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.

Tests du Khi deux

Généralités

En pratique, les qualiticiens ont parfois besoin de collecter des données de catégories pour évaluer un procédé lorsqu'il est impossible ou peu pratique pour eux de collecter des données continues. Par exemple, il est possible de classer un produit dans deux catégories (défectueux/non défectueux) ou dans plus de deux catégories (excellent, bon, passable et mauvais). Nous pourrions également prendre l'exemple d'un service financier classant ses factures arriérées en fonction du nombre de jours : 15 jours ou moins, 16 à 30 jours, 31 à 45 jours, ou 45 jours ou plus. Ainsi, la variable à tester correspond au nombre d'éléments dans chaque catégorie.

Grâce à leur polyvalence, les tests du Khi deux conviennent à de nombreuses applications impliquant des données de catégories. Dans l'Assistant, nous utilisons les tests du Khi deux pour :

- Tester l'adéquation de l'ajustement d'une loi multinomiale

  Vous pouvez utiliser ce test pour déterminer si les données suivent la même distribution qu'initialement. La distribution est définie comme une loi multinomiale avec un ensemble de pourcentages historiques, ou cible, qui définissent le pourcentage d'éléments dans chaque catégorie de résultats. Le test du Khi deux teste conjointement si un pourcentage diffère de manière significative de ses pourcentages cible ou historiques respectifs.

- Tester l'égalité des pourcentages de défectueux dans plus de 2 groupes

  Vous pouvez utiliser ce test pour déterminer s'il existe une différence entre les pourcentages de défectueux de différents groupes. Les groupes diffèrent au niveau d'une caractéristique à tester, notamment un produit fabriqué par des opérateurs différents, par des usines différentes ou à différents moments. Le test du Khi deux teste conjointement si un pourcentage de défectueux diffère de manière significative d'un autre pourcentage de défectueux.

- Tester l'association entre deux variables de catégories

Vous pouvez utiliser ce test pour déterminer si une variable de résultat de catégorie (Y) est liée ou associée à une autre variable de prédiction de catégorie (X). Le test du Khi deux teste conjointement s'il existe une association entre la variable de résultat et une variable de prédiction. Dans l'Assistant, vous pouvez effectuer un test d'association du Khi deux avec une variable de prédiction (X) qui comprend au moins deux valeurs distinctes (au moins deux échantillons).

Pour plus d'informations sur le test du Khi deux, reportez-vous à l'Annexe A.

Pour ce qui est des méthodes impliquant des tests d'hypothèse, il est généralement conseillé de veiller à ce que les hypothèses du test soient satisfaites, à ce que la puissance du test soit adaptée et à ce que toutes les approximations utilisées pour analyser les données produisent des résultats valides. Pour ce qui est des tests du Khi deux, les hypothèses sont inhérentes à la collecte de données et nous ne les abordons pas dans le contrôle des données.

Nous nous concentrons sur la puissance et la validité des méthodes d'approximation. L'Assistant utilise ces méthodes d'approximation pour effectuer les contrôles suivants sur vos données et présente les résultats dans le rapport :

- Effectif d'échantillon
- Validité du test
- Validité des intervalles

Dans ce document, nous étudions la relation entre ces contrôles de données et les tests du Khi deux en pratique, et nous décrivons la façon dont nous avons établi les indications relatives aux contrôles de données dans l'Assistant.

Contrôles de données

Effectif d'échantillon

En général, le principal objectif d'un test statistique d'hypothèse est de collecter des données 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 pourcentages de défectueux, ce qui entraîne une erreur de 2ème espèce. Il est donc essentiel de s'assurer que les effectifs d'échantillon sont suffisamment grands pour détecter des différences importantes en pratique avec une probabilité élevée.

Le test d'effectif d'échantillon dépend de la puissance du test. Ce calcul nécessite que l'utilisateur spécifie une différence significative entre un paramètre de population réel et la valeur nulle hypothétisée. Etant donné qu'il s'est avéré trop difficile de déterminer et d'exprimer cette différence pratique pour le test d'ajustement du Khi deux et les tests d'association du Khi deux, l'Assistant contrôle uniquement l'effectif d'échantillon du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons.

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 des effectifs d'échantillons vise à garantir que les échantillons sont suffisamment grands pour détecter d'importantes différences avec une probabilité élevée, ces échantillons ne doivent pas être grands au point que des différences sans importance deviennent statistiquement significatives avec une probabilité élevée.

Méthode

L'analyse de la puissance et de l'effectif d'échantillon est basée sur les formules présentées à l'Annexe B.

Les résultats

Lorsque les données ne permettent pas de réfuter l'hypothèse nulle et que vous ne spécifiez pas de différence pratique, l'Assistant calcule les différences pratiques susceptibles d'être détectées avec une probabilité de 80 % et 90 % en fonction des effectifs d'échantillons. En outre, si l'utilisateur fournit une différence pratique à tester particulière, l'Assistant calcule les effectifs d'échantillons qui produisent respectivement 80 et 90 % de chances de détecter cette différence.

Lors de la vérification de la puissance et de l'effectif d'échantillon, le rapport de l'Assistant du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons présente les indicateurs d'état suivants :

Etat	Condition
	Le test détecte une différence entre les pourcentages de défectueux, la puissance n'est donc pas un problème. OU La puissance est suffisante. Le test n'a pas détecté de différence entre les pourcentages de défectueux, mais l'échantillon est suffisamment grand pour fournir une probabilité d'au moins 90 % de détecter la différence donnée
	Il se peut que la puissance soit suffisante. Le test n'a pas détecté de différence entre les pourcentages de défectueux, 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.
	Il se peut que la puissance ne soit pas suffisante. Le test n'a pas détecté de différence entre les pourcentages de défectueux, 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.
	La puissance n'est pas suffisante (< 60 %). Le test n'a détecté aucune différence entre les pourcentages de défectueux. Les effectifs d'échantillon nécessaires pour atteindre une puissance de 80 % et de 90 % sont indiqués.
	Le test n'a pas détecté de différence entre les pourcentages de défectueux. Vous n'avez pas précisé de différence pratique à détecter entre les pourcentages de défectueux ; par conséquent, le rapport présente les différences que vous pouvez détecter avec 80 et 90 % de chances, en fonction de vos effectifs d'échantillons et du niveau d'alpha.

