# Chapitre 2 Régression linéaire simple logistique non linéaire


- Modèles 2-5
- Modèle linéaire simple 6-26
- Modèles linéarisables 24-27
- Modèles non linéaires 28-32
- Modèle Logistique 33-43
- Critères comparaison modèles 44-47
- Exemples avec JMP 48-52

# Objectif de ANALYSE STATISTIQUE : comprendre / prédire / optimiser SYSTÈME / PROCESSUS

**TOUTE analyse statistique repose sur un MODÈLE**

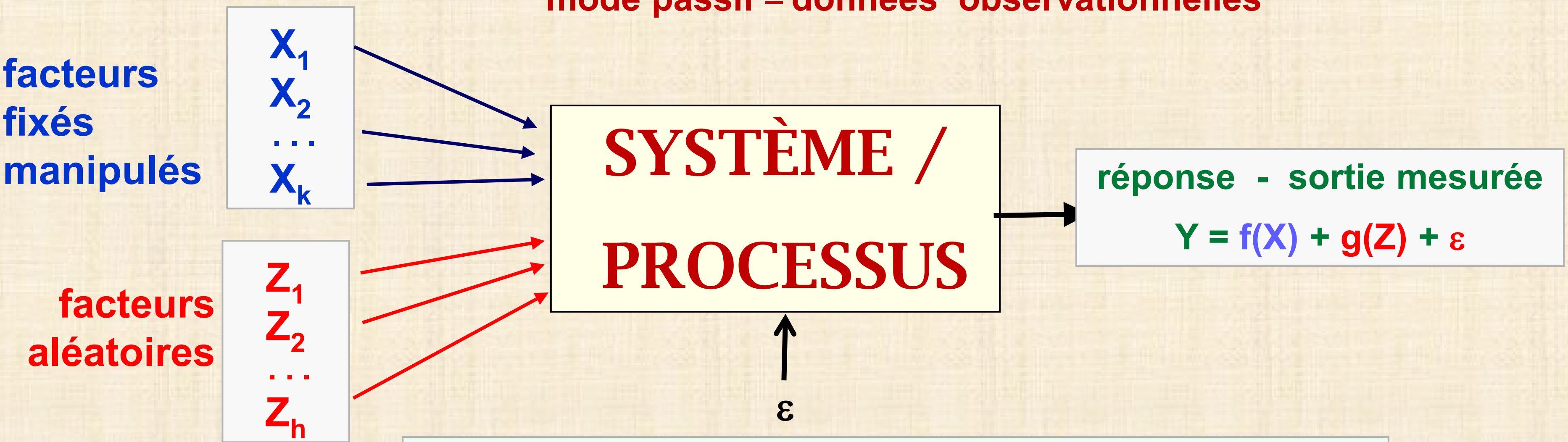
- relation entre input X (fixe) Z (aléatoire) et output Y
- connaissance de la structure des données:

  plan de collecte des données / nature variables / rôle des variables

  type d'influence des variables / unités statistiques (expérimentales)

$X_1, X_2, \ldots, Z_1, Z_2, \ldots$: facteurs, variables contrôlées en expérimentation
mode actif : données expérimentales

variables observées / mesurées en observation
mode passif – données observationnelles




erreur expérimentale = sources inconnues de variabilité

$\varepsilon \sim (0, \sigma^2)$

## Objectif de ANALYSE STATISTIQUE : comprendre / prédire / optimiser
## SYSTÈME / PROCESSUS

**VARIABLES**

**RÔLE**

Y : réponse , output

peut être: binaire (0, 1), multinomiale (classification), continue, multidimensionnelle

X, Z : explicatives, régresseurs, input, fixées, aléatoires

**NATURE**

X (fixes) : continues, catégoriques

Z (aléatoires) : continues, catégoriques

**INFLUENCE**

X : affecte la centralité (moyenne) de Y : effets fixes

Z : affecte la dispersion (variance) de Y : effets aléatoires

**MODÈLES à effets fixés seulement**

$$Y = f(X_1, X_2, \dots, X_k ; \beta_0, \beta_1, \beta_2, \dots) + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2)$$

**MODÈLES à effets mixtes : effets fixés + effets aléatires**

$$Y = f(X_1, X_2, \dots, X_k ; \beta_0, \beta_1, \beta_2, \dots) + g(Z_1, Z_2, \dots, Z_h ; \sigma_1^2, \sigma_2^2, \dots) + \varepsilon$$

## MODÈLES d'analyse de régression : $X_1$, $X_2$, …facteurs quantitatifs

- **Linéaire simple :** $Y = \beta_0 + \beta_1 X_1 + \varepsilon$
- **Régression multiple :** $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$
- **Effets principaux + effets d'interaction :**
  $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \ldots + \varepsilon$$
- **Quadratiques (response surface avec facteurs quantitatifs)**
  $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \ldots + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \ldots + \varepsilon$$
- **Régression logistique :** Y à valeurs catégoriques (oui / non)
- **Régression PLS:** nombre variables X > nombre d'observations ($k > n$)
- **Modèle d'analyse de la variance:** variables catégoriques A
  codage à effets $X = 0 / 1 / -1$ selon les modalités de A
- **Modèles avec 2 types de variables continues + catégoriques**
- **Mixtes:** facteurs modalités fixes + facteurs modalités aléatoires
- **Polynomial :** $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \bullet\bullet\bullet + \beta_k X^k + \varepsilon$

observations : $Y_i = \varphi(X_{i1}, X_{i2}, \ldots, X_{ik} ; \beta_0, \beta_1, \beta_2, \ldots) + \varepsilon_i \quad i = 1, 2, \ldots, n$

implicitement : une observation par unité (individu, sujet) statistique

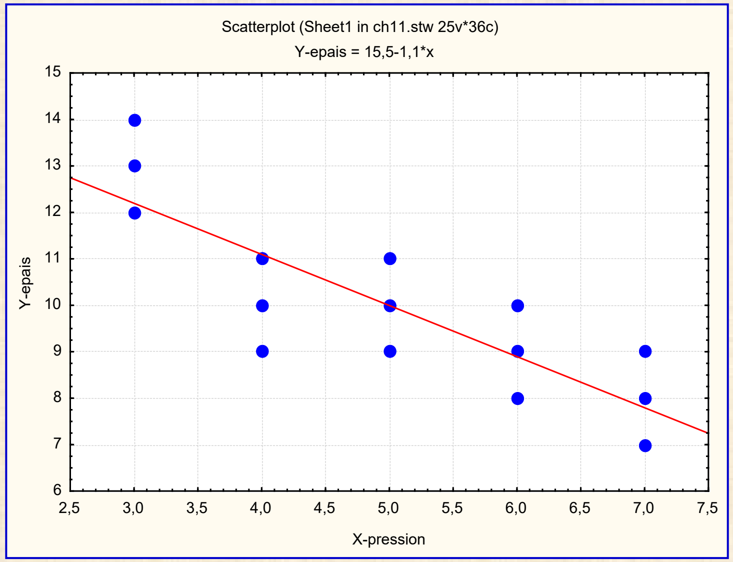
indépendance des $Y_i$ ?

# ANALYSE STATISTIQUE : étapes

1. Spécification d'un modèle statistique
2. Estimation des paramètres du modèle
3. Décomposition de la variabilité : ANOVA
4. Tests d'hypothèses sur les paramètres
5. Analyse diagnostique des résidus
   - vérification des hypothèses de base
   - identification d'observations influentes
   - transformation Y ? (si nécessaire)
6. Si nécessaire : itération des étapes 1 à 5
7. Optimisation de la réponse (s'il y a lieu)
8. Graphiques de la réponse

**Exemple : relation entre**
**Y : épaisseur substrat**
**X : pression**

| X-pression | Y-epais | ordre |
|---|---|---|
| 3 | 14 | 4 |
| 3 | 12 | 3 |
| 3 | 13 | 8 |
| 4 | 10 | 2 |
| 4 | 9 | 12 |
| 4 | 11 | 14 |
| 5 | 9 | 6 |
| 5 | 10 | 15 |
| 5 | 11 | 7 |
| 6 | 9 | 10 |
| 6 | 8 | 5 |
| 6 | 10 | 1 |
| 7 | 8 | 9 |
| 7 | 9 | 11 |
| 7 | 7 | 13 |




## Regression Linéaire Simple

$Y = \beta_0 + \beta_1 X_1 + \varepsilon$

**Modèle** $\quad Y = \beta_0 + \beta_1 x + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2)$

**Données** $\quad (x_i, y_i) \qquad i = 1, 2, \ldots, n$

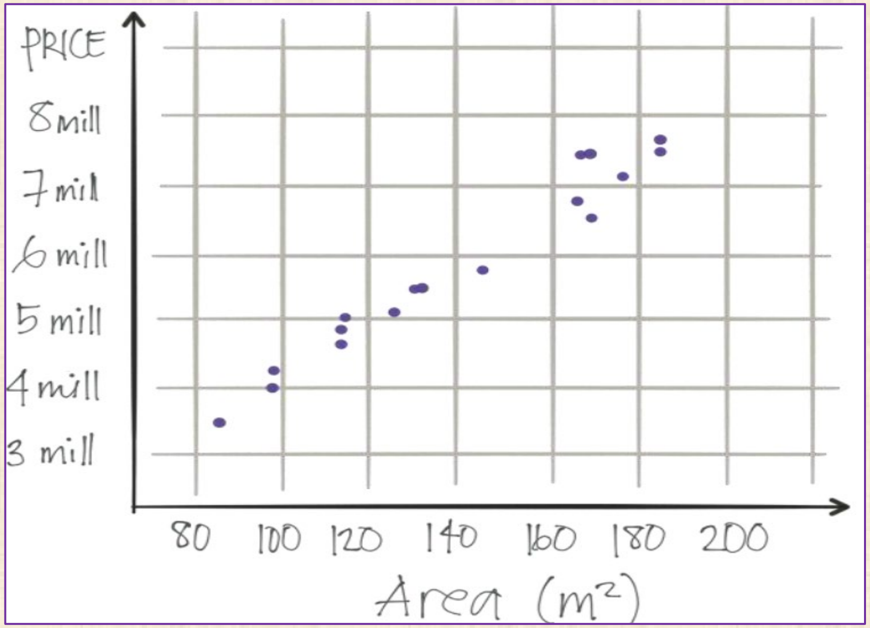
**Étapes**

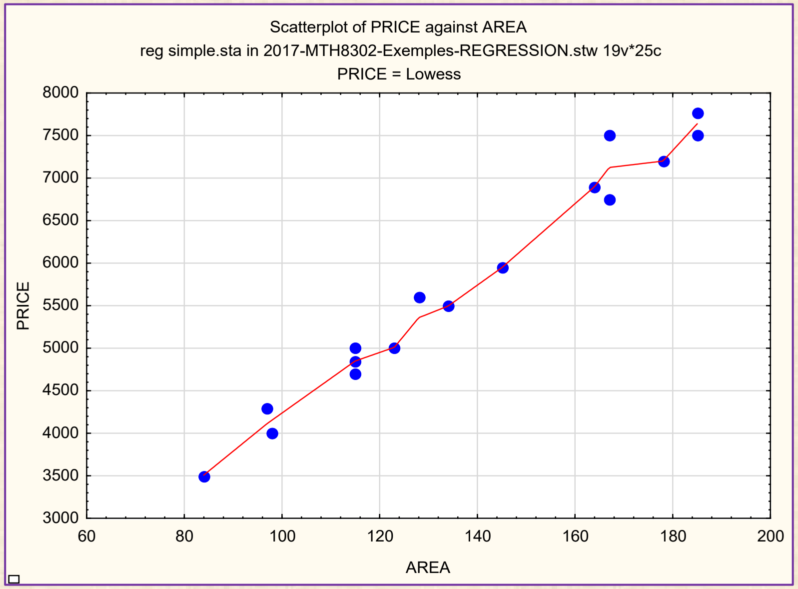
- estimation des paramètres : $\beta_0 \quad \beta_1 \quad \sigma^2$
- décomposition de la variabilité :
  = analyse de la variance
  = ANalysis Of VAriance = ANOVA
- tests - intervalles de confiance
- prédiction de Y et graphiques
- validation du modèle: analyse résidus

# Autre exemple : prix vs superficie

| AREA $m^2$ | PRICE Kkr |
|---|---|
| 134 | 5 495 |
| 115 | 4 700 |
| 167 | 7 500 |
| 185 | 7775 |
| 84 | 3500 |
| 98 | 40000 |
| 115 | 4 850 |
| 185 | 7500 |
| 164 | 6900 |
| 145 | 5950 |
| 123 | 5010 |
| 128 | 5600 |
| 167 | 6750 |
| 115 | 5000 |
| 178 | 7200 |
| 97 | 4290 |

| AREA | PRICE |
|---|---|
| 84 | 3500 |
| 97 | 4290 |
| 98 | 4000 |
| 115 | 4700 |
| 115 | 4850 |
| 115 | 5000 |
| 123 | 5010 |
| 128 | 5600 |
| 134 | 5495 |
| 145 | 5950 |
| 164 | 6900 |
| 167 | 7500 |
| 167 | 6750 |
| 178 | 7200 |
| 185 | 7775 |
| 185 | 7500 |



# Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

## NOTATION

$\bar{x} = \sum x_i / n$ : moyenne de X  $\bar{y} = \sum y_i / n$ : moyenne de Y

$SPxy = \sum (x_i - \bar{x})(y_i - \bar{y})$ : somme des produits XY

$SSxx = \sum (x_i - \bar{x})^2$ : somme des carrés de X

$SSyy = \sum (y_i - \bar{y})^2 = SStot$ = somme totale des carrés de Y

## ESTIMATION principe des moindres carrés: minimiser $S(\beta_0 , \beta_1)$

$S(\beta_0 , \beta_1) = \sum (y_i - \beta_0 - \beta_1 x_i)^2$ : écart par rapport à la droite

solution $b_1 = \hat{\beta}_1 = SPxy / SSxx = \sum c_i y_i$ où $c_i = (x_i - \bar{x}) / SSxx$

$b_0 = \hat{\beta}_0 = \bar{y} - b_1 \bar{x}$

$\sum c_i = 0 \quad \sum c_i^2 = 1$

prédiction $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = \bar{y} + b_1 (x - \bar{x})$ : droite de moindres carrés

résidu brut $e_i = \hat{y}_i - y_i$

somme des carrés résiduels $SSresid = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$

carré résiduel moyen $MSresid = SSresid / (n - 2)$

estimation de $\sigma^2$ $\hat{\sigma}^2 = MSresid$ $\hat{\sigma} = (MSresid)^{0.5}$

## Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

**DÉCOMPOSITION DE LA VARIABILITÉ : tableau d'analyse de la variance**

$$SSreg = \sum (\hat{y}_i - \bar{y})^2 = (SPxy)^2 / SSxx = \hat{\beta}_1^2 \, SSxx$$

= somme des carrés de régression (modèle, expliquée par X )

**ÉQUATION FONDAMENTALE DE DÉCOMPOSITION**

somme de carrés (SS) : SStot = SSreg + SSresid

variabilité : totale = modèle + résiduelle

degrés de liberté (DDL) : $n - 1 = 1 + (n - 2)$

**TABLEAU D'ANALYSE VARIANCE : modèle de régression linéaire simple**

| SOURCE | DDL | SS | MS = SS / DDL | F-ratio | p-valeur |
|---|---|---|---|---|---|
| régression | 1 | SSreg | MSreg = SSreg /1 | $F_0$ = MSreg/ MSresid | $P(F \geq F_0)$ |
| résiduelle | $n - 2$ | SSresid | MSresid = SSresid /( n - 2) $= \hat{\sigma}^2$ | ------ | ------ |
| totale | $n - 1$ | SStot | ------- | | |

