

**CHAPITRE 5 - COMPARAISON DE PLUSIEURS TRAITEMENTS:
L'ANALYSE DE LA VARIANCE**

•BUT:

Maîtriser les concepts, postulats et calculs de l'analyse de la variance et être capable de les appliquer correctement à une situation expérimentale donnée.

•OBJECTIFS D'APPRENTISSAGE:

- 5.1. L'étudiant doit être capable de définir en ses propres termes, les concepts de "variabilité intra échantillons" et de "variabilité inter échantillons".
- 5.2. L'étudiant doit être en mesure d'expliquer la relation existant entre la différence réelle entre les moyennes des échantillons à comparer et la différence entre les estimés de la variance (intra et inter) utilisés pour calculer la valeur F .
- 5.3 Une situation expérimentale étant présentée, l'étudiant doit être capable d'identifier les différentes sources de variations et de décomposer les degrés de liberté correspondants
- 5.4. Pour une situation expérimentale donnée, l'étudiant doit être en mesure de formuler les hypothèses de recherche pertinentes.
- 5.5. Une situation expérimentale et les résultats en découlant étant présentés, l'étudiant doit être en mesure d'effectuer les calculs relatifs à une analyse de la variance.
- 5.6. Les résultats de l'analyse de la variance étant obtenus, l'étudiant doit être en mesure de construire correctement un tableau d'ANOVA.
- 5.7. L'étudiant doit être capable de lire dans une table de F la valeur critique à un niveau de probabilité donné, afin d'interpréter la valeur F calculée de l'ANOVA et de tirer les conclusions appropriées.
- 5.8. L'étudiant doit être en mesure d'expliquer la relation existant entre le test de F de l'ANOVA et le test de T .
- 5.9. Les résultats d'une ANOVA étant présentés, l'étudiant doit être en mesure d'effectuer le calcul de l'erreur-type de la moyenne d'un traitement ou d'une différence entre deux moyennes de traitements.
- 5.10. L'étudiant doit être capable d'expliquer la différence entre un modèle mathématique de type I et un modèle de type II de l'analyse de variance.
- 5.11. L'étudiant doit être en mesure de nommer et d'expliquer les différents postulats de l'analyse de variance.
- 5.12. L'étudiant doit être capable de nommer les différents tests de vérification de l'homogénéité de la variance et d'énoncer les principales règles relatives à la décision de transformer ou non les données.
- 5.13. L'étudiant doit être en mesure de nommer les principales actions à prendre pour remédier à un problème d'hétérogénéité de la variance.
- 5.14. L'étudiant doit être en mesure de nommer les principaux types de transformations des données possibles et de dire en quelles occasions chacune de ces transformations doit être utilisée.

LECTURES SUGGÉRÉES:

- Steel and Torrie, sections 7.3, 7.4, 7.5 (modèles fixe vs aléatoire) et 7.10 (postulats)

5.1 - COMPARAISON DE L'ANALYSE DE VARIANCE ET DU TEST DE T

Afin de comparer l'analyse de variance à ce que nous connaissons déjà, le test de t, voyons comment s'effectue l'analyse de variance à partir de deux échantillons seulement. L'exemple suivant est tiré de Little and Hills, p. 31:

Tableau 5.1) Rendement (100 lbs/acre) de deux cultivars de blé
(exemple tiré de Little and Hills, p. 31)

Cultivar	Répétitions						$Y_i.$	$\bar{Y}_i.$
1	19	14	15	17	20		85	17 = $\bar{Y}_1.$
2	23	19	19	21	18		<u>100</u>	<u>20</u> = $\bar{Y}_2.$
							<u>185</u>	<u>18.5</u> = $\bar{Y}_{..}$

- 1) On calcule d'abord la variance de chacun des échantillons, S_1^2 et S_2^2 , et on assume que ces deux variances d'échantillons estiment en fait une variance commune, σ^2 , la variance de la population dont ont été tirés les deux échantillons de façon aléatoire. On a donc:

$$S_1^2 = \frac{\sum (Y_{1j} - \bar{Y}_{1.})^2}{r - 1}$$

$$= \frac{(19 - 17)^2 + \dots + (20 - 17)^2}{5 - 1} = \frac{26}{4} = 6.5 \text{ , et}$$

$$S_2^2 = \frac{\sum (Y_{2j} - \bar{Y}_{2.})^2}{r - 1}$$

$$= \frac{(23 - 20)^2 + \dots + (18 - 20)^2}{5 - 1} = \frac{16}{4} = 4.0$$

et on met en commun ces deux estimés de variance par l'utilisation d'une moyenne, puisque les échantillons sont de dimensions égales, et on obtient un estimé de la variance intra traitements:

$$S^2_{intra} = \frac{S_1^2 + S_2^2}{2} = \frac{6.5 + 4.0}{2} = 5.25 \text{ (Variance intra)}$$

- 2) Si on assume H_0 vraie: les deux échantillons sont des échantillons aléatoires tirés d'une même population,
• alors $\bar{Y}_1.$ et $\bar{Y}_2.$ estiment la même moyenne de la population, μ
-> on estime donc la variance des moyennes des échantillons 1 et 2.

$$S_y^2 = \frac{\sum (\bar{Y}_i. - \bar{Y}_{..})^2}{n-1}$$

$$= \frac{(17 - 18.5)^2 + (20 - 18.5)^2}{2-1} = 4.5$$

• puisque l'on sait que $S_y^2 = \frac{S^2}{r}$

alors $S_{\text{inter}}^2 = r S_y^2$

$= 5 \times 4.5 = 22.5$ (Variance inter)

- 3) On a donc deux mesures de la variance

S_{inter}^2 -> variabilité inter échantillons.

S_{intra}^2 -> variabilité intra échantillons.

- si $H_0: \mu_1 = \mu_2$ (ou $\mu_1 - \mu_2 = 0$) est vraie,

alors $\frac{S_{\text{inter}}^2}{S_{\text{intra}}^2} \approx 1$ puisque S_{inter}^2 et S_{intra}^2 estiment la même variance σ^2

- on peut alors calculer la probabilité d'obtenir des estimés différents de σ^2 en estimant la valeur F calculée et en comparant la valeur obtenue à la valeur F théorique des tables.
- on a donc:

$$F = \frac{S_{\text{inter}}^2 \text{ (numérateur)}}{S_{\text{intra}}^2 \text{ (dénominateur)}}$$

- si les deux échantillons de populations ont des moyennes différentes,

alors $S_{\text{inter}}^2 > S_{\text{intra}}^2$ et $F \uparrow$

Pour l'exemple de comparaison de deux cultivars de blé,

$$F = \frac{22.5}{5.25} = 4.29 \text{ avec } \begin{cases} 1 \text{ d.l. au numérateur} \\ 8 \text{ d.l. au dénominateur} \end{cases}$$

Chaque échantillon a 5 observations

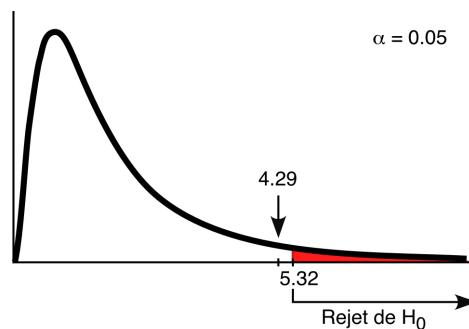
$\Rightarrow 5 - 1 = 4 \text{ d.l.}$

donc pour 2 échantillons: $4 + 4 = 8$

- Table de F -> on cherche, à partir de la courbe de distribution théorique, la probabilité de trouver une valeur $F > 4.29$

$F(1, 8) = 4.29$ avec une probabilité d'environ 7%

- La figure suivante illustre la position de la valeur 4.29 par rapport à l'ensemble des valeurs de F possibles. On peut constater que la valeur 4.29 se situe dans la zone d'acceptation de H_0 , soit en deçà de la valeur de F critique (F tables) au niveau alpha de 0.05.



Question: les deux moyennes diffèrent-elles réellement ou si la différence observée est uniquement l'effet du hasard?

- Si on prenait de nombreuses mesures sur des échantillons similaires, avec le même niveau de précision (le même nombre de répétitions), on devrait s'attendre à observer, juste par chance, une différence de l'ordre de 300 lbs/acre dans environ 7% des cas.
- On prend donc plus de risques en disant que les deux moyennes sont réellement différentes. Par contre, si elle est réelle, une différence de 300 lbs/acre peut représenter un gain économique non négligeable pour un producteur de grains. Il faut donc tenir compte du rapport risques-bénéfices dans la prise de décision.
- Décision: -> Il faudra vraisemblablement effectuer une expérience supplémentaire afin d'obtenir plus d'information. Le simple fait d'ajouter des répétitions ou de répéter l'expérience aura pour effet d'accroître la précision avec laquelle on estime la différence entre les deux moyennes. Avec plus de répétitions, on pourra probablement détecter de plus petites différences significatives.

5.2 - LE TABLEAU DE L'ANALYSE DE LA VARIANCE

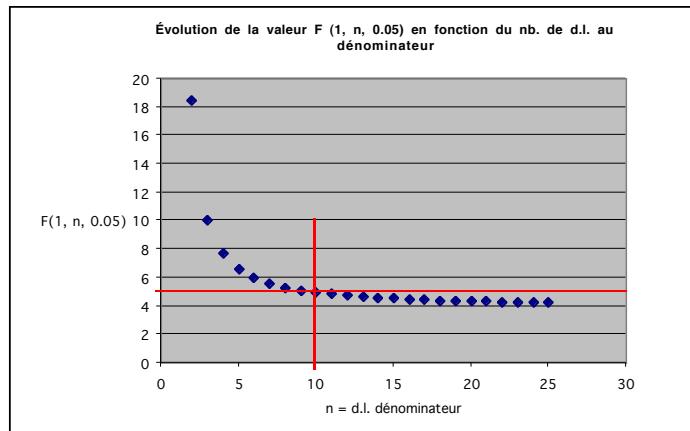
- technique introduite par R. A. Fisher.
- le tableau n'est rien de plus qu'une façon pratique de présenter les résultats des différents calculs.
- ANOVA: essentiellement un procédé arithmétique visant à répartir (partitionner) une somme des carrés totale en ses différentes composantes associées à des sources de variations connues.

Tableau 5.2) Tableau de l'ANOVA pour l'exemple de comparaison de 2 cultivars de blé.

<u>Source de variation</u>	<u>d.l.</u>	<u>SCécarts</u>	<u>MCécarts</u>	<u>Fcalc.</u>	<u>Fthéor.</u>
Cultivars	1	22.5 ÷ 1	22.5	4.29	5%: 5.32
Erreur	8	42.0 ÷ 8	5.25	(5)	10%: 3.46
Total	9	64.5	(4)	(6)	

- IMPORTANT : avant de débuter toute expérience, on devrait toujours effectuer la décomposition des sources de variation et des degrés de liberté afin de s'assurer qu'on disposera bien d'un terme d'erreur valide pour tester les différences entre les traitements. On s'assurera aussi, dans la mesure du possible, d'avoir un minimum de 10 degrés de liberté associés au terme d'erreur, de façon à obtenir une puissance minimale pour les comparaisons des traitements. Ces vérifications doivent être effectuées au moment de la planification de l'expérience, car après il sera trop tard pour apporter des correctifs.

Pourquoi un minimum de 10 d.l. à l'erreur?



