

# Auto-évaluation – Blocs complets aléatoires

## Solutions

1. Donnez un avantage et un inconvénient associés à l'utilisation du dispositif en blocs complets aléatoires.

*Réponse :*

Le dispositif en blocs complets aléatoires comporte des avantages tels que :

- il contrôle l'hétérogénéité environnementale (différences entre blocs) ;
- l'ANOVA en blocs complets aléatoires est plus puissante que l'ANOVA complètement aléatoire en présence de gradients ;
- il est logistiquement intéressant lorsque la répétition est contrainte dans l'espace ou le temps ;
- le bloc peut être un individu soumis à différents traitements ou des individus provenant d'une même portée (famille).

Les désavantages sont les suivants :

- l'ANOVA en blocs complets aléatoires est moins puissante que l'ANOVA à un critère lorsque l'effectif ( $n$ ) est petit et que l'effet du bloc est faible ;
- la dépendance potentielle des observations soumises aux différents traitements d'un même bloc dans des blocs trop petits ;
- il ne permet pas de données manquantes ;
- il ne permet pas de tester l'interaction entre le bloc et le facteur d'intérêt.

2. Voici le jeu de données `Temps.txt` d'un design en blocs complets aléatoires destiné à évaluer le temps (jours) que mettent les invertébrés (coléoptères, vers de terre, nématodes) à

recoloniser des parcelles de 1 x 1 m qui ont reçu un de deux types d'herbicides ou un contrôle (sans application d'herbicide) le long d'un gradient d'humidité du sol (non mesuré). Les blocs sont disposés perpendiculairement à ce gradient.

a) Effectuez une ANOVA en blocs complets aléatoires. Si vous rejetez  $H_o$  pour le facteur herbicide, déterminez où se trouvent les différences. Interprétez les résultats et présentez les résultats sous forme graphique.

*Réponse :*

On importe le jeu de données :

```
> ##importation
> temps <- read.table("Temps.txt", header = TRUE)
> ##premières observations
> head(temps)

  Temps Herbicide Bloc
1  4.59  Controle    1
2  8.15    Type1    1
3  8.06    Type2    1
4  7.38    Type1    2
5  4.88  Controle    2
6  5.85    Type2    2

> ##structure du jeu de données
> str(temps)

'data.frame':      15 obs. of  3 variables:
 $ Temps      : num  4.59 8.15 8.06 7.38 4.88 5.85 5.89 3.29 2.74 5.46 ...
 $ Herbicide: Factor w/ 3 levels "Controle","Type1",...: 1 2 3 2 1 3 2 1 3 1 ...
 $ Bloc       : int   1 1 1 2 2 2 3 3 3 4 ...

> ##niveau du facteur
```

```
> levels(temps$Bloc)
```

```
NULL
```

On remarque que la variable Bloc n'est pas reconnue comme facteur et on doit la convertir en facteur avant de continuer :

```
> temps$Bloc <- as.factor(temps$Bloc)
```

```
> ##maintenant reconnu comme facteur
```

```
> levels(temps$Bloc)
```

```
[1] "1" "2" "3" "4" "5"
```

On note que le dispositif est complètement balancé (aucune donnée manquante) :

```
> ##nombre d'observations par combinaison de facteurs
```

```
> tapply(X = temps$Temps, INDEX = list(temps$Herbicide, temps$Bloc),
```

```
      FUN = length)
```

```
      1 2 3 4 5
```

```
Controle 1 1 1 1 1
```

```
Type1    1 1 1 1 1
```

```
Type2    1 1 1 1 1
```

Le diagramme de boîtes et moustaches montre les données brutes (fig. 1).

On exécute l'ANOVA en blocs complets aléatoires :

```
> ##ANOVA à deux critères sans répétitions
```

```
> aov.bloc <- aov(Temps ~ Herbicide + Bloc, data = temps)
```

Les suppositions de l'ANOVA en blocs complets aléatoires sont respectées (fig. 2).

```
> boxplot(Temps ~ Herbicide, data = temps,  
          ylab = "Temps (jours)",  
          xlab = "Herbicide")
```

