

ASSISTANT MINITAB - LIVRE BLANC

Ce livre blanc fait partie d'une série de documents qui expliquent les recherches menées par les statisticiens de Minitab pour développer les méthodes et les outils de vérification des données utilisés dans l'Assistant de Minitab Statistical Software.

Test t à 1 échantillon

Généralités

Le test t à 1 échantillon permet d'estimer la moyenne de votre procédé et de la comparer à une valeur cible. Ce test est considéré comme robuste parce qu'il est particulièrement insensible à l'hypothèse de normalité lorsque l'échantillon est modérément grand. Selon la plupart des manuels de statistiques, le test t à 1 échantillon et l'intervalle de confiance t pour la moyenne sont adéquats pour tout échantillon d'effectif 30 ou plus.

Dans cet article, nous décrivons les simulations que nous avons effectuées pour évaluer la validité de cette règle générale des 30 échantillons. Nos simulations ont porté sur l'influence de la non-normalité sur le test t à 1 échantillon. Nous souhaitons également évaluer l'effet des données aberrantes sur les résultats des tests.

Conformément aux résultats de notre recherche, l'Assistant applique automatiquement les contrôles suivants à vos données, et affiche les résultats dans le rapport :

- Données aberrantes
- Normalité (l'échantillon est-il suffisamment grand pour que vous n'ayez pas à vous soucier de la normalité ?)
- Effectif d'échantillon

Pour obtenir des informations générales sur la méthodologie relative au test t à 1 échantillon, reportez-vous aux études menées par Arnold (1990), Casella et Berger (1990), Moore et McCabe (1993) et Srivastava (1958).

Remarque Les résultats de cette étude sont valables pour le test t pour données appariées de l'Assistant, qui applique la méthode du test t à 1 échantillon à un échantillon de différences pour données appariées.

Vérification des données

Données aberrantes

Les données aberrantes sont des valeurs extrêmement grandes ou extrêmement petites, également connues sous le nom de valeurs aberrantes. Ces données aberrantes peuvent avoir une grande influence sur les résultats de l'analyse. Lorsque l'échantillon est faible, elles peuvent réduire les chances d'obtenir des résultats statistiquement significatifs. Les données aberrantes peuvent venir de problèmes de collecte de données ou être dues à un comportement inhabituel du procédé étudié. 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 la présence de données aberrantes, inspirée de la méthode décrite par Hoaglin, Iglewicz et Tukey (1986), qui permet d'identifier les valeurs aberrantes dans les boîtes à moustaches.

Les résultats

L'Assistant identifie un point de données comme aberrant s'il se trouve à une distance 1,5 fois supérieure à l'étendue interquartile au-delà du quartile inférieur ou supérieur de la distribution. Les quartiles inférieur et supérieur sont les 25^{ème} et 75^{ème} 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, le rapport de l'Assistant sur le test t à 1 échantillon affiche les indicateurs d'état suivants :

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

Normalité

Le test t à 1 échantillon part de l'hypothèse selon laquelle la population est distribuée normalement. Par chance, si les échantillons sont suffisamment grands, cette méthode offre de bons résultats même si les données ne sont pas distribuées normalement.

Objectif

Nous voulions évaluer l'effet de la non-normalité sur les taux d'erreur de 1^{ère} et 2^{ème} espèces du test, afin de fournir aux utilisateurs des conseils concernant l'effectif de l'échantillon et la normalité.

Méthode

Nous avons effectué des simulations pour déterminer l'effectif d'échantillon pour lequel l'hypothèse de normalité peut être ignorée lors de la réalisation d'un test t à 1 échantillon ou du calcul d'un intervalle de confiance t pour la moyenne d'une population.



L'objectif de la première étude était d'évaluer l'effet de la non-normalité sur le taux d'erreur de 1^{ère} espèce du test. Plus spécifiquement, nous voulions déduire l'effectif d'échantillon minimal nécessaire pour que le test soit insensible à la loi de distribution de la population. Nous avons appliqué le test t à 1 échantillon à des échantillons petits, modérés et grands, issus de populations normales et non normales. Les populations non normales incluaient des populations présentant une asymétrie moyenne ou prononcée, des populations symétriques à queues lourdes et légères et des populations normales contaminées. La population normale servait de population de contrôle, à des fins de comparaison. Pour chaque cas, nous avons calculé les seuils de signification simulés, que nous avons comparés au seuil de signification cible (ou nominal) de 0,05. Si le test est adéquat, le seuil de signification simulé doit être proche de 0,05. Nous avons étudié les seuils de signification simulés dans toutes les conditions afin de déterminer l'effectif d'échantillon minimal avec lequel ces seuils demeurent proches de la valeur cible, quelle que soit la loi de distribution. Pour plus de détails, reportez-vous à l'Annexe A.

Dans la deuxième étude, nous avons évalué l'effet de la non-normalité sur le taux d'erreur de 2^{ème} espèce du test. Le plan de la simulation est identique à celui de la première étude. Seulement, cette fois, nous avons comparé les niveaux de puissance simulée dans différentes conditions à des niveaux de puissance cible calculés à l'aide de la fonction puissance théorique du test t à 1 échantillon. Pour plus de détails, reportez-vous à l'Annexe B.

