



**4 De nouvelles méthodes pour améliorer  
Estimation d'une seule population  
Proportion dans Minitab**

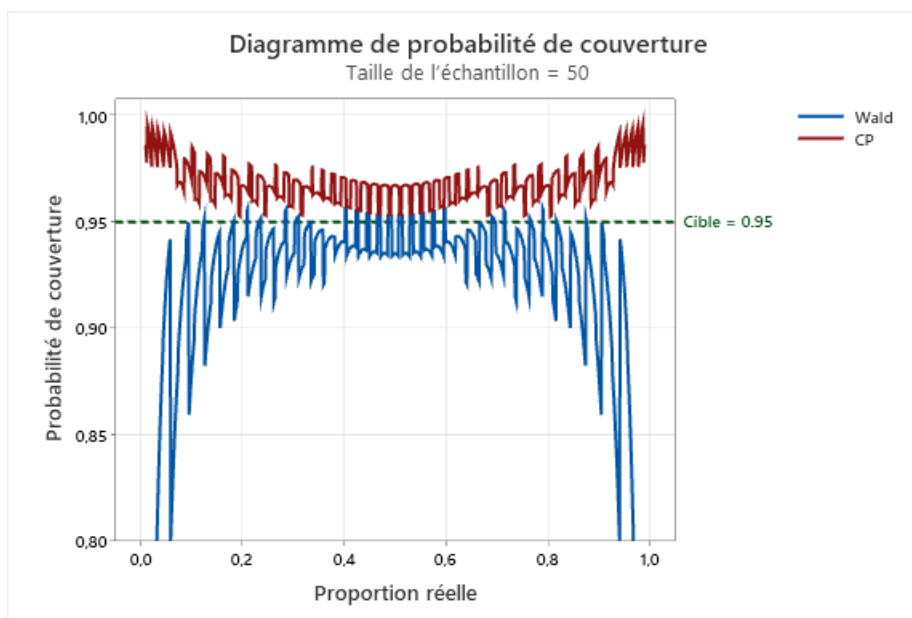


Un problème courant dans les statistiques de base est l'estimation de la proportion d'individus présentant une certaine caractéristique d'intérêt dans une population. Par exemple, un ingénieur qualité peut vouloir estimer la proportion de défauts dans un grand lot d'unités produites en série un jour donné ; Un scientifique médical voudra peut-être étudier la proportion de personnes dans une communauté qui ont été vaccinées contre un agent pathogène spécifique, mais qui ont néanmoins souffert de la maladie connexe ; Un directeur de campagne peut s'intéresser à la proportion d'électeurs inscrits qui ont l'intention de voter pour son candidat.

Les méthodes d'estimation d'intervalle les plus connues pour ce problème sont la méthode d'approximation normale des manuels appelée intervalle de confiance de Wald (IC) et l'IC exact de Clopper-Pearson (1934). D'une part, l'IC de Wald est extrêmement libéral en ce sens que le niveau de confiance réel (ou la probabilité de couverture) de l'IC est bien inférieur au niveau nominal visé, en particulier lorsque la proportion réelle est proche de 0 ou de 1 (voir Figure 1). D'autre part, l'IC Clopper-Pearson exact est

excessivement conservateur en ce sens que le niveau de confiance réel (ou la probabilité de couverture) de l'IC est bien supérieur au niveau nominal cible. Ces deux méthodes ne devraient plus être utilisées pour des applications pratiques (voir Agresti-Coull, 1998 ; Brown et coll., 2001).

Au cours des dernières années, cependant, ils ont joué un rôle majeur dans le développement de meilleures méthodes d'IC avec de meilleures probabilités de couverture intermédiaire. Par exemple, l'IC approximatif d'Agresti-Coull est un ajustement sur l'IC de Wald ; l'IC exact de Blaker (2000, 2001) utilise les bornes de confiance de Clopper-Pearson comme estimations de départ dans un algorithme numérique itératif. Compte tenu de ces méthodes nouvellement améliorées, Minitab a mis à jour l'outil statistique d'estimation d'une proportion de population unique afin d'y inclure les 4 méthodes suivantes : l'IC de Blaker ajusté et les méthodes de test, l'IC de Wilson/score et les méthodes de test (avec et sans correction de continuité), et l'IC d'Agresti-Coull et les méthodes de test. De plus, pour chacune de ces méthodes, Minitab s'assure que l'intégration continue et le test donnent des résultats cohérents.



