



**4 Neue Methoden zur Verbesserung Ihrer
Schätzung einer einzelnen Grundgesamtheit
Anteil in Minitab**



Ein häufiges Problem in der Basisstatistik ist die Schätzung des Anteils von Personen mit einem bestimmten Merkmal von Interesse an einer Grundgesamtheit. Beispielsweise kann ein Qualitätsingenieur den Anteil der Fehler in einer großen Charge von Massenproduktionseinheiten an einem bestimmten Tag schätzen. Ein medizinischer Wissenschaftler möchte möglicherweise den Anteil der Personen in einer Gemeinschaft untersuchen, die gegen einen bestimmten Krankheitserreger geimpft wurden, aber dennoch die verwandte Krankheit erlitten haben. Ein Wahlkampfmanager kann sich für den Anteil der registrierten Wähler interessieren, die beabsichtigen, für ihren Kandidaten zu stimmen.

Die bekanntesten Intervallschätzungsmethoden für dieses Problem sind die Lehrbuchmethode der Normalapproximation, die als Wald-Konfidenzintervall (KI) bezeichnet wird, und das Clopper-Pearson-exakte (1934) KI. Einerseits ist das Wald-KI insofern äußerst liberal, als das tatsächliche Konfidenzniveau (oder die Abdeckungswahrscheinlichkeit) des KI deutlich unter dem angestrebten nominalen

Niveau liegt, insbesondere wenn der tatsächliche Anteil nahe 0 oder 1 liegt (siehe Abbildung 1). Auf der anderen Seite ist das genaue Clopper-Pearson-KI insofern übermäßig konservativ, als das tatsächliche Konfidenzniveau (oder die Abdeckungswahrscheinlichkeit) des KI deutlich über dem angestrebten nominalen Niveau liegt. Beide Methoden sollten nicht mehr für praktische Anwendungen verwendet werden (vgl. Agresti-Coull, 1998; Brown et al., 2001).

In den letzten Jahren haben sie jedoch eine große Rolle bei der Entwicklung besserer CI-Methoden mit besseren Zwischenabdeckungswahrscheinlichkeiten gespielt. Zum Beispiel ist das Agresti-Coull-Näherungs-CI eine Anpassung des Wald-CI; Das exakte CI von Blaker (2000, 2001) verwendet Clopper-Pearson-Konfidenzschranken als Ausgangsschätzungen in einem iterativen numerischen Algorithmus. Unter Berücksichtigung dieser neu verbesserten Methoden hat Minitab das statistische Werkzeug zur Schätzung eines einzelnen Anteils der Grundgesamtheit aktualisiert und umfasst nun die folgenden 4 Methoden: das angepasste Blaker-CI und die Testmethoden, das Wilson/Score-CI und die Testmethoden (mit und ohne Kontinuitätskorrektur) sowie das Agresti-Coull-CI und die Testmethoden. Darüber hinaus stellt Minitab für jede dieser Methoden sicher, dass CI und Test konsistente Ergebnisse liefern.