$R^2$ = SSreg / SStot : coefficient de détermination

$0 \leq R^2 \leq 1$ : fraction variabilité Y expliquée par X

$r = \pm (R^2)^{0.5}$ : coefficient de corrélation linéaire entre Y et X

remarque : le signe de r ( + ou - ) sera celui de $\hat{\beta}_1$

## TEST et INTERVALLE de CONFIANCE

**résultat** $(\hat{\beta}_1 - \beta_1) / (\hat{\sigma} / SSxx^{0.5}) \sim T_{n-2}$ (loi de Student)

## applications

(a) test de $\beta_1$ $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

rejeter $H_0$ au seuil $\alpha$ si $|\hat{\beta}_1| (SSxx)^{0.5} / \hat{\sigma} > t_{n-2,\ 1-\alpha/2}$

remarque : le test est équivalent au test F du tableau ANOVA

(b) intervalle de confiance de $\beta_1$ : $\hat{\beta}_1 \pm t_{n-2,\ 1-\alpha/2}\ \hat{\sigma} / (SSxx)^{0.5}$

coefficient de confiance = $1 - \alpha$

(c) INTERVALLE de CONFIANCE : MOYENNE de Y à $X = x^*$

$E(Y | X = x^*)$ : $\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{n-2,\ 1-\alpha/2}\ \hat{\sigma}\ [ (1/n) + (( x^* - \bar{x} )^2 / SSxx) ]^{0.5}$

remarque : un intervalle de confiance pour $\beta_0$ s'obtient avec $x^* = 0$

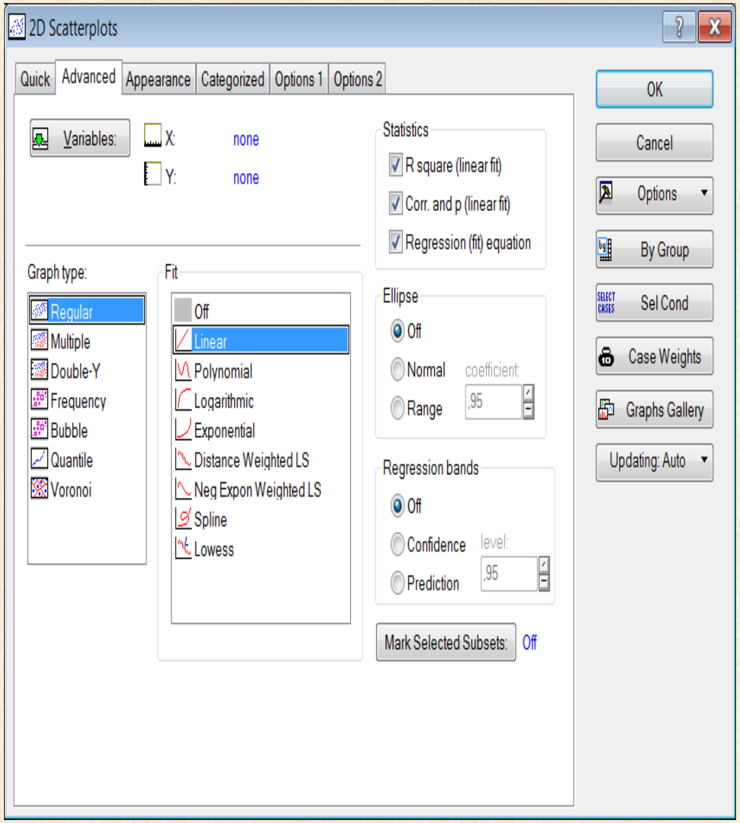
(d) INTERVALLE de PRÉDICTION : VALEUR de Y à $X = x^*$

$Y | X = x^*$ : $\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{n-2,\ 1-\alpha/2}\ \hat{\sigma}\ [ 1 + ( 1/n) + (( x^* - \bar{x} )^2 / SSxx)]^{0.5}$

remarque : ne pas confondre (c) et (d)

# Mise en œuvre avec *STATISTICA*

**avec GRAPHICS**
**… 2D Scatterplots**

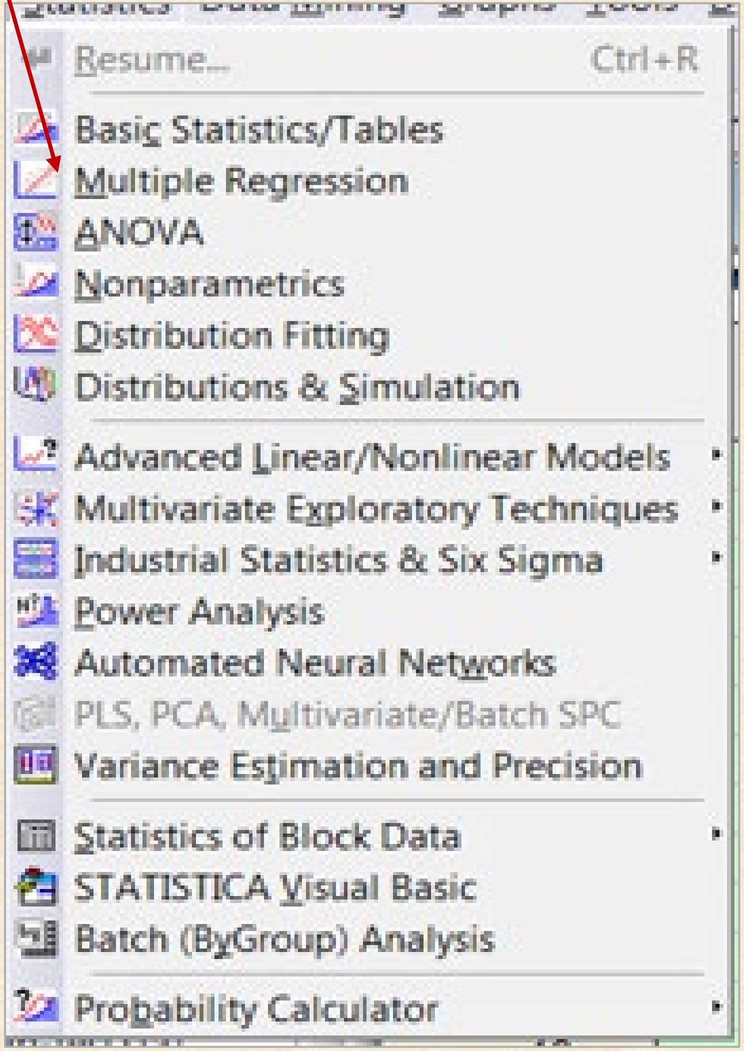



**pas d'analyse de la variance**

**pas d'analyse de résidus**

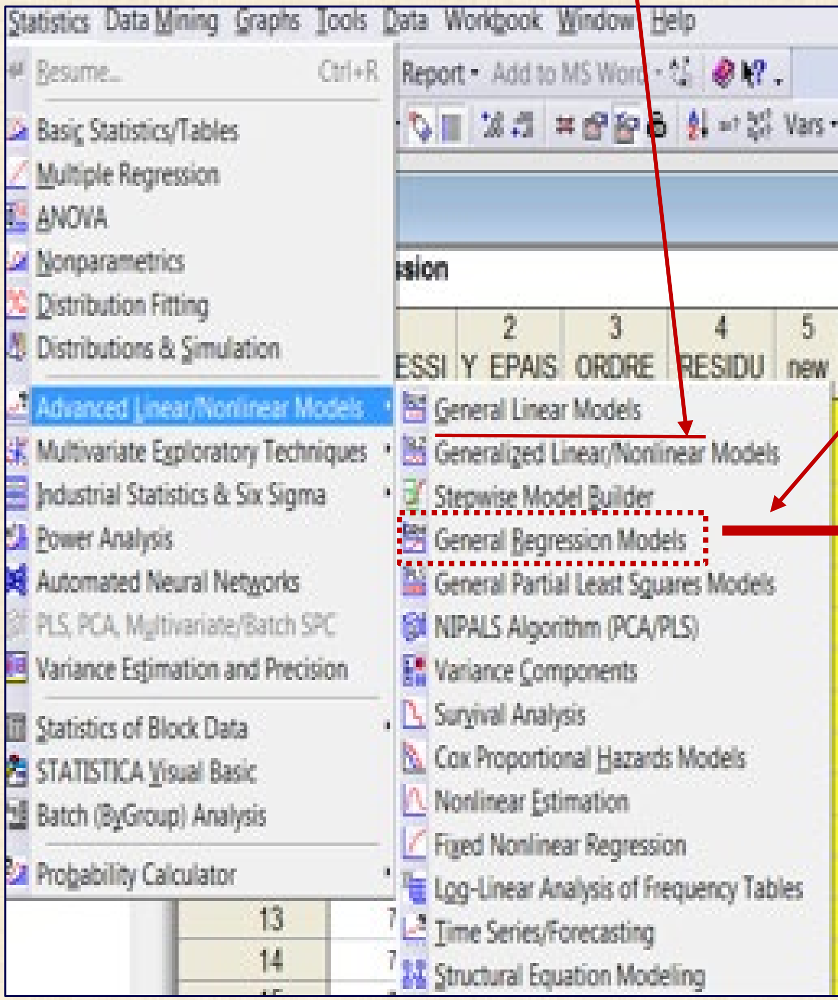
**avec STATISTICS…**
**… Multiple Linear Regression**

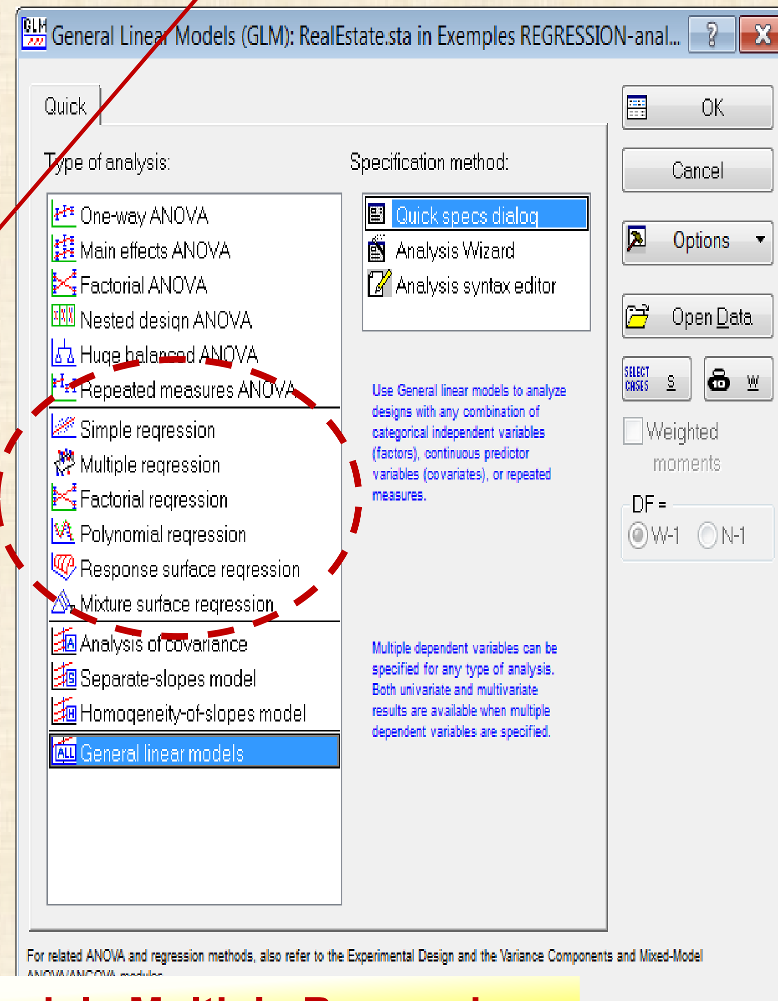
**variables continues seulement**



# Régression linéaire avec *STATISTICA*

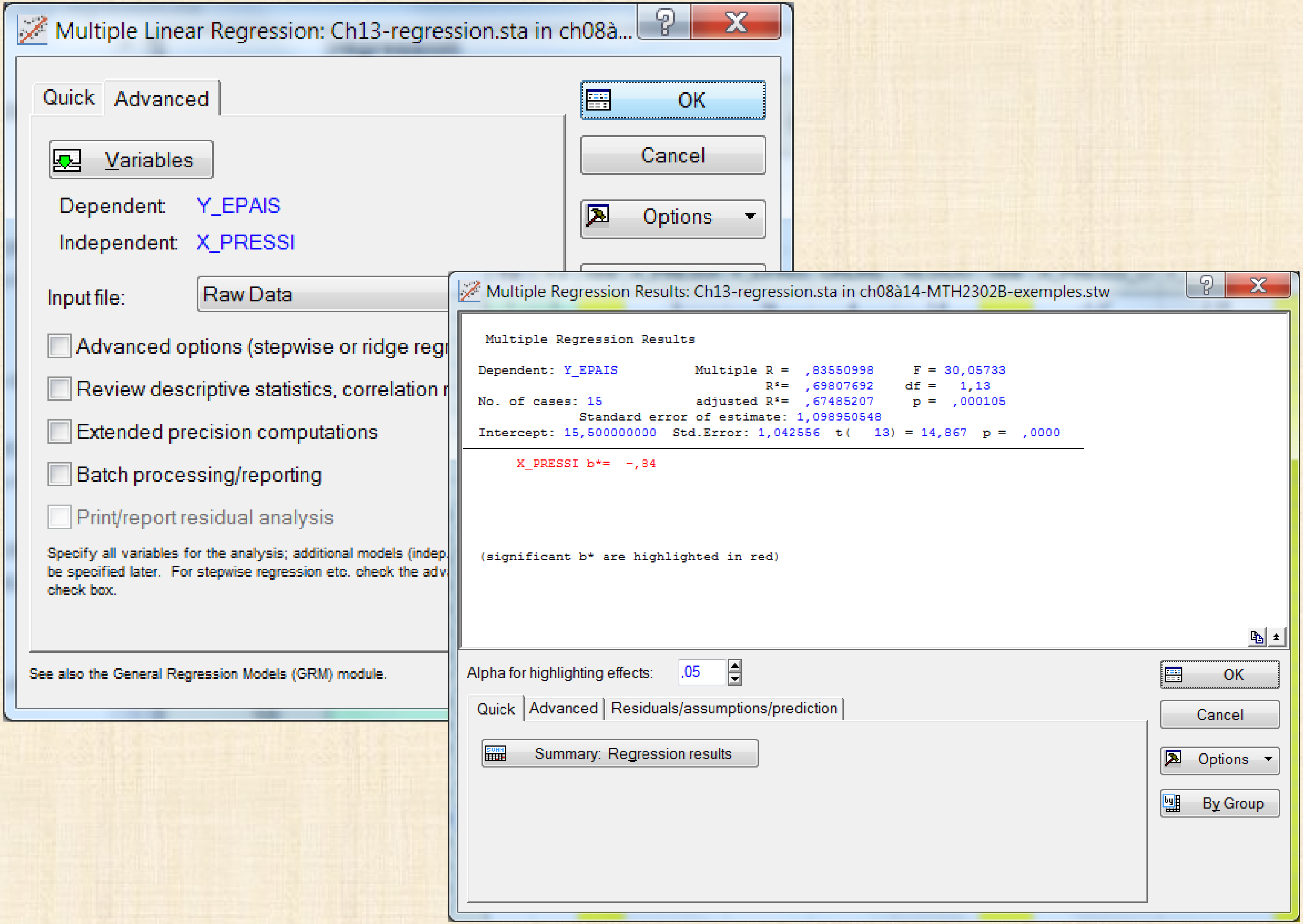
**Statistics … Advanced Linear/Nonlinear Models**
**… General Linear Models et General Regression Models**