Les résultats

L'effet de la non-normalité sur les taux d'erreur de 1^{ère} et 2^{ème} espèces du test est minime avec des effectifs d'échantillon aussi petits que 20. Toutefois, lorsque la population parent de l'échantillon est très asymétrique, de plus grands échantillons peuvent être requis. Dans ces cas, nous recommandons un effectif d'échantillon d'environ 40. Pour plus de détails, reportez-vous aux annexes A et B.

Etant donné que le test obtient de bons résultats avec des échantillons relativement petits, l'Assistant ne teste pas la normalité des données. Il vérifie l'effectif de l'échantillon et affiche les indicateurs d'état suivants dans le rapport :

Etat	Condition
	L'effectif de l'échantillon est d'au moins 20, la normalité n'est donc pas un problème.
	L'effectif de l'échantillon est inférieur à 20, ce qui peut poser un problème de normalité.

Effectif échantillon

Par définition, un test d'hypothèse vise à collecter des preuves permettant de rejeter l'hypothèse nulle de "non-différence". Lorsque les échantillons sont trop petits, la puissance du test peut ne pas être adaptée pour détecter une différence existante entre les moyennes, ce qui entraîne une erreur de 2^{ème} espèce. Il est donc essentiel de s'assurer que l'effectif de l'échantillon est suffisamment grand 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 l'effectif de l'échantillon est suffisamment grand pour que le test détecte des différences pratiques avec une probabilité élevée. Même si la planification l'effectif d'échantillon vise à garantir que l'effectif de l'échantillon est suffisamment grand pour détecter d'importantes différences avec une probabilité élevée, cet effectif ne doit pas être grand au point que des différences sans importance deviennent statistiquement significatives avec une probabilité élevée.

Méthode






L'analyse de puissance et d'effectif d'échantillon s'appuie sur la fonction puissance théorique propre au test utilisé pour effectuer l'analyse statistique. Comme indiqué précédemment, la fonction puissance du test t à 1 échantillon est insensible à l'hypothèse de normalité lorsque l'effectif d'échantillon est d'au moins 20. La fonction puissance dépend de l'effectif d'échantillon, de la différence entre la moyenne cible et la moyenne de la population, ainsi que de la variance de la population. Pour plus de détails, reportez-vous à l'Annexe B.

Les résultats

Lorsque les données ne fournissent pas suffisamment de preuves invalidant l'hypothèse nulle, l'Assistant calcule les différences pratiques pouvant être détectées avec une probabilité de 80 % et de 90 % pour les effectifs d'échantillon donnés. De plus, si l'utilisateur indique une différence pratique présentant un intérêt particulier, l'Assistant calcule l'effectif d'échantillon qui offre une probabilité de 80 % et de 90 % de détecter cette différence.

Nous ne pouvons pas vous présenter de résultat général, car les résultats dépendent de l'échantillon spécifique de l'utilisateur. Toutefois, vous pouvez vous reporter à l'annexe B pour plus d'informations sur la puissance du test t à 1 échantillon.

Lors du test de puissance et d'effectif d'échantillon, le rapport de l'Assistant sur le test t à 1 échantillon affiche les indicateurs d'état suivants :

Etat	Condition
	<p>Le test détecte une différence entre la moyenne et la cible, par conséquent la puissance n'est pas un problème.</p> <p>OU</p> <p>La puissance est suffisante. Le test n'a pas détecté de différence entre la moyenne et la cible, mais l'échantillon est suffisamment grand pour fournir une probabilité d'au moins 90 % de détecter la différence donnée.</p>
	<p>Il se peut que la puissance soit suffisante. Le test n'a pas détecté de différence entre la moyenne et la 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>Il se peut que la puissance ne soit pas suffisante. Le test n'a pas détecté de différence entre la moyenne et la cible, mais l'échantillon est suffisamment grand pour fournir une probabilité de 80 % à 90 % 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 la moyenne et la 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 la moyenne et la cible. Vous n'avez pas indiqué de différence pratique à détecter entre la moyenne et la cible. Par conséquent, le rapport indique les différences ayant 80 et 90 % de chances d'être détectées avec les effectifs d'échantillon, les écarts types et la valeur alpha utilisés.</p>

Références

- Arnold, S.F. (1990), *Mathematical statistics*, Englewood Cliffs, NJ : Prentice-Hall, Inc.
- Casella, G. et Berger, R. L. (1990), *Statistical inference*, Pacific Grove, CA : Wadsworth, Inc.
- Hoaglin, D. C., Iglewicz, B. et Tukey, J. W. (1986), Performance of some resistant rules for outlier labeling, *Journal of the American Statistical Association*, 81, 991-999.
- Moore, D.S. et McCabe, G.P. (1993), *Introduction to the practice of statistics*, 2ème éd. New York, NY : W. H. Freeman and Company.
- Neyman, J., Iwaskiewicz, K. et Kolodziejczyk, S. (1935), Statistical problems in agricultural experimentation, *Journal of the Royal Statistical Society, Series B*, 2, 107-180.
- Pearson, E.S. et Hartley, H.O. (Eds.) (1954), *Biometrika tables for statisticians*, Vol. I, London : Cambridge University Press.
- Srivastava, A. B. L. (1958), Effect of non-normality on the power function of t-test, *Biometrika*, 45, 421-429.

