

Intervalles de confiance sur écarts-types emboîtés

Louis-Jean Hollebecq
Directeur scientifique et technique

Sommaire

1	Introduction	2
2	Symboles et abréviations.....	2
3	Exemples de situations où un IC sur un ET imbriqué peut être nécessaire.....	3
4	Bases pour l'estimation d'une variance avec des niveaux imbriqués	4
4.1	Méthodes classiques pour 2 niveaux	4
4.2	Estimation d'une variance avec 2 niveaux imbriqués	5
4.3	Equation pour calculer une valeur approchée de $(v_2/V_2)_\alpha$ pour un α donné	6
4.4	Estimation d'une variance avec plus de 2 niveaux imbriqués.....	7
5	Estimation d'une variance avec niveaux imbriqués et rapport inconnu entre elles	12
5.1	Introduction.....	12
5.2	Estimation d'une variance de niveaux imbriqués quand la variance de niveau 1 est connue	13
5.3	Estimation d'une variance de niveaux imbriqués quand $v_1/n_1 \ll w_2$	13
5.4	Estimation d'une variance de niveaux imbriqués quand $v_2 < 0$	14
5.5	Situations pour lesquelles les résultats d'estimation ne sont pas satisfaisants en regard de son utilisation prévue.....	15
6	Conclusions	16
7	Références	16

Résumé :

Dans de nombreuses situations, les écarts types (ET) doivent être calculés à partir d'un plan emboîté. Cela se produit généralement dans les procédures de contrôle de la qualité et dans les comparaisons interlaboratoires. Dans de telles situations, les bases du calcul des écarts-types sont bien connues, mais les méthodes existantes pour calculer l'intervalle de confiance (IC) sur ces écarts-types sont plutôt insatisfaisantes. En particulier, elles ne tiennent pas du tout compte des valeurs négatives souvent rencontrées pour les variances estimées correspondantes. Cet article fournit des équations qui décrivent bien les distributions des variances des niveaux emboîtés et de leur dispersion, à condition que les vraies valeurs de ces variances soient connues. Les cas de 2 niveaux imbriqués et de plus de 2 niveaux imbriqués sont pris en compte. L'inversion de ces équations pour trouver l'IC sur les valeurs réelles des variances en fonction de leurs estimations est malheureusement impossible lorsque les variances des niveaux inférieurs sont inconnues. Toutefois, cet article propose des équations approchées qui peuvent être utilisées lorsque l'on peut s'attendre à ce que l'impact des variances des niveaux inférieurs soit faible. Des méthodes permettant de vérifier si cette condition est remplie sont également proposées. Si ce n'est pas le cas, le nombre de répétitions aux niveaux inférieurs doit être augmenté pour obtenir une détermination acceptable de l'IC.

1 Introduction

Dans de nombreuses situations, les écarts types doivent être calculés à partir d'un plan d'expériences imbriqué. C'est généralement le cas dans les essais d'aptitude des laboratoires, lorsque l'écart-type interlaboratoire ou l'écart-type d'homogénéité est calculé à partir d'une série d'essais répétés. C'est également le cas dans de nombreuses autres situations, par exemple lorsqu'un niveau de qualité à long terme d'une production est calculé à partir d'une série de résultats de tests de libération.

Dans de telles situations, les bases du calcul de l'écart-type sont bien connues, mais les méthodes existantes pour calculer l'IC à partir de ces bases sont plutôt insatisfaisantes. En particulier, elles ne prennent pas du tout en compte les valeurs négatives souvent rencontrées pour les variances estimées correspondantes.

Ce document fournit :

- ✚ Des exemples où de tels IC sont nécessaires ;
- ✚ Les bases techniques de leur calcul pour 2 niveaux et pour plus de 2 niveaux ;
- ✚ Des formules approchées pour calculer l'IC sur les variances imbriquées estimées lorsque le rapport des variances imbriquées est connu ;
- ✚ Des méthodes pour calculer l'IC sur les variances imbriquées estimées lorsque le rapport des variances imbriquées est inconnu.

2 Symboles et abréviations

Les symboles utilisés dans ce document sont listés dans le Tableau 1.

Tableau 1. Liste des symboles utilisés dans ce document.

Symbole	Désignation et commentaires
i	Niveau de variances imbriquées (les niveaux sont ordonnés des plus profonds aux plus globaux du plan d'expériences)
j	Indice pour un niveau donné i
k	Nombre total de niveaux imbriqués
m_1	Valeur moyenne des $x_{1,j}$ ($m_1 = \sum x_j/n_1$)
m_i	Valeur moyenne des valeurs de m_{i-1} ($m_i = \sum m_{i-1}/n_i$)
sd_i	Estimateur de SD_i
SD_i	Ecart-type du niveau i
T	Valeur seuil pour $V_1/(n_1 \cdot V_2)$ associée à une valeur α pour l'IC avec les paramètres n_1 , n_2 , V_1 et V_2 pour laquelle la limite inférieure de l'IC sur v_2 est négative
v_i	Estimateur de V_i
V_i	Variance du niveau i ($V_i = SD_i^2$)
w_i	Variance globale du niveau i , telle que définie dans l'équation (1) et l'équation (7)
x_{ij}	Valeurs de base à partir desquelles m_i et sd_i sont calculés
α	Niveau de confiance des intervalles de confiance

Abréviations :

-  IC : intervalle de confiance bilatéral. Par exemple, IC95% signifie l'intervalle means de confiance bilatéral [2,5%;97,5%]
-  MMC : méthode de Monte-Carlo
-  EA : essais d'aptitude
-  ET : écart-type (dénommé SD dans les équations)

3 Exemples de situations où un IC sur un ET imbriqué peut être nécessaire

Chaque fois qu'un ET est calculé à partir d'une série de résultats répétés, la connaissance de l'IC correspondant peut être nécessaire. C'est le cas, par exemple, dans les situations suivantes

-  L'attestation de conformité sur la base du niveau de qualité à long terme, comme spécifié par exemple dans la norme EN 10080 [1]. Le niveau de qualité du produit (dans ce cas, des aciers d'armature pour béton) est attesté non seulement par des essais internes de contrôle sur lots et des essais par des laboratoires indépendants, mais aussi par une surveillance du niveau de qualité à long terme de la production de l'usine. Pour ce faire, un écart-type à long terme des lots produits est calculé à partir de la série d'essais de contrôle sur lots. L'écart-type interne des lots doit être pris en compte pour obtenir cet écart-type à long terme de la production ;
-  Les expériences d'essais de fidélité spécifiées dans la norme ISO 5725-2 [2] sont basées sur l'ANOVA (analyse des variances) qui décompose les écarts types imbriqués correspondant aux différents niveaux de fidélité (typiquement la répétabilité, les niveaux intermédiaires et la reproductibilité). Le calcul de l'écart-type de fidélité intermédiaire peut alors demander de prendre en compte l'écart-type de sous-niveau ;

- ✚ Dans le cas des EA tel que spécifiés dans la norme ISO 13528 [3], la vérification de l'homogénéité des éléments utilisés pour l'EA (telle que décrite à l'annexe B) nécessite de prendre en compte l'écart-type de répétabilité des essais utilisés pour réaliser cette vérification ;
- ✚ Dans les deux cas ci-dessus, il peut y avoir plus de deux niveaux, par exemple, le niveau 1 est la répétition des essais dans des conditions de répétabilité, le niveau 2 est la répétition des essais sur des échantillons différents, le niveau 3 est la répétition des essais en utilisant des opérateurs différents et le niveau 4 est la répétition des essais sur des échantillons différents.