**module GRM préférable module Multiple Regression**

# Mise en œuvre avec *STATISTICA*



## Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

## résultats

**Exemple** : épaisseur substrat

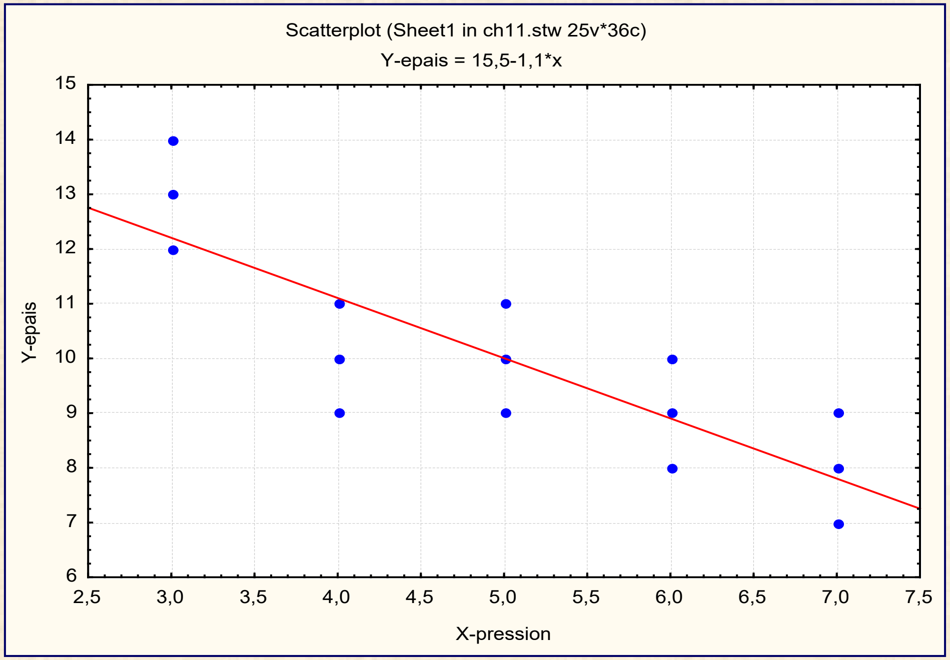
analyse avec le module

**MULTIPLE REGRESSION**

**de *STATISTICA***

$r = -0.8355$ $R^2 = 0.698$

$F(1,13)=30.057$ $p < 0.00011$




| | b* | Std. Err. | $b = \hat{\beta}$ | Std. Err. b | t(13) | p-level |
|---|---|---|---|---|---|---|
| Intercept | | | 15.500 | 1.0426 | 14.867 | 0.0000 |
| X-pression | - 0.8355 | 0.1524 | - 1.100 | 0.2006 | -5.482 | 0.0001 |

$b_0 = \hat{\beta}_0$

test de signification

$b_1 = \hat{\beta}_1$

coefficients
en variables
centrées - réduites

## Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

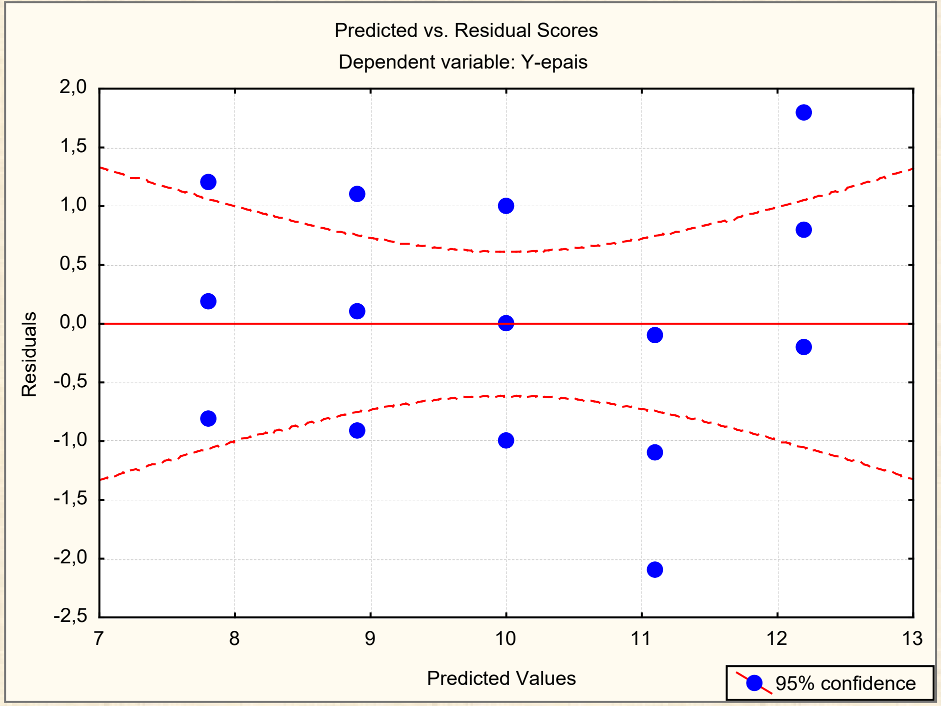
## résultats

**Exemple** : épaisseur substrat

| observés | prédits | résidus |
|---|---|---|
| 14 | 12.2 | 1.8 |
| 12 | 12.2 | -0.2 |
| 13 | 12.2 | 0.8 |
| 10 | 11.1 | -1.1 |
| 9 | 11.1 | -2.1 |
| 11 | 11.1 | -0.1 |
| 9 | 10.0 | -1.0 |
| 10 | 10.0 | 0.0 |
| 11 | 10.0 | 1.0 |
| 9 | 8.9 | 0.1 |
| 8 | 8.9 | -0.9 |
| 10 | 8.9 | 1.1 |
| 8 | 7.8 | 0.2 |
| 9 | 7.8 | 1.2 |
| 7 | 7.8 | -0.8 |

### Tableau d'analyse de la variance

| | SS | DDL | MS | F | p-valeur |
|---|---|---|---|---|---|
| Regress. | 36.3 | 1 | 36.3 | 30.06 | 0.00011 |
| Residuel | 15.7 | 13 | 1.21 | | |
| Total | 52 | 14 | | | |

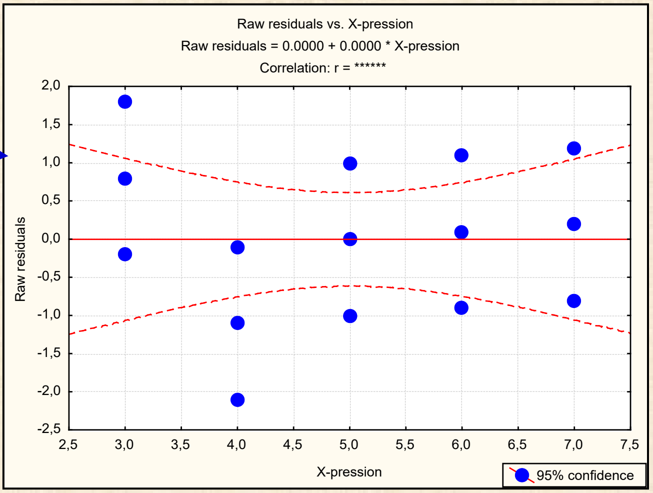


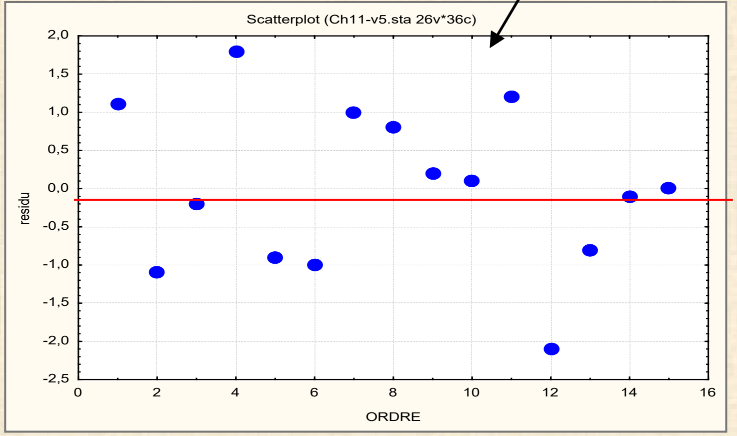
# Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

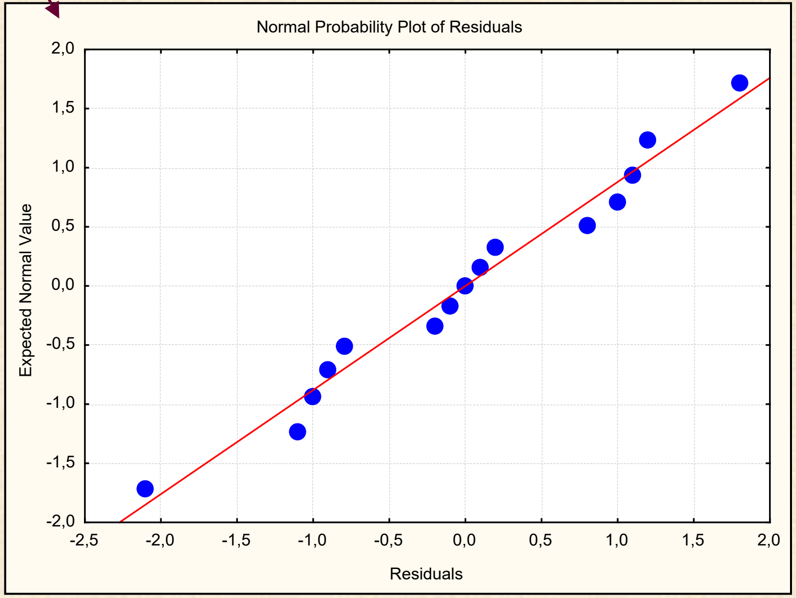
## résultats

### VALIDATION du MODÈLE : ANALYSE des RÉSIDUS

- résidus VS valeurs prédites $\hat{y}$ : variance constante ?
- résidus VS variable explicative X : variance constante ?
- résidus VS l'ordre des données : dépendance ?
- résidus sur échelle gaussienne : valeurs aberrantes?




**Résidus standardisés**

$z_i = e_i / \hat{\sigma}$ $\bar{z} = 0$ , écart type (z) = 1

**Résidus studentisés internes**

$t_i = e_i / \hat{\sigma} \left\{ [ (n-1)/n) + (x_i - \bar{x})^2 / SSxx ]^2 \right\}$

remarque : l'analyse des résidus peut se faire avec les

- résidus / résidus standardisés / résidus studentisés internes
- résidus studentisés externes ('deleted')
- préférablement : avec les résidus studentisés externes

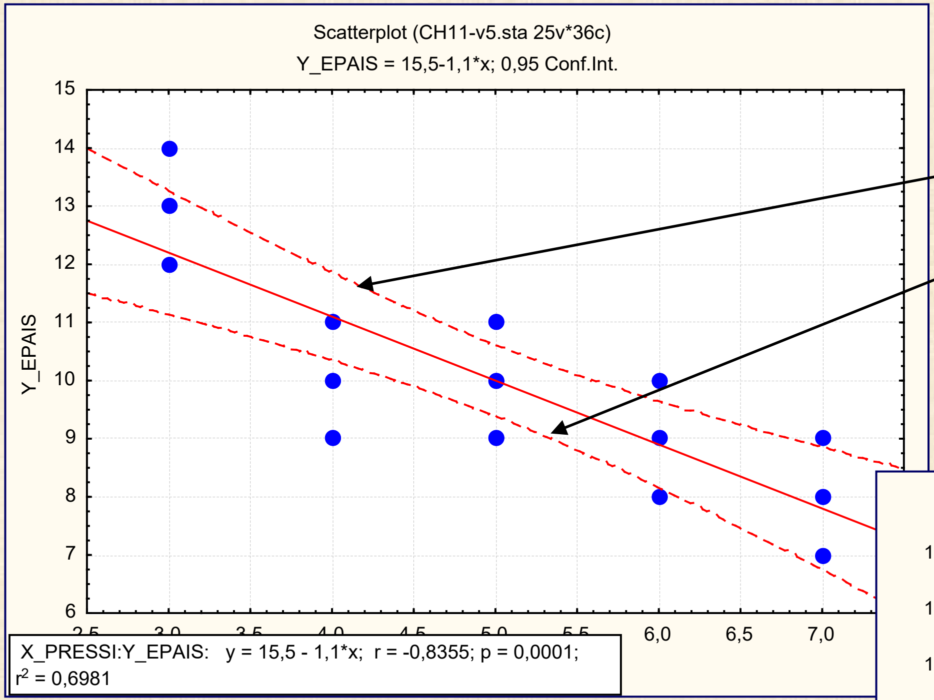
## Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

## résultats

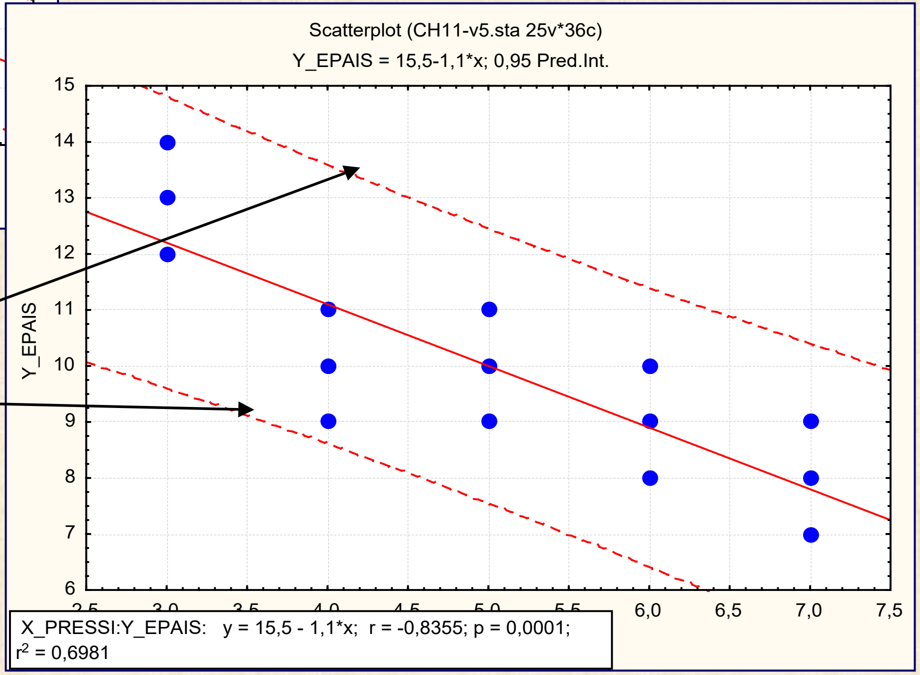
| | Y observé | Y prédit | résidu E | écart type Y prédit | Int. Prédiction à 95% lim inf Y | Int. Prédiction à 95% lim sup Y | Int. confiance à 95% lim inf Moy Y | Int. confiance à 95% lim suf Moy Y |
|---|---|---|---|---|---|---|---|---|
| 1 | 14 | 12.2 | 1.8 | 0.49 | 9.60 | 14.8 | 11.1 | 13.3 |
| 2 | 12 | 12.2 | -0.2 | 0.49 | 9.60 | 14.8 | 11.1 | 13.3 |
| 3 | 13 | 12.2 | 0.8 | 0.49 | 9.60 | 14.8 | 11.1 | 13.3 |
| 4 | 10 | 11.1 | -1.1 | 0.35 | 8.61 | 13.6 | 10.3 | 11.9 |
| 5 | 9 | 11.1 | -2.1 | 0.35 | 8.61 | 13.6 | 10.3 | 11.9 |
| 6 | 11 | 11.1 | -0.1 | 0.35 | 8.61 | 13.6 | 10.3 | 11.9 |
| 7 | 9 | 10.0 | -1.0 | 0.28 | 7.55 | 12.5 | 9.4 | 10.6 |
| 8 | 10 | 10.0 | -0.0 | 0.28 | 7.55 | 12.5 | 9.4 | 10.6 |
| 9 | 11 | 10.0 | 1.0 | 0.28 | 7.55 | 12.5 | 9.4 | 10.6 |
| 10 | 9 | 8.9 | 0.1 | 0.35 | 6.41 | 11.4 | 8.1 | 9.7 |
| 11 | 8 | 8.9 | -0.9 | 0.35 | 6.41 | 11.4 | 8.1 | 9.7 |
| 12 | 10 | 8.9 | 1.1 | 0.35 | 6.41 | 11.4 | 8.1 | 9.7 |
| 13 | 8 | 7.8 | 0.2 | 0.49 | 5.20 | 10.4 | 6.7 | 8.9 |
| 14 | 9 | 7.8 | 1.2 | 0.49 | 5.20 | 10.4 | 6.7 | 8.9 |
| 15 | 7 | 7.8 | -0.8 | 0.49 | 5.20 | 10.4 | 6.7 | 8.9 |

