

Plans d'expérience

J.Lejeune

Journée de l'A.P.M. *L.E.G.T.A. Le Robillard*

lejeune@math.unicaen.fr

D'après les données de Jacques PAVY, professeur de mathématiques au
L.E.G.T.A.

5 mars 2008

1. Objectif recherché :

Objectif : à la suite d'une expérimentation sur des parcelles pouvant être irriguées différemment, décider si les rendements **moyens** de différentes variétés de R.G.I. tétraploïde (4r) en 3ème année sur la coupe n°1 sont significativement différents ou non.

Elements de vocabulaire "officiel" (ISO 3534-1 1.11).

1. Objectif recherché :

Objectif : à la suite d'une expérimentation sur des parcelles pouvant être irriguées différemment, décider si les rendements **moyens** de différentes variétés de R.G.I. tétraploïde (4r) en 3ème année sur la coupe n°1 sont significativement différents ou non.

Elements de vocabulaire "officiel" (ISO 3534-1 1.11).

- ▶ **Variable de réponse** : "Variable représentant le résultat d'une expérience."

Variable de réponse = **rendement de R.G.I. sur une parcelle.**

1. Objectif recherché :

Objectif : à la suite d'une expérimentation sur des parcelles pouvant être irriguées différemment, décider si les rendements **moyens** de différentes variétés de R.G.I. tétraploïde (4r) en 3ème année sur la coupe n°1 sont significativement différents ou non.

Elements de vocabulaire "officiel" (ISO 3534-1 1.11).

- ▶ **Variable de réponse** : "Variable représentant le résultat d'une expérience."
Variable de réponse = **rendement de R.G.I. sur une parcelle.**
- ▶ **Variable de prédiction** : "Variable susceptible de contribuer à l'explication du résultat d'une expérience."
Première variable de prédiction = **type de variété.**
Deuxième variable de prédiction = **niveau d'irrigation.**
....il peut en exister beaucoup d'autres.....

- ▶ **Facteur** : "Variable de prédiction qui varie en vue de l'évaluation de son effet sur la variable de réponse"
Facteur V (étudié) = type de variété avec $I = 5$ niveaux V1, V2, V3, V4 et V5.
Facteur B (contrôlé) = niveau d'irrigation avec $J = 4$ niveaux B1, B2, B3 et B4. (*S'il n'y avait qu'un niveau pour la variable de prédiction, celle-ci ne serait pas un facteur*).

- ▶ **Facteur** : "Variable de prédiction qui varie en vue de l'évaluation de son effet sur la variable de réponse"
Facteur V (étudié) = type de variété avec $I = 5$ niveaux V1, V2, V3, V4 et V5.
Facteur B (contrôlé) = niveau d'irrigation avec $J = 4$ niveaux B1, B2, B3 et B4. (*S'il n'y avait qu'un niveau pour la variable de prédiction, celle-ci ne serait pas un facteur*).
- ▶ **Traitement** "Ensemble constitué d'un niveau pour chaque facteur et qui sera appliqué à une unité expérimentale".
Traitement = affectation à une parcelle irriguée au niveau B2 de la variété V3.

- ▶ **Unité expérimentale** : "Entité qui reçoit un traitement particulier produisant ensuite une valeur de la variable de réponse"

$N = I \times J = 20$ unités expérimentales = 20 parcelles

B4	101	102	103	104	105
B3	201	202	203	204	205
B2	301	302	303	304	305
B1	401	402	403	404	405

- ▶ **Unité expérimentale** : "Entité qui reçoit un traitement particulier produisant ensuite une valeur de la variable de réponse"

$N = I \times J = 20$ unités expérimentales = 20 parcelles

B4	101	102	103	104	105
B3	201	202	203	204	205
B2	301	302	303	304	305
B1	401	402	403	404	405

- ▶ **Plan expérimental** : "Affectation de traitements à chaque unité expérimentale"

Exemple 1 :

B4	V5	V4	V3	V1	V2
B3	V3	V1	V4	V2	V5
B2	V2	V5	V1	V4	V3
B1	V1	V2	V5	V3	V4

.....mais celui-ci est aussi un plan d'expériences

Exemple 2 :

<i>B4</i>	<i>V3</i>	<i>V3</i>	<i>V3</i>	<i>V3</i>	<i>V3</i>
<i>B3</i>	<i>V3</i>	<i>V4</i>	<i>V4</i>	<i>V1</i>	<i>V1</i>
<i>B2</i>	<i>V2</i>	<i>V2</i>	<i>V2</i>	<i>V2</i>	<i>V2</i>
<i>B1</i>	<i>V1</i>	<i>V2</i>	<i>V5</i>	<i>V2</i>	<i>V2</i>

et aussi celui-ci !

Exemple 3 :

<i>B4</i>					
<i>B3</i>		<i>V1</i>		<i>V2</i>	
<i>B2</i>					
<i>B1</i>					

2. Cas de 2 variétés : comparaison Student - Fisher

2.1. Cas sans répétitions : résultats pour l'exemple 3 :

Variété	Rendement
V1	4.27
V2	3.86

2. Cas de 2 variétés : comparaison Student - Fisher

2.1. Cas sans répétitions : résultats pour l'exemple 3 :

Variété	Rendement
V1	4.27
V2	3.86

μ_1 = moyenne des rendements pour V1 et μ_2 = moyenne des rendements pour V2

On suppose :

Si Y_1 = rendement d'une parcelle avec V1, alors $Y_1 \sim \mathcal{N}(\mu_1, \sigma^2)$.

Si Y_2 = rendement d'une parcelle avec V2, alors $Y_2 \sim \mathcal{N}(\mu_2, \sigma^2)$.

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ avec le risque $\alpha = 5\%$.

2. Cas de 2 variétés : comparaison Student - Fisher

2.1. Cas sans répétitions : résultats pour l'exemple 3 :

Variété	Rendement
V1	4.27
V2	3.86

μ_1 = moyenne des rendements pour V1 et μ_2 = moyenne des rendements pour V2

On suppose :

Si Y_1 = rendement d'une parcelle avec V1, alors $Y_1 \sim \mathcal{N}(\mu_1, \sigma^2)$.

Si Y_2 = rendement d'une parcelle avec V2, alors $Y_2 \sim \mathcal{N}(\mu_2, \sigma^2)$.

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ avec le risque $\alpha = 5\%$.

