Chapitre 2 Expériences avec un facteur

Utilisation de Statistica

Méthode de l'analyse variance - tests d'hypothèses

Exemple procédé de gravure (électronique)
(« wet etching ») enlèvement du silicium sur des «puces »
variable de réponse Y : taux d'enlèvement du procédé
comparaison de solution1 et solution2

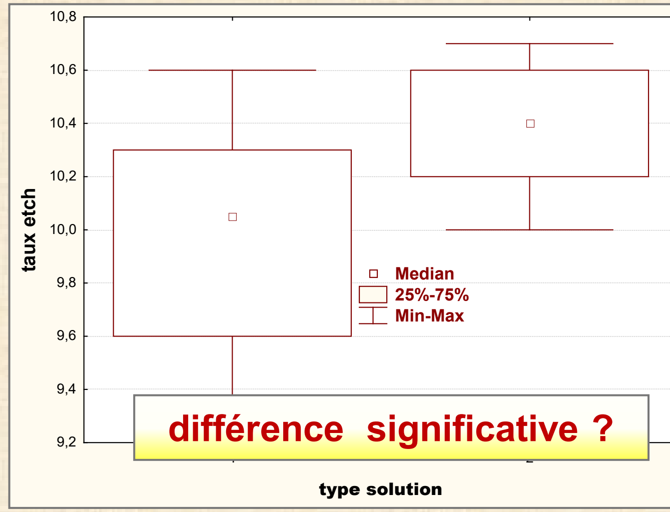
Données expérimentales Y n = 10 moyennes ($\overline{Y}$)

solution 1 : 9.9 10.6 9.4 10.3 9.3
10.0 9.6 10.3 10.2 10.1 — 9.97

solution 2 : 10.2 10.0 10.6 10.2 10.7
10.7 10.4 10.4 10.5 10.3 — 10.04

Un facteur à 2 modalités

X-solution	rep	Y-taux etch
1	1	9,9
1	2	9,4
1	3	9,3
1	4	9,6
1	5	10,2
1	6	10,6
1	7	10,3
1	8	10,0
1	9	10,3
1	10	10,1
2	1	10,2
2	2	10,6
2	3	10,7
2	4	10,4
2	5	10,5
2	6	10,0
2	7	10,2
2	8	10,7
2	9	10,4
2	10	10,3



différence significative ?

Exemple recherche nouvelle composition de fibres synthétique tissus

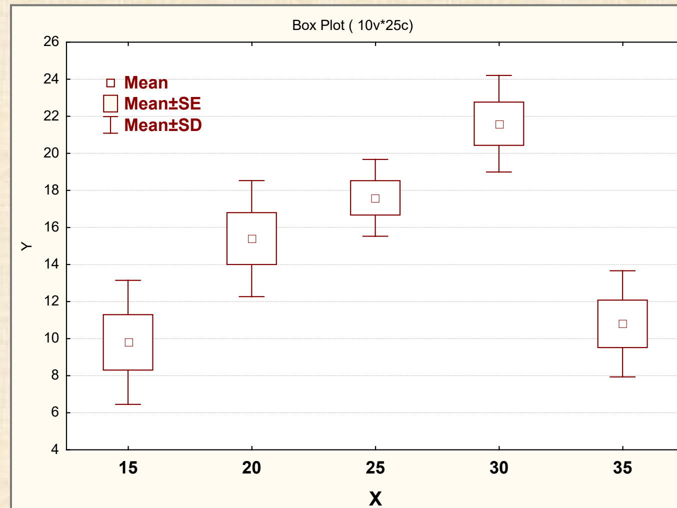
- facteur X : % coton varie entre 15 et 35
- réponse Y : force de tension tissu
- un facteur avec 5 modalités de X fixées :

15 20 25 30 35

Données expérimentales Y n = 5 répétitions

X	i	Y 1	2	3	4	5	moyennes ($\overline{Y}$)
15	1	7	7	15	11	9	9.8
20	2	12	17	12	18	18	15.4
25	3	14	18	18	19	19	17.6
30	4	19	25	22	19	23	21.6
35	5	7	10	11	15	11	10.8

X-coton	rep	Y-tension	x^2
15	1	7	225
15	2	7	225
15	3	15	225
15	4	11	225
15	5	9	225
20	1	12	400
20	2	17	400
20	3	12	400
20	4	18	400
20	5	18	400
25	1	14	625
25	2	18	625
25	3	18	625
25	4	19	625
25	5	19	625
30	1	19	900
30	2	25	900
30	3	22	900
30	4	19	900
30	5	23	900
35	1	7	1225
35	2	10	1225
35	3	11	1225
35	4	15	1225
35	5	11	1225



Influence du facteur X ?

différences significatives ?

Méthode des tests d'hypothèses

Ex 2.1 procédure à employer **Test t de Student**

- cadre pour des expériences de comparaison simple :
  1 facteur variant à 2 modalités
- utilisable dans tous les plans expérimentaux avec
  p (2 ou plus) facteurs variant à 2 modalités

Ex 2.2 procédure **ANOVA** ANALYSIS OF VARIANCE

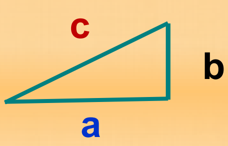
- avec 1 facteur variant avec k modalités (niveaux) (k>=2)
- p facteurs (p >=2) variant avec plusieurs modalités
- ne pas appliquer le test t à toutes les paires
- analyse de la variance : décomposition de la variabilité

…. théorème de Pythagore de la variabilité

a = variabilité expliquée (facteurs X)

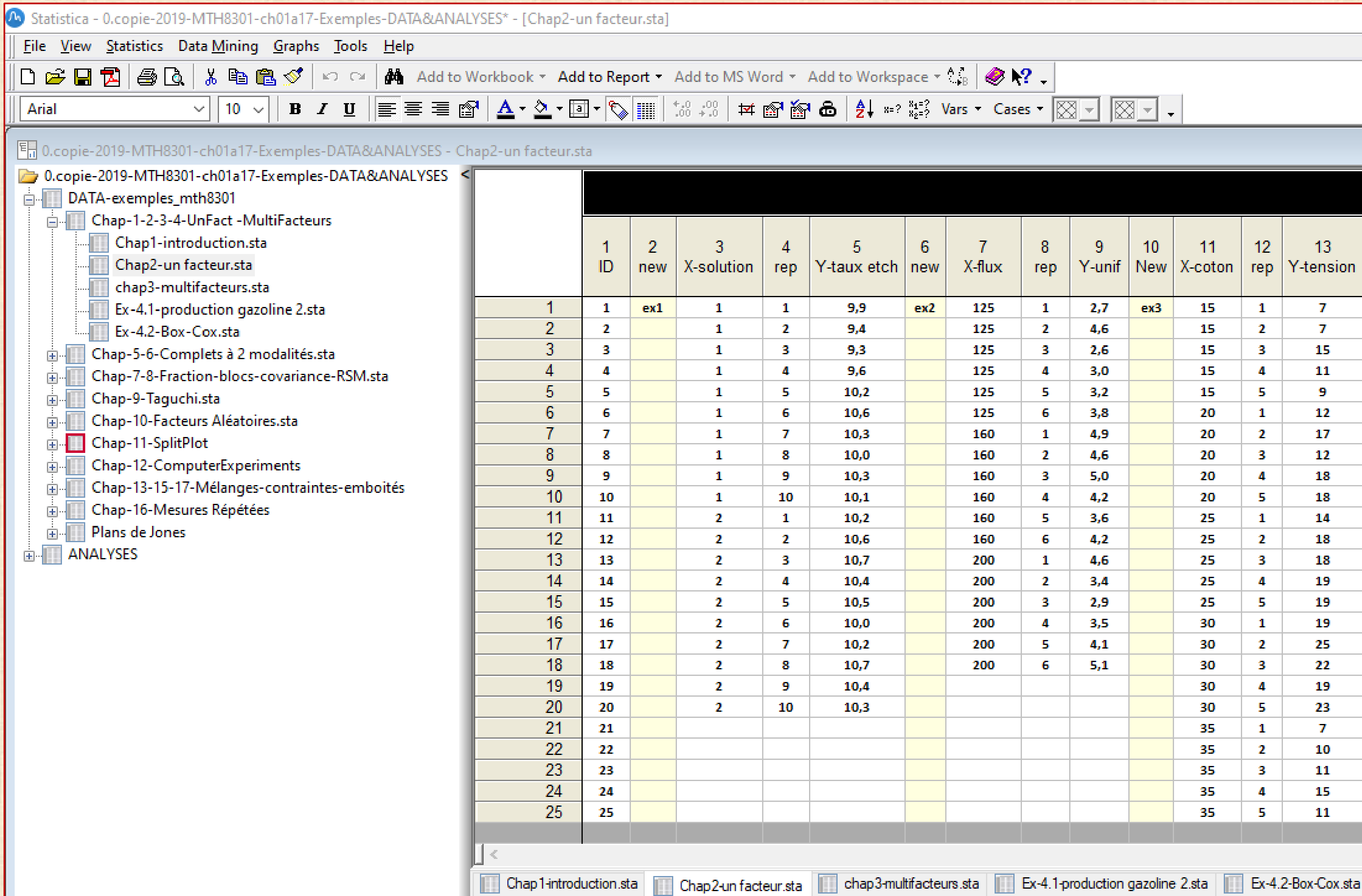
b = variabilité non expliquée (erreur expérimentale Ɛ)

c = variabilité totale (réponse Y)



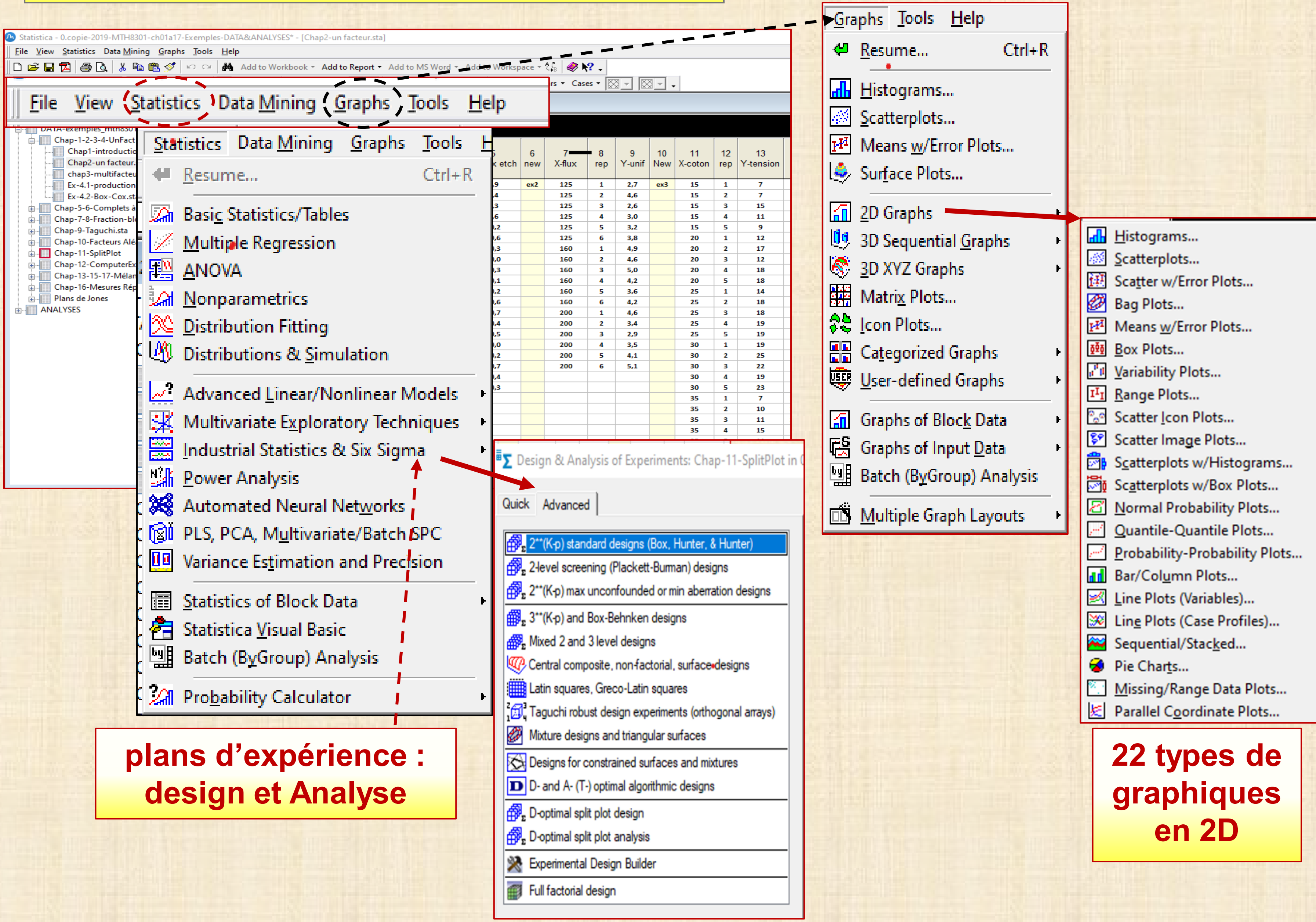
- **ANOVA** est la méthode générale employée pour analyser toutes les expériences industrielles et scientifiques
- si les écarts sont significatifs: identifier lesquels avec les méthodes des comparaisons multiples

Utilisation Statistica : classic view



	1 ID	2 new	3 X-solution	4 rep	5 Y-taux etch	6 new	7 X-flux	8 rep	9 Y-unif	10 New	11 X-coton	12 rep	13 Y-tension
1	1	ex1	1	1	9,9	ex2	125	1	2,7	ex3	15	1	7
2	2		1	2	9,4		125	2	4,6		15	2	7
3	3		1	3	9,3		125	3	2,6		15	3	15
4	4		1	4	9,6		125	4	3,0		15	4	11
5	5		1	5	10,2		125	5	3,2		15	5	9
6	6		1	6	10,6		125	6	3,8		20	1	12
7	7		1	7	10,3		160	1	4,9		20	2	17
8	8		1	8	10,0		160	2	4,6		20	3	12
9	9		1	9	10,3		160	3	5,0		20	4	18
10	10		1	10	10,1		160	4	4,2		20	5	18
11	11		2	1	10,2		160	5	3,6		25	1	14
12	12		2	2	10,6		160	6	4,2		25	2	18
13	13		2	3	10,7		200	1	4,6		25	3	18
14	14		2	4	10,4		200	2	3,4		25	4	19
15	15		2	5	10,5		200	3	2,9		25	5	19
16	16		2	6	10,0		200	4	3,5		30	1	19
17	17		2	7	10,2		200	5	4,1		30	2	25
18	18		2	8	10,7		200	6	5,1		30	3	22
19	19		2	9	10,4						30	4	19
20	20		2	10	10,3						30	5	23
21	21										35	1	7
22	22										35	2	10
23	23										35	3	11
24	24										35	4	15
25	25										35	5	11