# Regression Linéaire Simple : $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

**résultats**



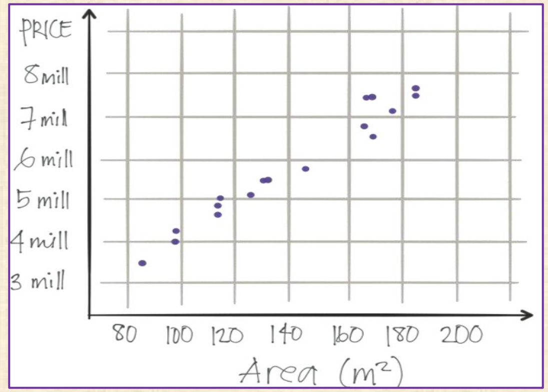

**intervalle de confiance : moyenne de Y**

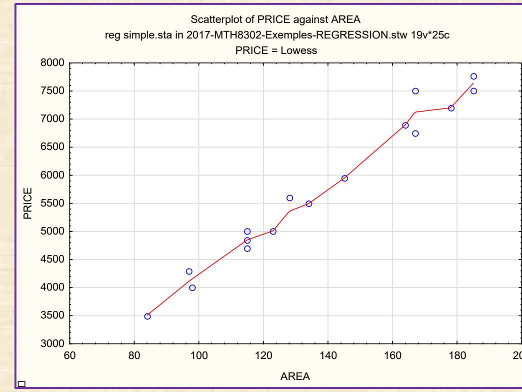



**intervalle de prédiction de Y**

# Autre exemple : prix maisons vs superficie

| AREA $m^2$ | PRICE Kkr |
|---|---|
| 134 | 5495 |
| 115 | 4700 |
| 167 | 7500 |
| 185 | 7775 |
| 84 | 3500 |
| 98 | 40000 |
| 115 | 4850 |
| 185 | 7500 |
| 164 | 6900 |
| 145 | 5950 |
| 123 | 5010 |
| 128 | 5600 |
| 167 | 6750 |
| 115 | 5000 |
| 178 | 7200 |
| 97 | 4290 |



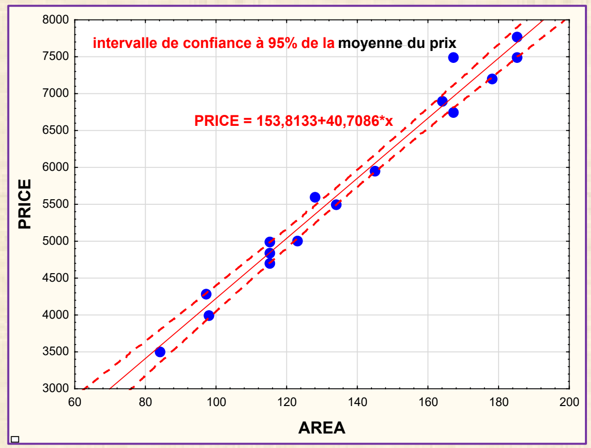



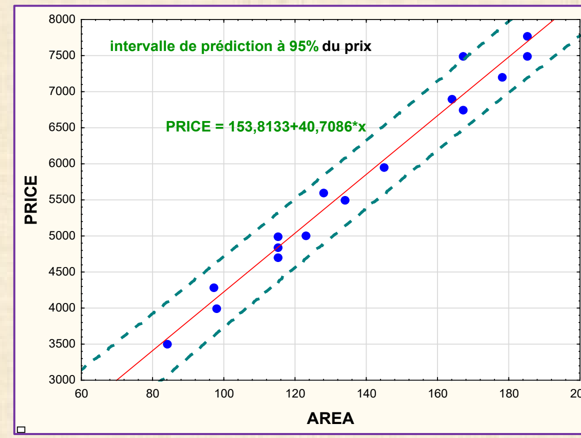

Regression Summary for Dependent Variable: PRICE
(reg simple.sta in 2017-MTH8302-Exemples-REGRESSION.stw)
R= ,98857556 R²= ,97728164 Adjusted R²= ,97565890
F(1,14)=602,24 p<,00000 Std.Error of estimate: 213,38

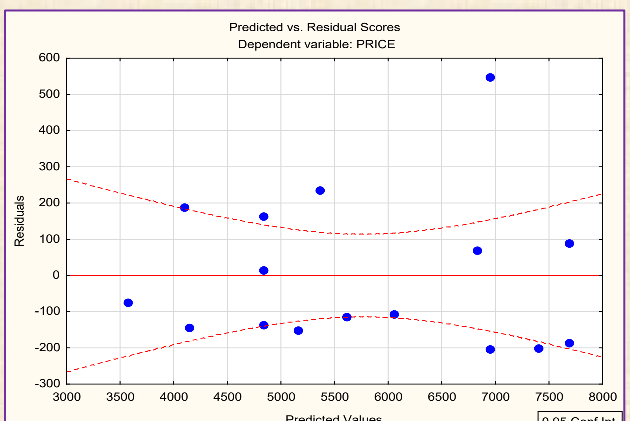
| N=16 | b* | Std.Err. of b* | b | Std.Err. of b | t(14) | p-value |
|---|---|---|---|---|---|---|
| Intercept | | | 153,8133 | 234,2436 | 0,65664 | 0,522061 |
| AREA | 0,988576 | 0,040283 | 40,7086 | 1,6588 | 24,54061 | 0,000000 |

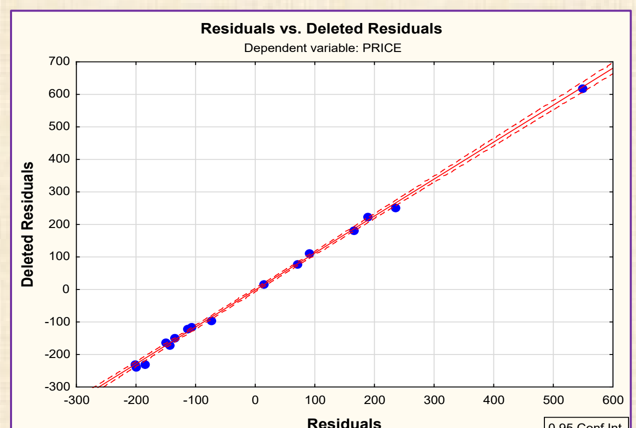
Analysis of Variance; DV: PRICE
(reg simple.sta in 2017-MTH8302-Exemples-REGRESSION.stw)

| Effect | Sums of Squares | df | Mean Squares | F | p-value |
|---|---|---|---|---|---|
| Regress. | 27419909 | 1 | 27419909 | 602,2418 | 0,000000 |
| Residual | 637416 | 14 | 45530 | | |
| Total | 28057325 | | | | |



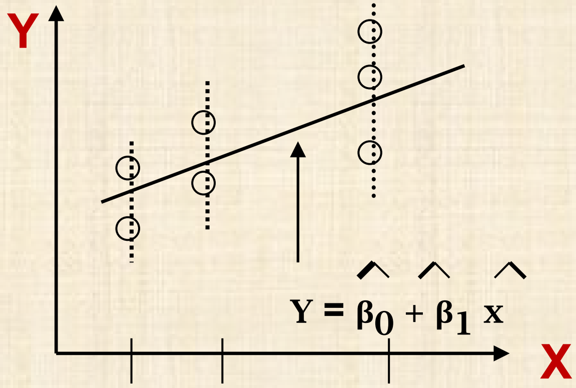
**JUSTESSE du MODÈLE:** avec observations répétées de Y à des valeurs de X

**Estimation de $\sigma^2$ :** - calculée avec le modèle ajusté (Droite de Moindres Carrés)
- dépend du modèle postulé

**question :** peut – on estimer $\sigma^2$ indépendamment du modèle postulé ($y = \beta_0 + \beta_1 x + \varepsilon$) ?

**réponse :** oui, si on a au moins 2 observations de Y à au moins 3 valeurs distinctes de X

**utilisation :** tester le manque d'ajustement (maj) du modèle postulé




| Données | | | | | | somme carrés | ddl |
|---|---|---|---|---|---|---|---|
| $x_1 \rightarrow$ | $y_{11}$ | $y_{12}$ | .... | $y_{1\,n1} \rightarrow$ | | $(n_1 - 1)\ s_1^2$ | $n_1 - 1$ |
| $x_2 \rightarrow$ | $y_{21}$ | $y_{22}$ | .... | $y_{2\,n2} \rightarrow$ | | $(n_2 - 1)\ s_2^2$ | $n_2 - 1$ |
| ..... | .... | ..... | ..... | ..... | | ............. | .......... |
| $x_k \rightarrow$ | $y_{k1}$ | $y_{k2}$ | .... | $y_{k\,nk} \rightarrow$ | | $(n_k - 1)\ s_k^2$ | $n_k - 1$ |
| | | | | | | $SSE = \sum (n_j - 1)\ s_j^2$ | $n - k$ |

**Nouvelle décomposition de la somme totale de carrés totale**

avant : SStot (totale) = SSreg (modèle ) + SSresid (résiduelle)

maintenant : SStot = SSreg + SSerpu (erreur pure ) + SSmaj (manque d'ajustement)

deg. liberté : $n - 1 = 1 + (n - k) + (k - 2)$

remarque : SSmaj est calculée par différence SSmaj = SSresid – SSerpu

# RÉGRESSION LINÉAIRE SIMPLE : observations répétées

Théorie

## tableau d'analyse de la variance modifié

| SOURCE | DDL | SS | MS = SS / DDL | F-ratio | p-valeur |
|---|---|---|---|---|---|
| modèle | 1 | SSreg | MSreg = SSreg/1 | $F_1$ = MSreg / MSresid | $p_1 = P(F \geq F_1)$ |
| résiduelle | n – 2 | SSresid | MSresid = SSresid/(n-2) | ---------- | ---------- |
| LOF | k – 2 | SSmaj | MSmaj = SSmaj/(k -2) | $F_2$ = MSmaj / MSerpu | $p_2 = P(F \geq F_2)$ |
| pure | n – k | SSerpu | MSerpu = Sserpu / (n – k) = $\hat{\sigma}^2$ | ----------- | ----------- |
| Totale | n – 1 | SStot | ----------- | | |

## test du manque d'ajustement du modèle linéaire

$$H_{0M} : E(Y \mid x) = \beta_0 + \beta_1 x \quad \text{versus} \quad H_1 : \text{non } H_{0M}$$

test : rejeter $H_0$ au seuil $\alpha$ si $F_2 > F_{n-k\,,\,k-2\,,\,1-\alpha}$

**remarques**

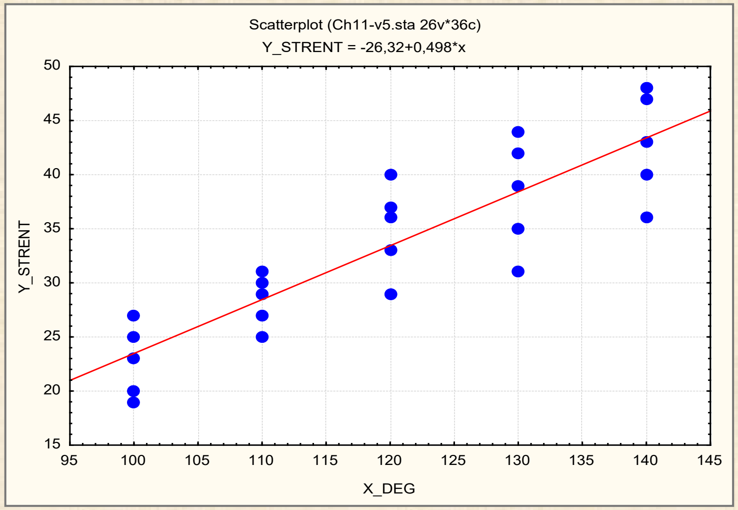
- on rejette $H_{0M}$ si $p_2 < \alpha$ (seuil)
- si on rejette $H_{0M}$, il faut postulé une autre équation ( modèle ) que la droite
- si on ne rejette pas $H_{0M}$, on conserve le tableau d'analyse de la variance (p 6)

# RÉGRESSION LINÉAIRE SIMPLE : observations répétées

**Exemple** : expérience avec un facteur X contrôlé

X = température (degrés F)    Y = résistance matériau

données:
- X = 100 : Y = 20 – 25 - 23 – 27 – 19
- X = 110 : Y = 25 – 29 – 31 – 30 – 27
- X = 120 : Y = 36 – 37 – 29 – 40 – 33
- X = 130 : Y = 35 – 39 – 31 – 42 – 44
- X = 140 : Y = 43 – 40 – 36 – 48 – 47




R = 0,879    $R^2$ = 0,773    F(1,23)=78,323    p < 0,000

| | b* | Std. Error | b | Std Errorr | t(23) | p-level |
|---|---|---|---|---|---|---|
| Interct | | | -26.320 | 6.799 | -3.87 | 0.0007 |
| X_DEG | 0.879 | 0.099 | 0.498 | 0.056 | 8.85 | 0.0000 |

| ANOVA | SS | DF | MS | F | p-level |
|---|---|---|---|---|---|
| Regress. | 1240.02 | 1 | 1240.02 | 78.323 | 0.0000 |
| Residual | 364.14 | 23 | 15.832 | | |
| Total | 1604.16 | 24 | | | |

**Exemple** : expérience avec un facteur X contrôlé (expérience)

X = température (degrés F)        Y = résistance matériau

données: X = 100 : Y = 20 – 25 - 23 – 27 – 19        X = 110 : Y = 25 – 29 – 31 – 30 – 27

X = 120 : Y = 36 – 37 – 29 – 40 – 33        X = 130 : Y = 35 – 39 – 31 – 42 – 44

X = 140 : Y = 43 – 40 – 36 – 48 – 47

| X_DEG | Y_STRENT Means | Y_STRENT Std.Dev. |
|---|---|---|
| 100 | 22,8 | 3,35 |
| 110 | 28,4 | 2,41 |
| 120 | 35,0 | 4,18 |
| 130 | 38,2 | 5,26 |
| 140 | 42,8 | 4,97 |
| All Grps | 33,4 | 8,18 |

SSerpu = 4 ($3{,}35^2 + 2{,}41^2 + \cdots + 4{,}97^2$)