Validité du test

La statistique du test du χ^2 ne suit qu'approximativement une loi du Khi deux. L'approximation s'améliore au fur et à mesure de l'accroissement des effectifs d'échantillons. Dans cette section, nous évaluons l'approximation utilisée pour déterminer l'effectif d'échantillon minimum nécessaire pour obtenir des résultats précis.

L'approximation du Khi deux par rapport aux statistiques du test est évaluée via l'examen de l'impact de petits dénombrements attendus pour les cellules sur le taux d'erreur de 1ère espèce (alpha). En utilisant l'erreur de 1ère espèce pour évaluer la validité du test, nous avons mis au point une règle afin de nous assurer que :

- La probabilité de rejeter à tort l'hypothèse nulle est faible et proche du taux d'erreur de 1ère espèce souhaité.
- Il est possible d'obtenir une approximation raisonnable de la queue de la loi de distribution nulle, ce qui est important pour le calcul exact de la valeur de p.

A l'aide d'une approche standard, nous avons défini un petit dénombrement attendu pour les cellules : un dénombrement inférieur ou égal à 5.

Nous avons élaboré deux modèles pour définir les proportions dans le cadre de l'hypothèse nulle : le modèle à proportions perturbées et le modèle à proportions égales. Pour plus de détails, reportez-vous à l'Annexe C. Les deux modèles sont utilisés dans les simulations présentées ultérieurement dans ce document. Les modèles sont utilisés pour chacun des tests du Khi deux à une exception : le modèle à proportions perturbées ne s'applique pas au test du Khi deux du pourcentage de défectueux en présence de plus de deux échantillons.

La validité du contrôle de données du test s'applique à tous les tests du Khi deux de l'Assistant. Chaque contrôle de données est décrit ci-dessous.

Test d'ajustement du Khi deux

Objectif

Nous avons évalué l'approximation du Khi deux par rapport aux statistiques du test en étudiant l'impact de l'importance et de la fréquence des petits dénombrements attendus sur le taux d'erreur de 1ère espèce.

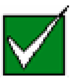
Méthode

Les échantillons d'effectif n ont été tirés d'une loi multinomiale, avec les proportions décrites dans les modèles à proportions perturbées ou les modèles à proportions égales (voir l'Annexe C). Pour chaque condition, nous avons effectué 10 000 tests d'ajustement du Khi deux avec un seuil de signification cible de 0,05. Pour chaque test, nous avons calculé l'erreur de 1ère espèce réelle de la façon suivante : $\frac{Nombre\ de\ tests\ rejetés}{Nombre\ de\ répliques\ (10\ 000)}$. Nous avons défini la plage des taux d'erreur de 1ère espèce acceptables sur [0,03 – 0,07] et relevé l'effectif d'échantillon minimal avec un taux d'erreur de 1ère espèce compris dans cette plage.

Les résultats

Les résultats de la simulation ont démontré que les dénombrements de cellules cible inférieurs à 1,25 peuvent donner lieu à des valeurs de p incorrectes lorsque le pourcentage de petits dénombrements de cellules cible est inférieur ou égal à 50 %. De même, des dénombrements de cellules cible inférieurs à 2,5 peuvent donner lieu à des valeurs de p incorrectes lorsque le pourcentage de petits dénombrements de cellules cible est supérieur à 50 %. Pour plus de détails, reportez-vous à l'Annexe D.

Lors de la vérification de la validité du test d'ajustement du Khi deux, le rapport de l'Assistant présente les indicateurs d'état suivants :

Etat	Condition
	Le dénombrement de cellules cible minimum est supérieur ou égal à 1,25 lorsque le pourcentage de petits dénombrements de cellules cible est inférieur ou égal à 50 %. OU Le dénombrement de cellules cible minimum est supérieur ou égal à 2,5 lorsque le pourcentage de petits dénombrements de cellules cible est supérieur à 50 %. Votre échantillon est assez grand pour que vous obteniez des dénombrements cible suffisants. La valeur de p du test devrait être exacte.

Etat	Condition
	Si les conditions ci-dessus ne sont pas remplies.

Test d'association du Khi deux

Objectif

Nous avons évalué l'approximation du Khi deux par rapport aux statistiques du test en étudiant l'impact de l'importance et de la fréquence des petits dénombrements attendus sur le taux d'erreur de 1ère espèce.

Méthode

Les échantillons d'effectif n_i ont été tirés d'une loi multinomiale, avec les proportions définies par les modèles à proportions perturbées ou à proportions égales (voir l'Annexe C). Pour simplifier, nous avons choisi $n_i = n \; \forall i$. Pour chaque condition, nous avons effectué 10 000 tests d'association du Khi deux avec un seuil de signification cible de 0,05. Pour chaque test, nous avons calculé le taux d'erreur de 1ère espèce réel de la façon suivante : $\frac{Nombre\ de\ tests\ rejetés}{Nombre\ de\ répliques\ (10\ 000)}$. Nous avons défini la plage des taux d'erreur de 1ère espèce acceptables sur [0,03 – 0,07] et relevé l'effectif d'échantillon minimal avec un taux d'erreur de 1ère espèce compris dans cette plage.

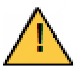
Les résultats

Nous avons constaté que le dénombrement minimum attendu pour les cellules dépend du nombre de valeurs X et du pourcentage de petits dénombrements attendus pour les cellules.

