



UNIVERSITÉ DE MONTPELLIER
Département de Mathématiques Appliquées

Outstanding 3 : Correction de test multiple et calcule de puissance

Atelier Projet — HAX916X

Réalisé par :
DIALLO Ousmane
ATTOUMANI Ibrahim



Année Universitaire 2025 – 2026

Table des matières

1. Simulation de données	2
1.1. Moyenne et erreurs type	2
1.2. Calcule de la puissance d'un test	3
1.3. Correction de test multiples	7
2. Annexe	9
2.1. Script: Données simulés	9
2.2. Script: Moyenne et erreurs type	11
2.3. Modélisation: Modèle mixte sous R	11

1. Simulation de données

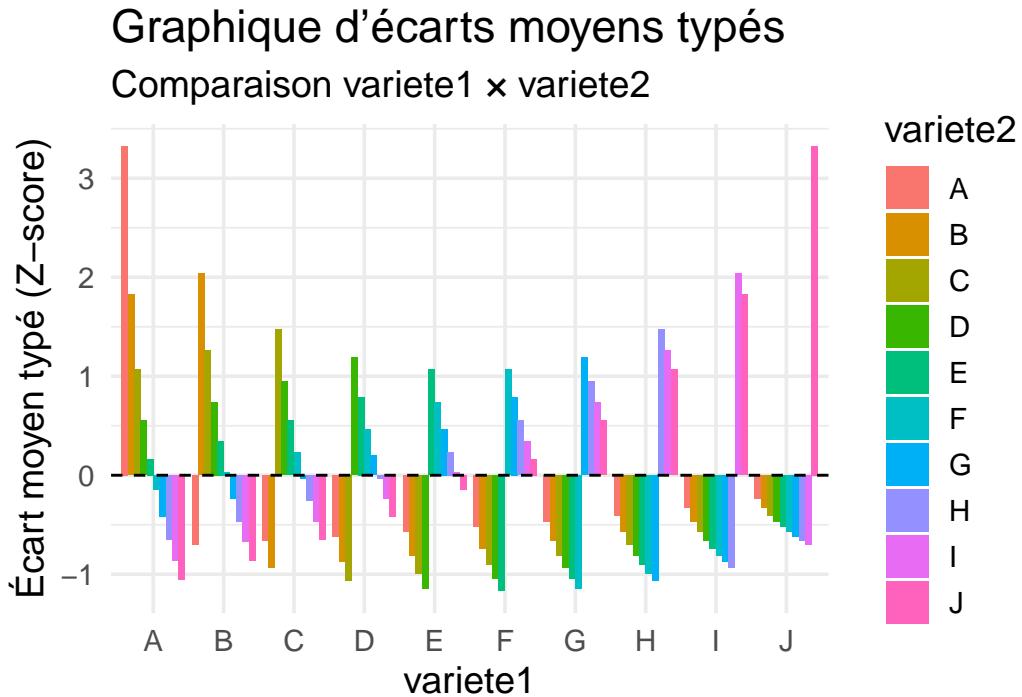
Le jeu de données simulé par le **script R** en annexe contient les mesures de croissance de différentes variétés d'avoine, seules ou en combinaison, sur trois répétitions expérimentales. Les variables **var1** et **var2** indiquent les variétés présentes dans chaque traitement, tandis que les variables numériques enregistrent la taille, le nombre de feuilles et la biomasse. Ce cadre permet d'étudier l'effet des interactions entre variétés sur la biomasse ou la taille, en tenant compte de la variabilité entre répétitions grâce à un modèle mixte.

