**Comparaisons de mesures répétées sur plusieurs échantillons**

# I Anova à mesures répétées

## I.1 Contexte et modèle

On considère une expérience durant laquelle on étudie une variable (par exemple, performances de sportifs ou degrés d'alcool dans des vins) à diverses reprises (au cours de plusieurs semaines, par exemple).

| | Temps 1 | Temps 2 | $\cdots$ | Temps $p$ |
|---|---|---|---|---|
| Sujet 1 | $y_{11}$ | | $\cdots$ | $y_{1p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| Sujet q | $y_{q1}$ | | $\cdots$ | $y_{qp}$ |

Dans le cadre d'une étude de mesures répétées, on considère que les échantillons associées aux différentes périodes sont appariés (mesures non indépendantes).
En outre, on suppose qu'il puisse exister un effet "Sujet", aléatoire, conduisant au modèle suivant :

$$Y_{ij} = \mu + \alpha_i + \tau_j + \varepsilon_{ij}$$

où $i$ varie de 1 à $p$ (nombre de niveaux du facteur contrôlé), $j$ variant de 1 à $q$ (nombre d'individus).
On note $N$ le nombre d'observations définies par $N = pq$.
On suppose, en outre, que :

- le facteur contrôlé (Temps) est à effets fixes $\alpha_i$ avec la contrainte $\sum_{i=1}^{p} \alpha_i = 0$ ;
- les effets aléatoires $\tau_j$ sont indépendants et identiquement distribués de loi $\mathcal{N}(0; \sigma_S)$ ;
- les variables $\tau_j$ et $\varepsilon_{ij}$ sont indépendantes ;
- les erreurs $\varepsilon_{ij}$ sont idépendantes et identiquement distribuées de loi $\mathcal{N}(0; \sigma_R)$.

On a donc

- $\mathbb{V}\left(Y_{ij}\right) = {\sigma_S}^2 + \sigma_R^2$
- $\mathbb{C}ov\left(Y_{ij}, Y_{kj}\right) = \mathbb{V}\left(\tau_j\right) = {\sigma_S}^2$ pour $j \neq k$

On a donc

$$\mathbb{V}\left(Y_{ij} - Y_{kj}\right) = \mathbb{V}\left(Y_{ij}\right) + \mathbb{V}\left(Y_{ik}\right) - 2\mathbb{C}ov\left(Y_{ij}, Y_{kj}\right) = 2\left({\sigma_S}^2 + \sigma_R^2\right) - 2{\sigma_S}^2 = 2\sigma_R^2$$

Ainsi, l'hypothèse d'homogénéité des covariances (appelée postulat de sphéricité) conduira à considérer que toutes les différences entre deux variables quantitatives observées (cf. tests de Student dans le cas d'échantillons appariés) ont la même variance.
Ces conditions de sphéricité peuvent être vérifiées à l'aide du test de Mauchly.

**Conséquences :**
L'ANOVA à mesures répétées peut être considérée comme la généralisation des tests de Student par paires.
Comme dans l'ANOVA classique il faut s'assurer de la normalité et l'homoscédasticité de la distribution des résidus.
Il faudrait également vérifier les conditions liés à l'effet "Intra-Sujet" mais il est souvent difficile de le faire en raison d'un faible nombre de valeurs.
En outre, il faut tester la sphéricité à l'aide du test de Mauchly.

Sous l'hypothèse $\mathcal{H}_{\prime}$ : $\mu_1 = \cdots = \mu_p$, on a alors

$$Y_{ij} \sim \mathcal{N}(\mu; \sigma)$$

où $\sigma^2 = {\sigma_S}^2 + {\sigma_R}^2$. On a

$$SCE_{tot} = SCE_{Facteur} + SCE_{Sujets} + SCE_{res}$$

La $SCE_{tot}$ est associée à $N - 1 = pq - 1$ degrés de liberté.
La $SCE_{Facteur} = SCE_{Date}$ est associée à $p - 1$ degrés de liberté.
La $SCE_{Sujet}$ est associée à $q - 1$ degrés de liberté.
La $SCE_{res}$ est associée à $(p - 1) \times (q - 1)$ degrés de liberté.