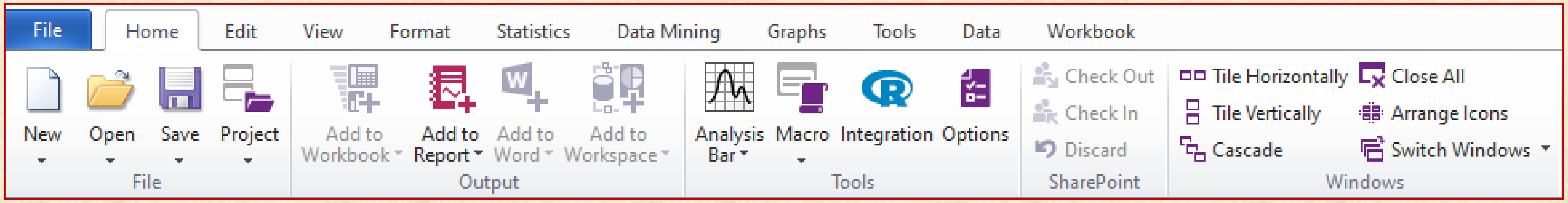
Utilisation Statistica : classic view

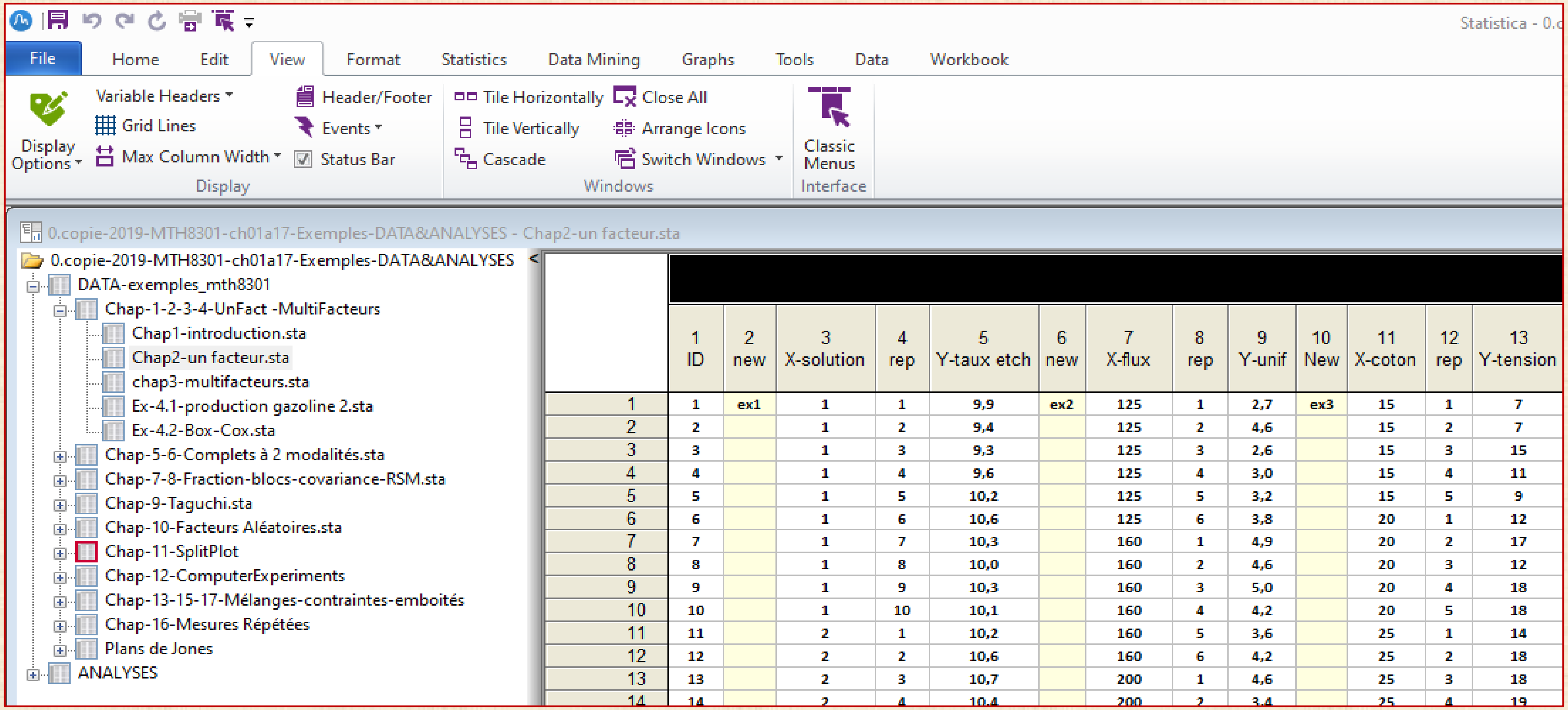


plans d'expérience : design et Analyse

22 types de graphiques en 2D

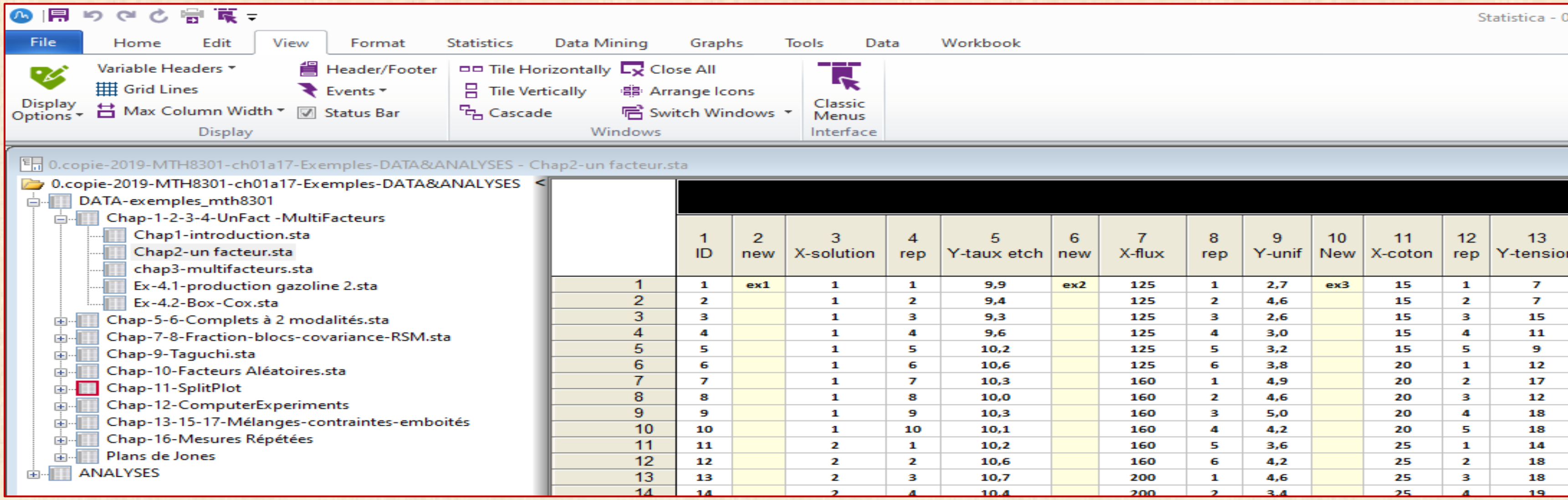
Utilisation Statistica : ribbon bar view

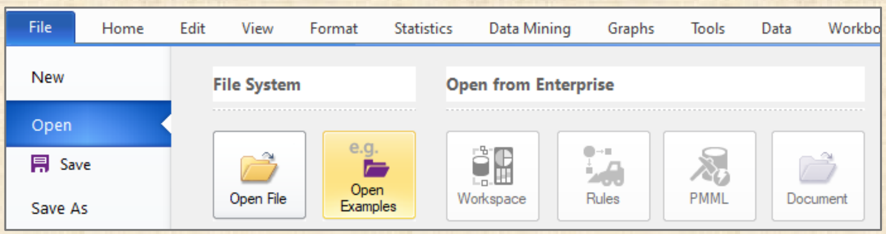


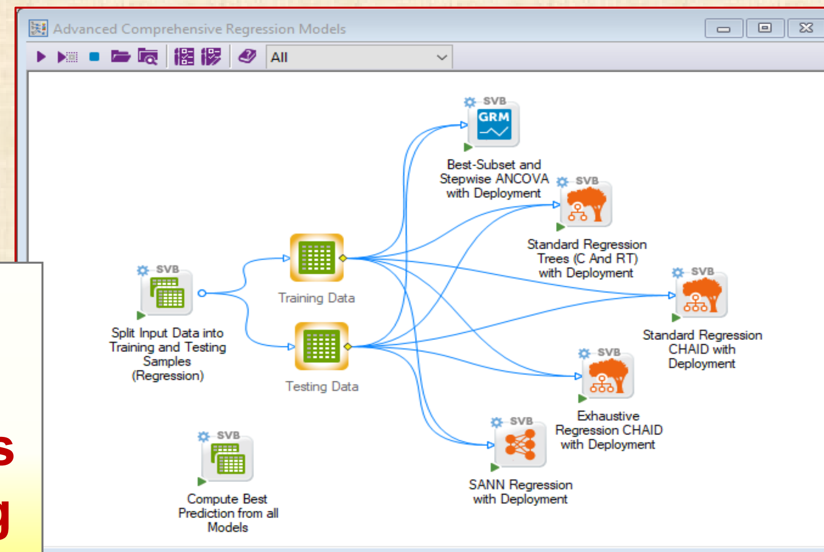


	1 ID	2 new	3 X-solution	4 rep	5 Y-taux etch	6 new	7 X-flux	8 rep	9 Y-unif	10 New	11 X-coton	12 rep	13 Y-tension
1	1	ex1	1	1	9,9	ex2	125	1	2,7	ex3	15	1	7
2	2		1	2	9,4		125	2	4,6		15	2	7
3	3		1	3	9,3		125	3	2,6		15	3	15
4	4		1	4	9,6		125	4	3,0		15	4	11
5	5		1	5	10,2		125	5	3,2		15	5	9
6	6		1	6	10,6		125	6	3,8		20	1	12
7	7		1	7	10,3		160	1	4,9		20	2	17
8	8		1	8	10,0		160	2	4,6		20	3	12
9	9		1	9	10,3		160	3	5,0		20	4	18
10	10		1	10	10,1		160	4	4,2		20	5	18
11	11		2	1	10,2		160	5	3,6		25	1	14
12	12		2	2	10,6		160	6	4,2		25	2	18
13	13		2	3	10,7		200	1	4,6		25	3	18
14	14		2	4	10,4		200	2	3,4		25	4	19

Utilisation Statistica : ribbon bar view







Statistica Workspace:

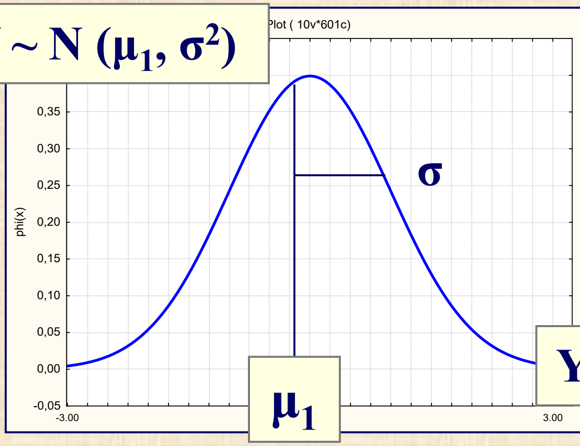
espace de travail pour traiter données sources diverses et Data Mining et Machine Learning

Expérience la plus simple

1 facteur à 2 modalités : test t Student (1/6)

modalité a1 ⟵ facteur A ⟶ modalité a2

$Y \sim N(\mu_1, \sigma^2)$

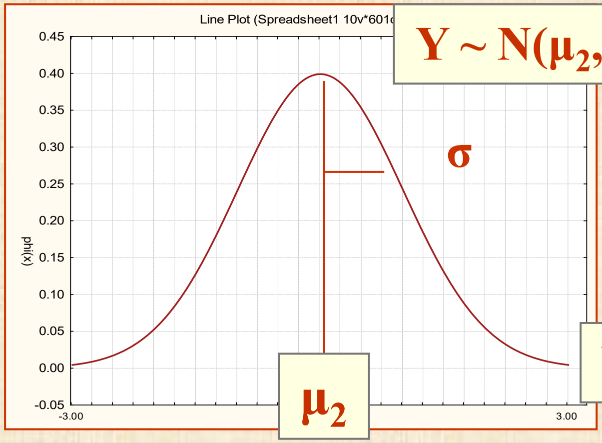


μ_1 Y

Hypothèse nulle
$H_0 : \mu_1 = \mu_2$

Hyp. alternative
$H_1 : \mu_1 \neq \mu_2$

$Y \sim N(\mu_2, \sigma^2)$



μ_2 Y

$y_{11}\ \ y_{12} \cdots\ \ y_{1\,n1}$ ⟵ échantillon ⟶ $y_{21}\ \ y_{22}\ \cdots\ \ y_{2\,n2}$

facteur A affecte t-il la variable de réponse Y ?

$\bar{y}_1 = \sum y_{1j} / n1$ ⟵ **moyennes** ⟶ $\bar{y}_2 = \sum y_{2j} / n2$

$s_1^2 = \sum (y_{1j} - \bar{y}_1)^2 / (n1 - 1)$ ⟵ **variances** ⟶ $s_2^2 = \sum (y_{2j} - \bar{y}_2)^2 / (n2 - 1)$

$\widehat{\sigma^2} = [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2] / (n_1 + n_2 - 2)$ **estimation erreur expérimentale σ**

décision basée sur écart $\bar{y}_1 - \bar{y}_2$

1 facteur à 2 modalités : test t Student (2/6)

test de comparaison $\longleftrightarrow$ effet du facteur A

$$t = \frac{\text{différence des moyennes}}{\text{écart type (différences des moyennes)}}$$

Statistique t de Student

$$t = \frac{\overline{y}_1 - \overline{y}_2}{\hat{\sigma}\ [1/n_1 + 1/n_2]^{0.5}}$$

loi Student avec df = $n_1 + n_2 - 2$ degrés de liberté

t « près de zéro » supporte
H_0 : pas de différence - le facteur A n'affecte pas Y

t « très différente de zéro » supporte
H_1 : le facteur A affecte la moyenne de Y

t est un rapport signal / bruit

t distance entre les moyennes en unités d'écart types

1 facteur à 2 modalités : test t Student (3/6)

- procédure objective pour décider si t est « grand »
- En 1908, W. S. Gosset (pseudonyme Student) obtient la distribution t

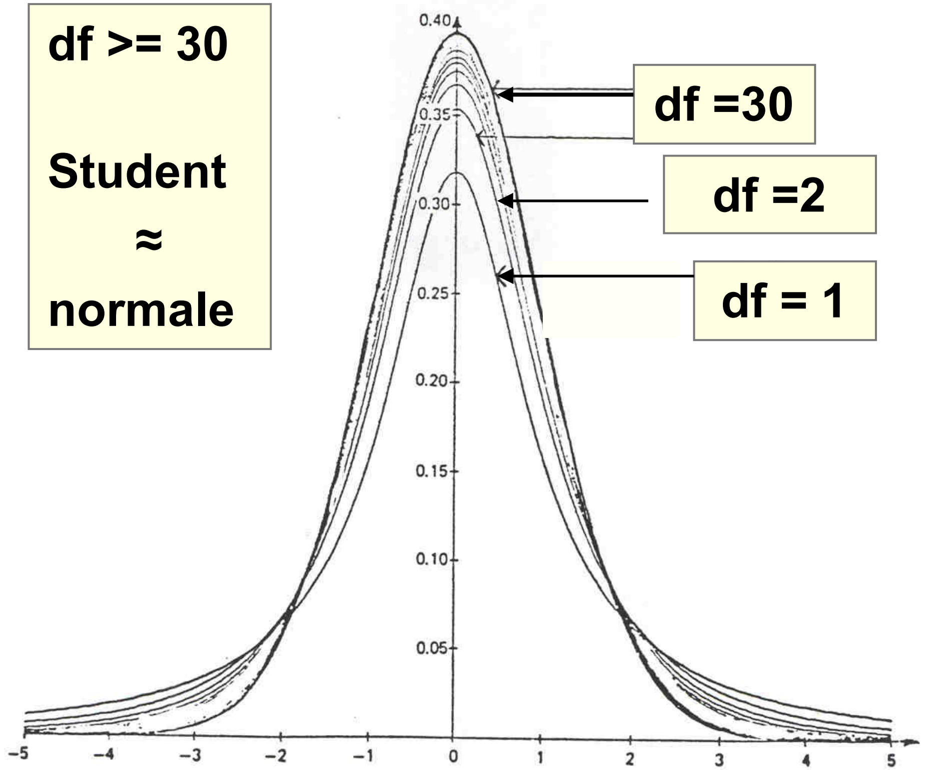
  appelé « Student »
- Tables
- logiciel statistique

  concept de « p-value »

distribution Student

df >= 30

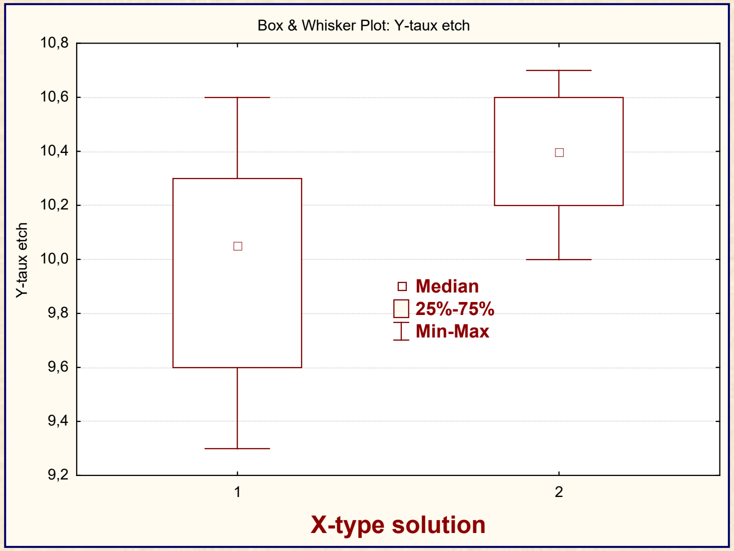
Student ≈ normale



1 facteur à 2 modalités : test t Student (4/6)

Ex 2.1 : analyse

solution	1	2
$\bar{y}$	9.97	10.40
S	0.42	0.23



T-tests; Grouping: X-type solution
Group 1: 1 Group 2: 2

	Mean	Mean	t-value	df	p	Valid N	Valid N	Std Dev.	Std Dev.	F-ratio	p
Y-taux etch	9.97	10.40	- 2.8278	18	0.011	10	10	0.42	0.23	3.3354	0.087

p -value = risque rejeter une hypothèse nulle H_0 vraie

interprétation : si p est « petit » (disons < 0.05) on rejette H_o

méthode : intervalle de confiance

$\mu_1 - \mu_2$: (- 0.76 , - 0.10)

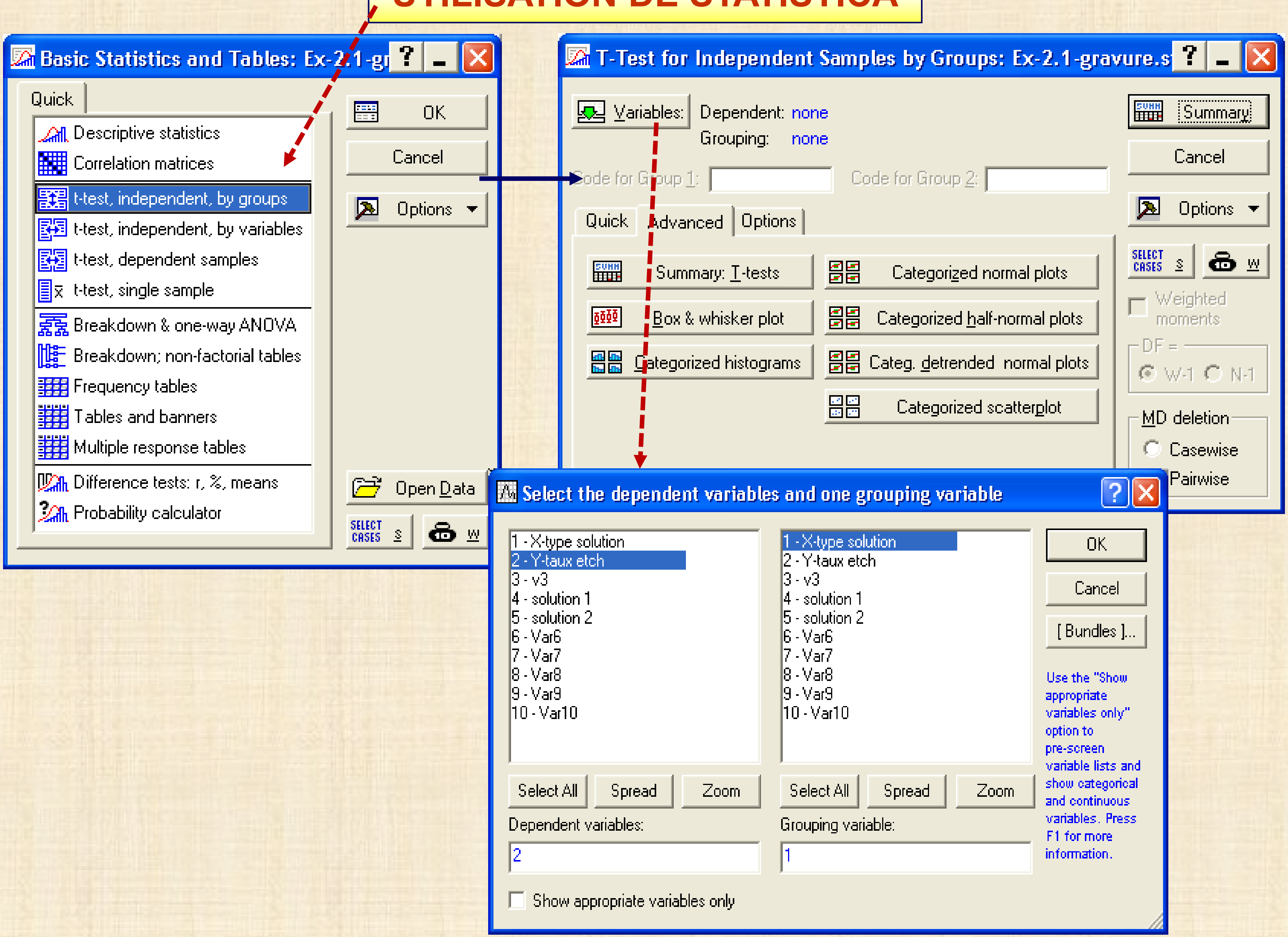
application 1 : les moyennes sont différentes car p = 0.011

application 2 : les variances sont égales car p = 0.087

test F de Fisher $F = s_1^2 / s_2^2$

1 facteur à 2 modalités : test t Student (5/6)

