



**4 Novos métodos para melhorar o seu  
Estimativa de uma única população  
Proporção no Minitab**

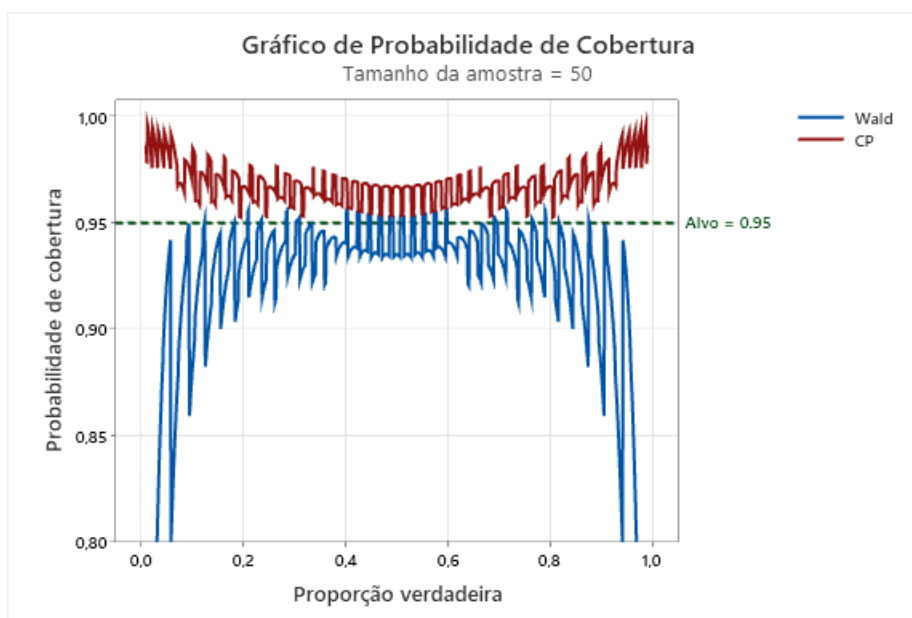


Um problema comum na estatística básica é a estimativa da proporção de indivíduos com uma determinada característica de interesse em uma população. Por exemplo, um engenheiro de qualidade pode querer estimar a proporção de defeitos em um grande lote de unidades produzidas em massa em um determinado dia; um cientista médico pode querer investigar a proporção de indivíduos em alguma comunidade que foram vacinados contra um patógeno específico, mas experimentaram a doença relacionada mesmo assim; Um gestor de campanha pode se interessar pela proporção de eleitores cadastrados que pretendem votar em seu candidato.

Os métodos de estimação intervalar mais conhecidos para esse problema são o método de aproximação normal do livro didático referido como intervalo de confiança (IC) de Wald e o IC exato de Clopper-Pearson (1934). Por um lado, o IC de Wald é extremamente liberal na medida em que o nível de confiança real (ou probabilidade de cobertura) do IC está bem abaixo do nível nominal visado, particularmente quando a proporção real é próxima de 0 ou 1 (ver Figura 1). Por outro lado, o IC exato de Clopper-Pearson é

excessivamente conservador na medida em que o nível de confiança real (ou probabilidade de cobertura) do IC está bem acima do nível nominal visado. Ambos os métodos não devem mais ser usados para quaisquer aplicações práticas (ver Agresti-Coull, 1998; Brown et al., 2001).

Nos últimos anos, no entanto, eles têm desempenhado um papel importante no desenvolvimento de melhores métodos de IC com melhores probabilidades de cobertura intermediária. Por exemplo, o IC aproximado de Agresti-Coull é um ajuste no IC de Wald; o IC exato de Blaker (2000, 2001) usa limites de confiança de Clopper-Pearson como estimativas iniciais em um algoritmo numérico iterativo. Atento a esses métodos recentemente aprimorados, o Minitab atualizou a ferramenta estatística para estimar uma única proporção populacional para incluir os 4 métodos a seguir: o IC Blaker ajustado e os métodos de teste, o IC de Wilson/escore e os métodos de teste (com e sem correção de continuidade) e o IC de Agresti-Coull e os métodos de teste. Além disso, para cada um desses métodos, o Minitab garante que o IC e o teste produzam resultados consistentes.