Si H_0 est vrai, alors $Y_1 - Y_2 \sim \mathcal{N}(0, 2\sigma^2)$ donc rejet de H_0 si $|Y_1 - Y_2| > 1.96 \times \sqrt{2} \times \sigma$.

2. Cas de 2 variétés : comparaison Student - Fisher

2.1. Cas sans répétitions : résultats pour l'exemple 3 :

Variété	Rendement
V1	4.27
V2	3.86

μ_1 = moyenne des rendements pour V1 et μ_2 = moyenne des rendements pour V2

On suppose :

Si Y_1 = rendement d'une parcelle avec V1, alors $Y_1 \sim \mathcal{N}(\mu_1, \sigma^2)$.

Si Y_2 = rendement d'une parcelle avec V2, alors $Y_2 \sim \mathcal{N}(\mu_2, \sigma^2)$.

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ avec le risque $\alpha = 5\%$.

Si H_0 est vrai, alors $Y_1 - Y_2 \sim \mathcal{N}(0, 2\sigma^2)$ donc rejet de H_0 si $|Y_1 - Y_2| > 1.96 \times \sqrt{2} \times \sigma$.

→ nécessité d'estimer σ .

Comment estimer σ ?

Soient n_1 répétitions avec V1 dans les mêmes conditions, donnant comme résultats : $y_{1,1}, y_{1,2}, \dots, y_{1,n_1}$

Soient n_2 répétitions avec V2 dans les mêmes conditions, donnant comme résultats : $y_{2,1}, y_{2,2}, \dots, y_{2,n_2}$

Soit $s_1^2 = \frac{\sum_{i=1}^{n_1} (y_{1,i} - \bar{y}_1)^2}{n_1 - 1}$ la variance d'échantillonnage pour V1.

Soit $s_2^2 = \frac{\sum_{i=1}^{n_2} (y_{2,i} - \bar{y}_2)^2}{n_2 - 1}$ la variance d'échantillonnage pour V2.

Comment estimer σ ?

Soient n_1 répétitions avec V1 dans les mêmes conditions, donnant comme résultats : $y_{1,1}, y_{1,2}, \dots, y_{1,n_1}$

Soient n_2 répétitions avec V2 dans les mêmes conditions, donnant comme résultats : $y_{2,1}, y_{2,2}, \dots, y_{2,n_2}$

Soit $s_1^2 = \frac{\sum_{i=1}^{n_1} (y_{1,i} - \bar{y}_1)^2}{n_1 - 1}$ la variance d'échantillonnage pour V1.

Soit $s_2^2 = \frac{\sum_{i=1}^{n_2} (y_{2,i} - \bar{y}_2)^2}{n_2 - 1}$ la variance d'échantillonnage pour V2.

On montre qu'une bonne estimation de σ^2 est :

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

Donc il faut au moins 2 répétitions dans les mêmes conditions, par variété.

Remarque : Si $n_1 = n_2$, alors $s^2 = \frac{s_1^2 + s_2^2}{2}$.

2.2. Cas de répétitions : comment conclure ?

Soit \bar{y}_1 (resp. \bar{y}_2) la moyenne des résultats pour V1 (resp. V2).

On rejette $H_0 : \mu_1 = \mu_2$ si $|\bar{y}_1 - \bar{y}_2|$ est "trop" grand.

Test de Student pour échantillons indépendants.

Si on prend le risque $\alpha = 5\%$ de se tromper, on calcule :

$$t_{obs} = \frac{\bar{y}_1 - \bar{y}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

2.2. Cas de répétitions : comment conclure ?

Soit \bar{y}_1 (resp. \bar{y}_2) la moyenne des résultats pour V1 (resp. V2).

On rejette $H_0 : \mu_1 = \mu_2$ si $|\bar{y}_1 - \bar{y}_2|$ est "trop" grand.

Test de Student pour échantillons indépendants.

Si on prend le risque $\alpha = 5\%$ de se tromper, on calcule :

$$t_{obs} = \frac{\bar{y}_1 - \bar{y}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

On rejette H_0 si

$$|t_{obs}| > t_{0.975}$$

où $t_{0.975}$ est le quantile d'ordre 0.975 d'une loi de Student à

$(n_1 - 1) + (n_2 - 1)$ degrés de liberté,

ou si la p.value associée \mathbf{P} est inférieure à 0.05.

Exemple 4 :

$B4$					
$B3$					
$B2$					
$B1$	$V1 = 3.97$	$V2 = 4.5$	$V2 = 4.4$		$V1 = 4.01$

Exemple 4 :

B4					
B3					
B2					
B1	V1 = 3.97	V2 = 4.5	V2 = 4.4		V1 = 4.01

$$\bar{y}_1 - \bar{y}_2 = 3.99 - 4.45 = -0.46$$

$$s_1^2 = 0.0008, \quad s_2^2 = 0.005, \quad s^2 = 0.0029$$

$$s = \sqrt{0.0029} = 0.05385165$$

$$t_{obs} = \frac{-0.46}{0.05385165} = -8.542$$

Exemple 4 :

B4					
B3					
B2					
B1	V1 = 3.97	V2 = 4.5	V2 = 4.4		V1 = 4.01

$$\bar{y}_1 - \bar{y}_2 = 3.99 - 4.45 = -0.46$$

$$s_1^2 = 0.0008, \quad s_2^2 = 0.005, \quad s^2 = 0.0029$$

$$s = \sqrt{0.0029} = 0.05385165$$

$$t_{obs} = \frac{-0.46}{0.05385165} = -8.542$$

Rejet de H_0 car $|t_{obs}| = 8.542 > 4.302653 = t_{0.975}$

$$P = 0.01343 < 0.05$$

Exemple 5 :

$B4$					
$B3$					
$B2$		$V2 = 4.95$	$V1 = 4.52$		
$B1$		$V1 = 3.46$			$V2 = 3.95$

Exemple 5 :

B4					
B3					
B2		V2 = 4.95	V1 = 4.52		
B1		V1 = 3.46			V2 = 3.95

$$\bar{y}_1 - \bar{y}_2 = 3.99 - 4.45 = -0.46$$

Calcul du t_{obs} comme dans l'exemple 4 :

$$s_1^2 = 0.5618, \quad s_2^2 = 0.5, \quad s^2 = 0.5309, \quad s = 0.7286288$$

Exemple 5 :

B4					
B3					
B2		V2 = 4.95	V1 = 4.52		
B1		V1 = 3.46			V2 = 3.95

$$\bar{y}_1 - \bar{y}_2 = 3.99 - 4.45 = -0.46$$

Calcul du t_{obs} comme dans l'exemple 4 :

$$s_1^2 = 0.5618, \quad s_2^2 = 0.5, \quad s^2 = 0.5309, \quad s = 0.7286288$$

$$t_{obs} = \frac{-0.46}{\left(\frac{1}{2} + \frac{1}{2}\right) \times 0.7286288} = -0.6313$$

Non rejet de H_0 car :

$$|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$$

$$\text{et } P = 0.5924 > 0.05$$

Il n'y a pas eu répétitions **dans les mêmes conditions** car on peut soupçonner B2 d'être plus irrigué que B1.