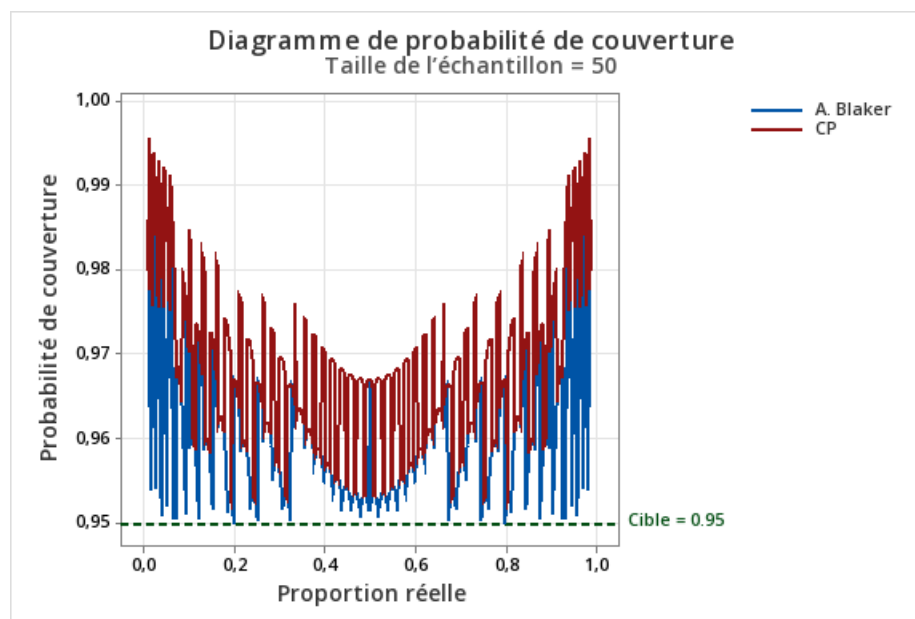
**Graphique 1 :** Comparaison des probabilités de couverture pour les IC de Wald et de Clopper-Pearson (CP) en fonction de la proportion réelle lorsque la taille de l'échantillon est de 50. Le graphique montre que les IC de Wald et de Clopper Pearson sont excessivement libéraux et conservateurs, respectivement, en particulier lorsque la proportion réelle est proche de 0 ou de 1. En supposant que les proportions réelles sont uniformément distribuées dans l'intervalle (0, 1), les probabilités moyennes de couverture basées sur un échantillon de taille 50 sont de 0,901 et 0,969 pour l'IC de Wald et l'IC de Clopper-Pearson, respectivement.

## Découvrez les 4 nouvelles méthodes d'estimation

Les 4 nouvelles méthodes se composent d'une IC exacte et d'une méthode de test appelée méthode de Blaker ajustée et de 3 méthodes d'IC et de test approximatives, y compris la méthode Wilson/score (Wilson), la méthode Wilson/score avec la méthode de correction de continuité de Yates (Wilson CC) et la méthode Agresti-Coull (AC). Une méthode exacte dans ce contexte signifie qu'il n'y a pas d'approximation utilisée dans la dérivation de la méthode, contrairement aux méthodes approximatives qui sont obtenues à l'aide de certaines formes de procédures d'approximation normales.