- Dans le cas du modèle à proportions perturbées, lorsque le pourcentage de petits dénombrements attendus pour les cellules est inférieur ou égal à 50 %, les dénombrements minimaux attendus pour les cellules sont ≤ 2 et ≤ 1 pour un nombre de valeurs X respectivement égal à (2 ou 3) et (4, 5 ou 6). En outre, lorsque le pourcentage de petits dénombrements attendus pour les cellules est > 50 %, les dénombrements minimaux attendus pour les cellules sont ≤ 3 et ≤ 1,5 pour un nombre de valeurs X respectivement égal à (2 ou 3) et (4, 5 ou 6).
- Dans le cas du modèle à proportions égales, le dénombrement minimum attendu pour les cellules est ≤ 2 lorsque le nombre de valeurs X est égal à (2 ou 3) et le dénombrement minimum attendu pour les cellules est ≤ 1,5 lorsque le nombre de valeurs X est égal à (4, 5 ou 6).

Pour plus de détails, reportez-vous à l'Annexe E.

Lors de la vérification de la validité du test d'association du Khi deux, le rapport de l'Assistant présente les indicateurs d'état suivants :

Etat	Nombre de valeurs de variables X	Condition
	2 ou 3	Le dénombrement minimum attendu pour les cellules est supérieur ou égal à 2 lorsque le pourcentage de petits dénombrements attendus pour les cellules (inférieur ou égal à 5) est inférieur ou égal à 50 %. Le dénombrement minimum attendu pour les cellules est supérieur ou égal à 3 lorsque le pourcentage de petits dénombrements attendus pour les cellules (inférieur ou égal à 5) est supérieur à 50 %.
	4, 5 ou 6	Le dénombrement minimum attendu pour les cellules est supérieur ou égal à 1 lorsque le pourcentage de petits dénombrements attendus pour les cellules (inférieur ou égal à 5) est inférieur ou égal à 50 %. Le dénombrement minimum attendu pour les cellules est supérieur ou égal à 2 (par commodité, 1,5 arrondi à 2) lorsque le pourcentage de petits dénombrements attendus pour les cellules (inférieur ou égal à 5) est supérieur à 50 %.
	Tous les cas	Si les conditions ci-dessus ne sont pas valables.

Test du Khi deux du pourcentage de défectueux en présence de plus de deux échantillons

Objectif

Nous avons évalué l'approximation du Khi deux par rapport aux statistiques du test en étudiant l'impact de l'importance et de la fréquence des petits dénombrements attendus sur le taux d'erreur de 1ère espèce.

Méthode

Nous avons défini les modèles p = $p_i = p_j \; \forall i, j$, p étant égal à 0,001, 0,005, 0,01, 0,025 et 0,25. Les échantillons d'effectif n_i ont été tirés d'une loi binomiale avec les valeurs de p_i décrites ci-dessus. Pour simplifier, nous avons choisi $n_i = n \; \forall i$. Pour chaque condition, nous avons effectué 10 000 tests du Khi deux du pourcentage de défectueux avec un seuil de signification de 0,05. Pour chaque test, nous avons calculé l'erreur de 1ère espèce réelle de la façon suivante : $\frac{Nombre\ de\ tests\ rejetés}{Nombre\ de\ répliques\ (10\ 000)}$. Nous avons défini la plage des taux d'erreur de 1ère espèce acceptables sur [0,03 – 0,07] et relevé l'effectif d'échantillon minimal avec un taux d'erreur de 1ère espèce compris dans cette plage.

Les résultats

Lorsqu'il existe 3 à 6 valeurs X, un nombre attendu minimum de défectueux et de non-défectueux supérieur ou égal à 1,5 produit un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07]. En présence de 7 à 12 valeurs X, un nombre attendu minimum de défectueux ou de non-défectueux supérieur ou égal à 1 produit un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07].

Pour plus de détails, reportez-vous à l'Annexe F.

Lors de la vérification de la validité du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons, le rapport de l'Assistant présente les indicateurs d'état suivants :

Etat	Nombre de valeurs X	Condition
	3 à 6	Le nombre minimum attendu de défectueux et de non-défectueux est supérieur ou égal à 1,5.
	7 à 12	Le nombre minimum attendu de défectueux et de non-défectueux est supérieur ou égal à 1.
	Tous les cas	Si les conditions ci-dessus ne sont pas valables.

Validité des intervalles

Les intervalles de comparaison du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons et du test d'ajustement du Khi deux sont basés sur l'approximation par une loi normale. En outre, les intervalles de confiance individuels du test d'ajustement du Khi deux sont basés sur l'approximation par une loi normale. Dans cette section, nous évaluons la validité de l'approximation par une loi normale. Selon la règle mentionnée généralement dans la plupart des manuels de statistiques, l'intervalle de confiance par approximation est exact si les dénombrements observés sont au moins égaux à 5.

La validité du contrôle de données des intervalles s'applique au test du Khi deux du pourcentage de défectueux pour plus de deux échantillons et au test d'ajustement du Khi deux.

Test du Khi deux du pourcentage de défectueux pour plus de deux échantillons

Objectif

Nous souhaitions évaluer la règle générale relative au nombre minimum de défectueux et de non-défectueux observés dans chaque échantillon afin de nous assurer de l'exactitude des intervalles de confiance par approximation.

Méthode

Nous définissons d'abord les intervalles utilisés dans le tableau comparatif. Les bornes des intervalles sont définies de telle sorte que, avec un taux d'erreur global d'environ $\propto$, tout intervalle ne chevauchant aucune de ces bornes indique une différence du pourcentage de défectueux. Pour connaître les formules utilisées, reportez-vous à l'Annexe G.

Les intervalles de comparaison sont basés sur des intervalles de confiance de comparaison des données appariées. Pour plus de détails, reportez-vous à la section Intervalles de comparaison du livre blanc de l'Assistant sur l'ANOVA à un facteur contrôlé. Nous utilisons un intervalle de confiance d'approximation par une loi normale pour chaque paire ($p_i – p_j$), puis nous utilisons une comparaison multiple de Bonferroni pour contrôler le taux d'erreur global de l'expérience. Par conséquent, nous n'avons besoin d'évaluer la validité que de l'un des intervalles de la procédure de comparaison de données appariées pour comprendre l'effet de l'approximation par une loi normale sur les intervalles de comparaison.