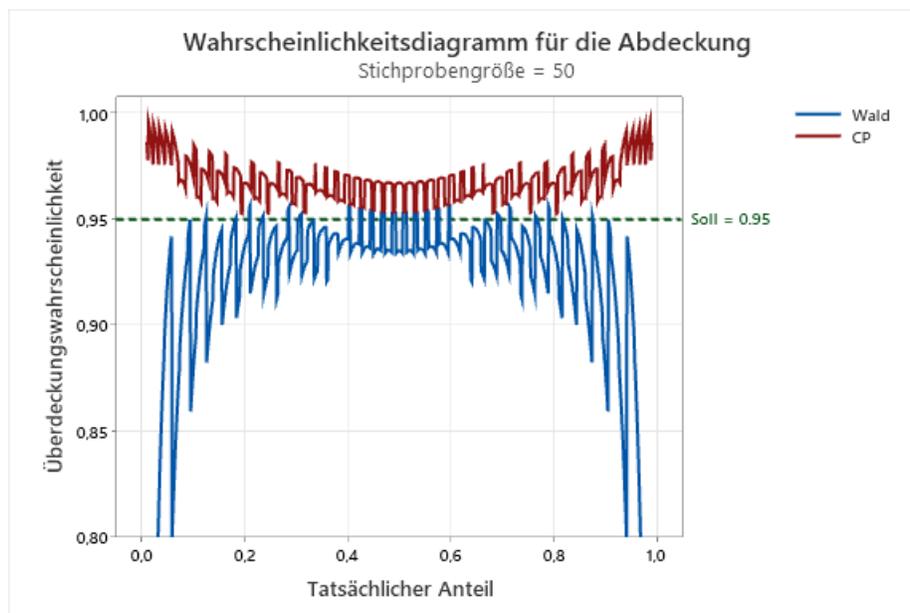


Abbildung 1: Vergleich der Abdeckungswahrscheinlichkeiten für die Wald-CIs und Clopper-Pearson-CIs (CP) als Funktionen des tatsächlichen Anteils bei einer Stichprobengröße von 50. Die Grafik zeigt, dass die Wald-CIs und die Clopper-Pearson-CIs übermäßig liberal bzw. konservativ sind, insbesondere wenn der tatsächliche Anteil nahe 0 oder 1 liegt. Unter der Annahme, dass die wahren Anteile gleichmäßig im Intervall (0, 1) verteilt sind, betragen die mittleren Abdeckungswahrscheinlichkeiten basierend auf einer Stichprobe der Größe 50 0,901 bzw. 0,969 für das Wald-KI bzw. das Clopper-Pearson-KI.

Lernen Sie die 4 neuen Schätzmethoden kennen

Die 4 neuen Methoden bestehen aus 1 exakten KI und einer Testmethode, die als angepasste Blaker-Methode bezeichnet wird, und 3 ungefähren KI- und Testmethoden, darunter die Wilson/Score-Methode (Wilson), die Wilson/Score-Methode mit Yates-Kontinuitätskorrektur (Wilson CC) und die Agresti-Coull-Methode (AC). Eine exakte Methode bedeutet in diesem Zusammenhang, dass bei der Herleitung der Methode keine Näherung verwendet wird, im Gegensatz zu Näherungsmethoden, die mit einigen Formen normaler Näherungsverfahren erhalten werden.

1. Die angepasste Blaker-Methode

Die angepasste Blaker-Methode, in Anlehnung an Klaschka und Reiczigel (2021), ist eine Modifikation der exakten CI- und Testmethoden von Blaker (2000, 2001). Die Modifikation behebt die rechenintensive Natur des ursprünglichen Blaker-Algorithmus und die gelegentlichen Inkonsistenzen zwischen seinem CI und den Testergebnissen. Wie das ursprüngliche Blaker-CI ist das resultierende angepasste CI exakt, verschachtelt und im Clopper-Pearson-CI enthalten. Infolgedessen ist der angepasste Blaker-CI weniger konservativ als der Clopper-Pearson-CI. Das CI ist in dem Sinne geschachtelt, dass ein CI mit einem höheren Konfidenzniveau immer ein CI mit einem niedrigeren Konfidenzniveau enthält. Zum Beispiel enthält ein zweiseitiges 95 % (bereinigtes) Blaker-CI immer das entsprechende zweiseitige 90 %-KI. Die Verschachtelung ist eine ansprechende Eigenschaft exakter CI-Methoden, die von einer diskreten Verteilung wie dem Binom abgeleitet werden. Beispielsweise ist das Clopper-Pearson-CI verschachtelt. Es gibt jedoch exakte CI-Methoden, die nicht unbedingt verschachtelt sind. Zum Beispiel die sogenannte Blyth-Still-Casella-CI (Blyth und Still, 1983; Casella, 1986) ist garantiert das kürzeste exakte CI, aber nicht verschachtelt. Crow (1956) CI ist ebenfalls nicht verschachtelt. Die Berechnungen von CIs auf Basis von Blaker oder der angepassten Blaker-Methode sind komplexer als die oben genannten klassischen CI-Methoden, da sie numerische Algorithmen erfordern. Mit den aktuellen Innovationen in der Computertechnik sollten wir uns jedoch nicht mehr scheuen, komplexe Algorithmen zu implementieren, die zu besseren Ergebnissen führen. Abbildung 2 zeigt die Verbesserungen des angepassten Blaker-CI gegenüber dem Clopper-Pearson-CI.