### 1. La méthode Blaker ajustée

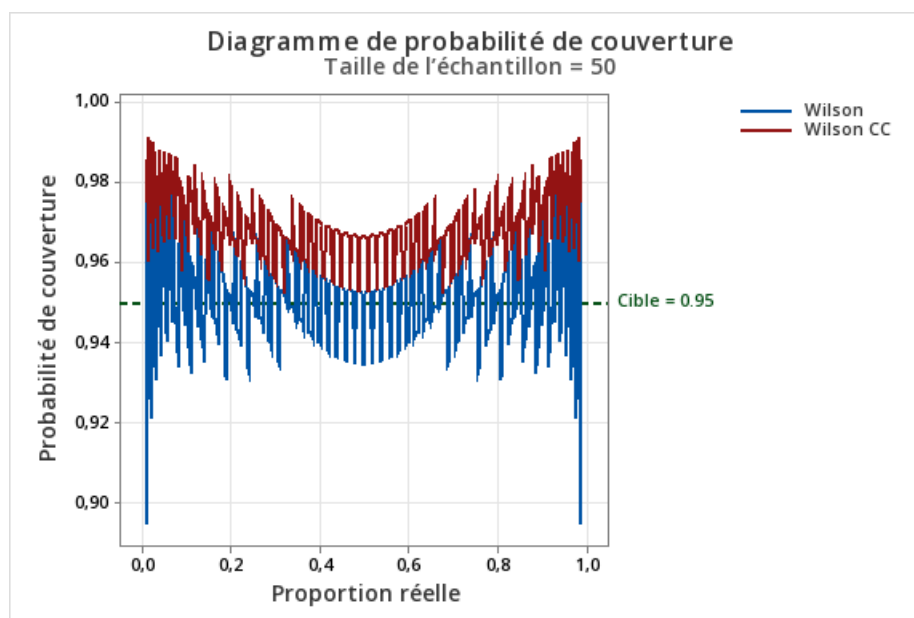
La méthode de Blaker ajustée, due à Klaschka et Reiczigel (2021), est une modification des méthodes exactes d'IC et d'essai de Blaker (2000, 2001). La modification tient compte de la nature à forte intensité de calcul de l'algorithme de Blaker d'origine et des incohérences occasionnelles entre son IC et les résultats des tests. Comme le CI Blaker d'origine, l'IC ajusté résultant est exact, imbriqué et contenu dans le CI Clopper-Pearson. Par conséquent, l'IC de Blaker ajusté est moins prudent que l'IC de Clopper-Pearson. Le CI est imbriqué dans le sens où un CI avec un niveau de confiance plus élevé contient toujours un CI avec un niveau de confiance plus faible. Par exemple, un IC Blaker bilatéral à 95 % (ajusté) contient toujours l'IC bilatéral à 90 % correspondant. L'imbrication est une propriété intéressante des méthodes d'IC exactes dérivées d'une distribution discrète telle que le binôme. Par exemple, le CI Clopper-Pearson est imbriqué. Il existe cependant des méthodes d'IC exactes qui ne sont pas nécessairement imbriquées. Par exemple, l'IC de Blyth-Still-Casella (Blyth et Still, 1983 ; Casella, 1986) est garanti d'être l'IC exact le plus court, mais n'est pas imbriqué. Crow (1956) CI n'est pas non plus imbriqué. Les calculs d'IC basés sur la méthode de Blaker ou la méthode de Blaker ajustée sont plus complexes que les méthodes d'IC classiques susmentionnées car elles nécessitent des algorithmes numériques. Cependant, avec les innovations actuelles en matière de technologie informatique, nous ne devons plus hésiter à mettre en œuvre des algorithmes complexes qui donnent de meilleurs résultats. La figure 2 illustre les améliorations de l'IC de Blaker ajusté par rapport à l'IC de Clopper-Pearson.



**Graphique 2 :** Comparaison des probabilités de couverture pour les IC ajustés de Blaker (A. Blaker) et de Clopper-Pearson (CP) en fonction de la proportion réelle lorsque la taille de l'échantillon est de 50. Le graphique indique que la probabilité de couverture de l'IC de Clopper-Pearson est au moins celle de l'IC de Blaker ajusté. Cela concorde avec le fait que les IC Blaker ajustés sont contenus dans les IC Clopper-Pearson. Pour un échantillon donné de taille 50, les probabilités moyennes de couverture (en supposant que la proportion réelle est uniformément répartie sur l'intervalle unitaire) sont de 0,960 et 0,969 pour l'IC de Blaker ajusté et l'IC de Clopper-Pearson, respectivement.