Les résultats

Pour évaluer la validité de l'approximation par une loi normale, nous avons seulement besoin d'examiner l'impact de l'approximation sur un intervalle pour connaître la différence entre les pourcentages de défectueux. Par conséquent, nous pouvons simplement utiliser la règle générale établie pour le cas des pourcentages de défectueux à 2 échantillons. Pour plus de détails, reportez-vous à la section sur les méthodes de test du pourcentage de défectueux à 2 échantillons dans le livre blanc de l'Assistant sur le test du pourcentage de défectueux à 2 échantillons. Les résultats de simulation du test du pourcentage de défectueux à 2 échantillons indiquent que l'exactitude de l'intervalle de confiance par approximation pour connaître la différence entre les pourcentages de défectueux est généralement fiable lorsque les échantillons sont assez grands, c'est-à-dire lorsque le nombre observé de défectueux et le nombre observé de non-défectueux dans chaque échantillon sont au moins égaux à 5.

Lors de la vérification de la validité des intervalles du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons, le rapport de l'Assistant présente les indicateurs d'état suivants :

Etat	Condition
	Tous les échantillons ont au moins 5 défectueux et 5 non-défectueux. Les intervalles de comparaison devraient être exacts.
	Si la condition ci-dessus n'est pas valable.

Test d'ajustement du Khi deux

Objectif

Nous souhaitions évaluer la règle générale relative au nombre minimum de défectueux et de non-défectueux observés dans chaque échantillon afin de nous assurer de l'exactitude des intervalles de confiance par approximation.

Méthode

Le test d’ajustement du Khi deux de l'Assistant comprend des comparaisons et des intervalles de confiance individuels. Nous utilisons les intervalles d'approximation par la loi normale standard pour les proportions et corrigeons les intervalles multiples à l'aide de la correction de Bonferroni (Goodman, 1965). Ainsi, les intervalles simultanés de Bonferroni sont calculés comme suit :

$$p_{iInférieure} = p_i - Z_{\alpha/2k}\sqrt{\frac{\mathrm{p_i}(1 - \mathrm{p_i})}{\mathrm{N}}}$$

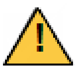
$$p_{iSupérieure} = p_i + Z_{\alpha/2k}\sqrt{\frac{\mathrm{p_i}(1 - \mathrm{p_i})}{\mathrm{N}}}$$

Les bornes des intervalles sont définies de sorte que, avec un taux d'erreur global d'environ ∝, tout intervalle ne contenant pas la valeur de proportion cible indique que la proportion réelle est différente de sa proportion cible correspondante. Les intervalles individuels utilisent la même forme que les intervalles de Bonferroni, mais ne corrigent pas les intervalles multiples à l'aide de la valeur $Z_{\alpha/2}$.

Les résultats

Les deux approches décrites ci-dessus suivent une méthodologie semblable à celle définie dans le test du pourcentage de défectueux à 2 échantillons de l'Assistant. Par conséquent, pour la validité de l'approximation par une loi normale, nous pouvons utiliser des règles semblables à celles que nous avons mises au point pour ce test. Pour plus de détails, reportez-vous à la section sur les méthodes de test du pourcentage de défectueux à 2 échantillons dans le livre blanc de l'Assistant sur ce test. Dans ce document, nous avons conclu que les intervalles de comparaison et les intervalles de confiance individuels peuvent ne pas être exacts lorsque les dénombrements d'échantillons sont inférieurs à 5.

Lors de la vérification de la validité des intervalles du test d'ajustement du Khi deux, le rapport de l'Assistant présente les indicateurs d'état suivants :

Etat	Condition
	Tous les dénombrements d'échantillons sont égaux à au moins 5. Les intervalles devraient être exacts.
	Il existe des dénombrements d'échantillons inférieurs à 5.

Références

Agresti, A. (1996), An introduction to categorical data analysis, New York, NY : Wiley.

Read, T. et Cressie, N. (1988), Goodness-of-fit statistics for discrete multivariate data, New York, NY : Springer-Verlag.

Fienberg, S. (1980), The analysis of cross-classified categorical data, Cambridge, MA : MIT Press.

Goodman, L. (1965), On simultaneous confidence intervals for multinomial proportions, *Technometrics, 7*, 247-254.

Annexe A : statistiques du test du Khi deux

L'Assistant utilise un test du Khi deux de la forme suivante :

$$x^2 = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

où

O_{ij} = dénombrements observés, conformément au tableau ci-dessous :

Cas	$\boldsymbol{O_{ij}}$
Test d'adéquation de l'ajustement d'une loi multinomiale	Le dénombrement observé de l'$i^{ème}$ résultat est défini comme suit : O_{i1}.
Test de l'égalité de plus de 2 pourcentages de défectueux	Le nombre observé d'éléments défectueux et d'éléments non défectueux de l'$i^{ème}$ échantillon est défini, respectivement, comme suit : O_{i1} et O_{i2} .
Test de l'association entre deux variables de catégories	Les dénombrements observés de l'$i^{ème}$ valeur de la variable X et de la $j^{ème}$ valeur de la variable Y sont définis comme suit : O_{ij}.

E_{ij} = dénombrement attendu conformément au tableau ci-dessous :

Cas	$\boldsymbol{E_{ij}}$
Test d'adéquation de l'ajustement d'une loi multinomiale	E_{i1} = np_i $i = 1, \ldots, k$ (k = nombre de résultats) n = effectif de l'échantillon p_i = proportions historiques $\sum_i p_i = 1$
Test de l'égalité de plus de 2 pourcentages de défectueux	E_{i1} = $n_i p$ (pour les défectueux) E_{i2} = $n_i(1 - p)$ (pour les non-défectueux) $i = 1, \ldots, k$ (k = nombre d'échantillons) n_i = $i^{ème}$ effectif d'échantillon p = proportion globale de défectueux

Cas	$\boldsymbol{E_{ij}}$
Test de l'association entre deux variables de catégories	$E_{ij} = \frac{(n_{i.} n_{.j})}{n_{..}}$ $i = 1, \ldots, m$ (m = nombre de valeurs X) $j = 1, \ldots, k$ (k = nombre de valeurs Y) $n_{i.}$ = dénombrement total relatif à l'$i^{ème}$ valeur de la variable X $n_{.j}$ = dénombrement total relatif à la $j^{ème}$ valeur de la variable Y $n_{..}$ = effectif d'échantillon global

Annexe B : puissance du test du Khi deux du pourcentage de défectueux en présence de plus de deux échantillons

Nous utilisons une loi du Khi deux non centrée pour calculer la puissance du test ($p_i = p_j = p\ \forall i, j$). Le paramètre de non-centralité dépend de n_i et $p_i \forall i$

où

n_i = l'effectif de l'$i^{\text{ème}}$ échantillon

Chaque p_i représente une proportion alternative (voir la section suivante de cette Annexe, Calcul des proportions alternatives) calculée à partir de la différence des proportions = δ.

