

# Test d'écart type à 1 échantillon

## Généralités

Le test d'écart type à 1 échantillon est utilisé pour évaluer la variabilité de votre procédé et la comparer à une valeur cible. En général, la variabilité est mesurée à l'aide de la variance, ou, de façon équivalente, de l'écart type.

Plusieurs méthodes statistiques ont été développées pour évaluer la variance d'une population, chacune avec ses avantages et ses limites. La méthode classique du Khi deux est probablement la plus utilisée pour tester la variance, mais elle est extrêmement sensible à l'hypothèse de normalité et peut produire des résultats très inexacts lorsque les données sont asymétriques ou présentent des queues lourdes. D'autres méthodes ont également été développées, mais elles présentent elles aussi des inconvénients. Par exemple, certaines méthodes ne sont valides que pour de grands échantillons ou pour des données obéissant à une loi de distribution symétrique (voir Annexe A).

Dans Minitab 15, nous utilisons une méthode alternative pour grands échantillons inspirée de la méthode de Box (1953) d'approximation de la distribution de la variance d'échantillon à partir d'une loi du Khi deux mise à l'échelle. Cette méthode, appelée méthode des degrés de liberté ajustés (DL(ajust)), est moins sensible à l'hypothèse de normalité pour des échantillons suffisamment grands et il a été démontré qu'elle fournit des intervalles de confiance plus exacts que d'autres méthodes (Hummel, Banga et Hettmansperger, 2005). Plus récemment, cependant, une méthode statistique améliorée a été développée par Bonett (2006) et semble fournir de meilleures approximations pour les intervalles de confiance.

Ce document vise à évaluer l'efficacité de la méthode de Bonett. En outre, pour la planification de l'effectif d'échantillon, nous examinons la puissance de la procédure de test utilisant les intervalles de confiance de Bonett. Au vu des résultats obtenus, nous avons décidé d'utiliser la méthode Bonett pour le test d'écart type à 1 échantillon de l'Assistant. Nous étudions également les vérifications des données suivantes, qui sont automatiquement effectuées et affichées dans le rapport de l'Assistant, et nous expliquons comment elles influent sur les résultats :

- Données aberrantes
- Validité du test
- Effectif de l'échantillon

# Méthodes de test d'écart type à 1 échantillon

## Méthode de Bonett contre méthode DL(ajust)

Avant l'apparition de la méthode de Bonett (2006), la procédure la plus robuste pour la création d'inférences sur la variance d'une population était certainement la méthode DL(ajust). Néanmoins, les résultats publiés par Bonett montrent que sa méthode fournit des niveaux de confiance stables et proches du niveau cible lorsque des échantillons de taille modérée sont prélevés dans des populations non normales. Aussi la méthode de Bonett est-elle sans doute préférable pour le calcul d'inférences sur l'écart type ou la variance d'une population.

### Objectif

Nous souhaitons comparer les performances de la méthode de Bonett à celles de la méthode DL(ajust) pour la création d'inférences sur la variance d'une population unique. De manière plus spécifique, nous voulions déterminer quelle méthode produisait les intervalles de confiance les plus exacts pour la variance (ou l'écart type) lorsque des échantillons ayant des effectifs différents étaient générés à partir de populations non normales.

Nous effectuons une comparaison des intervalles de confiance car la méthode de Bonett s'applique directement aux intervalles de confiance. Il est possible de déduire une procédure de test d'hypothèse à partir des intervalles de confiance de Bonett. Cela, dit pour comparer directement nos résultats à ceux publiés par Bonett (2006), nous avons examiné les intervalles de confiance plutôt que les tests d'hypothèse.

### Méthode

Les méthodes DL(ajust) et de Bonett sont toutes deux décrites en détail dans l'Annexe B. Pour comparer l'exactitude des intervalles de confiance obtenus avec chaque méthode, nous avons effectué les simulations suivantes. Nous avons d'abord généré des échantillons aléatoires présentant des effectifs différents et obéissant à des lois de distribution diverses, telles que des lois asymétriques à queues lourdes, des lois symétriques à queues lourdes et des lois symétriques à queues légères. Pour chaque type de loi et chaque effectif d'échantillon, 10 000 répliques d'échantillons ont été prélevées et des intervalles de confiance bilatéraux à 95 % de la variance ont été calculés à l'aide de chacune des deux méthodes. Nous avons ensuite calculé la proportion des 10 000 intervalles qui contenaient la variance véritable, aussi appelée probabilité de couverture simulée. Si les intervalles de confiance sont exacts, la probabilité de couverture simulée sera proche de la probabilité de couverture cible de 0,95. En outre, nous avons calculé la largeur moyenne des intervalles de confiance pour chaque méthode. Si les intervalles de confiance obtenus avec les deux méthodes présentent approximativement les mêmes probabilités de couverture simulée, la méthode qui permet d'obtenir les intervalles les plus courts (en moyenne) est la plus précise. Pour plus de détails, reportez-vous à l'Annexe C.

## Les résultats

De façon générale, la méthode de Bonett produit de meilleures probabilités de couverture et des intervalles de confiance plus précis que la méthode DL(ajust). Par conséquent, les tests statistiques de la variance effectués à l'aide de la méthode de Bonett présentent des taux d'erreur de 1ère et 2ème espèce inférieurs. Aussi, le test d'écart type à 1 échantillon de l'Assistant utilise la méthode de Bonett.

Ceci dit, nos résultats indiquent que si la loi présente des queues modérées ou lourdes, la méthode de Bonett requiert des effectifs d'échantillons plus grands pour parvenir au niveau d'exactitude cible :

- Pour des lois à queues normales ou légères, un effectif d'échantillon de 20 est suffisant.
- Pour des lois avec des queues modérément lourdes, l'effectif d'échantillon doit être d'au moins 80.
- Pour des lois avec des queues lourdes, l'effectif d'échantillon doit être d'au moins 200.

Aussi, pour vérifier la validité du test d'écart type à 1 échantillon ou des résultats des intervalles de confiance pour vos données, l'Assistant teste ces dernières afin d'évaluer à la fois l'effectif de l'échantillon et les queues de la loi de distribution des données (reportez-vous à la partie consacrée à la vérification des données pour validité de test ci-dessous).

## Efficacité de la puissance théorique

La méthode de Bonett sert avant tout à calculer les intervalles de confiance de la variance (ou de l'écart type). Cependant, la relation statistique qui existe entre les tests d'hypothèse et les intervalles de confiance nous permet d'utiliser les intervalles de confiance de Bonett obtenus par approximation pour réaliser un test d'hypothèse équivalent. Comme il n'existe aucune fonction puissance exacte pour ce test, il nous fallait en adapter une. Nous voulions également évaluer la sensibilité de cette fonction puissance théorique à l'hypothèse de normalité.

### Objectif

Nous souhaitons déterminer si nous pouvions utiliser cette fonction puissance théorique dérivée des intervalles de confiance de Bonett pour évaluer la puissance et l'effectif d'échantillon requis pour le test d'écart type à 1 échantillon de l'Assistant. Pour ce faire, il nous fallait déterminer si cette fonction puissance théorique reflétait la puissance réelle du test lors de l'analyse de données normales et non normales.

### Méthode

La fonction puissance théorique du test obtenue avec la méthode de Bonett est calculée dans l'Annexe C. Nous avons réalisé des simulations afin d'estimer les niveaux de puissance réels (que nous appelons niveaux de puissance simulés) obtenus avec la méthode de Bonett. Nous avons d'abord généré des échantillons aléatoires présentant différents effectifs, obéissant aux lois de distribution décrites dans l'étude précédente : asymétriques à queues

lourdes, symétriques à queues lourdes et symétriques à queues légères. Pour chaque loi, nous avons effectué le test sur chacune des 10 000 répliques d'échantillons. Pour chaque effectif d'échantillon, nous avons calculé la puissance simulée du test pour la détection d'une différence donnée, exprimée comme la fraction des 10 000 échantillons pour laquelle le test est significatif. Pour effectuer la comparaison, nous avons également calculé le niveau de puissance correspondant obtenu avec la fonction puissance théorique du test. Si la fonction puissance théorique n'est pas trop sensible à la normalité, les niveaux de puissance simulés et théorique doivent être proches pour les données normales et non normales. Pour plus de détails, reportez-vous à l'Annexe D.

## Les résultats

Nos simulations ont montré que lorsque l'échantillon obéit à une loi à queues normales ou légères, les puissances simulée et théorique du test obtenues avec la méthode de Bonett sont approximativement égales. Lorsque l'échantillon obéit à une loi à queues lourdes, en revanche, la fonction puissance théorique peut se montrer trop prudente et surestimer l'effectif d'échantillon requis pour obtenir une puissance donnée. Par conséquent, la fonction puissance théorique du test permet d'être certain que l'effectif des échantillons est assez grand pour détecter une différence pratique importante dans l'écart type, quelle que soit la loi de distribution. En revanche, si les données obéissent à une loi à queues lourde, l'effectif d'échantillon estimé peut être plus grand que celui réellement nécessaire, ce qui peut entraîner des dépenses inutiles lors de l'échantillonnage des éléments.

# Vérification des données

## Données aberrantes

Les données aberrantes sont des valeurs extrêmement grandes ou extrêmement petites, également connues sous le nom de valeurs aberrantes. Les données aberrantes peuvent avoir une forte influence sur les résultats de l'analyse et peuvent compromettre la possibilité de trouver des résultats statistiquement significatifs, notamment avec de petits échantillons. Les données aberrantes peuvent venir de problèmes de collecte de données ou être dues à un comportement inhabituel du procédé étudié. Ainsi, il vaut souvent la peine d'examiner ces points de données plus en profondeur et de les corriger lorsque cela est possible.

### Objectif

Nous souhaitons développer une méthode pour vérifier les valeurs très grandes ou très petites par rapport à l'échantillon global et susceptibles d'influer sur les résultats de l'analyse.



### Méthode

Nous avons développé une méthode pour vérifier les données aberrantes inspirée de la méthode décrite par Hoaglin, Iglewicz et Tukey (1986), qui permet d'identifier les valeurs aberrantes dans les boîtes à moustache.

### Les résultats

L'Assistant identifie un point de données comme aberrant s'il se trouve à une distance 1,5 fois supérieur à l'étendue interquartile au-delà du quartile inférieur ou supérieur de la distribution. Les quartiles inférieur et supérieur sont les 25ème et 75ème percentiles des données. L'étendue interquartile représente la différence entre les deux quartiles. Cette méthode donne de bons résultats même lorsqu'il existe plusieurs valeurs aberrantes car elle permet de détecter chaque valeur aberrante spécifique.