4 Bases pour l'estimation d'une variance avec des niveaux imbriqués

4.1 Méthodes classiques pour 2 niveaux

L'équation de base pour calculer une variance avec 2 niveaux imbriqués est rappelée dans l'équation (1) :

$$v_2 = w_2 - v_1/n_1 \quad (1)$$

Où v_2 est l'estimateur de V_2 ,

$$w_2 = \left(\sum_{j=1}^{n_2} (m_{1j} - m_2)^2 \right) / (n_2 - 1)$$

v_1 est l'estimateur de V_1 ,

et n_1 est le nombre de résultats utilisés pour calculer v_1 .

Exemple 1 :

Lors d'une vérification de l'homogénéité des échantillons réalisée en conformité avec l'annexe B de la norme ISO 13528 [3], les résultats d'essais du Tableau 2 ont été trouvés.

Tableau 2. Exemple de calcul de v_2 .

	Ech. 1	Ech.2	Ech.3	Ech.4	Ech.5	Ech.6	Ech.7	Ech.8	Ech.9	Ech.10
Essai 1	103,2	99,6	99,3	100,6	101,7	101,6	97,1	97,7	106,8	97,4
Essai 2	99,8	96,5	100,7	100,6	100,4	105,1	99,4	101,3	102,4	97,8
m_{1j}	101,5	98,05	100	100,6	101,05	103,35	98,25	99,5	104,6	97,6
v_{1j}	5,78	4,805	0,98	0	0,845	6,125	2,645	6,48	9,68	0,08

Donc, $w_2 = \sum_{j=1}^{10} m_{1j}/10 = 5,216$, $v_1 = \sum_{j=1}^{10} v_{1j}/10 = 3,742$, $n_1 = 2$, $v_2 = 5,216 - 3,742/2 = 3,345$

Plusieurs solutions sont classiquement proposées pour calculer l'IC sur V_2 comme par exemple :

- ✚ Dans la référence [4], l'IC est calculé au moyen de l'équation $IC = (w_2 \cdot F)/n_1$ où F est la valeur de Fisher-Snedecor (dépendante de n_1 , n_2 et α). Dans cet exemple, l'IC90% serait donc [-1,74;0,61];
- ✚ L'IC peut également être calculé avec l'équation $IC = (w_2 - v_1/n_1)/\chi_{\alpha, n_2-1}^2$. Dans cet exemple, l'IC90% serait donc [1,24;6,29].

Malheureusement, aucune d'entre elles ne cadre avec les simulations que l'on peut réaliser par la MMC (voir [5]). En particulier, elles ne s'accrochent pas bien des résultats négatifs souvent trouvés, notamment par la deuxième méthode. Cela s'explique par le fait que, dans les deux cas, plus ou moins une seule source de dispersion est prise

en compte au lieu de deux. Ceci est particulièrement évident dans le second cas, où la quantité $w_2 - v_1/n_1$ joue le même rôle que s^2 dans l'estimation simple de la variance selon l'équation (2) (origine : ISO 2854 [6]).

$$(n-1) \cdot \frac{v}{V} \approx \chi_{n-1}^2 \quad (2)$$

4.2 Estimation d'une variance avec 2 niveaux imbriqués

Nous avons pu vérifier par la MMC que l'équation (3), équivalente à l'équation (2), décrit exactement la distribution des estimateurs de variances avec 2 niveaux imbriqués.

$$\frac{v_2}{V_2} \approx \left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{n_2-1}^2}{n_2-1} - \frac{V_1}{n_1 \cdot V_2} \cdot \frac{\chi_{n_2 \cdot (n_1-1)}^2}{n_2 \cdot (n_1-1)} \quad (3)$$

Note : Lorsqu'elle est utilisée en MMC, cette équation (3) doit être utilisée avec des valeurs α indépendantes pour les deux fonctions de distribution χ^2 . Pour cette raison, elle ne peut pas être réécrite sous la forme $\frac{v_2}{V_2} \approx \frac{\chi_{n_2-1}^2}{n_2-1} \cdot \left(1 + \frac{V_1}{n_1 \cdot V_2} \cdot (1 - F_{n_2 \cdot (n_1-1), n_2-1})\right)$ en raison des effets de covariances.

Note : V_2 , qui est la valeur vraie de la variance, ne peut pas être négative, mais v_2/V_2 peut évidemment l'être, ce qui conduit à des valeurs négatives pour v_2 . Dans ces cas, $sd_2 = \sqrt{v_2}$ ne peut pas être calculé et sd_2 est généralement considéré comme égal à 0.

Note : Comme la valeur moyenne d'une distribution du χ^2 est son nombre de degrés de liberté, nous pouvons facilement vérifier que la valeur moyenne de la distribution décrite par l'équation (3) est égale à 1. Ceci confirme que l'équation (3) fournit des estimations sans biais de v_2/V_2 .

Cette équation (3) soulève les commentaires suivants :

Commentaire 1 :

Le premier terme de l'équation (3), c.à.d. $\left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{n_2-1}^2}{n_2-1}$ correspond au terme w_2 de l'équation (1) alors que le terme $\frac{V_1}{n_1 \cdot V_2} \cdot \frac{\chi_{n_2 \cdot (n_1-1)}^2}{n_2 \cdot (n_1-1)}$ de l'équation (3) correspond au terme v_1/n_1 de l'équation (1).

Commentaire 2 :

Le terme $V_1/(n_1 \cdot V_2)$ joue un rôle fondamental dans la précision de l'estimation de v_2/V_2 :

- ✚ Quand $V_1/(n_1 \cdot V_2) \ll 1$, alors $\frac{v_2}{V_2} \approx \frac{\chi_{n_2-1}^2}{n_2-1}$, c'est à dire une situation similaire à celle de l'équation (2);
- ✚ Quand $V_1/(n_1 \cdot V_2) \gg 1$, alors $\frac{v_2}{V_2} \approx \frac{V_1}{n_1 \cdot V_2} \cdot \left(\frac{\chi_{n_2-1}^2}{n_2-1} - \frac{\chi_{n_2 \cdot (n_1-1)}^2}{n_2 \cdot (n_1-1)}\right)$, c.à.d. Une valeur aléatoire dont l'ordre de grandeur est $[-k \cdot V_1/(n_1 \cdot V_2); +k \cdot V_1/(n_1 \cdot V_2)]$, k dépendant de n_1 et n_2 , (voir l'équation (4) plus loin), et aucune estimation valable de v_2/V_2 ne peut être calculée ;
- ✚ Les situations où $V_1/(n_1 \cdot V_2) \cong 1$ sont des situations intermédiaires pour lesquelles les estimations de v_2/V_2 peuvent être calculées, mais peuvent être de faible qualité.

Commentaire 3 :

L'équation (3) implique une combinaison de 2 valeurs aléatoires indépendantes (c.à.d. $\chi_{n_2-1}^2$ et $\chi_{n_2 \cdot (n_1-1)}^2$). Malheureusement, si une somme de deux χ^2 est facile à manipuler (par définition du χ^2), une différence de deux χ^2 ne l'est pas du tout.