Nous calculons le paramètre de non-centralité de la loi du Khi deux comme suit :

$$\chi^2 = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$O_{i1} = n_i p_i$$

$$O_{i2} = n_i (1 - p_i)$$

et calculons la puissance du test comme suit

$$\text{Prob}(X \geq x_{1-\alpha} \mid \chi^2)$$

X = est une variable aléatoire issue d'une loi du Khi deux non centrée avec le paramètre de non-centralité χ^2.

$x_{1-\alpha}$ = fonction cdf inverse évaluée à $1 - \alpha$ pour une loi du Khi deux centrée.

Calcul des proportions alternatives

Nous avons défini les proportions alternatives comme suit :

$$p_i = p_c + \frac{n_j}{n_i + n_j}\delta$$

$$p_j = p_c - \frac{n_i}{n_i + n_j}\delta$$

$$p_m = p_c \forall m \neq i, j$$

$$0 < \delta < 1$$

$$p_c = \frac{1}{N_T}\sum_{i=1}^{k} n_i \widehat{p_\iota}$$

$\widehat{p_\iota}$ = éléments défectueux des proportions d’échantillons pour l'$i^{\text{ème}}$ échantillon.

N_T = nombre total d'observations.

n_i = effectif de l'$i^{\text{ème}}$ échantillon.

Pour certaines différences δ, $p_i > 1$ ou $p_j < 0$. Par conséquent, nous élaborons les règles suivantes :

Si $p_j <$ 0	$p_i = \delta$ $p_j = 0$ $p_m = \frac{\delta}{2}\ \forall m \neq i, j$
Si $p_i >$ 1	$p_i = 1$ $p_j = 1 - \delta$ $p_m = 1 - \frac{\delta}{2}\ \forall m \neq i, j$

L'utilisation des deux plus petites valeurs de n_i donne lieu à la puissance minimale et l'utilisation des deux plus grandes valeurs de n_i donne lieu à la puissance maximale.

Annexe C : modèle à proportions perturbées et modèle à proportions égales

Modèle à proportions perturbées

A l'instar des auteurs Read et Cressie (1988), nous définissons l'ensemble de proportions dans le cadre de l'hypothèse nulle comme suit :

Nous choisissons une valeur δ proche de k - 1 (k = nombre de proportions pour chaque échantillon) et définissons un ensemble de valeurs p_i faibles comme suit :

$$p_i = \frac{(1 - \frac{\delta}{k-1})}{k} \text{ pour } i = 1, \dots, r$$

et la valeur restante p_i comme suit :

$$p_i = \frac{(1 - \sum_{i=1}^{r} p_i)}{(k-r)} \text{ pour } i = r+1, \dots, k$$

Les valeurs que nous avons utilisées pour δ dans les simulations figurent dans le tableau 1.

Tableau 1 : Valeur de δ utilisée dans les simulations avec la petite valeur p_i obtenue

k	δ	$p_{i=1,\dots,r}$
3	1,95	0,008
4	2,95	0,004
5	3,90	0,005
6	4,90	0,003

Pour chaque valeur de k, nous avons fait varier r = 1,..., k – 1 pour modifier l'ensemble des petites valeurs p_i's. . Par exemple, pour k = 3, nous avons obtenu les deux modèles suivants décrits dans le tableau 2.

Tableau 2 : Valeurs de p_i pour k = 3 à l'aide du modèle à proportions perturbées

r	p1	p2	p3
1	0,008	0,496	0,496
2	0,008	0,008	0,984

Modèles à proportions égales

Pour obtenir un modèle où 100 % des dénombrements attendus pour les cellules sont faibles, nous utilisons un modèle à proportions égales défini par

$$p_i = \frac{1}{k} \forall\ i$$

A l'aide de ce modèle, avec un effectif d'échantillon très faible, tous les dénombrements attendus pour les cellules sont considérés comme petits. Avec un modèle à proportions égales, les effectifs d'échantillons doivent être très faibles pour obtenir un petit dénombrement attendu pour les cellules, ce qui, en pratique, ne se produira probablement pas.

Annexe D : validité du test d'ajustement du Khi deux

Pour le modèle à proportions perturbées, nous avons représenté sous forme graphique le dénombrement minimum attendu pour les cellules, nécessaire pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au pourcentage de petits dénombrements attendus pour les cellules, conformément à la figure 1.