## 2. Les méthodes Wilson et Wilson CC

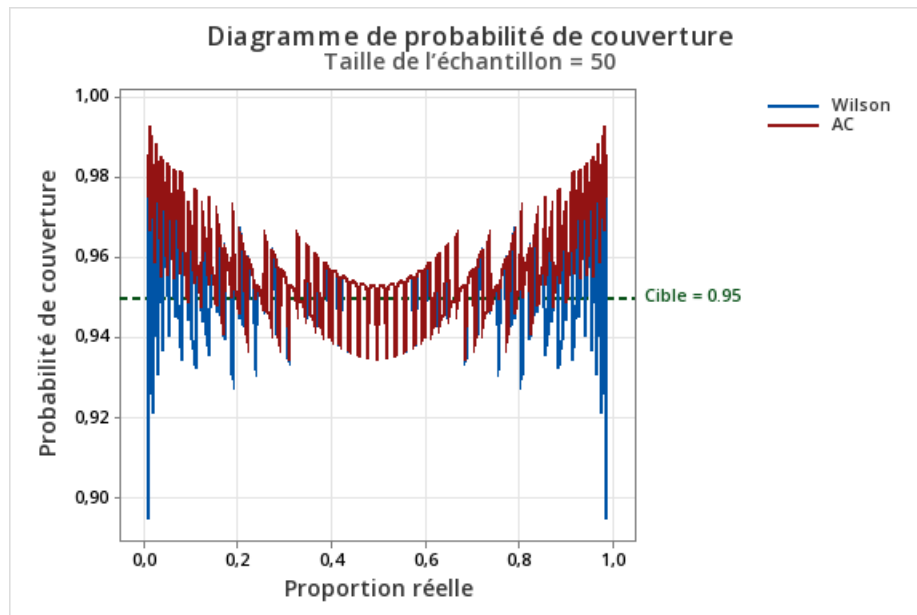
La méthode de l'IC de Wilson (1927) est dérivée de l'IC qui correspond au test de score, le test qui utilise l'erreur type nulle,  $\sqrt{p_0(1-p_0)/n}$ , par opposition à l'erreur type classique  $\sqrt{\hat{p}(1-\hat{p})/n}$ , sur le dénominateur de la statistique de test. C'est pour cette raison qu'on l'appelle aussi l'IC Wilson/score. Sa probabilité réelle de couverture peut être inférieure ou supérieure au niveau nominal visé, mais rester proche de celui-ci, sauf lorsque la proportion réelle est proche de 0 ou de 1 (voir figure 3). Un ajustement peut être effectué pour rendre l'IC de Wilson/score conservateur en utilisant la correction de continuité de Yates. Minitab fournit à la fois les versions de CI (avec et sans correction de continuité de Yates) et leurs tests d'hypothèses d'appariement.



**Graphique 3 :** Comparaison des probabilités de couverture pour les IC Wilson/score (Wilson) et le Wilson/score avec la correction de continuité de Yates (Wilson CC) en fonction de la proportion réelle lorsque la taille de l'échantillon est de 50. Le graphique montre que les IC CC de Wilson sont toujours conservateurs tandis que les IC Wilson sont conservateurs et libéraux en fonction de l'ampleur de la proportion réelle. En particulier, les IC de Wilson ont tendance à être trop libéraux lorsque la proportion réelle est très proche de 0 ou 1. Pour un échantillon donné de taille 50, les probabilités moyennes de couverture sont de 0,952 et 0,969 pour l'IC de Wilson et l'IC CC de Wilson, respectivement.