Cependant, comme la variance d'une distribution du χ_n^2 est $2n$ et grâce à la loi de combinaison des variances, nous pouvons calculer facilement la valeur exacte de l'ET des estimations de v_2/V_2 avec l'équation (4), comme suit :

$$SD\left(\frac{v_2}{V_2}\right) = \sqrt{\left(1 + \frac{V_1}{n_1 \cdot V_2}\right)^2 \cdot \frac{2}{n_2 - 1} + \left(\frac{V_1}{n_1 \cdot V_2}\right)^2 \cdot \frac{2}{n_2 \cdot (n_1 - 1)}} \quad (4)$$

De plus, il a pu être confirmé avec la MMC que la moyenne quadratique de la distribution résultante est égale à 1, confirmant que l'équation (3) fournit une estimation sans biais de v_2/V_2 .

Commentaire 4 :

En pratique, le rapport qui a un intérêt pratique est V_2/v_2 et non pas v_2/V_2 , parce que nous connaissons v_2 et nous recherchons un IC sur V_2 . Obtenir un IC sur V à partir de v en utilisant l'équation (2) est facile, mais obtenir un IC sur V_2 à partir de v_2 en utilisant l'équation (3) n'est pas simple du tout, et même impossible quand le rapport V_1/V_2 n'est pas connu. Nous reviendrons sur ce sujet au § 5.

Commentaire 5 :

Nous avons pu vérifier par la MMC que, même si l'équation (3) produit des distributions asymétriques, les quantiles $(v_2/V_2)_\alpha$ (où α est la probabilité cumulée de la distribution) sont distribués de sorte que les valeurs des différences $(v_2/V_2)_{1-\alpha} - (v_2/V_2)_\alpha$ sont proches des différences entre les quantiles correspondants des distributions gaussiennes dont l'ET est le même. Par exemple, les différences $(v_2/V_2)_{0,975} - (v_2/V_2)_{0,025} \cong 4 \cdot SD(v_2/V_2)$ calculés avec l'équation (4).

Commentaire 6 :

De la même façon, nous avons pu vérifier avec la MMC que, même si l'équation (3) produit des distributions asymétriques, la valeur centrale pour une valeur α donnée peut être approchée par l'équation (5) comme suit.

$$CV_\alpha\left(\frac{v_2}{V_2}\right) \cong \left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{1-\alpha, n_2-1}^2 + \chi_{\alpha, n_2-1}^2}{2 \cdot (n_2 - 1)} - \frac{V_1}{n_1 \cdot V_2} \cdot \frac{\chi_{1-\alpha, n_2 \cdot (n_1-1)}^2 + \chi_{\alpha, n_2 \cdot (n_1-1)}^2}{2 \cdot n_2 \cdot (n_1 - 1)} \quad (5)$$

Où CV_α représente une valeur centrale de la distribution des estimations de $(v_2/V_2)_\alpha$ pour un α donné.

4.3 Equation pour calculer une valeur approchée de $(v_2/V_2)_\alpha$ pour un α donné

Les commentaires 5 et 6 nous permettent de calculer des valeurs approchées de $(v_2/V_2)_\alpha$ en utilisant l'équation (6), comme suit :

$$(v_2/V_2)_\alpha = CV_\alpha + k_\alpha \cdot SD \quad (6)$$

Où CV_α représente une valeur centrale de la distribution estimations de $(v_2/V_2)_\alpha$ pour un α donné, calculé avec l'équation (5),
 k est le quantile de la distribution gaussienne,

SD est l'écart-type de la distribution des estimations de la variance, calculé avec l'équation (4).

Exemple 2 :

Plan d'expériences utilisé dans l'exemple 1, c.à.d. $n_1 = 2$, $n_2 = 10$ pour lequel V_1/V_2 serait égal à 4 (c.à.d. $SD_1 = 2 \cdot SD_2$, ce qui fait sens dans le contexte de la norme ISO 13528 annexe B), et qui a été utilisé pour construire l'exemple 1. Donc $V_1/(n_1 \cdot V_2) = 2$, $SD\left(\frac{v_2}{V_2}\right) = \sqrt{(1 + 2)^2 \cdot \frac{2}{10-1} + (2)^2 \cdot \frac{2}{10 \cdot (2-1)}} \cong 1,673$ et les calculs qui en découlent sont présentés dans le Tableau 3.

Tableau 3. Exemple de calcul d'une valeur approchée $(v_2/V_2)_\alpha$.

α	$(v_2/V_2)_\alpha$ calculé par MMC	$\frac{\chi_{\alpha, n_2-1}^2}{n_2-1}$	$\left(\frac{\chi_{\alpha, n_2-1}^2}{n_2-1} + \frac{\chi_{1-\alpha, n_2-1}^2}{n_2-1}\right)/2$	$\frac{\chi_{\alpha, n_2, (n_1-1)}^2}{n_2 \cdot (n_1-1)}$	$\left(\frac{\chi_{\alpha, n_2, (n_1-1)}^2}{n_2 \cdot (n_1-1)} + \frac{\chi_{1-\alpha, n_2, (n_1-1)}^2}{n_2 \cdot (n_1-1)}\right)/2$	$CV_\alpha \left(\frac{v_2}{V_2}\right)$	k de la distribution gaussienne	$(v_2/V_2)_\alpha = CV_\alpha + k \cdot SD$
0,025	-2,02	0,300	1,207	0,325	1,187	1,248	-1,960	-2,03
0,05	-1,54	0,369	1,125	0,394	1,112	1,149	-1,645	-1,60
0,1	-1,00	0,463	1,047	0,487	1,043	1,057	-1,282	-1,09
0,2	-0,36	0,598	0,979	0,618	0,981	0,975	-0,842	-0,43
0,5	0,87	0,927	0,927	0,934	0,934	0,913	0,000	0,91
0,8	2,29	1,360	0,979	1,344	0,981	0,975	0,842	2,38
0,9	3,15	1,632	1,047	1,599	1,043	1,057	1,282	3,20
0,95	3,92	1,880	1,125	1,831	1,112	1,149	1,645	3,90
0,975	4,64	2,114	1,207	2,048	1,187	1,248	1,960	4,53

Notez que les α sont choisis de sorte que les valeurs du tableau peuvent être trouvées à la ligne $1-\alpha$. Par exemple, pour $\alpha = 0,025$, $\frac{\chi_{1-\alpha, n_2-1}^2}{n_2-1} = 2,114$. Donc le terme de la 4^e colonne est $(0,300 + 2,114)/2 = 1,207$.

Nous pouvons conclure de ce tableau que, quand $n_1 = 2$, $n_2 = 10$ and $SD_1/SD_2 = 2$, 95% des estimations sd_2 de l'écart-type SD_2 sont inclus dans l'intervalle $[0; 2,13]$, environ 35% d'entre étant égaux à 0.

Nous constatons avec cet exemple que l'équation (6) fournit des valeurs approchées de $(v_2/V_2)_\alpha$ qui ne sont pas mauvaises. De plus, lors de simulations par MMC pour un grand nombre de valeurs de n_1 , n_2 and V_1/V_2 , l'équation (6) s'est montrée plus précise lorsque soit $V_1/(n_1 \cdot V_2) \ll 1$ soit $V_1/(n_1 \cdot V_2) \gg 1$. En d'autres termes, cet exemple pour lequel $V_1/(n_1 \cdot V_2) \cong 1$ n'est pas un exemple favorable.