Annexe A : influence de la non-normalité sur le seuil de signification (validité du test)

Sous l'hypothèse de normalité, le test t à 1 échantillon constitue le test de taille α non biaisé uniformément le plus puissant (UPP). Cela signifie que ce test est au moins aussi puissant que n'importe quel autre test non biaisé de la moyenne de taille α . Toutefois, lorsque la population parent de l'échantillon n'est pas distribuée normalement, les propriétés d'optimalité énoncées ci-dessus sont également valables si l'effectif d'échantillon est suffisamment grand. En d'autres termes, lorsque les échantillons sont suffisamment grands, le seuil de signification réel du test t à 1 échantillon est approximativement égal à la valeur cible pour les données normales comme pour les données non normales, et la fonction puissance du test est également insensible à l'hypothèse de normalité (Srivastava, 1958).

Nous voulions déterminer l'effectif d'échantillon requis pour que le test t soit insensible à l'hypothèse de normalité. Dans de nombreuses publications, il est mentionné que si l'effectif d'échantillon $n \geq 30$, l'hypothèse de normalité peut être ignorée dans des applications plus pratiques (Arnold, 1990 ; Casella et Berger, 1990 ; Moore et McCabe &, 1993). L'objectif de l'étude décrite dans ces annexes est d'effectuer des simulations pour déterminer la validité de cette règle générale en évaluant l'influence de différentes lois non normales sur le test t à 1 échantillon.

Etude par simulation A

Nous voulions évaluer l'influence de la non-normalité sur le taux d'erreur de 1^{ère} espèce du test pour déterminer l'effectif d'échantillon minimal pour lequel ce taux d'erreur est stable et demeure proche de la valeur cible, quelle que soit la loi de distribution.

Pour ce faire, nous avons réalisé des tests t bilatéraux avec $\alpha = 0,05$ sur des échantillons aléatoires de différents effectifs ($n = 10, 15, 20, 25, 30, 35, 40, 50, 60, 80, 100$) et obéissant à plusieurs lois aux propriétés diverses. Ces lois sont les suivantes :

- La loi normale standard ($N(0, 1)$)
- Des lois de distribution symétriques à queues lourdes, à savoir les lois T à 5 et 10 degrés de liberté ($T(5)$, $T(10)$)
- La loi de Laplace avec un emplacement 0 et une échelle 1 (Lpl)
- Des lois de distribution asymétriques à queues lourdes, représentées par la loi exponentielle avec une échelle 1 (Exp) et des lois du Khi deux à 3, 5 et 10 degrés de liberté ($Khi(3)$, $Khi(5)$, $Khi(10)$)

- Des lois de distribution symétriques à queues plus légères, telles que la loi de distribution uniforme (U(0, 1)) et la loi bêta, avec les deux paramètres définis sur 3 (B(3, 3))
- Une loi de distribution à queues lourdes présentant une asymétrie vers la gauche (B(8, 1))

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

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

où p est défini comme le paramètre de mélange et $1 - p$ la proportion de contamination (ou la proportion de valeurs aberrantes). Nous avons sélectionné deux populations normales contaminées pour l'étude : CN(0,9, 3) (10 % de la population est constituée de valeurs aberrantes) et CN(0,8, 3) (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 effectif d'échantillon, 10 000 répliques d'échantillons ont été créées pour chaque loi et un test t à 1 échantillon avec l'hypothèse nulle $\mu = \mu_o$ et l'hypothèse alternative $\mu \neq \mu_o$ a été réalisé pour chacun des 10 000 échantillons. Pour chaque test, nous avons défini la moyenne hypothétisée μ_o sur la moyenne réelle de la population parent de l'échantillon. Par conséquent, pour un effectif d'échantillon donné, la fraction des 10 000 répliques d'échantillons qui produisent un rejet de l'hypothèse nulle représente le taux d'erreur de 1ère espèce simulé ou le seuil de signification du test. Le seuil de signification cible étant de 5 %, l'erreur de simulation est d'environ 0,2 %.

Les résultats de la simulation sont présentés dans les tableaux 1 et 2 et représentés graphiquement dans les figures 1 et 2.

Résultats et récapitulatif

Les résultats (voir le tableau 1 et la figure 1) indiquent que lorsque les échantillons sont issus de populations symétriques, les seuils de signification simulés du test sont proches du seuil de signification cible, même pour de petits échantillons. Toutefois, les résultats sont légèrement prudents pour les lois symétriques à queues lourdes lorsque les échantillons sont petits, y compris dans le cas de petits échantillons générés à partir de distributions normales contaminées. En outre, nous pouvons remarquer que les valeurs aberrantes réduisent le seuil de signification du test lorsque les échantillons sont petits. Toutefois, cet effet est inversé lorsque de petits échantillons sont générés à partir de populations parent symétriques à queues légères (loi bêta (3, 3) et lois uniformes). Dans ces cas, les seuils de signification simulés sont légèrement plus élevés.