Recette pour calculer l'ANOVA à la mitaine

1) Faire le plan de l'ANOVA

-> identifier les sources de variations

2) Décomposer les degrés de liberté

3) Calculer les sommes des carrés des écarts

- SC_{tot} .
- SC_{trait} .
- $SC_{\text{erreur}} = SC_{\text{tot}} - SC_{\text{trait}}$.

4) Calculer les moyennes des carrés des écarts

$$MC_{\text{cult}^{\text{rs}}} = \frac{SC_{\text{cult}^{\text{rs}}}}{d.l._{\text{cult}^{\text{rs}}}}$$

$$MC_{\text{erreur}} = \frac{SC_{\text{erreur}}}{d.l._{\text{erreur}}}$$

5) Obtenir la valeur F calculée.

$$F_{\text{calc}} = \frac{MC_{\text{cultivar (inter)}}}{MC_{\text{erreur (intra)}}}$$

6) Vérifier la valeur F requise pour le niveau de signification α

Table de F

d.l. numérateur	d.l. numérateur
d.l. dénominateur	

Tableau 5.3) Forme générale de l'ANOVA à une voie (plan entièrement aléatoire)

Source de variation	d.l.	S.C.	M.C.	Fcalc	Ftables
Traitement	#trait - 1	S.C. trait	$\frac{S.C. trait}{d.l. trait}$	$\frac{M.C. trait}{M.C.E.}$	$F(\alpha, d.l_{trait}, d.l_{erreur})$
Erreur	différence	S.C. erreur	$\frac{S.C. erreur}{d.l. erreur}$		
Total	#tot - 1	S.C. tot			

Suite à l'examen du tableau précédent, il faut noter les remarques suivantes:

- les degrés de liberté (d. l.) sont additifs
- les sommes de carrés (S.C.) sont additifs: $S.C. tot = S.C. trait + S.C. erreur$
- les moyennes de carrés (M.C.) ne sont pas additives

Donc, on ne peut pas additionner ou soustraire les moyennes de carrés: il faut toujours utiliser les degrés de liberté et les sommes de carrés pour calculer les moyennes de carrés et les valeurs F correspondantes.

Comme mentionné auparavant, la valeur F calculée résulte du rapport de deux variances. Les variances (carrés moyens) sont distribuées selon des khi-deux. Comme on l'a vu dans le chapitre 3, un rapport de deux variables distribuées selon des khi-deux est distribué selon une F .

$$F_{calc} = \frac{MC_{traitements}}{MC_{erreur}} = \frac{\text{Mesure de variance inter - traitements}}{\text{Mesure combinée de variance intra - traitements}}$$

A titre de conclusion, il faut se rappeler que l'ANOVA n'est qu'un procédé mathématique visant à répartir une somme des carrés totale en ses différentes composantes associées à des sources de variation connues. La présentation des résultats sous forme de tableau ne vise qu'à simplifier l'interprétation des résultats.

Tableau 5.4) Notation utilisée pour décrire le calcul des différentes sommes de carrés

		Traitement				
Observation		1	2	3	... i...	... t
1		Y ₁₁	Y ₂₁	Y ₃₁	Y _{i1}	Y _{t1}
2		Y ₁₂	Y ₂₂	Y ₃₂	Y _{i2}	Y _{t2}
3		Y ₁₃	Y ₂₃	Y ₃₃	Y _{i3}	Y _{t3}
j		Y _{1j}	Y _{2j}	Y _{3j}	Y _{ij}	Y _{tj}
n		Y _{1n}	Y _{2n}	Y _{3n}	Y _{in}	Y _{tn}
Total:		Y _{1..}	Y _{2..}	Y _{3..}	Y _{i..}	Y _{t..}
Moyenne:		$\overline{Y}_{1..}$	$\overline{Y}_{2..}$	$\overline{Y}_{3..}$	$\overline{Y}_{i..}$	$\overline{Y}_{t..}$

Y_{ij} = j^{ème} observation du i^{ème} traitement.

Y_{in} = n^{ière} observation du i^{ème} traitement.

$Y_{1..}$ = somme des observations pour le traitement 1.

$\bar{Y}_{1..}$ = moyenne des observations pour le traitement 1.

$Y_{..}$ = grand total (observations et traitements) $\sum_i Y_{i..} = Y_{..}$

$\bar{Y}_{..}$ = grande moyenne = $\frac{Y_{..}}{t_n}$ (si n égaux).

Tableau 5.5). Analyse de la variance à une voie, avec un nombre égal de répétitions (Traduit du Tableau 7.2 de Steel and Torrie)

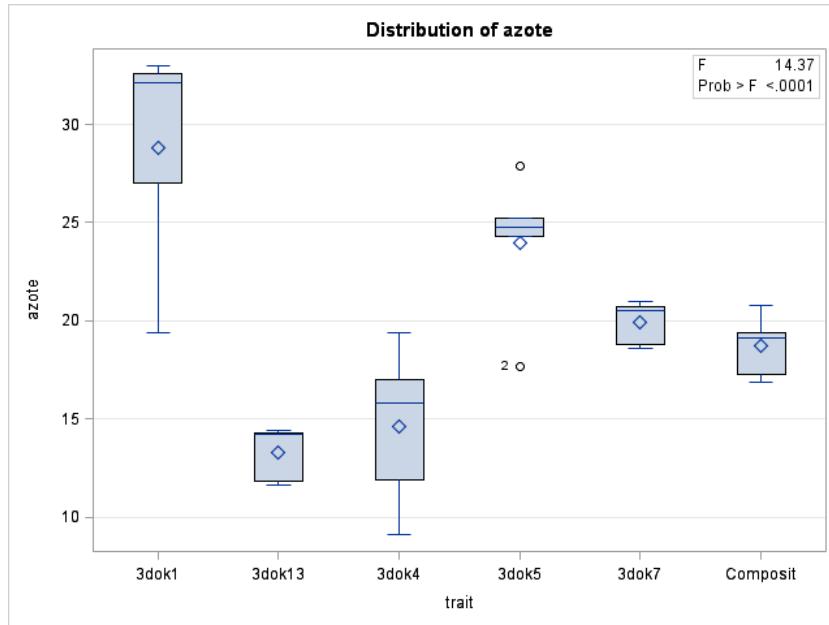
Source de variation	d.l.	Somme	des carrés	Moyennes	<i>F</i>
		Définition	Pratique	des carrés	
Traitements	t - 1	$r \sum_i (\bar{y}_{i..} - \bar{y}_{..})^2$	$\sum_{i,j} \frac{y_{ij}^2}{r} - \frac{y_{..}^2}{rt}$	$\frac{SC_{trait.}}{t - 1}$	
Erreur	t (r - 1)	$\sum_{i,j} (y_{ij} - \bar{y}_{i..})^2$	par soustraction	$\frac{SC_{Erreur}}{t(r - 1)}$	
Total	rt - 1	$\sum_{i,j} (y_{ij} - \bar{y}_{..})^2$	$\sum_{i,j} y_{ij}^2 - \frac{y_{..}^2}{rt}$		

Pour une source de variation de l'ANOVA, la moyenne des carrés (M.C., aussi appelée carré moyen) est obtenue en divisant la somme des carrés (S.C.) par le nombre de degrés de liberté (d.l.) correspondant.

Calcul de l'ANOVA à une voie ...

Tableau 5.6). Contenu en azote de plantes de trèfle rouge inoculées de combinaisons de *R. trifolii* et *R. meliloti* (mg). (Adaptation du Tableau 7.1 de Steel and Torrie)

Compila- tion	3DOK1	3DOK5	3DOK4	3DOK7	3DOK13	Mé- lange	Total
	19.4	17.7	17.0	20.7	14.3	17.3	
	32.6	24.8	19.4	21.0	14.4	19.4	
	27.0	27.9	9.1	20.5	11.8	19.1	
	32.1	25.2	11.9	18.8	11.6	16.9	
	33.0	24.3	15.8	18.6	14.2	20.8	
$\sum_j y_{ij} = y_{i..}$	144.1	119.9	73.2	99.6	66.3	93.5	596.6 = $y_{..}$
$\sum_j y_{ij}^2$	4 287.53	2 932.27	1 139.42	1 989.14	887.29	1 758.71	12 994.36 = $\sum_{ij} y_{ij}^2$
$\bar{y}_{i..}$	$\bar{y}_{1..} = 28.8$	$\bar{y}_{2..} = 24.0$	14.6	19.9	13.3	18.7	Moyennes



Calcul de l'ANOVA

- Calcul du facteur de correction:

$$C = \frac{Y_{..}^2}{rt} = \frac{\left(\sum_{ij} Y_{ij}\right)^2}{rt}$$

$$= \frac{(596.6)^2}{(5)(6)} = 11,864.38$$

- Calcul de la somme des carrés totale:

$$SC_{\text{tot.}} = \sum_{ij} Y_{ij}^2 - C$$

$$= 12,994.36 - 11,864.38 = 1,129.98$$

- Calcul de la somme des carrés attribuable à l'effet des traitements (inter groupes):

$$SC_{\text{trait.}} = \sum_i \frac{Y_i^2}{r} - \frac{Y_{..}^2}{rt}$$

$$= \left(\frac{Y_1^2 + Y_2^2 + \dots + Y_t^2}{r} \right) - C$$

$$= \left(\frac{144.1^2 + 119.9^2 + \dots + 93.5^2}{5} \right) - 11,864.38 = 847.05$$

- Calcul de la somme des carrés attribuable à la variation à l'intérieur des traitements (intra groupes):

$$S.C._{\text{Erreur}} = S.C._{\text{totale}} - S.C._{\text{traitements}}$$

$$= 1,129.98 - 847.05 = 282.93$$

- La somme des carrés de l'erreur peut aussi être obtenue en mettant en commun les sommes de carrés intra traitements:

$$\begin{aligned} S.C.\text{Erreur} &= \sum_i (\sum_j Y_{ij}^2 - \frac{Y_i^2}{r}) \\ &= (4,287.53 - \frac{144.1^2}{5}) + (2,932.27 - \frac{119.9^2}{5}) + \dots \\ &\quad + (1,758.71 - \frac{93.5^2}{5}) = 282.93 \end{aligned}$$

- Les résultats numériques de l'ANOVA sont présentés sous forme de tableau:

Tableau 5.7) Analyse de variance pour les données du tableau 5.6

Source de variation	d.l.	S.C.	M.C	F
Inter cultures	5	847.05	169.41	14.37**
Intra cultures	24	282.93	11.79 ¹	
Total	29	1,129.98		

¹ Estimé de S^2 , terme d'erreur généralisé (estimé de σ^2 commun)

- Calcul de l'erreur-type d'une moyenne de traitement ($r = \text{nb. de répétitions}$)

$$S_{\bar{y}} = \sqrt{\frac{S^2}{r}} = \sqrt{\frac{11.79}{5}} = 1.54 \text{ mg}$$

- Calcul de l'erreur-type d'une différence entre deux moyennes de traitements:

$$S_{\bar{y}_i - \bar{y}_{i'}} = \sqrt{\frac{2S^2}{r}} = \sqrt{\frac{2(11.79)}{5}} = 2.17 \text{ mg} \quad (i \neq i')$$

- Calcul du coefficient de variation de l'expérience:

$$CV (\%) = \frac{S}{\bar{y}} \times 100 = \frac{\sqrt{11.79}}{19.9} \times 100 = 17.2\%$$

L'analyse de cet exemple à l'aide du progiciel SAS est présentée à la section 3.2- Comparaisons multiples pour un plan entièrement aléatoire du cahier de laboratoire (fichier excast71.sas).