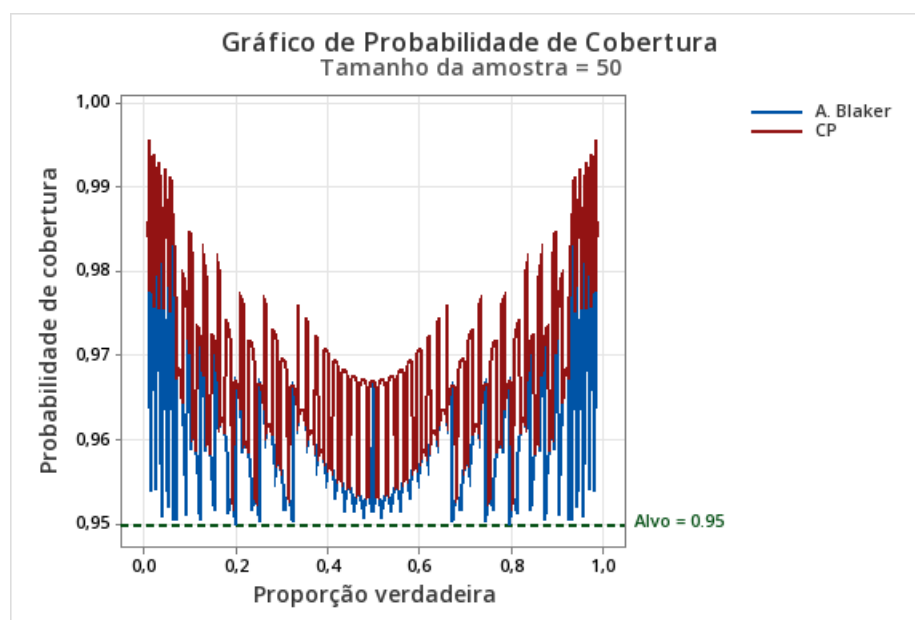
**Figura 1:** Comparação das probabilidades de cobertura para os ICs de Wald e Clopper-Pearson (CP) em função da proporção verdadeira quando o tamanho da amostra é 50. O gráfico ilustra que os IC de Wald e os IC de Pearson de Clopper são excessivamente liberais e conservadores, respectivamente, particularmente quando a proporção real está próxima de 0 ou 1. Assumindo que as proporções verdadeiras estão uniformemente distribuídas no intervalo (0, 1), as probabilidades médias de cobertura com base em uma amostra de tamanho 50 são de 0,901 e 0,969 para o IC de Wald e IC de Clopper-Pearson, respectivamente.

## Conheça os 4 Novos Métodos de Estimação

Os 4 novos métodos consistem em 1 IC exato e método de teste chamado método Blaker ajustado e 3 IC aproximados e métodos de teste incluindo o método de Wilson/escore (Wilson), Wilson/escore com o método de correção de continuidade de Yates (Wilson CC) e o método de Agresti-Coull (AC). Um método exato neste contexto significa que não há aproximação usada na derivação do método, em oposição aos métodos aproximados que são obtidos usando algumas formas de procedimentos de aproximação normais.

### 1. O Método Blaker Ajustado

O método Blaker ajustado, devido a Klaschka e Reiczigel (2021), é uma modificação do IC exato de Blaker (2000, 2001) e métodos de teste. A modificação aborda a natureza computacional intensiva do algoritmo Blaker original e as inconsistências ocasionais entre seu IC e os resultados do teste. Como o IC Blaker original, o IC ajustado resultante é exato, aninhado e está contido no IC de Clopper-Pearson. Como resultado, o IC Blaker ajustado é menos conservador que o IC Clopper-Pearson. O IC é aninhado no sentido de que um IC com um nível de confiança mais alto sempre contém um IC com um nível de confiança mais baixo. Por exemplo, um IC Blaker de 95% (ajustado) de dois lados sempre contém o CI de 90% bilateral correspondente. A aninhamento é uma propriedade atraente de métodos exatos de IC derivados de uma distribuição discreta como o binômio. Por exemplo, o CI Clopper-Pearson é aninhado. Há, no entanto, métodos exatos de IC disponíveis que não são necessariamente aninhados. Por exemplo, o chamado Blyth-Still-Casella CI (Blyth e Still, 1983; Casella, 1986) é garantido como o IC exato mais curto, mas não está aninhado. Crow (1956) CI também não está aninhado. Os cálculos de ICs baseados em Blaker ou no método Blaker ajustado são mais complexos do que os métodos clássicos de IC acima mencionados, pois requerem algoritmos numéricos. Com as inovações atuais na tecnologia computacional, no entanto, não devemos mais nos furtar de implementar algoritmos complexos que produzam melhores resultados. A Figura 2 ilustra as melhorias do IC Blaker ajustado em relação ao IC Clopper-Pearson.