Tableau 1 Seuils de signification simulés pour un test t à 1 échantillon bilatéral pour des échantillons issus de populations symétriques. Le seuil de signification cible est de $\alpha = 0,05$.

Loi	N(0,1)	T(5)	T(10)	Lpl	CN(0,9,3)	CN(0,8,3)	B(3, 3)	U(0, 1)
N	Symétriques à queues lourdes						Symétriques à queues plus légères	
10	0,050	0,046	0,048	0,044	0,043	0,039	0,057	0,057
15	0,051	0,050	0,049	0,049	0,043	0,043	0,053	0,054
20	0,047	0,051	0,051	0,047	0,043	0,044	0,051	0,052
25	0,050	0,047	0,050	0,046	0,046	0,046	0,048	0,050
30	0,053	0,050	0,048	0,043	0,049	0,046	0,050	0,048
35	0,052	0,047	0,049	0,050	0,047	0,045	0,051	0,054
40	0,046	0,052	0,054	0,048	0,046	0,049	0,044	0,050
50	0,050	0,049	0,051	0,048	0,047	0,051	0,053	0,050
60	0,052	0,049	0,053	0,050	0,051	0,056	0,054	0,052
80	0,049	0,050	0,051	0,047	0,047	0,052	0,049	0,049
100	0,050	0,052	0,049	0,051	0,052	0,054	0,051	0,054

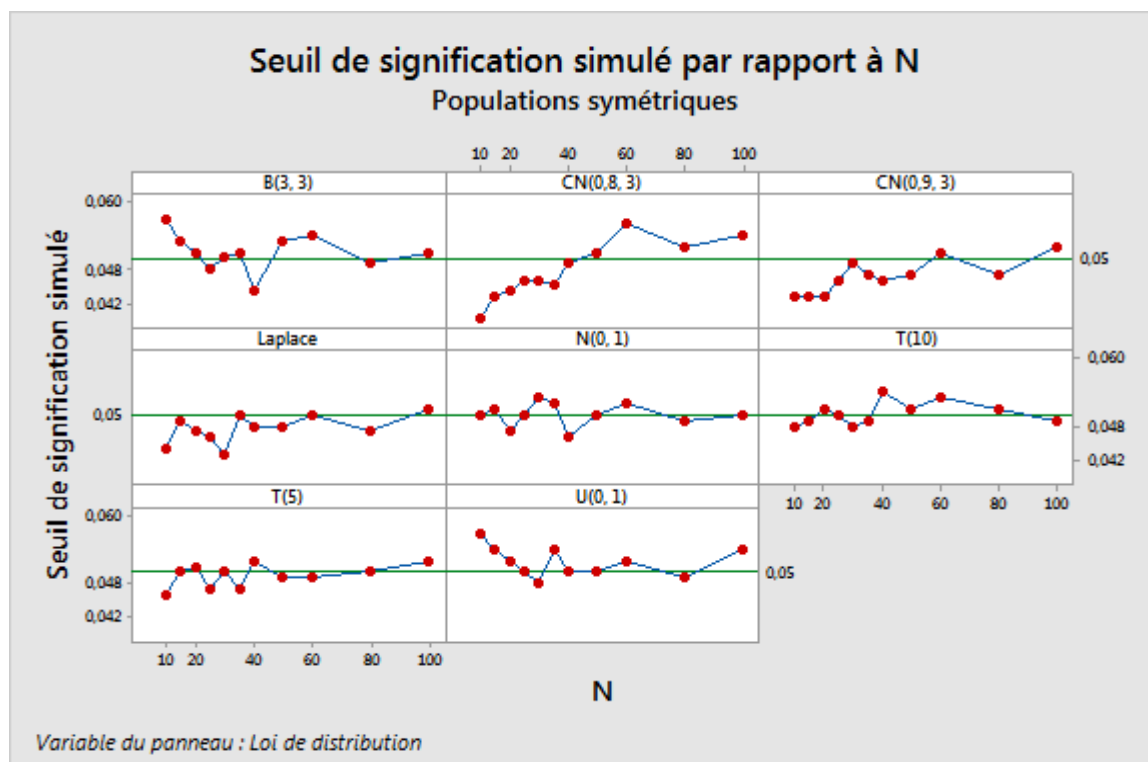


Figure 1 Graphique des seuils de signification simulés pour des tests t à 1 échantillon bilatéraux en fonction des effectifs d'échantillon issus de populations symétriques. Le seuil de signification cible est de $\alpha = 0,05$.

Lorsque les échantillons générés obéissent à des lois de distribution asymétriques, les performances du test dépendent du degré de l'asymétrie. Les résultats du tableau 2 et de la figure 2 indiquent que le test t à 1 échantillon est sensible à l'asymétrie dans de petits échantillons. Pour les populations hautement asymétriques (exponentielle, Khi(3) et bêta(8,1)), de plus grands échantillons sont requis pour que les seuils de signification simulés soient proches du seuil de signification cible. Toutefois, pour les populations modérément asymétriques (Khi(5) et khi(10)), un effectif d'échantillon d'au moins 20 est suffisant pour que les seuils de signification simulés soient proches du seuil cible. Avec un effectif d'échantillon d'au moins 20, le seuil de signification simulé est d'environ 0,063 pour la loi du Khi deux à 5 degrés de liberté et d'environ 0,056 pour la loi du Khi deux à 10 degrés de liberté.