= 347,6 avec 4*5 = 20 ddl

SSmaj = SSresid – SSerpu = 364,14 – 347,60

= 16,54 avec 23-20 = 3 ddl

MSmaj = 16,54 / 3 = 5,51

MSerpu = SSerpu / 20 = 17,38

$F_2$ = 5,51 / 17,38 = 0,32

modèle linéaire OK : pas rejeté

## DÉFICIENCES DÉTECTÉES à l'analyse de résidus

**correctifs et transformations**

- rendre la **variance plus constante** (stabilisation de la variance)
- obtenir une **distribution gaussienne** pour le terme d'erreur
- **transformer** certains modèles non linéaires en modèles linéaires

### TRANSFORMATIONS pour stabiliser la variance : Y' = h( Y )

| Cas | lien $\sigma^2$ = var(Y) et $\mu$ = moy(Y) | transformation Y' |
|---|---|---|
| Y ~ Poisson | $\sigma^2 \;\alpha\; \mu$ | $Y' = Y^{0.5}$ |
| Y ~ Binomiale | $\sigma^2 \;\alpha\; \mu ( 1 - \mu )$ | $Y' = \arcsin(Y^{0.5})$ |
| plusieurs ordres de grandeurs pour Y | $\sigma^2 \;\alpha\; \mu^2$ | $Y' = \log( Y )$ |
| autres cas: transformation Box-Cox | | $Y' = Y^{\lambda}$ $\quad -2 \le \lambda \le 2$ <br> $\lambda = ?$ |

# RÉGRESSION SIMPLE : transformations

## TRANSFORMATIONS sur X ou Y pour rendre linéaire certains modèles

| modèle originel | transformation | | | modèle linéaire | figure |
|---|---|---|---|---|---|
| $Y = \beta_0 X^{\beta_1}$ | $X' = \ln(X)$ | $Y' = \ln(Y)$ | $\beta_0' = \ln(\beta_0)$ | $Y' = \beta_0' + \beta_1 X'$ | a |
| $Y = \beta_0 \exp(\beta_1 X)$ | $X' = X$ | $Y' = \ln(Y)$ | $\beta_0' = \ln(\beta_0)$ | $Y' = \beta_0' + \beta_1 X'$ | b |
| $Y = 1 / (\beta_0 + \beta_1 X)$ | $X' = X$ | $Y' = 1 / Y$ | | $Y' = \beta_0 + \beta_1 X'$ | c |

Régression logistique : Y à 2 valeurs catégoriques (oui/non) (0, 1) d

$$Y = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 X)}$$

linéarisation: $\log(Y/(1-Y)) = \beta_0 + \beta_1 x$

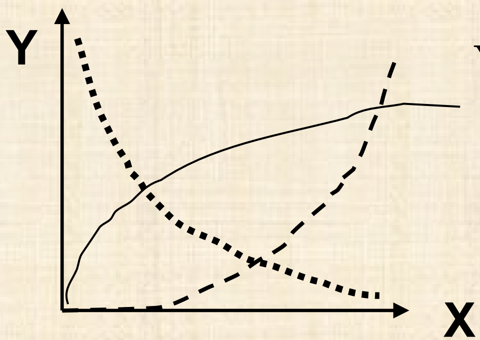



figure a

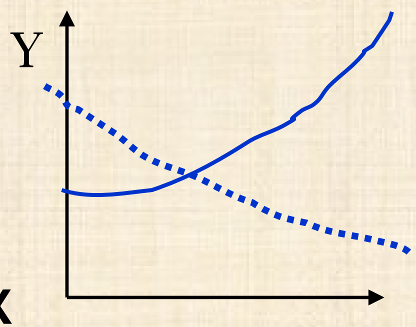



figure b

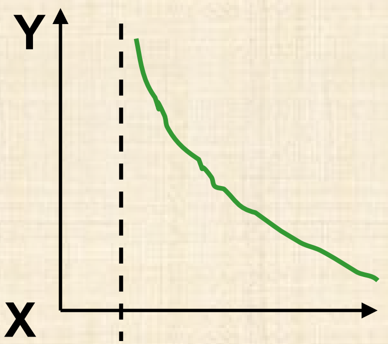



figure c

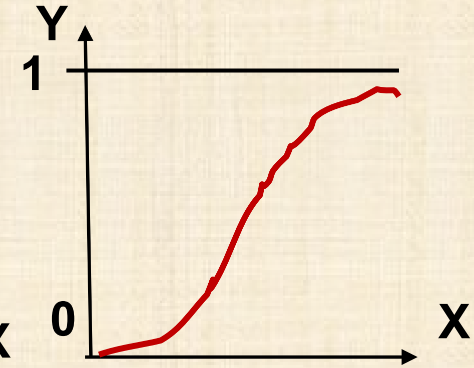



figure d

fonction logistique

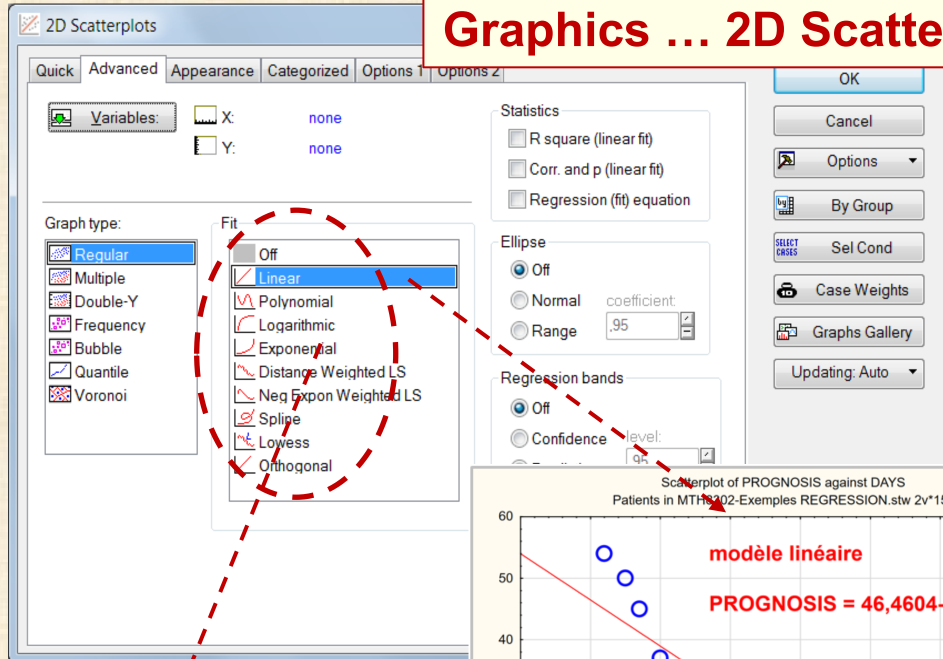
# RÉGRESSION SIMPLE : transformations pour linéariser

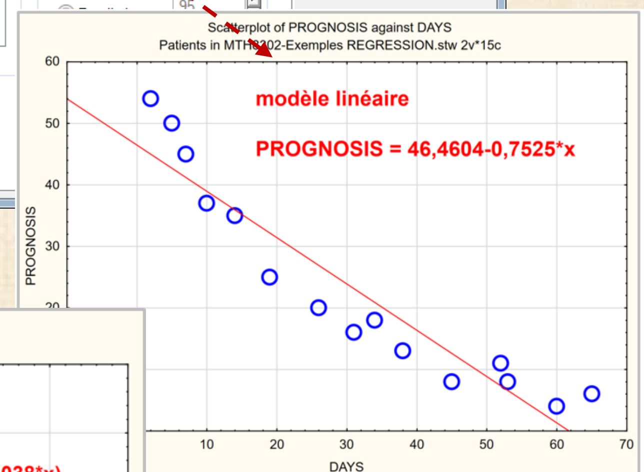
**Exemple**: patients.sta
DAYS number of days that each patient was hospitalized
PROGNOSIS index of the prognosis for long-term recovery
larger values reflect a better prognosis

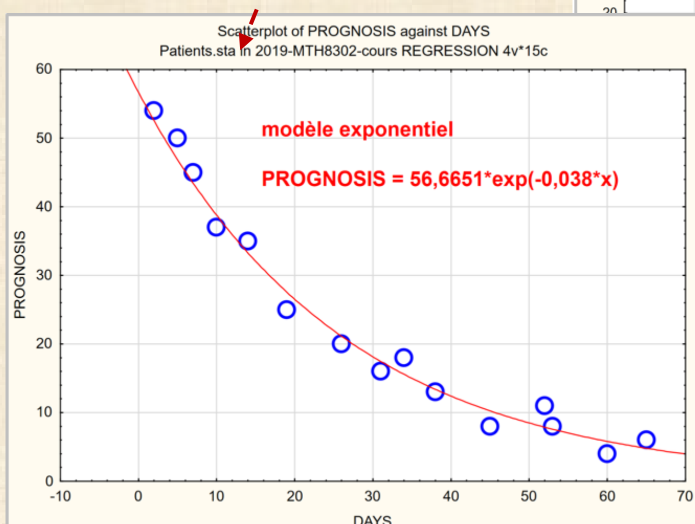
Long term prognosis of severely injured patients

| 1 ID | 2 Patient | 3 DAYS | 4 Y_PROGNOSIS |
|---|---|---|---|
| 1 | A | 2 | 54 |
| 2 | B | 5 | 50 |
| 3 | C | 7 | 45 |
| 4 | D | 10 | 37 |
| 5 | E | 14 | 35 |
| 6 | F | 19 | 25 |
| 7 | G | 26 | 20 |
| 8 | H | 31 | 16 |
| 9 | I | 34 | 18 |
| 10 | J | 38 | 13 |
| 11 | K | 45 | 8 |
| 12 | L | 52 | 11 |
| 13 | M | 53 | 8 |
| 14 | N | 60 | 4 |
| 15 | O | 65 | 6 |

