobertura cuando la proporción real es moderada. El IC de Agresti-Coull, sin embargo, es generalmente menos liberal cuando la proporción real es cercana a 0 o 1. Como se muestra en la Figura 4, para una muestra de tamaño 50 el IC de Agresti-Coull se vuelve conservador a medida que la proporción real se acerca a 0 o 1. Otro atractivo de la IC de Agresti-Coull es su simplicidad de implementación heredada de la IC de Wald. Además, es fácil de enseñar y recordar, especialmente cuando el nivel de confianza es del 95%. Para este nivel de confianza, se conoce comúnmente como el método de IC de "sumar 2 éxitos y 2 fracasos" como un memorándum al ajuste realizado al IC de Wald para derivarlo.

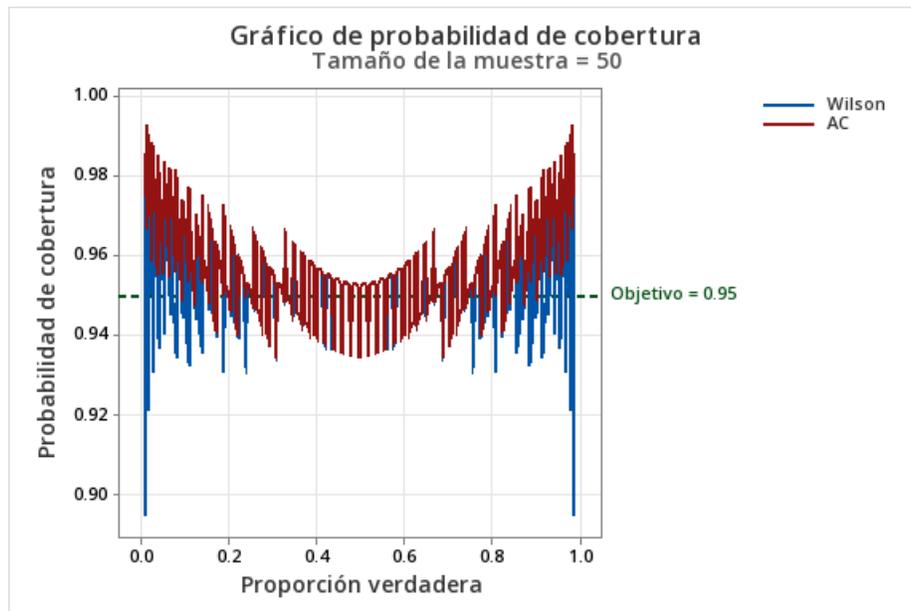


Figura 4: Comparación de las probabilidades de cobertura de los IC Wilson/score (Wilson) y los IC Agresti-Coull (AC) en función de la proporción real cuando el tamaño de la muestra es 50. El gráfico indica que los dos métodos producen esencialmente las mismas probabilidades de cobertura para valores moderados de la proporción verdadera en el intervalo (0,3, 0,75). Sin embargo, para valores de proporción real cercanos a 0 o 1, los IC de Wilson son liberales, mientras que los IC de Agresti-Coull son conservadores. Esto es consistente con el hecho de que los IC de Wilson están contenidos en los IC de Agresti-Coull. Para cualquier muestra dada de tamaño 50, las probabilidades medias de cobertura son de 0,952 y 0,958 para el IC de Wilson y el IC de Agresti-Coull, respectivamente.

Algunos ejemplos ilustrativos sencillos

Un ingeniero de calidad en una fábrica de producción en masa ha seleccionado una muestra aleatoria de 1465 productos producidos en masa en un día determinado. Después de pruebas independientes de los 1465 artículos, se encontró que 53 eran defectuosos. El ingeniero quiere saber si la proporción de artículos defectuosos producidos en un día determinado es significativamente diferente del 2,5%.

Aunque este problema se plantea como una pregunta de prueba de hipótesis estadística, se anima cada vez más a los estadísticos aplicados a incluir en el resultado del análisis una estimación puntual y un intervalo de confianza junto con el valor p de la prueba. Minitab sigue esta convención tanto como sea posible, particularmente en los Módulos de Estadísticas Básicas. Por ejemplo, utilizando Minitab, los resultados del análisis para la pregunta anterior basados en el método de Blaker ajustado son los siguientes.

Prueba e IC para una proporción

Método

p proporción de eventos

Método Se ha ajustado el método exacto de Blaker.

Estadísticos descriptivos

N Muestra de eventos p IC 95% para p

1465 53 0.036177 (0.027353 , 0.046822)

Prueba

Hipótesis nula: $p = 0.025$

Hipótesis alternativa H_1 : $p \neq 0,025$

Valor p

0.009

Se puede generar una salida similar para cada uno de los 4 métodos. Las características y ventajas de cada método se resumen en la tabla siguiente.