## 3. La méthode Agresti-Coull

L'IC d'Agresti-Coull est obtenu à partir d'un ajustement de l'IC de Wald classique excessivement libéral. L'IC qui en résulte a des propriétés de couverture similaires à celles des CI de Wilson, mais un peu plus conservatrices en général. De plus, les deux types de CI ont le même point médian, mais les CI de Wilson sont toujours contenus dans les CI d'Agresti-Coull. Comme l'illustre la figure 4, ils ont essentiellement les mêmes probabilités de couverture lorsque la proportion réelle est modérée. Cependant, l'IC d'Agresti Coull est généralement moins libéral lorsque la proportion réelle est proche de 0 ou de 1. Comme le montre la figure 4, pour un échantillon de taille 50, l'IC d'Agresti-Coull devient prudent à mesure que la proportion réelle se rapproche de 0 ou de 1. Un autre attrait de l'Agresti-Coull CI est sa simplicité de mise en œuvre héritée de l'IC de Wald. De plus, il est facile à enseigner et à retenir, en particulier lorsque le niveau de confiance est de 95%. Pour ce niveau de confiance, on parle communément de méthode de l'IC « ajouter 2 succès et 2 échecs » en tant que mémorandum à l'ajustement effectué sur l'IC de Wald pour le dériver.



**Graphique 4 :** Comparaison des probabilités de couverture pour les IC de Wilson/score (Wilson) et les IC d'Agresti-Coull (AC) en fonction de la proportion réelle lorsque la taille de l'échantillon est de 50. Le graphique indique que les deux méthodes donnent essentiellement les mêmes probabilités de couverture pour des valeurs modérées de la proportion vraie dans l'intervalle (0,3, 0,75). Cependant, pour les valeurs de la proportion réelle proche de 0 ou 1, les IC de Wilson sont libéraux tandis que les IC d'Agresti-Coull sont conservateurs. Ceci est cohérent avec le fait que les IC de Wilson sont contenus dans les IC d'Agresti-Coull. Pour un échantillon donné de taille 50, les probabilités moyennes de couverture sont de 0,952 et 0,958 pour l'IC de Wilson et l'IC d'Agresti-Coull, respectivement.

## Quelques exemples illustratifs simples

Un ingénieur qualité d'une usine de production de masse a sélectionné un échantillon aléatoire de 1465 produits de masse un jour donné. Après des tests indépendants sur les 1465 articles, 53 se sont avérés défectueux. L'ingénieur veut savoir si la proportion d'articles défectueux produits le jour donné est significativement différente de 2,5 %.

Bien que ce problème soit présenté comme une question de test d'hypothèse statistique, les statisticiens appliqués sont de plus en plus encouragés à inclure dans le résultat de l'analyse une estimation ponctuelle et un intervalle de confiance ainsi que la valeur p du test. Minitab suit cette convention autant que possible, en particulier dans les modules de statistiques de base. Par exemple, à l'aide de Minitab, les résultats de l'analyse de la question ci-dessus basés sur la méthode de Blaker ajustée sont les suivants.

Test et IC pour 1 proportion

### Méthode

p proportion d'événement

Méthode ajustée Méthode exacte de Blaker

### Statistiques descriptives

**N Échantillon d'événement p IC à 95 % pour p**

---

1465 53 0.036177 (0.027353, 0.046822)

### Test

Null hypothesis  $H_0: p = 0.025$

Hypothèse alternative  $H_1: p \neq 0,025$

### Valeur de P

0,009

Une sortie similaire peut être générée pour chacune des 4 méthodes. Les caractéristiques et avantages de chaque méthode sont récapitulés dans le tableau suivant.