Solution ?

$H_0 : \mu_1 = \mu_2$ équivalent à $H_0 : d = \mu_1 - \mu_2 = 0$

On calcule, pour chaque bloc B_j , les différences d_j entre les variétés.

Nouvelle présentation :

	V1	V2	d_j
B1	3.46	3.95	-0.49
B2	4.52	4.95	-0.43

Solution ?

$H_0 : \mu_1 = \mu_2$ équivalent à $H_0 : d = \mu_1 - \mu_2 = 0$

On calcule, pour chaque bloc B_j , les différences d_j entre les variétés.

Nouvelle présentation :

	V1	V2	d_j
B1	3.46	3.95	-0.49
B2	4.52	4.95	-0.43

$$\bar{d} = -0.46 \quad s_d = 0.04242641$$

$$t_{obs} = \sqrt{n} \frac{\bar{d}}{s_d} = \sqrt{2} \frac{-0.46}{0.04242641} = -15.3333$$

Solution ?

$H_0 : \mu_1 = \mu_2$ équivalent à $H_0 : d = \mu_1 - \mu_2 = 0$

On calcule, pour chaque bloc B_j , les différences d_j entre les variétés.

Nouvelle présentation :

	V1	V2	d_j
B1	3.46	3.95	-0.49
B2	4.52	4.95	-0.43

$$\bar{d} = -0.46 \quad s_d = 0.04242641$$

$$t_{obs} = \sqrt{n} \frac{\bar{d}}{s_d} = \sqrt{2} \frac{-0.46}{0.04242641} = -15.3333$$

Rejet de H_0 car : $|t_{obs}| = 15.3333 > 12.7 = t_{0.975}$

Ici $12.7 = t_{0.975}$ est le quantile d'une loi de Student à

$n - 1 = 2 - 1 = 1$ degrés de liberté. $\mathbf{P} = 0.04146 < 0.05$.

C'est le test de Student pour échantillons appariés.

2.3. Autre présentation :

2.3.1. On ne tient pas compte d'un possible effet de l'irrigation :

Soit μ moyenne générale inconnue. On pose :

$v_i = \mu_i - \mu$ effet principal de la variété i sur la moyenne générale.

On a :

$$\mu_i = \mu + v_i$$

$$y_{ik} = \mu + v_i + \epsilon_{ik} \text{ avec } \epsilon_{ik} \sim \mathcal{N}(0, \sigma^2)$$

2.3. Autre présentation :

2.3.1. On ne tient pas compte d'un possible effet de l'irrigation :

Soit μ moyenne générale inconnue. On pose :

$v_i = \mu_i - \mu$ effet principal de la variété i sur la moyenne générale.

On a :

$$\mu_i = \mu + v_i$$

$$y_{ik} = \mu + v_i + \epsilon_{ik} \text{ avec } \epsilon_{ik} \sim \mathcal{N}(0, \sigma^2)$$

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ revient à tester :

$$H_0 : \forall i, v_i = 0 \text{ contre } H_1 : \exists v_i \neq 0$$

2.3. Autre présentation :

2.3.1. On ne tient pas compte d'un possible effet de l'irrigation :

Soit μ moyenne générale inconnue. On pose :

$v_i = \mu_i - \mu$ effet principal de la variété i sur la moyenne générale.

On a :

$$\mu_i = \mu + v_i$$

$$y_{ik} = \mu + v_i + \epsilon_{ik} \text{ avec } \epsilon_{ik} \sim \mathcal{N}(0, \sigma^2)$$

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ revient à tester :

$$H_0 : \forall i, v_i = 0 \text{ contre } H_1 : \exists v_i \neq 0$$

On estime :

$$\hat{v}_1 = \bar{y}_1 - \bar{y} = -0.23 \text{ et } \hat{v}_2 = 0.23.$$

Idée : rejet de H_0 si $\sum_{i=1}^2 \hat{v}_i^2$ "trop grand " par rapport à la variance σ^2

Tableau D(données) :

V1	V2
3.46	3.95
4.52	4.95

ou par exemple

V1	V2
4.52	3.95
3.46	4.95

Tableau D(données) :

V1	V2
3.46	3.95
4.52	4.95

ou par exemple

V1	V2
4.52	3.95
3.46	4.95

Tableau M - Tableau MT = Tableau F (écarts dûs au facteur)

V1	V2		V1	V2	=	V1	V2
3.99	4.45	-	4.22	4.22	=	-0.23	0.23
3.99	4.45		4.22	4.22		-0.23	0.23

Tableau D - Tableau M = Tableau R(écarts résiduels)

V1	V2		V1	V2	=	V1	V2
3.46	3.95	-	3.99	4.45	=	-0.53	-0.50
4.52	4.95		3.99	4.45		0.53	0.50

Tableau D - Tableau MT = Tableau T (écarts à la moyenne totale)

V1	V2		V1	V2	=	V1	V2
3.46	3.95	-	4.22	4.22	=	-0.76	-0.27
4.52	4.95		4.22	4.22		0.30	0.73

Pour les tableaux T,F et R, calcul de la somme des carrés des écarts :

- ▶ $SCE_T = 1.2734$
- ▶ $SCE_F = 0.2116$
- ▶ $SCE_R = 1.0618$

Equation dite de l'analyse de la variance : $SCE_T = SCE_F + SCE_R$
 vérifiée car $1.2734 = 0.2116 + 1.0618$.