Terme d'erreur généralisé de l'ANOVA

Le terme S^2 est un terme d'erreur généralisé. S^2 est un estimé de σ^2 , la vraie variance de l'erreur, qui est une mesure de la variation observée entre les observations recevant le même traitement. Le calcul du terme d'erreur généralisé est basé sur un postulat selon lequel il existe une variance commune σ^2 . S^2 est un estimé valide de σ^2 uniquement si ce postulat est vrai.

Les composantes individuelles de la variance intra traitement sont basées sur un petit nombre de degrés de liberté, et peuvent varier considérablement autour de σ^2 . Les composantes individuelles de la variance intra traitement peuvent donc varier considérablement autour de σ^2 ,

et ne constituent donc pas d'aussi bons estimés de σ^2 que le regroupement de l'ensemble des composantes.

La précision de l'estimation de la variance de l'erreur fonctionne un peu sur le même principe que pour les moyennes. La précision d'une moyenne évaluée à partir de 24 observations est beaucoup plus grande que la précision d'une moyenne de 4 observations. Plus le nombre d'observations servant au calcul de la moyenne est élevé, plus la précision avec laquelle la moyenne est évaluée est grande. Donc, l'erreur-type de la première moyenne sera inférieure à celle de la deuxième.

- exemple: supposons que l'expérience comporte 6 traitements et qu'on veuille comparer les traitements A et B. On aurait donc les rapports suivants pour le test de t et une comparaison simple construite sur un test de F , selon qu'il s'agit d'un test de F dans un contexte d'analyse de la variance (ANOVA) ou d'un test de F utilisé pour un contraste:

$$\text{test de } t = \frac{\text{Différence entre les 2 trait}^s}{\text{Variance due au hasard (2 trait}^s)}$$

$$\text{test de } F (\text{ANOVA}) = \frac{\text{Variance entre les 6 trait}^s}{\text{Variance due au hasard (6 trait}^s)}$$

$$\text{test de } F (\text{contraste qualitatif}) = \frac{\text{Différence entre les 2 groupes}}{\text{Variance due au hasard (6 trait}^s)}$$

Dans le cas du test de F , la variation due au hasard est donc estimée à partir d'un plus grand nombre d'observations, et fournit un meilleur estimé de la variance de l'erreur. On recommande habituellement un minimum de 10 degrés à l'erreur pour un minimum de précision. On peut facilement vérifier que les valeurs de F critiques nécessaires pour déclarer une différence significative augmentent rapidement lorsque le nombre de degrés de liberté est inférieur à 10, qui correspond à un point d'inflexion dans la courbe.

Le test de t et le test de F donnent des résultats équivalents dans une situation où on compare deux traitements. On a déjà montré que:

$$t_{\alpha/2(n_2)}^2 = F_{\alpha}(1, n_2)$$

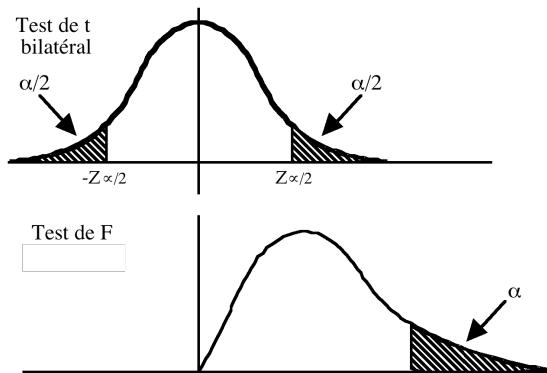


Fig. 5.1) Relation entre le test de t et le test de F (les courbes ne sont pas tracées sur une échelle commune)

Comme discuté ci-haut, l'ANOVA présente l'avantage, lorsque les variances des traitements sont homogènes, de n'utiliser qu'un seul critère de décision, le terme d'erreur généralisé. Le test de t pourra cependant être utile dans des situations de variances hétérogènes ou de fichiers de données déséquilibrés.

Le fondement de l'ANOVA: la partition des sources de variation

Le fondement de l'ANOVA

On a donc vu que: $S.C.\text{tot.} = S.C.\text{trait.} + S.C.\text{Erreur}$
 $= S.C.\text{inter} + S.C.\text{intra}$

C'est-à-dire: $\sum_{i=1}^t \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^t \sum_{j=1}^n (\bar{Y}_i - \bar{Y}_{..})^2 + \sum_{i=1}^t \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2$

The equation $\sum_{i=1}^t \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^t \sum_{j=1}^n (\bar{Y}_i - \bar{Y}_{..})^2 + \sum_{i=1}^t \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2$ is shown. Three red arrows point from the terms $(\bar{Y}_i - \bar{Y}_{..})^2$, $(Y_{ij} - \bar{Y}_i)^2$, and $(Y_{ij} - \bar{Y}_{..})^2$ respectively to the terms \bar{Y}_i , Y_{ij} , and $\bar{Y}_{..}$ in the equation $Y_{ij} = \mu + T_i + \varepsilon_{ij}$.

$$Y_{ij} = \mu + T_i + \varepsilon_{ij}$$

et où $\varepsilon_{ij} \sim NID(0, \sigma^2)$

où: Y_{ij} = réponse observée pour la $j^{\text{ème}}$ obsⁿ du $i^{\text{ème}}$ trait.

μ = moy. générale

T_i = effet du $j^{\text{ème}}$ trait.

ε_{ij} = erreur aléatoire associée à la $j^{\text{ème}}$ obsⁿ du $i^{\text{ème}}$ trait.

donc, dans l'équation du modèle $Y_{ij} = \mu + t_i + \varepsilon_{ij}$

- la variabilité des Y_{ij} correspond à la S.C. totale
- μ est invariable. On associe généralement la perte de 1 d.l. au total à la correction pour la moyenne dans le modèle surparamétré.
- la variabilité des T_i correspond à la S.C. des traitements
- la variabilité des ε_{ij} correspond à la S.C. erreur.

• Modèle fixe ou aléatoire?

Il existe trois types d'analyses de variance, selon le modèle mathématique utilisé:

-> le modèle à effets fixes (type I - descriptif):

- l'intérêt porte sur la détection et l'estimation de différences entre les moyennes des traitements **du ou des groupes soumis à l'analyse**; l'inférence porte sur ces traitements particuliers.
- les traitements sont fixes et constituent une population finie.
- une répétition de l'expérience portera sur le même ensemble de traitements étudiés auparavant; on concentre l'attention sur ces traitements.
- ex: expérience visant à comparer 5 génotypes bien adaptés à une région.

-> le modèle à effets aléatoires (type II - prédictif):

- l'intérêt porte sur la généralisation des résultats de l'analyse **à la population d'où sont tirés les groupes** soumis à l'analyse; l'intérêt peut aussi porter sur l'estimation de l'importance relative des variances associées aux facteurs agissant sur le groupe soumis à l'étude. L'inférence porte sur la population de traitements.
- les traitements sont une sélection aléatoire d'une population de traitements.
- une répétition de l'expérience portera sur un nouvel ensemble de traitements, mais tirés aléatoirement de la même population de traitements.

- ex: expérience destinée à évaluer une collection génétique par échantillonnage aléatoire d'un certain nombre de représentants.

-> le modèle mixte:

- ce type de modèle hybride comporte à la fois des effets fixes et des effets aléatoires. Les variations du modèle mixte sont nombreuses.

5.3 - LES POSTULATS DE L'ANALYSE DE VARIANCE.

Les tests statistiques ne sont valides que dans des conditions déterminées, car ceux-ci sont construits en partant de postulats bien définis. Sans ces postulats, ou conditions d'utilisation, il serait tout simplement impossible de construire des tests statistiques. Il faut donc connaître les limites de validité des différents tests, et prendre conscience des dangers reliés à «l'automatisme» dans le traitement statistique des données.

Comme on l'a vu précédemment, l'ANOVA est basée sur un modèle mathématique. Il importe donc de voir quels sont les postulats émis lors de la construction du modèle en question, c'est-à-dire d'examiner les conditions que les données doivent remplir pour que l'ANOVA soit valide.

Le respect des postulats de l'analyse de variance est surtout important dans le cas où les conclusions de notre expérience sont généralisées à l'ensemble des individus formant la population d'où provient l'échantillon (effets aléatoires). Dans le cas où on tient strictement à décrire le matériel sur lequel on travaille (effets fixes: les conclusions portent uniquement sur l'échantillon lui-même) les conséquences du non respect des postulats sont moins importantes, sauf si on désire effectuer des comparaisons simples (dites à un degré de liberté). La partie algébrique de l'ANOVA, c'est-à-dire la décomposition de la somme des carrés totale en des sommes des carrés associées à des effets connus, est toujours valide comme une description des données, ou comme un sommaire de leurs propriétés, quelle que soit la nature des données. C'est quand l'ANOVA est utilisée comme méthode d'inférence statistique, c'est-à-dire pour tirer des conclusions au sujet de la population échantillonnée, que les postulats doivent être respectés, et ce dans la mesure de leur importance respective.

Il y a donc deux types d'analyses de variance, selon le modèle utilisé:

- **le modèle à effets fixes (type I):** l'intérêt porte sur la détection et l'estimation de différences entre les moyennes des traitements ou des groupes soumis à l'analyse.
- **le modèle à effets aléatoires (type II):** l'intérêt porte sur la généralisation des résultats de l'analyse à la population d'où sont tirés les groupes soumis à l'analyse. L'intérêt peut aussi porter sur l'estimation de l'importance des variances dues aux facteurs de variation agissant sur le groupe soumis à l'étude.

1) Indépendance des erreurs expérimentales

- les erreurs des différents traitements doivent être indépendantes. Par le terme «erreur», on n'entend pas «faute», mais plutôt **variation non contrôlée**.
- l'indépendance des erreurs expérimentales est assurée par la distribution aléatoire des unités expérimentales aux différents traitements.
- par distribution ou répartition aléatoire, on entend distribution réellement au hasard (ie. de façon objective, indépendamment de la connaissance ou de la volonté de l'expérimentateur), et non au recours à une forme quelconque de pseudo-hasard. Les façons acceptables d'effectuer une distribution aléatoire sont les suivantes:

- par **tirage aléatoire physique**. Par exemple, on peut utiliser une série de disques numérotés selon le nombre de traitements à distribuer, les mélanger dans un contenant, et effectuer un tirage aveugle.
- par **l'utilisation d'une table de nombres aléatoires**. Lorsqu'on utilise une telle table, il est important de choisir un point de départ de façon aléatoire!
- par **tirage aléatoire informatique**. De plus en plus de programmes informatiques permettent d'effectuer la distribution aléatoire des unités expérimentales. Par exemple, la procédure PLAN de SAS permet de facilement générer des plans d'expérience.
- l'indépendance des erreurs expérimentales est une condition essentielle à la réalisation de l'analyse de variance: **sans distribution au hasard des unités expérimentales, il n'y a pas de calcul des valeurs P possible!!!** (Même en présence d'effets fixes). Le calcul mathématique des sommes des carrés est toujours possible, à titre indicatif, mais la non validité des valeurs P qui y sont associées empêche la généralisation des résultats).