ID	var1	var2	taille1_t1	taille1_t2	taille1_t3	taille2_t1	taille2_t2	taille2_t3
AA_1	A	A	3.5	6.5	10	1	3	7.5
BB_1	B	B	3.5	5.5	8	2	3	7.5
CC_1	C	C	0.5	5.5	9	3	6	7.5
DD_1	D	D	3.5	4.5	9	2	3	6.5
EE_1	E	E	1.5	4.5	8	3	2	6.5
FF_1	F	F	1.5	4.5	10	1	3	7.5
GG_1	G	G	2.5	4.5	10	2	4	7.5
HH_1	H	H	1.5	4.5	8	1	6	7.5
II_1	I	I	1.5	6.5	9	1	5	6.5
JJ_1	J	J	1.5	6.5	9	1	4	8.5
AB_1	A	B	1.5	5.5	9	3	5	8.5
AC_1	A	C	2.5	6.5	9	1	4	7.5
AD_1	A	D	2.5	3.5	10	2	4	6.5
AE_1	A	E	2.5	6.5	10	1	3	6.5
AF_1	A	F	2.5	4.5	8	2	6	8.5
AG_1	A	G	0.5	6.5	10	2	5	7.5
AH_1	A	H	0.5	4.5	9	3	5	8.5
AI_1	A	I	3.5	3.5	8	2	4	8.5
AJ_1	A	J	3.5	6.5	10	3	3	6.5
BC_1	B	C	1.5	4.5	10	1	3	8.5
BD_1	B	D	3.5	4.5	8	2	5	8.5
BE_1	B	E	1.5	6.5	9	2	2	8.5

Table 1: Extrait de 22 lignes du jeu de données sur la croissance d'avoine.

1.1. Moyenne et erreurs type

Dans cette partie, nous allons créer un graphique affichant les écarts à la moyenne des variables à l'aide de **ggplot2**. Ce graphique permettra d'observer, avant modélisation, si les variables présentent des relations ou des tendances entre elles, facilitant ainsi l'interprétation des interactions potentielles. Le script affichant le graphique des écart à la moyenne sera en annexe.



Le graphique ci-dessus comparant des écarts moyens typés (sous forme de Z-score) entre différentes variétés étiquetées de A à J. L'axe des abscisses indique la *variete1*, tandis que chaque couleur correspond à une *variete2* différente.

Pour chaque *variete1*, les barres montrent l'écart moyen typé par rapport à chacune des autres *variete2*. Une valeur positive indique que *variete1* présente, en moyenne, des valeurs plus élevées que *variete2*, tandis qu'une valeur négative montre l'inverse.

Par exemple, pour la colonne correspondant à *variete1* = *J*, la comparaison avec *variete2* = *J* (barre rose) montre un écart fortement positif, tandis que pour *variete2* ≠ *J*, les écarts sont négatifs.

Le graphique permet d'identifier facilement les variétés généralement meilleures ou moins performantes par rapport aux autres.

1.2. Calcule de la puissance d'un test

On considère deux échantillons i.i.d. :

$$\begin{cases} X_{1,n} = (X_{1,1}, \dots, X_{1,n}) & \text{la biomasse de la variété 1,} \\ X_{2,n} = (X_{2,1}, \dots, X_{2,n}) & \text{la biomasse de la variété 2} \end{cases}$$

d'espérance respective m_1 et m_2 .

Les tests suivants seront basés sur le TCL et seront asymptotiques, c'est-à-dire applicables si n est assez grand ($n > 30$).

Soit $(\hat{\sigma}_{1,n}^2, \hat{\sigma}_{2,n}^2)$ les variances respectives de la biomasse des variétés 1 et 2 telles que :

$$\hat{\sigma}_{1,n}^2 \xrightarrow{n \rightarrow \infty} \sigma_1^2, \quad \hat{\sigma}_{2,n}^2 \xrightarrow{n \rightarrow \infty} \sigma_2^2.$$

Dans la suite, si $\hat{\sigma}_{1,n}^2$ et $\hat{\sigma}_{2,n}^2$ sont connues, on prendra directement :

$$(\hat{\sigma}_{k,n}^2)_{k=1,2} = (\sigma_{k,n}^2)_{k=1,2}.$$

Dans le cas contraire, on estimera celles-ci :

$$\hat{\sigma}_{k,n}^2 = \frac{1}{n} \sum_{i=1}^n (X_{k,i} - \bar{X}_{k,n})^2, \quad k = 1, 2.$$

On sait que :

$$\mathbb{E}[\bar{X}_{1,n}] = m_1, \quad \text{Var}(\bar{X}_{1,n}) = \frac{\hat{\sigma}_{1,n}^2}{n},$$

$$\mathbb{E}[\bar{X}_{2,n}] = m_2, \quad \text{Var}(\bar{X}_{2,n}) = \frac{\hat{\sigma}_{2,n}^2}{n}.$$

On pose: $X_n = X_{1,n} - X_{2,n}$ avec $X_{1,n} \perp\!\!\!\perp X_{2,n}$.