En conclusion, l'équation (6) semble globalement fournir des valeurs $(v_2/V_2)_\alpha$ à 0,1 ou 10% près, le plus grand des deux, et bien mieux lorsque $V_1/(n_1 \cdot V_2)$ n'est pas proche de 1 ou lorsque n_1 ou n_2 sont grands. De plus, lorsque n_1 ou n_2 sont grands, $\left(\frac{\chi_{\alpha, n_2-1}^2}{n_2-1} + \frac{\chi_{1-\alpha, n_2-1}^2}{n_2-1}\right)/2$ et $\left(\frac{\chi_{\alpha, n_2, (n_1-1)}^2}{n_2 \cdot (n_1-1)} + \frac{\chi_{1-\alpha, n_2, (n_1-1)}^2}{n_2 \cdot (n_1-1)}\right)/2$ deviennent proches de 1 quel que soit α , et l'équation (6) peut être simplifiée en can $(v_2/V_2)_\alpha = 1 + k_\alpha \cdot SD$.

4.4 Estimation d'une variance avec plus de 2 niveaux imbriqués

En accord avec les principes de l'ANOVA, quand plus de 2 niveaux imbriqués sont utilisés, les équations (1),(3) et (4) doivent être étendues en équations (7) à (11) comme suit :

L'équation (7) est l'extension de l'équation (1), comme suit :

$$v_i = w_i - \frac{v_{i-1}}{n_{i-1}} - \frac{v_{i-2}}{n_{i-1} \times n_{i-2}} - \dots - \frac{v_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1} \quad (7)$$

Où i est le rang du niveau considéré,
 k est le nombre total de niveaux imbriqués,
 V_i est la valeur vraie de la variance du niveau i ,

v_i est l'estimation de V_i ,

$$w_i = \left(\sum_1^{n_i} (m_{ij} - m_i)^2 \right) / (n_i - 1),$$

n_i est le nombre total de séries de résultats utilisés pour calculer v_i .

Par transformation algébrique, l'équation (7) peut aussi être exprimée sous la forme de l'équation (8) comme suit, qui permet de calculer v_i sans qu'il soit besoin de calculer avant toutes les valeurs v_{i-1} à v_1 :

$$v_i = w_i - \frac{\sum \text{Var}(m_{i-2})}{n_k \times n_{k-1} \times \dots \times n_{i-1}} \quad (8)$$

L'équation (9) est l'extension de l'équation (3), comme suit :

$$\frac{v_i}{V_i} \approx \left(1 + \frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right) \times \frac{\chi_{(n_i-1), n_{i+1}, n_{i+2}, \dots, n_k}^2}{(n_i - 1) \times n_{i+1} \times n_{i+2} \times \dots \times n_k} \quad (9)$$

$$- \left(\frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right) \times \frac{\chi_{n_k \times n_{k-1} \times \dots \times (n_{i-1}-1)}^2}{n_k \times n_{k-1} \times \dots \times (n_{i-1} - 1)}$$

L'équation (10) est l'extension de l'équation (4), comme suit :

$$SD\left(\frac{v_i}{V_i}\right) = \left(\left(1 + \frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right)^2 \times \frac{2}{(n_i - 1) \times n_{i+1} \times n_{i+2} \times \dots \times n_k} \right. \quad (10)$$

$$\left. + \left(\frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right)^2 \times \frac{2}{n_k \times n_{k-1} \times \dots \times (n_{i-1} - 1)} \right)^{1/2}$$

L'équation (11) est l'extension de l'équation (5), comme suit :

$$CV_\alpha\left(\frac{v_i}{V_i}\right) \cong \left(1 + \frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right) \quad (11)$$

$$\times \frac{\chi_{1-\alpha, (n_i-1), n_{i+1}, n_{i+2}, \dots, n_k}^2 + \chi_{\alpha, (n_i-1), n_{i+1}, n_{i+2}, \dots, n_k}^2}{2 \cdot (n_i - 1) \times n_{i+1} \times n_{i+2} \times \dots \times n_k}$$

$$- \left(\frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i} \right)$$

$$\times \frac{\chi_{1-\alpha, n_k \times n_{k-1} \times \dots \times (n_{i-1}-1)}^2 + \chi_{\alpha, n_k \times n_{k-1} \times \dots \times (n_{i-1}-1)}^2}{2 \cdot n_k \times n_{k-1} \times \dots \times (n_{i-1} - 1)}$$

Nous pouvons voir que le terme $\frac{V_{i-1}}{n_{i-1} \times V_i} + \frac{V_{i-2}}{n_{i-1} \times n_{i-2} \times V_i} + \dots + \frac{V_1}{n_{i-1} \times n_{i-2} \times \dots \times n_1 \times V_i}$ joue le même rôle que le terme $V_1/(n_1 \cdot V_2)$ dans le cas « 2 niveaux », et les commentaires 1 et 2 du § 4.2 s'appliquent de la même façon. Il s'ensuit que la qualité de l'estimation V_i dépend de beaucoup plus de paramètres que dans le cas à 2 niveaux. Cependant,

le nombre de degrés de liberté augmente beaucoup lorsque i décroît de sorte que, dans la plupart des cas, il est probable que l'influence des variances des niveaux les plus bas est faible.

Comme applications de ces équations :

- ✚ L'exemple 3 décrit le calcul de v_i en utilisant l'équation (8) dans le cas de 3 niveaux imbriqués et où c'est la variance du niveau 2 qui est recherchée. Dans ce cas, l'équation approchée (6) peut être utilisée en remplaçant n_2 par $n_2 \cdot n_3$;
- ✚ L'exemple 4 décrit le calcul de v_i en utilisant l'équation (8) dans le cas où plus de 3 niveaux sont imbriqués ;
- ✚ L'exemple 5 décrit le calcul des centiles et de l'écart-type des estimations de v_i en utilisant les équations (9) et (10) dans le cas où plus de 3 niveaux sont imbriqués.

Exemple 3 :

Considérons le plan d'expériences utilisé dans l'exemple 1, qui serait répété 5 ou 8 ou 13 ou 20 times (par exemple par 5 ou 8 ou 13 ou 20 laboratoires différents). Les calculs qui en découlent sont présentés dans le Tableau 4.

Tableau 4. Exemple de calcul de variances de niveau 2 pour 3 niveaux de variance imbriqués.

α	$(v_2/V_2)_\alpha$ calculé par MMC					$(v_2/V_2)_\alpha = CV_\alpha + k \cdot SD$				
	$n_3 = 1$	$n_3 = 3$	$n_3 = 8$	$n_3 = 13$	$n_3 = 20$	$n_3 = 1$	$n_3 = 3$	$n_3 = 8$	$n_3 = 13$	$n_3 = 20$
0,025	-2,02	-0,42	-0,13	0,11	0,28	-2,03	-0,42	-0,13	0,11	0,28
0,05	-1,54	-0,19	0,05	0,25	0,39	-1,60	-0,20	0,05	0,25	0,39
0,1	-1,00	0,06	0,25	0,42	0,53	-1,09	0,05	0,25	0,41	0,52
0,2	-0,36	0,37	0,50	0,61	0,69	-0,43	0,37	0,50	0,61	0,68
0,5	0,87	0,98	0,99	0,99	0,99	0,91	0,98	0,99	0,99	1,00
0,8	2,29	1,61	1,49	1,39	1,31	2,38	1,63	1,50	1,39	1,31
0,9	3,15	1,98	1,76	1,60	1,50	3,20	1,97	1,77	1,60	1,48
0,95	3,92	2,28	2,00	1,77	1,65	3,90	2,26	1,99	1,78	1,62
0,975	4,64	2,54	2,22	1,94	1,77	4,53	2,52	2,19	1,93	1,75

Sans surprise, la précision de la détermination de V_2 augmente significativement lorsque n_3 augmente.