Tableau 2 Seuils de signification simulés d'un test t à 1 échantillon bilatéral pour des échantillons issus de populations symétriques. Le seuil de signification cible est de $\alpha = 0,05$.

N	Exp	Khi(3)	B(8, 1)	Khi(5)	Khi(10)
	Asymétrie des populations				
	2,0	1,633	-1,423	1,265	0,894
Seuils de signification simulés					
10	0,101	0,089	0,087	0,069	0,060
15	0,088	0,076	0,072	0,068	0,057

N	Exp	Khi(3)	B(8, 1)	Khi(5)	Khi(10)
	Asymétrie des populations				
	2,0	1,633	-1,423	1,265	0,894
	Seuils de signification simulés				
20	0,083	0,073	0,069	0,063	0,056
25	0,075	0,068	0,067	0,067	0,056
30	0,069	0,067	0,066	0,058	0,054
35	0,075	0,067	0,062	0,062	0,056
40	0,067	0,067	0,061	0,059	0,056
50	0,064	0,057	0,062	0,057	0,054
60	0,063	0,056	0,061	0,054	0,055
80	0,059	0,058	0,053	0,052	0,052
100	0,060	0,055	0,055	0,047	0,053

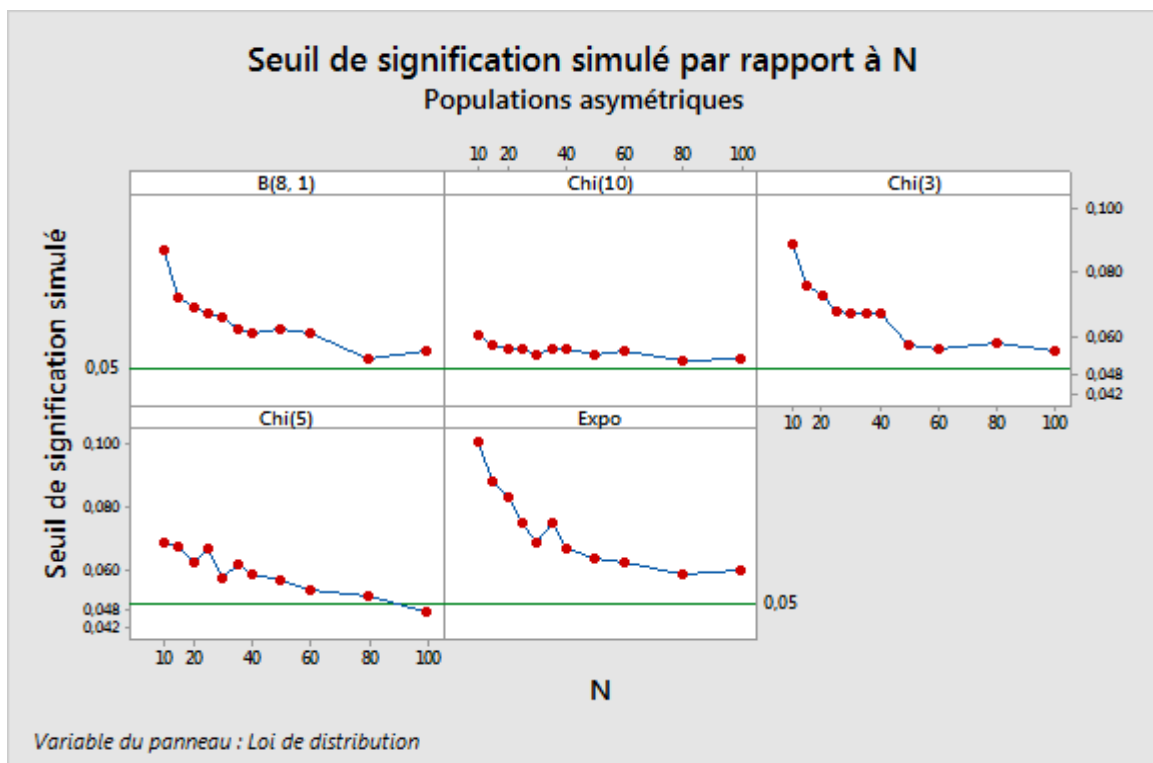


Figure 2 Graphique des seuils de signification simulés pour des tests t à 1 échantillon bilatéraux en fonction des effectifs d'échantillon issus de populations asymétriques. Le seuil de signification cible est de $\alpha = 0,05$.

Dans cette étude, nous avons examiné les tests d'hypothèse plutôt que les intervalles de confiance. Toutefois, ces résultats s'étendent naturellement aux intervalles de confiance, car les tests d'hypothèse et les intervalles de confiance peuvent tous deux être utilisés pour déterminer une signification statistique.

Annexe B : effectif d'échantillon et puissance du test

Nous voulions évaluer la sensibilité de la fonction puissance à l'hypothèse de normalité sous laquelle elle est dérivée. Notez que si β est l'erreur de 2ème espèce d'un test, $1 - \beta$ est la puissance du test. Par conséquent, l'effectif d'échantillon prévu est choisi de telle sorte que le taux d'erreur de 2ème espèce soit faible ou, autrement dit, pour que le niveau de puissance soit raisonnablement élevé.