C'est-à-dire:

$$\mathbb{E}[\bar{X}_n] = m_1 - m_2 = m,$$

et

$$\text{Var}(\bar{X}_n) = \text{Var}(\bar{X}_{1,n}) + \text{Var}(\bar{X}_{2,n}) = \frac{\hat{\sigma}_{1,n}^2 + \hat{\sigma}_{2,n}^2}{n}.$$

On note :

$$\frac{S^2}{n} = \frac{\hat{\sigma}_{1,n}^2 + \hat{\sigma}_{2,n}^2}{n}.$$

Par le TCL :

$$\frac{\sqrt{n}(\bar{X}_n - m)}{S} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

considérons le teste d'hypothèse sur la moyenne suivant:

$$(H_0) : m_1 = m_2 \quad \text{vs} \quad (H_1) : m_1 \neq m_2$$

c'est à dire :

$$(H_0) : \mu = m_1 - m_2 = 0 \quad \text{vs} \quad (H_1) : \mu = m_1 - m_2 = \delta \neq 0$$

On a alors:

$$\begin{aligned} Z_n &= \sqrt{n} \frac{\bar{X}_n - \mu_0}{S} = \sqrt{n} \frac{\bar{X}_n - \mu}{S} + \sqrt{n} \frac{\mu - \mu_0}{S} \\ &\iff Z_n \sim \mathcal{N}(M, 1), \quad \text{avec } M = \sqrt{n} \frac{\mu - \mu_0}{S} \end{aligned}$$

La puissance du test est la probabilité d'accepter (H_1) lorsque celui-ci a raison.

D'où:

$$\begin{cases} \mu = \mu_0 = 0 \iff Z_n \sim \mathcal{N}(0, 1), & \text{sous } (H_0) \\ \mu \neq \mu_0 = \delta \iff Z_n \sim \mathcal{N}(M, 1), & \text{sous } (H_1) \end{cases}$$

Le test bilatéral de niveau $\alpha = 5\%$ rejette (H_0) si $|Z_n| > z_{1-\alpha/2}$.

La puissance (probabilité de rejeter (H_0) sous (H_1)) s'écrit exactement :

$$\begin{aligned} 1 - \beta &= P(Z_n > z_{1-\alpha/2} \mid H_1) + P(Z_n < -z_{1-\alpha/2} \mid H_1) \\ &= 1 - \Phi(z_{1-\alpha/2} - M) + \Phi(-z_{1-\alpha/2} - M), \end{aligned}$$

où Φ est la Fonction de répartition de la loi $\mathcal{N}(0, 1)$.

Une approximation courante (utilisée pour obtenir une formule explicite pour n) consiste à imposer que la queue droite donne la probabilité $1 - \beta$.

$$P(Z_n > z_{1-\alpha/2} \mid H_1) \approx 1 - \beta,$$

ce qui revient à résoudre

$$\begin{aligned} 1 - \Phi(z_{1-\alpha/2} - M) &= 1 - \beta \\ \iff \Phi(z_{1-\alpha/2} - M) &= \beta \\ \iff z_{1-\alpha/2} - M &= z_\beta \\ \iff z_{1-\alpha/2} - M &= -z_{1-\beta} \end{aligned}$$

D'où

$$M = z_{1-\alpha/2} + z_{1-\beta}.$$

En remplaçant $M = \sqrt{n} \frac{\delta}{S}$, (où $\delta = m_1 + m_2$) on obtient:

$$n \approx \frac{S^2}{\delta^2} (z_{1-\alpha/2} + z_{1-\beta})^2$$

Cette formule est largement utilisée et donne une bonne estimation de l'ordre de grandeur de n .