UTILISATION DE STATISTICA



Méthode des intervalles de confiance

intervalle de confiance pour un paramètre θ

$L \leq \theta \leq U$ avec $P(L \leq \theta \leq U) = 1 - \alpha$

$1 - \alpha$: coefficient de confiance

Application

Intervalle de confiance pour la différence entre 2 moyennes $\mu_1 - \mu_2$

$$\mu_1 - \mu_2 : (\overline{y_1} - \overline{y_2}) \pm \widehat{\sigma} * t_{df,\, 1-\alpha/2} * (1/n_1 + 1/n_2)^{0.5}$$

$$df = n_1 + n_2 - 2$$

percentile $(1 - \alpha/2)$ distribution Student

avec df degrés de liberté

Ex 2-1

$\mu_1 - \mu_2$: (- 0.76 , - 0.10)

avec coefficient de confiance de 95%

Expériences avec un facteur : plus de 2 groupes

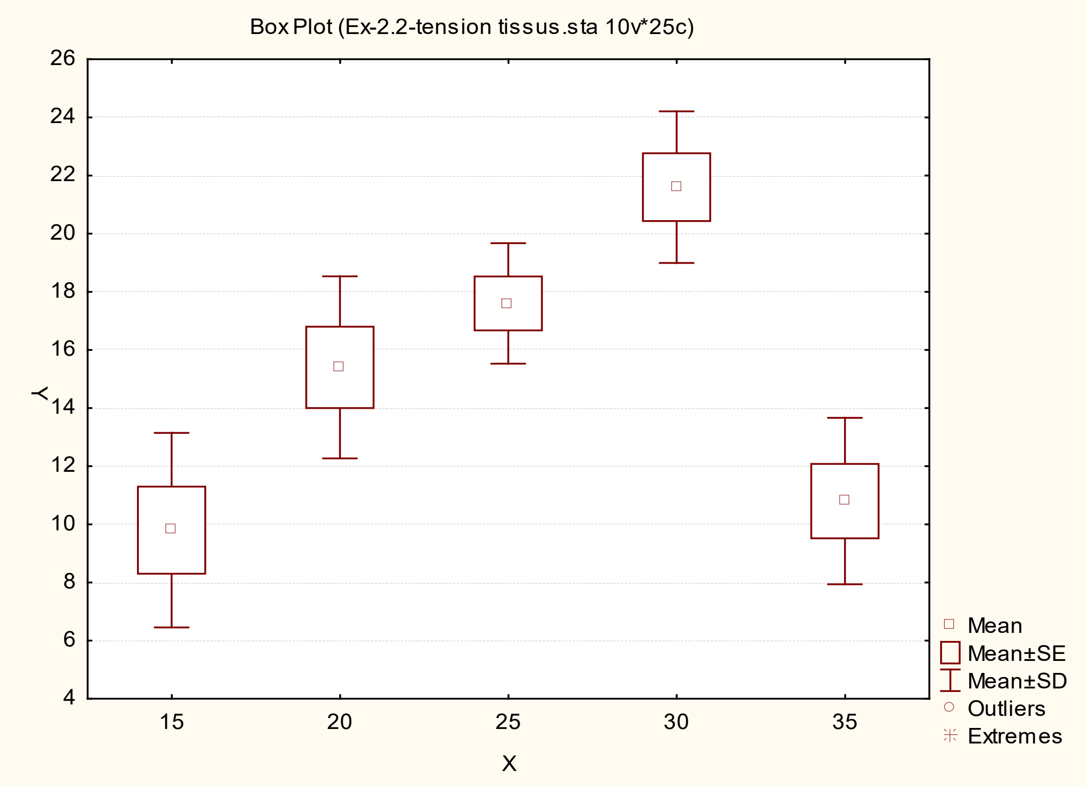
Ex-2.2–tension tissus

recherche nouvelle composition
de fibres synthétique tissus

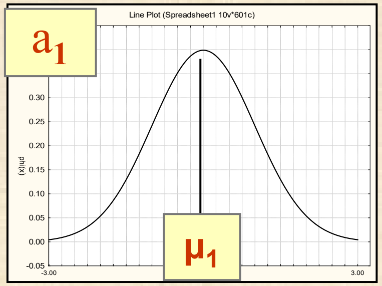
- facteur X : % coton varie entre 15 et 35
- réponse Y : force de tension tissu
- 5 modalités X = 15 20 25 30 35
- n = 5 répétitions

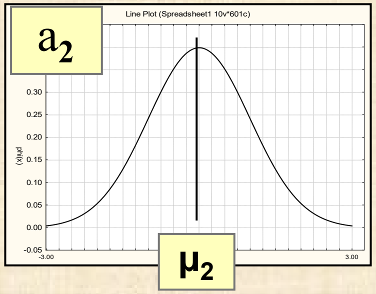
données n = 5 répétitions

X	i	1	2	3	4	5	moyenne
15	1	7	7	15	11	9	9.8
20	2	12	17	12	18	18	15.4
25	3	14	18	18	19	19	17.6
30	4	19	25	22	19	23	21.6
35	5	7	10	11	15	11	10.8



Expériences avec un facteur : plus de 2 groupes





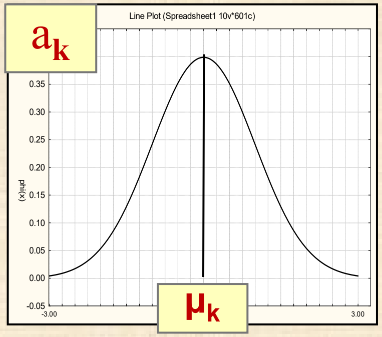


Tableau des données

modalité	observations	moyennes	variances
a_1	$y_{11}\ y_{12}\ y_{13} \cdots y_{1\,n1}$	$\bar{y}_{1.}$	s_1^2
a_2	$y_{21}\ y_{22}\ y_{23} \cdots y_{2\,n2}$	$\bar{y}_{2.}$	s_2^2
....	....	....	....
a_i	$y_{i1}\ y_{12}\ y_{i3} \cdots y_{i\,ni}$	$\bar{y}_{i.}$	s_i^2
....	....	....	....
a_k	$y_{k1}\ y_{k2}\ y_{k3} \cdots y_{k\,nk}$	$\bar{y}_{k.}$	s_k^2
toutes		$\bar{y}_{..}$	

$$y_{i.} = \sum y_{ij}$$
$$y_{..} = \sum \sum y_{ij}$$
$$\bar{y}_{i.} = y_{i.} / n_i$$
$$N = \sum n_i$$
$$\bar{y}_{..} = y_{..} / N$$
$$SS_i = \sum (y_{ij} - \bar{y}_{i.})^2$$
$$s_i^2 = SS_i / (n_i - 1)$$

- traitements à comparer : k modalités du facteur
- méthode d'assignation complètement aléatoire :
  les k traitements sont assignés aux N unités
  sans aucune restriction (au hasard, randomisation)
- préférable: choisir des tailles égales $n_i = n$ $n = ?$
  meilleure comparaison entre les traitements
- effets fixés - aussi cas avec effets aléatoires
- hypothèse nulle $H_0: \mu_1 = \mu_2 = \ldots = \mu_k$
  pas de différences entre les moyennes de groupes

design = protocole expérimental

Expériences avec un facteur : plus de 2 groupes

Modèle

$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ $\quad i = 1, 2, \ldots, k \quad j = 1, 2, \ldots, n_i$

μ : effet général

τ_i : effet différentiel traitement i

$\sum \tau_i = 0$

ε_{ij} : erreur $\sim N(0,\sigma^2)$

$\sim$: notation pour distribution

Autres modèles

Modèle à moyennes : $Y_{ij} = \mu_i + \varepsilon_{ij} \qquad \mu_i = \mu + \tau_i$

Modèle polynomial si facteur est quantitatif (A = X)

par exemple : $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$

Expériences avec un facteur : plus de 2 groupes

Analyse de la variance - ANOVA

$$\sum\sum(y_{ij}-\bar{y}_{..})^2 = \sum\sum[(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})]^2$$
$$= \sum n_i(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum\sum(y_{ij}-\bar{y}_{i.})^2$$

$$\sum\underbrace{(\bar{y}_{i.}-\bar{y}_{..})}_{0}\sum(y_{ij}-\bar{y}_{i.}) = 0$$

$$SS_{tot} = \sum\sum(y_{ij}-\bar{y}_{..})^2$$ variabilité totale

$$SS_{trait} = \sum n_i(\bar{y}_{i.}-\bar{y}_{..})^2$$ inter variabilité

$$SS_{erreur} = \sum\sum(y_{ij}-\bar{y}_{i.})^2$$ intra variabilité

$$SS_{tot} = SS_{trait} + SS_{erreur}$$

Expériences avec un facteur : plus de 2 groupes

Tableau d'analyse de la variance

Source	Somme Carrés (SS)	Deg. lib.(df)	Carré Moyen (MS)	F
Traitements	$SS_{trait} = \sum n_i(\overline{y}_{i.} - \overline{y}_{..})^2$	$k - 1$	$MS_{trait} = SS_{trait}/(k-1)$	$F_0 = \dfrac{MS_{trait}}{MS_{erreur}}$
Erreur	$SS_{erreur} = \sum\sum(y_{ij} - \overline{y}_{i.})^2$	$N - k$	$MS_{erreur} = SS_{erreur}/(N-k)$	
Totale	$SS_{tot} = \sum\sum(y_{ij} - \overline{y}_{..})^2$	$N - 1$		

SS: Sum of Square   df: degree of freedom   MS: Mean Square

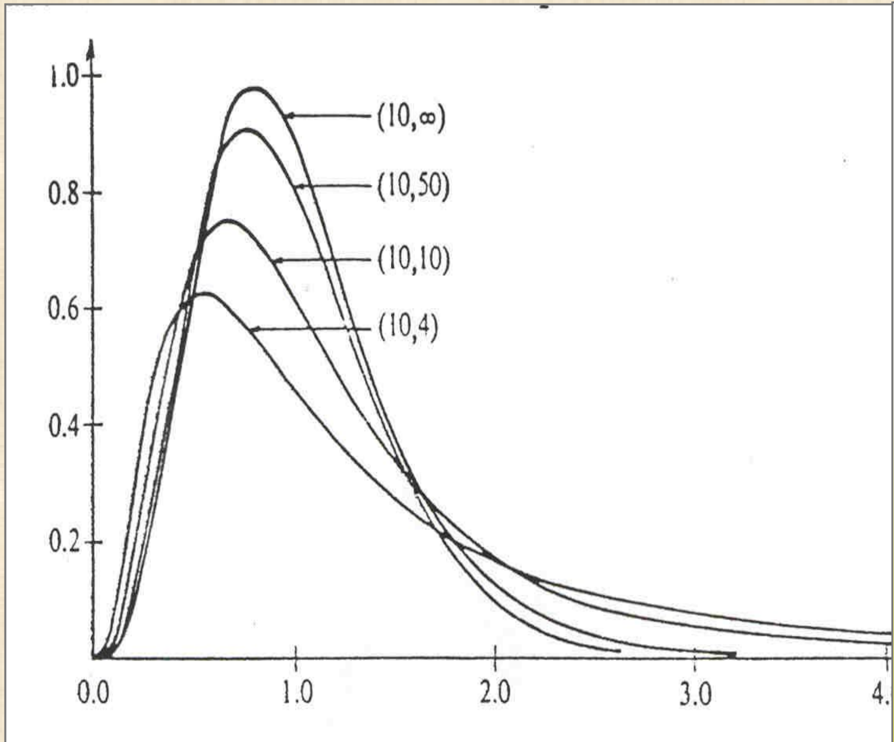
distribution de référence pour F_0 : distribution F de Fisher
avec $df_1 = k - 1$ degrés de liberté au numérateur
et $df_2 = N - k$ degrés de liberté au dénominateur

Test de H_0 : $\mu_1 = \mu_2 = \ldots. = \mu_k$

Rejet de H_0 au seuil (risque) α si $F_0 > F_{\alpha, df1, df2}$

Expériences avec un facteur : plus de 2 groupes

Distribution F de Fisher



- si X1 suit une loi Khi-deux avec df1 ddl
- X2 suit une loi Khi-deux avec df2 ddl
- X1 et X2 sont indépendantes

alors

(X1/df1) / (X2/df2) suit une loi F (df1,df2)

- t^2_{df} = F (1, df) : carré Student

  = Fisher F(df1 = 1, df2 = df)

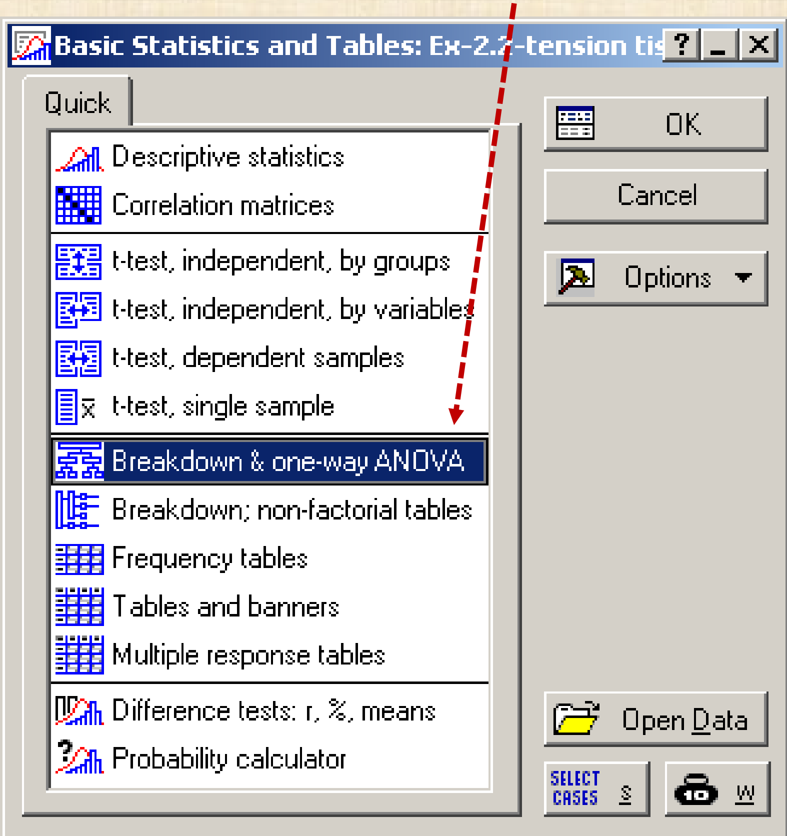
la distribution F est la plus importante de toutes les distributions

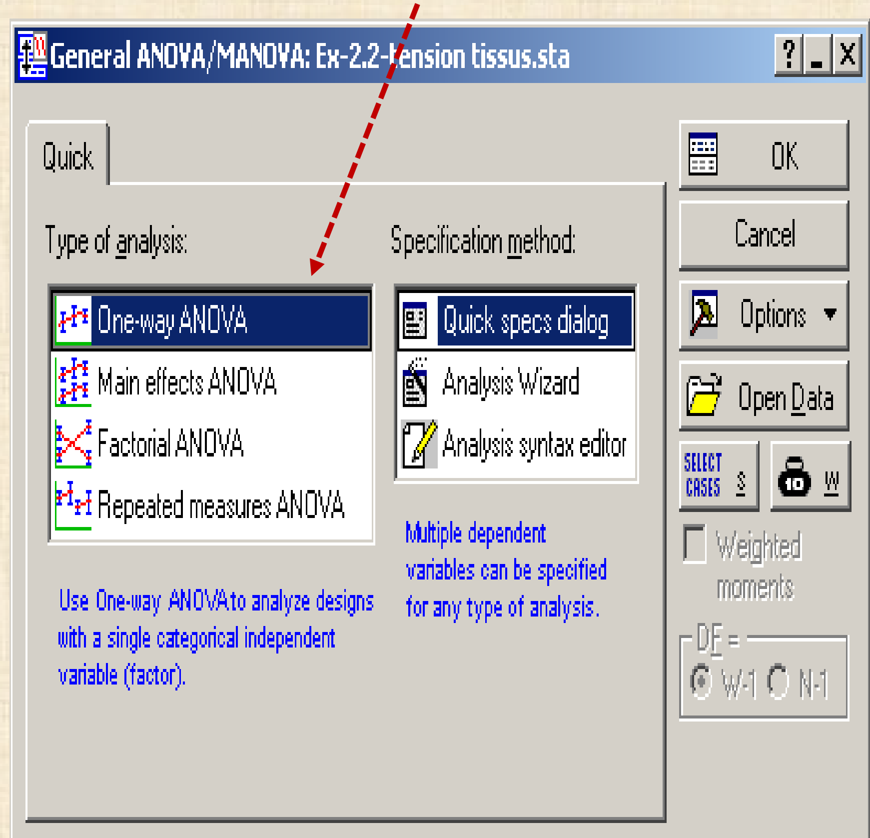
car elle est employée dans TOUTES les analyses de plans d'expériences