Théorie de Fisher :

Théorie de Fisher :

- ▶ Sous l'hypothèse d'égalité des moyennes :

$F_{obs} = \frac{SCE_F / (p - 1)}{SCE_R / (n - 1)p}$ suit une loi de Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

Rejet de l'égalité des moyennes et donc de la nullité des effets a_i du facteur sur la moyenne avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

Théorie de Fisher :

- ▶ Sous l'hypothèse d'égalité des moyennes :

$F_{obs} = \frac{SCE_F / (p - 1)}{SCE_R / ((n - 1)p)}$ suit une loi de Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

Rejet de l'égalité des moyennes et donc de la nullité des effets a_i du facteur sur la moyenne avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

- ▶ $SCE_R / ((n - 1)p)$ noté aussi $\widehat{\sigma^2}$ est un "bon" estimateur de σ^2 .

Théorie de Fisher :

- ▶ Sous l'hypothèse d'égalité des moyennes :

$$F_{obs} = \frac{SCE_F / (p - 1)}{SCE_R / ((n - 1)p)}$$

suit une loi de Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

Rejet de l'égalité des moyennes et donc de la nullité des effets a_i du facteur sur la moyenne avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $p - 1$ et $(n - 1)p$ degrés de liberté.

- ▶ $SCE_R / ((n - 1)p)$ noté aussi $\widehat{\sigma^2}$ est un "bon" estimateur de σ^2 .

$$n \sum_{i=1}^p (\widehat{a}_i)^2$$

- ▶ $F_{obs} = \frac{i=1}{\widehat{\sigma^2}}$ donc conforme à l'idée.

Ici $p = 2$, $n = 2$ et $F_{0.95} = 18.51282$

Tableau d'analyse de la variance à 1 facteur(issu du logiciel R)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
variete	1	0.2116	0.2116	0.3986	0.5924
Residuals	2	1.0618	0.5309		

Non-rejet de l'égalité des moyennes car :

$$F_{obs} = \frac{0.2116}{1.0618/2} = 0.3986 < 18.51282 = F_{0.95} \text{ et la p-value est}$$

$$P = 0.5924$$

Remarques :

Rappel : avec le test de Student pour échantillons indépendants,
non-rejet de H_0 car $|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$

Remarques :

Rappel : avec le test de Student pour échantillons indépendants,
non-rejet de H_0 car $|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$

$$\begin{aligned} \blacktriangleright 0.6313^2 = t_{obs}^2 &= \left(\frac{\bar{y}_1 - \bar{y}_2}{s\left(\frac{1}{2} + \frac{1}{2}\right)} \right)^2 = \frac{SCE_F/(2-1)}{SCE_R/(2-1)2} = F_{obs} = \\ &0.3986 \end{aligned}$$

Remarques :

Rappel : avec le test de Student pour échantillons indépendants,
non-rejet de H_0 car $|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$

$$\blacktriangleright 0.6313^2 = t_{obs}^2 = \left(\frac{\bar{y}_1 - \bar{y}_2}{s\left(\frac{1}{2} + \frac{1}{2}\right)} \right)^2 = \frac{SCE_F/(2-1)}{SCE_R/(2-1)2} = F_{obs} = 0.3986$$

$$\blacktriangleright 4.302653^2 = t_{0.975}^2 = F_{0.95} = 18.51282$$

Remarques :

Rappel : avec le test de Student pour échantillons indépendants, non-rejet de H_0 car $|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$

$$\blacktriangleright 0.6313^2 = t_{obs}^2 = \left(\frac{\bar{y}_1 - \bar{y}_2}{s\left(\frac{1}{2} + \frac{1}{2}\right)} \right)^2 = \frac{SCE_F / (2 - 1)}{SCE_R / (2 - 1) \cdot 2} = F_{obs} = 0.3986$$

$$\blacktriangleright 4.302653^2 = t_{0.975}^2 = F_{0.95} = 18.51282$$

\blacktriangleright La p-value est identique pour Student et pour Fischer $P = 0.5924$.

Remarques :

Rappel : avec le test de Student pour échantillons indépendants, non-rejet de H_0 car $|t_{obs}| = 0.6313 < 4.302653 = t_{0.975}$

- ▶ $0.6313^2 = t_{obs}^2 = \left(\frac{\bar{y}_1 - \bar{y}_2}{s(\frac{1}{2} + \frac{1}{2})} \right)^2 = \frac{SCE_F/(2-1)}{SCE_R/(2-1)2} = F_{obs} = 0.3986$
- ▶ $4.302653^2 = t_{0.975}^2 = F_{0.95} = 18.51282$
- ▶ La p-value est identique pour Student et pour Fischer $P = 0.5924$.
- ▶ Même estimation de σ^2 . Avec Student, $s^2 = 0.5309$ et avec Fischer $\widehat{\sigma}^2 = 0.5309$

2.3.2. Prise en compte d'un éventuel effet irrigation :

Soit μ_{ij} moyenne générale inconnue du rendement quand la variété i est dans le bloc j ;

$\mu_{i.} = \frac{1}{2}(\mu_{i1} + \mu_{i2})$ moyenne du rendement pour la variété i

$\mu_{.j} = \frac{1}{2}(\mu_{1j} + \mu_{2j})$ moyenne pour le bloc j

$\mu_{..} = \frac{1}{2}(\mu_{1.} + \mu_{2.})$ moyenne générale

$v_i = \mu_{i.} - \mu_{..}$ effet principal de la variété i sur la moyenne générale.

$b_j = \mu_{.j} - \mu_{..}$ effet principal du bloc j sur la moyenne générale

Modèle statistique :

$$y_{ij} = \mu + v_i + b_j + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

2.3.2. Prise en compte d'un éventuel effet irrigation :

Soit μ_{ij} moyenne générale inconnue du rendement quand la variété i est dans le bloc j ;

$\mu_{i.} = \frac{1}{2}(\mu_{i1} + \mu_{i2})$ moyenne du rendement pour la variété i

$\mu_{.j} = \frac{1}{2}(\mu_{1j} + \mu_{2j})$ moyenne pour le bloc j

$\mu_{..} = \frac{1}{2}(\mu_{1.} + \mu_{2.})$ moyenne générale

$v_i = \mu_{i.} - \mu_{..}$ effet principal de la variété i sur la moyenne générale.