Application directe sous R:

```
# individus avant repetition
ind = length(traitements)

# echantillon biomasse
X1 <- data$b_masse1; X2 <- data$b_masse2

# moyene
m1 <- mean(X1); m2 <- mean(X2); delta=m1-m2

# variance
hat_sigma1 <- var(X1); hat_sigma2 <- var(X2)
S = sqrt(hat_sigma1 + hat_sigma2)

# alpha et beta
alpha <- 0.05; beta <- 0.10

# quantile d'ordre 1-alpha/2 et 1-beta de N(0.1)
z1_alpha2 <- qnorm(1-(alpha/2)); z1_beta <- qnorm(1-beta)
z <- z1_alpha2 + z1_beta

# la taille de l'echantillon
n = round((S^2/delta^2)*z^2)

k = round(n/ind)

ind; n; k
```

Avec 55 **traitements uniques**, pour atteindre une **puissance statistique de 0.90**, il est nécessaire d'avoir un **total de 822 individus**, ce qui correspond à **environ 15 répétitions par traitement**.

1.3. Correction de test multiples

Lorsqu'un grand nombre de tests statistiques sont réalisés simultanément, le risque global d'erreur de type~1 (**erreur de première espèce**) augmente. Pour contrôler ce risque, il est nécessaire d'appliquer une correction dite *de tests multiples*. Parmi les méthodes existantes, on trouve par exemple Holm, Hochberg, ou encore la procédure de Benjamini–Hochberg (FDR). Dans cette étude, nous présentons la méthode de Bonferroni, qui constitue l'approche la plus simple et la plus conservatrice.

Méthode de Bonferroni

La correction de Bonferroni consiste à ajuster le seuil de signification~ α en le divisant par le nombre total de tests~ m :

$$\alpha_{\text{corr}} = \frac{\alpha}{m}.$$

Ainsi, un test individuel n'est déclaré significatif que si sa p-valeur vérifie :

$$p < \frac{\alpha}{m}.$$

Cette méthode garantit un contrôle strict du taux d'erreur de type-I pour l'ensemble des tests, mais elle peut devenir très conservatrice lorsque~ m est élevé.

Supposons qu'on a réalisé une ANOVA sur la biomasse1 en fonction de la variété1 (**var1**), puis que nous avons appliqué les corrections de Bonferroni et de Holm sur les p-valeurs obtenues lors des comparaisons multiples entre les niveaux du facteur.

La commande **p.adjust** permet d'obtenir les p-valeurs ajustées selon la méthode souhaitée. La significativité des comparaisons peut ensuite être évaluée en les comparant au seuil de décision~ $\alpha = 0.05$.

```
# p-valeurs initiales
pvals <- c(
  AA = 0.004,
  AB = 0.0048,
  BC = 0.78,
  BB = 0.19,
  CD = 0.89,
  DD = 0.0000099
)

# Correction de Bonferroni
pvals_bonf <- p.adjust(pvals, method = "bonferroni")

# Correction de Holm
pvals_holm <- p.adjust(pvals, method = "holm")
```

Affichage

pvals_bonf
pvals_holm

Les p-valeurs initiales des comparaisons multiples pour var1 sont : $AA = 0.004$, $AB = 0.0048$, $BC = 0.78$, $BB = 0.19$, $CD = 0.89$, $DD = 9.9 \times 10^{-6}$.

Après correction de Bonferroni : $AA = 0.024$, $AB = 0.0288$, $BC = 1.00$, $BB = 1.00$, $CD = 1.00$, $DD = 5.94 \times 10^{-5}$.

Avec la correction de Holm : $AA = 0.02$, $AB = 0.02$, $BC = 1.00$, $BB = 0.57$, $CD = 1.00$, $DD = 5.94 \times 10^{-5}$.

AA , AB et DD restent significatives pour un $\alpha = 0.05$ après ajustement, les autres ne le sont pas.

La correction de Bonferroni est plus conservatrice que Holm, donnant des p-valeurs légèrement plus élevées.