2) Homogénéité de la variance de l'erreur expérimentale

- la condition d'homogénéité de la variances de l'erreur (ou homoscédasticité) est sans doute le second plus important postulat. D'autre part, on fait souvent face au problème d'hétérogénéité des variances.
- il faut, pour que l'analyse soit valide, que tous les groupes expérimentaux aient la même variance (ie. variance vraie, dépourvue de l'erreur d'échantillonnage). Si cette condition n'est pas remplie, il s'ensuit une perte d'information sur les effets des traitements, perte la plus importante dans le cas de comparaisons individuelles entre traitements.
- comme on l'a vu précédemment, la condition d'homogénéité de la variance est fondamentale à la construction d'un test de F , dont le terme d'erreur (dénominateur) est constitué de la mise en commun de l'ensemble des sources de variation intra-traitement. Par exemple, la présence d'une seule variance anormalement élevée pour l'un des groupes à comparer va augmenter de façon anormale la variance de l'erreur, et ainsi diminuer la sensibilité du test appliquée à l'ensemble des moyennes.
- la non-homogénéité de la variance a aussi pour effet de perturber l'équilibre entre les différentes comparaisons possibles: les groupes expérimentaux n'auront pas le même poids si leurs variances diffèrent. En effet, la degré de précision rattaché à la détermination d'une moyenne dépend de sa variance. Plus la variance est grande, moins la moyenne est précise, et vice-versa. Par exemple, dans le cas extrême d'une variance nulle, une seule détermination suffit à déterminer la valeur prise par la moyenne du groupe, puisque la valeur est constante.
- de façon générale, le test de F de l'analyse de variance est assez robuste face à la présence de variances inégales **si le nombre de répétitions est égal pour tous les traitements** (dimensions des échantillons égales). Par contre, les comparaisons simples (à 1 degré de liberté) entre les moyennes des niveaux de facteurs peuvent être affectées par des variances inégales, de telle sorte que les niveaux de signification réels et spécifiés puissent différer de façon marquée. Aussi, la procédure de comparaison multiple de Scheffé, basée sur la distribution de F , n'est pas affectée de façon substantielle par des variances inégales si les échantillons sont de dimensions égales.

3) Distribution normale de l'erreur expérimentale

- l'échantillon faisant l'objet de l'étude doit provenir d'une population normale. La normalité n'est pas importante pour le modèle à effets fixes, pourvu que l'écart de la normalité ne soit

pas d'une forme extrême (et de façon générale le degré d'aplatissement de la courbe a un effet plus important que le degré d'asymétrie sur l'inférence).

- le principal problème relié à la normalité de l'erreur expérimentale réside dans la vérification du postulat. Dans la plupart des cas, l'échantillon utilisé est de trop petite taille pour obtenir une représentation adéquate de la distribution parentale.
- le test de F est assez robuste au manque de normalité.
- la normalité peut être améliorée par des transformations mathématiques des données. Toutefois, en pratique on transforme rarement les données pour améliorer la normalité. Les transformations sont plutôt effectuées pour rendre les variances des différents traitements plus homogènes. Cependant, la plupart des transformations utilisées pour rendre les variances plus homogènes ont aussi pour effet d'améliorer la normalité.

4) Additivité des effets

- le modèle mathématique utilisé pour effectuer l'analyse de variance est un modèle linéaire additif. Pour que l'ajustement du modèle aux données expérimentales soit satisfaisant, il faut donc que les effets des différentes sources de variation (ie. traitements, répétitions, etc...) soient additifs.
- l'examen préliminaire des données avant l'analyse statistique demeure le meilleur moyen de vérifier l'additivité des effets des traitements. Une source importante de non-additivité réside dans la présence de données aberrantes ou extrêmes. L'élimination de telles données, toutefois, requiert la plus grande prudence. Avant d'éliminer une donnée qui semble aberrante, il faut être en mesure de fournir une explication sûre de la raison de la présence d'une telle valeur. Par exemple, la démonstration d'une erreur de lecture ou de transcription des données pourrait justifier l'enlèvement d'une donnée aberrante; par contre, une simple suspicion au sujet d'une valeur qui semble extrême ne pourrait justifier l'élimination de la valeur en question. Il existe certains tests statistiques permettant de décider si une valeur est aberrante ou non (ex: test de Chauvenet, qui enlève les valeurs aberrantes aux deux extrémités). De même, le test de Tukey pour la non-additivité permet de vérifier le degré d'additivité observé dans les données.
- la forme la plus courante de non-additivité se présente lorsque les effets des traitements ou des blocs sont multiplicatifs au lieu d'être additifs. Par exemple, l'effet d'un traitement sur un bloc consiste alors à le multiplier par un certain facteur, au lieu de lui ajouter quelque chose.
- comme dans le cas de la normalité, plusieurs transformations destinées à améliorer l'homogénéité de la variance ont aussi un effet sur l'additivité. Des quatre conditions nécessaires pour assurer la validité de l'ANOVA, la condition d'additivité des effets est probablement celle dont on se soucie le moins, sauf si on sait à l'avance le comportement du matériel expérimental avec lequel on projette de travailler.

5.4 - LA VÉRIFICATION DES POSTULATS

Vérification de l'homogénéité des variances

L'homogénéité de la variance est sans aucun doute le postulat le plus facile et le plus important à vérifier en pratique. L'examen attentif des données devrait révéler le besoin de tester l'homogénéité de la variance.

L'examen attentif des données

La figure suivante présente un exemple d'examen des données révélant un lien étroit entre les variances et les moyennes des traitements. Les valeurs mesurées sont les temps de désintégration (secondes) de quatre types de comprimés médicamenteux autodissolvants (les traitements), lorsque plongées dans un verre d'eau.

Examen attentif des données ...

TABLE 1.3.1

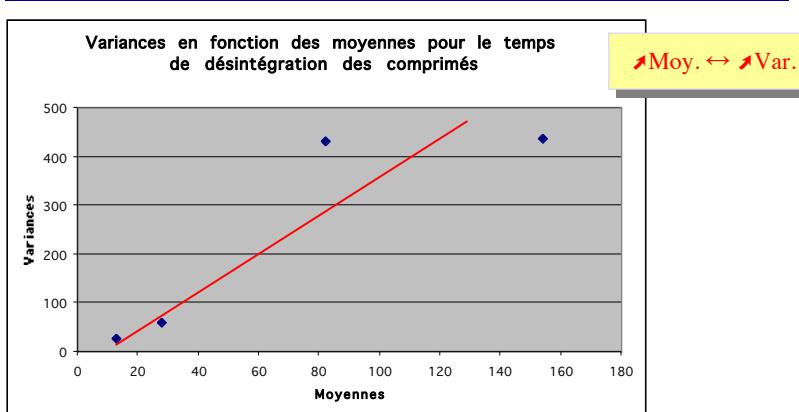
Time (seconds) of Disintegration of Tablets for Four Types of Tablets

Types of tablets							
	A	B	C	D			
20	42	8	12	50	124	151	178
28	25	10	24	67	72	125	151
36	24	12	10	90	78	180	152
16	31	16	19	103	70	149	161
25	33	9	10	90	76	175	118
Mean	28 ↘	13 ↘ ↘	82 ↗	154 ↗ ↗			
Variance	59.55 ↘	26.22 ↘ ↘	430.89 ↗	436.22 ↗ ↗			

Adapté de Anderson and McLean, 1974.

La mise en graphique des variances en fonction des moyennes constitue le meilleur outil pour détecter ce type de liaison entre les variances et les moyennes. La figure suivante présente le graphique des variances des temps de désintégration en fonction des moyennes correspondantes de chacun des types de comprimés

Examen attentif des données ...



Si l'examen préliminaire des variances des différents traitements le suggère, l'homogénéité de la variance peut être vérifiée au moyen de l'un des tests suivants:

- l'analyse graphique des résidus;
- le test de Levene;
- le test de Bartlett;
- le test de Burr-Foster;

- le test de Bartlett-Kendall (aussi appelé test de Box);
- le test de Hartley.

Les tests d'homogénéité de la variance précédents ont été développés au cours du siècle dernier, à une période où l'informatique était encore peu ou pas accessible. Ces tests ont été développés pour des situations expérimentales simples (ex. : plan entièrement aléatoire) et leur utilisation est douteuse avec des situations expérimentales plus complexes. Aujourd'hui, avec les logiciels et les ordinateurs modernes, on peut facilement procéder directement à l'analyse graphique des résidus, une technique utilisée en régression qui permet d'observer les patrons de distribution des résidus (termes d'erreur correspondant à chaque observation). On peut ainsi détecter tout patron de variation typique de situations problématiques susceptibles d'affecter la validité de l'ANOVA.

L'analyse graphique des résidus

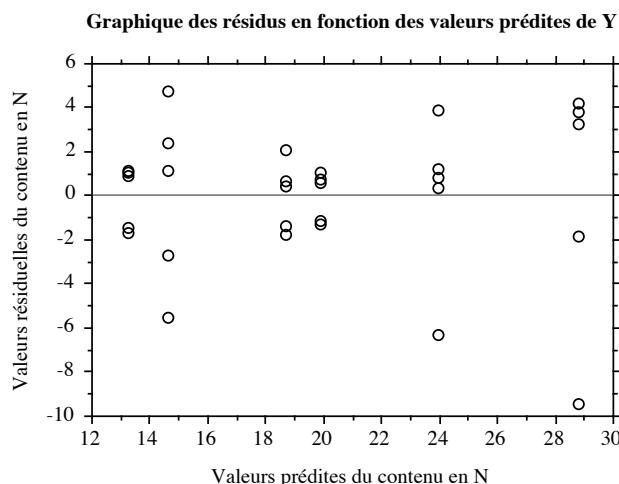
L'analyse graphique des résidus est basée sur le modèle mathématique sous-jacent à l'analyse de la variance. Ainsi, pour le plan d'expérience le plus simple, le plan entièrement aléatoire, le modèle mathématique est le suivant:

$$Y_{ij} = \mu + t_i + e_{ij};$$

chaque valeur prise par une observation (Y_{ij}) peut donc être expliquée comme étant la somme de la grande moyenne des observations de l'expérience (par rapport à laquelle les autres effets sont décrits sous forme de déviations), l'effet du $i^{\text{ème}}$ traitement (t_i) et un terme d'erreur correspondant à la $j^{\text{ème}}$ observation du $i^{\text{ème}}$ traitement (e_{ij}). Le graphique d'analyse des résidus le plus couramment utilisé consiste à mettre en évidence la relation entre les résidus et les valeurs prédites par le modèle. Les valeurs prédites sont obtenues à l'aide de l'équation suivante:

$$\hat{Y}_{ij} = \mu + t_i$$

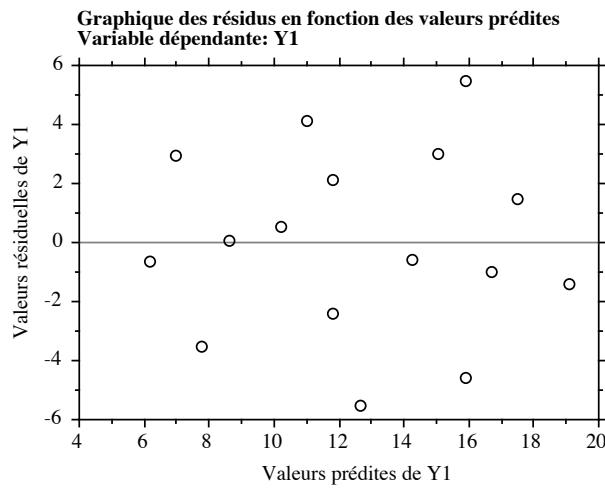
On obtient donc un graphique des valeurs résiduelles (e_{ij}) en fonction des valeurs prédites (\hat{Y}_{ij}), qui correspondent aux valeurs prédites par le modèle mathématique, excluant l'erreur expérimentale. Pour l'exemple des souches de *Rhizobium*, on obtient le graphique des résidus suivant, pour le contenu en azote:



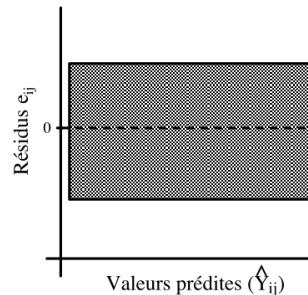
On peut facilement constater que le patron de distribution des résidus ne montre aucune tendance anormale. Les résidus sont dispersés de façon uniforme autour de la valeur zéro de l'ordonnée, qui représente les résidus exprimés en déviations par rapport à zéro. Même si

l'analyse des résidus présente l'avantage de permettre l'observation directe du comportement des valeurs résiduelles, elle présente aussi l'inconvénient d'une interprétation subjective. Dans certains cas il est très difficile de prendre une décision sur la seule base de l'observation du graphique des résidus; on recommande alors de compléter l'analyse des résidus avec un test d'homogénéité des variances comme ceux décrits plus loin.