Lors du test des données aberrantes, l'Assistant affiche les indicateurs d'état suivants dans le rapport :

Etat	Condition
	Il n'existe aucun point de données aberrant.
	Au moins un point de données est aberrant et peut avoir une forte influence sur les résultats.

## Validité du test

Dans la section relative aux méthodes de test d'écart type à 1 échantillon ci-dessus, nous avons démontré que la méthode de Bonett fournit généralement de meilleurs résultats que la méthode DL(ajust). En revanche, lorsque les queues d'une loi sont plus lourdes, la méthode de Bonett requiert des effectifs d'échantillon plus grands pour obtenir des résultats exacts. Par conséquent, pour évaluer la validité du test, il nous faut utiliser une méthode qui prenne non seulement en compte l'effectif d'échantillon, mais aussi la lourdeur des queues de la loi parent. Gel et al. (2007) ont développé un test permettant de déterminer si un échantillon provient d'une loi à queues lourdes. Ce test, appelé test SJ, utilise le rapport entre l'écart type de l'échantillon ( $s$ ) et l'estimateur de queue  $J$  (pour plus de détails, reportez-vous à l'Annexe E).

### Objectif

Pour un échantillon de données spécifique, il nous fallait développer une règle permettant de déterminer la validité de la méthode de Bonett en évaluant la lourdeur des queues des données.



### Méthode

Nous avons réalisé des simulations pour examiner la puissance du test SJ pour l'identification de lois à queues lourdes. Si le test SJ est suffisamment puissant avec des échantillons modérément grands, il peut être utilisé pour distinguer les lois à queues lourdes des lois à queues légères et servir ainsi nos besoins. Pour plus de détails, reportez-vous à l'Annexe F.

### Les résultats

Nos simulations ont démontré que lorsque les échantillons sont suffisamment grands, le test SJ peut être utilisé pour distinguer les lois à queues lourdes de celles à queues légères. Pour les échantillons à effectif modéré ou grand, des valeurs de  $p$  réduites indiquent des queues lourdes, tandis que des valeurs de  $p$  élevées indiquent des queues légères. Néanmoins, étant donné que les grands échantillons ont tendance à présenter des valeurs de  $p$  plus réduites que les petits échantillons, nous devons aussi prendre en considération l'effectif de l'échantillon lors de la détermination de la lourdeur des queues. Aussi, l'ensemble de règles de classement des queues de loi pour chaque échantillon que nous avons définies pour l'Assistant tiennent compte à la fois de l'effectif de l'échantillon et de la valeur de  $p$  du test SJ. Pour voir les étendues de valeurs de  $p$  et d'effectifs d'échantillons associées à des distributions à queues légères, modérées et lourdes, consultez l'Annexe F.

En fonction de ces résultats, le rapport de l'Assistant affiche les indicateurs d'état suivant pour évaluer la validité du test d'écart type à 1 échantillon (méthode de Bonett) pour vos données échantillon :

Etat	Condition
	<p>Votre échantillon ne semble pas posséder de queues lourdes. L'effectif de votre échantillon est suffisamment grand pour que vous puissiez vous fier à cette vérification.</p> <p>OU</p> <p>Les queues de votre échantillon sont modérément lourdes ou lourdes. Cependant, l'effectif de votre échantillon est suffisamment grand pour compenser, la valeur de p devrait donc être exacte.</p>
	<p>Les queues de votre échantillon sont modérément lourdes ou lourdes. L'effectif de votre échantillon n'est pas assez grand pour compenser. Faites preuve de prudence lors de l'interprétation des résultats.</p> <p>OU</p> <p>Votre échantillon n'est pas suffisamment grand pour vérifier l'existence de queues lourdes de façon fiable. Faites preuve de prudence lors de l'interprétation des résultats.</p>

## Effectif de l'échantillon

En général, un test d'hypothèse statistique permet de collecter des preuves permettant de rejeter l'hypothèse nulle de "non différence". Si l'échantillon est trop petit, la puissance de ce test peut ne pas être adéquate pour détecter une différence qui existe réellement, ce qui entraîne une erreur de 2ème espèce. Il est donc essentiel de s'assurer que les effectifs d'échantillons sont suffisamment grands pour détecter des différences importantes dans la pratique avec une probabilité élevée.

### Objectif

Si les données ne permettent pas de rejeter l'hypothèse nulle, il nous faut déterminer si les effectifs d'échantillons sont suffisamment grands pour que le test détecte des différences pratiques avec une probabilité élevée. Même si la planification d'effectif d'échantillon vise à garantir que les effectifs d'échantillons sont suffisamment grands pour détecter d'importantes différences avec une probabilité élevée, ces effectifs ne doivent pas être grands au point que des différences sans importance deviennent statistiquement significatives avec une probabilité élevée.

### Méthode






L'analyse de puissance et d'effectif d'échantillon pour le test d'écart type à 1 échantillon s'appuie sur la fonction puissance théorique du test. Cette fonction puissance fournit des estimations correctes lorsque les données présentent des queues normales ou légères, mais peuvent produire des estimations trop prudentes lorsque les données présentent des queues lourdes (reportez-vous aux résultats de simulation récapitulés dans la section Méthodes de test d'écart type à 1 échantillon ci-dessus, dans la partie relative à l'efficacité de la fonction puissance théorique).



## Les résultats

Lorsque les données ne fournissent pas suffisamment de preuves invalidant l'hypothèse nulle, l'Assistant utilise la fonction puissance du test d'approximation selon la loi normale pour calculer les différences pratiques pouvant être détectées avec une probabilité de 80 % et de 90 % pour l'effectif d'échantillon donné. De plus, si l'utilisateur indique une différence pratique spécifique présentant un intérêt particulier, l'Assistant utilise la fonction puissance du test d'approximation selon la loi normale pour calculer les effectifs d'échantillons qui offrent une probabilité de 80 % et de 90 % de détecter la différence.

Pour faciliter l'interprétation des résultats, le rapport de l'Assistant pour le test d'écart type à 1 échantillon affiche les indicateurs d'état suivants lors de la vérification de la puissance et de l'effectif d'échantillon :

Etat	Condition
	<p>Le test détecte une différence entre l'écart type et la valeur cible, la puissance n'est donc pas un problème.</p> <p>OU</p> <p>La puissance est suffisante. Le test n'a pas détecté de différence entre l'écart type et la valeur cible, mais l'échantillon est suffisamment grand pour fournir une probabilité d'au moins 90 % de détecter la différence donnée.</p>
	<p>La puissance est peut-être suffisante. Le test n'a pas détecté de différence entre l'écart type et la valeur cible, mais l'échantillon est suffisamment grand pour fournir une probabilité de 80 % à 90 % de détecter la différence donnée. L'effectif d'échantillon nécessaire pour atteindre une puissance de 90 % est indiqué.</p>
	<p>La puissance n'est peut-être pas suffisante. Le test n'a pas détecté de différence entre l'écart type et la valeur cible, et l'échantillon est suffisamment grand pour fournir une probabilité de 60 % à 80 % de détecter la différence donnée. Les effectifs d'échantillon nécessaires pour atteindre une puissance de 80 % et de 90 % sont indiqués.</p>
	<p>La puissance n'est pas suffisante. Le test n'a pas détecté de différence entre l'écart type et la valeur cible, et l'échantillon n'est pas suffisamment grand pour fournir une probabilité d'au moins 60 % de détecter la différence donnée. Les effectifs d'échantillon nécessaires pour atteindre une puissance de 80 % et de 90 % sont indiqués.</p>
	<p>Le test n'a pas détecté de différence entre l'écart type et la valeur cible. Vous n'avez pas spécifié de différence pratique à détecter. Selon vos données, le rapport peut indiquer les différences que vous pourriez détecter avec une probabilité de 80 % et de 90 %, en fonction de l'effectif de votre échantillon et de la valeur d'alpha.</p>

# Références

Bonett, D.G., (2006), Approximate confidence interval for standard deviation of nonnormal distributions, *Computational Statistics & Data Analysis*, 50, 775-782.

Box, G.E.P., (1953), Non-normality and tests on variances, *Biometrika*, 40, 318.

Efron, B. et Tibshirani, R. J., (1993), An introduction to the bootstrap, Boca Raton, FL, Chapman et Hall/CRC.

Gel, Y. R., Miao, W. et Gastwirth, J. L., (2007), Robust directed tests of normality against heavy-tailed alternatives, *Computational Statistics & Data Analysis*, 51, 2734-2746.

Hummel, R., Banga, S. et Hettmansperger, T.P. (2005), Better confidence intervals for the variance in a random sample, *Minitab Technical Report*.

Lee, S.J. et Ping, S., (1996), Testing the variance of symmetric heavy-tailed distributions, *Journal of Statistical Computation and Simulation*, 56, 39-52.

# Annexe A : méthodes de test de la variance (ou l'écart type)

Le tableau ci-dessous récapitule les avantages et les limites des diverses méthodes de test de la variance.

Méthode	Commentaire
Procédure Khi deux classique	Extrêmement sensible à l'hypothèse de normalité. Même de petits écarts par rapports à la normalité peuvent produire des résultats imprécis, quel que soit l'effectif de l'échantillon. En fait, lorsque les données s'écartent de la normalité, l'augmentation de l'effectif de l'échantillon diminue l'exactitude de la procédure.
Méthode pour grands échantillons fondée sur la loi normale asymptotique de la transformation logarithmique de la variance de l'échantillon	Généralement plus efficace que la méthode du Khi deux classique, mais exige que les plus grands échantillons soient insensibles à l'hypothèse de normalité.
Méthode pour grands échantillons fondée sur le développement d'Edgeworth pour les tests unilatéraux (queue supérieure) Reportez-vous à Lee et Ping (1996).	Produit des taux d'erreur de 1ère espèce acceptables mais exige que les données obéissent à une loi symétrique.
Méthode pour grands échantillons fondée sur une approximation de la distribution de la variance de l'échantillon par une loi du Khi deux mise à l'échelle. Cette méthode est appelée méthode des degrés de liberté ajustés (DL(ajust)). Reportez-vous à Hummel, Banga et Hettmansperger (2005).	Fournit une meilleure probabilité de couverture que la méthode basée sur la loi normale asymptotique de la transformation logarithmique de la variance de l'échantillon et que la méthode d'approximation de bootstrap ABC non paramétrique pour les intervalles de confiance (Efron et Tibshirani, 1993). Utilisée pour le test à 1 variance dans Minitab 15.
Loi asymptotique ajustée de Bonett de la transformation logarithmique de la variance de l'échantillon Reportez-vous à Bonett (2006).	Fournit une bonne probabilité de couverture pour les intervalles de confiance, même pour les échantillons modérément grands. Cette méthode exige cependant des échantillons beaucoup plus grands lorsque les données obéissent à des lois à queues lourdes. Utilisée pour le test à 1 variance et pour le test d'écart type à 1 échantillon de l'Assistant de Minitab 16.