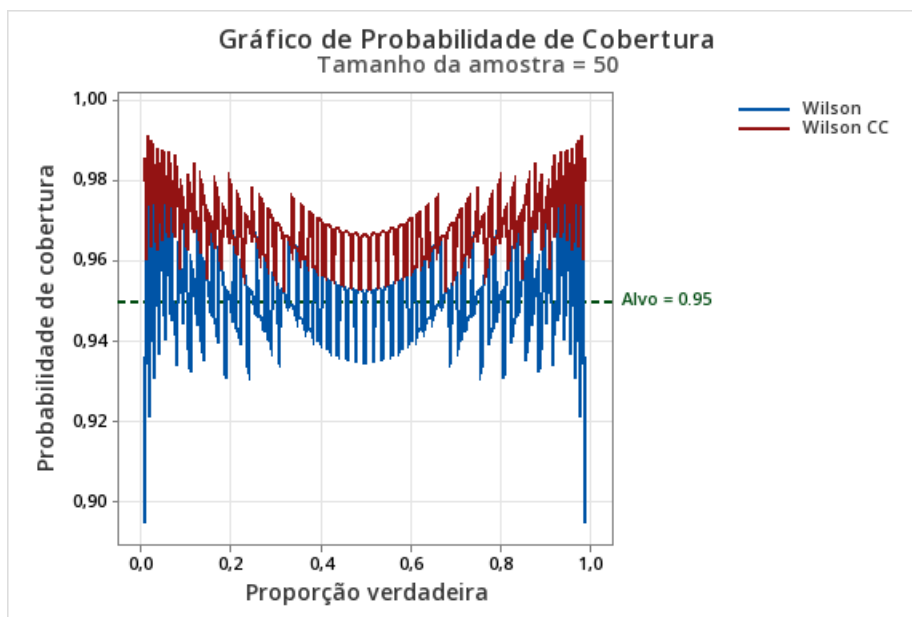


**Figura 2:** Comparação das probabilidades de cobertura para os ICs Blaker (A. Blaker) ajustados e os IC Clopper-Pearson (CP) em função da proporção verdadeira quando o tamanho da amostra é 50. O gráfico indica que a probabilidade de cobertura do IC Clopper-Pearson é pelo menos a do IC Blaker ajustado. Isso é consistente com o fato de que os ICs Blaker ajustados estão contidos nos IC Clopper-Pearson. Para qualquer amostra de tamanho 50, as probabilidades médias de cobertura (assumindo que a verdadeira proporção está uniformemente distribuída no intervalo unitário) são de 0,960 e 0,969 para o IC Blaker ajustado e IC Clopper-Pearson, respectivamente.

### 2. Os métodos Wilson e Wilson CC

O método do IC de Wilson (1927) é derivado como o IC que corresponde ao teste de pontuação, o teste que utiliza o erro padrão nulo,  $\sqrt{p_0(1-p_0)/n}$ , em oposição ao erro padrão clássico  $\sqrt{p(1-p)/n}$ , no denominador da estatística do teste. Por essa