Les fonctions puissance des tests t sont connues et ont fait l'objet de plusieurs études. Pearson et Hartley (1954), ainsi que Neyman, Iwazkiewicz et Kolodziejczyk (1935) fournissent des graphiques et des tableaux de ces fonctions puissance.

Pour un test t à 1 échantillon bilatéral de taille α , cette fonction peut être exprimée mathématiquement en fonction de l'effectif d'échantillon et de la différence δ entre la moyenne réelle μ et la moyenne hypothétisée μ_0 tel que

$$\pi(n, \delta) = 1 - F_{n-1, \lambda}(t_{n-1}^{\alpha/2}) + F_{n-1, \lambda}(-t_{n-1}^{\alpha/2})$$

où $F_{d, \lambda}(\cdot)$ est la fonction de distribution cumulative (DCF) de la loi T à $d = n - 1$ degrés de liberté avec le paramètre de non-centralité

$$\lambda = \frac{\delta\sqrt{n}}{\sigma}$$

et où t_d^α désigne le point de percentile supérieur 100α de la loi T à d degrés de liberté.

Pour les hypothèses alternatives unilatérales, la puissance est exprimée sous la forme

$$\pi(n, \delta) = 1 - F_{n-1, \lambda}(t_{n-1}^\alpha)$$

lorsque l'hypothèse alternative est $\mu > \mu_0$, et sous la forme

$$\pi(n, \delta) = F_{n-1, \lambda}(-t_{n-1}^\alpha)$$

lorsque l'hypothèse alternative est $\mu < \mu_0$.

Ces fonctions supposent que les données sont distribuées normalement et que le seuil de signification du test est défini sur une certaine valeur α .

Etude par simulation B

L'objectif de cette simulation est d'évaluer l'effet de la non-normalité sur la fonction puissance théorique du test t à 1 échantillon. Pour évaluer l'effet de la non-normalité, nous avons comparé les niveaux de puissance simulée aux niveaux de puissance cible calculés à l'aide de la fonction puissance théorique du test.

Nous avons effectué des tests t bilatéraux avec $\alpha = 0,05$ sur des échantillons aléatoires de différents effectifs ($n = 10, 15, 20, 25, 30, 35, 40, 50, 60, 80, 100$) générés à partir des populations décrites dans la première étude de simulation (voir l'annexe A).

Pour chacune des populations, l'hypothèse nulle du test est $\mu = \mu_0 - \delta$ et son hypothèse alternative est $\mu \neq \mu_0 - \delta$, où μ_0 est défini sur la moyenne réelle de la population et $\delta = \sigma/2$ (σ est l'écart type de la population parent). Ainsi, la différence entre la moyenne réelle et la moyenne hypothétisée est 0. L'hypothèse nulle devrait donc être rejetée.

Pour chaque effectif d'échantillon, 10 000 répliques d'échantillon sont créées pour chaque loi. Pour chaque effectif d'échantillon donné, la fraction des 10 000 répliques pour lesquelles l'hypothèse nulle est rejetée représente le niveau de puissance simulée du test avec un effectif d'échantillon et une différence δ donnés. Notez que nous avons choisi cette valeur de différence particulière parce qu'elle produit des valeurs de puissance relativement faibles lorsque les échantillons sont petits.

En outre, les valeurs de puissance théorique correspondantes, appelées valeurs de puissance cible, sont calculées avec la différence δ et avec différents effectifs d'échantillons afin de pouvoir être comparées aux valeurs de puissance simulée.

Les résultats de la simulation sont présentés dans les tableaux 3 et 4 et représentés graphiquement dans les figures 3 et 4.

Résultats et récapitulatif

Les résultats confirment que la puissance du test t à 1 échantillon est généralement insensible à l'hypothèse de normalité lorsque l'effectif d'échantillon est suffisamment grand. Pour les échantillons issus de populations symétriques, les résultats du tableau 3 indiquent que les niveaux de puissance cible et simulée sont proches, même pour de petits échantillons. Les courbes de puissance correspondantes affichées dans la figure 3 sont pratiquement indissociables. Pour les échantillons générés à partir de distributions normales contaminées, les valeurs de puissance sont quelque peu prudentes pour les effectifs d'échantillons petits ou modérés. Cela peut être dû au fait que le seuil de signification réel du test pour ces populations est légèrement plus élevé que le seuil de signification cible α défini.