Exemple 4 :

La MMC a été utilisée pour déterminer les IC95% des estimations des écarts-types d'homogénéité pour les cas où N_p est le nombre de participants, le nombre d'échantillons par participant est 3 et le nombre de répétitions par échantillon est 2. Donc, dans cet exemple, $n_3 = N_p$, $n_2 = 3$, $n_1 = 2$, $V_2 = \sigma_H^2$, $V_1 = \sigma_r^2$. Les « Khi2 values » représentent les IC lorsque $V_1/(n_1 \cdot V_2) \rightarrow 0$.

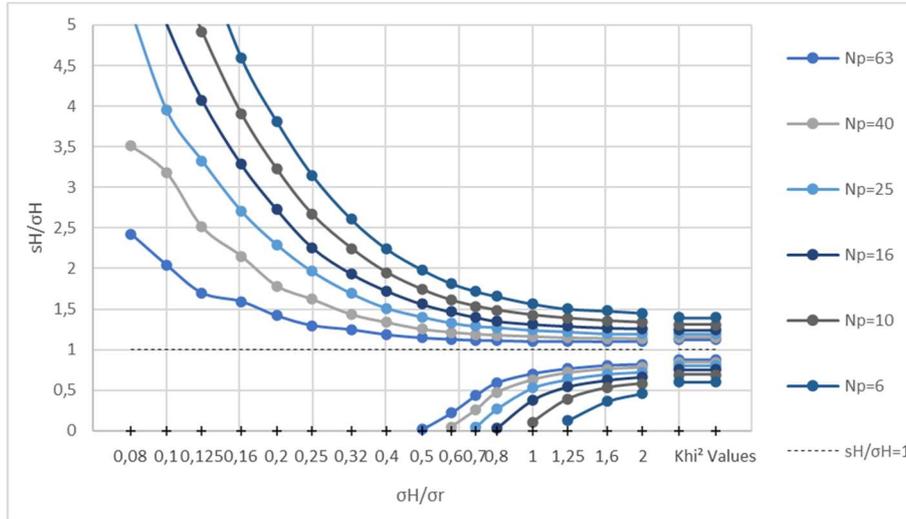


Figure 1: Limites inférieures et supérieures des IC95% de s_H/σ_H pour N_p participants, 3 échantillons par participant, 2 résultats d'essais par échantillon.

Nous pouvons très clairement voir qu'il existe une valeur seuil pour σ_H/σ_r , c.à.d. $V_1/(n_1 \cdot V_2)$, en dessous de laquelle v_H peut devenir négative et s_H doit être arrondi à 0.

Quand une variance d'un niveau intermédiaire doit être estimée, l'équation (1) doit être adaptée comme dans l'exemple 5 :

Exemple 5 :

Plan d'expérience utilisé dans l'exemple 1 avec 5 niveaux et pour lequel on cherche à estimer v_4 comme montré dans le Tableau 4.

Tableau 5. Exemple de calcul de variance du niveau 4 avec 5 niveaux imbriqués et avec $n_1 = 2, n_2 = 3, n_3 = 4, n_4 = 3, n_5 = 2,$

			Niveau 5					
			1			2		
			Niveau 4					
Niv. 3	Niv. 2	Niv. 1	1	2	3	1	2	3
1	1	1	104,5	91,2	96	92,3	86,2	95,2
1	1	2	87	87,3	81,2	95,9	99,1	97,8
1	2	1	111,4	98,1	113,1	105,8	101,8	98,1
1	1	1	103,9	102	102,3	89,8	98,7	100,4
1	1	2	101,9	102,2	99,3	88,8	99	99,7
1	2	1	103,7	101,5	105,8	87,9	99,1	102
1	2	2	101,1	102,8	105,4	90,4	96,3	101,8
1	3	1	102,7	102	99,8	90,1	104,8	102,4
1	3	2	102,8	99,7	98,4	92	104,3	101,6
2	1	1	104,6	98,3	101,6	93,8	97,7	96,9
2	1	2	104,5	98,6	104,3	96,2	100,3	97,1
2	2	1	103,9	98,4	107,2	98,8	98,8	91,1
2	2	2	103,9	99,3	106,6	98,6	97,7	88,7
2	3	1	104	99,8	105,3	98,9	98,6	93,6
2	3	2	101,6	100,7	105,1	97,4	97,8	96,5
3	1	1	94,2	99,2	102,9	96,9	91,2	94,8
3	1	2	95,6	98,4	100,8	97	91,5	93,3

			Niveau 5						
			1			2			
			Niveau 4						
Niv. 3	Niv. 2	Niv. 1	1	2	3	1	2	3	
3	2	1	93,6	96	104,8	96,7	98	93,5	
3	2	2	94	94,2	103,3	97,2	97,9	92	
3	3	1	94,3	99,1	103,7	95,1	93,6	93,6	
3	3	2	93,7	100,1	106,4	95,7	94,6	93,9	
4	1	1	100,9	89,1	102,4	105,2	102,1	98,2	
4	1	2	102,3	88,9	101,7	104,5	103,1	98,3	
4	2	1	99,2	88,9	102,5	101,8	100,2	97,5	
4	2	2	100,3	88,9	103,6	100,7	99,8	96,2	
4	3	1	101,7	94,1	103,5	97,2	100,2	104,9	
4	3	2	99,9	91,4	102,9	98,9	100,6	105,5	
			$m_{3,1}$	102,68	101,7	101,83	89,833	100,37	101,32
			$m_{3,2}$	103,75	99,183	105,02	97,283	98,483	93,983
			$m_{3,3}$	94,233	97,833	103,65	96,433	94,467	93,517
			$m_{3,4}$	100,72	90,217	102,77	101,38	101	100,1
			$w_{3,j}$	18,184	24,449	1,8346	22,875	8,6591	16,422
			$m_{4,j}$	100,35	97,233	103,32	96,233	98,579	97,229
			w_4	9,2534			1,3862		