# Annexe B : définition de la méthode de Bonett et de la méthode DL(ajust)

Soit  $x_1, \dots, x_n$  un échantillon aléatoire observé de taille  $n$ , issu d'une population avec un moment d'ordre 4 fini. Soit  $\bar{x}$  et  $s$  la moyenne et l'écart type de l'échantillon, respectivement. Soit également  $\gamma$  et  $\gamma_e$  le coefficient d'aplatissement de la population et l'excès d'aplatissement, respectivement, avec  $\gamma_e = \gamma - 3$ . Ainsi, pour une population normale,  $\gamma = 3$  et  $\gamma_e = 0$ . Enfin, soit  $\sigma^2$  la variance de population inconnue. Dans les sections suivantes, nous présentons deux méthodes pour la réalisation d'une inférence sur  $\sigma^2$ , la méthode des degrés de liberté ajustés (DL(ajust)) et la méthode de Bonett.

## Formule B1 : méthode DL(ajust)

La méthode DL(ajust) utilise sur une approximation de la distribution de la variance d'échantillon par une loi du Khi deux mise à l'échelle (voir Box, 1953). Plus spécifiquement, les deux premiers moments de la variance de l'échantillon sont comparés aux moments d'une loi du Khi deux mise à l'échelle pour déterminer l'échelle et les degrés de liberté inconnus. Cette approche produit l'approximation d'intervalle de confiance bilatéral à  $(1 - \alpha)100$  pour cent suivant pour la variance :

$$\left[ \frac{rs^2}{\chi_{r,\alpha/2}^2}, \frac{rs^2}{\chi_{r,1-\alpha/2}^2} \right]$$

avec

$$r = \frac{2n}{\hat{\gamma}_e + 2n/(n-1)}$$
$$\hat{\gamma}_e = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Cette estimation de l'excès d'aplatissement est identique à celle utilisée pour les commandes de statistiques élémentaires dans Minitab.

## Formule B2 : méthode de Bonett

La méthode de Bonett s'appuie sur l'approche classique bien connue utilisant le théorème de la limite centrale et sur la méthode  $\delta$  de Cramer pour obtenir une loi asymptotique de la transformation logarithmique de la variance de l'échantillon. La transformation logarithmique est utilisée pour accélérer la convergence vers la normalité. Avec cette approche, l'approximation de l'intervalle de confiance bilatéral à  $(1 - \alpha)100$  pour cent de la variance est définie comme :

$$\left[ s^2 \exp(-z_{\alpha/2} es), s^2 \exp(z_{\alpha/2} es) \right]$$

où  $z_\alpha$  est le percentile supérieur de la distribution normale standard et  $es$  est une estimation asymptotique de l'erreur standard de la variance de l'échantillon transformée par logarithme, exprimée de la façon suivante :

$$es = \sqrt{\frac{\hat{\gamma} - (n-3)/n}{n-1}} = \sqrt{\frac{\hat{\gamma}_e + 2 + 3/n}{n-1}}$$

Dans un premier temps, Hummel et al. (2005) avaient réalisé des études de simulation démontrant que la méthode DL(ajust) était plus efficace que cette approche classique. Cependant, Bonett apporte deux ajustements à l'approche classique pour en surmonter les limites.

Le premier ajustement concerne l'estimation de l'aplatissement. Pour évaluer l'aplatissement, Bonett utilise la formule suivante :

$$\hat{\gamma}_e = \frac{n}{(n-1)^2} \sum_{i=1}^n \left( \frac{x_i - m}{s} \right)^4 - 3$$

où  $m$  est une moyenne tronquée avec une proportion de troncage égale à  $1/2\sqrt{n-4}$ . L'estimation de l'aplatissement a tendance à améliorer l'exactitude des niveaux de confiance pour les lois à queues lourdes (symétriques ou asymétriques).

Quant au second ajustement, Bonett détermine de façon empirique un multiplicateur constant pour la variance de l'échantillon et l'erreur standard. Ce multiplicateur constant égalise approximativement les probabilités des queues lorsque l'échantillon est petit ; il est exprimé de la façon suivante :

$$c = \frac{n}{n - z_{\alpha/2}}$$

Ces ajustements produisent l'approximation de l'intervalle de confiance bilatéral de Bonett à  $(1 - \alpha)100$  pour cent de la variance :

$$[cs^2 \exp(-c z_{\alpha/2} es), cs^2 \exp(c z_{\alpha/2} es)]$$

# Appendice C : exactitude comparée de la méthode de Bonett et de la méthode DL(ajust)

## Simulation C1 : comparaison des intervalles de confiance

Nous voulions comparer l'exactitude des intervalles de confiance de la variance lorsqu'ils sont calculés à l'aide de la méthode DL(ajust) et de la méthode de Bonett. Nous avons généré des échantillons aléatoires avec différents effectifs ( $n = 20, 30, 40, 50, 60, 80, 100, 150, 200, 250, 300$ ) et obéissant à différentes lois, puis nous avons calculé les intervalles de confiance à l'aide de chaque méthode. Les lois de distribution des échantillons étaient les suivantes :

- La loi normale standard ( $N(0, 1)$ )
- Des lois de distribution symétriques à queues légères, à savoir la loi de distribution uniforme ( $U(0, 1)$ ) et la loi bêta, avec les deux paramètres définis sur 3 ( $B(3,3)$ )
- Des lois de distribution symétriques à queues lourdes, à savoir des lois  $t$  à 5 et 10 degrés de liberté ( $t(5), t(10)$ ) et la loi de Laplace avec un emplacement 0 et une échelle 1 ( $Lpl$ )
- Des lois de distribution asymétriques à queues lourdes, à savoir la loi exponentielle avec une échelle 1 ( $Exp$ ) et des lois du Khi deux avec des degrés de liberté de 3, 5 et 10 ( $Khi(3), Khi(5), Khi(10)$ )
- Une loi de distribution à queues lourdes et présentant une asymétrie vers la gauche, plus précisément une loi bêta avec des paramètres définis sur 8 et 1, respectivement ( $B(8, 1)$ )

En outre, pour évaluer l'effet direct des valeurs aberrantes, nous avons généré des échantillons à partir de distributions normales contaminées tels que

$$CN(p, \sigma) = pN(0, 1) + (1 - p)N(0, \sigma)$$

où  $p$  est le paramètre de mélange et  $1 - p$  est la proportion de contamination (qui est égale à la proportion des valeurs aberrantes). Nous avons sélectionné deux populations normales contaminées pour l'étude :  $CN(0,9, 3)$ , où 10 % de la population est constituée de valeurs aberrantes, et  $CN(0,8, 3)$ , où 20 % de la population est constituée de valeurs aberrantes. Ces deux distributions sont symétriques et présentent de longues queues à cause des valeurs aberrantes.

Pour chaque loi et chaque effectif d'échantillons, 10 000 répliques d'échantillons ont été générées et des intervalles de confiance bilatéraux à 95 % ont été calculés à l'aide de chaque méthode. Les valeurs de départ entrées dans le générateur d'échantillons aléatoires sont définies de telle sorte que les deux méthodes soient appliquées à des échantillons

équivalents. A partir de ces intervalles de confiance, nous avons ensuite calculé la probabilité de couverture simulée (PCouv) et la largeur moyenne des intervalles (LMoy) pour chaque méthode. Si les intervalles de confiance obtenus avec les deux méthodes présentent approximativement les mêmes probabilités de couverture simulée, la méthode qui permet d'obtenir les intervalles les plus courts (en moyenne) est la plus précise. Etant donné que nous avons utilisé un niveau de confiance de 95 %, l'erreur de simulation était de  $\sqrt{0,95(0,05)/10\,000} = 0,2\%$ .

Les résultats de la simulation sont enregistrés dans les tableaux 1 et 2 ci-dessous.

Tableau 1 : Probabilité de couverture simulée des intervalles de confiance bilatéraux à 95 % de la variance, calculée à l'aide des méthodes DL(ajust) et de Bonett. Les échantillons générés obéissent à des lois de distribution symétriques à queues légères, normales, presque normales ou lourdes.

Loi de distribution	Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
	U(0, 1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	T(5)
Asymétrie	0	0	0	0	0	0	0	0
Aplatissement ( $r_e$ )	-1,20	-0,667	0	1,00	3,00	4,544	5,333	6,00

<b>n = 10</b>									
DL(ajust)	PCouv	0,91	0,909	0,903	0,883	0,853	0,793	0,815	0,858
	LMoy	0,154	0,087	3,276	5,16	13,924	21,658	14,913	11,742
Bonett	PCouv	0,972	0,967	0,962	0,952	0,919	0,891	0,92	0,935
	LMoy	0,242	0,115	3,71	5,134	10,566	15,335	10,367	8,578

<b>n = 20</b>									
DL(ajust)	PCouv	0,937	0,937	0,923	0,909	0,881	0,819	0,817	0,868
	LMoy	0,080	0,045	1,572	2,463	5,781	9,265	6,539	5,151
Bonett	PCouv	0,953	0,954	0,946	0,934	0,909	0,856	0,864	0,904
	LMoy	0,10	0,051	1,683	2,422	4,932	7,282	4,945	4,026

<b>n = 30</b>									
DL(ajust)	PCouv	0,946	0,942	0,933	0,917	0,894	0,851	0,823	0,882
	LMoy	0,061	0,034	1,17	1,764	4,117	6,33	4,557	3,667
Bonett	PCouv	0,951	0,95	0,947	0,933	0,909	0,869	0,852	0,907
	LMoy	0,070	0,037	1,221	1,75	3,654	5,383	3,736	2,997

<b>n = 40</b>									
---------------	--	--	--	--	--	--	--	--	--

Loi de distribution		Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0, 1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	T(5)
Asymétrie		0	0	0	0	0	0	0	0
Aplatissement ( $\gamma_e$ )		-1,20	-0,667	0	1,00	3,00	4,544	5,333	6,00
DL(ajust )	PCouv	0,953	0,947	0,932	0,922	0,904	0,867	0,833	0,89
	LMoy	0,051	0,028	0,971	1,489	3,246	5,131	3,654	3,024
Bonett	PCouv	0,954	0,951	0,941	0,936	0,914	0,879	0,856	0,907
	LMoy	0,057	0,030	1,002	1,469	2,994	4,519	3,128	2,542