Tableau 3 Niveaux de puissance simulée avec une différence $\delta = \sigma/2$ pour un test t à 1 échantillon bilatéral de taille $\alpha = 0,05$ lorsque les échantillons sont issus de populations symétriques. Les niveaux de puissance simulée sont comparés aux niveaux de puissance théorique cible obtenus sous l'hypothèse de normalité.

n	Puissance cible	N(0, 1)	T(5)	T(10)	Lpl	CN(0,9, 3)	CN(0,8, 3)	B(3, 3)	U(0, 1)
		Niveau de puissance simulée pour $\delta = \sigma/2$ (populations symétriques)							
10	0,293	0,299	0,334	0,311	0,357	0,361	0,385	0,28	0,269
15	0,438	0,438	0,48	0,45	0,491	0,512	0,511	0,423	0,421
20	0,565	0,57	0,603	0,578	0,60	0,629	0,623	0,557	0,548
25	0,67	0,674	0,695	0,68	0,691	0,712	0,70	0,665	0,67
30	0,754	0,756	0,77	0,756	0,767	0,768	0,765	0,754	0,75
35	0,82	0,819	0,827	0,815	0,82	0,819	0,812	0,822	0,818
40	0,869	0,87	0,871	0,868	0,862	0,869	0,868	0,875	0,867
50	0,934	0,933	0,929	0,93	0,929	0,923	0,925	0,932	0,94
60	0,968	0,967	0,963	0,965	0,964	0,955	0,955	0,968	0,971
80	0,993	0,993	0,989	0,992	0,991	0,988	0,989	0,994	0,994
100	0,999	0,998	0,996	0,998	0,999	0,998	0,996	0,999	0,999

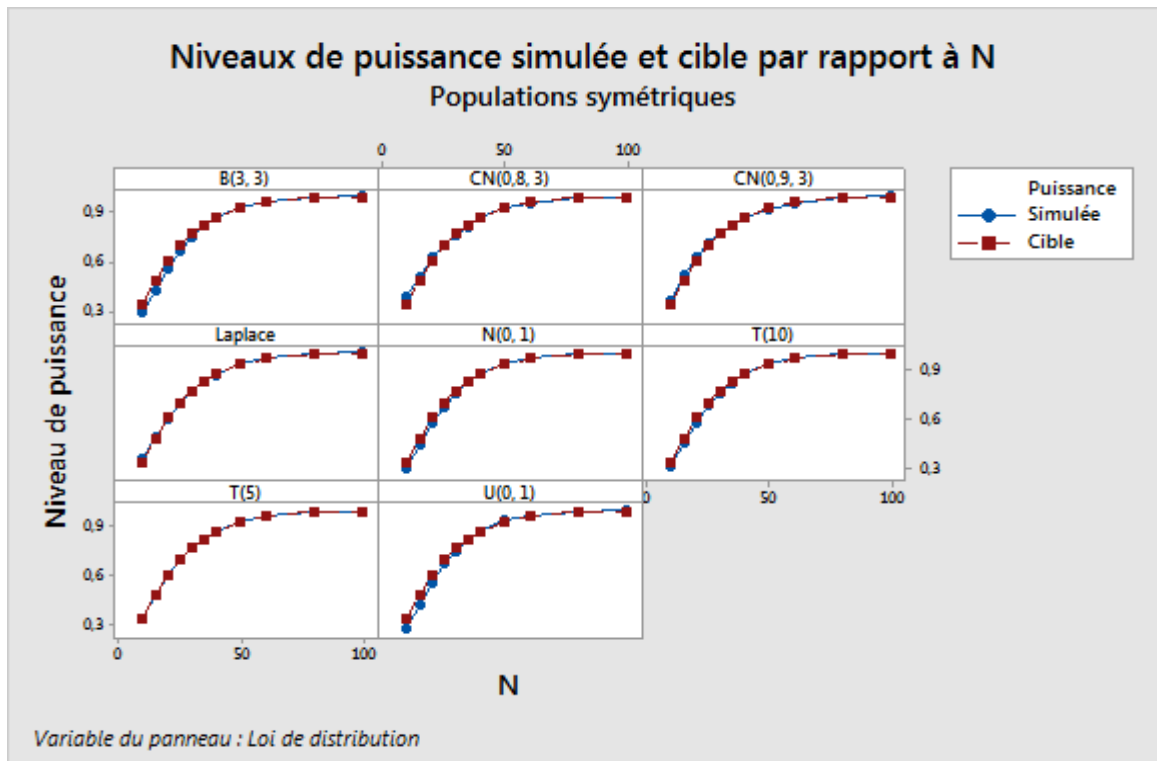


Figure 3 Courbes de puissance simulée comparées aux courbes de puissance théorique cible avec $\alpha = 0,05$ pour un test t à 1 échantillon bilatéral lorsque les échantillons sont issus de populations symétriques. Les valeurs de puissance sont évaluées avec une différence de $\delta = \sigma/2$.

Toutefois, lorsque les échantillons sont issus de populations asymétriques, les valeurs de puissance simulée sont hors cible avec de petits échantillons, comme l'indiquent le tableau 4 et la figure 4. Pour les populations obéissant à des lois modérément asymétriques, telles que les lois du Khi deux à 5 et 10 degrés de liberté, lorsque l'effectif d'échantillon est d'au moins 20, les niveaux de puissance simulée et cible sont proches. Par exemple, pour $n = 20$, le niveau de puissance cible est de 0,565 lorsque les niveaux de puissance simulée sont de 0,576 et 0,577 pour les lois du Khi deux (5) et du Khi deux (10), respectivement. Pour les lois très asymétriques, de plus grands échantillons sont requis pour que les niveaux de puissance simulée soient proches du seuil de signification cible. Cela peut être dû au fait que le test t à 1 échantillon ne contrôle pas le taux d'erreur de 1ère espèce de façon appropriée lorsque les échantillons sont petits et que les populations parent sont très asymétriques.