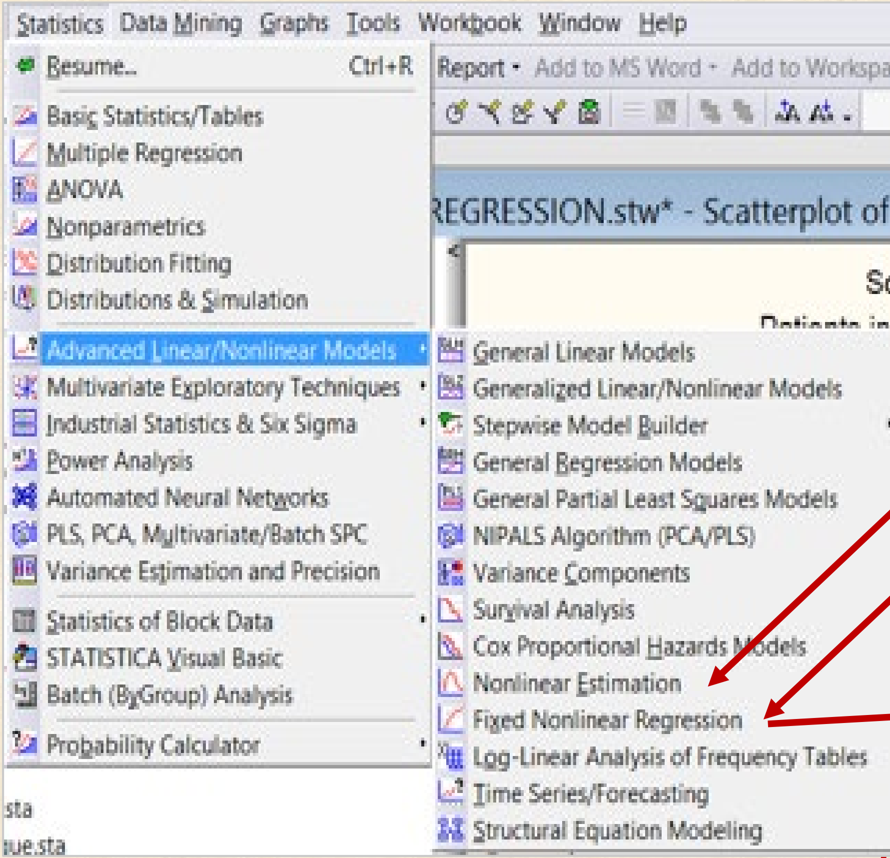
## Graphics … 2D Scatterplots



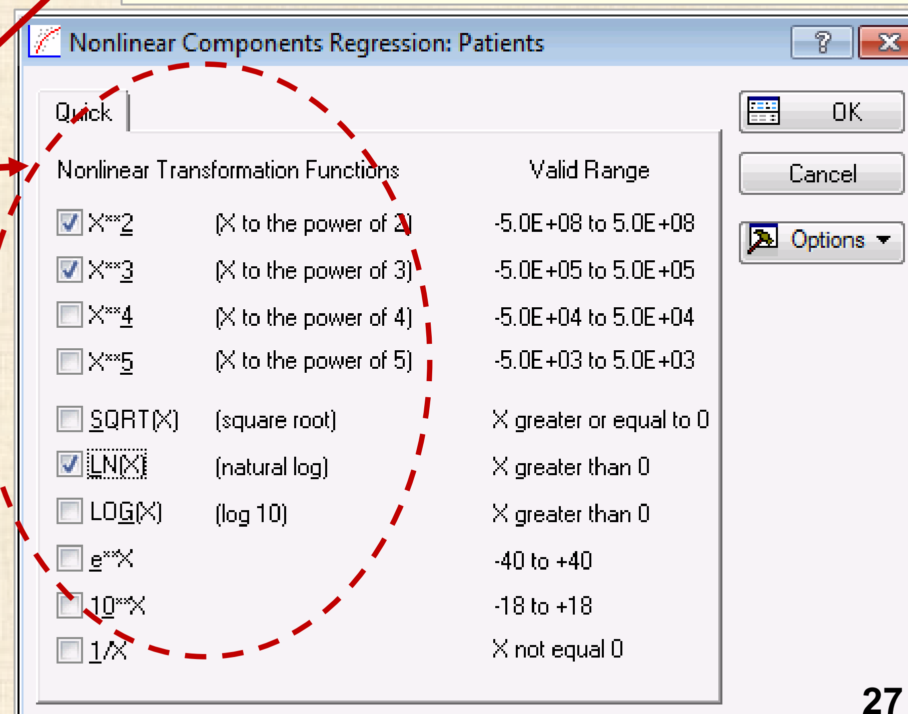



# RÉGRESSION non linéaire avec Statistica

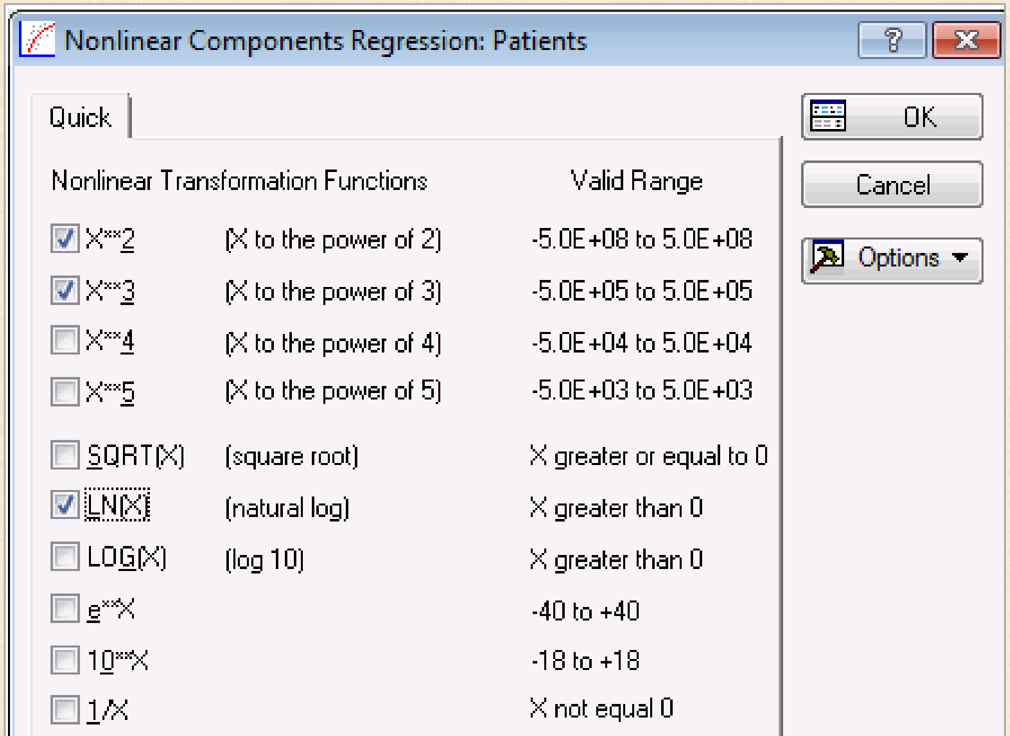



**Nonlinear Estimation**

**Fixed Nonlinear Regression:**
**transformations programmées**



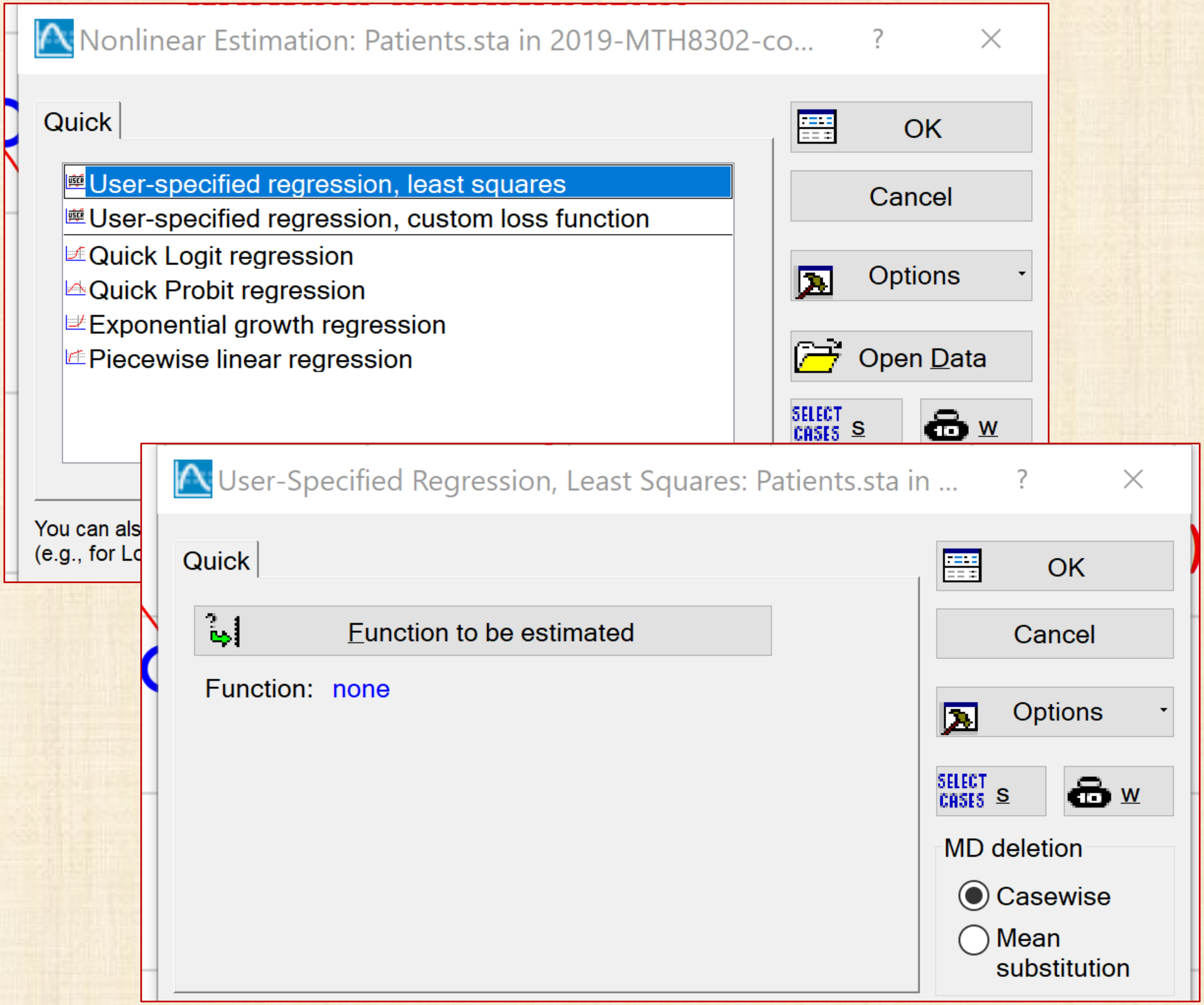
## Exemple: patients.sta … Recherche une transformation des variables




**identification de la transformation la plus utile avec la matrice de corrélation**

| Variable | Correlations (Patients.sta in Exemples REGRESSION-analyses.stw) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | DAYS | PROGNOSIS | V1**2 | V1**3 | LN-V1 | V2**2 | V2**3 | LN-V2 |
| DAYS | 1,000000 | -0,941053 | 0,968107 | 0,914647 | 0,919686 | -0,858993 | -0,786973 | -0,977280 |
| PROGNOSIS | -0,941053 | 1,000000 | -0,833385 | -0,742322 | -0,982535 | 0,976196 | 0,934715 | 0,949181 |
| V1**2 | 0,968107 | -0,833385 | 1,000000 | 0,985486 | 0,806789 | -0,715873 | -0,630328 | -0,937609 |
| V1**3 | 0,914647 | -0,742322 | 0,985486 | 1,000000 | 0,718699 | -0,613618 | -0,528271 | -0,880520 |
| LN-V1 | 0,919686 | -0,982535 | 0,806789 | 0,718699 | 1,000000 | -0,980910 | -0,958517 | -0,910681 |
| V2**2 | -0,858993 | 0,976196 | -0,715873 | -0,613618 | -0,980910 | 1,000000 | 0,988570 | 0,864646 |
| V2**3 | -0,786973 | 0,934715 | -0,630328 | -0,528271 | -0,958517 | 0,988570 | 1,000000 | 0,793656 |
| LN-V2 | -0,977280 | 0,949181 | -0,937609 | -0,880520 | -0,910681 | 0,864646 | 0,793656 | 1,000000 |

# Nonlinear Estimation : specification du modèle par l'usager



## Exemple: patients.sta avec Fixed Nonlinear Regression

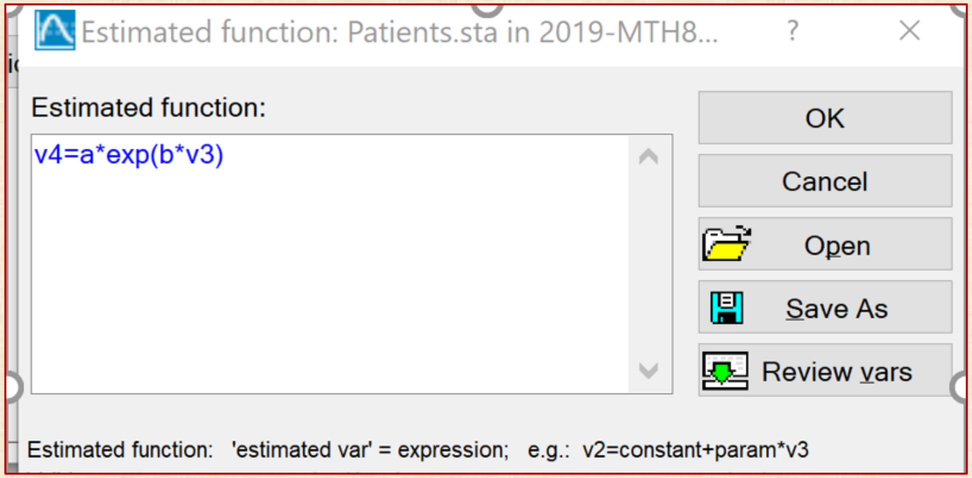
Long term prognosis of severely injured patients

| 1<br>ID | 2<br>Patient | 3<br>DAYS | 4<br>Y_PROGNOSIS |
|---|---|---|---|
| 1 | A | 2 | 54 |
| 2 | B | 5 | 50 |
| 3 | C | 7 | 45 |
| 4 | D | 10 | 37 |
| 5 | E | 14 | 35 |
| 6 | F | 19 | 25 |
| 7 | G | 26 | 20 |
| 8 | H | 31 | 16 |
| 9 | I | 34 | 18 |
| 10 | J | 38 | 13 |
| 11 | K | 45 | 8 |
| 12 | L | 52 | 11 |
| 13 | M | 53 | 8 |
| 14 | N | 60 | 4 |
| 15 | O | 65 | 6 |

**En Statistica, les variables, ont 2 noms**

- celui qui est donné par l'usager comme
  *ID , patient, DAYS, Y_PROGNOSIS*
- alias v0, v1, v2, v3, v4 donné par Statistica
  - v0 = numéro de ligne
  - v1 = ID
  - v2 = patient
  - v3 = DAYS
  - v4 = Y_PROGNOSIS

utilité : simplifie l'écriture des équations comme dans l'exemple suivant



**Exemple**: patients.sta avec Fixed Nonlinear Regression

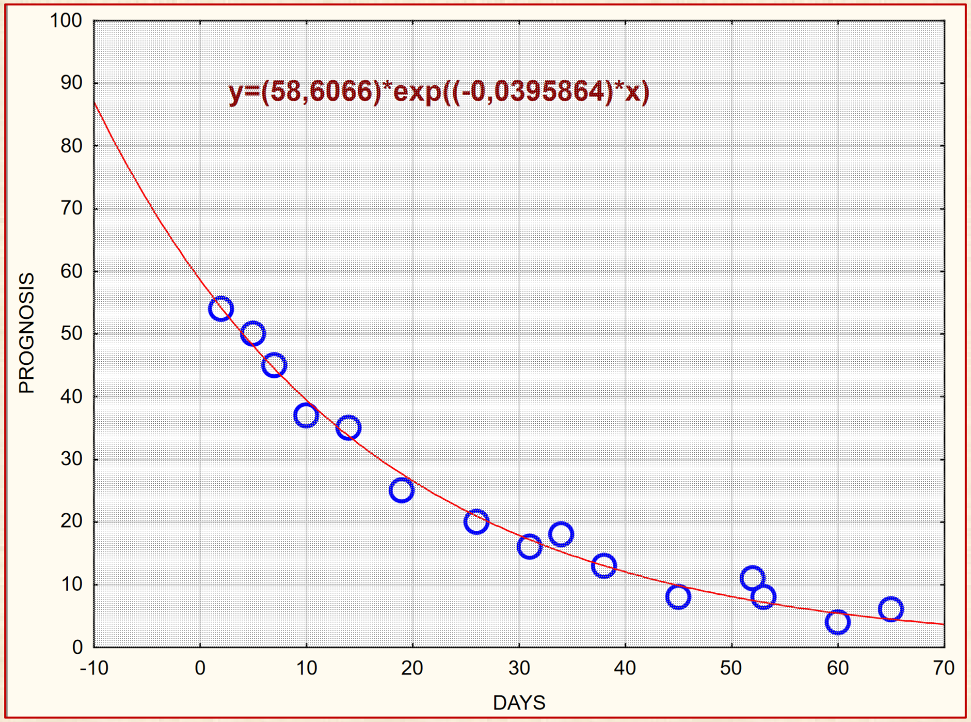
Model is: v4=a*exp(b*v3) (Patients.sta in 2019-MTH8302-cours REGRESS
Dep. Var. : PROGNOSIS
Level of confidence: 95.0% ( alpha=0.050)

| | Estimate | Standard error | t-value df = 13 | p-value | Lo. Conf Limit | Up. Conf Limit |
|---|---|---|---|---|---|---|
| a | 58,60656 | 1,472155 | 39,8101 | 0,000000 | 55,42616 | 61,78696 |
| b | -0,03959 | 0,001711 | -23,1326 | 0,000000 | -0,04328 | -0,03589 |

Model is: v4=a*exp(b*v3) (Patients.sta in
Dep. Var. : PROGNOSIS

| | Loss Function | a | b |
|---|---|---|---|
| 1 | 128,6318 | 0,10000 | 0,100000 |
| 2 | 107,5159 | 0,48400 | 0,024948 |
| 3 | 106,0166 | 0,55084 | 0,041731 |
| 4 | 104,7397 | 0,80603 | 0,043587 |
| 5 | 103,0540 | 1,29923 | 0,032370 |
| 6 | 99,9276 | 2,29417 | 0,023877 |
| 7 | 94,5864 | 4,28583 | 0,013376 |
| 8 | 85,1784 | 8,27081 | 0,002636 |
| 9 | 69,0339 | 16,24156 | -0,010138 |
| 10 | 41,8540 | 32,18361 | -0,026108 |
| 11 | 10,6020 | 57,10354 | -0,043433 |
| 12 | 7,0460 | 58,59191 | -0,039305 |
| 13 | 7,0327 | 58,60049 | -0,039576 |
| 14 | 7,0327 | 58,60635 | -0,039586 |
| 15 | 7,0327 | 58,60656 | -0,039586 |

**Exemple**: patients.sta avec Fixed Nonlinear Regression




| Effect | Model is: v4=a*exp(b*v3) (Patients.sta in 2019-MTH8302-cours REGRESS Dep. Var. : PROGNOSIS<br>1<br>Sum of Squares | 2<br>DF | 3<br>Mean Squares | 4<br>F-value | 5<br>p-value |
|---|---|---|---|---|---|
| **Regression** | 12060,54 | 2,00000 | 6030,270 | 1585,011 | 0,000000 |
| Residual | 49,46 | 13,00000 | 3,805 | | |
| Total | 12110,00 | 15,00000 | | | |
| Corrected Total | 3943,33 | 14,00000 | | | |
| Regression vs.Corrected Total | 12060,54 | 2,00000 | 6030,270 | 21,409 | 0,000055 |

# RÉGRESSION SIMPLE : modèle logistique

Si Y = 0 ou 1 peut-on faire de la régression?
Réponse: oui , régression logistique

**Exemple** : burns.sta
435 obs.
tableau 1 : données

| patient | X_log(surf+1) | Y |
|---|---|---|
| 1 | 1,35 | 0 |
| ........ | | |
| 34 | 1,75 | 1 |
| ........ | | |
| 434 | 2,35 | 0 |
| 435 | 2,35 | 1 |

Y = 0 si survit
= 1 si décès

n = 435 adultes traités pour des brulures au 3ième degré en fonction de log(surface +1) de brulure