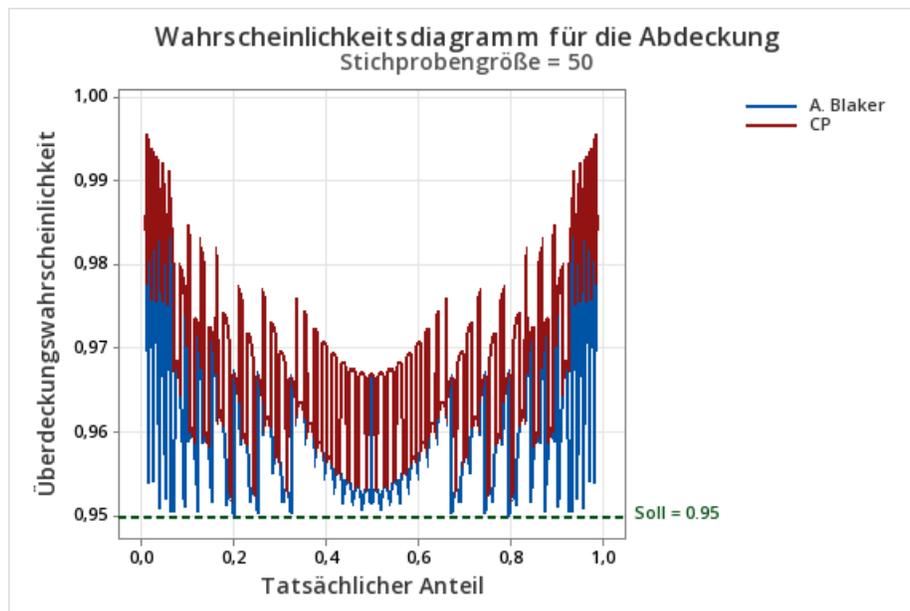


Abbildung 2: Vergleich der Abdeckungswahrscheinlichkeiten für die adjustierten Blaker (A. Blaker) CIs und die Clopper-Pearson (CP) CIs als Funktionen des tatsächlichen Anteils bei einer Stichprobengröße von 50. Die Grafik zeigt, dass die Abdeckungswahrscheinlichkeit des Clopper-Pearson-CI mindestens der des angepassten Blaker-CI entspricht. Dies steht im Einklang mit der Tatsache, dass angepasste Blaker-CIs in Clopper-Pearson-CIs enthalten sind. Für jede Stichprobe der Größe 50 betragen die durchschnittlichen Abdeckungswahrscheinlichkeiten (unter der Annahme, dass der wahre Anteil gleichmäßig auf dem Einheitsintervall verteilt ist) 0,960 bzw. 0,969 für das bereinigte Blaker-KI bzw. das Clopper-Pearson-KI.

2. Die Wilson- und Wilson-CC-Methoden

Die CI-Methode von Wilson (1927) wird als das KI abgeleitet, das dem Score-Test entspricht, dem Test, der den Null-Standardfehler $\sqrt{p_0(1-p_0)/n}$ verwendet, im Gegensatz zum klassischen Standardfehler $\sqrt{\hat{p}(1-\hat{p})/n}$ auf dem Nenner der Teststatistik. Aus diesem Grund wird es auch als Wilson/Score-CI bezeichnet. Die tatsächliche Abdeckungswahrscheinlichkeit kann kleiner oder größer sein als der angestrebte nominale Betrag, bleibt aber nahe daran, es sei denn, der tatsächliche Anteil liegt nahe 0 oder 1 (siehe Abbildung 3). Eine Anpassung kann vorgenommen werden, um das Wilson/Score-CI konservativ zu machen, indem die Kontinuitätskorrektur von Yates verwendet wird. Minitab stellt beide CI-Versionen (mit und ohne Yates-Kontinuitätskorrektur) und ihre übereinstimmenden Hypothesentests zur Verfügung.