**n = 50**

DL(ajust )	PCouv	0,951	0,945	0,937	0,925	0,911	0,878	0,838	0,893
	LMoy	0,045	0,025	0,849	1,291	2,789	4,357	3,091	2,603
Bonett	PCouv	0,951	0,947	0,944	0,938	0,918	0,888	0,855	0,908
	LMoy	0,049	0,026	0,87	1,28	2,613	3,939	2,729	2,24

**n = 60**

DL(ajust )	PCouv	0,949	0,943	0,938	0,926	0,913	0,89	0,853	0,899
	LMoy	0,040	0,022	0,766	1,155	2,49	3,857	2,768	2,283
Bonett	PCouv	0,949	0,947	0,943	0,935	0,918	0,896	0,868	0,91
	LMoy	0,043	0,023	0,781	1,147	2,354	3,552	2,498	2,023

**n = 70**

DL(ajust )	PCouv	0,948	0,945	0,94	0,93	0,913	0,89	0,858	0,896
	LMoy	0,037	0,020	0,701	1,056	2,283	3,458	2,475	2,049
Bonett	PCouv	0,947	0,946	0,944	0,938	0,918	0,894	0,868	0,905
	LMoy	0,039	0,021	0,713	1,049	2,174	3,227	2,272	1,828

**n = 80**

DL(ajust )	PCouv	0,947	0,949	0,938	0,929	0,918	0,905	0,869	0,902
	LMoy	0,034	0,019	0,652	0,988	2,089	3,205	2,30	1,906
Bonett	PCouv	0,946	0,95	0,942	0,935	0,923	0,907	0,877	0,911
	LMoy	0,036	0,019	0,662	0,982	2,005	3,014	2,133	1,716

**n = 90**

DL(ajust )	PCouv	0,946	0,947	0,948	0,929	0,918	0,908	0,869	0,901
	LMoy	0,032	0,018	0,611	0,921	1,951	2,982	2,124	1,874



Loi de distribution		Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0, 1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	T(5)
Asymétrie		0	0	0	0	0	0	0	0
Aplatissement ( $\gamma_e$ )		-1,20	-0,667	0	1,00	3,00	4,544	5,333	6,00
Bonett	PCouv	0,945	0,948	0,952	0,936	0,92	0,91	0,874	0,909
	LMoy	0,034	0,018	0,618	0,916	1,882	2,822	1,984	1,646

**n = 100**

DL(ajust )	PCouv	0,947	0,951	0,945	0,933	0,92	0,91	0,885	0,912
	LMoy	0,030	0,017	0,576	0,873	1,83	2,801	2,017	1,658
Bonett	PCouv	0,946	0,953	0,948	0,937	0,923	0,912	0,891	0,916
	LMoy	0,032	0,017	0,583	0,869	1,772	2,666	1,899	1,522

**n = 150**

DL(ajust )	PCouv	0,949	0,951	0,947	0,936	0,932	0,925	0,896	0,912
	LMoy	0,024	0,014	0,464	0,70	1,47	2,228	1,602	1,325
Bonett	PCouv	0,948	0,952	0,949	0,939	0,933	0,924	0,898	0,915
	LMoy	0,025	0,014	0,467	0,698	1,438	2,156	1,539	1,251

**n = 200**

DL(ajust )	PCouv	0,943	0,949	0,948	0,938	0,927	0,93	0,914	0,918
	LMoy	0,021	0,012	0,40	0,605	1,265	1,906	1,373	1,178
Bonett	PCouv	0,942	0,951	0,949	0,94	0,928	0,93	0,915	0,92
	LMoy	0,021	0,012	0,402	0,603	1,245	1,86	1,333	1,106

**n = 250**

DL(ajust )	PCouv	0,952	0,952	0,949	0,942	0,938	0,929	0,909	0,915
	LMoy	0,019	0,010	0,355	0,538	1,12	1,69	1,219	1,037
Bonett	PCouv	0,951	0,952	0,949	0,944	0,941	0,929	0,909	0,916
	LMoy	0,019	0,010	0,357	0,537	1,106	1,657	1,19	0,986

**n = 300**

DL(ajust )	PCouv	0,95	0,948	0,951	0,94	0,938	0,936	0,92	0,914
	LMoy	0,017	0,009	0,324	0,49	1,019	1,544	1,115	0,933
Bonett	PCouv	0,95	0,947	0,951	0,942	0,937	0,929	0,92	0,916
	LMoy	0,017	0,010	0,325	0,489	1,009	1,657	1,093	0,897

Tableau 2 : Probabilité de couverture simulée des intervalles de confiance bilatéraux à 95 % de la variance, calculée à l'aide des méthodes DL(ajust) et de Bonett. Les échantillons générés obéissent à des lois de distribution asymétriques à queues presque normales, modérément lourdes ou lourdes.

Loi de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
Asymétrie		0,894	-1,423	1,265	1,633	2
Aplatissement ( $\gamma_e$ )		1,20	2,284	2,40	4,00	6
<b>n = 10</b>						
DL(ajust)	PCouv	0,869	0,815	0,836	0,797	0,758
	LMoy	93,383	0,065	61,994	47,821	10,711
Bonett	PCouv	0,95	0,917	0,938	0,911	0,882
	LMoy	91,006	0,058	53,83	38,137	7,498
<b>n = 20</b>						
DL(ajust)	PCouv	0,889	0,862	0,862	0,833	0,811
	LMoy	41,497	0,026	25,479	20,099	4,293
Bonett	PCouv	0,932	0,912	0,913	0,893	0,877
	LMoy	41,60	0,026	24,094	17,232	3,37
<b>n = 30</b>						
DL(ajust)	PCouv	0,901	0,881	0,88	0,864	0,838
	LMoy	30,021	0,018	18,182	13,63	2,844
Bonett	PCouv	0,931	0,92	0,914	0,906	0,885
	LMoy	30,462	0,019	17,858	12,634	2,441
<b>n = 40</b>						
DL(ajust)	PCouv	0,909	0,882	0,885	0,867	0,862
	LMoy	24,459	0,015	14,577	10,649	2,193
Bonett	PCouv	0,93	0,915	0,913	0,904	0,898
	LMoy	24,952	0,015	14,504	1,991	1,991
<b>n = 50</b>						
DL(ajust)	PCouv	0,912	0,90	0,892	0,871	0,868
	LMoy	21,373	0,013	12,694	9,115	1,861

Loi de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
Asymétrie		0,894	-1,423	1,265	1,633	2
Aplatissement ( $\gamma_e$ )		1,20	2,284	2,40	4,00	6
Bonett	PCouv	0,93	0,927	0,916	0,903	0,901
	LMoy	21,814	0,013	12,741	8,897	1,735

**n = 60**

DL(ajust)	PCouv	0,915	0,908	0,901	0,89	0,875
	LMoy	18,928	0,011	11,338	8,211	1,645
Bonett	PCouv	0,93	0,933	0,923	0,917	0,90
	LMoy	19,369	0,012	11,456	8,093	1,554

**n = 70**

DL(ajust)	PCouv	0,915	0,91	0,904	0,898	0,881
	LMoy	17,513	0,010	10,307	7,461	1,488
Bonett	PCouv	0,932	0,932	0,922	0,919	0,906
	LMoy	17,906	0,011	10,464	7,408	1,429

**n = 80**

DL(ajust)	PCouv	0,92	0,916	0,911	0,904	0,89
	LMoy	16,157	0,009	9,604	6,892	1,349
Bonett	PCouv	0,935	0,936	0,929	0,924	0,915
	LMoy	16,537	0,010	9,765	6,882	1,314

**n = 90**

DL(ajust)	PCouv	0,924	0,918	0,911	0,897	0,894
	LMoy	15,25	0,009	9,007	6,323	1,255
Bonett	PCouv	0,938	0,936	0,929	0,918	0,913
	LMoy	15,609	0,009	9,175	6,366	1,23

**n = 100**

DL(ajust)	PCouv	0,926	0,919	0,915	0,908	0,895
	LMoy	14,332	0,008	8,451	6,016	1,171
Bonett	PCouv	0,935	0,936	0,931	0,924	0,916
	LMoy	14,664	0,009	8,625	6,063	1,158

**n = 150**

Loi de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
Asymétrie		0,894	-1,423	1,265	1,633	2
Aplatissement ( $\gamma_e$ )		1,20	2,284	2,40	4,00	6
DL(ajust)	PCouv	0,933	0,925	0,923	0,913	0,911
	LMoy	11,606	0,007	6,781	4,792	0,933
Bonett	PCouv	0,943	0,941	0,936	0,929	0,928
	LMoy	11,846	0,007	6,942	4,875	0,937
<b>n = 200</b>						
DL(ajust)	PCouv	0,935	0,934	0,926	0,916	0,915
	LMoy	9,973	0,006	5,849	4,127	0,799
Bonett	PCouv	0,942	0,948	0,936	0,93	0,931
	LMoy	10,185	0,006	5,991	4,212	0,808
<b>n = 250</b>						
DL(ajust)	PCouv	0,938	0,939	0,934	0,926	0,922
	LMoy	8,899	0,005	5,231	3,652	0,705
Bonett	PCouv	0,946	0,951	0,944	0,936	0,931
	LMoy	9,078	0,005	5,355	3,735	0,716
<b>n = 300</b>						
DL(ajust)	PCouv	0,942	0,938	0,934	0,931	0,922
	LMoy	8,156	0,005	4,749	3,344	0,64
Bonett	PCouv	0,947	0,948	0,943	0,941	0,933
	LMoy	8,314	0,005	4,862	3,419	0,651

Nos résultats corroborent largement ceux publiés par Bonett (2006). Comme indiqué dans les tableaux 1 et 2, les intervalles de confiance calculés à l'aide de la méthode de Bonett sont meilleurs que ceux calculés à l'aide de la méthode DL(ajust) car ils produisent des probabilités de couverture plus proches du niveau cible de 0,95 et des intervalles de confiance plus étroits en moyenne. Si les intervalles de confiance obtenus avec les deux méthodes présentent approximativement les mêmes probabilités de couverture simulée, la méthode qui permet d'obtenir les intervalles les plus courts (en moyenne) est la plus précise. Cela signifie que le test statistique de la variance à l'aide la méthode de Bonett fonctionne mieux et qu'il présente des taux d'erreur de 1ère et 2ème espèce plus faibles. Lorsque les effectifs d'échantillon sont grands, les deux méthodes produisent des résultats quasiment identiques, mais pour les effectifs d'échantillon petits ou modérés, la méthode de Bonett est supérieure.