Expériences avec un facteur : plus de 2 groupes

Analyse avec STATISTICA

Basic Statistics and Tables ou **Statistics … ANOVA …**



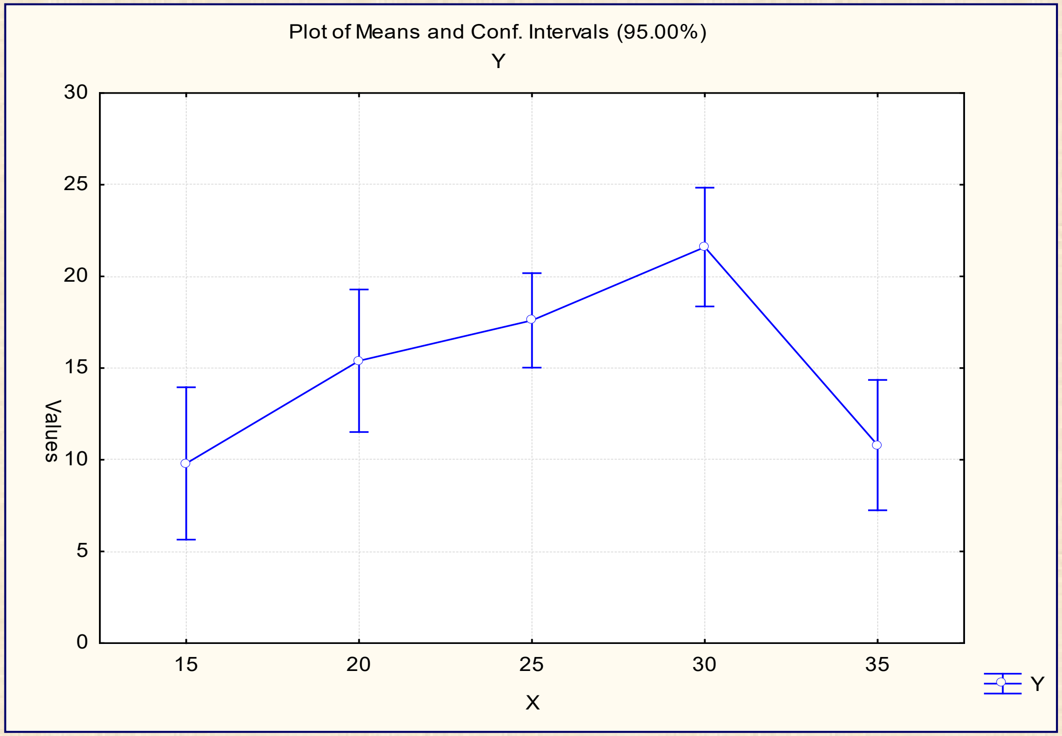


Expériences avec un facteur : plus de 2 groupes

Ex. 2.2 : analyse avec STATISTICA

avec Basic Statistics and Tables

	SS effect	df effect	MS effect	SS error	df error	MS error	F	p
Y	475.76	4	118.94	161.20	20	8.06	14.76	0.000009



différences significatives

lesquelles ?

méthode des comparaisons multiples

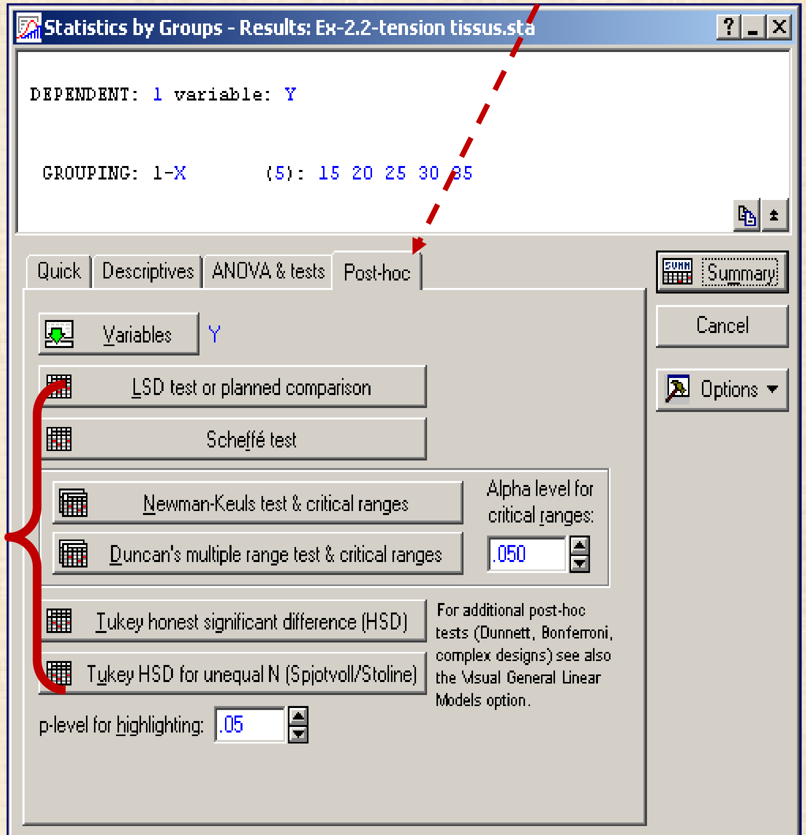
Expériences avec un facteur : plus de 2 groupes

Ex. 2.2 : analyse avec STATISTICA

Post-hoc : comparaisons multiples

LSD – Scheffé - Newman&Keuls

Duncan - Tukey



Méthode recommandée

Tukey *Honest Significant Differences HSD*

comparisons de toutes les différences entre les paires de modalités

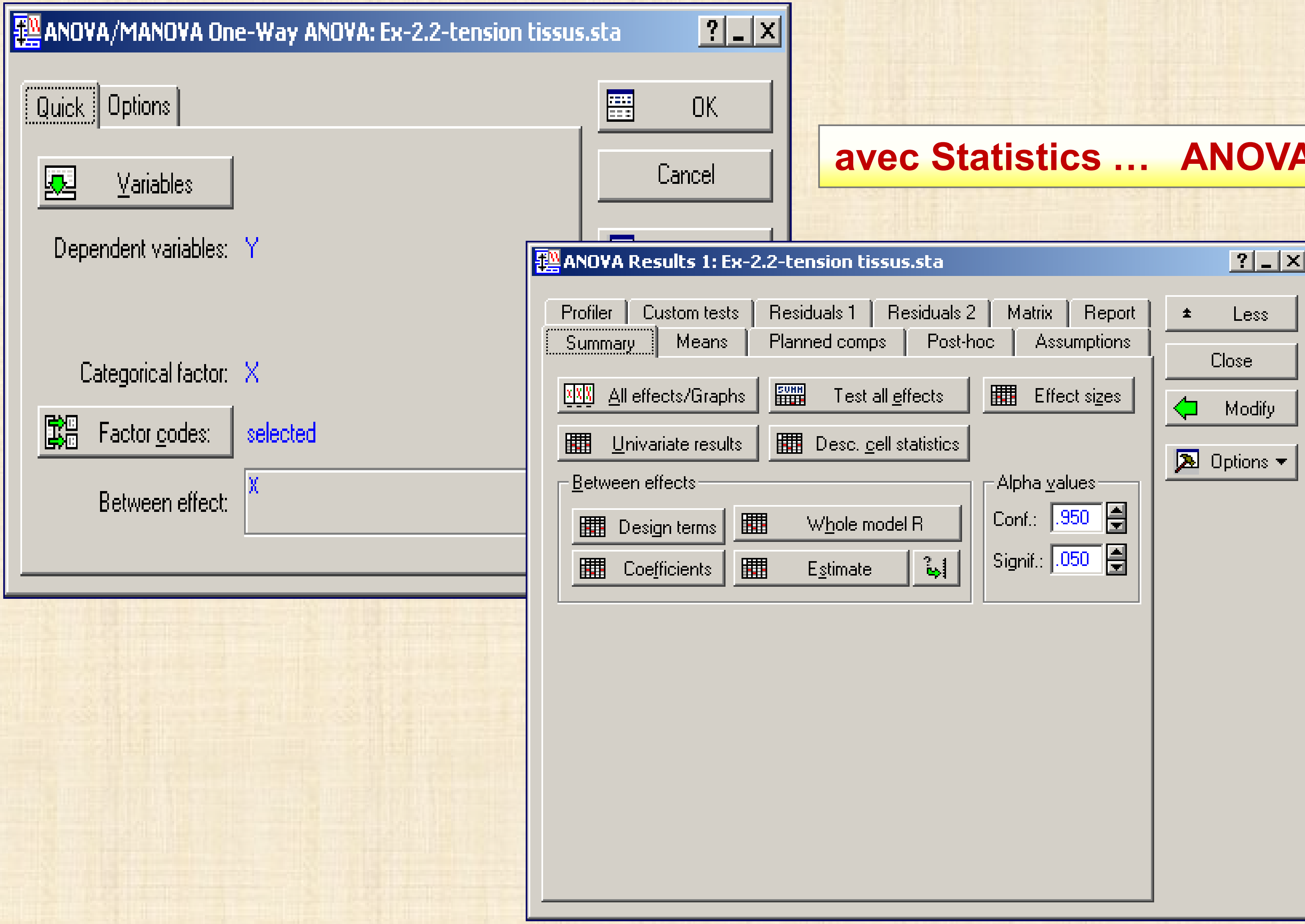
Tukey HSD test; Variable: Y

marked differences are significant at $p < 0.05000$

moy	{1} 9.8	{2} 15.4	{3} 17.6	{4} 21.6	{5} 10.8
15 {1}		0.0386	0.0027	0.0001	0.9798
20 {2}	0.0386		0.7373	0.0190	0.1164
25 {3}	0.0027	0.7373		0.2102	0.0092
30 {4}	0.0001	0.0190	0.2102		0.0002
35 {5}	0.9798	0.1164	0.0092	0.0002	

Expériences avec un facteur : plus de 2 groupes

Ex. 2.2 : analyse avec STATISTICA



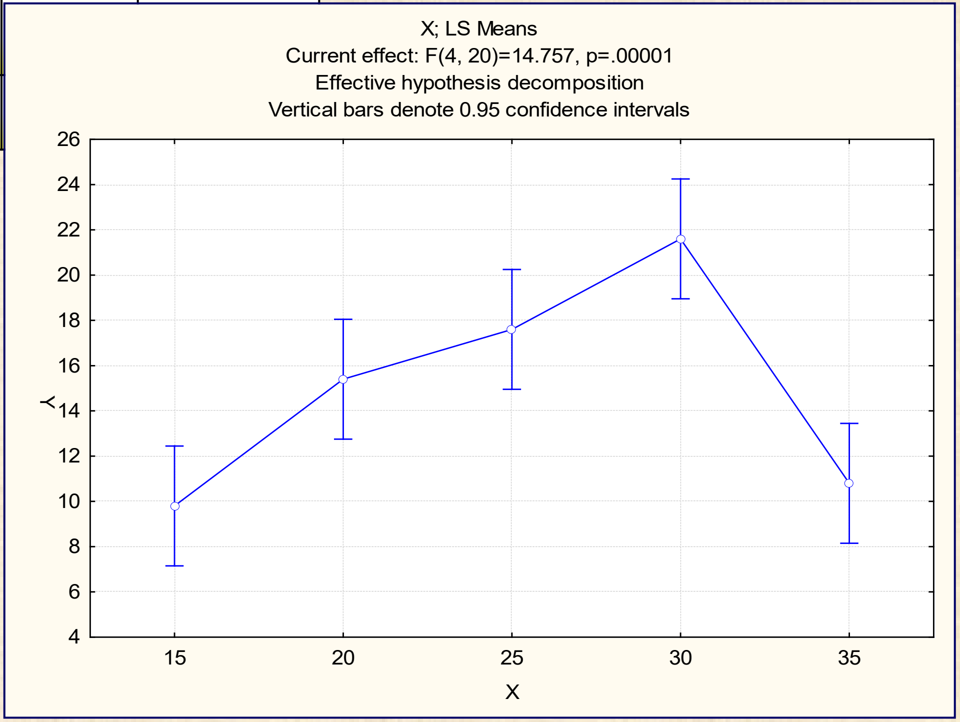
avec Statistics … ANOVA …

Expériences avec un facteur : plus de 2 groupes

Ex. 2.2 : analyse avec STATISTICA … Statistics … ANOVA

Univariate Results for Each DV (Ex-2.2-tension tissus)

	Degr. Of freedom	Y SS	Y MS	Y F	Y p
Intercept	1	5655.04	5655.04	701.62	0.000000
X	4	475.76	118.940	14.76	0.000009
Error	20	161.20	8.060		
Total	24	636.96			



Analyse des résidus doit être exécutée avec tout modèle d'analyse statistique

- vérifier les hypothèses de base après ajustement d'un modèle

HYPOTHESES de BASE (ordre d'importance) COMMENT?

H1 - variance constante? …… tests de Hartley /Cochran / Bartlett / Levene

H2 - données aberrantes? ('outliers') … résidus vs résidus 'deleted'

H3 - distribution normale? …… résidus sur échelle gaussienne (normale)

H4 - « bon » modèle? …… R^2 et R^2 ajusté élevé: pas absolument nécessaire

H5 - observations indépendantes? ….. test Durbin-Watson

- Si hypothèses de base violées et / ou certaines formes réponse
  - transformation de la réponse Y : Box-Cox Y^{λ} $-2 < \lambda < 2$ $\lambda = ?$
  - test non paramétrique de Kruskall-Wallis

- H1 critique seulement si max var(Y) / min var(Y) > 10 (inter groupes)
- H2 on ne veut pas que l'analyse soit dominée par quelques observations
- H3 le test F est robuste vis-à-vis la non normalité
- H4 est utile en analyse de régression
- pas utile en modèle d' analyse de variance (variables X catégoriques)
- H5 surtout pour données observationnelles collectées à forte cadence dans le temps : pas le cas si les données proviennent d'expériences

Transformations de la variable de réponse : cas de variances inégales

CONDITION	TRANSFORMATION
réponse Y est un comptage : distribution Poisson	$Y' = \sqrt{Y}$ ou $Y' = \sqrt{Y} + \sqrt{Y+1}$
réponse Y est une proportion : distribution binomiale	$Y' = 2 \arcsin(\sqrt{Y})$
(ecart type)^2 proportionnel à la moyenne (s^2/m)	$Y' = \sqrt{Y}$
écart type proportionnel à la moyenne (s/m)	$Y' = \log(Y)$
écart type proportionnel à la (moyenne)^2 (s/m^2)	$Y' = 1 / Y$

Recommandation 1

examiner les quantités $s_i^2 / Y_{i.}$, $s_i / Y_{i.}$, $s_i / Y_{i.}^2$
pour chaque niveau du facteur et choisir la transformation dont le coefficient de variation (CV) est le plus petit

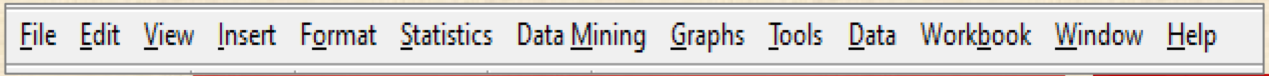
Recommandation 2

transformation Box-Cox sur Y

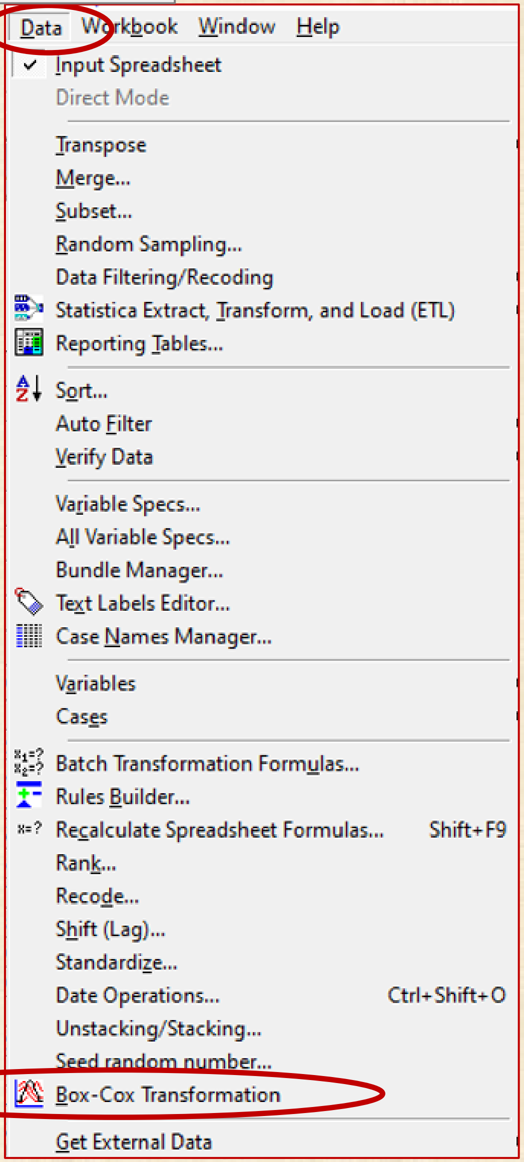
$$Y' = Y^{\lambda} \qquad -2 < \lambda < 2$$

$\lambda = ?$ on choisit λ tel que $SSE(\lambda)$ soit minimum
on prend une valeur arrondie
si $\lambda = 0$ $Y' = \log(Y)$