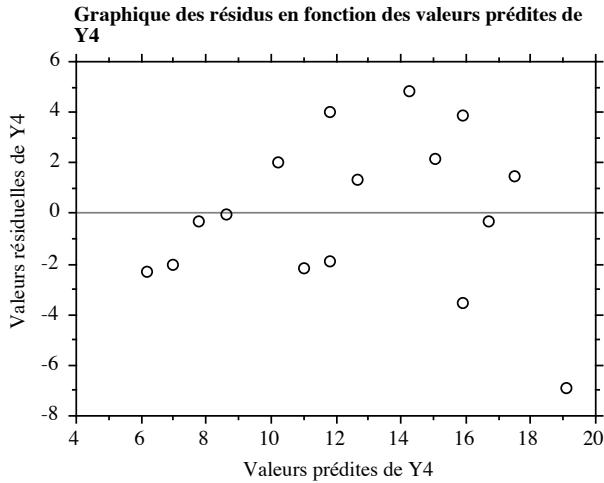
Certains patrons typiques de distribution des résidus *pathologiques* ont été identifiés. Voici quelques exemples; le premier exemple illustre un patron de distribution normal des résidus pour une expérience où les variances sont homogènes:



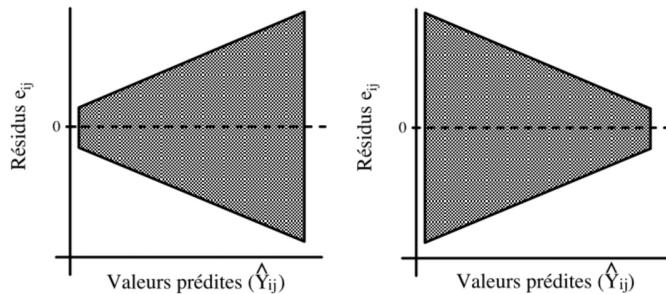
On peut voir que la variance des résidus autour de zéro est à peu près constante, peu importe la valeur des y ajustés (qui représentent habituellement les moyennes des traitements). Un tel patron de résidus peut être illustré schématiquement comme suit:



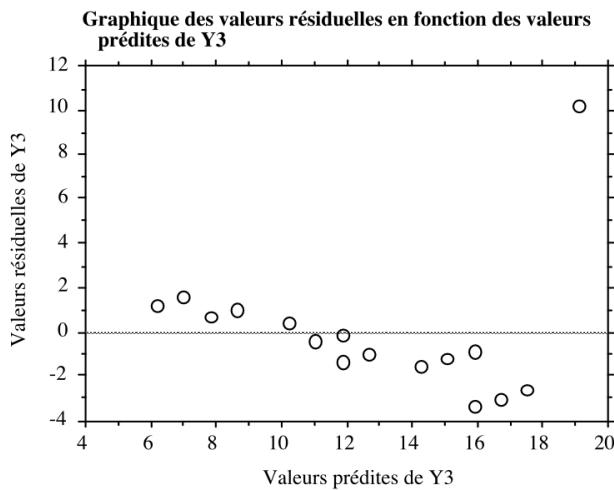
Le 2^e exemple représente une situation typique d'hétérogénéité des variances. On remarque que la variance des résidus est faible du côté gauche (faibles valeurs prédictes), et qu'elle s'accroît à mesure qu'on se déplace vers la droite (valeurs prédictes élevées):



Le patron des résidus ressemble alors généralement à un entonnoir s'ouvrant vers la droite. On peut aussi observer un entonnoir s'ouvrant vers la gauche dans certaines circonstances (ex. surtransformation). Les 2 figures ci-dessous schématisent les patrons de distribution caractéristiques de l'hétérogénéité des variances:

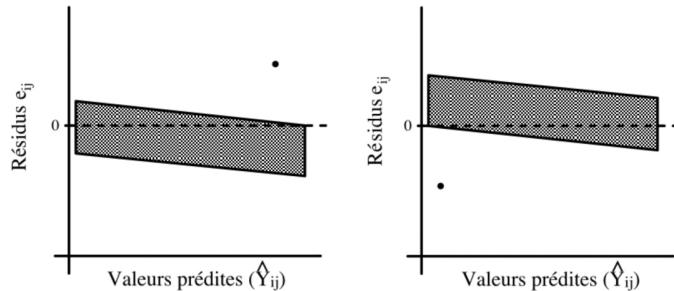


Le troisième exemple illustre l'effet d'une donnée extrême sur le patron des résidus. On peut observer la donnée extrême dans le coin supérieur droit de la figure. Puisque la somme de tous les résidus égale zéro, la présence d'une donnée extrême a pour effet de repousser l'ensemble des données dans la direction opposée à la donnée extrême :

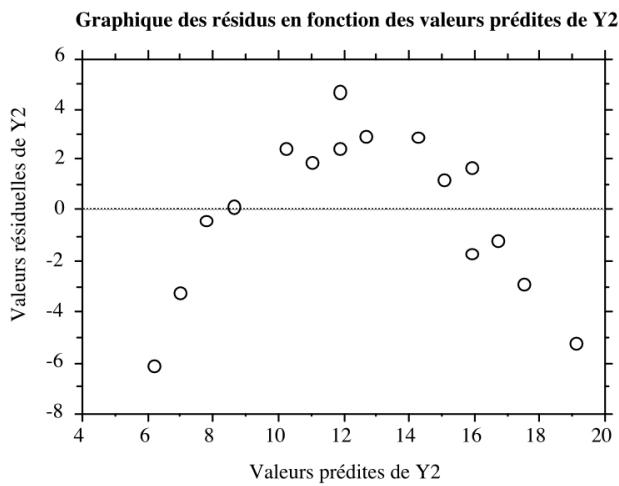


Une donnée extrême n'est pas forcément « aberrante ». Il pourrait s'agir d'une donnée rare, mais tout à fait plausible. Pour retirer une telle donnée du fichier de données, il faudrait être certain que la donnée est aberrante (ce qui peut être facile dans le cas d'erreur causée par la dyslexie ; ex. une valeur de 92 au lieu de 29, pour une variable qui montre habituellement des écarts entre 0 et 30). Le retrait d'une valeur extrême plausible (non aberrante) pourrait biaiser l'analyse des résultats.

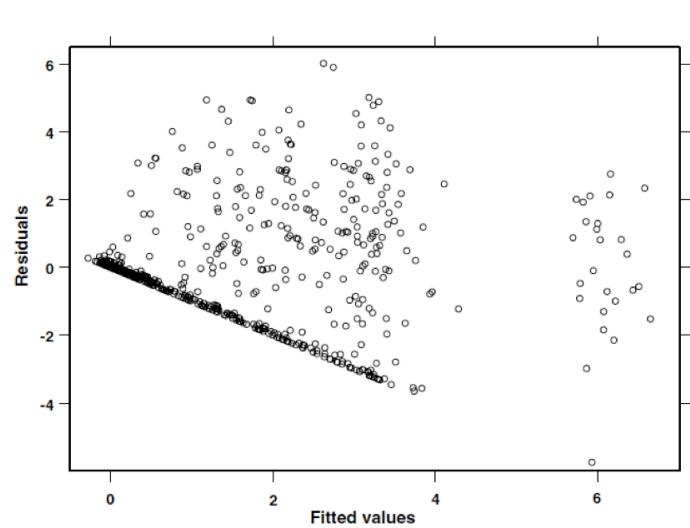
Ce type de patron de résidus peut être schématisé par l'une des 2 figures suivantes:



Le 4^e exemple illustre le patron des résidus pour une expérience où un effet quadratique important aurait été omis du modèle. Les résidus décrivent donc une courbe caractéristique d'un effet quadratique (ce type de phénomène est surtout rencontré en régression et donc beaucoup moins fréquent en analyse de la variance):



Enfin, le 5^e exemple ci-dessous illustre le patron des résidus pour une expérience où l'on retrouve un trop grand nombre de valeurs « zéro ». La seule solution à ce problème consiste à retirer de l'analyse statistique tous les traitements qui ne donnent que des zéros. La variance de ces traitements est alors pratiquement nulle et devient incompatible dans un contexte « d'analyse de la variance ». Ces traitements se distinguent ainsi nettement des autres traitements. Une note en bas de tableau (ou de figure) peut être utilisée pour présenter l'essentiel de l'information concernant ces traitements.



Bien qu'elle soit quelque peu subjective, l'analyse des résidus permet de vérifier rapidement la qualité de l'ajustement du modèle mathématique utilisé aux données expérimentales. L'utilisation des nouveaux logiciels de la micro-informatique facilite grandement l'utilisation de ce genre d'outil diagnostic; il ne fait nul doute que l'analyse graphique gagnera en popularité.

Les tests d'homogénéité de la variance

Les tests d'homogénéité de la variance

- test de **Levene** (le + recommandé)
 - approx., mais robuste et puissant
- test de **Bartlett** (le grand classique)
 - sensible à la non normalité (aplatissement) \Leftrightarrow erreur type I
 - invalidé par une variance = 0
- test de **Burr-Foster**
 - table de référence peu diffusée (Anderson et McLean)
- test de **Bartlett-Kendall (test de Box)**
 - nécessite beaucoup de réps^{ns}
- test de **Hartley**
 - si plupart des d.l. < 5
 - $S^2 \text{ max.} / S^2 \text{ min.}$

- utiles comme appui à l'analyse graphique
- ANOVA à une voie seulement (situations simples)

Test de Levene

$$\frac{\sum (X_{ij} - \bar{X}_{i.})^2}{n-1} \Rightarrow \frac{\sum |X_{ij} - \bar{X}_{i.}|}{n-1}$$

ANOVA des $| \text{dév}^{\text{ns}} |$ p/r aux moy^s des trait^s

Somme des
 $(\text{dév}^{\text{ns}})^2$

Somme des
 $| \text{dév}^{\text{ns}} |$

- \Leftrightarrow le fait d'éviter d'élever au carré rend le test beaucoup moins sensible aux distr^{ns} à longues queues
- \Leftrightarrow test approx., car les $| \text{dév}^{\text{ns}} |$ ne sont pas distr. ~ normale

Test de Levene ...