Bien que la méthode de Bonett fonctionne généralement mieux que la méthode DL(ajust), elle produit invariablement des probabilités de couverture qui se trouvent en dessous de la couverture cible de 0,95 pour les distributions à queues lourdes (symétriques ou asymétriques), même pour de très grands échantillons ( $n > 100$ ). Cela se voit dans la figure 1 ci-dessous, qui représente les probabilités de couverture simulée obtenus avec la méthode de Bonett en fonction de l'excès d'aplatissement de la population pour les effectifs d'échantillon petits, modérés et grands.

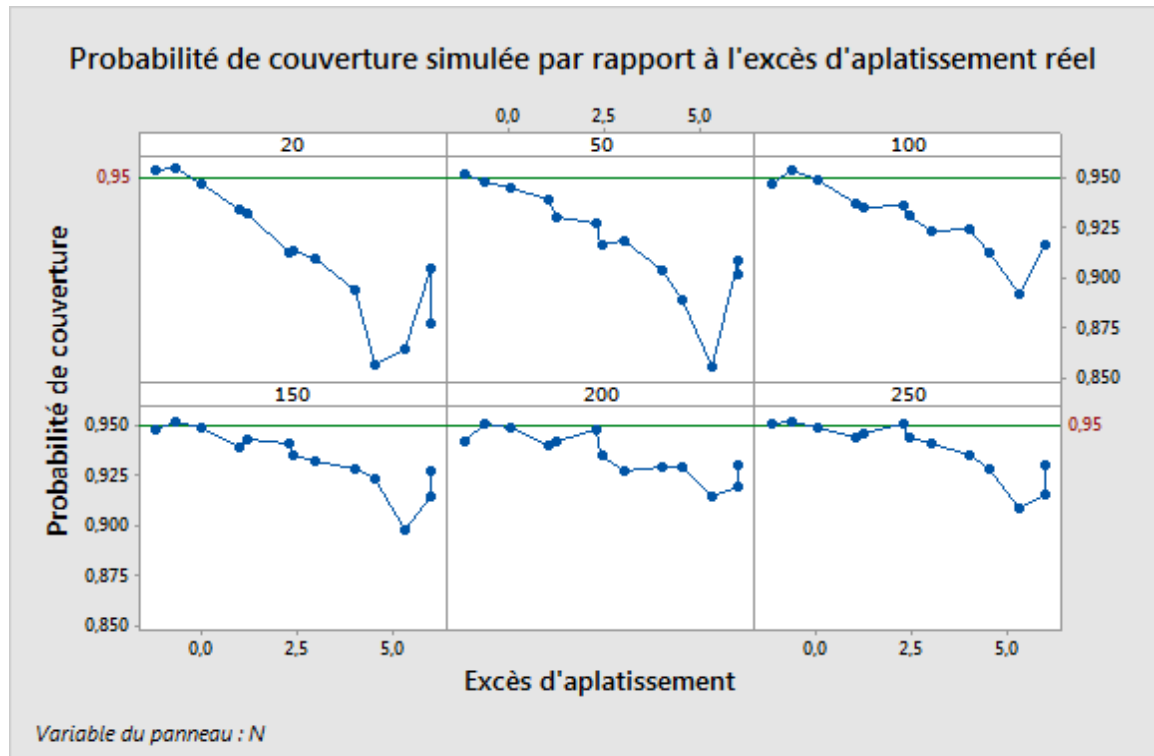


Figure 1 : Probabilité de couverture simulée des intervalles de confiance à 95 % de Bonett en fonction de l'excès d'aplatissement de chaque loi pour différents effectifs d'échantillon.

Comme indiqué à la figure 1, plus l'aplatissement est important, plus l'effectif d'échantillon nécessaire pour que les probabilités de couverture simulée soient proches du niveau cible doit être élevé. Comme nous l'avons remarqué précédemment, les probabilités de couverture simulée obtenues avec la méthode de Bonett sont faibles pour les distributions à queues lourdes. En revanche, pour les lois de distribution à queues légères, comme la loi uniforme et la loi bêta(3,3), ces probabilités de couverture simulée sont stables et sur la cible, même avec des effectifs d'échantillon aussi petits que 20. Par conséquent, notre critère de détermination de la validité de la méthode de Bonett doit tenir compte de l'effectif de l'échantillon et de la lourdeur des queues de la loi utilisée pour créer l'échantillon.

La première étape du développement de ce critère consiste à classer ces lois en trois catégories, selon la lourdeur de leurs queues :

- Lois à queues légères ou normales (type L) : il s'agit des lois pour lesquelles les intervalles de confiance de Bonett produisent des probabilités de couverture stables et proches du niveau de couverture cible. Pour ces lois, les effectifs d'échantillon aussi bas que 20 produisent des résultats exacts. Elles regroupent par exemple la loi

uniforme, la loi bêta(3, 3), la loi normale, la loi t à 10 degrés de liberté et la loi du Khi deux à 10 degrés de liberté.

- Lois à queues modérément lourdes (type M) : pour ces lois, la méthode de Bonett requiert un effectif d'échantillon d'au moins 80 pour que les probabilités de couverture simulée soient proches de la couverture cible. Elles regroupent par exemple la loi du Khi deux à 5 degrés de liberté et la loi bêta(8, 1).
- Lois à queues lourdes (type H) : il s'agit des lois pour lesquelles les intervalles de confiance de Bonett produisent des probabilités de couverture bien en dessous de la couverture cible, à moins que les effectifs d'échantillons soient très grands ( $n \geq 200$ ). Elles regroupent par exemple la loi t à 5 degrés de liberté, la loi de Laplace, la loi du Khi deux à 3 degrés de liberté, la loi exponentielle et les deux lois normales de contamination, CN(0,9, 3) et CN(0,8, 3).

Aussi, pour définir une règle générale d'évaluation de la validité de la méthode de Bonett, il nous faut développer une procédure permettant d'identifier le type de loi auquel obéissent les données échantillons, parmi les trois que nous venons de définir. Nous avons développé cette procédure dans le cadre de la vérification des données pour l'estimation de la validité du test. Pour plus de détails, reportez-vous à l'Annexe E.

# Annexe D : puissance théorique

Nous avons défini la fonction puissance théorique du test associée à la méthode de Bonett et réalisé des simulations afin de comparer la puissance simulée et la puissance théorique du test. Si les courbes de puissance simulée et théorique sont proches l'une de l'autre, cela signifie que l'analyse de puissance et d'effectif d'échantillon réalisée à l'aide de la fonction puissance théorique produit des résultats exacts.

## Formule D1 : fonction puissance théorique de la méthode de Bonett

Comme nous l'avons expliqué précédemment, la méthode de Bonett est fondée sur l'approche classique bien connue qui consiste à utiliser le théorème de la limite centrale et la méthode  $\delta$  de Cramer pour définir une loi de distribution asymptotique de la variance de l'échantillon transformé par logarithme. Plus spécifiquement, il est établi que dans de grands échantillons,  $\frac{\ln S^2 - \ln \sigma^2}{es}$  obéit plus ou moins à la loi normale standard. Le dénominateur,  $es$ , est l'erreur standard de la variance de l'échantillon transformé par logarithme pour les grands échantillons, exprimé de la façon suivante

$$es = \sqrt{\frac{\gamma - (n-3)/n}{n-1}}$$

où  $\gamma$  est l'aplatissement de la population parent inconnue.

Ainsi, une fonction puissance par approximation utilisant un niveau alpha par approximation pour le test bilatéral effectué à l'aide de la méthode de Bonett peut être exprimée comme fonction de l'effectif d'échantillon, du rapport  $\rho = \sigma/\sigma_0$  et de l'aplatissement de la population parent  $\gamma$ , tel que

$$\pi(n, \rho, \gamma) = 1 - \Phi\left(z_{\alpha/2} - \frac{\ln \rho^2}{\sqrt{\frac{\gamma - 1 + 3/n}{n-1}}}\right) + \Phi\left(-z_{\alpha/2} - \frac{\ln \rho^2}{\sqrt{\frac{\gamma - 1 + 3/n}{n-1}}}\right)$$

où  $\sigma_0$  est la valeur hypothétisée de l'écart type inconnu,  $\Phi$  est la fonction de répartition (CDF) de la loi normale standard et  $z_\alpha$  est le point de percentile  $\alpha$  supérieur de la loi normale standard. Les fonctions de puissance unilatérales peuvent également être obtenues à partir de ces calculs.

Notez que lors de la planification de l'effectif d'échantillon dans le cadre d'une étude, vous pouvez utiliser une estimation de l'aplatissement au lieu du véritable aplatissement. Cette estimation est généralement définie à partir de l'opinion d'experts ou des résultats d'expériences précédentes. Si cette information n'est pas disponible, il est conseillé de mener une petite étude pilote afin de développer les plans pour la véritable étude. En utilisant un échantillon de l'étude pilote, l'aplatissement peut être estimé de la façon suivante

$$\hat{\gamma} = \frac{n}{(n-1)^2} \sum_{i=1}^n \left(\frac{x_i - m}{s}\right)^4$$

où  $m$  est une moyenne tronquée avec une proportion de troncage égale à  $1/2\sqrt{n-4}$ .

## Simulation D1 : comparaison de la puissance réelle et de la puissance théorique

Nous avons mis en place une simulation pour comparer les niveaux de puissance réelle estimés (appelés niveaux de puissance simulée) aux niveaux de puissance théorique (appelés niveaux de puissance par approximation) lorsque l'on teste la variance à l'aide de la méthode de Bonett.

Pour chaque expérience, nous avons généré 10 000 répliques d'échantillons, avec chacun un effectif de  $n$ , où  $n = 20, 30, 40, 50, \dots, 120$ , pour chacune des lois décrites dans la simulation C1 (voir Annexe C). Pour chaque loi et chaque effectif d'échantillon  $n$ , nous avons calculé le niveau de puissance simulée, exprimé comme la fraction des 10 000 répliques d'échantillon aléatoire pour laquelle le test bilatéral était significatif au niveau d'alpha  $\alpha = 0,05$ . Lors du calcul de la puissance simulée, nous avons utilisé  $\rho = \sigma/\sigma_0 = 1,25$  pour obtenir des niveaux de puissance relativement petits. Nous avons ensuite calculé les niveaux de puissance correspondants pour la fonction puissance théorique, afin de pouvoir comparer.

Les résultats sont affichés dans les tableaux 3 et 4 et représentés graphiquement dans la figure 2 ci-dessous.

Tableau 3 : Niveaux de puissance simulée (évalués à  $\rho = \sigma/\sigma_0 = 1,25$ ) d'un test de la variance bilatéral utilisant la méthode de Bonett, comparés aux niveaux de puissance théorique (approximation normale). Les échantillons générés obéissent à des lois de distribution symétriques à queues légères, normales, presque normales ou lourdes.