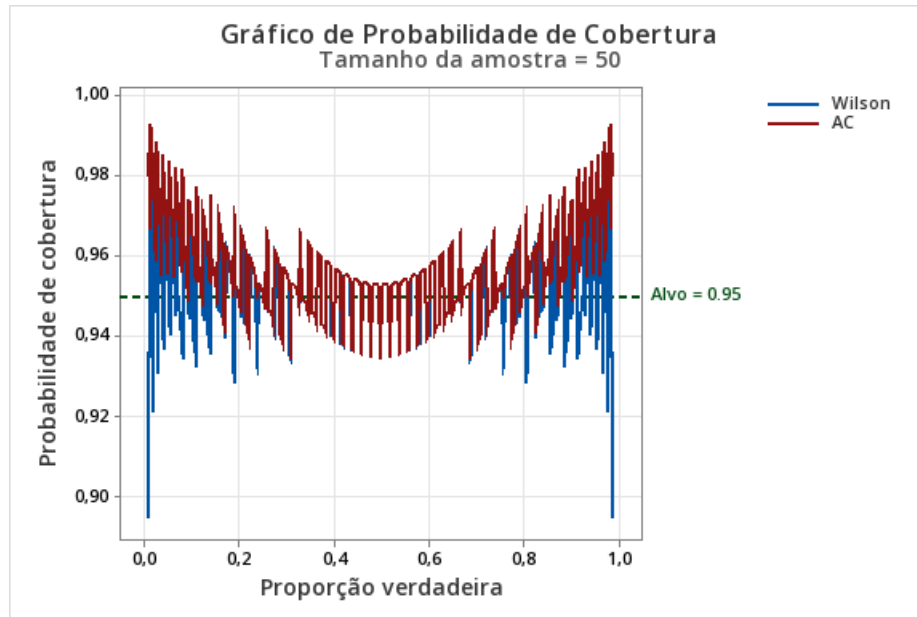
razão, também é chamada de IC Wilson/escore. Sua probabilidade real de cobertura pode ser menor ou maior do que o nível nominal visado, mas permanece próxima a ele, exceto quando a proporção verdadeira é próxima de 0 ou 1 (ver Figura 3). Um ajuste pode ser feito para tornar o IC Wilson/escore conservador usando a correção de continuidade de Yates. O Minitab fornece as versões do IC (com e sem a correção de continuidade de Yates) e seus testes de hipótese correspondentes.



**Figura 3:** Comparação das probabilidades de cobertura para os ICs Wilson/escore (Wilson) e Wilson/escore com correção de continuidade de Yates (Wilson CC) em função da proporção verdadeira quando o tamanho da amostra é 50. O gráfico mostra que os IC de Wilson são sempre conservadores, enquanto os IC de Wilson são conservadores e liberais, dependendo da magnitude da verdadeira proporção. Em particular, os ICs de Wilson tendem a ser muito liberais quando a proporção real é muito próxima de 0 ou 1. Para qualquer amostra de tamanho 50, as probabilidades médias de cobertura são de 0,952 e 0,969 para o IC de Wilson e o IC de Wilson CC, respectivamente.

### 3. O Método Agresti-Coull

A IC de Agresti-Coull é obtida a partir de um ajuste da IC clássica de Wald excessivamente liberal. O IC resultante tem propriedades de cobertura semelhantes às IC de Wilson, mas um pouco mais conservador em geral. Além disso, os dois tipos de IC têm o mesmo ponto médio, mas os IC de Wilson estão sempre contidos nos IC de Agresti-Coull. Como ilustrado na Figura 4, eles têm essencialmente as mesmas probabilidades de cobertura quando a proporção verdadeira é moderada. O IC Agresti Coull, no entanto, é geralmente menos liberal quando a proporção real é próxima de 0 ou 1. Como mostrado na Figura 4, para uma amostra de tamanho 50 o IC de Agresti-Coull torna-se conservador à medida que a proporção real se aproxima de 0 ou 1. Outro apelo da IC Agresti-Coull é a simplicidade de sua implementação herdada da IC Wald. Além disso, é fácil de ensinar e lembrar, especialmente quando o nível de confiança é de 95%. Para esse nível de confiança, é comumente referido como o método de IC "adicionar 2 sucessos e 2 fracassos" como um memorando ao ajuste feito no IC de Wald para derivá-lo.



**Figura 4:** Comparação das probabilidades de cobertura para os IC Wilson/escore (Wilson) e os IC Agresti-Coull (AC) em função da verdadeira proporção quando o tamanho da amostra é 50. O gráfico indica que os dois métodos produzem essencialmente as mesmas probabilidades de cobertura para valores moderados da verdadeira proporção no intervalo (0,3; 0,75). Para valores da proporção verdadeira próximos de 0 ou 1, no entanto, os IC de Wilson são liberais, enquanto os IC de Agresti-Coull são conservadores. Isso é consistente com o fato de que os IC de Wilson estão contidos nos IC de Agresti-Coull. Para qualquer amostra de tamanho 50, as probabilidades médias de cobertura são de 0,952 e 0,958 para o IC de Wilson e IC de Agresti-Coull, respectivamente.

## Alguns exemplos ilustrativos simples

Um engenheiro de qualidade em uma fábrica de produção em massa selecionou uma amostra aleatória de 1465 produtos produzidos em massa em um determinado dia. Após testes independentes dos 1465 itens, 53 foram encontrados com defeito. O engenheiro quer saber se a proporção de itens defeituosos produzidos no dia determinado é significativamente diferente de 2,5%.