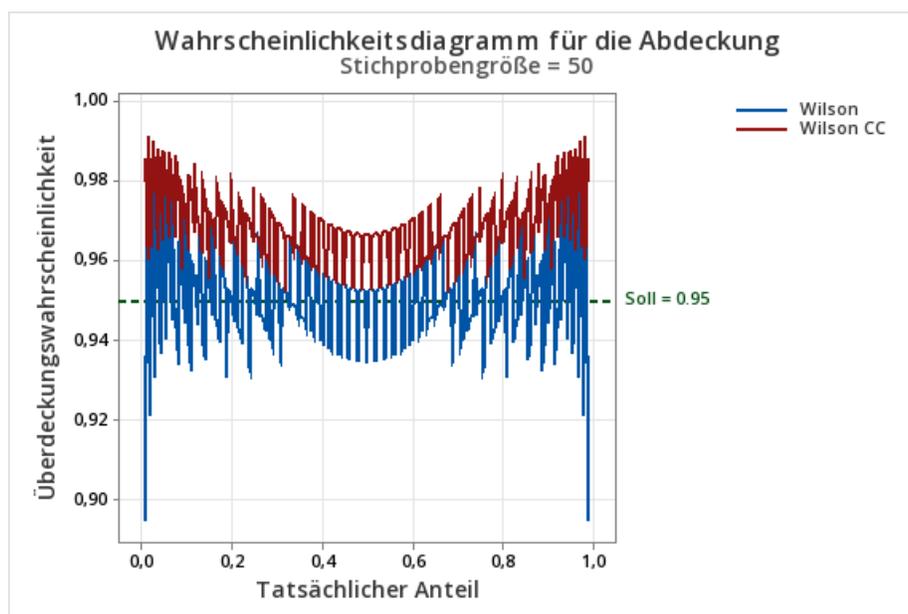


Abbildung 3: Vergleich der Abdeckungswahrscheinlichkeiten für die Wilson/Score (Wilson) CIs und die Wilson/Score mit der Yates-Kontinuitätskorrektur (Wilson CC) als Funktionen des wahren Anteils bei einer Stichprobengröße von 50. Die Grafik zeigt, dass Wilson-CC-CIs immer konservativ sind, während Wilson-CIs je nach Größe des tatsächlichen Anteils konservativ und liberal sind. Insbesondere Wilson-CIs neigen dazu, zu liberal zu sein, wenn das wahre Verhältnis sehr nahe bei 0 oder 1 liegt. Für jede Stichprobe der Größe 50 betragen die mittleren Abdeckungswahrscheinlichkeiten 0,952 bzw. 0,969 für das Wilson-CI bzw. das Wilson-CC-CI.

3. Die Agresti-Coull-Methode

Die Agresti-Coull-CI ergibt sich aus einer Anpassung der übermäßig liberalen klassischen Wald-CI. Das resultierende CI hat ähnliche Abdeckungseigenschaften wie die Wilson-CIs, ist aber im Allgemeinen etwas konservativer. Darüber hinaus haben die beiden Arten von CIs den gleichen Mittelpunkt, aber Wilson-CIs sind immer in Agresti-Coull-CIs enthalten. Wie in Abbildung 4 dargestellt, haben sie im Wesentlichen die gleichen Abdeckungswahrscheinlichkeiten, wenn der tatsächliche Anteil moderat ist. Das Agresti Coull CI ist jedoch im Allgemeinen weniger liberal, wenn das wahre Verhältnis nahe bei 0 oder 1 liegt. Wie in Abbildung 4 dargestellt, wird das Agresti-Coull-KI bei einer Stichprobe der Größe 50 konservativ, wenn sich der tatsächliche Anteil 0 oder 1 nähert. Ein weiterer Vorteil der Agresti-Coull-CI ist die einfache Implementierung, die von der Wald-CI übernommen wurde. Darüber hinaus ist es leicht zu lehren und zu merken, insbesondere wenn das Konfidenzniveau 95 % beträgt. Für dieses Konfidenzniveau wird es allgemein als "addieren Sie 2 Erfolge und 2 Misserfolge" CI-Methode bezeichnet, als Memorandum zu der Anpassung, die an der Wald-CI vorgenommen wurde, um es abzuleiten.