TABLE 13.11.1
EXAMPLE OF LEVENE'S TEST OF HOMOGENEITY OF VARIANCE

Traitements	Données originales Data for Class				Absolute Deviations from Class Mean			
	1	2	3	4	1	2	3	4
7 rép ^s	7.40	8.84	8.09	7.55	0.54	2.08	1.89	0.71
	6.18	6.69	7.96	5.65	0.68	0.07	1.76	1.19
	6.86	7.12	5.31	6.92	0.00	0.36	0.89	0.08
	7.76	7.42	7.39	6.50	0.90	0.66	1.19	0.34
	6.39	6.83	0.51	5.46	0.47	0.07	5.69	1.38
	5.95	5.06	7.84	7.40	0.91	1.70	1.64	0.56
	7.48	5.35	6.28	8.37	0.62	1.40	0.08	1.53
Total	48.02	47.31	43.38	47.85	4.12	6.34	13.14	5.79
Mean	6.86	6.76	6.20	6.84	0.589	0.906	1.877	0.827
s_i^2	0.500	1.630	7.325	1.100	0.095	0.668	3.214	0.302

Test de Levene ...

TABLE 13.11.2
ANALYSIS OF VARIANCE OF MEAN DEVIATIONS

Source	df	Sum of Squares	Mean Squares	F
Between classes	3	6.773	2.258	2.11
Within classes	24	25.674	1.070	

$$\text{Bartlett} \Leftrightarrow \chi^2 = 11.22 (P = 0.01)$$

$$\text{Levene} \Leftrightarrow F = 2.11 (P = 0.13)$$

```
data test;
p=1-probf(2.11, 3, 24);
run;
proc print data=test;
run;
```

Test de Bartlett

S'il y a a estimations s_i^2 , chacune avec le même nb. f de d.l., le test est:

$$M = 2.3026 f \left(a \log \bar{s}^2 - \sum \log s_i^2 \right) \quad \left(\bar{s}^2 = \frac{\sum s_i^2}{a} \right)$$

L'hypothèse nulle veut que chaque s_i^2 soit une estimation de la même variance σ^2 , la quantité M/C est alors distribuée approx. selon une χ^2 avec $(a-1)$ d.l. où:

$$C = 1 + \frac{a+1}{3af} \quad \begin{array}{l} a = \text{nb. de var}^s (\text{pop}^{\text{ns}}) \\ f = \text{nb. de d.l. (égaux)} \end{array}$$

$$\chi^2 = \frac{M}{C}$$

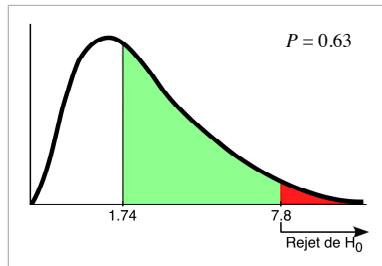
Test de Bartlett ...

Quantités de matières grasses absorbées par des beignets selon 4 catégories de matières grasses:

TABLEAU 10.21.1
CALCUL DU TEST D'HOMOGENÉITÉ DES VARIANCES DE BARTLETT
TOUTES LES ESTIMATIONS ONT $f = 5$ DEGRÉS DE LIBERTÉ

Matières grasses	s_i^2	$\log s_i^2$
1	$S_1^2 = 178$	2,2504
2	60	1,7781
3	98	1,9912
4	$S_4^2 = 68$	1,8325
Total	$\bar{s}^2 = \frac{404}{100,9}$ log $\rightarrow \log \bar{s}^2 = 2,0038$	7,8522
	$M = (2,3026) (5) [4(2,0038) — 7,8522] = 1,88, \quad (d.l. = 3)$	
	$C = 1 + \frac{a+1}{3af} = 1 + \frac{5}{(3)(4)(5)} = 1,083$	
$\chi^2 = M/C$	$\chi^2 = 1,88/1,083 = 1,74 \quad (d.l. = 3), \quad P > 0,5$	

$$\bar{S}^2 = \frac{\sum S_i^2}{a}$$



Test de Bartlett et Kendall - Test de Box

TABLE 1.3.3 Log s^2 Analysis of Homogeneity									
Type of tablet									
A		B		C		D			
s^2	$\log s^2$	s^2	$\log s^2$	s^2	$\log s^2$	s^2	$\log s^2$		
59	1.7709	10	1.0000	449.5	2.6527	493.0	2.6928		
52.5	1.7202	39	1.5911	510.0	2.7076	478.5	2.6799		

ANOVA of $\log s^2$										
Source	df	MS	F_{calc}	$F_{\text{crit}}(3,4)$	(0.01)					
Types of tablet	3	0.96876	21.8	16.7	P = 0.006					
Within error	4	0.04439								
Total	7									

Adapté de Snedecor and Cochran, 1989.

Test de Tukey pour l'additivité

Il existe un test de Tukey pour vérifier l'additivité des effets, mais il est peu utilisé en pratique parce que peu de logiciels permettent de calculer le test.

Causes de l'hétérogénéité de la variance et mesures correctrices

Quelques cas fréquents de non homogénéité de la variance sont:

- quand on tente de mettre ensemble plusieurs expériences (ex: essais de cultivars à plusieurs sites).
- quand il y a utilisation d'un témoin sans traitement. Dans ce cas l'enlèvement du témoin améliore souvent l'homogénéité de la variance.
- quand il y a présence d'une donnée aberrante.
- utilisation d'échantillons très variables.

Plusieurs moyens d'action s'offrent à nous lorsqu'on est confronté à des variances non homogènes:

- 1- on peut transformer les données de façon à rendre homogènes les variances des différents traitements;
- 2- on peut séparer les données en groupes de telle sorte que les variances à l'intérieur de chaque groupe soient homogènes. Ainsi, chaque groupe peut être analysé séparément;
- 3- élimination du traitement responsable de l'hétérogénéité ou des données aberrantes;
- 4- on peut effectuer une pondération des moyennes en fonction de leurs variances, méthode communément utilisée en régression;
- 5- utilisation de tests non paramétriques, ou d'équivalents utilisant la transformation de rang;
- 6- recours aux modèles mixtes (ex.: procédure Mixed de SAS; plus complexe).

Quand l'une des approches ci-haut démontre l'hétérogénéité de la variance, et qu'aucune transformation ne donne de résultats satisfaisants, alors l'approche la plus simple consiste à identifier et éliminer le traitement responsable de l'hétérogénéité. Dans certains cas, la présence de données aberrantes pourra aussi affecter l'homogénéité des variances. L'examen des variances des traitements révèlera généralement l'identité du ou des traitements responsables de l'hétérogénéité. De même, si on tente de mettre en commun plusieurs expériences, une analyse de variance pour chacune des expériences indiquera souvent la source d'hétérogénéité. La plupart du temps, l'enlèvement du traitement responsable de l'hétérogénéité suffira à rendre la variance de l'erreur expérimentale plus homogène.

Il existe aussi un test vérifier si les données sont additives. Le test de Tukey pour la non-additivité fournit un test à un degré de liberté. Toutefois, son utilisation n'est pas très répandue.

5.5 - LA TRANSFORMATION DES DONNÉES

Une façon de pallier au non respect des postulats de base de l'analyse de variance consiste à effectuer des transformations mathématiques sur les données, de façon à en améliorer les propriétés. La transformation des données équivaut à changer l'échelle de mesure d'un phénomène, dans le but de modifier des caractéristiques des données susceptibles d'invalider plus ou moins l'analyse de variance.

La non-additivité des effets, la non-normalité de l'erreur expérimentale, et la non-homogénéité de la variance de l'erreur sont les trois conditions pour lesquelles l'utilisation des transformations peut s'avérer bénéfique. Aucune transformation, par contre, ne peut compenser pour la non-indépendance des erreurs.

• La transformation logarithmique

- permet d'améliorer l'homogénéité de la variance et l'additivité.
- indiquée lorsque
- les effets des traitements (sources de variation) sont multiplicatifs au lieu d'être additifs (une multiplication est ainsi remplacée par une addition: $A \times B$ est remplacé par $\log A + \log B$).
- la variabilité augmente avec la moyenne, de telle sorte que l'écart type augmente proportionnellement à la moyenne.
- il est généralement préférable d'utiliser **log (x + 1)**. Si la variable x prend des valeurs comprises entre zéro (inclusivement) et un, il y aura problème puisque le logarithme de zéro n'existe pas et que le logarithme d'une proportion située entre zéro et un est négatif.

La transformation logarithmique ...

Effets additifs vs multiplicatifs:

		Blocs		
		1	2	
		1	10	20
		2	30	40
			+ 10	

Effets multiplicatifs

		Blocs		
		1	2	
		1	10	20
		2	30	60
			x 2	

Effets additifs

		Blocs		
		1	2	
		1	1.00	1.30
		2	1.48	1.78
			+ 0.30	

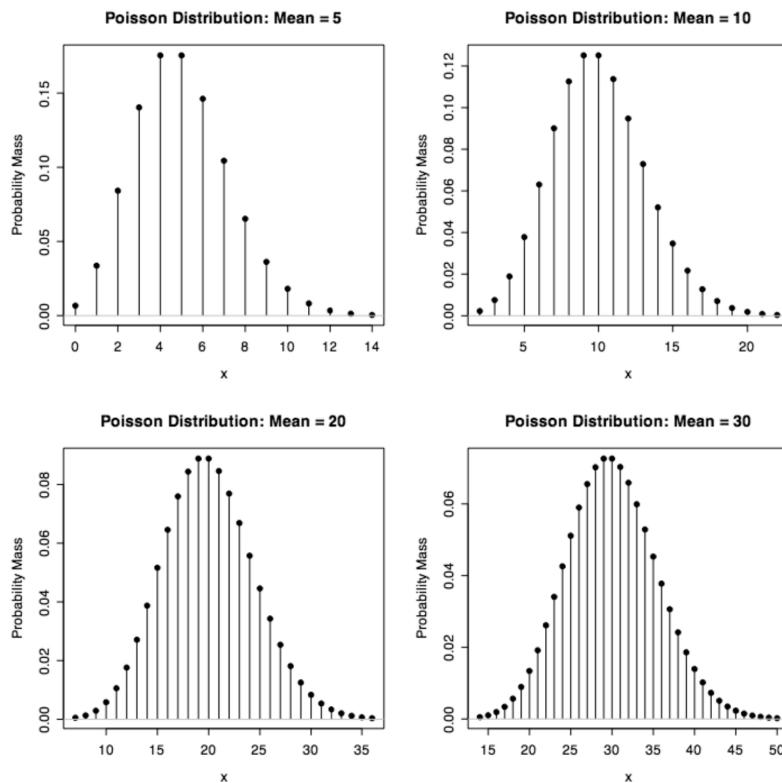
(Voir cahier de laboratoire, section 2.3- Exemple de transformation logarithmique; fichier exbclh12.sas. Cet exemple est présenté à titre complémentaire et n'est pas vu systématiquement au laboratoire).

• La transformation racine carrée

- indiquée lorsque les données suivent une distribution de Poisson. La loi de Poisson décrit la distribution d'événements survenant au hasard dans l'espace ou le temps, lorsque ces événements sont mesurés dans des unités égales d'espace ou de temps, c'est-à-dire lorsque les conditions de mesure sont homogènes. Il s'agit donc de données énumératives (nombres entiers), ou comptes, et non pas de données de mesures.
- Exemple:
 - nombre de plantes par unité de surface
 - nombre de bactéries par plat de Pétri
- quand les données sont distribuées selon la loi de Poisson, on peut montrer que les résultats de comptes répétés sur des unités expérimentales sont distribués autour de la moyenne avec une variance égale à la moyenne elle-même (l'écart type égale la racine carrée de la

moyenne). Pour effectuer une analyse de variance sur de telles données, il faut donc les transformer de façon à rendre la variance indépendante de la moyenne. C'est ce que la transformation racine carrée permet de faire, bien que de façon approximative.