$n$	Puissance	Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0, 1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	t(5)
20	Simul.	0,521	0,39	0,31	0,237	0,178	0,152	0,139	0,172
	Approx.	0,514	0,359	0,264	0,195	0,137	0,117	0,109	0,104
30	Simul.	0,707	0,551	0,441	0,337	0,225	0,186	0,169	0,228
	Approx.	0,717	0,519	0,382	0,276	0,186	0,154	0,143	0,135
40	Simul.	0,831	0,679	0,526	0,427	0,285	0,266	0,203	0,285
	Approx.	0,846	0,651	0,49	0,356	0,236	0,192	0,176	0,165
50	Simul.	0,899	0,753	0,621	0,505	0,332	0,255	0,238	0,34
	Approx.	0,921	0,754	0,586	0,431	0,284	0,229	0,21	0,196
60	Simul.	0,942	0,822	0,701	0,57	0,38	0,285	0,274	0,384
	Approx.	0,961	0,83	0,668	0,501	0,332	0,266	0,243	0,227
70	Simul.	0,964	0,866	0,757	0,632	0,424	0,327	0,314	0,439
	Approx.	0,981	0,885	0,737	0,566	0,379	0,303	0,276	0,257



n	Puissance	Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0, 1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	t(5)
80	Simul.	0,981	0,909	0,815	0,689	0,481	0,372	0,347	0,483
	Approx.	0,991	0,923	0,794	0,624	0,423	0,34	0,309	0,288
90	Simul.	0,988	0,937	0,851	0,724	0,514	0,40	0,377	0,523
	Approx.	0,996	0,95	0,84	0,676	0,467	0,375	0,342	0,318
100	Simul.	0,994	0,961	0,88	0,779	0,558	0,43	0,411	0,566
	Approx.	0,998	0,967	0,876	0,722	0,508	0,41	0,373	0,347
110	Simul.	0,997	0,967	0,909	0,803	0,591	0,471	0,449	0,592
	Approx.	0,999	0,979	0,905	0,763	0,547	0,443	0,404	0,376
120	Simul.	0,999	0,982	0,929	0,844	0,629	0,502	0,476	0,63
	Approx.	1,00	0,987	0,928	0,799	0,584	0,476	0,434	0,405

Tableau 4 : Niveaux de puissance simulée (évalués à  $\rho = \sigma/\sigma_0 = 1,25$ ) d'un test de la variance bilatéral utilisant la méthode de Bonett, comparés aux niveaux de puissance théorique (approximation normale). Les échantillons générés obéissent à des lois de distribution asymétriques à queues presque normales, modérément lourdes ou lourdes.

n	Puissance	Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
20	Simul.	0,222	0,166	0,172	0,139	0,128
	Approx.	0,186	0,152	0,149	0,123	0,104
30	Simul.	0,314	0,216	0,234	0,19	0,151
	Approx.	0,263	0,263	0,205	0,164	0,135
40	Simul.	0,387	0,266	0,292	0,223	0,186
	Approx.	0,338	0,266	0,261	0,204	0,165
50	Simul.	0,455	0,324	0,349	0,263	0,208
	Approx.	0,409	0,323	0,316	0,245	0,196
60	Simul.	0,521	0,376	0,399	0,302	0,239
	Approx.	0,477	0,377	0,369	0,286	0,227
70	Simul.	0,583	0,419	0,463	0,361	0,269
	Approx.	0,539	0,43	0,42	0,325	0,257
80	Simul.	0,646	0,473	0,499	0,394	0,299
	Approx.	0,597	0,479	0,469	0,365	0,288

n	Puissance	Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
90	Simul.	0,688	0,517	0,561	0,428	0,327
	Approx.	0,649	0,526	0,516	0,403	0,318
100	Simul.	0,738	0,561	0,591	0,469	0,368
	Approx.	0,695	0,571	0,56	0,44	0,347
110	Simul.	0,779	0,608	0,637	0,495	0,394
	Approx.	0,737	0,611	0,60	0,475	0,376
120	Simul.	0,81	0,635	0,679	0,538	0,416
	Approx.	0,774	0,65	0,638	0,509	0,405

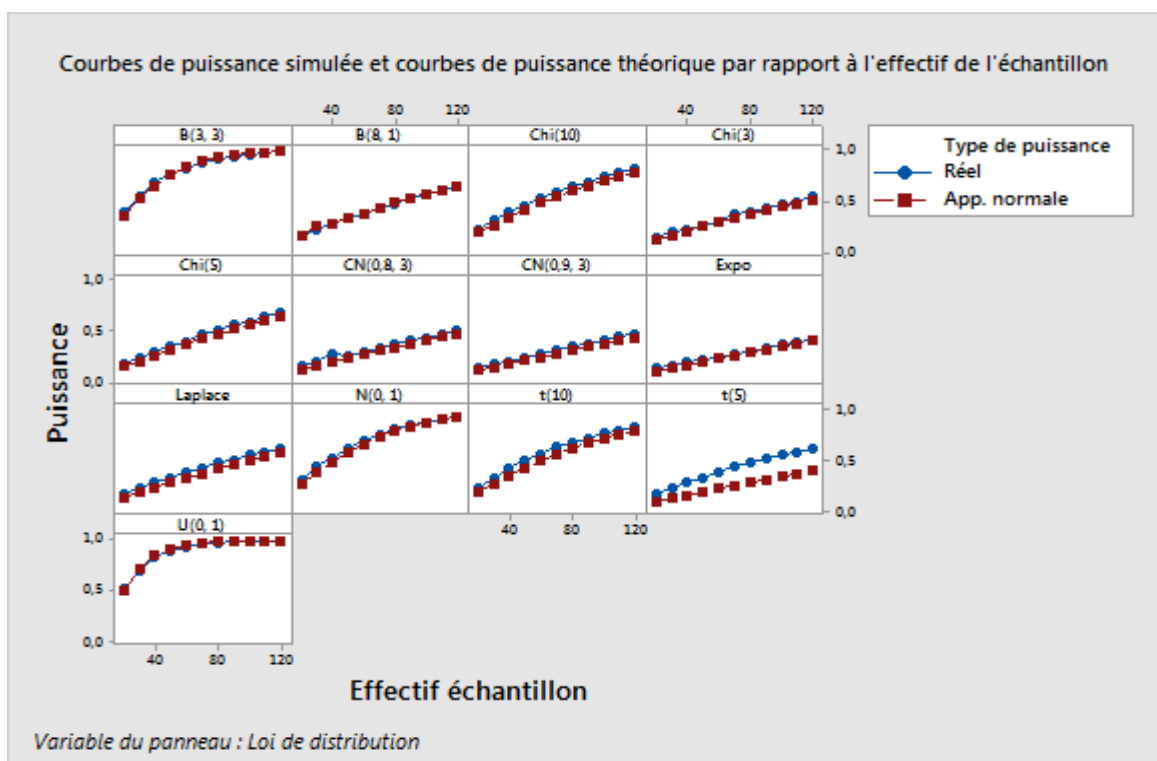


Figure 2 : Courbes de puissance simulée comparées aux courbes de puissance théorique pour plusieurs lois

Les résultats des tableaux 3 et 4 et la figure 2 montrent que lorsque les échantillons sont générés à partir de lois à queues légères (lois de type L, comme définies dans l'Annexe C), comme la loi uniforme, la loi bêta (3,3), la loi normale, la loi t à 10 degrés de liberté et la loi du Khi deux à 10 degrés de liberté, les valeurs de puissance théorique et les niveaux de puissance simulée sont pratiquement indissociables.

En revanche, en ce qui concerne les lois à queues lourdes, (lois de type H), les courbes de puissance simulée sont nettement au-dessus des courbes de puissance théorique lorsque les

échantillons sont petits. Ces lois à queues lourdes sont la loi  $t$  à 5 degrés de liberté, la loi de Laplace, la loi du Khi deux à 3 degrés de liberté, la loi exponentielle et les deux lois normales de contamination,  $CN(0,9;3)$  et  $CN(0,8;3)$ . Ainsi, lors de la planification de l'effectif d'échantillon pour une étude, si l'échantillon provient d'une loi à queues lourdes, l'effectif d'échantillon estimé par la fonction puissance théorique peut être plus grand que l'effectif d'échantillon réellement nécessaire pour atteindre une puissance cible donnée.

# Annexe E : test SJ de détermination de queues normales ou de queues lourdes

Les résultats de l'étude de simulation de l'Annexe C ont démontré que lorsque les queues de la loi sont plus lourdes, il est nécessaire d'utiliser des effectifs d'échantillon plus grands pour obtenir une probabilité de couverture simulée des intervalles de confiance de Bonett proche du niveau cible. L'asymétrie, en revanche, ne semble pas avoir d'effet significatif sur les probabilités de couverture simulée.

Par conséquent, pour évaluer la validité de la méthode de Bonett, il nous fallait développer un critère qui prenne en compte l'effectif de l'échantillon et la lourdeur des queues de la loi utilisée pour créer l'échantillon. Par chance, Get et al. (2007) fournissent un test raisonnablement puissant pour tester directement l'hypothèse nulle selon laquelle la loi a des queues normales, contre l'hypothèse alternative selon laquelle la loi a des queues lourdes. Le test, que nous appelons test SJ, se fonde sur la statistique suivante :

$$\hat{R} = \frac{S}{\hat{j}}$$

où  $S$  est l'écart type de l'échantillon et  $\hat{j}$  l'estimation de l'écart absolu de la moyenne de l'échantillon par rapport à la médiane,  $m$ , calculée comme suit

$$\hat{j} = \frac{\sqrt{\pi/2}}{n} \sum_{i=1}^n |X_i - m|$$

Avec la présence de queues lourdes comme hypothèse alternative, un test par approximation de taille  $\alpha$  rejette l'hypothèse nulle de queues normales si

$$\frac{\sqrt{n}(\hat{R} - 1)}{\sigma_R} \geq z_\alpha$$

où  $z_\alpha$  est le percentile  $\alpha$  supérieur d'une loi normale standard et  $\sigma_R = (\pi - 3)/2$ .

Gel et al. (2007) ont montré que le remplacement du percentile  $\alpha$  supérieur de la loi normale standard par celui de la loi  $t$  à  $(\sqrt{n} + 3)/2$  degrés de liberté fournit de meilleures approximations pour les effectifs d'échantillons modérés. Par conséquent, lors de l'application du test de vérification des données SJ pour la validité du test, nous remplaçons  $z_\alpha$  par  $t_{d,\alpha}$ , le percentile  $\alpha$  supérieur de la loi  $t$  à  $d = (\sqrt{n} + 3)/2$  degrés de liberté.