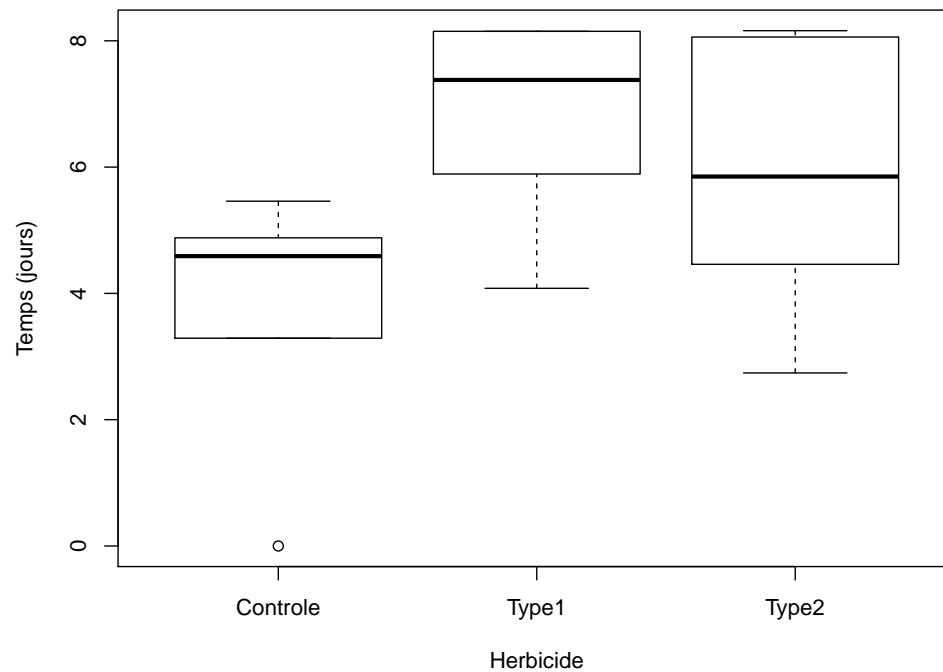


FIGURE 1 – Diagramme de boîtes et moustaches présentant les données de temps de recolonisation de parcelles ayant reçu un traitement d’herbicide ou non.

```
> ##préparation de la fenêtre graphique  
> ##pour accomoder 4 graphiques sur la même page  
> par(mfrow = c(2, 2))  
> ##présentation de quatre graphiques diagnostiques  
> plot(aov.bloc)
```

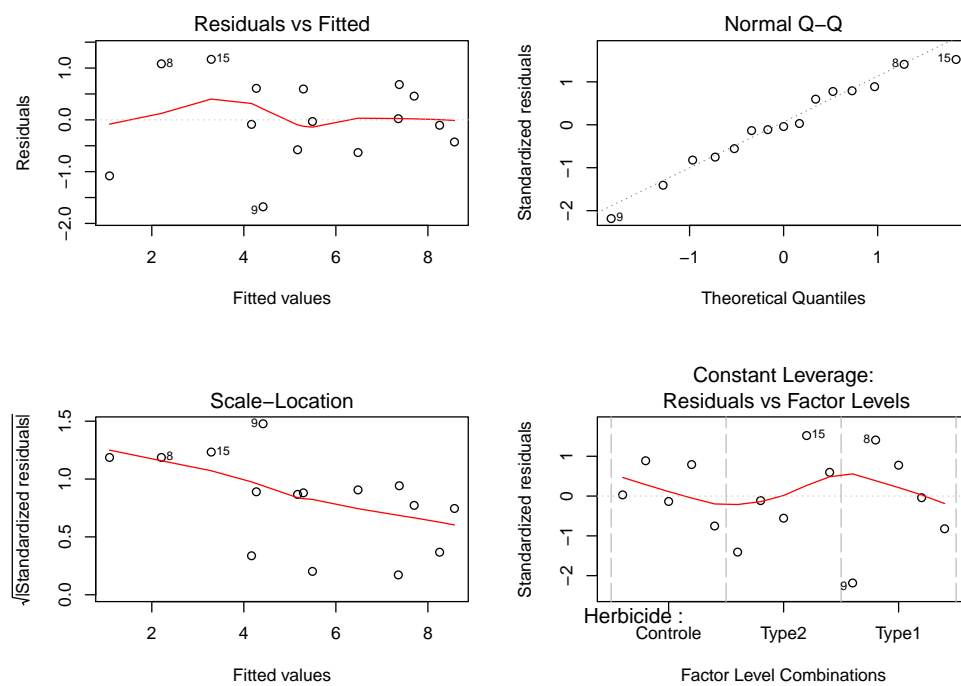


FIGURE 2 – Diagnostics des suppositions d’homogénéité de la variance et de normalité des résidus à la suite de l’ANOVA en blocs complets aléatoires.

Le tableau d'ANOVA nous indique :

```
> summary(aov.bloc)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Herbicide	2	25.3	12.65	11.4	0.0045 **
Bloc	4	44.3	11.07	10.0	0.0033 **
Residuals	8	8.9	1.11		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

On conclut que les blocs et le type d'herbicide ont un effet sur le temps de recolonisation. On peut effectuer des comparaisons multiples pour déterminer où se trouve les différences entre les groupes d'herbicide. On ne fait aucune comparaisons multiples pour le bloc puisque ce facteur ne nous intéresse pas vraiment. On inclut le bloc dans l'analyse pour tenir compte du dispositif expérimental que nous avons utilisé.