Méthode	95%	Valeur P du test d'appariement
Blaker ajusté	(2,74%,4,68%)	0,009
Wilson avec la correction de Yates	(2,75%,4,74%)	0,008
Wilson/score	(2,78%,4,70%)	0,006
Agresti-Coull	(2,77%,4,74%)	0,007

Dans cet exemple, toutes les méthodes aboutissent à la même conclusion selon laquelle le % défectueux diffère de 2,5 % au niveau de signification de 0,05 puisque toutes les valeurs de p sont inférieures à 0,05. Les intervalles de confiance et les valeurs p correspondantes pour toutes les méthodes sont similaires, en partie parce que la taille de l'échantillon est très grande. De plus, l'IC de chaque méthode ne couvre pas la valeur de proportion hypothétique (2,5 %), ce qui est cohérent avec la valeur de p de chacun des tests d'hypothèse d'appariement.

Dans l'exemple ci-dessus, supposons maintenant que l'ingénieur qualité n'ait testé que 50 articles et ait constaté que 2 étaient défectueux. De plus, supposons que l'ingénieur veuille savoir si la proportion de défauts est significativement différente de 1,0 %. Les caractéristiques et avantages de chaque méthode sont récapitulés dans le tableau suivant.

Méthode	95%	Valeur P du test correspondant
Blaker ajusté	(0,72%,13,35%)	0,089
Wilson avec la correction de Yates	(0,70%,14,86%)	0,155
Wilson/score	(1,10%,13,46%)	0,033
Agresti-Coull	(0,34%,14,22%)	0,124

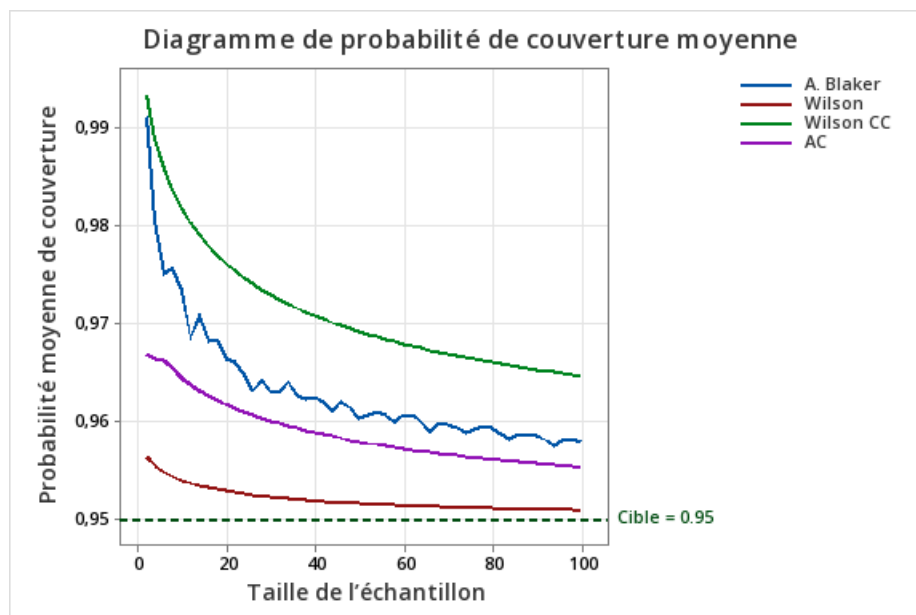
Dans ce cas, seule la méthode Wilson/score permet de tirer une conclusion significative en ce sens que le % défectueux diffère de 1,0 % au niveau de signification de 0,05. Au même niveau de signification, toutes les autres méthodes aboutissent à la conclusion

opposée d'une preuve insuffisante pour décider qu'il y a une différence. Les incohérences dans les résultats d'une méthode à l'autre sont, en grande partie, dues au fait que la taille de l'échantillon est modérée. En moyenne, les probabilités de couverture de ces méthodes se rapprochent du niveau nominal à mesure que la taille de l'échantillon augmente (voir la figure 5 ci-dessous). Pour les plans d'échantillonnage petits à modérés, cependant, les disparités dans les probabilités moyennes de couverture associées à chaque méthode sont plus prononcées, ce qui fait que les IC correspondants ont des largeurs sensiblement différentes. Une question importante, cependant, est de savoir quel résultat rapporter au patron ? Il n'y a pas de réponse simple à cette question, car une meilleure réponse peut dépendre de facteurs tels que la connaissance préalable de l'ampleur de la proportion réelle ou même du domaine d'application. Nous fournirons quelques lignes directrices générales dans la section suivante.