# Annexe F : validité du test

## Simulation F1 : utilisation de la puissance simulée du test SJ pour déterminer des types de lois

Nous avons réalisé des simulations pour déterminer la puissance du test SJ. Nous avons généré des échantillons de tailles variées ( $n = 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200$ ) avec des lois de distribution différentes. Les lois avaient des queues normales, légères, modérées ou lourdes, et étaient identiques à celles décrites dans la simulation C1 (voir Annexe C). Pour chaque effectif d'échantillon, 10 000 répliques d'échantillon ont été créées pour chaque loi. Nous avons calculé la puissance simulée du test SJ, exprimée comme la proportion de cas pour lesquels l'hypothèse nulle (selon laquelle la loi parent a des queues normales) est rejetée. En outre, nous avons calculé les valeurs moyennes de  $R$  (RMoy) et les valeurs moyennes de  $p$  (PMoy).

Les résultats de la simulation sont indiqués dans les tableaux 5 et 6 ci-dessous.

Tableau 5 : Niveaux de puissance simulée du test SJ. Les échantillons générés obéissent à des lois de distribution symétriques à queues légères, normales, presque normales ou lourdes.

Loi de distribution		Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0,1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	t(5)
<b>n</b>	RVérit	0,921	0,965	1,0	1,032	1,128	1,152	1,118	1,085
10	Puissance	0,021	0,041	0,075	0,103	0,249	0,264	0,198	0,161
	RMoy	1,010	1,036	1,060	1,073	1,129	1,131	1,106	1,096
	PMoy	0,482	0,401	0,341	0,314	0,219	0,228	0,272	0,278
15	Puissance	0,009	0,027	0,071	0,121	0,35	0,389	0,283	0,215
	RMoy	0,986	1,018	1,043	1,063	1,13	1,14	1,11	1,093
	PMoy	0,572	0,44	0,357	0,302	0,171	0,181	0,24	0,247
20	Puissance	0,002	0,016	0,066	0,144	0,428	0,465	0,331	0,253
	RMoy	0,966	1,001	1,030	1,054	1,127	1,137	1,104	1,086
	PMoy	0,669	0,503	0,382	0,311	0,147	0,161	0,236	0,244
25	Puissance	0,002	0,011	0,065	0,153	0,50	0,55	0,397	0,293
	RMoy	0,959	0,995	1,025	1,050	1,128	1,141	1,107	1,086
	PMoy	0,721	0,535	0,391	0,305	0,12	0,128	0,208	0,223
30	Puissance	0,001	0,010	0,060	0,17	0,561	0,603	0,431	0,334
	RMoy	0,951	0,989	1,019	1,046	1,127	1,141	1,106	1,084
	PMoy	0,773	0,57	0,409	0,304	0,103	0,112	0,197	0,209

Loi de distribution		Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0,1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	t(5)
<b>n</b>	RVérit	0,921	0,965	1,0	1,032	1,128	1,152	1,118	1,085
40	Puissance	0,00	0,006	0,058	0,19	0,665	0,709	0,513	0,401
	RMoy	0,944	0,984	1,015	1,043	1,126	1,145	1,109	1,084
	PMoy	0,84	0,616	0,42	0,287	0,073	0,076	0,162	0,179
50	Puissance	0,00	0,004	0,058	0,208	0,746	0,785	0,59	0,462
	RMoy	0,939	0,98	1,012	1,040	1,126	1,146	1,111	1,084
	PMoy	0,886	0,654	0,427	0,279	0,053	0,055	0,131	0,156
60	Puissance	0,00	0,002	0,060	0,231	0,813	0,836	0,647	0,518
	RMoy	0,936	0,978	1,010	1,039	1,127	1,146	1,112	1,084
	PMoy	0,913	0,686	0,43	0,267	0,039	0,039	0,109	0,134
70	Puissance	0,00	0,002	0,054	0,247	0,863	0,879	0,702	0,554
	RMoy	0,934	0,975	1,009	1,037	1,127	1,147	1,112	1,083
	PMoy	0,935	0,716	0,437	0,259	0,028	0,029	0,091	0,123
80	Puissance	0,00	0,001	0,054	0,265	0,896	0,912	0,729	0,591
	RMoy	0,933	0,974	1,007	1,037	1,128	1,147	1,111	1,083
	PMoy	0,95	0,74	0,44	0,241	0,021	0,021	0,079	0,105
90	Puissance	0,00	0,001	0,054	0,281	0,933	0,934	0,771	0,633
	RMoy	0,932	0,973	1,007	1,036	1,128	1,148	1,113	1,083
	PMoy	0,962	0,759	0,445	0,237	0,014	0,016	0,067	0,093
100	Puissance	0,00	0,001	0,057	0,301	0,947	0,954	0,805	0,661
	RMoy	0,93	0,972	1,006	1,036	1,127	1,148	1,113	1,083
	PMoy	0,971	0,779	0,446	0,224	0,012	0,011	0,055	0,083
120	Puissance	0,00	0,00	0,052	0,334	0,974	0,974	0,852	0,732
	RMoy	0,929	0,971	1,005	1,035	1,128	1,149	1,114	1,083
	PMoy	0,982	0,809	0,452	0,206	0,006	0,007	0,041	0,064
140	Puissance	0,00	0,00	0,052	0,336	0,986	0,988	0,894	0,785
	RMoy	0,928	0,971	1,004	1,034	1,127	1,15	1,116	1,084
	PMoy	0,989	0,834	0,454	0,192	0,004	0,003	0,027	0,048
160	Puissance	0,00	0,00	0,054	0,402	0,993	0,992	0,916	0,819
	RMoy	0,927	0,97	1,004	1,034	1,128	1,15	1,114	1,084
	PMoy	0,993	0,858	0,457	0,177	0,002	0,002	0,021	0,040

Loi de distribution		Lois symétriques à queues légères, normales ou presque normales				Lois symétriques à queues lourdes			
		U(0,1)	B(3, 3)	N(0, 1)	t(10)	Lpl	CN(0,8, 3)	CN(0,9, 3)	t(5)
<b>n</b>	RVérit	0,921	0,965	1,0	1,032	1,128	1,152	1,118	1,085
180	Puissance	0,00	0,00	0,052	0,416	0,998	0,996	0,934	0,853
	RMoy	0,926	0,969	1,003	1,034	1,128	1,149	1,115	1,084
	PMoy	0,995	0,874	0,461	0,167	0,001	0,001	0,016	0,033
200	Puissance	0,00	0,00	0,053	0,448	0,998	0,998	0,954	0,884
	RMoy	0,926	0,969	1,003	1,034	1,127	1,15	1,116	1,083
	PMoy	0,997	0,89	0,461	0,153	0,001	0,001	0,011	0,025

Tableau 6 : Niveaux de puissance simulée du test SJ. Les échantillons générés obéissent à des lois de distribution asymétriques à queues presque normales, modérément lourdes ou lourdes.

Lois de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
<b>n</b>	RVérit	1,028	1,075	1,059	1,098	1,151
10	Puissance	0,12	0,213	0,161	0,218	0,283
	RMoy	1,072	1,105	1,088	1,108	1,136
	PMoy	0,326	0,284	0,304	0,279	0,251
15	Puissance	0,139	0,27	0,205	0,292	0,377
	RMoy	1,062	1,105	1,082	1,11	1,141
	PMoy	0,32	0,261	0,286	0,245	0,209
20	Puissance	0,152	0,295	0,223	0,328	0,449
	RMoy	1,051	1,089	1,070	1,101	1,142
	PMoy	0,335	0,26	0,296	0,242	0,186
25	Puissance	0,16	0,336	0,255	0,374	0,515
	RMoy	1,043	1,084	1,068	1,101	1,144
	PMoy	0,337	0,236	0,281	0,219	0,156
30	Puissance	0,171	0,37	0,285	0,414	0,564
	RMoy	1,043	1,084	1,065	1,097	1,142
	PMoy	0,329	0,228	0,274	0,206	0,139

Lois de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
<b>n</b>	RVérit	1,028	1,075	1,059	1,098	1,151
40	Puissance	0,193	0,44	0,331	0,49	0,651
	RMoy	1,039	1,085	1,064	1,098	1,143
	PMoy	0,321	0,188	0,246	0,171	0,106
50	Puissance	0,215	0,484	0,37	0,556	0,72
	RMoy	1,037	1,081	1,064	1,10	1,143
	PMoy	0,314	0,173	0,22	0,14	0,080
60	Puissance	0,224	0,527	0,395	0,607	0,778
	RMoy	1,035	1,079	1,062	1,099	1,146
	PMoy	0,303	0,152	0,208	0,119	0,062
70	Puissance	0,241	0,568	0,438	0,648	0,822
	RMoy	1,034	1,079	1,061	1,098	1,146
	PMoy	0,292	0,134	0,191	0,104	0,048
80	Puissance	0,259	0,612	0,474	0,689	0,855
	RMoy	1,034	1,079	1,062	1,098	1,148
	PMoy	0,28	0,115	0,17	0,089	0,036
90	Puissance	0,284	0,643	0,501	0,733	0,89
	RMoy	1,034	1,079	1,060	1,099	1,148
	PMoy	0,27	0,104	0,163	0,075	0,028
100	Puissance	0,285	0,675	0,527	0,757	0,912
	RMoy	1,032	1,078	1,060	1,098	1,147
	PMoy	0,267	0,094	0,151	0,067	0,022
120	Puissance	0,323	0,728	0,572	0,816	0,942
	RMoy	1,032	1,077	1,060	1,098	1,149
	PMoy	0,246	0,074	0,129	0,050	0,014
140	Puissance	0,344	0,769	0,621	0,852	0,963
	RMoy	1,031	1,077	1,060	1,099	1,148
	PMoy	0,232	0,060	0,112	0,036	0,009
160	Puissance	0,363	0,815	0,666	0,887	0,978
	RMoy	1,031	1,077	1,060	1,098	1,15
	PMoy	0,217	0,047	0,093	0,027	0,005



Lois de distribution		Lois asymétriques à queues presque normales ou modérément lourdes			Lois asymétriques à queues lourdes	
		Khi(10)	B(8, 1)	Khi(5)	Khi(3)	Exp
<b>n</b>	RVérit	1,028	1,075	1,059	1,098	1,151
180	Puissance	0,385	0,843	0,692	0,91	0,986
	RMoy	1,031	1,077	1,059	1,099	1,148
	PMoy	0,209	0,039	0,083	0,021	0,004
200	Puissance	0,41	0,877	0,727	0,931	0,989
	RMoy	1,030	1,077	1,059	1,098	1,149
	PMoy	0,196	0,030	0,071	0,016	0,003

Les résultats de nos simulations présentés dans les tableaux 5 et 6 corroborent ceux publiés par Gel et al. (2007). Lorsque les échantillons sont issus de populations normales, les niveaux de puissance simulée (qui dans ce cas représentent le niveau de signification réel du test) ne sont pas loin du niveau cible, même pour des effectifs d'échantillon aussi bas que 25. Lorsque les échantillons obéissent à des lois à queues lourdes, la puissance du test est faible pour les petits effectifs d'échantillon mais augmente d'au moins 40 % lorsque l'effectif d'échantillon atteint 40. Plus précisément, pour un effectif d'échantillon de 40, la puissance du test est d'environ 40,1 % pour la loi t à 5 degrés de liberté, d'environ 66,5 % pour la loi de Laplace et d'environ 65,1 % pour la loi exponentielle.