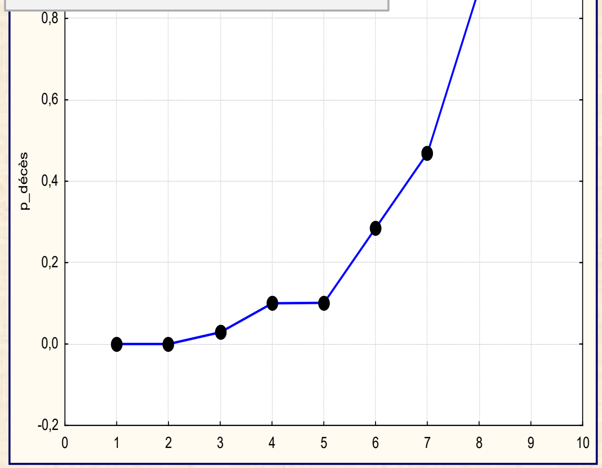
tableau 2: résumé (sans perte d'info)

| id | X | Y | nombre |
|---|---|---|---|
| 1 | 1,35 | 0 | 13 |
| 2 | 1,35 | 1 | 0 |
| 3 | 1,60 | 0 | 19 |
| 4 | 1,60 | 1 | 0 |
| 5 | 1,75 | 0 | 67 |
| 6 | 1,75 | 1 | 2 |
| 7 | 1,85 | 0 | 45 |
| 8 | 1,85 | 1 | 5 |
| 9 | 1,95 | 0 | 71 |
| 10 | 1,95 | 1 | 8 |
| 11 | 2,05 | 0 | 50 |
| 12 | 2,05 | 1 | 20 |
| 13 | 2,15 | 0 | 35 |
| 14 | 2,15 | 1 | 31 |
| 15 | 2,25 | 0 | 7 |
| 16 | 2,25 | 1 | 49 |
| 17 | 2,35 | 0 | 1 |
| 18 | 2,35 | 1 | 12 |

tableau 3 : réorganisation

| | X | n_ survit | n_ décès | n_ total | p_décès = n_décès / n_total |
|---|---|---|---|---|---|
| 1 | 1,35 | 13 | 0 | 13 | 0,0000 |
| 2 | 1,60 | 19 | 0 | 19 | 0,0000 |
| 3 | 1,75 | 67 | 2 | 69 | 0,0290 |
| 4 | 1,85 | 45 | 5 | 50 | 0,1000 |
| 5 | 1,95 | 71 | 8 | 79 | 0,1013 |
| 6 | 2,05 | 50 | 20 | 70 | 0,2857 |
| 7 | 2,15 | 35 | 31 | 66 | 0,4697 |
| 8 | 2,25 | 7 | 49 | 56 | 0,8750 |
| 9 | 2,35 | 1 | 12 | 13 | 0,9231 |
| | total | 308 | 127 | 435 | |

p_décès
= n_décès / n_total




## 3 méthodes d'ajustement

- Linéarisation
- Nonlinear estimation
- GLZ (linéaire généralisé)

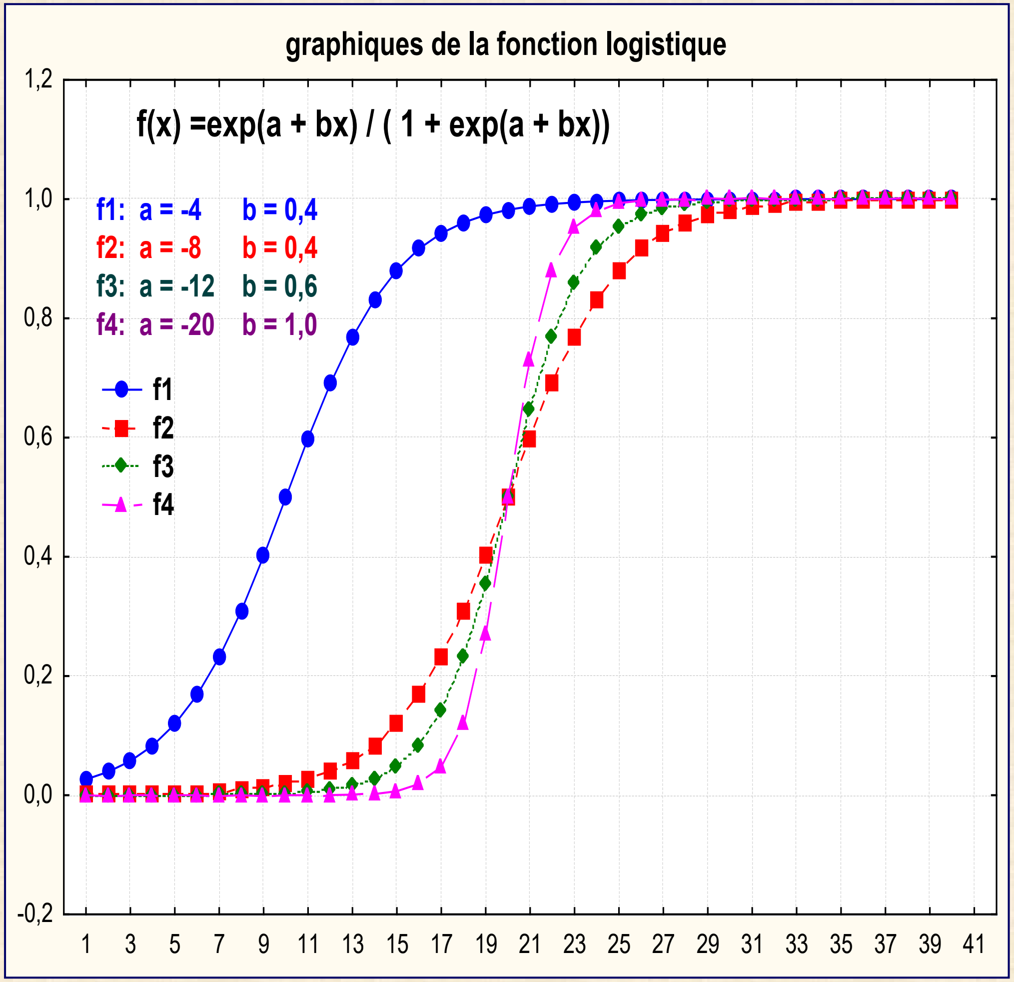
# RÉGRESSION SIMPLE : modèle logistique

## fonction logistique

$$\pi(x) = \exp(\alpha + \beta x)/(1 + \exp(\alpha + \beta x))$$

- courbe en forme sigmoïdale
- $0 \leq \pi(x) \leq 1$ …. probabilités
- si x tends vers $-\infty$ $\pi(x)$ tend vers 0
- si x tends vers $\infty$ $\pi(x)$ tend vers 1
- si $x = -\alpha/\beta$ alors $\pi(x) = 0,5$

  et la pente de f(x) vaut $\beta/4$




$$\text{Prob}(Y = 1 \text{ si } X = x) = \pi(x) = \exp(\alpha + \beta x)/(1 + \exp(\alpha + \beta x))$$

$$\text{Prob}(Y = 0 \text{ si } X = x) = 1 - \pi(x) = 1/(1 + \exp(\alpha + \beta x))$$

$$\log[\pi(x)/(1 - \pi(x))] = \alpha + \beta x$$

$$\log[\pi(x)/(1 - \pi(x))] = \text{logit}\,(\pi(x))$$

$\pi(x) / (1 - \pi(x))$: rapport des cotes (odd ratio)

## Interprétation du paramètre β

$\text{logit}(\pi(x)) = \log[\pi(x) / (1 - \pi(x))] = \alpha + \beta x$

$\text{logit}(\pi(x+1)) = \log[\pi(x+1) / (1 - \pi(x+1))] = \alpha + \beta(x+1) = \alpha + \beta x + \beta$

$\beta = \text{logit}(\pi(x+1)) - \text{logit}(\pi(x))$

$= \log \{[\pi(x+1) / (1 - \pi(x+1))] / \pi(x) / (1 - \pi(x))]\}$

$= \log \{ \text{odds}(x+1) / \text{odds}(x) \}$

$e^{\beta} = \text{odds}(x+1) / \text{odds}(x) = \text{odds ratio}$ (ratio des cotes)

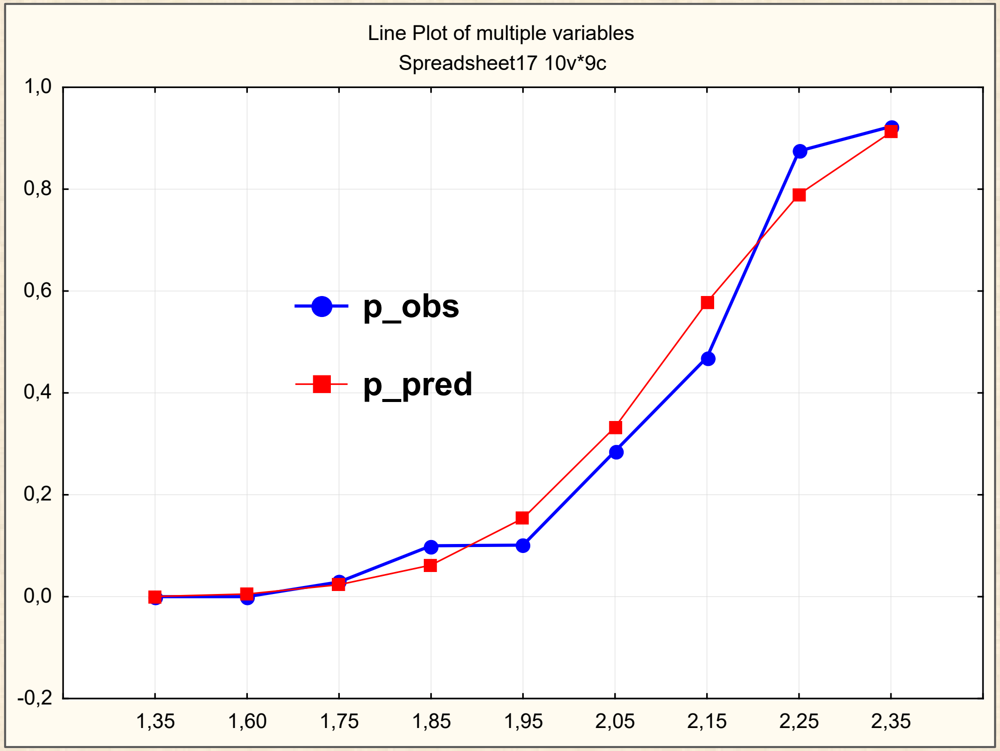
accroissement de la probabilité que Y = 1

si x augmente de 1

## RÉGRESSION SIMPLE : modèle logistique

## Tableau

| i | X | p décès | logit(p) (obs) | logit(p) (pred) | p décès (pred) |
|---|---|---|---|---|---|
| 1 | 1,35 | 0 | indéfini | -7,816 | 0,0004 |
| 2 | 1,60 | 0 | Indéfini | -5,248 | 0,0053 |
| 3 | 1,75 | 0,0290 | -3,540 | -3,708 | 0,0235 |
| 4 | 1,85 | 0,1000 | -2,197 | -2,67 | 0,0621 |
| 5 | 1,95 | 0,1013 | -2,183 | -1,64 | 0,1538 |
| 6 | 2,05 | 0,2857 | -0,916 | -0,621 | 0,3329 |
| 7 | 2,15 | 0,4697 | -0,092 | 0,406 | 0,5781 |
| 8 | 2,25 | 0,8750 | 2,10 | 1,434 | 0,7900 |
| 9 | 2,35 | 0,9231 | 2,485 | 2,461 | 0,9117 |




**logit(p) = ln[ p/(1-p) ]**

**logit(0) n'est pas défini**

**régression de logit(p)(obs) sur X**
**cas logit(p)(obs) défini i = 3, 4,.., 7**
**u= logit(p)(pred) = -21,68 + 10,27*X**

**utilisation de l'équation obtenue**
**pour tous les cas i = 1,2,…,7**
**p_pred = exp (u) / (1 + exp(u))**

**2 possibilités pour le calcul des prédictions**

**ajustement par linéarisation**
**équations linéaires – cas traité ici**

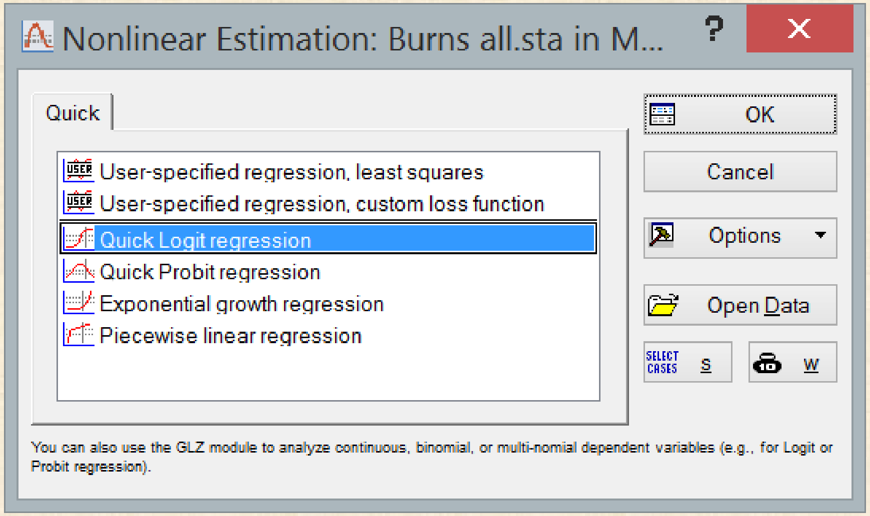
**ajustement par estimation non linéaire**
**équations non linéaires – page suivante**

# RÉGRESSION logistique : ajustement par estimation non linéaire

## Exemple : burns.sta

tableau 1 : données

| patient | X_log(surf+1) | Y | YC |
|---|---|---|---|
| 1 | 1,35 | 0 | 1 |
| ....... | ....... | ....... | ....... |
| 34 | 1,75 | 1 | 0 |
| ....... | ....... | ....... | ....... |
| 434 | 2,35 | 0 | 1 |
| 435 | 2,35 | 1 | 0 |

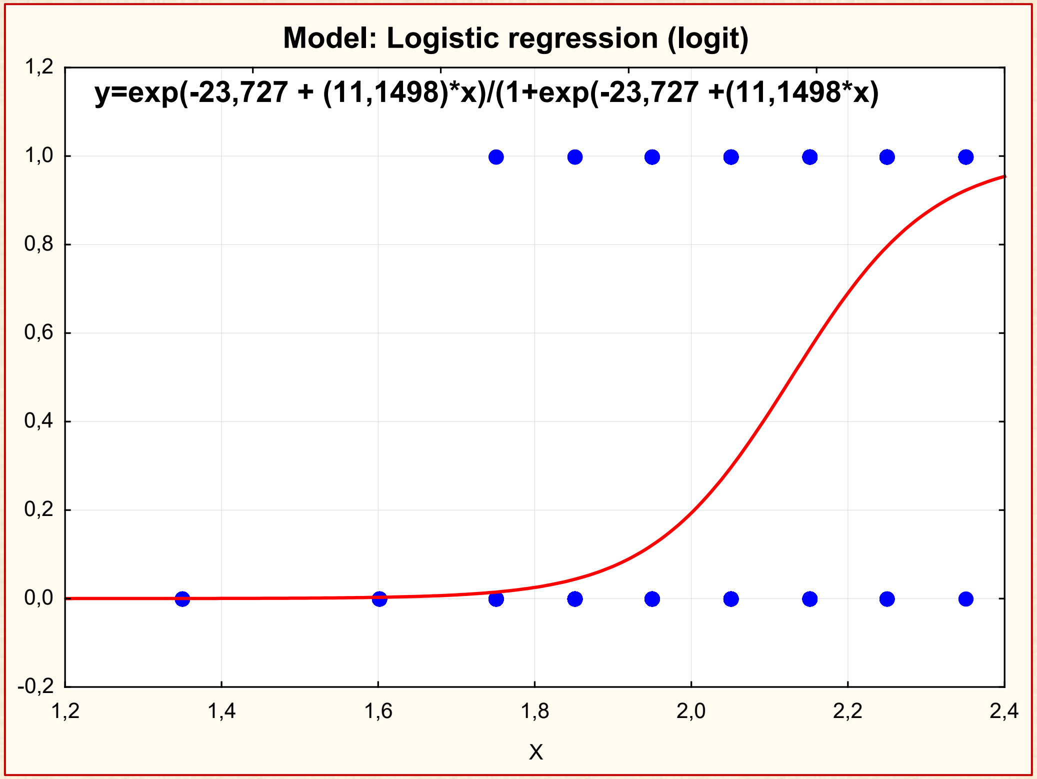



**Model: Logistic regression (logit)**

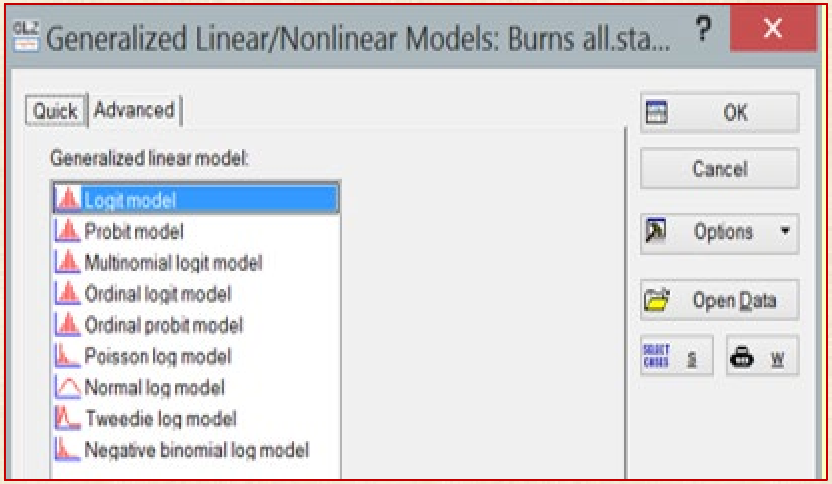
nombre de 0 : 308 cas
nombre de 1 : 127 cas