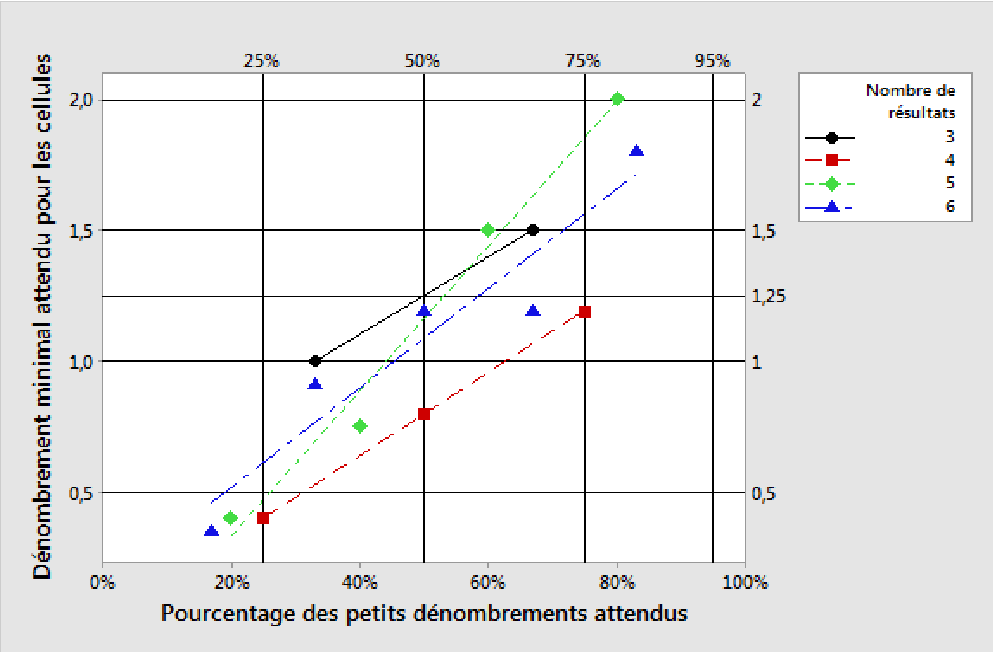


Figure 1 : Dénombrements minimaux attendus pour les cellules, nécessaires pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au pourcentage de petits dénombrements attendus pour les cellules

Dans la figure 1, lorsque le pourcentage de petits dénombrements attendus pour les cellules est inférieur à 50 %, les dénombrements minimaux attendus pour les cellules sont inférieurs ou égaux à 1,25. Tous les dénombrements minimaux attendus pour les cellules sont inférieurs ou égaux à 2. Sur la base de ces résultats de simulation, les règles que nous utilisons dans le rapport de l'Assistant sont prudentes.

Nous avons ensuite procédé à la même simulation à l'aide du modèle à proportions égales pour définir la distribution nulle. Le tableau 4 récapitule les résultats de la simulation à l'aide d'un modèle à proportions égales.

Tableau 4 : Dénombrement minimum attendu dans les cellules pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07]

k	Dénombrement minimum attendu pour les cellules
3	2,5
4	1,25
5	1
6	1,4

Comme indiqué ci-dessus, le modèle à proportions égales donne lieu à des cas où 100 % des dénombrements de cellules sont petits. Le tableau 4 montre que tous les dénombrements minimaux attendus pour les cellules sont inférieurs ou égaux à 2,5, ce qui justifie les règles que nous utilisons dans le rapport de l'Assistant.

Annexe E : validité du test d'association du Khi deux

Pour le modèle à proportions perturbées, nous avons représenté sous forme graphique le dénombrement minimum attendu pour les cellules, nécessaire pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au pourcentage de petits dénombrements attendus pour les cellules, pour chaque nombre de valeurs X, conformément à la figure 2.

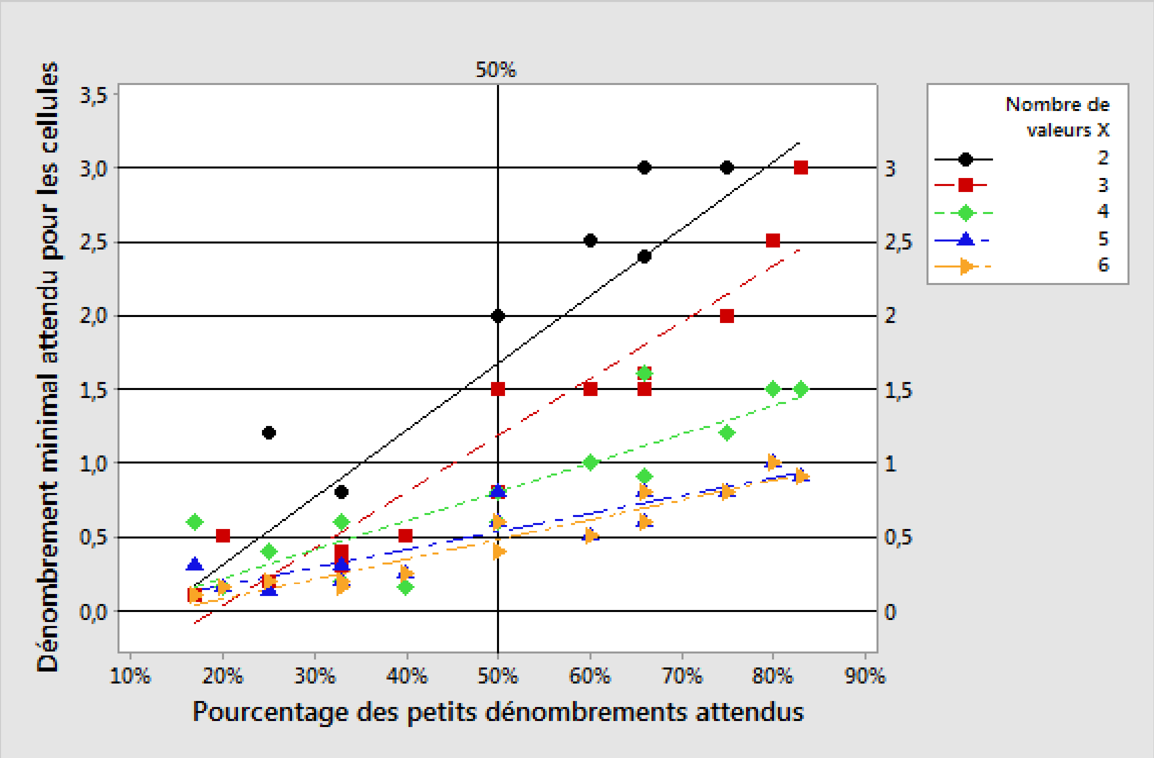


Figure 2 : Dénombrements minimaux attendus pour les cellules, nécessaires pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au pourcentage de petits dénombrements attendus pour les cellules.

La figure 2 indique que le dénombrement minimum attendu pour les cellules dépend du nombre de valeurs X et du pourcentage de petits dénombrements attendus pour les cellules.