- la transformation $\sqrt{x + 0.5}$ donne généralement de meilleurs résultats, spécialement si les données sont constituées de très petits nombres (inférieurs à 10) et que les zéros sont fréquents (événements rares). La transformation logarithmique peut aussi être utile dans certains cas.



Famille de courbes de distribution de fréquence de la fonction Poisson pour différentes valeur de lambda.

Une approche moderne dans le traitement des comptes consiste à utiliser la régression logistique ou les modèles loglinéaires.

(Voir cahier de laboratoire, section 2.2- Exemple de transformation racine carrée; fichier exbcli7b.sas).

• La transformation angulaire

- utilisée lorsque les données sont des proportions entre individus, et seulement si les pourcentages sont établis sur une base constante (ie. nombre égal d'individus dans chaque groupe: la dimension de l'échantillon est la même pour tous les traitements).

$\arcsin \sqrt{p}$, où $-1 \leq p \leq 1$

ou

$\arcsin \sqrt{\frac{p}{100}}$, si p exprimé en pourcentage

- les données exprimées sous forme de pourcentages ne doivent pas être transformées automatiquement à l'aide de la transformation angulaire. On peut rencontrer 3 situations courantes:
 - 1- les valeurs sont comprises entre 30% et 70%: dans ce cas, il n'est généralement pas nécessaire de transformer les données originales;
 - 2- les valeurs sont comprises entre 0 et 30%, ou entre 70 et 100%: à cause du comportement des données de ce type, la transformation racine carrée ou le logarithme vont habituellement donner de bons résultats. Le comportement des données s'apparente à celui des données de dénombrement ou des données de croissance corrigé par ces transformations.
 - 3- si les situations 1 et 2 ne s'appliquent pas, alors la transformation arcsinus de la racine carrée des données donne généralement de bons résultats. Dans ce cas on peut constater que la variance d'échantillonnage de la proportion, pour un échantillon donné, varie suivant la proportion elle-même. Ainsi la variance est maximale quand la proportion est de 50%, et elle tend vers zéro à mesure que la proportion s'approche de 0% ou de 100%. Le poids des observations diffère donc d'une observation à l'autre: il est maximal quand la proportion se situe autour de 0% ou de 100%, et il est minimal autour de 50%.

Il existe différentes versions de la transformation angulaire améliorée pour les proportions binomiales, où les valeurs zéro et 100 peuvent être remplacées par diverses valeurs avant l'application de la transformation. Par exemple, si $n < 50$, où n est la taille de l'échantillon sur lequel sont observées les proportions (exprimées en pourcentage), une proportion de zéro devrait être remplacée par $1/(4n)$ et une proportion de 100% par $(n - 1/4)/n$ avant de transformer en angles (Snedecor et Cochran, 1989). Cette modification améliore l'égalité des variances dans l'échelle transformée. D'autres variantes existent.

(Voir cahier de laboratoire, section 2.1- Exemple de transformation angulaire ; fichier excagg72.sas).

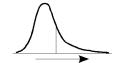
• La transformation réciproque:

- utilisée principalement dans le cas où la variable est le temps de survie d'organismes vivants (ex: toxicologie). La variance est alors proportionnelle au temps de survie.
- $1/x$, ou $100,000/x$, si on désire se débarrasser des nombres décimaux.
- la transformation réciproque permet d'améliorer l'homogénéité de la variance. Une caractéristique intéressante de cette transformation est qu'elle permet d'inclure dans l'analyse les individus n'ayant pas succombé à l'effet du traitement, et dont le temps de survie est par conséquent indéfini (L'inverse de l'infini devient zéro, ce qui rend le calcul de la moyenne et de la variance possible).

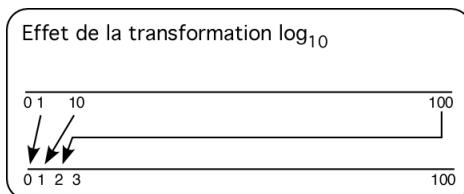
• **Les transformations par l'échelle des puissances.**

On peut montrer qu'il existe une relation entre la plupart des transformations décrites précédemment. Ces transformations peuvent être exprimées sous la forme d'une valeur «y» élevée à une puissance «p», où p peut prendre les valeurs énumérées dans le tableau ci-dessous. On peut, à partir de l'échelle des puissances, choisir n'importe laquelle transformation d'intérêt.

Tableau 5.8) Transformations issues de l'échelle des puissances

Puissance (p)	Transformation effectuée	Terminologie	Remarques
>3	y^n		Dissymétrie négative 
3	y^3	Cube	
2	y^2	Carré	
1	y^1	Données brutes	Aucune transformation
1/2	\sqrt{y}	Racine carrée	Indiquée pour les comptes
(0)	$\log(y)$	Logarithme	$\log(y)$ prend la place de zéro dans l'échelle des puissances
-1/2	$1/\sqrt{y}$	Racine réciproque	
-1	$1/y$	Réciproque	Indiquée pour les temps de survie
-2	$1/y^2$	Réciproque du carré	
<-2			Dissymétrie positive 

Une propriété intéressante de l'échelle des puissances est que chacune des transformations ainsi obtenues préserve l'ordre des données (les transformations obtenues à partir de puissances négatives sont précédées d'un signe négatif, afin de préserver l'ordre). Un problème relié à l'utilisation des transformations issues de l'échelle des puissances est la difficulté de trouver une signification théorique justifiant leur utilisation.



Effet de la transformation logarithmique sur l'échelle de mesure

• **La transformation de rang**

La transformation de rang consiste simplement à remplacer les valeurs originales d'une variable par leur rang dans la séquence ordonnée des observations selon les valeurs prises par cette même variable. La transformation permet ainsi de réaliser un test qui est pratiquement l'équivalent d'un test non paramétrique. Une telle transformation permet alors de contourner un problème de variances hétérogènes ou de distribution anormale des observations lorsqu'aucune

des transformations mentionnées ci-haut ne donne de résultats acceptables. Par exemple, le fait de faire une ANOVA sur les rangs des observations pour une variable bien précise est à peu près l'équivalent de faire le test non paramétrique correspondant (en l'occurrence, le *Kruskal-Wallis* pour un plan entièrement aléatoire, ou le *test de Friedman* pour le plan en blocs complets).

• La transformation de rang et les tests non paramétriques

Un test non paramétrique est un test où le postulat de normalité de l'ANOVA est remplacé par un postulat de distribution symétrique. Il n'est donc plus nécessaire que la variable dépendante soit distribuée selon une courbe normale pour que les résultats du test soient valides; il suffit simplement que la distribution soit symétrique et la même pour tous les traitements. Il est important de réaliser que le postulat d'indépendance des erreurs de l'ANOVA doit être respecté: l'échantillonnage doit être aléatoire pour chaque traitement.

La plupart des tests non paramétriques sont construits sur les statistiques de rang, ce qui permet aussi de contourner le problème d'hétérogénéité des variances. En négligeant les espacements entre les observations et en ne retenant que l'information ayant trait au classement relatif de celles-ci, la transformation de rang permet d'homogénéiser les variances des traitements. Les tests non paramétriques, ou leur équivalent, sont donc tout indiqués pour les expériences où les remèdes habituels contre l'hétérogénéité des variances (transformation mathématique, élimination d'un traitement dont la variance est différente, regroupement de traitements ayant des variances du même ordre de grandeur) n'offrent aucune solution acceptable. En cas de doute, l'utilisation d'un test non paramétrique peut même permettre de confirmer les résultats d'une analyse paramétrique normale dont l'application semble douteuse (ex. variable dépendante montrant une distribution douteuse, variances légèrement hétérogènes, etc.).

Plusieurs tests non paramétriques consistent à ordonner les observations selon les valeurs de la variable dépendante et à analyser les rangs au lieu des valeurs originales. Le tableau suivant présente quelques tests paramétriques exécutés sur les rangs et leur équivalent non paramétrique:

Test avec transformation de rang	Test non paramétrique
Test de t avec échantillons indépendants appliqué aux rangs † (PROC RANK + test de t)	Test de Mann-Whitney, ou test de la somme des rangs de Wilcoxon (<i>«Wilcoxon rank sum test»</i>)
Test de t avec données appariées (<i>«paired t test»</i>)	Test des signes de Wilcoxon (<i>«Wilcoxon signed ranks»</i>)
ANOVA à 1 voie appliquée aux rangs ‡ (PROC RANK + ANOVA à 1 voie)	Test de Kruskal-Wallis (<i>«Kruskal-Wallis k-sample test»</i>)
ANOVA à 2 voies pour un plan en blocs complets appliquée aux rangs intra-blocs (PROC RANK avec énoncé BY BLOC + ANOVA à 2 voies)	Test de Friedman
Régression appliquée aux rangs §	Régression non paramétrique

† Si le test de t est appliqué aux scores normaux (valeurs z) plutôt qu'aux rangs, le test devient équivalent au test de van der Waerden; si le test est appliqué aux scores médians (2 groupes), le test devient le test de la médiane (*«median test»*).

‡ Le test de F utilisé par la procédure paramétrique appliquée aux rangs est souvent meilleur que l'approximation par le khi-carré utilisée par le Kruskal-Wallis.

§ La méthode des moindres carrés est non paramétrique; seuls les test d'hypothèses et la construction d'intervalles de confiance requièrent la normalité. Le coefficient de corrélation de Spearman est l'équivalent non paramétrique du coefficient de corrélation de Pearson.

Les tests non paramétriques sont moins bien connus et présentent une approche beaucoup moins unifiée que leur contrepartie paramétrique. Pour ces raisons, l'utilisation de la transformation de rang avec les tests paramétriques présente un intérêt certain pour la réalisation de tests équivalents aux tests non paramétriques.

La transformation de rang peut être utile dans l'analyse de plans d'expérience et en régression multiple. Pour analyser un plan d'expérience avec la transformation de rang, il suffit d'ordonner les observations selon les valeurs prises par la variable dépendante, de la valeur la plus petite à la valeur la plus grande (plusieurs logiciels permettent de réaliser cette opération à l'aide d'une transformation qui remplace les valeurs originales par leur rang), et d'appliquer l'analyse de variance aux rangs. Ce test résulte en une procédure conditionnellement indépendante de la distribution de la variable. Un tel test est cependant robuste, ce qui signifie que la valeur P obtenue est une très bonne approximation de la valeur P exacte du test non paramétrique correspondant, peu importe la distribution de la variable dépendante.