Tableau 4 Valeurs de puissance simulée avec une différence $\delta = \sigma/2$ pour un test t à 1 échantillon bilatéral de taille $\alpha = 0,05$ lorsque les échantillons sont issus de populations asymétriques. Les valeurs de puissance simulée sont comparées aux valeurs de puissance cible obtenues sous l'hypothèse de normalité.

N	Puissance cible	Exp		Khi(3)	B(8, 1)	Khi(5)	Khi(10)
		Asymétrie des populations					
		2,0		1,633	-1,423	1,265	0,894
		Niveaux de puissance simulée					
10	0,293	0,206		0,212	0,39	0,225	0,238
15	0,438	0,416		0,413	0,484	0,409	0,407
20	0,565	0,604		0,591	0,566	0,576	0,577
25	0,67	0,763		0,734	0,657	0,709	0,695
30	0,754	0,859		0,834	0,729	0,808	0,785
35	0,82	0,917		0,895	0,776	0,874	0,835
40	0,869	0,955		0,935	0,823	0,925	0,905
50	0,934	0,987		0,981	0,90	0,973	0,96
60	0,968	0,997		0,994	0,937	0,991	0,985
80	0,993	1,00		0,999	0,98	0,999	0,997
100	0,999	1,00		1,00	0,994	1,00	1,00

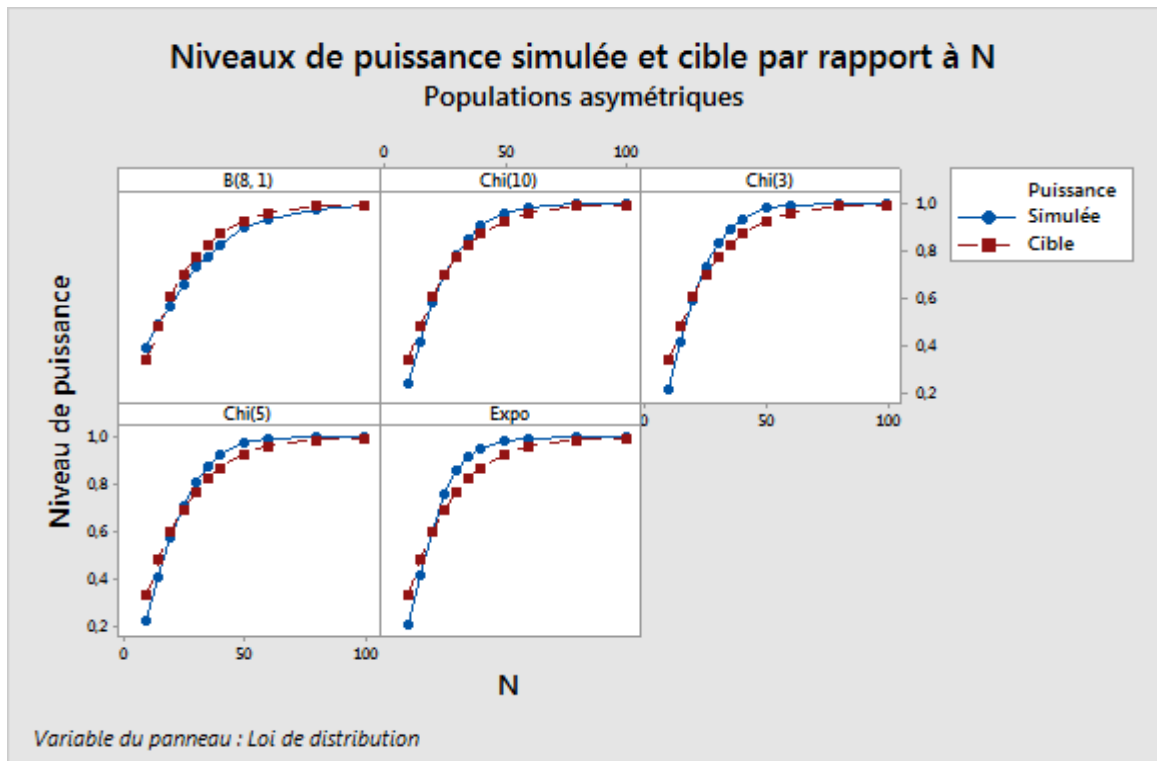


Figure 4 Courbes de puissance simulée comparées aux courbes de puissance théorique cible avec $\alpha = 0,05$ pour un test t à 1 échantillon bilatéral lorsque les échantillons sont issus de populations symétriques. Les valeurs de puissance sont évaluées avec une différence de $\delta = \sigma/2$.

En résumé, pour les lois modérément asymétriques, la fonction puissance est fiable lorsque l'effectif d'échantillon est d'au moins 20, quelle que soit la population parent à partir de laquelle l'échantillon est créé. Pour les populations très asymétriques, un effectif d'échantillon plus grand (environ 40) est requis pour obtenir une puissance simulée proche de la puissance cible.