La figure 2 indique que, lorsque le pourcentage de petits dénombrements attendus pour les cellules est ≤ 50 %, les dénombrements minimaux attendus pour les cellules sont ≤ 2 et ≤ 1 pour un nombre de valeurs X respectivement égal à 2 ou 3 et 4, 5 ou 6. En outre, lorsque le pourcentage de petits dénombrements attendus pour les cellules est > 50 %, les dénombrements minimaux attendus pour les cellules sont ≤ 3 et ≤ 1,5 pour un nombre de valeurs X respectivement égal à 2 ou 3 et 4, 5, ou 6.

Pour le modèle à proportions égales, nous avons représenté sous forme graphique le dénombrement minimum attendu pour les cellules par rapport au nombre de valeurs X (m) et au nombre de valeurs Y (k), comme représenté dans la figure 3.

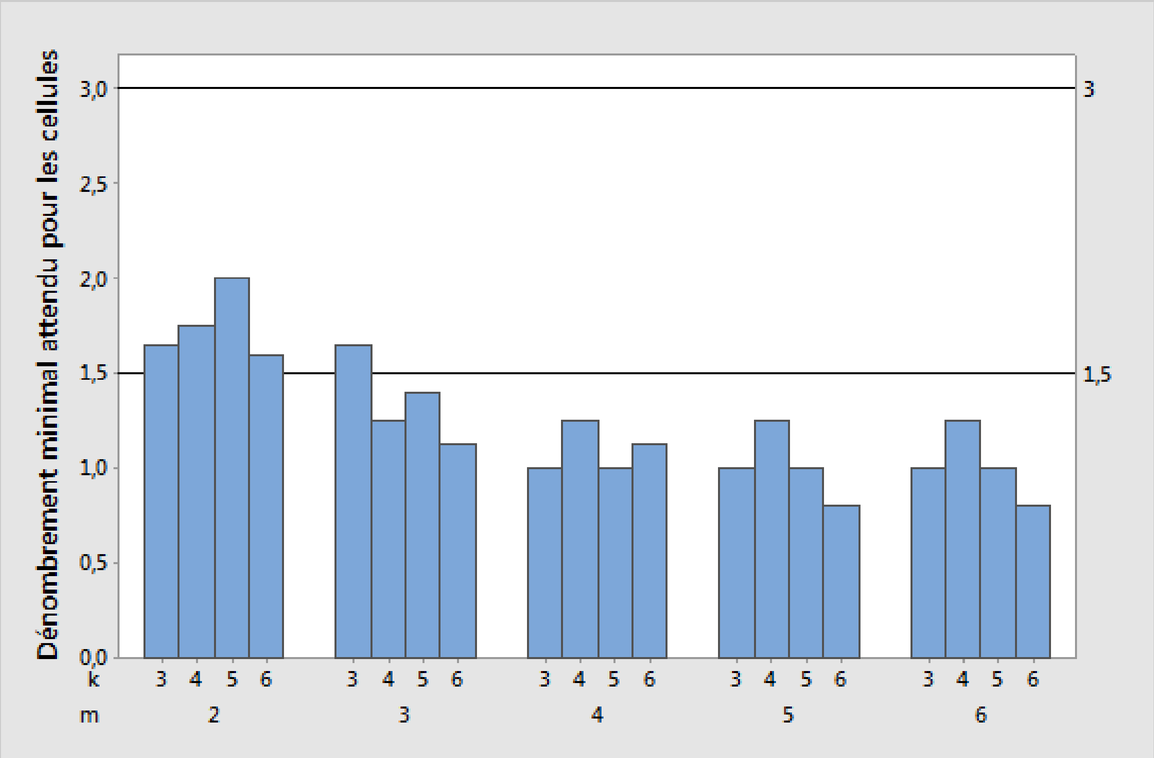


Figure 3 : Dénombrement minimum attendu pour les cellules, nécessaire pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport aux valeurs X (m) et Y (k)

La figure 3 indique que le dénombrement minimum attendu pour les cellules est ≤ 2 lorsque le nombre de valeurs X est égal à 2 ou 3 et que le dénombrement minimum attendu pour les cellules est ≤ 1,5 lorsque le nombre de valeurs X est égal à 4, 5 ou 6. Sur la base de ces résultats de simulation, les règles du rapport de l'Assistant sont prudentes.

Annexe F : validité du test du Khi deux du pourcentage de défectueux en présence de plus de deux échantillons

Pour chaque valeur de p et chaque valeur m = 3, 4, 5,..., 12, nous avons représenté sous forme graphique le dénombrement minimum attendu pour les cellules. Les résultats sont présentés dans les figures 4 et 5.