TRANSFORMATION RÉPONSE Y : utiliser Box-Cox



mettre en œuvre la transformation de Box-Cox
onglet Data du menu principal de Statistica



Transformation de Box Cox si l'écart type de Y σ = ET(Y)
est proportionnel à la moyenne μ = E(Y) : $\sigma \;\alpha\; \mu^{\alpha}$
examen de log ($\bar{Y}_i$) versus log (s_i) : pente α

Transformation pour stabiliser la variance : Y ' = h(Y)

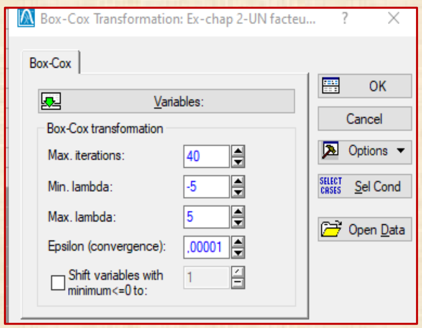
cas	relation		transformation Y'
Y ~ attribut	$\sigma^2 \;\alpha\; \mu(1-\mu)$		Y ' = arcsin($Y^{0.5}$)
Box-Cox	$\sigma \;\alpha\; \mu^{\alpha}$	$\lambda = 1 - \alpha$	$Y' = Y^{\lambda}$ $-2 \le \lambda \le 2$
σ constant	$\alpha = 0$	$\lambda = 1$	Y '= Y pas de transformation
$\sigma \;\alpha\; \mu^{0.5}$ (Poisson)	$\alpha = 0.5$	$\lambda = 0.5$	$Y' = Y^{0.5}$
$\sigma \;\alpha\; \mu$	$\alpha = 1$	$\lambda = 0$	Y ' = log(Y)
$\sigma \;\alpha\; \mu^{1.5}$	$\alpha = 1.5$	$\lambda = -0.5$	$Y' = Y^{-0.5}$
$\sigma \;\alpha\; \mu^{2}$	$\alpha = 2$	$\lambda = -1$	$Y' = Y^{-1}$

Exemple : Y-durée temps panne 3 villes A B C

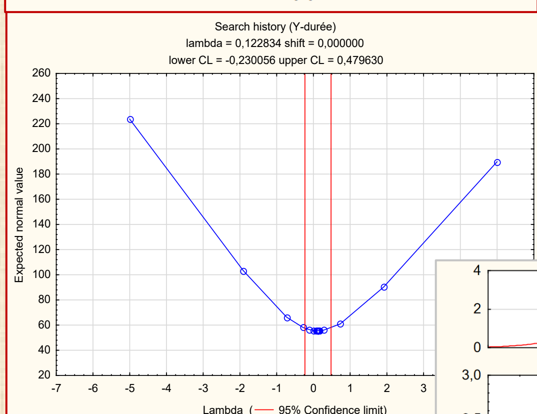
24 new	25 ville	26 rep	27 Y-durée	28 Y_rang	29 Y_Box-Cox	30 Y_log10(Y-durée)
transformation	A	1	4,41	2	1,6277	0,6444
Box-Cox	A	2	100,65	13	6,2038	2,0028
	A	3	14,45	6	3,1609	1,1599
ex4	A	4	47,13	9	4,9272	1,6733
Y	A	5	85,21	12	5,9133	1,9305
durée (min)	B	1	8,24	4	2,4074	0,9159
pannes	B	2	81,16	11	5,8295	1,9093
serveurs	B	3	7,35	3	2,2603	0,8663
villes	B	4	12,29	5	2,9383	1,0896
A B C	B	5	1,61	1	0,4904	0,2068
	C	1	106,19	14	6,2985	2,0261
	C	2	33,83	7	4,4057	1,5293
	C	3	78,88	10	5,7807	1,8970
	C	4	342,81	15	8,5342	2,5351
	C	5	44,33	8	4,8293	1,6467

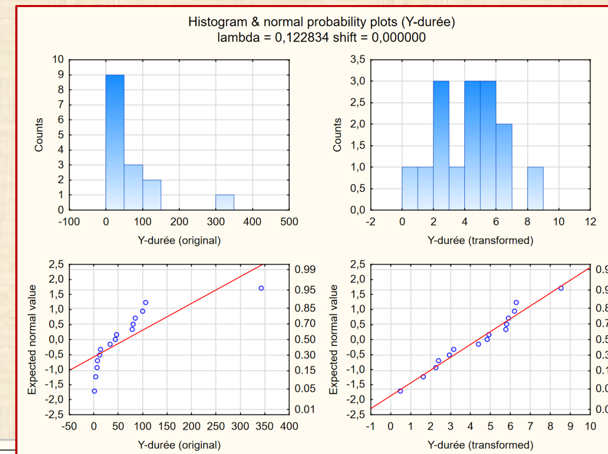
ville	obs	m Moyenne	s écart type	s^2 / m	s / m	s / m^2
A	5	50.37	42.29	35.51	0.84	0.017
B	5	22.13	33.22	49.86	1.50	0.068
C	5	121.21	127.15	133.39	1.05	0.009
			moyenne	72,92	1,13	0,031
			std dev	52,854	0,338	0,032
			CV(%)	72	30	103

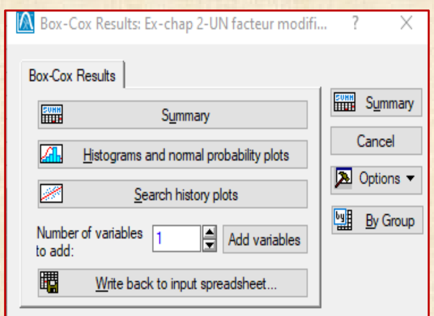
minimum cv(%) obtenu avec s / m
choix:
transformation logarithmique de Y

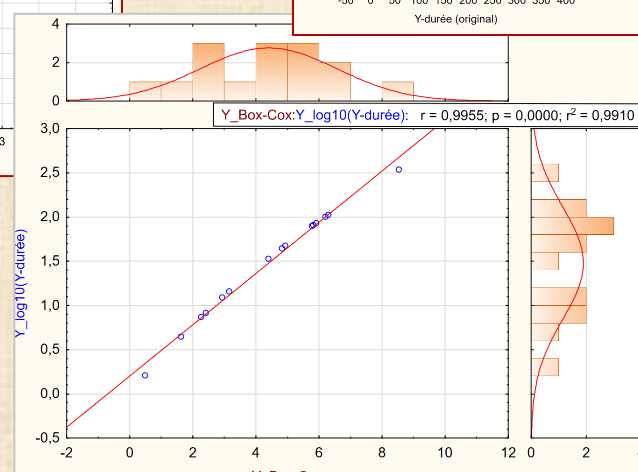


Search history (Y-durée)
lambda = 0,122834 shift = 0,000000
lower CL = -0,230056 upper CL = 0,479630





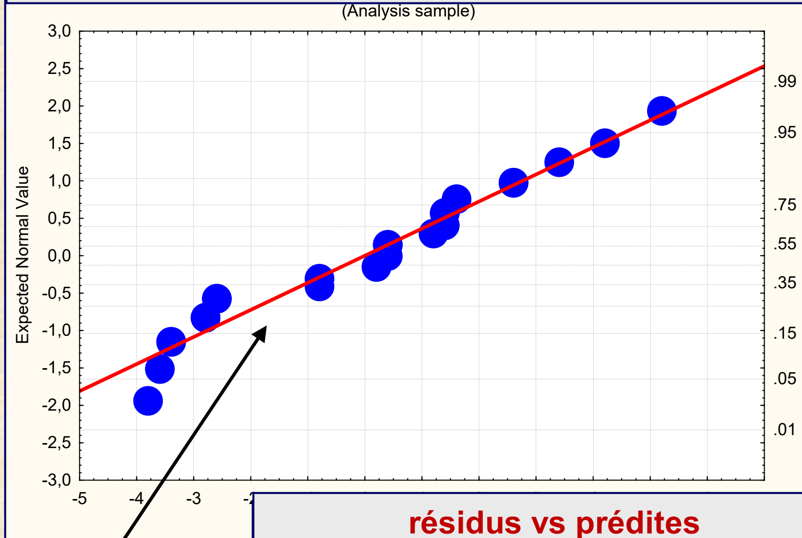




Expériences avec un facteur : plus de 2 modalités

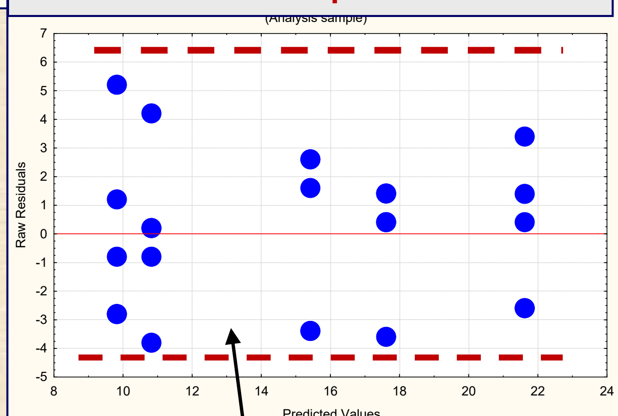
Analyse des résidus : avec jugement visuel des résidus

résidus sur échelle probabilité gaussienne



points alignés

résidus vs prédites



points bande horizontale

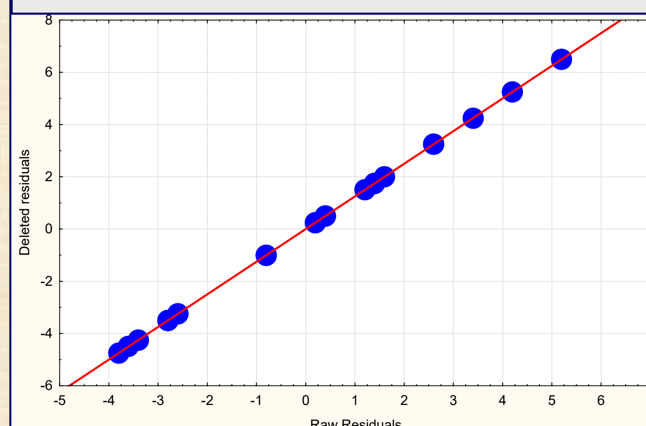
Tests of homogénéité variances

	Hartley	Cochran	Bartlett	df	p
Y	2.605	0.278	0.934	4	0.9198

Levene's Test for Homogeneity of Variances
Effect: X Degrees of freedom for all F's: 4, 20

	MS	MS	F	p
Y	1.37	2.13	0.64	0.6372

résidus vs résidus ‘deleted’



Expériences avec un facteur : plus de 2 groupes

n = nombre de répétitions de chaque essai = ?

n dépend de

alpha (α) : taux (probabilité) de fausse détection ($0 < \alpha < 1$)
risque de rejeter une hypothèse vraie
valeurs souvent employées: 0,05 / 0,01

beta (β) : taux (probabilité) de manque de détection ($0 < \beta < 1$)
risque de ne pas rejeter une hypothèse fausse
valeurs souvent employées: 0,20 / 0,10
$1 - \beta$: puissance du test

σ : erreur expérimentale

$\Delta = \lambda \sigma$: écart de moyenne à détecter

$\lambda = \Delta/\sigma$: facteur de proportionnalité
en pratique $0.5 < \lambda < 3$

k : nombre de modalités (groupes) à comparer

n : nombre de répétitions de chaque sous groupe (modalité)

$$n = F(\alpha , \beta, \sigma, \lambda , k)$$

expériences avec plusieurs facteurs :
n entre 2 et 5 : généralement suffisant

Expériences avec un facteur

nombre de répétitions n = ?

* : > 100

k = 2 — n

alpha	0.10		0.05		0.01	
beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	70	88	86	*	*	*
1.0	18	23	23	27	32	38
1.6	8	10	10	12	14	16
2.0	6	7	7	8	10	11
3.0	3	4	4	5	6	6

k = 3 — n

alpha	0.10		0.05		0.01	
beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	85	*	*	*	*	*
1.0	22	27	27	32	37	43
1.6	10	12	11	14	16	18
2.0	7	8	8	9	11	12
3.0	4	4	5	5	6	7

k = 4 — n

alpha	0.10		0.05		0.01	
beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	70	88	86	*	*	*
1.0	25	30	30	36	40	47
1.6	11	13	13	15	17	20
2.0	7	9	9	10	12	13
3.0	4	5	5	5	6	7

k = 5 — n

Alpha	0.10		0.05		0.01	
Beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	*	*	*	*	*	*
1.0	27	33	32	39	43	50
1.6	11	14	14	16	18	21
2.0	8	9	9	11	12	14
3.0	4	5	5	6	7	7

k = 6 — n

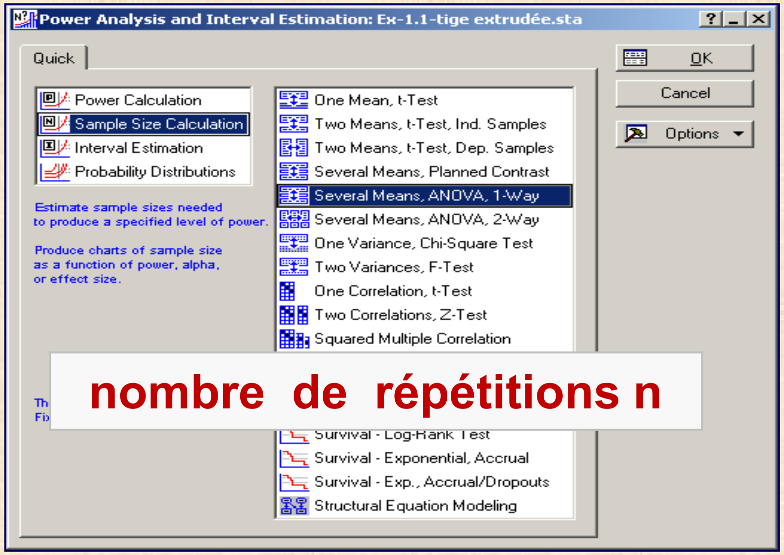
Alpha	0.10		0.05		0.01	
Beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	*	*	*	*	*	*
1.0	29	35	34	41	46	53
1.6	12	15	14	17	19	22
2.0	8	10	10	11	13	15
3.0	4	5	5	6	7	8

k = 7 — n

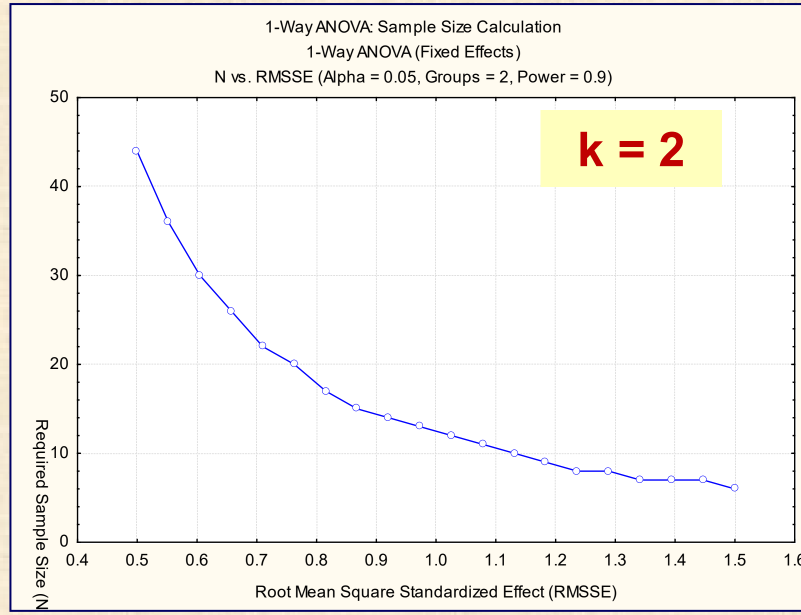
Alpha	0.10		0.05		0.01	
Beta	0.10	0.05	0.10	0.05	0.10	0.05
λ 0.5	*	*	*	*	*	*
1.0	31	37	36	43	48	56
1.6	13	15	15	18	19	23
2.0	9	10	10	12	13	15
3.0	5	5	5	6	7	8

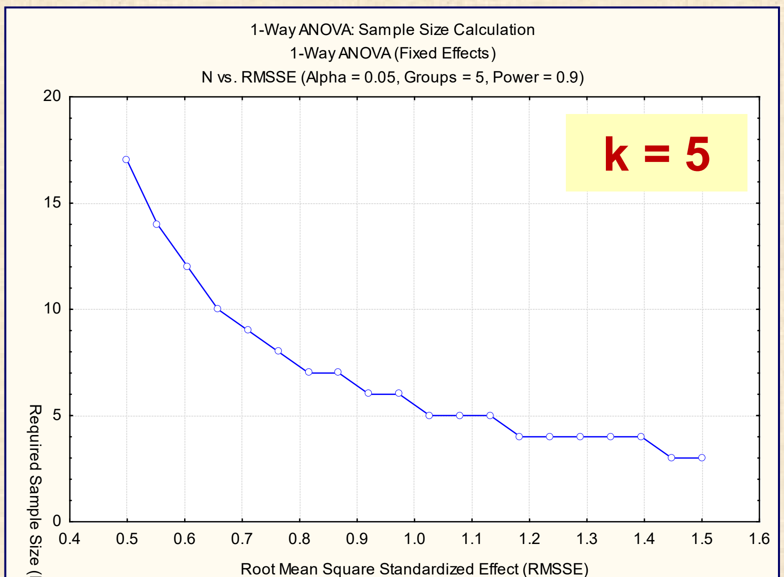
combien de répétitions ?