Les calculs sont comme suit :

$$m_{3,j} = \sum_1^{n_1 n_2} m_{3,k} / n_3.$$

Par exemple, $m_{3,1,1} = (103,9 + 101,9 + 103,7 + 101,1 + 102,7 + 102,8) / 6 = 102,68$

$$w_{3,j} = \text{Var}(m_{3,j}) = \sum_1^{n_3} (m_{3,k} - m_3)^2 / n_3.$$

Par exemple, $w_{3,1,1} = \text{Var}(102,68; 103,75; 94,233; 100,72) = 18,184$

$$w_3 = \sum_1^{n_4 n_5} w_{3,j}^2 / (n_4 \cdot n_5).$$

Par exemple, $w_3 = (18,184^2 + 24,449^2 + 1,8346^2 + 22,875^2 + 8,6591^2 + 16,422^2) / (3 \times 2) = 3,8510$

$$m_{4,j} = \sum_1^{n_1 n_2 n_3} x_k / n_1 n_2 n_3.$$

Par exemple, $m_{4,1} = (103,9 + 101,9 + 103,7 + 101,1 + \dots + 99,2 + 100,3 + 101,7 + 99,9) / 24 = 100,35$

$$w_{4,j} = \text{Var}(m_{4,j}) = \sum_1^{n_4} (m_{4,k} - m_4)^2 / n_4.$$

Par exemple, $w_{4,1} = \text{Var}(100,35; 97,233; 103,32) = 9,2534$

$$w_4 = \text{Var}(m_4) = \sum_1^{n_5} w_{4,j}^2 / n_5.$$

Par exemple, $w_4 = (9,2534^2 + 1,3862^2) / 2 = 5,3198$

$$v_4 = 5,3198 - 3,8510 = 1,4688$$

$$sd_4 = \sqrt{v_4} = 1,2119$$

Il faut noter que l'équation intuitive $v_4 = w_4 - w_3 / (n_1 \cdot n_2 \cdot n_3)$ est fautive et ne fournit pas une estimation sans biais de v_4 même si, dans la plupart des cas, la différence entre les deux n'est pas très importante.

Les mêmes calculs peuvent être réalisés pour trouver que :

$$v_1 = 1,0177$$

$$v_2 = 4,1425$$

$$v_3 = 13,854$$

$$v_5 = 2,5821$$

Exemple 6 :

Calcul d'IC sur v_4 (estimation de V_4) dans le cas où $n_1 = 2$; $V_1 = 0,1$; $n_2 = 3$; $V_2 = 0,2$; $n_3 = 4$; $V_3 = 0,3$; $n_4 = 5$; $V_4 = 0,4$; $n_5 = 3$; $V_5 = 0,5$. Les résultats des calculs qui en découlent sont exposés dans le Tableau 6.

$$SD\left(\frac{v_4}{V_4}\right) = \left(\left(1 + \frac{0,3}{4 \times 0,4} + \frac{0,2}{4 \times 3 \times 0,4} + \frac{0,1}{4 \times 3 \times 2 \times 0,4}\right)^2 \times \frac{2}{(5-1) \times 3} + \left(\frac{0,3}{4 \times 0,4} + \frac{0,2}{4 \times 3 \times 0,4} + \frac{0,1}{4 \times 3 \times 2 \times 0,4}\right)^2 \times \frac{2}{3 \times 5 \times (4-1)} \right)^{1/2}$$

$$= \left(\frac{(1 + 0,2396)^2}{6} + \frac{2 \times 0,2396^2}{45} \right)^{1/2} = 0,5086$$

Tableau 6. Exemple de calculs de variances de niveau 4 avec 5 niveaux imbriqués.

α	0,025	0,05	0,1	0,2	0,5	0,8	0,9	0,95	0,975
CV_α	1,1831	1,1103	1,0418	0,9814	0,9354	0,9814	1,0418	1,1103	1,1831
$(v_4/V_4)_\alpha$ (MMC)	0,21±0,02	0,29±0,02	0,41±0,01	0,57±0,01	0,94±0,01	1,4±0,02	1,67±0,02	1,92±0,03	2,16±0,04
$(v_4/V_4)_\alpha = CV_\alpha + k.SD$	0,1864	0,2738	0,3900	0,5533	0,9354	1,4094	1,6935	1,9468	2,1799

5 Estimation d'une variance avec niveaux imbriqués et rapport inconnu entre elles

5.1 Introduction

Parce qu'elles fournissent une description précise du phénomène, les équations (3) à (6) sont très précieuses sur le plan théorique, mais elles ne sont d'aucune utilité pratique parce que le rapport V_1/V_2 est normalement inconnu lorsque le but est de déterminer une estimation de V_2 . Pour la même raison, les équations (8) à (11) ne peuvent pas non plus être utilisées pour les cas où plus de deux niveaux de variance sont imbriqués

De plus, l'équation (3) montre clairement qu'une seule équation n'est disponible pour une estimation qui requiert la détermination de deux valeurs (càd. V_1 and V_2).

Pour essayer de faire face à cette question, nous avons utilisé l'équation (3) pour produire l'équation (12) à partir de laquelle nous pouvons déduire l'équation (13) par manipulation algébrique.

$$\frac{n_1 \cdot w_2}{v_1} \approx \frac{1 + \frac{V_1}{n_1 \cdot V_2}}{\frac{V_1}{n_1 \cdot V_2}} \cdot F_{(n_2-1, n_2 \cdot (n_1-1))} \quad (12)$$

$$\frac{V_1}{n_1 \cdot V_2} \approx \frac{1}{\frac{n_1 \cdot w_2}{v_1} \times F_{(n_2 \cdot (n_1-1), n_2-1)} - 1} \quad (13)$$

Par conséquent, nous pourrions espérer que le rapport $1/(n_1 \cdot w_2/v_1 - 1)$ est un estimateur du rapport clé $V_1/(n_1 \cdot V_2)$, mais il n'en est pas un bon, en particulier lorsque $V_1/(n_1 \cdot V_2) > 0,1$, càd. dans les cas il serait intéressant de pouvoir l'utiliser. De plus, $1/\left(\frac{n_1 \cdot w_2}{v_1} \times \frac{1}{F_{(n_2-1, n_2 \cdot (n_1-1))}} - 1\right)$ n'est pas une fonction continue (elle diverge pour la valeur α pour laquelle $F_{(n_2-1, n_2 \cdot (n_1-1))} = n_1 \cdot w_2/v_1$, et ses centiles ne peuvent pas être calculés à partir des centiles correspondants de la fonction F. L'utilisation de la MMC est nécessaire pour les déterminer. Il s'en suit également proportion α des estimations est inférieure à 0, lesquelles sont erronées puisque le rapport $V_1/(n_1 \cdot V_2)$ ne peut jamais être inférieur à 0.

Pour toutes ces raisons, l'équation (3) seule ne fournit pas une solution efficace pour déterminer un IC sur V_2 . Pour faire face à cette question, les § 5.2 à 5.4 envisagent plusieurs possibilités qui peuvent être mises en œuvre pour résoudre ce problème, comme suit :

- ✚ Les situations où V_1 peut être considéré comme connu ;
- ✚ Les situations où $v_1/n_1 \ll w_2$, pour lesquelles il est probable que l'influence du terme $V_1/(n_1 \cdot V_2)$ est faible ;
- ✚ Les situations où $v_2 < 0$, pour lesquelles il est probable que V_2 est petit lorsqu'il est comparé à V_1/n_1 ;
- ✚ Les situations où les résultats de l'équation (1) ne sont pas satisfaisants par rapport à l'utilisation qui est prévue pour V_2 .

Pour simplification, les § 5.2 à 5.4 considère le cas de 2 niveaux imbriqués. Bien sûr, pour chacune de ces situations, les conclusions du § 4.3 peuvent être utilisées quand plus de 2 niveaux sont présents dans le processus d'estimation.