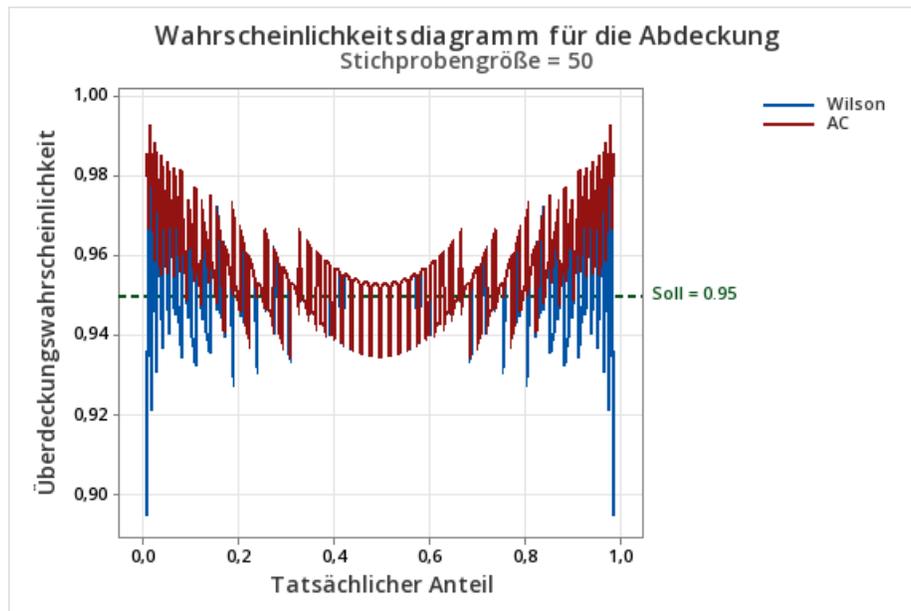


Abbildung 4: Vergleich der Abdeckungswahrscheinlichkeiten für die Wilson/Score (Wilson) CIs und die Agresti-Coull (AC) CIs als Funktionen des tatsächlichen Anteils bei einer Stichprobengröße von 50. Die Grafik zeigt, dass die beiden Methoden im Wesentlichen die gleichen Abdeckungswahrscheinlichkeiten für moderate Werte des wahren Anteils im Intervall (0,3, 0,75) liefern. Für Werte mit dem wahren Anteil nahe 0 oder 1 sind Wilson-CIs jedoch liberal, während Agresti-Coull-CIs konservativ sind. Dies steht im Einklang mit der Tatsache, dass Wilson-CIs in Agresti-Coull-CIs enthalten sind. Für jede Stichprobe der Größe 50 beträgt die mittlere Abdeckungswahrscheinlichkeit 0,952 bzw. 0,958 für das Wilson-KI bzw. das Agresti-Coull-KI.

Einige einfache anschauliche Beispiele

Ein Qualitätsingenieur in einer Massenproduktionsfabrik hat eine Zufallsstichprobe von 1465 Massenprodukten an einem bestimmten Tag ausgewählt. Nach unabhängiger Prüfung der 1465 Artikel wurden 53 als fehlerhaft befunden. Der Ingenieur möchte wissen, ob der Anteil der fehlerhaften Artikel, die an einem bestimmten Tag produziert werden, signifikant von 2,5 % abweicht.

Obwohl es sich bei dieser Aufgabe um eine Testfrage für statistische Hypothesen handelt, werden angewandte Statistiker zunehmend dazu angehalten, neben dem p-Wert des Tests auch eine Punktschätzung und ein Konfidenzintervall in das Analyseergebnis einzubeziehen. Minitab hält sich so weit wie möglich an diese Konvention, insbesondere in den grundlegenden Statistikmodulen. Wenn Sie z. B. Minitab verwenden, sehen die Analyseergebnisse für die obige Frage basierend auf der angepassten Blaker-Methode wie folgt aus.

Test von Anteilen und KI bei einer Stichprobe

Methode

p Ereignisanteil

Methode Angepasste Blaker's exakte Methode

Deskriptive Statistik

N-Ereignis-Stichprobe p 95%-KI für p

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

Test

Nullhypothese: p

= 0,025

Alternativhypothese H_1 : $p \neq 0,025$

p-Wert

0,009

Für jede der 4 Methoden kann eine ähnliche Ausgabe erzeugt werden. Die Merkmale und Vorteile der einzelnen Methoden sind in der folgenden Tabelle zusammengefasst.

Methode	95% CI	p-Wert des Matching-Tests
Angepasster Blaker	(2,74%,4,68%)	0,009
Wilson mit Yates' Korrektur	(2,75%,4,74%)	0,008
Wilson/Partitur	(2,78%,4,70%)	0,006
Agresti-Coull	(2,77%,4,74%)	0,007

In diesem Beispiel führen alle Methoden zur gleichen Schlussfolgerung, dass der Prozentsatz der fehlerhaften Werte von 2,5 % auf dem Signifikanzniveau 0,05 abweicht, da alle p-Werte kleiner als 0,05 sind. Die Konfidenzintervalle und die entsprechenden p-Werte für alle Methoden sind ähnlich, was zum Teil darauf zurückzuführen ist, dass die Stichprobe sehr groß ist. Darüber hinaus deckt das KI für jede Methode nicht den hypothetischen Anteilswert (2,5 %) ab, der mit dem p-Wert der einzelnen übereinstimmenden Hypothesentests übereinstimmt.