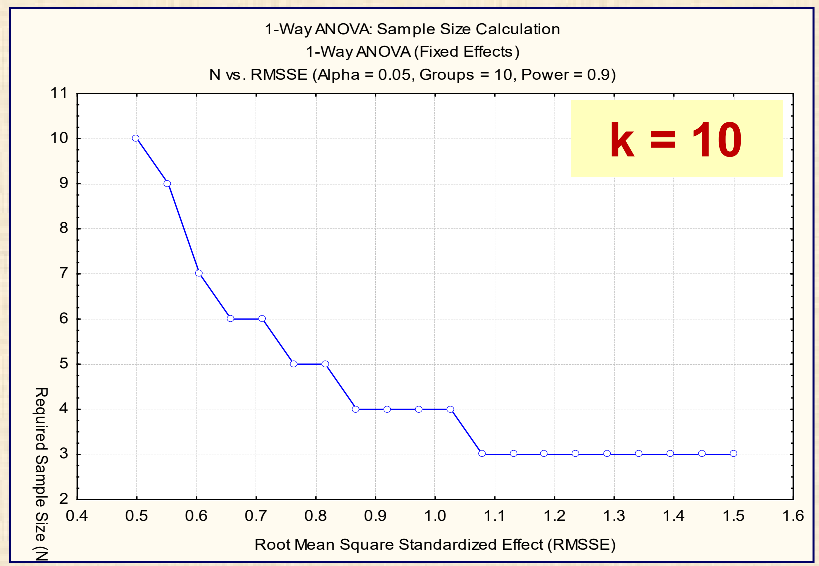
Statistics … Power Analysis



nombre de répétitions n



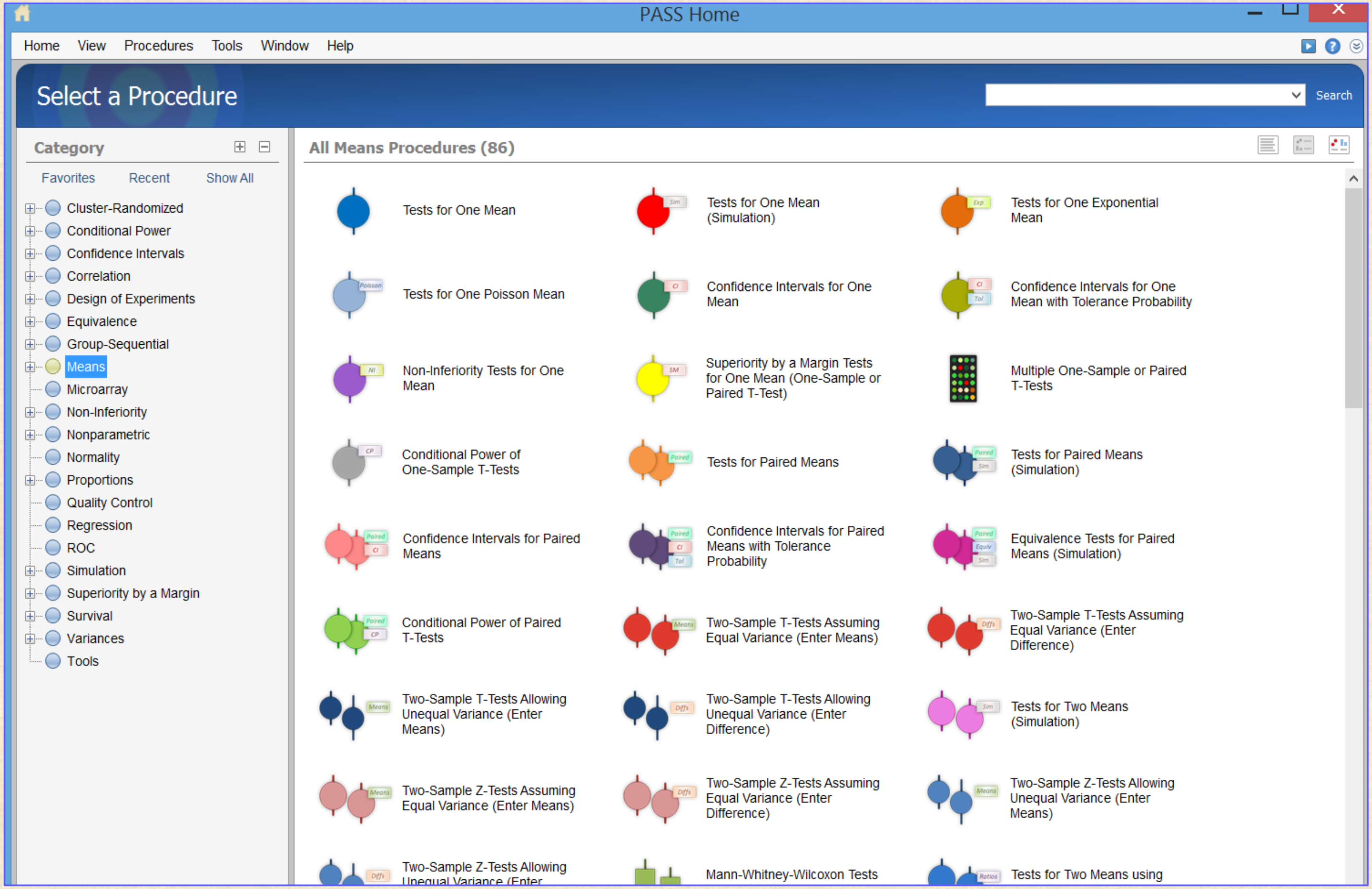




PASS 14
Sample Size
Enjoy the Best Sample Size Software on the Market

PASS : Power Analysis and Sample Size

http://www.ncss.com/



AUTRES SUJETS

- **effet de dispersion**

  Montgomery 8ième ed. p. 116

  analyse faite avec Y = log(s)   s : écart type

- **ANOVA non paramétrique**

  test de Kruskall-Wallis

  Montgomery 8ième ed. p. 128

  analyse ANOVA faite avec les rangs

  $$r_{ij} = \text{rang}(Y_{ij})$$

ANOVA non paramétrique : test de Kruskall-Wallis

Les méthodes paramétriques sont basées sur les rangs de la variable de réponse plutôt que les valeurs observées. On assigne aux observations Y_{ij} le rang R_{ij} des valeurs ordonnées en ordre croissant de 1 à N. On procède comme dans le test F usuel que l'on applique aux rangs R_{ij}.

Test de Kruskall-Wallis

$$F_{KW} = MSTR / MSE$$

$$MSTR = \sum n_i (\overline{R}_{i.} - \overline{R}_{..})^2 / (g - 1)$$

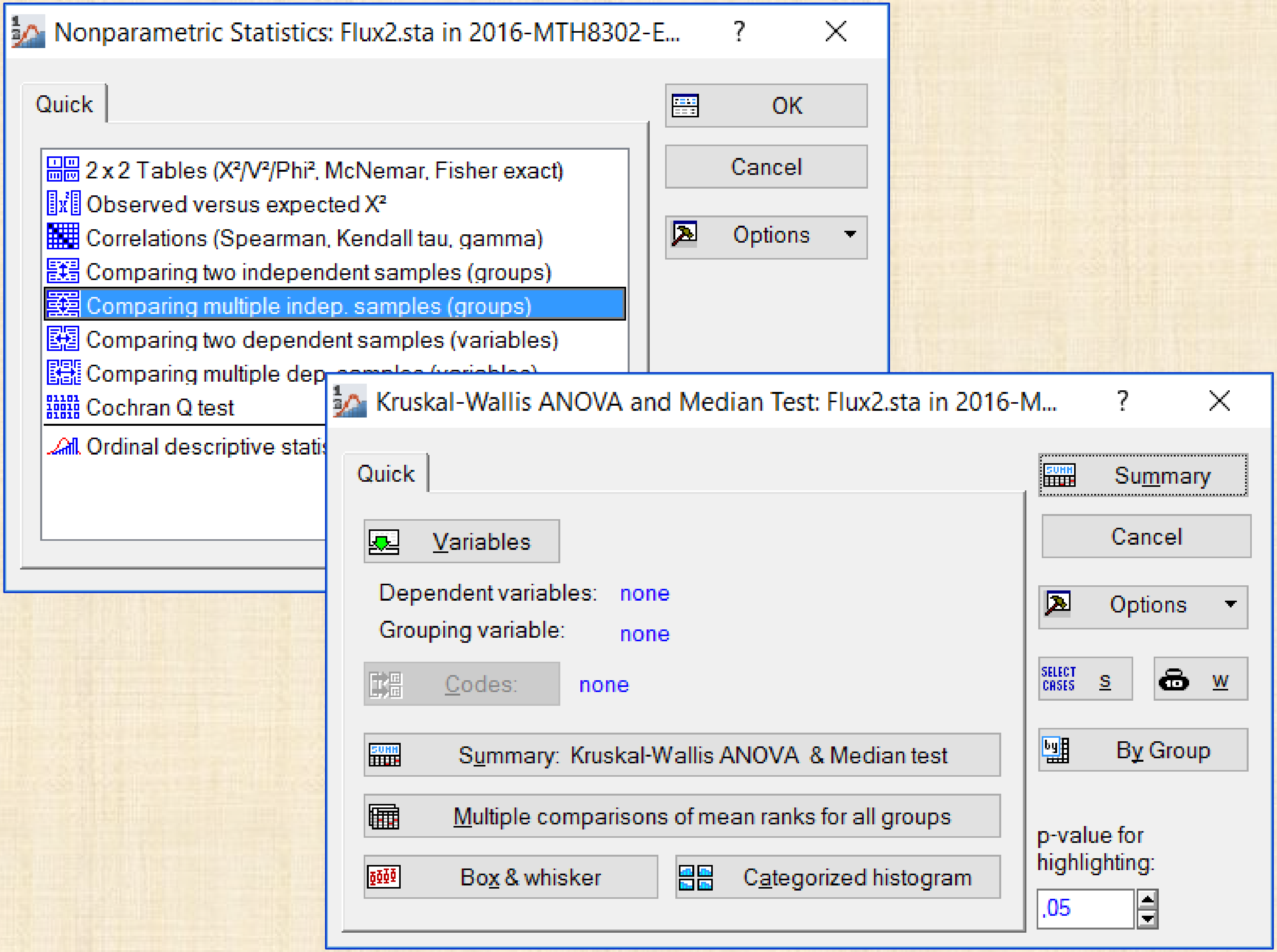
$$MSE = \sum\sum (R_{ij} - R_{i.})^2 / (N - g)$$

$$\overline{R}_{i.} = \sum R_{ij} \ / \ n_i$$

$$\overline{R}_{..} = \sum\sum R_{ij} / N = (N + 1) / 2$$

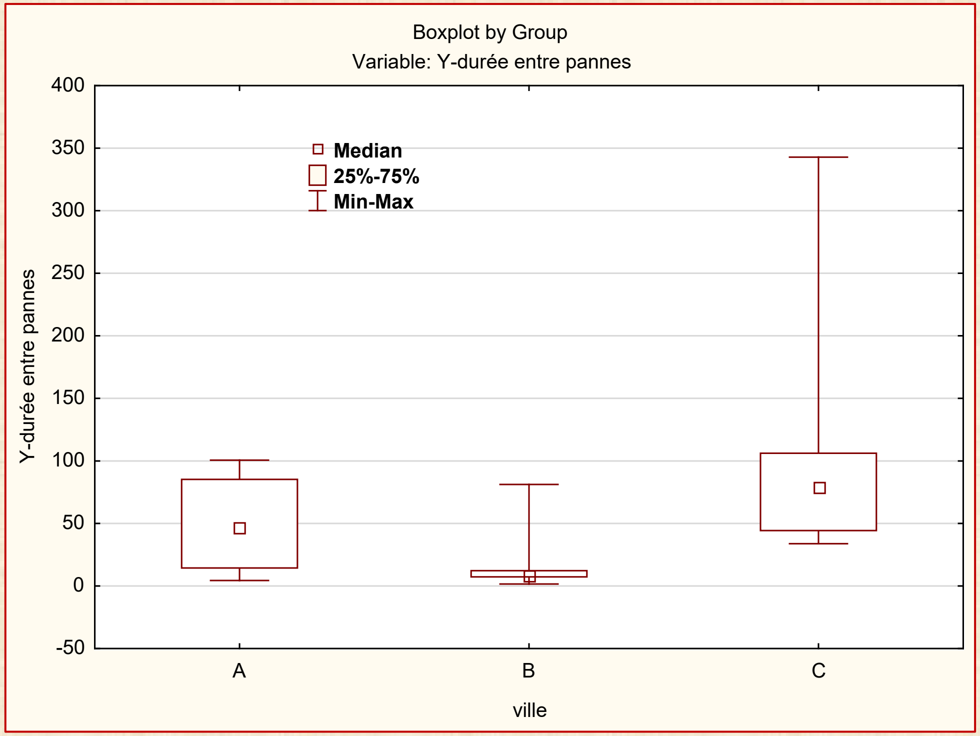
	Kutner et all 5 ed. - p. 791				
	1 ville	2 inter valle	3 Y-durée entre pannes	4 Y-Rang	5 logY
1	A	1	4,41	2	0,644
2	A	2	100,65	13	2,003
3	A	3	14,45	6	1,160
4	A	4	47,13	9	1,673
5	A	5	85,21	12	1,930
6	B	1	8,24	4	0,916
7	B	2	81,16	11	1,909
8	B	3	7,35	3	0,866
9	B	4	12,29	5	1,090
10	B	5	1,61	1	0,207
11	C	1	106,19	14	2,026
12	C	2	33,83	7	1,529
13	C	3	78,88	10	1,897
14	C	4	342,81	15	2,535
15	C	5	44,33	8	1,647

ANOVA non paramétrique : test de Kruskall-Wallis



ANOVA non paramétrique : test de Kruskall-Wallis

Test F basé sur les rangs Y-rang					
	SS	DF	MS	F	p
Intercept	960,00	1	960,00	61,02	0,0000
ville	91,20	2	45,60	2,90	0,0940
Error	188,80	12	15,73		



Kruskal-Wallis ANOVA by Ranks; Y-durée entre pannes
Independent (grouping) variable: ville
Kruskal-Wallis test: H (2, N= 15) = 4,560000 p =,1023

ville	Code	Valid N	Sum of Ranks	Mean Rank
A	1	5	42	8,4
B	2	5	24	4,8
C	3	5	54	10,8

pas de différence significative entre les villes

que faire avec des facteurs nuisibles ?

connu	contrôlable	méthode de contrôle
oui ……	oui …………	1. bloquer les essais plans en blocs – chapitre 7
oui …….	mesurable ….	2. analyse de covariance : enlever effet sur la réponse
non ……	non ……….	3. randomiser les essais

« block what you can and randomize what you cannot »

blocs : peuvent être définis par <u>plusieurs</u> facteurs nuisibles
par exemple, plans en carrés latins

Plans en blocs

- Facteur « nuisible » ou secondaire : a probablement un effet sur la réponse Y mais il ne présente aucun intérêt en soi

- Facteur nuisible typique : lots de matière première, opérateurs, pièces d'équipement, temps, unités expérimentales, …

- Stratégie : contrôler / minimiser l'impact des facteurs nuisibles
  blocage de essais

- Beaucoup d'expériences industrielles implique le blocage

- Absence de blocage
  - peut conduire l'expérience à un échec
  - manque de détection des effets des facteurs primaires

Exemple : 1 facteur principal + 1 facteur secondaire (=bloc et aléatoire)

Y mesure dureté Rockwell

	type matériau (bloc)			
pointe	**m1**	**m2**	**m3**	**m4**
type A	7.69	10.77	15.75	12.82
B	10.61	9.81	12.03	12.22
C	12.22	13.02	13.55	15.36
D	13.18	14.65	14.96	17.15

- **type pointe : facteur primaire**
- **comparaison des pointes: différentes?**
- **matériau = facteur secondaire**
  = facteur bloc = facteur aléatoire
- **éliminer influence facteur bloc avant de comparer pointes**

- **matériau représente un facteur secondaire (nuisible)**
- **blocage avec le type de matériau est nécessaire**
- **Plan en blocs complètement aléatoire**

  « Randomized Complete Block Design » = RCBD

  assignation des pointes (traitements) au hasard (aléatoire) à l'intérieur de chaque bloc **(restriction)**

  versus

  assignation sans restriction = complètement aléatoire

Exemple : 2 facteurs secondaires carré latin 4 x 4

étude de 4 additifs essence A – B – C – D = facteur principal

réponse : Y = consommation (m/g)

facteurs blocs : voiture v1–v2–v3-v4 conducteur : c1–c2–c3–c4

plan en carré latin 4 x 4

facteur principal / y

	voiture			
conducteur	v1	v2	v3	v4
c1	A / 21	B / 26	D / 20	C / 25
c2	D / 23	C / 26	A / 20	B / 27
c3	B / 15	D / 13	C / 16	A / 16
c4	C / 17	A / 15	B / 20	D / 20

seulement 16 des 4 x 4 x 4 = 64 combinaisons de Additif x Voiture x Cond

propriétés chaque ligne (conducteur) – 4 additifs présents

chaque colonne (voiture) – 4 additifs présents

Exemple : plans en blocs incomplets

- nombre k d'unités chaque bloc (taille du bloc) plus petit que nombre a de traitements du facteur : $k < a$ cause fréquente : manque de matériel expérimental

- Exemple

trait	bloc (matière première) 1	2	3	4
1	-	73	71	74
2	67	-	72	75
3	68	73	-	75
4	72	75	75	-

$a = 4$ traitements $b = 4$ blocs $k = 3$ unités par blocs

- N = nombre total d'observations = $b\,k = a\,r$

  a : nombre de traitements k : nombre d'unités chaque bloc
  b : nombre de blocs r : nombre répétitions traitement

- BIBD : Balanced Incomplete Bloc Design

  - chaque traitement apparaît le même nombre de fois
  - chaque paire de traitements apparaît le même nombre de fois

COMPARAISONS MULTIPLES

Si le test F est significatif : les moyennes sont statistiquement différentes

Peut-on dire plus? Sont-elles toutes différentes ?
Sinon, quelle moyenne diffère de quelle autre?
Peut-on faire des comparaisons (contrastes) entre des groupes de moyennes?

Toutes ces questions constituent l'analyse a posteriori (post-hoc) des moyennes. Elles font intervenir le problème de comparaisons multiples sur le même ensemble de données.
Il faut contrôler les risques associés à ces comparaisons multiples.

On veut contrôler le risque et avoir un coefficient de confiance global de $1 - \alpha$ sur l'ensemble des comparaisons (tests).