5.2 Estimation d'une variance de niveaux imbriqués quand la variance de niveau 1 est connue

Quand V_1 , la variance du niveau 1, est connue, l'équation (3) peut être simplifiée en équation (14) comme suit :

$$\frac{v_2}{V_2} \approx \left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{n_2-1}^2}{n_2 - 1} - \frac{V_1}{n_1 \cdot V_2} \quad (14)$$

Contrairement à l'équation (3), l'équation (14) peut être facilement inverse pour obtenir le rapport V_2/v_2 , qui est le seul utile en pratique, comme suit dans l'équation (15) :

$$\frac{V_2}{v_2} \approx \frac{(n_2 - 1) \cdot \left(1 + \frac{V_1}{n_1 \cdot v_2}\right)}{\chi_{n_2-1}^2} - \frac{V_1}{n_1 \cdot v_2} \quad (15)$$

L'ET des estimations est donc réduit, l'équation (4) devient alors l'équation (16) comme suit :

$$SD\left(\frac{v_2}{V_2}\right) = \left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \sqrt{\frac{2}{n_2 - 1}} \quad (16)$$

Conclusion :

Connaître V_1 est d'une grande aide pour calculer un IC on V_2 car :

1. Une équation est disponible pour le calculer exactement ;
2. L'ET de la distribution des estimations est significativement réduit comparativement aux situations où V_1 est inconnue.

5.3 Estimation d'une variance de niveaux imbriqués quand $v_1/n_1 \ll w_2$

Quand $v_1/n_1 \ll w_2$, il est probable que l'influence du terme $V_1/(n_1 \cdot V_2)$ est faible, même s'il n'est pas impossible que, par hasard, le terme $\left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{n_2-1}^2}{n_2 - 1}$ est dans la partie la plus élevée de son IC et que le terme $\frac{V_1}{n_1 \cdot V_2} \cdot \frac{\chi_{n_2, (n_1-1)}^2}{n_2 \cdot (n_1 - 1)}$ est dans la partie la plus basse de son IC, de sorte que $v_1/(n_1 \cdot v_2) \ll V_1/(n_1 \cdot V_2)$ et par conséquent, l'estimation de v_2/V_2 est de moindre qualité que ce qui est attendu.

Pour vérifier cela, nous avons utilisé la MMC pour générer des valeurs V_1/V_2 aléatoires et nous avons observé les valeurs résultantes de V_2/v_2 .

A ce propos, en utilisant cette méthode et contrairement au cas de l'équation (3), nous avons accès au V_2/v_2 plutôt qu'à v_2/V_2 , qui est le rapport dont nous avons besoin en pratique. Cependant, les résultats dépendent de la distribution des valeurs aléatoires $v_1/(n_1 \cdot v_2)$ qui sont effectivement utilisées pour les calculs. Par conséquent il est nécessaire d'apporter du soin dans le choix de cette distribution. Pour ce qui nous concerne, nous avons sélectionné une distribution log-normale pour les valeurs de V_1/V_2 avec une valeur moyenne et un écart-type égaux à 1, ce qui produit une distribution avec 68% des valeurs d'entrée dans l'intervalle $[0,1;10]$, 95% dans l'intervalle $[0,01;100]$ et 99,7% dans l'intervalle $[0,001;1000]$. Ce choix permet aux calculs de couvrir l'ensemble du champ pour lequel $V_1/(n_1 \cdot V_2) \ll 1 \ll V_1/(n_1 \cdot v_2)$, en ayant à l'esprit que les situations hors de l'intervalle $[0,01;100]$ sont 1 – très rares dans la pratique 2 – bien traitées par l'équation (3) car dans ces cas, l'un des de 2 termes de l'équation est significativement plus grand que l'autre et que la manipulation de l'équation en est grandement simplifiée. Nous pouvons ainsi sélectionner tous les cas où $v_1/n_1 < 0,1 \cdot w_2$ (le rapport 0,1 est apparu comme une limite adéquate pour séparer les cas où $v_1/n_1 \ll w_2$ et les autres) et nous avons vérifié les rapports v_2/V_2 , en fonction de v_2/v_1 . La Figure 2 montre un exemple de résultats de cette mise en œuvre (exprimé en ET et non V), avec $n_1 = 2$, $n_2 = 10$ de l'exemple 1.

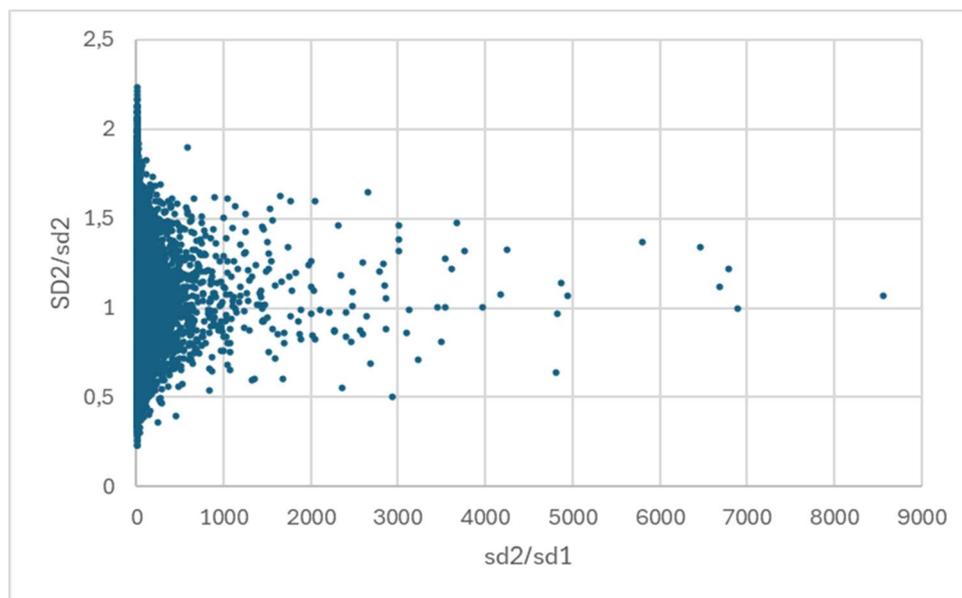


Figure 2: rapports et_2/ET_2 en fonction de et_2/et_1 pour les situations où $v_1/n_1 < 0,1 \cdot w_2$.

Les résultats montrent que, lorsque $v_1/n_1 < 0,1 \cdot w_2$, l'équation (2) peut être utilisée pour évaluer correctement l'IC sur SD_2/sd_2 correspondant. Cependant, en de très rares occasions (non montrées sur la figure pour qu'elle reste lisible) cette équation peut produire des estimations de ET_2 extravagantes (de très grandes valeurs de SD_2/sd_2 , c.à.d. de très fortes sous-estimations de ET_2).

5.4 Estimation d'une variance de niveaux imbriqués quand $v_2 < 0$

Nous avons vu au § 4.2 qu'il arrive que l'équation (3) produise des valeurs négatives v_2 . Nous avons aussi vu dans l'exemple 4 du § 4.4 qu'il existe une valeur seuil T pour $n_1 \cdot V_2/V_1$ associée à une valeur α pour l'IC avec les paramètres n_1 , n_2 , V_1 and V_2 . Cela signifie que, quand la valeur trouvée v_2 est négative, une valeur efficace de la limite supérieure de l'IC sur V_2 peut être calculée à partir de cette valeur seuil. Même s'il est probable que cette valeur calculée de la limite supérieure est très supérieure à sa valeur réelle, dans bcp de cas cela fournit suffisamment d'information pour résoudre le problème pratique pose, voir l'exemple 7 comme suit.