Nehmen wir im obigen Beispiel an, dass der Qualitätsingenieur nur 50 Artikel getestet hat und festgestellt hat, dass 2 fehlerhaft sind. Nehmen wir außerdem an, dass der Ingenieur wissen möchte, ob der Anteil der fehlerhaften Defekte signifikant von 1,0 % abweicht. Die Merkmale und Vorteile der einzelnen Methoden sind in der folgenden Tabelle zusammengefasst.

Methode	95% CI	p-Wert des entsprechenden Tests
Angepasster Blaker	(0,72%,13,35%)	0,089
Wilson mit Yates' Korrektur	(0,70%,14,86%)	0,155
Wilson/Partitur	(1,10%,13,46%)	0,033
Agresti-Coull	(0,34%,14,22%)	0,124

In diesem Fall führt nur die Wilson/Score-Methode zu einer signifikanten Schlussfolgerung, da der Prozentsatz fehlerhaft von 1,0 % auf dem Signifikanzniveau von 0,05 abweicht. Auf der gleichen Signifikanzebene führen alle anderen Methoden zu der gegenteiligen Schlussfolgerung, dass die Evidenz nicht ausreicht, um zu entscheiden, dass es einen Unterschied gibt. Die Inkonsistenzen in den Ergebnissen zwischen den Methoden sind zum großen Teil darauf zurückzuführen, dass die Stichprobe moderat ist. Im Durchschnitt nähern sich die Erfassungswahrscheinlichkeiten dieser Methoden mit zunehmender Stichprobengröße dem nominalen Niveau an (siehe Abbildung 5 unten). Bei kleinen bis moderaten Stichprobendesigns sind die Unterschiede in den mittleren Abdeckungswahrscheinlichkeiten, die mit den einzelnen Methoden verbunden sind, jedoch ausgeprägter, was dazu führt, dass die entsprechenden KIs deutlich unterschiedliche Breiten aufweisen. Eine wichtige Frage ist jedoch, welches Ergebnis dem Chef zu berichten ist. Auf diese Frage gibt es keine einfache Antwort, denn eine bessere kann von Faktoren wie dem Vorwissen über die Größe des tatsächlichen Anteils oder sogar dem Anwendungsbereich abhängen. Im nächsten Abschnitt werden wir einige grobe allgemeine Richtlinien bereitstellen.

Schlussfolgerung

Die Abbildungen 2, 3 und 4 zeigen, dass die 4 CI-Methoden, Adjusted Blaker, Wilson, Wilson CC und Agresti-Coull, unterschiedliche Eigenschaften der Abdeckungswahrscheinlichkeit haben. Der Wilson CC ist der konservativste, gefolgt vom angepassten Blaker. Die Agresti Coull und die Wilson sind oft liberal und konservativ, je nachdem, wie groß die wahren Proportionen sind. Insgesamt ist die Wilson-CI-Methode die liberalste aller 4 Methoden. Darüber hinaus zeigt Abbildung 5, dass im Durchschnitt alle 4 Methoden konservativ sind, wobei die Wilson-CC-Methode die konservativste ist, gefolgt von der adjustierten Blaker-, der Agresti-Coull- und der Wilson-Methode. Für ein gegebenes Problem kann die geeignete Methode von der jeweiligen Anwendung, der Stichprobengröße und der Verfügbarkeit von Vorkenntnissen über die Größe des tatsächlichen Anteils abhängen.

