# STT-1920: Méthodes statistiques

## Exercices additionnels sur l'analyse de la variance

NUMÉRO 1. On considère un problème d'anova à un facteur avec 4 niveaux. Voici nos moyennes échantillonnales:

$$\overline{y}_{1\cdot} = 23.5 \qquad \overline{y}_{2\cdot} = 17.2 \qquad \overline{y}_{3\cdot} = 25.7 \qquad \overline{y}_{4\cdot} = 29.7$$

La moyenne échantillonnale globale $\overline{y}_{\cdot\cdot}$ peut-elle être égale à 17.3? Peut-elle être égale à 16.8? Expliquez.

NUMÉRO 2. [Suite du numéro 1] Supposons qu'au numéro 1 les tailles des échantillons sont

$$n_1 = 7 \qquad n_2 = 18 \qquad n_3 = 11 \qquad n_4 = 8.$$

Calculez la moyenne échantillonnale globale $\overline{y}_{\cdot\cdot}$. Calculez la somme des carrés inter-groupes $SST_R$. Calculez la moyenne des carrés inter-groupes $MST_R$.

NUMÉRO 3. [Suite des numéros 1 et 2] Voici les variances échantillonnales:

$$s_1^2 = 4.30 \qquad s_2^2 = 6.15 \qquad s_3^2 = 8.84 \qquad s_4^2 = 3.61.$$

Calculez l'erreur quadratique moyenne $MSE$.

NUMÉRO 4. [Suite des numéros 1, 2 et 3] Calculez la statistique de Bartlett. Que peut-on conclure?

NUMÉRO 5. [Suite des numéros 1, 2, 3 et 4] En supposant l'homogénéité des variances théoriques, obtenez un intervalle de confiance de niveau 90% pour $\sigma^2$, la variance théorique commune aux quatre populations.

NUMÉRO 6. [Suite des numéros 1, 2, 3, 4 et 5] Toujours en supposant l'homogénéité des variances théoriques, obtenez un intervalle de confiance de niveau 95% pour $\frac{\mu_1+\mu_3+\mu_4}{3} - \mu_2$.

NUMÉRO 7. On fait une anova pour comparer 5 populations. Nos tailles déchantillons sont toutes égales à 10. Quels sont l'espérance et l'écart-type de la statistique $MST_R/MSE$?

NUMÉRO 8. [Suite du numéro 7] On obtient $SST_R = 109.2$ et $SST = 820.7$. Complétez la table d'anova. Avec le logiciel R ou avec une calculatrice scientifique munie de la loi de Fisher, obtenez le *p-value*. Tracez le graphe de la loi de Fisher et illustrez votre *p-value* sur ce graphe.

NUMÉRO 9. Dans un problème à deux échantillons indépendants issus de populations avec lois normales de même variance, on veut tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$. Le statisticien A affirme qu'on devrait utiliser le test de Student:

$$\text{On rejette } H_0 \text{ si } \left| \frac{\overline{Y}_1 - \overline{Y}_2}{S_c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \geq t_{n_1+n_2-2,\alpha/2}.$$

Le statisticien B affirme qu'on devrait utiliser le test de Fisher:

$$\text{On rejette } H_0 \text{ si } \frac{MST_R}{MSE} \geq F_{I-1,N-I,\alpha}$$

avec $I = 2$. Qui a raison?

Numéro 10. Dix-sept patients ont participé à une expérience visant à comparer 3 médicaments pour réduire la pression sanguine. On a mesuré la baisse de pression après 2 semaines. Voici les résultats:

Médicament A: 25.6, 32.1, 29.3, 28.8, 22.7

Médicament B: 31.6, 28.1, 23.4, 33.3, 24.8, 28.6

Médicament C: 28.7, 36.4, 31.3, 39.5, 33.7, 36.2

Ces trois médicaments sont-ils tous aussi bons les uns que les autres? Faites le test approprié au seuil 5%

Numéro 11. [Suite du numéro 10]. L'hypothèse d'égalité des variances est-elle raisonnable? Faites le test de Bartlett au seuil 5%.

Numéro 12. [Suite des numéros 10 et 11]. L'hypothèse de normalité est-elle raisonnable? Obtenez les 17 résidus et faites les analyses appropriées.

Numéro 13. On a fait une expérience avec 4 types d'engrais (a, b, c et d) et 3 espèces de blé d'Inde (I, II et III). Pour notre expérience, on disposait de 32 plants de l'espèce I, 32 plants de l'espèce II et 32 plants de l'espèce III. Pour chacune des 3 espèces, on a formé 4 groupes de 8 plants et chaque groupe a reçu un des 4 engrais. À la fin de l'expérience, on a mesuré la hauteur des plants en cm. On a donc en tout 96 observations c'est-à-dire 8 obvervations pour chacune des 12 combinaisons possibles engrais-espèce. Les données sont dans le fichier Excel `numero-13.xls` disponible sur le site web du cours. Avec R-Commander, faites l'anova à deux facteurs. Y a-t-il un effet engrais? Y a-t-il un effet espèce? Y a-t-il un effet d'interaction?