$b_j = \mu_{.j} - \mu_{..}$ effet principal du bloc j sur la moyenne générale

Modèle statistique :

$$y_{ij} = \mu + v_i + b_j + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

Tester $H_0 : \mu_{1.} = \mu_{2.}$ contre $H_1 : \mu_{1.} \neq \mu_{2.}$ revient à tester :

$H_0 : \forall i, v_i = 0$ contre $H_1 : \exists v_i \neq 0$

Tableau MV - Tableau MT = Tableau V (\hat{v}_i = écarts dûs au facteur variété)

	V1	V2	-		V1	V2	=		V1	V2
B1	3.99	4.45		B1	4.22	4.22		B1	-0.23	0.23
B2	3.99	4.45		B2	4.22	4.22		B2	-0.23	0.23

Tableau MB - Tableau MT = Tableau B (\hat{b}_j = écarts dûs au facteur bloc)

	V1	V2	-		V1	V2	=
B1	3.705	3.705		B1	4.22	4.22	
B2	4.735	4.735		B2	4.22	4.22	

	V1	V2
B1	-0.515	-0.515
B2	0.515	0.515

MT+V+B=Tableau ME (\hat{y}_{ij} = moyennes estimées)

	V1	V2		V1	V2	
B1	4.22	4.22	+	B1	-0.23	0.23
B2	4.22	4.22		B2	-0.23	0.23

	V1	V2		V1	V2	
B1	-0.515	-0.515	=	B1	3.475	3.935
B2	0.515	0.515		B2	4.505	4.965

Tableau D - Tableau ME=Tableau R ($\hat{\epsilon}_{ij}$ = écarts résiduels)

	V1	V2		V1	V2		V1	V2	
B1	3.46	3.95	-	B1	3.475	=	B1	-0.015	0.015
B2	4.52	4.95		B2	4.505		B2	0.015	-0.015

Pour les tableaux T,V,B et R, calcul de la somme des carrés des écarts :

▶ $SCE_T = 1.2734$

▶ $SCE_V = 0.2116$

▶ $SCE_B = 1.0609$

▶ $SCE_R = 0.0009$

Equation dite de l'analyse de la variance :

$SCE_T = SCE_V + SCE_B + SCE_R$ vérifiée car
 $1.2734 = 0.2116 + 1.0609 + 0.0009.$

Théorie de Fisher :

I variété et J blocs

Théorie de Fisher :

I variété et J blocs

- ▶ $F_{obs} = \frac{SCE_V / (I - 1)}{SCE_R / ((I - 1)(J - 1))}$ suit une loi de Fisher à $I - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

Rejet de $H_0 : \forall i, v_i = 0$ avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $I - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

Théorie de Fisher :

I variété et J blocs

- ▶ $F_{obs} = \frac{SCE_V/(I-1)}{SCE_R/(I-1)(J-1)}$ suit une loi de Fisher à $I-1$ et $(I-1)(J-1)$ degrés de liberté.

Rejet de $H_0 : \forall i, v_i = 0$ avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $I-1$ et $(I-1)(J-1)$ degrés de liberté.

- ▶ $F_{obs} = \frac{SCE_B/(J-1)}{SCE_R/(I-1)(J-1)}$ suit une loi de Fisher à $J-1$ et $(I-1)(J-1)$ degrés de liberté.

Rejet de $H_0 : \forall j, b_j = 0$ avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $J-1$ et $(I-1)(J-1)$ degrés de liberté.

Théorie de Fisher :

I variété et J blocs

- ▶ $F_{obs} = \frac{SCE_V / (I - 1)}{SCE_R / ((I - 1)(J - 1))}$ suit une loi de Fisher à $I - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

Rejet de $H_0 : \forall i, v_i = 0$ avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $I - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

- ▶ $F_{obs} = \frac{SCE_B / (J - 1)}{SCE_R / ((I - 1)(J - 1))}$ suit une loi de Fisher à $J - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

Rejet de $H_0 : \forall j, b_j = 0$ avec le risque $\alpha = 5\%$ si

$F_{obs} > F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $J - 1$ et $(I - 1)(J - 1)$ degrés de liberté.

- ▶ $SCE_R / ((I - 1)(J - 1))$ noté aussi $\widehat{\sigma^2}$ est un "bon" estimateur de σ^2 .

Ici $I = 2$, $J = 2$ et $F_{0.95} = 161.4476$

Tableau d'analyse de la variance à 2 facteurs (avec le logiciel R)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
variete	1	0.2116	0.2116	0.2116	235.11	0.04146 *
bloc	1	1.0609	1.0609	1.0609	1178.78	0.01854 *
Residuals	1	0.0009	0.0009			

Conclusions :

$F_{obs} = 235.11 > 161.4476 = F_{0.95}$ donc rejet de l'égalité des moyennes de rendement pour les variétés

$F_{obs} = 1178.78 > 161.4476 = F_{0.95}$ donc rejet de l'égalité des moyennes de rendement pour les blocs

Rappel : avec le test de Student pour échantillons appariés, rejet de H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$
Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à $n - 1 = 2 - 1 = 1$ degrés de liberté. $\mathbf{P} = 0.04146 < 0.05$.

Rappel : avec le test de Student pour échantillons appariés, rejet de H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$
Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à $n - 1 = 2 - 1 = 1$ degrés de liberté. $P = 0.04146 < 0.05$.

$$\begin{aligned} \blacktriangleright (-15.3333)^2 = t_{obs}^2 &= \left(\frac{\bar{d}}{(s_d/\sqrt{2})} \right)^2 = \left(\frac{-0.46}{(0.04242641/\sqrt{2})} \right)^2 = \\ &= \frac{SCE_V/(2-1)}{SCE_R/(2-1)(2-1)} = F_{obs} = 0.3986 \end{aligned}$$

Rappel : avec le test de Student pour échantillons appariés, rejet de

H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$

Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à

$n - 1 = 2 - 1 = 1$ degrés de liberté. $P = 0.04146 < 0.05$.

$$\blacktriangleright (-15.3333)^2 = t_{obs}^2 = \left(\frac{\bar{d}}{(s_d/\sqrt{2})} \right)^2 = \left(\frac{-0.46}{(0.04242641/\sqrt{2})} \right)^2 =$$

$$\frac{SCE_V/(2-1)}{SCE_R/(2-1)(2-1)} = F_{obs} = 0.3986$$

$$\blacktriangleright 12.70620^2 = t_{0.975}^2 = F_{0.95} = 161.4476$$

Rappel : avec le test de Student pour échantillons appariés, rejet de H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$
 Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à $n - 1 = 2 - 1 = 1$ degrés de liberté. $P = 0.04146 < 0.05$.