Método	IC de 95%	Valor P de la prueba de emparejamiento
Blaker ajustado	(2.74%,4.68%)	0.009
Wilson con la corrección de Yates	(2.75%,4.74%)	0.008
Wilson/puntuación	(2.78%,4.70%)	0.006
Agresti-Coull	(2.77%,4.74%)	0.007

En este ejemplo, todos los métodos arrojan la misma conclusión de que el % defectuoso difiere del 2,5% en el nivel de significación de 0,05, ya que todos los valores p son inferiores a 0,05. Los intervalos de confianza y los valores p correspondientes para todos los métodos son similares, en parte, porque el tamaño de la muestra es muy grande. Además, el IC para cada método no cubre el valor de proporción hipotético (2,5%), lo cual es consistente con el valor p de cada una de las pruebas de hipótesis de emparejamiento.

En el ejemplo anterior, supongamos ahora que el ingeniero de calidad probó solo 50 artículos y descubrió que 2 estaban defectuosos. Además, supongamos que el ingeniero quisiera saber si la proporción de defectuosos era significativamente diferente del 1,0%. Las características y ventajas de cada método se resumen en la tabla siguiente.

Método	IC de 95%	Valor P de la prueba correspondiente
Blaker ajustado	(0.72%,13.35%)	0.089
Wilson con la corrección de Yates	(0.70%,14.86%)	0.155
Wilson/puntuación	(1.10%,13.46%)	0.033
Agresti-Coull	(0.34%,14.22%)	0.124

En este caso, solo el método Wilson/score arroja una conclusión significativa en el sentido de que el % defectuoso difiere del 1,0% en el nivel de significación de 0,05. Al mismo nivel de significación, todos los demás métodos arrojan la conclusión opuesta de

evidencia insuficiente para decidir que existe una diferencia. Las inconsistencias en los resultados entre los métodos se deben, en gran parte, a que el tamaño de la muestra es moderado. En promedio, las probabilidades de cobertura de estos métodos se aproximan al nivel nominal a medida que aumenta el tamaño de la muestra (véase la Figura 5 a continuación). Sin embargo, para diseños de muestras pequeñas a moderadas, las disparidades en las probabilidades medias de cobertura asociadas con cada método son más pronunciadas, lo que hace que los IC correspondientes tengan anchos notablemente diferentes. Una pregunta importante, sin embargo, es ¿qué resultado reportar al jefe? No hay una respuesta directa a esta pregunta, ya que una mejor pregunta puede depender de factores como el conocimiento previo de la magnitud de la proporción real o incluso el área de aplicación. Proporcionaremos algunas pautas generales aproximadas en la siguiente sección.

Conclusión

Las figuras 2, 3 y 4 muestran que los 4 métodos de IC, Blaker ajustado, Wilson, Wilson CC y Agresti-Coull tienen diferentes propiedades de probabilidad de cobertura. El Wilson CC es el más conservador, seguido por el ajustado Blaker. El Agresti Coull y el Wilson son a menudo liberales y conservadores, dependiendo de la magnitud de la verdadera proporción. En general, el método Wilson IC es el más liberal de los 4 métodos. Además, la Figura 5 indica que, en promedio, los 4 métodos son conservadores, siendo el CC de Wilson el más conservador, seguido por el método ajustado de Blaker, el método de Agresti-Coull y el método de Wilson. Para un problema dado, el método apropiado a elegir puede depender de la aplicación particular, el tamaño de la muestra y si se dispone o no de algún conocimiento previo de la magnitud de la proporción real.

Por ejemplo, las agencias reguladoras a menudo están interesadas en métodos conservadores para proteger a los consumidores. Un método demasiado conservador, sin embargo, puede producir regulaciones estrictas, mientras que un método demasiado liberal puede producir regulaciones laxas. En general, para diseños de muestras moderadas a grandes, el método ajustado de Blaker o el método de Willson CC pueden ser buenas opciones. Para diseños de muestras pequeñas, el Wilson o el Agresti-Coull pueden ser apropiados si se dispone de un conocimiento previo del valor de la proporción real. Este conocimiento previo a menudo se basa en experiencias similares previas o en un pequeño estudio piloto diseñado específicamente para obtener una estimación aproximada de la proporción. Por ejemplo, en el área de control de calidad de las aplicaciones, donde la proporción de defectos suele ser cercana a 0, el método a elegir depende de si el investigador desea o no un resultado conservador. Un resultado conservador puede basarse en el método de Agresti-Coull, mientras que un resultado liberal puede basarse en el método de Wilson/puntuación. Por último, pero más importante, siempre es una buena práctica planificar con anticipación el tamaño de la muestra para protegerse contra el error de tipo II o para controlar el ancho de los IC. Minitab tiene "una aplicación para eso también". Hay herramientas disponibles en Minitab para determinar el tamaño adecuado de la muestra para controlar el error de tipo II asociado con las pruebas de hipótesis o para controlar el ancho de los IC.