Si on fait un nombre de k comparaisons, chacune avec un coefficient de confiance de $1 - \alpha$, alors le coefficient de confiance global sur l'ensemble des k comparaisons diminue.

Plus on augmente le nombre de comparaisons (tests), plus on augmente les chances de conclure à tort.

Le tableau suivant illustre le problème.

COMPARAISONS MULTIPLES

nombre de modalités g	nombre de comparaisons $k = g*(g-1)/2$	coefficent de confiance global $(1 - \alpha)^k$	$1 - \alpha = 0,95$
2	1	$1 - \alpha$	0,95
3	3	$(1 - \alpha)^3$	0,86
4	6	$(1 - \alpha)^6$	0,735
5	10	$(1 - \alpha)^{10}$	0,60
6	15	$(1 - \alpha)^{15}$	0,46
8	28	$(1 - \alpha)^{28}$	0,24
10	45	$(1 - \alpha)^{45}$	0,10

catégories de tests

- tests (comparaisons) planifiés **AVANT** l'exécution des calculs
- tests suggérés **APRÈS** l'analyse

  post hoc **a posteriori** (« data snooping »)

COMPARAISONS MULTIPLES

A - intervalle de confiance pour une moyenne particulière

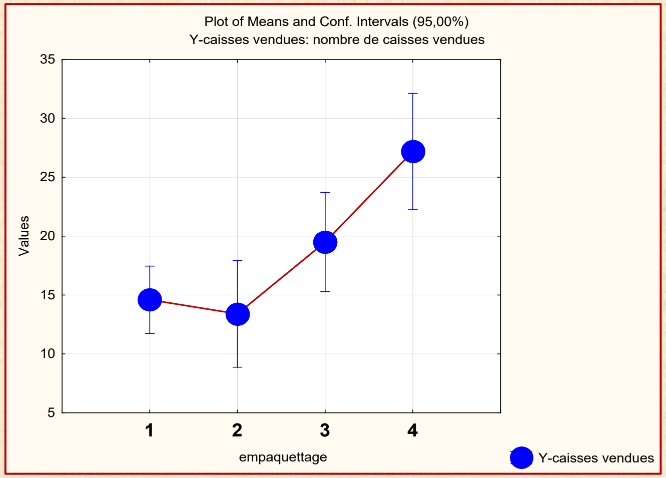
$$\mu_i : \overline{Y}_{i.} \pm t(1-\alpha/2, N-g) * MSE^{0.5} * [1 / n_i]^{0,5}$$

$t(1-\alpha/2, N-g)$: $(1-\alpha/2)^{\text{ième}}$ percentile loi T de Student avec (N - g) degrés de liberté $1-\alpha$: coefficient de confiance

	Kutner et all 5 ed. p.686		
	1 empaquetage	2 magasin	3 Y-caisses vendues
1	1	1	11
2	1	2	17
3	1	3	16
4	1	4	14
5	1	5	15
6	2	1	12
7	2	2	10
8	2	3	15
9	2	4	19
10	2	5	11
11	3	1	23
12	3	2	20
13	3	3	18
14	3	4	17
15	4	1	27
16	4	2	33
17	4	3	22
18	4	4	26
19	4	5	28

Y-caisses vendues

Emp	Means	N	SD
1	14,6	5	2,30
2	13,4	5	3,65
3	19,5	4	2,65
4	27,2	5	3,96
All	18,63	19	6,44



exemple

$\overline{Y}_1. = 14,6$ MSE = 10,55 $1-\alpha = 0.95$ $t(0.95, 15) = 2,13$

$$11,5 \leq \mu_1 \leq 17,7$$

COMPARAISONS MULTIPLES

B - intervalle de confiance pour la différence entre 2 moyennes

$$\mu_i - \mu_{i'} : (\bar{Y}_{i.} - \bar{Y}_{i'.}) \pm t(1 - \alpha/2, N - g) * MSE^{0.5} * [(1/n_i) + (1/n_{i'})]^{0.5}$$

exemple $\mu_3 - \mu_4 : (19{,}5 - 27{,}2) \pm 2{,}13 * 10{,}55^{0.5} * [(1/5) + (1/4)]^{0.5}$

$$-12{,}3 \leq \mu_3 - \mu_4 \leq -3{,}7$$

C - contraste = comparaison

$L = \sum c_i \mu_i$ $\quad \sum c_i = 0$ $\quad \hat{L} = \sum c_i \bar{Y}_{i.}$ $\quad s(\hat{L}) = MSE^{0.5} * [\sum c_i^2 / n_i]^{0.5}$

exemple 1&2 vs 3&4 $\quad L = 0{,}5*(\mu_1 + \mu_2) - 0{,}5*(\mu_3 + \mu_4)$

$\hat{L} = -9{,}35$ $\quad \sum c_i^2 / n_i = 0{,}2125$ $\quad s(\hat{L}) = 2{,}24$

$$-12{,}5 \leq L \leq -6{,}2$$

Procédures d'inférences simultanées (comparaisons multiples)

Les méthodes A – B – C ont **deux limitations** :

- le coefficient de confiance 1 – α et le seuil α d'un test s'applique à **UN** test seulement.
- Le test ou la comparaison ne doit **pas être suggéré par les données** (« data snooping »).

La solution de ce problème est **d'utiliser une procédure de comparaison multiple qui inclut toutes les inférences possibles qui peuvent être anticipées et d'intérêt après que les données furent examinées.**

Par exemple, on peut s'intéresser à toutes les comparaisons définies par les **différences entre toutes les paires de moyennes**.

Il existe **3 procédures** pour faire de l'inférence **après avoir vu** les données et en contrôlant le coefficient de confiance:

- méthode de Tukey («HSD = Honest Significant Differences »)
- méthode de Scheffé pour les contrastes
- méthode de Bonferronni pour les comparaisons prédéfinies

distribution « Studentized Range »

dédiée sur les comparaisons (contrastes spécifiques) définies
par les différences entre toutes les moyennes prises 2 à 2

basée sur distribution « Studentized Range »

$Y_1, Y_2, \ldots, Y_g$: g observations indépendantes d'une population $N(\mu , \sigma^2)$

$W = \max(Y_1, Y_2,..,Y_g) - \min(Y_1, Y_2,..,Y_g)$: étendue ("range")

S^2 : estimation de σ^2 basée sur υ degrés de liberté

$Q(g, \upsilon) = W / S$ est le « studentized range »

valeurs $q(0{,}95; g; \upsilon)$ extrait table Kutner et all 5 ed. p. 1334

υ	g = 2	g = 3	g = 4	g = 5	g = 10	g = 15	g = 20
2	6,08	8.33	9.80	10.9	14.0	15.7	16.8
5	3,64	4.60	5.22	5.67	6.80	7.72	8.21
10	3,15	3.88	4.33	4.65	5.60	6.11	6.47
20	2,95	3.58	3.96	4.23	5.01	5.43	5.71
40	2,86	3.44	3.79	4.04	4.73	5.11	5.36
60	2,83	3.40	3.74	3.98	4.65	5.00	5.24
120	2,80	3.36	3.68	3.92	4.56	4.90	5.13
infini	2,77	3.31	3.63	3.86	4.47	4.80	5.01

COMPARAISONS MULTIPLES

méthode de Tukey

$D = \mu_i - \mu_{i'}$ $\hat{D} = \bar{Y}_{i.} - \bar{Y}_{i'.}$

$s^2(\hat{D}) = MSE * [(1/n_i) + (1/n_{i'})]$

$D : \hat{D} \pm 0.707 * q(1 - \alpha; g, N - g) * s(\hat{D})$

Tukey HSD test; variable Y-caisses vendues probabilities for Post Hoc Tests Error: Between MS = 10.547, df = 15				
Empaquetage	{1} 14.6	{2} 13.4	{3} 19.5	{4} 27.2
1		0.9354	0.1550	0.0003
2	0.9354		0.0584	0.0002
3	0.1550	0.0584		0.0143
4	0.0003	0.0002	0.0143	

Méthode de Scheffé

$L = \sum c_i \, \mu_i \quad \sum c_i = 0 \qquad \hat{L} = \sum c_i \, Y_{i.} \qquad s(\hat{L}) = MSE^{0.5} * [\sum c_i^2 / n_i]^{0.5}$

$L : \hat{L} \pm (g-1) * F(1-\alpha, g-1, N-g) * s(\hat{L})$

empaquettage		{1} M=14,600	{2} M=13,400	{3} M=19,500	{4} M=27,200
Scheffe Test; Variable: Y-caisses vendues Marked differences are significant at p < ,05000					
1	{1}		0,950675	0,212530	0,000229
2	{2}	0,950675		0,089489	0,000086
3	{3}	0,212530	0,089489		0,024782
4	{4}	0,000229	0,000086	0,024782	

Méthode de Bonferroni $\quad L : \hat{L} \pm t(1-(\alpha/2g), N-g) * s(\hat{L})$

Bonferroni test; variable Y-caisses vendues
Probabilities for Post Hoc Tests
Error: Between MS = 10.547, df = 15.000

Cell No.	empaquetage	{1} 14.600	{2} 13.400	{3} 19.500	{4} 27.200
1	1		1,0000	0,2397	0,0001
2	2	1,0000		0,0808	0,0000
3	3	0,2397	0,0808		0,0180
4	4	0,0001	0,0000	0,0180	

Scheffé et Bonferroni donne même conlusion et aussi à celle de Tukey

recommandations sur les méthodes de comparaison multiples

- Si on veut seulement faire des comparaisons entre les paires, la procédure de Tukey est supérieure et elle est recommandée.

- Si le test F rejette l'égalité des moyennes alors il existe au moins un contraste qui diffère de zéro parmi tous les contrastes.

- La procédure de Bonferroni est préférable à la procédure de Scheffé si le nombre de contrastes d'intérêt est à peu près le même que le nombre de modalités.

- Il existe d'autres procédures pour des fonctions spécialisées. Par exemple, la procédure de Dunnett pour comparer chaque traitement vis-à-vis un contrôle ;

- Procédure de Hsu : pour choisir choisir le « meilleur » traitement.

ANOM : Analyis Of Means (Ott)

ANOM : Analyis Of Means (Ott)

méthode alternative au test F.   Basée sur l'ensemble des tests de l'effet différentiel de chaque modalité.

avantage : représentation graphique semblable à une carte contrôle

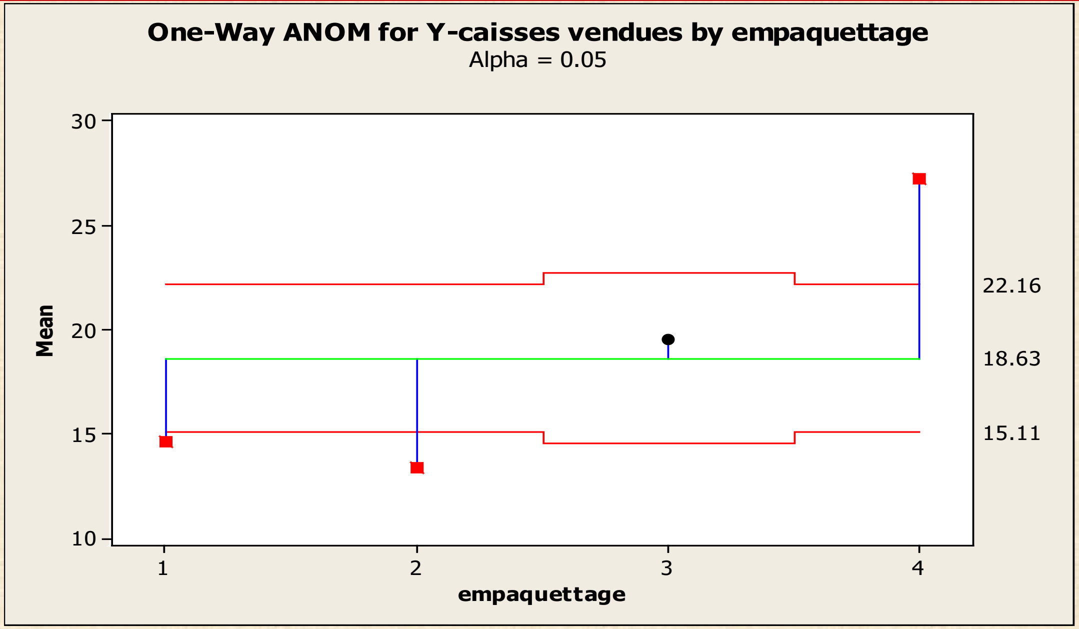
$$\hat{\tau}_i = \bar{Y}_{i.} - \bar{Y}_{..}$$

$$s^2(\hat{\tau}_i) = MSE\ [\ ((g-1)/g)^2\ (1/n_i) + (1/g^2)\ (\sum_{h \neq i} (1/n_h))\ )$$

ANOM :   test si le moyennes diffèrent de la moyenne globale

ANOVA : test si les moyennes sont différentes

avec Minitab



DIAGNOSTICS ET MESURES CORRECTIVES

- **Diagnostic : écarts importants par rapport aux hypothèses de base**
- **Si oui, qu'elles sont les mesures correctives?**

Analyse diagnostique : basée sur les résidus et des graphiques (comme en régression)

<u>4 types de résidus</u>

$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \bar{Y}_{i.}$ — **résidu brut**

$e_{ij}^* = e_{ij} / MSE^{0.5}$ — **résidu semi studentisé**

$r_{ij} = e_{ij}^* / [(n_i - 1) / n_i]^{0.5}$ — **résidu studentisé**

$t_{ij} = e_{ij}^* [(N - g - 1) / (SSE [1 - (1/n_i)] - e_{ij}^2)]^{0.5}$ — **résidu studentisé avec observation supprimée**

Écarts du modèle d'ANOVA en ordre <u>d'importance décroissante</u>

1. variance non constante
2. erreurs (observations) non indépendantes
3. présence de valeurs aberrantes
4. normalité du terme d'erreur
5. omission de variables explicatives importantes

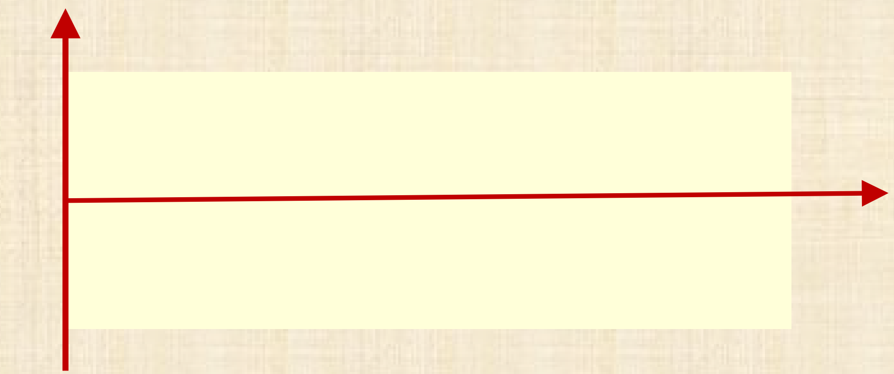
DIAGNOSTICS ET MESURES CORRECTIVES

HYPOTHÈSES	VÉRIFICATION	DIAGNOSTIC
variance non constante	graphique de résidus studentisés VS valeurs prédites	- bande horizontale - tests : Hartley Brown-Forsythe
non indépendance	si l’ordre temporel est connu	- résidus VS temps - test d’indépendance sérielle
valeurs aberrantes	t_{ij} VS valeurs prédites	-
normalité	résidus sur échelle de probabilité gaussienne	écart par rapport à la droite-
omisssion	résidus VS valeurs prédites	résidus corrélés avec autres facteurs non tenu en compte

graphique des résidus (axe vertical) VS : Y prédits, Y observés, X
critère recherché = bande horizontale

Résidu

0

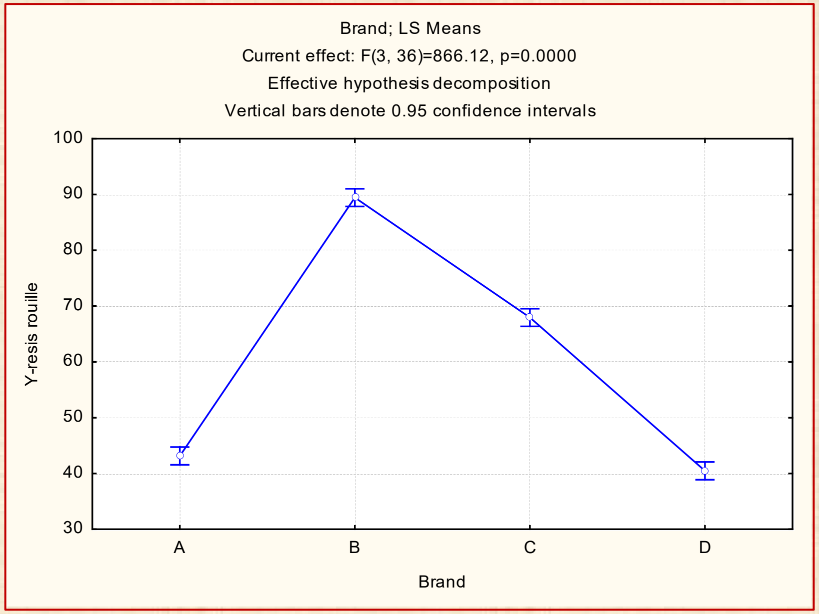


Y observé
Y prédit
chaque X

Exemple : rouille

	Kutner et all - 5 ed p.735		
	1 Brand	2 rep	3 Y-resis rouille
1	A	1	43,9
2	A	2	39,0
3	A	3	46,7
4	A	4	43,8
5	A	5	44,2
6	A	6	47,7
7	A	7	43,6
8	A	8	38,9
9	A	9	43,6
10	A	10	40,0
11	B	1	89,8
12	B	2	87,1
13	B	3	92,7
14	B	4	90,6
15	B	5	87,7
16	B	6	92,4
17	B	7	86,1
18	B	8	88,1
19	B	9	90,8
20	B	10	89,1