Embora esse problema seja lançado como uma questão de teste de hipótese estatística, os estatísticos aplicados são cada vez mais encorajados a incluir no resultado da análise uma estimativa pontual e um intervalo de confiança juntamente com o valor de  $p$  do teste. O Minitab segue essa convenção tanto quanto possível, especialmente nos Módulos Básicos de Estatística. Por exemplo, usando o Minitab, os resultados da análise para a pergunta acima com base no método Blaker ajustado são os seguintes.

Teste e IC para Uma Proporção

### Método

p proporção de eventos  
Método ajustado método exato de Blaker

### Estatísticas Descritivas

**N Amostra de Eventos p IC 95% para p**

---

1465, 53 0,036177( 0,027353, 0,046822)

### Teste

Hipótese nula:  $C \leq 0,025$

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

**Valor-P**

0,009

Uma saída semelhante pode ser gerada para cada um dos 4 métodos. As características e vantagens de cada método são resumidas na tabela a seguir.

<b>Método</b>	<b>95%</b>	<b>Valor de p do teste de pareamento</b>
Blaker ajustado	(2,74%,4,68%)	0,009
Wilson com correção de Yates	(2,75%,4,74%)	0,008
Wilson/escore	(2,78%,4,70%)	0,006
Agresti-Coull	(2,77%,4,74%)	0,007

Neste exemplo, todos os métodos produzem a mesma conclusão de que o % defeituoso difere de 2,5% no nível de significância de 0,05, uma vez que todos os valores de p são menores que 0,05. Os intervalos de confiança e os valores de p correspondentes para todos os métodos são semelhantes, em parte, porque o tamanho da amostra é muito grande. Além disso, o IC para cada método não cobre o valor da proporção hipotética (2,5%), o que é consistente com o valor de p de cada um dos testes de hipóteses pareadas.

No exemplo acima, suponha agora que o engenheiro de qualidade testou apenas 50 itens e descobriu que 2 estavam com defeito. Além disso, suponha que o engenheiro quisesse saber se a proporção de defeituosos era significativamente diferente de 1,0%. As características e vantagens de cada método são resumidas na tabela a seguir.

<b>Método</b>	<b>95%</b>	<b>Valor de p do teste correspondente</b>
Blaker ajustado	(0,72%,13,35%)	0,089
Wilson com correção de Yates	(0,70%,14,86%)	0,155
Wilson/escore	(1,10%,13,46%)	0,033
Agresti-Coull	(0,34%,14,22%)	0,124

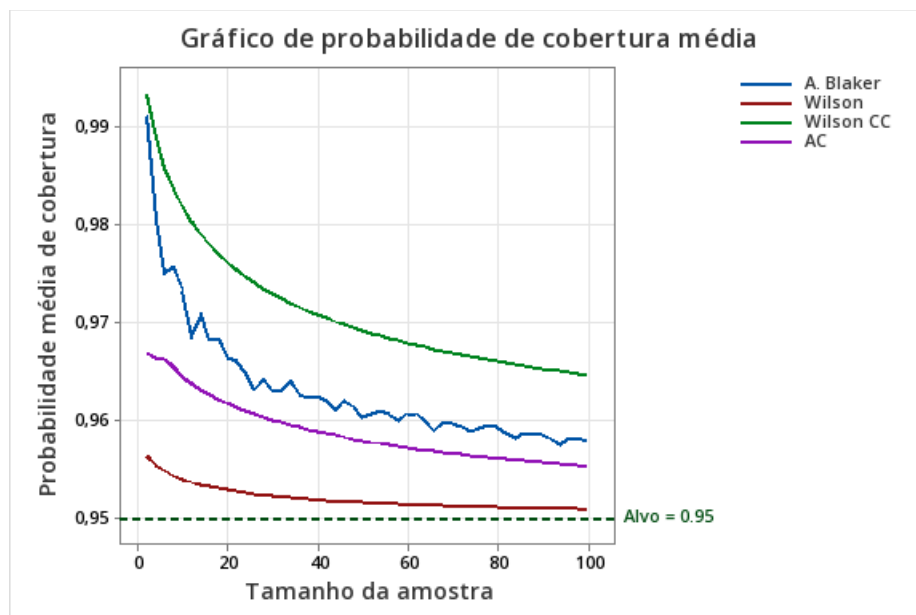
Nesse caso, apenas o método de Wilson/escore permite concluir significativamente que o percentual de defeituoso difere de 1,0% ao nível de significância de 0,05. No mesmo nível de significância, todos os outros métodos produzem a conclusão oposta de evidência insuficiente para decidir que há uma diferença. As inconsistências nos resultados entre os métodos devem-se, em grande parte, ao tamanho da amostra ser moderado. Em média, as probabilidades de cobertura desses métodos se aproximam