## Conclusion

Les figures 2, 3 et 4 montrent que les 4 méthodes d'IC, Blaker ajusté, Wilson, Wilson CC et Agresti-Coull ont des propriétés de probabilité de couverture différentes. Le Wilson CC est le plus conservateur, suivi du Blaker ajusté. L'Agresti, le Coull et le Wilson sont souvent libéraux et conservateurs, selon l'ampleur de la proportion réelle. Dans l'ensemble, la méthode Wilson CI est la plus libérale des 4 méthodes. De plus, la figure 5 indique qu'en moyenne, les 4 méthodes sont conservatrices, la CC de Wilson étant la plus conservatrice, suivie de la méthode ajustée de Blaker, de la méthode Agresti-Coull et de la méthode Wilson. Pour un problème donné, la méthode appropriée à choisir peut dépendre de l'application particulière, de la taille de l'échantillon et de l'existence ou non d'une connaissance préalable de l'ampleur de la proportion réelle.

Par exemple, les organismes de réglementation s'intéressent souvent à des méthodes conservatrices afin de protéger les consommateurs. Cependant, une méthode trop conservatrice peut donner lieu à des réglementations strictes, tandis qu'une méthode trop libérale peut donner lieu à des réglementations souples. En général, pour les plans d'échantillonnage modérés à grands, la méthode Blaker ajustée ou la méthode Willson CC peuvent être de bons choix. Pour les petits plans d'échantillonnage, le Wilson ou l'Agresti-Coull peut être approprié si l'on dispose d'une connaissance préalable de la valeur réelle des proportions. Ces connaissances préalables sont souvent fondées sur des expériences similaires antérieures ou sur une petite étude pilote spécialement conçue pour obtenir une estimation approximative de la proportion. Par exemple, dans le domaine du contrôle de la qualité des applications où la proportion de défauts est généralement proche de 0, la méthode à choisir dépend du fait que l'investigateur souhaite ou non un résultat conservateur. Un résultat conservateur peut être basé sur la méthode Agresti-Coull tandis qu'un résultat libéral peut être basé sur la méthode Wilson/score. Enfin, et c'est le plus important, il est toujours bon de planifier à l'avance la taille de l'échantillon afin de se prémunir contre l'erreur de type II ou de contrôler la largeur des IC. Minitab dispose d'une application pour cela aussi. Des outils sont disponibles dans Minitab pour déterminer la taille d'échantillon adéquate afin de contrôler l'erreur de type II associée aux tests d'hypothèses ou de contrôler la largeur des IC.



**Graphique 5** : Probabilité moyenne de couverture des 4 méthodes d'IC en fonction de la taille de l'échantillon. Les probabilités moyennes de couverture sont calculées en supposant que la proportion réelle est uniformément répartie sur l'intervalle unitaire. Les courbes de couverture moyenne illustrent qu'en moyenne, la méthode de l'IC Wilson/score avec correction de continuité de Yates (Wilson CC) est la plus prudente, suivie des méthodes ajustées de Blaker (A. Blaker), d'Agresti-Coull (AC) et de Wilson/score (Wilson). Les courbes de probabilité de couverture moyenne se rapprochent du niveau de couverture nominal visé à mesure que la taille de l'échantillon augmente. De plus, alors que les courbes de couverture moyenne pour les méthodes approximatives (Wilson CC, Wilson, Agresti-Coull) sont lisses, la courbe de couverture pour le Blaker ajusté exact présente quelques mouvements de fluctuation à mesure qu'elle se rapproche de la couverture nominale. Cela indique, peut-être, que la méthode de l'IC de Blaker ajustée peut encore être améliorée.

## Référence

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.





Vous avez des données. Nous avons Solutions Analytics.

Télécharger une version  
de [www.minitab.com](https://www.minitab.com)