

Erratum

“Why psychologists Should by Default Use Welch’s *t*-test Instead of Student’s *t*-test” (Chapitre 2)

Erreurs conceptuelles

- 1) **p.10**: “the *F*-ratio statistic is obtained by computing $SD2/SD1$ (standard deviation ratio, *SDR*)”
 - D’abord, le ratio entre les 2 écart-types d’échantillons ne correspond pas au *SDR*, mais à l’estimation du *SDR* ($SDR = \frac{\sigma_2}{\sigma_1}$).
 - Ensuite, ce n’est pas nécessairement le deuxième écart-type d’échantillon qui se trouve au numérateur, mais le plus grand des deux échantillons (si bien que le *F*-ratio est toujours supérieur à 1; Hayes & Cai, 2007).
- 2) **p.10**: “When $SDR = 1$, the equal variances assumption is true, when $SDR > 1$ the standard deviation of the second **sample population** is bigger than the standard deviation of the first **sample population**, and when $SDR < 1$ the standard deviation of the second **sample population** is smaller than the standard deviation of the first **sample population**”.
- 3) **p.14** (partie “simulations”) : “As long as the variances are equal between ~~groups~~ **populations** or sample sizes are equal, the distribution of Student’s *p*-values is uniform... ”.

Note: à divers autres endroits, j’ai remplacé le mot “groupe” par le mot “population” pour éviter toute ambiguïté (car j’utilisais parfois le mot “groupe” pour décrire les échantillons, parfois pour décrire les populations).

- 4) **p.14** (partie “simulations”) : “Differences are small, except in three scenarios (See table A5.2, A5.5, and A5.6 in the additionnal file). En réalité, la plus grande différence entre les tests se trouve en A5.7 (quand les deux distributions présentent une asymétrie positive) et vaut 4.29%. Dans tous les autres cas, elle est inférieure. Le constat erroné d’après laquelle la différence dans les tables A5.2, A5.5 et A5.6 était plus marquée provenait d’une erreur dans mes scripts, qui a été corrigée (également sur le Supplemental Material attaché en ligne à l’article du chapitre 2).
- 5) **Test de Yuen**: nous sous-entendons à plusieurs reprises que le test de Yuen a un taux inacceptable d’erreur de type I lorsque les distributions de population sont asymétriques. Il s’agit d’une erreur conceptuelle qui a été expliquée dans les limites de la thèse.

Ambiguïtés possibles

- 1) **p.13**: “When both variances and sample sizes are the same in each independent group, the *t*-values, degrees of freedom, and the *p*-values in Student’s *t*-test and Welch’s *t*-test are the same (see Table 1).” Les statistiques, degrés de liberté et *p*-valeurs seront identiques à condition que les **estimations de variances soient identiques** et soient obtenues sur base **d’échantillons de tailles égales**. Or, la phrase peut donner l’impression que ces paramètres seront identiques à condition que la condition d’homoscédasticité soit respectée (ce qui n’est pas vrai). Cette information fournie dans la table est peu pertinente en soi, dans la mesure où très fréquemment, on obtiendra des estimations de variance différentes pour chaque groupe lorsque la condition d’homoscédasticité est respectée (et qu’à l’inverse, il est possible, bien que peu probable, d’obtenir des estimations de variance identiques même en cas d’hétéroscédasticité).

- 2) **p.16** (à propos du test de Levene) : “*Because the statistical power for this test is often low, researchers will inappropriately choose Student’s t -test instead of more robust alternatives.*” Cette phrase peut amener à comprendre que si le test de Levene était toujours très puissant, il serait approprié de l’utiliser en vue de choisir entre les tests t de Student et t de Welch. Pourtant, privilégier le test t de Student lorsque l’on ne peut rejeter l’hypothèse d’égalité des variances (autrement dit, lorsque les résultats du test de Levene sont non significatifs) reviendrait à confondre le non-rejet de l’hypothèse d’égalité des variances avec l’acceptation de l’hypothèse d’égalité des variances (cf. chapitre 5).

Mise en forme et Notations

- 1) Les lettres utilisées pour décrire les statistiques (t ou F , par exemple) doivent toujours être inscrites en *italique*. Il en est de même pour les notations mathématiques. Or, cela a été omis à plusieurs reprises dans l’article. Par exemple, il aurait fallu écrire . . . :

- **p.10** : “... as the Mann-Whitney U -test...”;
- **p.10** : “... the F -ratio statistic...”;
- **p.10** : “... x_{ij} ” au lieu de “ x_{ij} ”;
- **p.10** : “... $|x_{ij} - \hat{\theta}_j|$ ” au lieu de “ $|x_{ij} - \hat{\theta}_j|$ ”.