Exemple 7 :

Lors d'une vérification d'homogénéité d'échantillons comme décrit dans la norme ISO 13528 Annexe B [3], il doit être vérifié si la variance d'homogénéité des échantillons est significativement plus faible que la variance interlaboratoires prévue (voir exemple 1). Cependant, dans la plupart des cas, la variance de répétabilité (V_1 dans l'expérience) est souvent significativement plus grande que la variance d'homogénéité (V_2 dans l'expérience). Il arrive donc souvent que $v_2 < 0$, sans possibilité d'avoir une idée de l'IC correspondante. Cependant, s'il apparaît que $V_2 < T \cdot v_1 \ll V_3$, et donc que la condition d'acceptation des échantillons peut être considérée comme satisfaite, même si la valeur vraie de V_2 reste complètement inconnue.

Cette valeur de T peut être calculée en recherchant les 0 de l'équation (3) en fonction de α , c.à.d. les valeurs de α pour lesquelles $\left(1 + \frac{V_1}{n_1 \cdot V_2}\right) \cdot \frac{\chi_{n_2-1}^2}{n_2-1} = \frac{V_1}{n_1 \cdot V_2} \cdot \frac{\chi_{n_2 \cdot (n_1-1)}^2}{n_2 \cdot (n_1-1)}$. La manipulation de cette équation mène à l'équation (17) :

$$T \approx \frac{1}{F_{(n_2 \cdot (n_1-1), n_2-1)} - 1} \quad (17)$$

Exemple 8 :

Lors d'une vérification d'homogénéité d'échantillons comme décrit dans la norme ISO 13528 Annex B [3], les résultats suivants ont été trouvés : $w_2 = \sum_{j=1}^{10} m_{1j}/10 = 0,216$, $v_1 = \sum_{j=1}^{10} v_{1j}/10 = 3,742$, $n_1 = 2$, $v_2 = 0,216 - 3,742/2 = -1,655$, inférieur à 0.

Dans ce cas, $T_{0,95} = 1/(F_{(n_2 \cdot (n_1-1), n_2-1)} - 1) = 1/(3,1373 - 1) = 0,468$. D'autre part, la valeur supérieure de l'IC90(V_1) (où IC90(V_1) est l'intervalle de confiance bilatéral sur V_1) peut être calculé comme $MaxV_1 = v_1/\chi_{n_2 \cdot (n_1-1)}^2 = 3,742/0,394 = 9,50$. Nous pouvons donc en conclure que $V_2 < T \cdot MaxV_1/n_1 = 0,468 \times 9,50/2 = 2,22$ avec un niveau de confiance meilleur que 95%. Comme il est évident que $V_2 > 0$, nous pouvons en déduire que l'IC90(V_2) est meilleur que]0;2,22[.

Cette méthode de calcul semble très conservatrice car elle inclut 2 étapes avec une petite valeur de α . Cependant, elle ne peut être appliquée que quand $v_2 < 0$, c.à.d. dans les cas où $n_1 \cdot w_2/v_1 < 0$. Par conséquent, une covariance s'applique pour les valeurs consécutives de α et le α global est inférieur à $\alpha_1 \cdot \alpha_2$.

5.5 Situations pour lesquelles les résultats d'estimation ne sont pas satisfaisants en regard de son utilisation prévue

Quand aucune des possibilités exposées dans les § 5.2 à 5.4 ne fournit une solution satisfaisante pour déterminer un IC pour une variance imbriquée, une nouvelle expérience avec un nombre n_1 accru doit être envisagé. Augmenter n_1 est le seul moyen de décroître le rapport $V_1/(n_1 \cdot V_2)$ qui commande la possibilité de la déterminer (voir plus haut), car ni V_1 ni V_2 ne sont sous le contrôle de l'expérimentateur.

Réduire $V_1/(n_1 \cdot V_2)$ à une valeur inférieure à 0,1 devrait être le but à atteindre, car nous avons vu au § 5.3 que remplir cette condition assure une détermination adéquate de l'IC.

Le problème est donc de déterminer une valeur optimale d'augmentation de n_1 . Cette nouvelle valeur n_1 doit être suffisamment grande pour obtenir une valeur adéquate $V_1/(n_1 \cdot V_2)$ mais elle doit aussi être compatible avec les conditions techniques, économiques et pratiques dans lesquelles l'expérimentation est conduite. Par exemple, la disponibilité du matériel, le coût des essais ou le temps nécessaire à la réalisation des essais peuvent limiter l'augmentation de la valeur n_1 . Parfois, en fonction des conditions économiques et pratiques de l'expérimentation, il peut être intéressant de réduire la valeur n_2 , de sorte que le nombre total d'essais $n_1 \cdot n_2$ n'augmente pas beaucoup du fait de la modification du schéma d'expérimentation.

Exemple 9 :

Lors d'une vérification d'homogénéité d'échantillons comme décrit dans la norme ISO 13528 Annex B [3], le plan d'expérience peut être change de 10 échantillons et 2 essais par échantillon en 3 échantillons et 7 essais par échantillon, quand cela est techniquement possible (en particulier s'il est possible de réaliser 7 essais sur un même échantillon). Le rapport $V_1/(n_1 \cdot V_2)$ est ainsi augmenté 3,5 fois.

Dans tous les cas, aucune méthode générale ne peut déterminer exactement la nouvelle valeur n_1 qui devrait être choisie car :

1. Le rapport $V_1/(n_1 \cdot V_2)$ ne peut pas être correctement connu dans ces cas (voir l'équation (13)) ;
2. Et les conditions limites techniques et économiques ne peuvent pas être exprimées sous forme d'équations.

Par conséquent, l'expérience technique de l'expérimentateur ou les essais pas à pas sont les meilleurs moyens de résoudre ce problème.

6 Conclusions

Les équations classiques ne fournissent pas de déterminations adéquates de l'IC sur les variances de niveaux emboîtés. Nous pourrions trouver des équations qui décrivent bien les distributions des variances de niveaux emboîtés et leur dispersion, à condition que les valeurs réelles de ces variances soient connues. Les inverser pour exprimer l'IC sur les valeurs réelles des variances en fonction de leurs estimations est malheureusement impossible lorsque les variances des niveaux inférieurs sont inconnues. Toutefois, des équations approchantes peuvent être utilisées lorsque l'on peut s'attendre à ce que l'impact des variances des niveaux inférieurs soit faible. Des équations permettant de le vérifier sont également proposées.

7 Références

- [1] EN 10080:2005, Aciers pour l'armature du béton - Aciers soudables pour béton armé - Généralités
- [2] ISO 5725-2:2019, Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 2: Méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée
- [3] ISO 13528:2022, Méthodes statistiques utilisées dans les essais d'aptitude par comparaison interlaboratoires
- [4] Llyod S. Nelson, "Using nested designs. II Confidence limits for standard deviations", Journal of Quality Technology, Vol. 27 n°3, July 1995
DOI: <https://doi.org/10.1080/00224065.1995.11979599>
- [5] David Luengo, Luca Martino, Mónica Bugallo, Víctor Elvira and Simo Särkkä, "A survey of Monte Carlo methods for parameter estimation" EURASIP Journal on Advances in Signal Processing, Article 25, May 2020
DOI: <https://doi.org/10.1186/s13634-020-00675-6>
- [6] ISO 2854:1976, Interprétation statistique des données — Techniques d'estimation et tests portant sur des moyennes et des variances