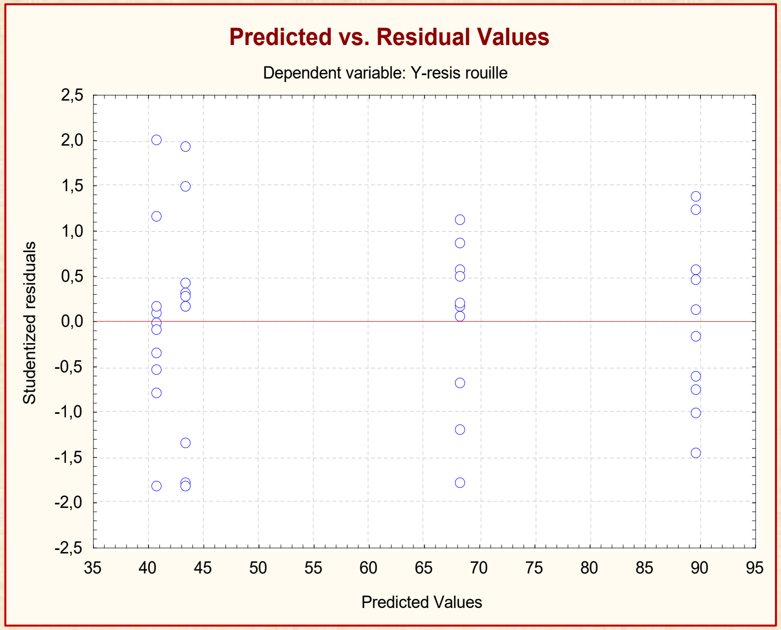
	Kutner et all - 5 ed p.735		
	1 Brand	2 rep	3 Y-resis rouille
21	C	1	68,4
22	C	2	69,3
23	C	3	68,5
24	C	4	66,4
25	C	5	70,0
26	C	6	68,1
27	C	7	70,6
28	C	8	65,2
29	C	9	63,8
30	C	10	69,2
31	D	1	36,2
32	D	2	45,2
33	D	3	40,7
34	D	4	40,5
35	D	5	39,3
36	D	6	40,3
37	D	7	43,2
38	D	8	38,7
39	D	9	40,9
40	D	10	39,7

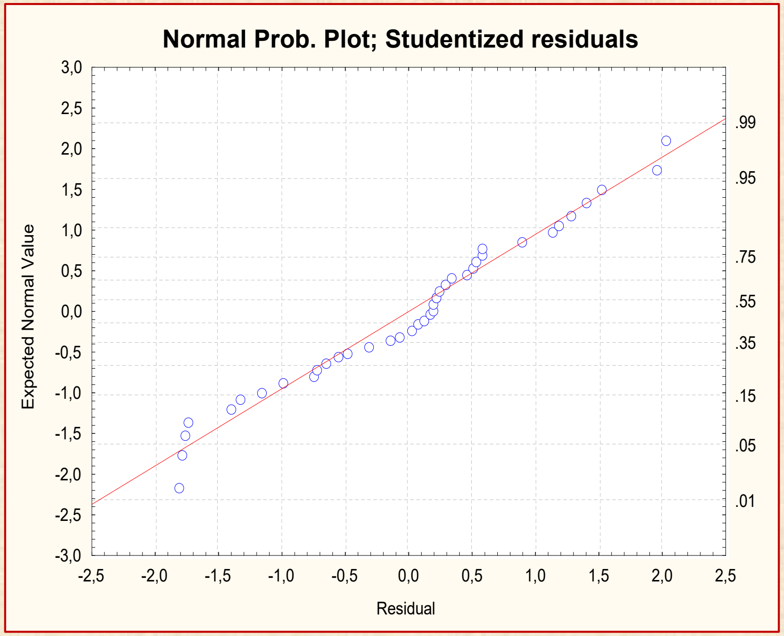


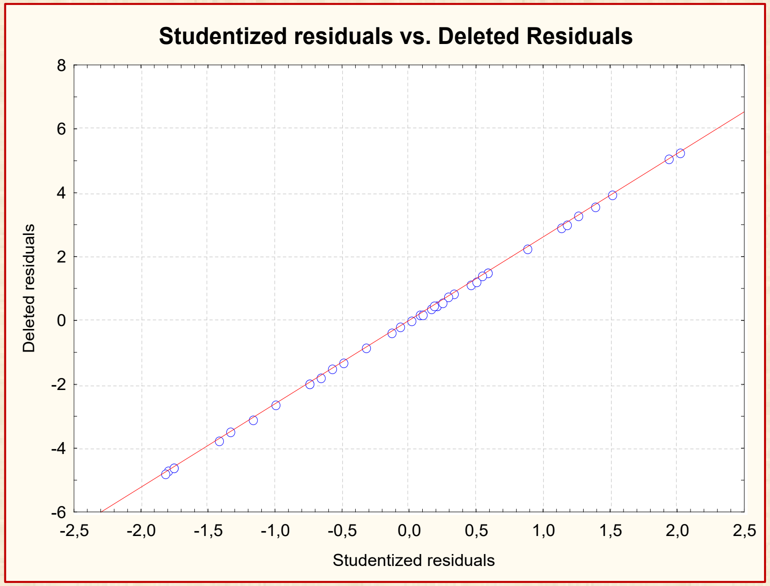
Effect	Univariate Results for Each DV (Rouille.sta) Sigma-restricted parameterization Effective hypothesis decomposition Degr. of Freedom	Y-resis rouille SS	Y-resis rouille MS	Y-resis rouille F	Y-resis rouille p
Intercept	1	145202,5	145202,5	23649,26	0,00
Brand	3	15953,5	5317,8	866,12	0,00
Error	36	221,0	6,1		
Total	39	16174,5			

DIAGNOSTICS ET MESURES CORRECTIVES

Exemple : rouille







DIAGNOSTICS ET MESURES CORRECTIVES

Tests : homogénéité de la variance

Hartley, Bartlett, Cochran, Brown-Forsythe, Levene

Test de Hartley exigence : $n_i = n$ + normalité

H_0 : $\sigma_1^2 = \sigma_2^2 = ... = \sigma_g^2$

H_a : les variances ne sont pas toutes égales

Hartley : $H^* = \max(s_i^2) / \min(s_i^2)$

Rejet de H_0 si $H > H(1-\alpha, g, n-1)$

$H(1-\alpha, g, df)$: $(1-\alpha)$ percentile distribution de Hartley

Exemple : rouille

marque	# obs	moyenne	écart type	variance
tous	40	60.25	20.36	414.53
A	10	43.14	3.00	9.00
B	10	89.44	2.22	4.93
C	10	67.95	2.17	4.71
D	10	40.47	2.44	5.95

$H^* = 9{,}00 / 4{,}17 = 1{,}91$

on ne rejette pas H_0 car p = 0,7532

DIAGNOSTICS ET MESURES CORRECTIVES

Test de Brown-Forsythe

n_i peuvent être inégaux
test robuste à la non normalité

$d_{ij} = | y_{ij} - \text{med}(y_{ij}) |$ med (y) = médiane (y)

$FBF = MSTR / MSE$

$MSTR = \sum n_i * (\bar{d}_{i.} - \bar{d}_{..})^2 / (g - 1)$

$MSE = \sum\sum(d_{ij} - \bar{d}_{i.})^2 / (N - g)$

$\bar{d}_{i.} = \sum d_{ij} / n_i$ $\bar{d}_{..} = \sum\sum d_{ij} / N$

FBF suit approximativement loi $F(g - 1, N - g)$

rejet H_0 si $FBF > F(1- \alpha, g - 1, N - g)$

Test de Levene $d_{ij} = |y_{ij} - \text{moy}(y_{ij})|$

moy (y) = moyenne (y)

test de Brown-Forsythe : modification test de Levene

DIAGNOSTICS ET MESURES CORRECTIVES

Test de Cochran $n_i = n$

$$C = \max(s_i^2) / \sum s_i^2$$

loi d'échantillonange de C dépend de g et de n

rejet H_0 si $C > C(1 - \alpha ; n, g)$

tableau des percentiles de la distribution de C : $C(1 - \alpha ; n, g)$
Statistical Principles in Experimental Design, 2 ed.,
B.J. Winer, 1971, Mc Graww-Hill, p. 876

n	percentile 1 - α	g = 2	g = 3	g = 4	g = 5	g = 8	g = 10
2	0.95	0.9985	0.9669	0.9065	0.8412	0.6798	0.6020
	0.99	0.9999	0.9933	0.9676	0.9279	0.7945	0.7175
3	0.95	0.9750	0.8709	0.7679	0.6838	0.5157	0.4450
	0.99	0.9950	0.9423	0.8643	0.7885	0.6152	0.5358
4	0.95	0.9392	0.7977	0.6841	0.5981	0.4377	0.3733
	0.99	0.9794	0.8831	0.7814	0.6957	0.5209	0.4469
5	0.95	0.9057	0.7457	0.6287	0.5441	0.3910	0.3311
	0.99	0.9586	0.8335	0.7212	0.6329	0.4627	0.3934
6	0.95	0.8772	0.7071	0.5895	0.5065	0.3595	0.3029
	0.99	0.9373	0.7933	0.6761	0.5875	0.4226	0.3572
8	0.95	0.8332	0.6530	0.5365	0.4564	0.3185	0.2666
	0.99	0.8988	0.7335	0.6129	0.5229	0.3704	0.3106
10	0.95	0.8010	0.6167	0.5017	0.4387	0.2926	0.2439
	0.99	0.8674	0.6912	0.5702	0.5037	0.3373	0.2813
17	0.95	0.7341	0.5466	0.4366	0.3645	0.2462	0.2032
	0.99	0.7949	0.6059	0.4884	0.4094	0.2779	0.2297
37	0.95	0.6602	0.4748	0.3720	0.3066	0.2022	0.1655
	0.99	0.7067	0.5153	0.4057	0.3351	0.2214	0.1811

DIAGNOSTICS ET MESURES CORRECTIVES

Test de Bartlett n_i peuvent être inégaux

$$c = 1 + (1/3*(g - 1))*[\sum (1/(n_i-1) - (1/N)]$$

$$B = (2.303/c)*[(N - g)*\log(MSE) - \sum (n_i - 1)*\log(s_i^2)]$$

B suit approximativement loi khi-deux

avec (g – 1) degrés de liberté

rejet H_0 si $B > \chi^2(1 - \alpha ; g - 1)$

Exemple

Hartley	Cochran	Bartlett	df	p
10.445	0.5865	12.985	4	0.0113

Test	SS Effect	df Effect	MS Effect	SS Error	Df Error	MS Error	F	p
Levene	8,69	4	2,17	24,8	35	0,71	3,07	0,029
Brown-Forsythe	9,35	4	2,34	27,9	35	0,80	2,94	0,034

tests concordent : variances inégales

DIAGNOSTICS ET MESURES CORRECTIVES

VARIANCES	NORMALITÉ	MESURE CORRECTIVE
hétérogènes	oui	régression pondérée
hétérogènes	non	transformation de Box-Cox
« gros » écarts	« gros » écarts	ANOVA non paramétrique Kruskall-Wallis

régression pondérée

$$Y_{ij} = \mu_i + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

modèle à moyennes de cellules

poids w $\quad w_{ij} = 1 / s_i^2$

on remplace le modèle d'ANOVA par un modèle de régression avec des variables indicatrices
+ on fait l'ajustement de moindres carrés pondérés
  avec les poids w

concept : applications pour les facteurs aléatoires

Exemple : soudure variances inégales

	Kuntner et all - 5 ed. p. 783		
	1 type flux	2 rep	3 Y-force soudure
1	A	1	14,87
2	A	2	16,81
3	A	3	15,83
4	A	4	15,47
5	A	5	13,60
6	A	6	14,76
7	A	7	17,40
8	A	8	14,62
9	B	1	18,43
10	B	2	18,76
11	B	3	20,12
12	B	4	19,11
13	B	5	19,81
14	B	6	18,43
15	B	7	17,16
16	B	8	16,40
17	C	1	16,95
18	C	2	12,28
19	C	3	12,00
20	C	4	13,18
21	C	5	14,99
22	C	6	15,76
23	C	7	19,35
24	C	8	15,52
25	D	1	8,59
26	D	2	10,90
27	D	3	8,60
28	D	4	10,13
29	D	5	10,28
30	D	6	9,98
31	D	7	9,41
32	D	8	10,04
33	E	1	11,55
34	E	2	13,36
35	E	3	13,64
36	E	4	12,16
37	E	5	11,62
38	E	6	12,39
39	E	7	12,05
40	E	8	11,95

groupe	i = 1	i = 2	i = 3	i = 4	i = 5
w_{ij}	0.653	0.637	0.162	1.449	1.689
s_i^2	1,531	1,570	6,185	0,667	0,592

	Kutner et all - 5 ed. p. 788								
	1 flux	2 rep	3 indA	4 indB	5 indC	6 indD	7 indE	8 poids	9 Y-soudure
1	A	1	1	0	0	0	0	0,653	14,87
2	A	2	1	0	0	0	0	0,653	16,81
3	A	3	1	0	0	0	0	0,653	15,83
4	A	4	1	0	0	0	0	0,653	15,47
5	A	5	1	0	0	0	0	0,653	13,60
6	A	6	1	0	0	0	0	0,653	14,76
7	A	7	1	0	0	0	0	0,653	17,40
8	A	8	1	0	0	0	0	0,653	14,62
9	B	1	0	1	0	0	0	0,637	18,43

	Kutner et all - 5 ed. p. 788								
	1 flux	2 rep	3 indA	4 indB	5 indC	6 indD	7 indE	8 poids	9 Y-soudure
31	D	7	0	0	0	1	0	1,499	9,41
32	D	8	0	0	0	1	0	1,499	10,04
33	E	1	0	0	0	0	1	1,689	11,55
34	E	2	0	0	0	0	1	1,689	13,36
35	E	3	0	0	0	0	1	1,689	13,64
36	E	4	0	0	0	0	1	1,689	12,16
37	E	5	0	0	0	0	1	1,689	11,62
38	E	6	0	0	0	0	1	1,689	12,39
39	E	7	0	0	0	0	1	1,689	12,05
40	E	8	0	0	0	0	1	1,689	11,95

Exemple : soudure + variances inégales

modèle complet (F) : $Y_{ij} = \mu_1 X_{ij1} + \mu_2 X_{ij2} + ... + \mu_g X_{ijg} + \varepsilon$

modèle réduit (R) : $Y_{ij} = \mu X_{ij1} + \mu X_{ij2} + ... + \mu X_{ijg} + \varepsilon$

$= \mu (X_{ij1} + X_{ij2} + ... + X_{ijg}) + \varepsilon$

$= \mu + \varepsilon$

avec l'hypothèse $H_0 : \mu_1 = \mu_2 = ... = \mu_g = \mu$

Test : $F_0 = [(SSE(R) - SSE(F)) / SSE(F)] * [(N - g) / (g - 1)]$ N=n*g

Modèle complet

Y = 15.4* indA + 18.5 * indB + 15.0 * indC +9.7 * indD + 12.3* indE

SSE(F) = 73,8 avec N - g = 40 - 5 = 35 degrés de liberté (ddl)

Modèle réduit

Y = 14.21 SSE(R) = 3,31*39 = 359,2 avec 39 degrés de liberté

Test $F_0 = (359,2 - 73,8) / 73,8 * (35 / 5) = 27,07$

$F_1 = F(0,99 ; 5, 35) = 3,59$ 99ième perc. dist. F avec (5 et 35) ddl

$F_0 > F_1$ donc rejet de H_0

conclusion : groupes sont de moyennes inégales

Transformations de la variable de réponse : cas de variances inégales

CONDITION	TRANSFORMATION
réponse Y est un comptage : distribution Poisson	$Y' = \sqrt{Y}$ ou $Y' = \sqrt{Y} + \sqrt{Y+1}$
réponse Y est une proportion : distribution binomiale	$Y' = 2 \arcsin(\sqrt{Y})$
(ecart type)^2 proportionnel à la moyenne (s^2/m)	$Y' = \sqrt{Y}$
écart type proportionnel à la moyenne (s/m)	$Y' = \log(Y)$
écart type proportionnel à la (moyenne)^2 (s/m^2)	$Y' = 1 / Y$

Recommandation 1

examiner les quantités $s_i^2 / Y_{i.}$ $s_i / Y_{i.}$ $s_i / Y_{i.}^2$
pour chaque niveau du facteur et choisir la transformation dont le coefficient de variation (CV) est le plus petit

Recommandation 2

transformation Box-Cox sur Y

$$Y' = Y^{\lambda} \qquad -2 < \lambda < 2$$

λ = ? on choisit λ tel que SSE(λ) soit minimum

on prend une valeur arrondie

si $\lambda = 0$ $Y' = \log(Y)$

Les écarts d'hypothèses de base sur le modèle statistique sont-ils importants ?

John Sall, Bradley Jones (2005).
Leptokurtosiphobia = peur irrationnelle de la non normalité
***Six Sigma Forum magazine*, vol 4, no 3, May 2005**

Conclusions

1. **Le manque de normalité n'est pas très important pour le cas de modèles à effets fixes.**

   Tester la normalité des résidus est une étape non nécessaire car

   - **pour de « grands échantillons » la non normalité est facile à détecter mais elle est sans conséquence**
   - **pour de « petits échantillons », la non normalité pourrait avoir des conséquences, mais la non normalité est quasiment impossible à détecter: aucun test est suffisamment puissant.**

   Pour le cas de modèles à effets aléatoires les conséquences sont plus importantes

Les écarts d'hypothèses de base sur le modèles statistiques sont-ils importants ?

conclusions

2. Le test F est robuste si les tailles n_i ne sont pas trop inégales.

3. Indépendance : conséquences importantes pour l'inférence.
   Par exemple, une forte autocorrélations dans les valeurs de la réponse Y a comme conséquence pratique que les tailles sont plus faibles en réalité qu'elles le paraissent, rendant ainsi plus difficile la détection des différences significatives.

4. Les mesures répétées (mesures longitudinales) sur une même unité d'observation constituent un cas fréquent de dépendance.
   Important de savoir reconnaître cette situation lorsqu'elle est présente dans la structure des données et de faire une analyse appropriée.
   Cette méthode importante sera vue plus loin.