```
> ##moyenne des groupes
```

```
> Moys <- tapply(X = temps$Temps, INDEX = temps$Herbicide, FUN = mean)
```

```
> Moys
```

Controle	Type1	Type2
3.644	6.730	5.854

```
> TukeyHSD(aov.bloc, which = "Herbicide")
```

Tukey multiple comparisons of means

95% family-wise confidence level

```
Fit: aov(formula = Temps ~ Herbicide + Bloc, data = temps)
```

```
$Herbicide
```

diff	lwr	upr	p adj
------	-----	-----	-------

```
Type1-Controle  3.086  1.18436 4.9876 0.00421
Type2-Controle  2.210  0.30836 4.1116 0.02545
Type2-Type1     -0.876 -2.77764 1.0256 0.42578
```

On note deux groupes, l'un formé par les groupes **Type1** et **Type2**, et l'autre formé par **Controle**. On peut représenter ces groupes à l'aide de traits :

Controle	Type2	Type1
3.64	5.85	6.73

---

Ou encore, on peut illustrer le résultat pour les différents niveaux du traitement **Herbicide** directement dans un graphique (fig. 3) :

```
> ##MSE de l'ANOVA
> MSE <- 1.11
> ##vecteur pour créer l'axe des x's
> Herb <- 1:3
> ##barres d'erreur
> inf <- Moys - sqrt(MSE)
> sup <- Moys + sqrt(MSE)
> ##graphique
> ##À noter que l'on crée le graphique avec suffisamment
> ##d'espace pour permettre le placement des barres d'erreurs
> ##ainsi que les lettres en bas.
> ##C'est pourquoi on étend la limite inférieure à 0.2 unités
> ##au bas du graphique au delà de la valeur minimale
> ##de la barre d'erreur.
> plot(Moys ~ Herb, ylab = "Temps de recolonisation (jours)",
```

```
xlab = "Type d'herbicide",
type = "p", ylim = c(min(inf) - 0.2, max(sup)),
xlim = c(0, 4), cex.lab = 1.2, xaxt = "n",
main = "Moyennes  $\pm$  racine carrée de MSE")
> ##ajout de l'axe des x's
> axis(side = 1, at = c(1, 2, 3),
      labels = c("Controle", "Type 1", "Type 2"))
> ##ajout de barres d'erreurs
> arrows(x0 = Herb, x1 = Herb, y0 = inf, y1 = sup,
      angle = 90, code = 3, length = 0.05)
> ##ajout des lettres
> text(x = 1, y = 2.46, labels = "a", cex = 1.2)
> text(x = 2, y = 5.54, labels = "b", cex = 1.2)
> text(x = 3, y = 4.67, labels = "b", cex = 1.2)
```

Finalement, on conclut que le temps de recolonisation ne diffère pas entre les groupes d'herbicide Type1 et Type2, mais que ce temps est nettement supérieur à celui du groupe témoin (Controle).

**b)** Était-ce justifié d'utiliser des blocs ? Justifiez votre réponse.

Oui, l'utilisation des blocs était justifiée puisque le terme Bloc expliquait une partie importante de la variance de la variable réponse, tel qu'indiqué par un  $F_{4,8}$  de 10 et un  $P = 0.0033$ . De plus, le dispositif a été créé en bloc et c'est toujours une bonne idée d'inclure ce genre de facteur dans l'analyse pour refléter la structure du dispositif.



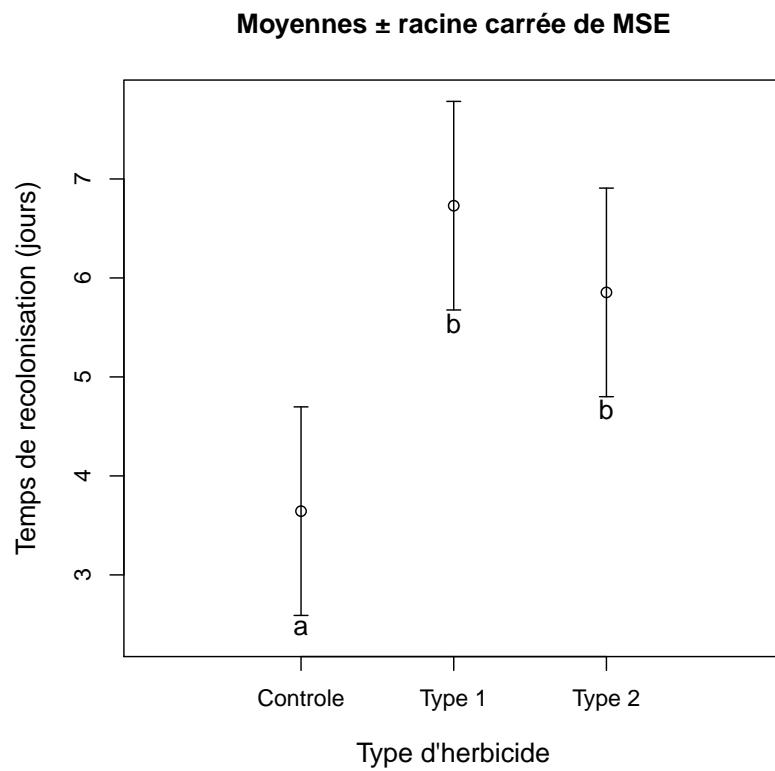


FIGURE 3 – Effets du type d'herbicide sur le temps de recolonisation.