Zum Beispiel sind Regulierungsbehörden oft an konservativen Methoden interessiert, um die Verbraucher zu schützen. Eine zu konservative Methode kann jedoch zu strengen Vorschriften führen, während eine zu liberale Methode zu lockeren Vorschriften führen kann. Im Allgemeinen kann für mittlere bis große Stichprobendesigns die angepasste Blaker- oder die Willson-CC-Methode eine gute Wahl sein. Für kleine Stichprobendesigns kann die Wilson- oder die Agresti-Coull geeignet sein, wenn eine Vorkenntnis über den wahren Proportionswert vorhanden ist. Dieses Vorwissen beruht häufig auf früheren ähnlichen Erfahrungen oder kleinen Pilotstudien, die speziell darauf ausgelegt sind, eine grobe Schätzung des Anteils zu erhalten. Im Bereich der Qualitätskontrolle von Anwendungen, bei denen der Anteil der Fehler in der Regel nahe 0 liegt, hängt die zu wählende Methode davon ab, ob der Prüfer ein konservatives Ergebnis wünscht oder nicht. Ein konservatives Ergebnis kann auf der Agresti-Coull-Methode basieren, während ein liberales Ergebnis auf der Wilson/Score-Methode basieren kann. Zu guter Letzt, aber noch wichtiger, ist es immer eine gute Praxis, die Stichprobengröße im Voraus zu planen, um sich vor Fehlern vom Typ II zu schützen oder die Breite der CIs zu kontrollieren. Minitab hat "auch dafür eine App". In Minitab stehen Werkzeuge zur Verfügung, mit denen der geeignete Stichprobenumfang bestimmt werden kann, um den mit Hypothesentests verbundenen Fehler vom Typ II oder die Breite von KIs zu kontrollieren.

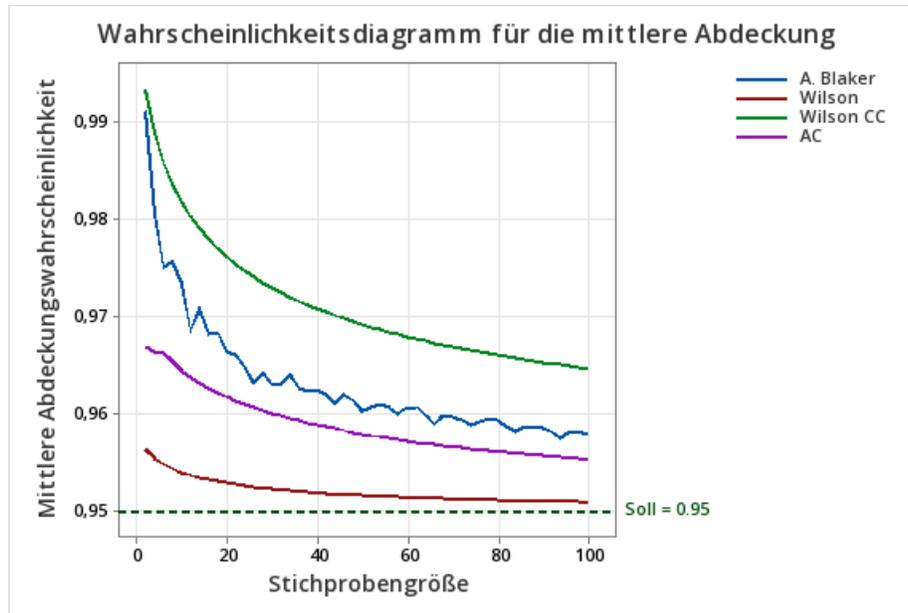


Abbildung 5: Mittlere Abdeckungswahrscheinlichkeit aller 4 CI-Methoden in Abhängigkeit von der Stichprobengröße. Die mittleren Abdeckungswahrscheinlichkeiten werden unter der Annahme berechnet, dass der wahre Anteil gleichmäßig auf dem Einheitenintervall verteilt ist. Die mittleren Abdeckungskurven zeigen, dass im Durchschnitt die Wilson/Score mit Yates' Continuity Correction (Wilson CC) CI-Methode die konservativste ist, gefolgt von der adjustierten Blaker (A. Blaker), der Agresti-Coull (AC) und der Wilson/Score (Wilson) CI-Methode. Die mittleren Abdeckungswahrscheinlichkeitskurven nähern sich mit zunehmender Stichprobengröße dem angestrebten nominalen Abdeckungs niveau an. Während die mittleren Abdeckungskurven für die Näherungsmethoden (Wilson CC, Wilson, Agresti-Coull) glatt sind, weist die Abdeckungskurve für den exakt angepassten Blaker einige wackelige Bewegungen auf, wenn er sich der nominalen Abdeckung nähert. Dies deutet möglicherweise darauf hin, dass die angepasste Blaker-CI-Methode weiter verbessert werden kann.

Referenz

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. und 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. und Pearson, E. S. (1934). The Use of Confidence or Fiducial Limits Illustrated in the Case of Binomial. *Biometrika* 26, S. 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.

Minitab 

Sie haben Daten. Wir haben Solutions Analytics.

Laden Sie eine kostenlose
T minitab.com T