- ▶ $(-15.3333)^2 = t_{obs}^2 = \left(\frac{\bar{d}}{(s_d/\sqrt{2})} \right)^2 = \left(\frac{-0.46}{(0.04242641/\sqrt{2})} \right)^2 = \frac{SCE_V/(2-1)}{SCE_R/(2-1)(2-1)} = F_{obs} = 0.3986$
- ▶ $12.70620^2 = t_{0.975}^2 = F_{0.95} = 161.4476$
- ▶ La p-value est identique pour Student et pour Fisher $P = 0.04146$.

Rappel : avec le test de Student pour échantillons appariés, rejet de H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$
 Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à $n - 1 = 2 - 1 = 1$ degrés de liberté. $P = 0.04146 < 0.05$.

- ▶ $(-15.3333)^2 = t_{obs}^2 = \left(\frac{\bar{d}}{(s_d/\sqrt{2})} \right)^2 = \left(\frac{-0.46}{(0.04242641/\sqrt{2})} \right)^2 = \frac{SCE_V/(2-1)}{SCE_R/(2-1)(2-1)} = F_{obs} = 0.3986$
- ▶ $12.70620^2 = t_{0.975}^2 = F_{0.95} = 161.4476$
- ▶ La p-value est identique pour Student et pour Fisher $P = 0.04146$.
- ▶ Avec Fisher $\hat{\sigma}^2 = 0.0009 = \frac{s_d^2}{2}$

Rappel : avec le test de Student pour échantillons appariés, rejet de H_0 car $|t_{obs}| = 15.3333 > 12.70620 = t_{0.975}$
 Ici $12.70620 = t_{0.975}$ est le quantile d'une loi de Student à $n - 1 = 2 - 1 = 1$ degrés de liberté. $P = 0.04146 < 0.05$.

$$\begin{aligned} \blacktriangleright (-15.3333)^2 = t_{obs}^2 &= \left(\frac{\bar{d}}{(s_d/\sqrt{2})} \right)^2 = \left(\frac{-0.46}{(0.04242641/\sqrt{2})} \right)^2 = \\ &= \frac{SCE_V/(2-1)}{SCE_R/(2-1)(2-1)} = F_{obs} = 0.3986 \end{aligned}$$

$$\blacktriangleright 12.70620^2 = t_{0.975}^2 = F_{0.95} = 161.4476$$

► La p-value est identique pour Student et pour Fisher $P = 0.04146$.

$$\blacktriangleright \text{Avec Fisher } \hat{\sigma}^2 = 0.0009 = \frac{s_d^2}{2}$$

► Pour 2 variétés, équivalence Student-Fisher

3. Cas de p variétés ($p > 2$) : analyse à 1 ou 2 facteurs ?

3.1. Pas de prise en compte d'un éventuel effet "irrigation"

N unités expérimentales.

Pour V_i , on fixe n_i le nombre de répétitions (qui seront ici supposées faites dans les mêmes conditions puisqu'on ne prend pas en compte un éventuel effet irrigation). Les n_i peuvent être différents mais $n_1 + n_2 + \dots + n_p = N$

$N = 20$ unités expérimentales et 5 variétés, on peut prendre

$n_i = n = 4$.

Plan en randomisation totale :

Affectation au hasard d'une variété, jusqu'à épuisement.

Exemple 6 :

V3	V1	V1	V2	V1
V4	V2	V4	V5	V3
V3	V5	V2	V2	V5
V1	V4	V3	V5	V4

V1	V2	V3	V4	V5
3.97	3.94	3.59	3.71	4.75
4.47	4.07	3.92	3.96	4.78
4.27	3.86	3.92	4.02	4.61
3.93	4.40	3.74	3.91	4.77
$\bar{y}_1 = 4.16$	$\bar{y}_2 = 4.0675$	$\bar{y}_3 = 3.79$	$\bar{y}_4 = 3.9$	$\bar{y}_5 = 4.72$

$Y_{ik} \sim \mathcal{N}(\mu_i, \sigma^2)$ valeur du rendement pour la variété V_i et la k -ème répétition

Tester $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ contre H_1 : "il existe au moins 2 moyennes différentes"

avec le risque $\alpha = 5\%$.

1ère possibilité :

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ puis

tester $H_0 : \mu_2 = \mu_3$ contre $H_1 : \mu_2 \neq \mu_3$ etc...

DANGERS !

- ▶ Non rejet de $H_0 : \mu_1 = \mu_2$ et non-rejet de $H_0 : \mu_2 = \mu_3$ n'entraîne pas non-rejet de $H_0 : \mu_1 = \mu_3$!
Il faudrait faire 10 tests de Student !

$Y_{ik} \sim \mathcal{N}(\mu_i, \sigma^2)$ valeur du rendement pour la variété V_i et la k -ème répétition

Tester $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ contre H_1 : "il existe au moins 2 moyennes différentes"
avec le risque $\alpha = 5\%$.

1ère possibilité :

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ puis
tester $H_0 : \mu_2 = \mu_3$ contre $H_1 : \mu_2 \neq \mu_3$ etc...

DANGERS !

- ▶ Non rejet de $H_0 : \mu_1 = \mu_2$ et non-rejet de $H_0 : \mu_2 = \mu_3$
n'entraîne pas non-rejet de $H_0 : \mu_1 = \mu_3$!
Il faudrait faire 10 tests de Student !
- ▶ Pour obtenir un risque global de 5%, il faut faire chaque test
avec un risque de 0,5%

$Y_{ik} \sim \mathcal{N}(\mu_i, \sigma^2)$ valeur du rendement pour la variété V_i et la k -ème répétition

Tester $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ contre H_1 : "il existe au moins 2 moyennes différentes"
avec le risque $\alpha = 5\%$.

1ère possibilité :

Tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$ puis
tester $H_0 : \mu_2 = \mu_3$ contre $H_1 : \mu_2 \neq \mu_3$ etc...

DANGERS !

- ▶ Non rejet de $H_0 : \mu_1 = \mu_2$ et non-rejet de $H_0 : \mu_2 = \mu_3$
n'entraîne pas non-rejet de $H_0 : \mu_1 = \mu_3$!
Il faudrait faire 10 tests de Student !
- ▶ Pour obtenir un risque global de 5%, il faut faire chaque test avec un risque de 0,5%
- ▶ Avec des tests séparés, on a 10 estimations différentes pour le même σ^2 !

2ème possibilité :

Analyse de la variance à un facteur. (ANOVA sur les calculatrices).

Tableau d'analyse de la variance avec R

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
variete	4	2.11447	0.52862	15.354	3.496e-05 ***
Residuals	15	0.51643	0.03443		

2ème possibilité :

Analyse de la variance à un facteur. (ANOVA sur les calculatrices).