2. Annexe

2.1. Script: Données simulés

```
set.seed(0808) # Pour reproductibilité

# Définition des variétés
varietes <- LETTERS[1:10]

# Toutes les combinaisons
pairs <- combn(varietes, 2, repeats.allowed = TRUE)

# éviter A-B et B-A en double
veg_pairs <- apply(pairs, 2, function(x) paste(sort(x), collapse = ""))

# variété avec elle-même
aut_pairs <- paste0(varietes, varietes)

# paires uniques + confrontations avec soi-même
traitements <- unique(c(aut_pairs, veg_pairs))

# Répétitions
rep <- 1:3

# Création des vecteurs pour le dataframe
ID <- c()
Traitement <- c()
var1 <- c()
var2 <- c()

# Boucle pour avoir toutes les obs de la répétition 1, puis 2, puis 3
for(r in rep){
  for(tr in traitements){
    ID <- c(ID, paste0(tr, "_", r))
    Traitement <- c(Traitement, tr)

    # Définir var1 et var2
    var1 <- c(var1, substr(tr,1,1))
    var2 <- c(var2, substr(tr,2,2))
  }
}

# Nombre total d'observations
n <- length(ID)
```

```
# Simulation des variables numériques et qualitatives

nb_fille1_t1 <- sample(1:6, n, replace=TRUE)
nb_fille1_t2 <- sample(6:9, n, replace=TRUE)
nb_fille1_t3 <- sample(9:12, n, replace=TRUE)
nb_fille2_t1 <- sample(2:7, n, replace=TRUE)
nb_fille2_t2 <- sample(7:10, n, replace=TRUE)
nb_fille2_t3 <- sample(10:12, n, replace=TRUE)
taille1_t1 <- sample(0.5:4, n, replace=TRUE)
taille1_t2 <- sample(3.5:7, n, replace=TRUE)
taille1_t3 <- sample(8:10, n, replace=TRUE)
taille2_t1 <- sample(1:3.5, n, replace=TRUE)
taille2_t2 <- sample(2:6, n, replace=TRUE)
taille2_t3 <- sample(6.5:9, n, replace=TRUE)
b_masse1 <- round(runif(n, 0.5,5),2)
b_masse2 <- round(runif(n, 0.5,5),2)

# Création du dataframe final
data <- data.frame(
  var1,
  var2,
  taille1_t1,
  taille1_t2,
  taille1_t3,
  taille2_t1,
  taille2_t2,
  taille2_t3,
  nb_fille1_t1,
  nb_fille1_t2,
  nb_fille1_t3,
  nb_fille2_t1,
  nb_fille2_t2,
  nb_fille2_t3,
  b_masse1,
  b_masse2
)
row.names(data)=ID

View(data)
```

2.2. Script: Moyenne et erreurs type

```

# Table de contingence
tab <- table(data$var1, data$var2)

# -----
# Calcul des écarts à la moyenne théorique
# -----
# Effectifs théoriques si indépendance
theo <- chisq.test(tab)$expected

# Ecart-type global (pour standardisation)
sd_global <- sqrt(theo)

# Ecart moyen type = (observé - théorique) / écart-type
ecarts <- (tab - theo) / sd_global

# Format long pour ggplot
library(reshape2)
ecarts_df <- melt(ecarts)
colnames(ecarts_df) <- c("variete1", "variete2", "EcartZ")

# Graphe d'écarts moyens
library(ggplot2)

ggplot(ecarts_df, aes(x = variete1, y = EcartZ, fill = variete2)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(
    title = "Graphique d'écarts moyens typés",
    subtitle = "Comparaison variete1 × variete2",
    y = "Écart moyen typé (Z-score)",
    x = "variete1"
  ) +
  theme_minimal(base_size = 14)

```

2.3. Modélisation: Modèle mixte sous R

```

# Identifier les variables qualitatives
vars_quali <- c("var1", "var2")

# selectionner les variables
df_qual <- data[ , vars_quali]
df_quant <- data[ , !(names(data) %in% vars_quali)]

```

```
# Créer les variables indicatrices
df_qual <- model.matrix(~ . - 1, data = df_qual)

# Fusionner les deux
df <- cbind(df_quant, df_qual)

# Transformer en data.frame
df <- as.data.frame(df)

data$rep <- rep(1:3, each = length(traitements))
data$var2 <- factor(data$var2)
data$var1 <- factor(data$var1)

library(lme4)
model <- lmer(b_masse1 ~ var1 * var2 + (1|rep), data=data)
summary(model)
```

References

- [1] Horizons Hémato, *Ajustement pour la multiplicité des tests – Quand est-ce nécessaire ou pas?*, Supplément n°2 au Volume9, numéro3, France, 2019. Disponible en ligne : https://horizonshemato.com/wp-content/uploads/2019/09/EHA_2019_-Ajustement-pour-la-mutliplicitÃl-des-tests-Quand-est-ce-nÃlcessaire-ou-pas-.pdf, consulté le November 25, 2025.