Pour les lois à queues légères (loi bêta (3,3) et lois uniformes), la puissance du test est proche de 0 pour les petits échantillons et diminue encore à mesure que l'effectif d'échantillon augmente. Ceci n'est pas surprenant, car le résultat obtenu pour ces lois soutient en réalité l'hypothèse alternative d'une loi à queues légères, et non celle d'une loi à queues lourdes.

Lorsque les échantillons obéissent à des lois à queues légèrement plus lourdes, comme la loi t à 10 degrés de liberté ou la loi du Khi deux à 10 degrés de liberté, les niveaux de puissance sont bas pour les effectifs d'échantillon modérés ou grands. Pour nous, ce résultat est positif, car le test pour une variance (écart type) offre de bons résultats pour ces lois et nous ne voulons pas qu'elles soient signalées comme des lois à queues lourdes. Cependant, lorsque l'effectif d'échantillon augmente, la puissance du test augmente aussi et ces lois à queues légèrement lourdes sont alors signalées comme des lois à queues lourdes.

Par conséquent, les règles d'évaluation du poids des queues de la loi pour ce test doivent aussi prendre en compte l'effectif de l'échantillon. Pour cela, deux méthodes sont envisageables : la première consiste à calculer un intervalle de confiance pour la mesure du poids de la queue. Cependant, la loi de distribution de la statistique SJ est très sensible à la loi de distribution parent de l'échantillon. La seconde méthode consiste à évaluer la lourdeur des queues de la loi à partir de la force du rejet de l'hypothèse nulle par le test SJ et de l'effectif de l'échantillon. Plus spécifiquement, de petites valeurs de p indiquent des queues lourdes et de grandes valeurs de p indiquent des queues légères. Cependant, les grands échantillons ont tendance à avoir des valeurs de p plus petites que les petits échantillons. C'est pourquoi, en prenant en compte les niveaux de puissance simulée, les effectifs d'échantillons et les valeurs de p moyennes du tableau 3, il nous faut concevoir un ensemble

de règles générales afin d'évaluer les queues de la loi pour chaque échantillon à l'aide du test SJ.

Pour les effectifs d'échantillon modérés ou grands ( $40 \leq n \leq 100$ ), si la valeur de  $p$  est comprise entre 0,01 et 0,05, nous considérons que l'hypothèse nulle est modérément contredite. Cela signifie que la loi de l'échantillon est classée comme loi à queue modérément lourde (de type M). En revanche, si la valeur de  $p$  est inférieure à 0,01, l'hypothèse nulle est largement invalidée et la loi parent de l'échantillon est classée comme loi à queues lourdes (de type H).

Pour les grands échantillons ( $n > 100$ ), la loi parent est classée comme une loi de type M si la valeur de  $p$  est comprise entre 0,005 et 0,01, et comme loi de type H si la valeur de  $p$  est très petite (inférieure à 0,005). Notez que lorsque l'effectif d'échantillon est inférieur à 40, la puissance du test SJ est généralement trop faible pour pouvoir évaluer la loi de l'échantillon de façon efficace.

Les règles de classification générales pour déterminer la validité du test de variance 1 lorsque l'on utilise la méthode de Bonett sont récapitulées dans le tableau 7 ci-dessous.

Tableau 7 : Règles de classification pour l'identification de la loi parent de chaque échantillon ( $p$  est la valeur de  $p$  du test SJ)

Condition	Type de loi
$n < 40$	Aucun type déterminé
$100 \geq n \geq 40$ et $p > 0,05$	Loi de type L
$n > 100$ et $p > 0,01$	Loi de type L
$40 \leq n \leq 100$ et $0,01 < p \leq 0,05$	Loi de type M
$n > 100$ et $0,005 < p \leq 0,01$	Loi de type M
$40 \leq n \leq 100$ et $p \leq 0,01$	Loi de type H
$n > 100$ et $p \leq 0,005$	Loi de type H

Comme indiqué précédemment et déterminé à partir des résultats des tableaux 1 et 2 de la simulation C1, l'effectif d'échantillon minimum requis pour atteindre une probabilité de couverture d'au moins 0,93 lorsque les échantillons obéissent à une loi de type L, M ou H, est d'environ 20, 80 et 200 respectivement. Cependant, en raison de la faible puissance du test SJ pour les petits échantillons, l'effectif d'échantillon minimum requis pour les lois de type L est défini sur 40.

## Simulation F2 : vérification des règles des lois de classification

Nous avons généré des échantillons avec certaines des lois de distribution décrites dans la simulation C1 et utilisé le test SJ pour déterminer la proportion d'échantillons classés dans

chacun des trois groupes de lois : type L, type M et type H. Les résultats de la simulation sont indiqués dans le tableau 8.

Tableau 8 : Fraction des 10 000 échantillons identifiée comme obéissant à une loi de type L, M ou H, par effectif et type de loi

n	Lois de distribution	Type L				Type M	Type H	
		B(3, 3)	N(0, 1)	t(10)	Khi(10)	Khi(5)	Lpl	Exp
40	%type L	99,6	94,0	81,5	80,3	66,6	33,0	34,4
	%type M	0,4	5,5	14,0	14,0	20,0	31,9	22,9
	%type H	0,0	0,5	4,5	5,7	13,4	35,1	42,8
50	%type L	99,7	94,4	78,7	79,1	64,0	25,1	28,0
	%type M	0,3	5,1	15,6	14,2	20,0	29,9	20,7
	%type H	0,0	0,5	5,7	6,7	16,0	45,0	51,3
60	%type L	99,7	94,5	77,3	77,3	59,1	18,5	22,6
	%type M	0,3	5,1	16,4	15,0	22,0	27,4	19,2
	%type H	0,0	0,5	6,3	7,7	18,9	54,1	58,2
70	%type L	99,8	94,4	74,5	75,2	55,9	14,0	18,1
	%type M	0,2	5,0	18,1	16,0	22,2	24,0	17,5
	%type H	0,0	0,6	7,4	8,8	21,9	62,0	64,4
80	%type L	99,9	94,3	74,1	74,4	53,0	10,0	13,9
	%type M	0,1	5,1	17,8	16,7	22,8	21,0	15,5
	%type H	0,0	0,6	8,2	8,9	24,2	69,0	70,6
90	%type L	99,9	94,4	71,2	72,1	49,5	7,5	11,1
	%type M	0,1	5,0	19,1	17,2	22,6	16,5	13,7
	%type H	0,0	0,6	9,7	10,7	27,9	76,0	75,3
100	%type L	99,9	94,5	70,8	70,3	47,3	4,8	8,9
	%type M	0,1	4,9	19,5	17,9	22,7	14,3	11,8
	%type H	0,0	0,6	9,7	11,8	30,0	80,9	79,4
120	%type L	100,0	99,4	87,4	87,2	64,8	12,0	14,4
	%type M	0,0	0,4	5,0	4,5	7,9	7,8	5,6
	%type H	0,0	0,2	7,6	8,4	27,4	80,4	80,0
140	%type L	100,0	99,3	86,0	85,1	60,5	7,0	9,9
	%type M	0,0	0,5	5,2	5,0	8,6	5,6	4,1
	%type H	0,0	0,2	8,8	9,9	30,9	87,4	86,0

n	Lois de distribution	Type L				Type M	Type H	
		B(3, 3)	N(0, 1)	t(10)	Khi(10)	Khi(5)	Lpl	Exp
160	%type L	100,0	99,4	83,4	83,0	55,6	4,0	6,9
	%type M	0,0	0,5	6,3	5,8	9,5	3,5	3,0
	%type H	0,0	0,1	10,4	11,2	34,9	92,5	90,1
180	%type L	100,0	99,3	81,1	81,7	51,0	2,5	4,6
	%type M	0,0	0,5	6,8	5,9	9,4	1,9	2,2
	%type H	0,0	0,2	12,1	12,4	39,6	95,6	93,2
200	%type L	100,0	99,5	79,0	80,5	47,2	1,3	3,0
	%type M	0,0	0,4	7,6	6,1	9,4	1,6	1,7
	%type H	0,0	0,1	13,4	13,4	43,4	97,1	95,3

Les résultats du tableau 8 montrent que lorsque les échantillons obéissent à des lois à queues légères (type L) et à queues lourdes (type H), une plus grande proportion d'échantillons est correctement classée. Par exemple, lorsque des échantillons d'effectif 40 sont générés avec la loi bêta(3,3), 99,6 % des échantillons sont correctement classés comme ayant des queues légères ; de même, lorsque des échantillons d'effectif 90 sont générés avec la loi de Laplace, 76,0 % d'entre eux sont correctement classés comme ayant des queues lourdes. Par conséquent, aucun message d'avertissement erroné concernant la validité du test n'apparaît dans le rapport lorsque les échantillons obéissent véritablement à des lois à queues légères ; en revanche, ils sont émis à juste titre lorsque l'échantillon obéit à une loi à queues lourdes et que l'effectif de l'échantillon est inférieur à la valeur nécessaire. En outre, pour les échantillons obéissant à des lois à queues modérément lourdes (type M), telles que la loi du Khi deux (5), la majorité des échantillons sont classés à tort comme obéissant à une loi à queues légères (type L) lorsque les échantillons sont petits (pour un effectif d'échantillon de 40, 66 % des échantillons sont classés à tort comme obéissant à une loi à queues légères). Ainsi, dans ces cas, les messages d'avertissement peuvent ne pas être émis dans le rapport, même si les lois parent ont des queues modérément lourdes. Cela dit, lorsque l'effectif d'échantillon est supérieur à 80, un classement erroné en tant que loi de type L n'a pas d'effet, car l'exigence d'effectif minimum a déjà été remplie.

© 2015, 2017 Minitab Inc. All rights reserved.

Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See [minitab.com/legal/trademarks](http://minitab.com/legal/trademarks) for more information.