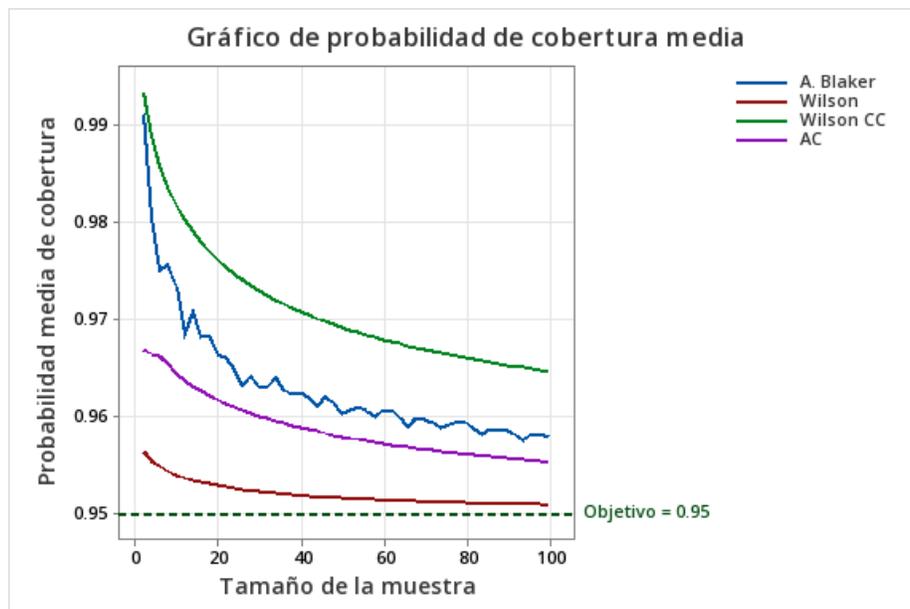


Figura 5: Probabilidad media de cobertura de los 4 métodos de IC en función del tamaño de la muestra. Las probabilidades medias de cobertura se calculan asumiendo que la proporción real se distribuye uniformemente en el intervalo unitario. Las curvas de cobertura media ilustran que, en promedio, el método de IC de Wilson/score con corrección de continuidad de Yates (Wilson CC) es el más conservador, seguido por los métodos ajustados de Blaker (A. Blaker), Agresti-Coull (AC) y Wilson/score (Wilson). Las curvas de probabilidad de cobertura media se acercan al nivel de cobertura nominal objetivo a medida que aumenta el tamaño de la muestra. Además, mientras que las curvas de cobertura media para los métodos aproximados (Wilson CC, Wilson, Agresti-Coull) son suaves, la curva de cobertura para el Blaker ajustado exacto tiene algunos movimientos de ondulación a medida que se acerca a la cobertura nominal. Esto indica, tal vez, que el método ajustado de IC de Blaker puede mejorarse aún más.

Referencia

- Agresti, A. and Coull, B. A. (1998). Approximate is better than “Exact” for interval Estimation of Binomial Proportion. *The American Statistician* 52, 119–125
- Blaker, H. (2000). Confidence Curves and Improved Exact Confidence Intervals for Discrete Distributions. *The Canadian Journal of Statistics*, 28, 783–798
- Blaker, H. (2001). Corrigenda: Confidence curves and improves exact confidence intervals for discrete distributions. *The Canadian Journal of Statistics*, 29, 681.
- Blyth, C. R. and Still, H. A. (1983). Binomial Confidence Intervals. *Journal of the American Statistical Association* 78, 108–116.
- Brown, L. D., Cai, T. and Das Gupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science* 16, 101–133.
- Casella, G., 1986. Refining binomial confidence intervals. *Canad. J. Statist.* 14, 113–129.
- Clopper, C. J. and Pearson, E. S. (1934). The Use of Confidence or Fiducial Limits Illustrated in the Case of Binomial. *Biometrika* 26, 404–413
- Crow, E.L., 1956. Confidence intervals for a proportion. *Biometrika* 43, 423–435.
- Klaschka, J. and Reiczigel, J. (2021). On matching confidence intervals and tests for some discrete distributions: methodological and computational aspects, *Computational Statistics*, Springer, vol. 36(3), 1775-1790.
- Wilson E. B. (1927) Probable Inference, the Law of Successions and Statistical Inference. *J. Amer. Statist. Assoc.* 22, 209–21

© 2024 Minitab, LLC. All rights reserved. Minitab®, Minitab Connect®, Minitab Model Ops®, Minitab Engage®, Minitab Workspace®, Salford Predictive Modeler®, SPM®, and the Minitab® logo are all registered trademarks of Minitab, LLC, in the United States and other countries.

Additional trademarks of Minitab, LLC can be found at www.minitab.com. All other marks referenced remain the property of their respective owners.



Tienes datos. Contamos con soluciones analíticas.

Descargue una versión de
www.minitab.com