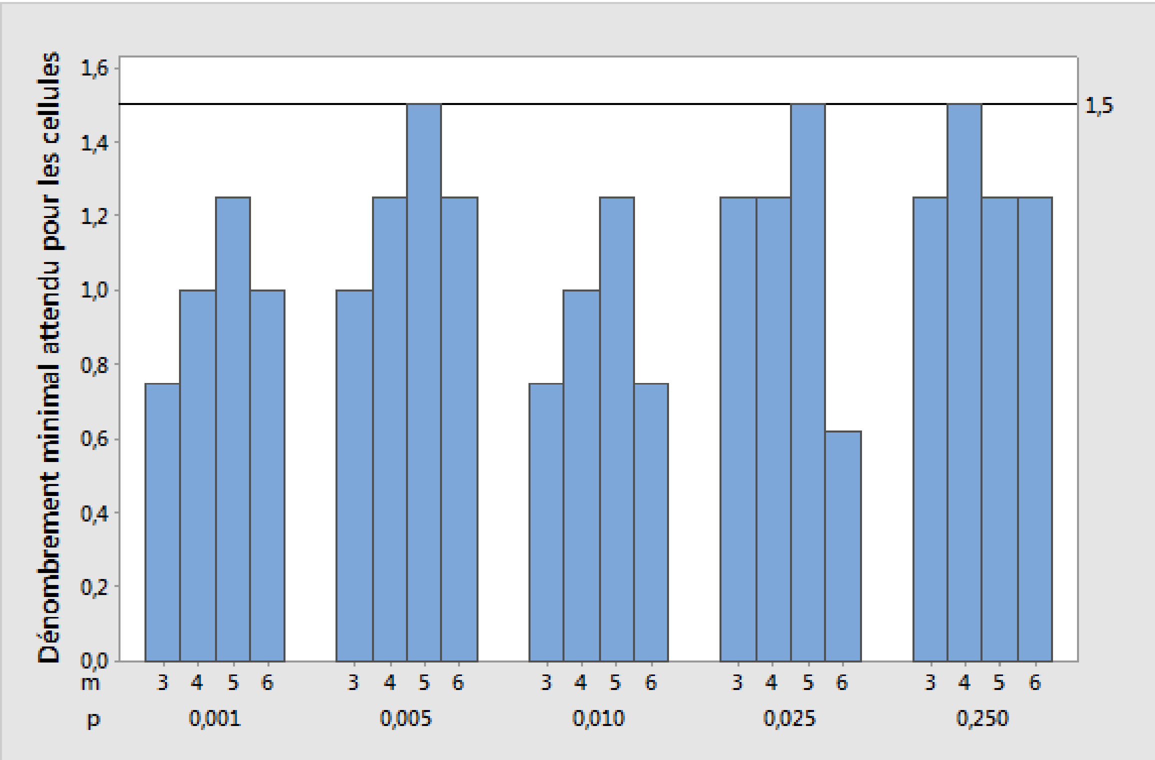


Figure 4 : Dénombrement minimum attendu pour les cellules, nécessaire pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au nombre de valeurs X (m = 3 à 6)

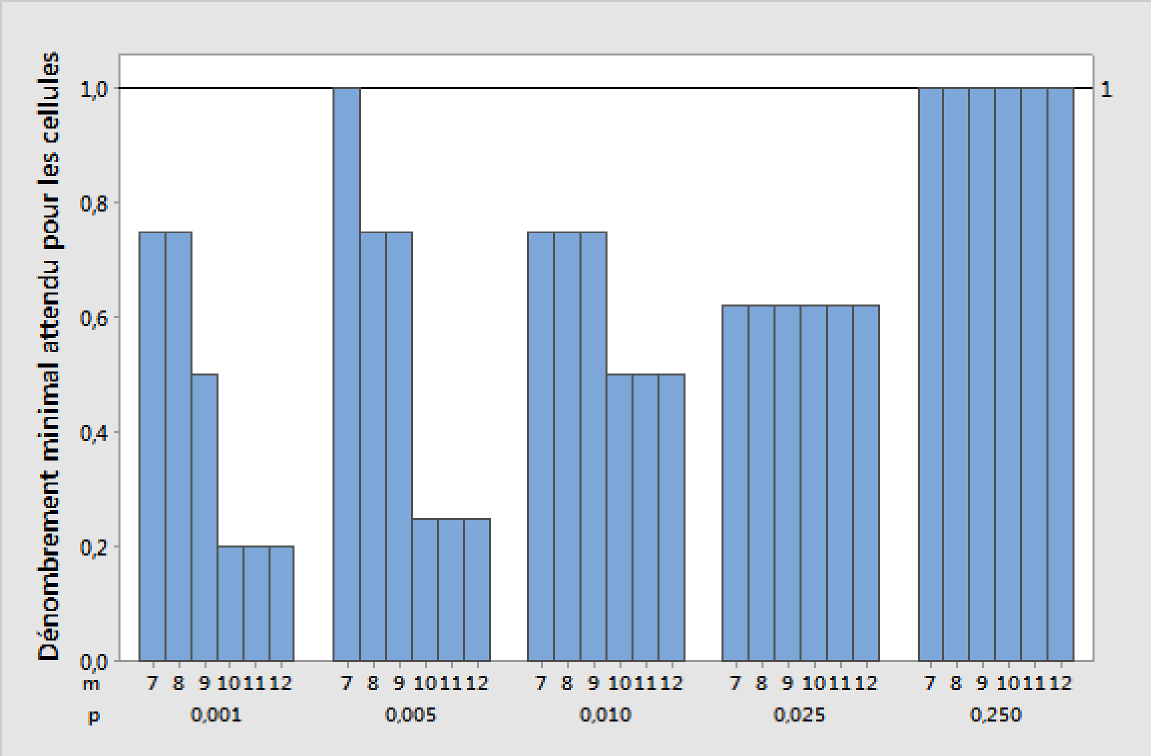


Figure 5 : Dénombrement minimum attendu pour les cellules, nécessaire pour obtenir un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07] par rapport au nombre de valeurs X (m = 7 à 12)

Lorsque le nombre de valeurs X est égal à 3, 4, 5 ou 6, un dénombrement attendu pour les cellules supérieur ou égal 1,5 produit un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07]. Lorsque le nombre de valeurs X est égal à 7, 8, 9,..., 12, un dénombrement attendu pour les cellules supérieur ou égal à 1 produit un taux d'erreur de 1ère espèce dans l'intervalle [0,03, 0,07].

Annexe G : intervalles de comparaison du test du Khi deux du pourcentage de défectueux pour plus de deux échantillons

Les bornes inférieure et supérieure de la valeur p_i sont définies comme suit :

$$p_{iInférieure} = p_i - Z_{\alpha/c} X_i$$

$$p_{iSupérieure} = p_i + Z_{\alpha/c} X_i$$

où

c = nombre de comparaisons = k (k - 1) /2

k correspondant au nombre d'échantillons.

$Z_{\alpha/c}$ = $(1 - \frac{\alpha}{2c})$ percentile d'une loi normale avec moyenne = 0 et écart type = 1

$$X_i = ((k-1)\sum_{j \neq i} b_{ij} - \sum\sum_{1 \leq j < l \leq k} b_{jl}) / ((k-1)(k-2))$$

où

$$b_{ij} = \sqrt{\frac{p_i(1 - p_i)}{n_i} + \frac{p_j(1 - p_j)}{n_j}}$$