- 2) Il est très important d’être consistant dans le choix des notations mathématiques, pour éviter toute confusion pour le lecteur. Or, nous ne l’avons pas toujours été. Par exemple, nous avons utilisé plusieurs notations différentes pour décrire l’écart-type et la variance :

- **p.10** : nous utilisons respectivement SD1 et SD2 pour décrire l’écart-type de chaque groupe;
- **p.12** (équation 1) : nous utilisons respectivement S_1^2 et S_2^2 pour décrire la variance de chaque groupe, alors que nous utilisons s_1^2 et s_2^2 (lettres minuscules) dans la légende de cette formule;
- **p.13** (équations 3 et 4) : nous utilisons respectivement s_1^2 et s_2^2 (lettres minuscules) pour décrire la variance de chaque groupe.

- 3) On parle normalement d’erreur de type I et II. Or, dans tout l’article du chapitre 2, j’ai parlé des erreurs de type 1 et 2. Par exemple, p.13: “*An increase in the Type 1 error rate leads to an inflation of... while an increase in the Type 2 error rate...*”

Fautes de frappe

- **p.14** : “see \forall Figure 2a”.
- **p.15** : “ p -values from Welch’s ~~t -test~~ and Student’s t -tests, shown separately...”
- **p.16** : Note 4 : “other variants have been proposed such as the 20% trimmed mean”

Bibliographie

Référence manquante: Bradley (1978).

Annexe B: erratum de l'article "Taking parametric assumptions very seriously : Arguments for the Use of Welch's W -test instead of the Classical F -test in One-Way ANOVA" (Chapitre 3)

Erreur conceptuelle

- **p.22** : "Formula (7) Equation 7 provides the computation of the W -test, or Welch's F -test Welch's statistic (W). In the numerator of the W -test W statistic, the squared deviation between group mean..."
- **p.22** : "... negative pairings (the group with the smallest largest sample size is extracted from the population with the smallest SD);
- **p.25** : "... which is either more liberal or more conservative, depending on the SD s and SD sample sizes pairing";
- **p.27** : "Moreover, there is one constant observation in our simulations : whatever the configuration of the n -ratio, the consistency of the three tests is closer to zero when there is a negative positive correlation between the SD and the mean sample size (meaning that the largest group with the highest mean has the lowest largest variance)."

Mise en forme et Notations

Une légende est manquante pour certaines notations mathématiques. Par exemple, en ce qui concerne l'équation (1), bien que n_j , k et s_j^2 aient été correctement définis, les définitions pour \bar{x}_j , $\bar{x}_{..}$ et N ne sont données que plus tard, en référence à d'autres équations. Cela peut rendre la lecture de l'article plus compliquée pour certaines personnes non familières avec ces notations.

Par ailleurs, comme dans l'article précédent sur le test t de Welch, on constate certaines incohérences en termes de notation. Par exemple, si la moyenne de chaque groupe est définie par \bar{x}_j dans l'équation (1), elle est définie par \bar{X}_j dans l'équation (7).

On a également omis d'italiser certaines lettres statistiques (comme dans les Figures par exemple).

p.21 : ($SD_{spanish} = .80 > SD_{english} = .50$) ($S_{spanish} = .80 > S_{english} = .50$, with S = sample standard deviation) [...] For men, the reverse was true ($SD_{spanish} = .97 < SD_{english} = 1.33$) ($S_{spanish} = .97 < S_{english} = 1.33$)

Enfin, dû à un manque de connaissance de Latex lors de mes premières tentatives d'écritures d'articles via Rmarkdown, certaines majuscules sont manquantes dans les références bibliographiques.

Faute(s) de frappe et grammaire et autres

- **p.20** : "which can leads to asymmetry in the distribution";
- **p.20** : "... we think that a first realistic first step towards progress would be to get researchers...";
- **p.21** : "Based on mathematical explanations and Montee Carlo simulations";
- **p.22** : "With with $N = \dots$ ";
- **p.22** : "Where where \bar{x}_j and s_j^2 are respectively the group mean and the group variance...";
- **p.23** : "the type I error rate of all tests";
- **p.23** : "When there are more than three two groups";
- **p.24** : "In Figures 2 to 6 (see Figure 1 for the legend)";
- **p.24** : "... whatever the correlation between the SD and the mean does not matter";

Annexe C: erratum de l'article "Why Hedge's g^* based on the non-pooled standard deviation should be reported with Welch's t -test" (Chapitre 4)

- **p.39:** titre de la Table 1: "~~Expectency~~ **Expectation**, bias and variance of Cohen's d_s d "
- **p.39:** dans cette même table, les indices " s " doivent être supprimés dans la première colonne.
- **p.41:** les mêmes remarques (concernant le titre et les indices) s'appliquent à la table 2.
- **p.46:** les mêmes remarques (concernant les indices) s'appliquent à la table 3. De plus, il convient d'écrire "~~Cohen's~~ **Hedges' g^*** ".