Les carrés moyens liés au Facteur (Date) et aux résidus sont définis par

$$CM_{Facteur} = \frac{SCE_{Facteur}}{p-1} \text{ et } CM_{res} = \frac{SCE_{res}}{(p-1) \times (q-1)}$$

En conséquence, sous $\mathcal{H}_0$, on a

$$F_{obs} = \frac{CM_{Facteur}}{CM_{res}} \sim \mathcal{F}\left(p - 1; (p-1) \times (q-1)\right)$$

Remarque :
Le principe de l'analyse de variance à mesures répétéees consiste à éliminer la variabilité due aux différences entre sujets et donc à soustraire la variabilité inter-sujets de la variabilité résiduelle ("within").
La variabilité erreur par rapport à laquelle on teste l'effet du traitement est réduite, il est plus facile d'identifier les effets du traitement.
Ce type d'expérience (plans intra-sujets) est donc plus puissant que les plans où chaque sujet est soumis à une seule observation (plans inter-sujets).
Une ANOVA à mesures répétées avec un critère de classification peut se rapporter au niveau calculatoire à une ANOVA à 2 critères de classification dans laquelle le second critère serait le critère sujet avec une seule observation (Anova à 2 facteurs sans répétition).

## I.2 Un exemple

On considère l'évolution du degré d'alcool de 6 vins relevés à 3 périodes différentes.
Les données sont repértoriées dans ce qui suit.

```
> data

   Sujets Dates Degres
1   Vin 1    T1   10.3
2   Vin 2    T1    8.7
3   Vin 3    T1   10.3
4   Vin 4    T1   10.3
5   Vin 5    T1    9.4
6   Vin 6    T1    7.5
7   Vin 1    T2    8.6
8   Vin 2    T2    9.2
9   Vin 3    T2    9.5
10  Vin 4    T2   11.9
11  Vin 5    T2   10.3
12  Vin 6    T2    8.7
13  Vin 1    T3    8.9
14  Vin 2    T3    9.7
15  Vin 3    T3    9.7
16  Vin 4    T3    9.6
17  Vin 5    T3   10.3
18  Vin 6    T3    9.1

> boxplot(data$Degres ~ data$Dates) # visualisation
> Matrice = matrix(data$Degres,ncol=3,byrow=F) # réorganisation pour test de Mauchly
> mauchly.test(lm(Matrice ~ 1)) # sphéricité OK

        Mauchly's test of sphericity

data:  SSD matrix from lm(formula = Matrice ~ 1)
W = 0.26206, p-value = 0.4316

> anova = aov(Degres~Dates+Sujets, data=data)
> par(mfrow=c(2,2))
> plot(anova)
> residus = resid(anova)
> bartlett.test(residus ~ data$Dates) # OK Homoscédasticté

        Bartlett test of homogeneity of variances

data:  residus by data$Dates
Bartlett's K-squared = 0.13881, df = 2, p-value = 0.933

> shapiro.test(residus) # OK Normalité des résidus
```

```
        Shapiro-Wilk normality test

data:  residus
W = 0.95886, p-value = 0.5797

> anova_rep <- aov(Degres~Dates+Error(Sujets), data=data)
> summary(anova_rep)

Error: Sujets
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  5  8.504   1.701

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Dates      2  0.241  0.1206   0.173  0.843
Residuals 10  6.959  0.6959

> # Pas de différence significative
```

Dans le cas d'un effet significatif, on peut utiliser des tests post-hoc de comparaisons multiples de moyennes comme, par exemple, les tests de Student (méthodes d'ajustement de probabilité de Bonferroni ou Holm) ...
Certains auteurs préconisent d'utiliser la correction de Benjamini et Yekutieli (BY pour R).

# II Une alternative non paramétrique : le test de Friedman

Il existe un test non paramétrique lorsque l'une (au moins) des conditions d'application de l'ANOVA à mesures répétées n'est pas vérifiée. Il s'agit du test de Friedman, basé sur des rangs, à rapprocher du test de Kruskal-Wallis pour des échantillons indépendants.
Ce test permet également, conformément à la norme NF ISO 8587, de classer des produits ou des lots les uns par rapport aux autres, en fonction des préférences des consommateurs.
Dans ce cas, on ne cherche pas ici à déterminer le niveau de satisfaction procuré par chaque produit ou lot mais plutôt à le hiérarchiser par rapport aux autres produits.
*tLes produits sont généralement présentés simultanément au consommateur qui doit leur attribuer un rang de préférence sur différents critères proposés (aspect, toucher, goût, etc.).*

Reprenons l'exemple précédent. Les notes pouvant induire un problème de normalité, l'usage du test de Frieman peut être envisagée.

```
> friedman.test(Degres, Dates, Sujets)

        Friedman rank sum test

data:  Degres, Dates and Sujets
Friedman chi-squared = 0.6087, df = 2, p-value = 0.7376
```

On retrouve l'absence d'effet (significatif) de la date sur le degré d'alcool.