La transformation de rang est spécialement utile pour les plans d'expérience pour lesquels il n'existe aucun test non paramétrique. La transformation de rang peut aussi être utile pour faciliter le choix d'une transformation mathématique appropriée. Dans ce cas, on recommande d'utiliser à la fois la procédure paramétrique habituelle sur les données brutes et la même procédure avec la transformation de rang. Si les deux procédures donnent des résultats presque identiques, on assume alors que les postulats de l'ANOVA sont raisonnablement respectés et que la procédure paramétrique régulière sur les données brutes est valide. Si les deux procédures donnent des résultats différents, on peut soit retenir la transformation mathématique donnant les résultats les plus similaires à ceux de la transformation de rang, soit considérer directement l'analyse sur les rangs, puisqu'elle est dans ce cas probablement plus précise que l'analyse sur les données originales. Dans un tel cas, on suggère aussi d'inspecter les données et les graphiques de résidus de façon à détecter d'éventuelles données extrêmes ou de vérifier la présence de distributions très asymétriques; ces deux conditions vont affecter les analyses paramétriques mais auront peu d'effets sur les analyses non paramétriques. Le *coefficient de détermination* (r^2), qui mesure la qualité d'ajustement d'un modèle à un ensemble de données (voir chapitre sur la régression), peut aussi être utile comme critère de choix dans la comparaison de modèles avec différentes transformations mathématiques.

Utilisation de la transformation de rang avec le logiciel SAS:

La procédure PROC RANK permet de créer un nouveau fichier de données SAS contenant les rangs au lieu des valeurs originales. Le nouveau fichier contient toutes les variables du fichier d'origine avec en plus les variables nommées dans l'énoncé RANKS. La procédure a la forme suivante:

```
PROC RANK DATA= ..... OUT= ..... <autres options disponibles>;
   BY liste-des-groupes;
   RANKS listes-des-nouvelles-variables;
   VAR liste-des-variables;
```

Par défaut, la méthode de classification des valeurs égales est TIES=MEAN, c'est-à-dire que les valeurs égales sont remplacées par la moyenne de leurs rangs dans la séquence ordonnée des valeurs (chaque valeur supplémentaire se voyant attribuer un rang différent). L'énoncé RANKS est utilisé si on veut que les valeurs originales de la ou des variable(s) originale(s) soi(en)t incluse(s) en plus des rangs. Si l'énoncé RANKS est omis de la procédure, les valeurs originales sont simplement remplacées par les rangs qu'elles occupent dans la séquence ordonnée.

Des travaux plus récents ont cependant démontré que la transformation de rang ne devrait pas être utilisée avec des expériences factorielles. Shah et Madden (2004) proposent une alternative pour l'analyse de données ordinaires d'expériences factorielles. Une autre approche consiste à utiliser les *rangs alignés* (« aligned rank transform ») (Richter and Payton, 1999 ; Mansouri, 1999 ; Wobbrock *et al.*, 2011).

Références sur la transformation de rang:

- Conover, W.J. 1980. Practical non parametric statistics. 2^e ed. John Wiley & Sons, New York. 493 pp.
- Conover, W.J. and Iman, R.L. 1976. On some alternative procedures using ranks for the analysis of experimental designs. Communications in statistics, A5, 14, 1348-1368
- Conover, W.J. and Iman, R.L. 1981. Rank transformations as a bridge between parametric and nonparametric statistics. The American Statistician, 35, 124-129.
- Iman, R.L. and Conover, W.J. 1979. The use of the rank transformation in regression. Technometrics, 21, 499-509.
- Mansouri, H. 1999. Multifactor analysis of variance based on the aligned rank transform technique. Comput. Statist. Data Anal. 29 : 177-189.
- Richter, S.J. and M.E. Payton 1999. Nearly exact tests in factorial experiments using the aligned rank transform. J. Appl Stat. 26 (2) : 203-217.
- Shah, D.A., and L.V. Madden 2004. Nonparametric analysis of ordinal data in designed factorial experiments. Phytopathology 94 : 33-43.
- SAS Institute Inc. 1990. SAS® Procedures Guide, Version 6, Third Edition, Cary, NC: SAS Institute Inc. 705 pp. (Chapter 29 - The rank procedure).
- Wobbrock *et al.*. 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. CHI 2011 - Session: Research Methods. May 7-12, Vancouver, BC, Canada.

• Quand transformer?

- le test de F de l'ANOVA est très «robuste» pour α dans le cas du modèle fixe avec des échantillons d'égale dimension (même n^{bre} . de reps);
 - => le niveau α du test de F est alors peu affecté par l'hétérogénéité des variances.
- > Box a démontré que dans une situation où on observe une différence de l'ordre de 9 X entre la plus petite variance et la plus grande variance de traitement (neuf fois la plus petite variance), la valeur α associée au test de F de l'ANOVA est relativement peu affectée.

La philosophie de la transformation des données diffère de celle de l'ANOVA. L'ANOVA vise à trouver le maximum de vraies différences (significatives) entre les traitements à l'étude. C'est la raison pour laquelle on choisit un niveau de signification plutôt faible (ex. : $\alpha\square= 0.05$). Dans le cas des tests d'homogénéité des variances (si l'usage d'un tel test est approprié), on désire transformer le moins possible les données, et uniquement quand la mesure est nécessaire pour valider l'analyse et l'interprétation des données. Même si la transformation des données permet de réaliser correctement des analyses autrement non valides, elle complique quelque peu l'analyse et l'interprétation des données, car la relation entre l'échelle transformée et l'échelle de départ (données brutes) n'est pas toujours évidente avec certaines transformations (ex. : arcsinus racine carrée).

Quand transformer?

ANOVA	Tests d'homogénéité de la variance
<ul style="list-style-type: none"> - trouver le max. de vraies différences sign. - $\alpha = 0.05$ 	<ul style="list-style-type: none"> - transformer le moins souvent possible - $\alpha = 0.05$ (Levene) - $\alpha = 0.01 - 0.001$ (Bartlett ou Burr-Foster)

• Règles pratiques de transformation des données suite à un test d'homogénéité des variances (Anderson et McLean, 1974):

L'observation d'une relation visuelle (par ex., entre la moyenne et la variance) indique qu'un test de l'homogénéité de la variance devrait être considéré. Anderson et McLean rapportent que la règle pratique suivante fonctionne relativement bien lorsqu'on doit prendre une décision sur la pertinence de transformer les données à partir des résultats d'un test d'homogénéité des variances:

- 1) Si l'homogénéité de la variances est acceptée au niveau $\alpha = 0.01$, ne pas transformer;
- 2) Si l'homogénéité de la variances est rejetée au niveau $\alpha = 0.001$, transformer;
- 3) Si le résultat du test de l'homogénéité des variances se situe entre $\alpha = 0.01$ et 0.001 , essayer de trouver la distribution théorique des données:
 - s'il y a une raison pratique de transformer, => transformer;
 - s'il n'y a aucune bonne raison de transformer, => ne pas transformer.

Suggestion: si l'expérimentateur n'a aucune connaissance théorique de la variable étudiée, utiliser le niveau $\alpha = 0.001$ comme seuil de décision quand la distribution de la variable semble avoir de très longues queues (de façon à prévenir les transformations inutiles).

• Interprétation des données transformées

Comme mentionné plus haut, la transformation des données complique quelque peu l’interprétation des données. Trois approches sont possibles :

- 1- Interprétation des **moyennes transformées**. La présentation des moyennes transformées est facile, mais l’effet des traitements sur les moyennes transformées est souvent plus difficile à visualiser qu’avec les données brutes;
- 2- Interprétation des **moyennes détransformées**. La détransformation consiste simplement à appliquer la transformation inverse sur les moyennes des traitements (calculées dans l’échelle transformée). La présentation des moyennes détransformées est un peu plus laborieuse, mais plus facile à visualiser, car la détransformation ramène les moyennes dans l’échelle originale (avec les unités de mesure originales). Il faut noter qu'il n'est pas possible de détransformer les erreurs types ou les écart types calculés dans une échelle transformée. On suggère alors de présenter les résultats sous forme d’intervalles de confiance, et de détransformer les limites inférieure et supérieure des intervalles de confiance. Le principal inconvénient de cette approche est que les limites de confiance deviennent ainsi asymétriques, mais le résultat est valide;
- 3- Interprétation des moyennes brutes. L’interprétation des moyennes brutes n’est pas recommandée, car la moyenne devient un moins bon indice de tendance centrale dans une situation avec variances hétérogènes (ou dissymétrie prononcée); des différences de classement sont alors possible lorsqu’on compare les traitements dans l’échelle transformée et dans l’échelle originale (données brutes).

• Détransformation des moyennes de traitements après une analyse statistiques des données transformées

On peut facilement détransformer les moyennes des traitements avec une calculatrice ou dans Excel. La colonne de droite du tableau ci-dessous présente les formules à entrer dans Excel. La cellule B2 (ou toute autre cellule d’un chiffrier) doit contenir la donnée à détransformer.

Transformation	Transformation inverse	Transformation Inverse dans Excel
$\log_{10}(Y)$	10 exposant Y	<i>Cliquer sur une cellule et entrer la formule suivante :</i> = 10^B2 où B2 contient la donnée à détransformer
$\log_{10}(Y+1)$	(10 exposant Y) - 1	= (10^B2) - 1
$\ln(Y)$	e exposant Y	= EXP^B2
racine carrée de Y	Y exposant 2	= B2^2
arcsin (racine carrée de Y)	((sin(Y)) exposant 2 (en radians)	= (SIN(B2))^2

• Commentaires sur les transformations:

Il est faux de prétendre que la transformation des données peut remédier à tous les problèmes, ou encore « améliorer » un ensemble de données. En réalité, il est toujours préférable d’effectuer

de bonnes observations et d'être soigneux lors de la réalisation de l'expérience et de la cueillette des données. Les transformations permettent seulement «d'améliorer» l'homogénéité et l'additivité de la variance, et elles peuvent s'avérer totalement inefficaces dans certains cas.

Une pratique courante consiste à «essayer» différentes transformations et à utiliser celle qui donne les meilleurs résultats. Bien que cette pratique ne soit pas fondamentalement mauvaise, le choix d'une transformation devrait d'abord être effectué sur une base théorique (à priori). La transformation utilisée devrait, dans la mesure du possible, avoir un sens. Transformer les données équivaut à modifier l'échelle de mesure d'une variable. Par exemple, la transformation logarithmique de la longueur en cm de la tige d'une plante équivaut à mesurer la tige au moyen d'une règle graduée selon une échelle logarithmique. D'autre part, on devrait associer à «l'essai» d'une transformation la perte d'un certain nombre de degrés de liberté (quelque part autour de 2/3 par transformation essayée, selon Tukey), mais les statisticiens ne s'entendent pas sur une procédure à suivre.

La décision de transformer les données doit être basée strictement sur la vérification des postulats de l'ANOVA. C'est donc l'observation du graphique d'homogénéité des variances et du graphique de vérification de la normalité qui aide à choisir la meilleure transformation. Le choix de la transformation ne doit nullement être influencé par les résultats obtenus.

Enfin, la transformation des données ne devrait être effectuée que lorsque jugée réellement nécessaire. La transformation des données complique souvent l'interprétation des analyses, surtout dans le cas où la nouvelle échelle de mesure ainsi créée n'a aucune signification théorique.

• Notes complémentaires:

Le test de **Brown-Forsythe** est une variante du test de Levene qui porte sur les déviations par rapport aux médianes des groupes, au lieu des déviations par rapport aux moyennes.

L'**ANOVA de Welch** (Welch's ANOVA, ou test de Welch) est une ANOVA à une voie pondérée robuste par rapport au postulat d'homogénéité des variances. Elle n'est cependant utilisable qu'avec une ANOVA à une voie, une situation correspondant au plan entièrement aléatoire.

Il est également possible de modéliser la structure hétérogène des variances à l'aide de la procédure **Mixed**. Par contre, il arrive que cette approche ne soit pas satisfaisante, quand la procédure ne converge pas vers une solution.