do nível nominal à medida que o tamanho da amostra aumenta (ver Figura 5 abaixo). Para desenhos amostrais pequenos a moderados, no entanto, as disparidades nas probabilidades médias de cobertura associadas a cada método são mais pronunciadas, fazendo com que os ICs correspondentes tenham larguras visivelmente diferentes. Uma questão importante, no entanto, é qual resultado relatar ao chefe? Não há uma resposta direta para essa pergunta, pois uma melhor pode depender de fatores como o conhecimento prévio da magnitude da verdadeira proporção ou mesmo da área de aplicação. Forneceremos algumas diretrizes gerais aproximadas na próxima seção.

## Conclusão

As Figuras 2, 3 e 4 mostram que os métodos 4 IC, Blaker Ajustado, Wilson, Wilson CC e Agresti-Coull apresentam diferentes propriedades de probabilidade de cobertura. O Wilson CC é o mais conservador, seguido pelo Blaker ajustado. O Agresti Coull e o Wilson costumam ser liberais e conservadores, dependendo da magnitude da verdadeira proporção. No geral, o método de Wilson CI é o mais liberal de todos os 4 métodos. Além disso, a Figura 5 indica que, em média, todos os 4 métodos são conservadores, sendo o CC de Wilson o mais conservador, seguido pelo Blaker ajustado, Agresti-Coull e Wilson. Para um determinado problema, o método apropriado a ser escolhido pode depender da aplicação em particular, do tamanho da amostra e da disponibilidade ou não de algum conhecimento prévio da magnitude da verdadeira proporção.

Por exemplo, as agências reguladoras estão frequentemente interessadas em métodos conservadores para proteger os consumidores. Um método demasiado conservador, no entanto, pode produzir regulamentos rigorosos, enquanto um método demasiado liberal pode produzir regulamentos frouxos. Em geral, para desenhos de amostras moderadas a grandes, o método Blaker ajustado ou o método CC de Willson podem ser boas escolhas. Para desenhos de amostras pequenas, o Wilson ou o Agresti-Coull podem ser apropriados se houver um conhecimento prévio do verdadeiro valor da proporção. Esse conhecimento prévio é frequentemente baseado em experiências anteriores semelhantes ou em pequenos estudos-piloto especificamente desenhados para obter uma estimativa aproximada da proporção. Por exemplo, na área de controle de qualidade de aplicações onde a proporção de defeitos é tipicamente próxima de 0, o método a ser escolhido depende se o investigador quer ou não um resultado conservador. Um resultado conservador pode ser baseado no método de Agresti-Coull enquanto um resultado liberal pode ser baseado no método de Wilson/score. Por último, mas mais importante, é sempre uma boa prática planejar com antecedência o tamanho da amostra, de modo a evitar o erro tipo II ou a controlar a largura dos IC. O Minitab tem "um aplicativo para isso também". Ferramentas estão disponíveis no Minitab para determinar o tamanho adequado da amostra para controlar o erro tipo II associado a testes de hipótese ou para controlar a largura dos ICs.



**Figura 5:** Probabilidade média de cobertura de todos os 4 métodos de IC em função do tamanho da amostra. As probabilidades médias de cobertura são calculadas assumindo que a proporção verdadeira é uniformemente distribuída no intervalo unitário. As curvas de cobertura média ilustram que, em média, o método de Wilson/escore com correção de continuidade de Yates (Wilson CC) é o mais conservador, seguido pelos métodos de Blaker (A. Blaker) ajustado, Agresti-Coull (AC) e Wilson/score (Wilson). As curvas de probabilidade de cobertura média aproximam-se do nível de cobertura nominal desejado à medida que o tamanho da amostra aumenta. Além disso, enquanto as curvas de cobertura média para os métodos aproximados (Wilson CC, Wilson, Agresti-Coull) são suaves, a curva de cobertura para o Blaker ajustado exato tem alguns movimentos oscilantes à medida que se aproxima da cobertura nominal. Isso indica, talvez, que o método Blaker CI ajustado possa ser melhorado.

## Referência

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](http://www.minitab.com). All other marks referenced remain the property of their respective owners.