Tableau d'analyse de la variance avec R

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
variete	4	2.11447	0.52862	15.354	3.496e-05 ***
Residuals	15	0.51643	0.03443		

Rejet de H_0 car $F_{obs} = 15.354 > 3.055568 = F_{0.95}$ où $F_{0.95}$ est le quantile d'ordre 0.95 d'une loi de Fisher à $p - 1 = 5 - 1 = 4$ et $N - p = 20 - 5 = 15$ degrés de liberté. De plus la p.value $\mathbf{P} = 3.496e - 05 < 0.05$

Ici, σ^2 est estimé par 0.03443

Et si cet effet irrigation existait alors que nous ne l'avions pas pris en compte, que pensez du plan suivant ?

Exemple 2 :

B4	V3	V3	V3	V3	V3
B3	V3	V4	V4	V1	V1
B2	V2	V2	V2	V2	V2
B1	V1	V2	V5	V2	V2

Risque d'attribuer un "bon" rendement à la variété V3 alors que celui-ci ne provient que d'un surcroît d'irrigation.

Plan factoriel en blocs aléatoires complets :

Exemple 1 :

B4	V5	V4	V3	V1	V2
B3	V3	V1	V4	V2	V5
B2	V2	V5	V1	V4	V3
B1	V1	V2	V5	V3	V4

et données fournies par Jacques PAVY (Plan utilisé au L.E.G.T.A.) :

	V1	V2	V3	V4	V5
B1	3.97	3.94	3.59	3.71	4.75
B2	3.93	4.40	3.74	3.91	4.77
B3	4.27	3.86	3.92	4.02	4.61
B4	4.47	4.07	3.92	3.96	4.78
	$\bar{y}_1 = 4.16$	$\bar{y}_2 = 4.0675$	$\bar{y}_3 = 3.79$	$\bar{y}_4 = 3.9$	$\bar{y}_5 = 4.72$

Modèle statistique :

$$y_{ij} = \mu_{ij} + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \text{ ou}$$

$$y_{ij} = \mu + v_i + b_j + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \text{ avec}$$

$$\sum_{i=1}^I v_i = 0 \quad \sum_{j=1}^J b_j = 0$$

Modèle statistique :

$$y_{ij} = \mu_{ij} + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \text{ ou}$$

$$y_{ij} = \mu + v_i + b_j + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \text{ avec}$$

$$\sum_{i=1}^I v_i = 0 \quad \sum_{j=1}^J b_j = 0$$

Tester $H_0 : \mu_{1.} = \mu_{2.} = \dots = \mu_{5.}$ contre H_1 : "il existe au moins deux moyennes différentes" revient à tester :

$$H_0 : \forall i, v_i = 0 \text{ contre } H_1 : \text{"il existe au moins un } v_i \neq 0\text{"}$$

1ère possibilité :

Pour le Bloc 1, tester $H_0 : \mu_{11} = \mu_{21} = \dots = \mu_{51}$ contre H_1 : " il existe au moins deux moyennes différentes"

Pour le Bloc 2, tester $H_0 : \mu_{12} = \mu_{22} = \dots = \mu_{52}$ contre H_1 : " il existe au moins deux moyennes différentes"

etc....

Problèmes :

- ▶ Pour un bloc donné, pas de répétitions. Il faudrait au moins 2 répétitions par variété et donc 10 parcelles. Au total, 40 parcelles au lieu de 20.

Problèmes :

- ▶ Pour un bloc donné, pas de répétitions. Il faudrait au moins 2 répétitions par variété et donc 10 parcelles. Au total, 40 parcelles au lieu de 20.
- ▶ Pour obtenir un risque global de 5%, il faut faire chacun des tests de Fisher avec un risque de 1,25%

Problèmes :

- ▶ Pour un bloc donné, pas de répétitions. Il faudrait au moins 2 répétitions par variété et donc 10 parcelles. Au total, 40 parcelles au lieu de 20.
- ▶ Pour obtenir un risque global de 5%, il faut faire chacun des tests de Fisher avec un risque de 1,25%
- ▶ La variabilité du rendement pour chaque variété est caractérisée par le même σ^2 . Avec des tests séparés, on a 4 estimations différentes pour le même σ^2 !

2ème possibilité :

Analyse de la variance à 2 facteurs sans répétitions :

Tableau issu du logiciel R

	Df	Sum Sq	Mean Sq	F value	Pr(> F)
variete	4	2.11447	0.52862	17.6928	5.692e-05 ***
blocs	3	0.15790	0.05263	1.7616	0.2079
Residuals	12	0.35853	0.02988		

2ème possibilité :

Analyse de la variance à 2 facteurs sans répétitions :

Tableau issu du logiciel R

	Df	Sum Sq	Mean Sq	F value	Pr(> F)
variete	4	2.11447	0.52862	17.6928	5.692e-05 ***
blocs	3	0.15790	0.05263	1.7616	0.2079
Residuals	12	0.35853	0.02988		

Rejet de $H_0 : \mu_{1.} = \mu_{2.} = \dots = \mu_{5.}$ car

$F_{obs} = 17.6928 > 3.259167 = F_{0.95}$ où

$F_{0.95}$ est le quantile d'ordre 0.95 d'une Fisher à $5 - 1$ et

$(5 - 1) \times (4 - 1)$ degrés de liberté.

De plus la p.value $\mathbf{P} = 5.692e - 05 < 0.05$

Ici, σ^2 est estimé par 0.02988

3.3. Discussion

Simulation avec $\mu = 4.13$ et $\sigma = 0.1728294$

On suppose que les moyennes réelles des variétés sont :

V1	V2	V3	V4	V5
4.03	4.03	4.03	4.03	4.53

Calcul de la probabilité de rejeter $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ en utilisant

- 1) l'analyse à 1 facteur P_1
 - 2) l'analyse à 2 facteurs P_2
- suivant les moyennes par blocs :

3.3. Discussion

Simulation avec $\mu = 4.13$ et $\sigma = 0.1728294$

On suppose que les moyennes réelles des variétés sont :

V1	V2	V3	V4	V5
4.03	4.03	4.03	4.03	4.53

Calcul de la probabilité de rejeter $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ en utilisant

- 1) l'analyse à 1 facteur P_1
 - 2) l'analyse à 2 facteurs P_2
- suivant les moyennes par blocs :

►

B1	B2	B3	B4
4.13	4.13	4.13	4.13

 alors $P_1 = 0.951$ $P_2 = 0.931$

3.3. Discussion

Simulation avec $\mu = 4.13$ et $\sigma = 0.1728294$

On suppose que les moyennes réelles des variétés sont :

V1	V2	V3	V4	V5
4.03	4.03	4.03	4.03	4.53

Calcul de la probabilité de rejeter $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ en utilisant

1) l'analyse à 1 facteur P_1

2) l'analyse à 2 facteurs P_2

suivant les moyennes par blocs :

▶	<table border="1"><tr><th>B1</th><th>B2</th><th>B3</th><th>B4</th></tr><tr><td>4.13</td><td>4.13</td><td>4.13</td><td>4.13</td></tr></table>	B1	B2	B3	B4	4.13	4.13	4.13	4.13	alors $P_1 = 0.951$	$P_2 = 0.931$
	B1	B2	B3	B4							
4.13	4.13	4.13	4.13								

▶	<table border="1"><tr><th>B1</th><th>B2</th><th>B3</th><th>B4</th></tr><tr><td>3.83</td><td>4.03</td><td>4.23</td><td>4.43</td></tr></table>	B1	B2	B3	B4	3.83	4.03	4.23	4.43	alors $P_1 = 0.282$	$P_2 = 0.946$
	B1	B2	B3	B4							
3.83	4.03	4.23	4.43								

3.3. Discussion

Simulation avec $\mu = 4.13$ et $\sigma = 0.1728294$

On suppose que les moyennes réelles des variétés sont :

V1	V2	V3	V4	V5
4.03	4.03	4.03	4.03	4.53

Calcul de la probabilité de rejeter $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$ en utilisant

1) l'analyse à 1 facteur P_1

2) l'analyse à 2 facteurs P_2

suivant les moyennes par blocs :

▶	<table border="1"><tr><th>B1</th><th>B2</th><th>B3</th><th>B4</th></tr><tr><td>4.13</td><td>4.13</td><td>4.13</td><td>4.13</td></tr></table>	B1	B2	B3	B4	4.13	4.13	4.13	4.13	alors $P_1 = 0.951$ $P_2 = 0.931$
	B1	B2	B3	B4						
4.13	4.13	4.13	4.13							

▶	<table border="1"><tr><th>B1</th><th>B2</th><th>B3</th><th>B4</th></tr><tr><td>3.83</td><td>4.03</td><td>4.23</td><td>4.43</td></tr></table>	B1	B2	B3	B4	3.83	4.03	4.23	4.43	alors $P_1 = 0.282$ $P_2 = 0.946$
	B1	B2	B3	B4						
3.83	4.03	4.23	4.43							

Supériorité de l'analyse à deux facteurs.

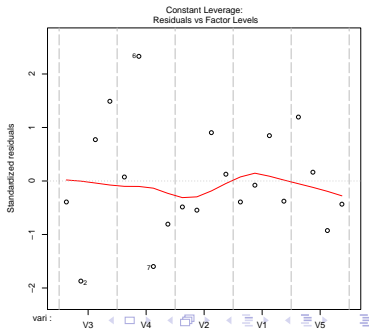
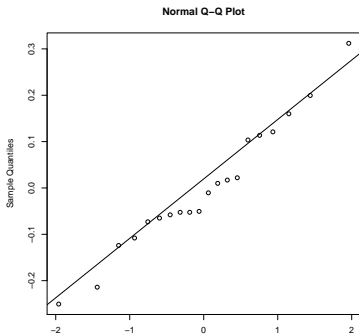
Validation des hypothèses :

Validation des hypothèses :

- ▶ **Egalité des variances ?** Test de Bartlett.
Même variance pour les variétés ? p -value = 0.4088
Même variance par bloc ? p -value = 0.8977
Dans les deux cas, non-rejet de l'égalité des variances.

Validation des hypothèses :

- ▶ **Egalité des variances ?** Test de Bartlett.
Même variance pour les variétés ? p -value= 0.4088
Même variance par bloc ? p -value= 0.8977
Dans les deux cas, non-rejet de l'égalité des variances.
- ▶ **Normalité des résidus ?**
Etude graphique :



Critique du plan en bloc aléatoire complet

Si le modèle était :

$$y_{ij} = \mu + v_i + b_j + (bv)_{ij} + \epsilon_{ij} \text{ avec } \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

Il ne serait pas possible d'estimer le terme d'interaction $(bv)_{ij}$ car pour un couple (V_i, B_j) pas de répétitions.

Quelle suite à l'étude ?

Si non-rejet de $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$, on ne met pas en évidence une différence entre les variétés.

Si rejet de $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$, on peut souhaiter connaître quelles sont les moyennes différentes.

Il existe une grande quantité de comparaisons multiples de moyennes. Toutes possèdent le défaut déjà signalé :

Non rejet de $H_0 : \mu_1 = \mu_2$ et non-rejet de $H_0 : \mu_2 = \mu_3$ n'entraîne pas non-rejet de $H_0 : \mu_1 = \mu_3$!

Méthode implantée par défaut dans R : test de Holm. Pairwise comparisons using t tests with pooled SD

	V1	V2	V3	V4
V2	0.85084	—	—	—
V3	0.08059	0.26728	—	—
V4	0.26728	0.66341	0.85084	—
V5	0.00420	0.00119	$3.5e - 05$	0.00013

$\mu_5 \neq \mu_1 = \mu_2 = \mu_3 = \mu_4$

4. Plus de 2 facteurs

$I = 4$ variétés, $J = 4$ blocs, $K = 4$ traitements chimiques

Pour faire un plan complet, il faut au moins $4 \times 4 \times 4 = 64$ unités expérimentales.

On peut réduire en utilisant un carré latin.

Par exemple :

	1	2	3	4
1	3	4	2	1
2	2	3	1	4
3	1	2	4	3
4	4	1	3	2

A un bloc j et une variété i , on attribue le traitement k qui se trouve à l'intersection de la ligne j et de la colonne i du carré latin.

$B4$	$(V4, T2)$	$(V3, T3)$	$(V1, T4)$	$(V2, T1)$
$B3$	$(V3, T4)$	$(V1, T1)$	$(V4, T3)$	$(V2, T2)$
$B2$	$(V2, T3)$	$(V1, T2)$	$(V4, T4)$	$(V3, T1)$
$B1$	$(V1, T3)$	$(V2, T4)$	$(V3, T2)$	$(V4, T1)$