Max likelihood

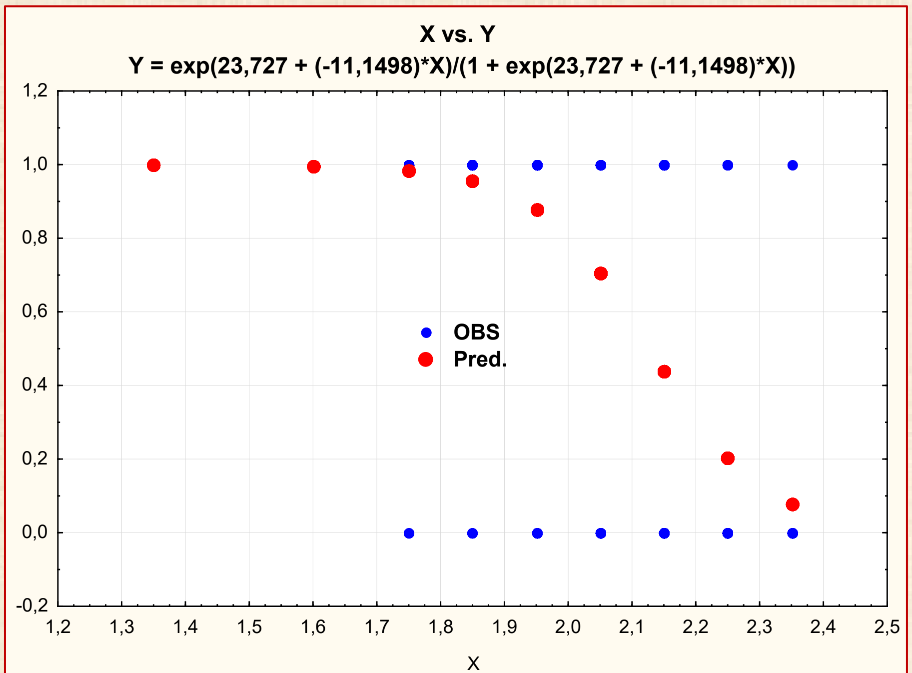
| | $\alpha$ | $\beta$ |
|---|---|---|
| Estimation | - 23,73 | 11,15 |

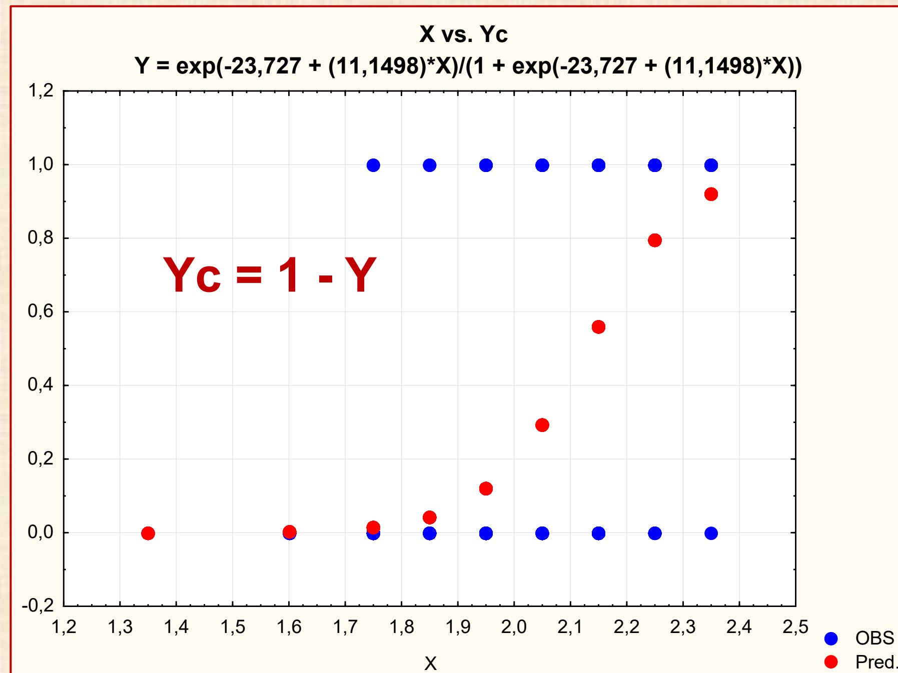


# RÉGRESSION logistique : ajustement par estimation non linéaire – module GLZ




| Effect | Y - Parameter estimates (Burns all.sta) Modeled probability that Y = 0 | | | |
|---|---|---|---|---|
| | Level of Effect | Column | Estimate | Standard Error |
| Intercept | | 1 | 23,7271 | 2,357973 |
| X | | 2 | -11,1498 | 1,125685 |
| Scale | | | 1,0000 | 0,000000 |




**remarque : codage employé par Statistica (page suivante)**

# Modèle logistique : remarque codage employé par Statistica

**RESPONSE CODES FOR 0 / 1 response variables** (source : Help de *Statistica*)

**When selecting (or specifying via syntax) response codes for binomial response variables (e.g., for logit models), the two selected values (numeric codes or text values) are recoded for the analysis,**

**so that the first code that is specified will be recoded to 1,**
**and the second value will be recoded to 0.**

**For example, if you specified two text values *No* and *Yes*, then the first value (*No*) will be recoded for the analysis to *1*, and the parameter estimates can be interpreted with respect to the probability of responding with *No*; so, if in such an analysis a positive parameter estimate is found for a particular predictor variable x¸ then it means that the more of x the greater is the likelihood of obtaining the response *No*.**

**Note that the default coding for binomial response variables in the *Nonlinear Estimation* module may not be the same as that performed in *GLZ*. One way to verify the particular coding that was performed is to review the spreadsheet of the predicted responses on the Resid 1 tab of the results dialog; that spreadsheet will report in the first column the recoded values (*0* or *1*) of the response variable.**
**Codes for dep. var. Logit and probit regression models, in a sense, predict probabilities underlying the dichotomous dependent variable, and these methods will produce predicted (expected) values in the range between 0 and 1**
**(for details, see Common Nonlinear Regression Models).**

**If the dependent variable is not coded in this way, that is, as 0 and 1, then you must specify the respective codes in the *Codes for dep. var.* boxes. To do this, double-click on this field (or press the F2 button on your keyboard) to display the variable dialog box or simply type in the two codes.**
**As the data are read, the dependent variable will then be transformed so that all values that match the first code become 0 (zero), and all values that match the second code become 1.**

# RÉGRESSION SIMPLE : estimation non linéaire

## Pharmacologie.sta : estimation réponse à une dose de médicament

| X DOSE | Y RESPONSE |
|---|---|
| 1 | 0,5 |
| 2 | 2,3 |
| 3 | 3,4 |
| 4 | 24,0 |
| 5 | 54,7 |
| 6 | 82,1 |
| 7 | 94,8 |
| 8 | 96,2 |
| 9 | 96,4 |

**Modèle**

dose X - response Y (en %)

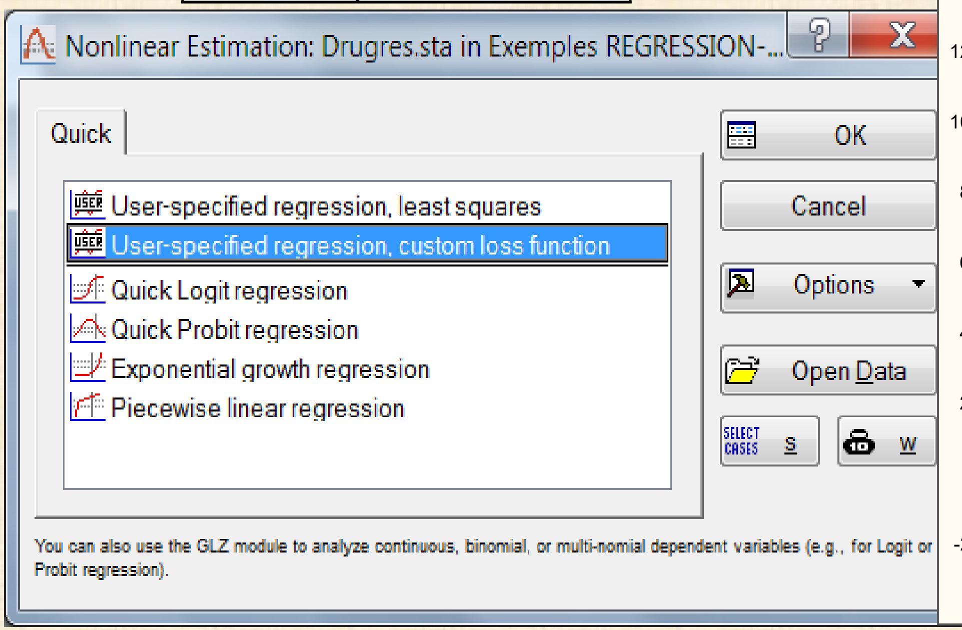
response = $\beta_0$ (1 - 1 / [1+(dose / $\beta_2$)**$\beta_1$]

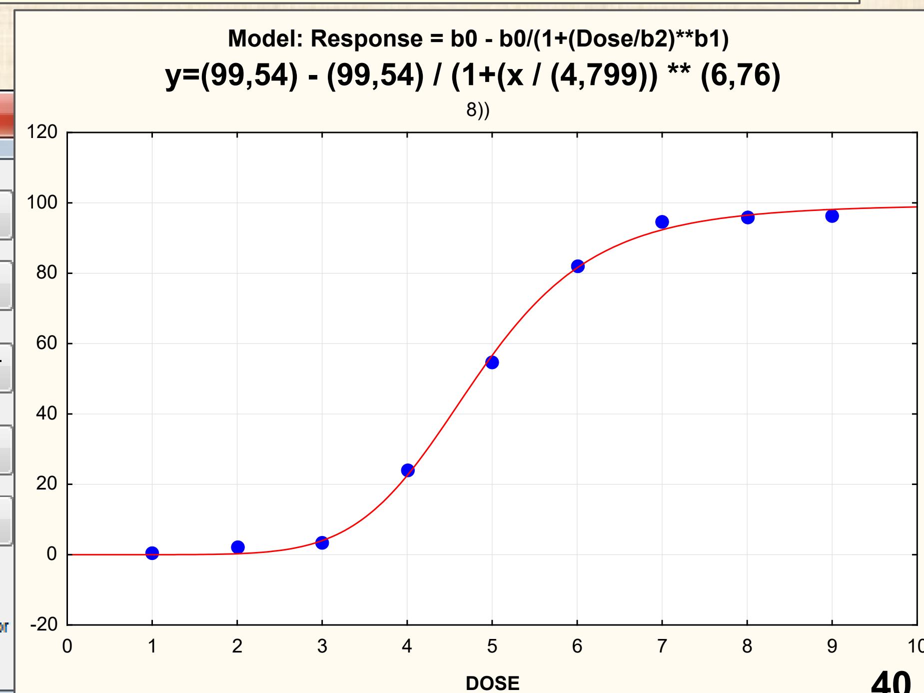
non linéarisable

$\beta_0$ valeur à saturation

$\beta_2$ concentration quand Y = 50%

$\beta_1$ pente function

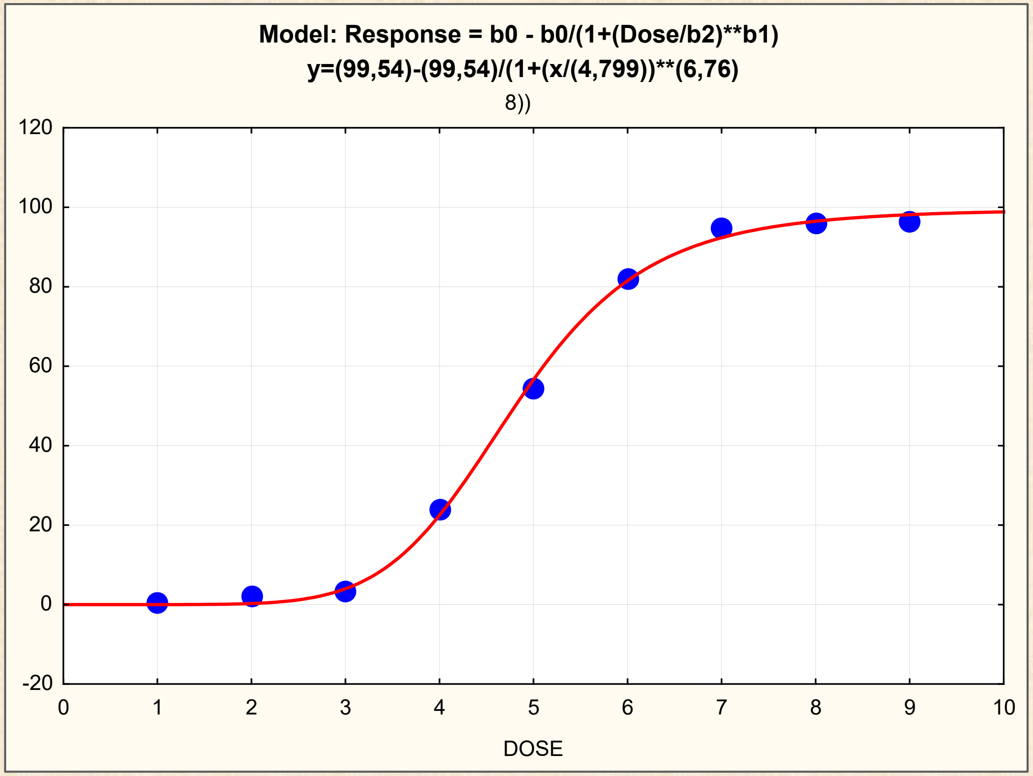


# RÉGRESSION SIMPLE : estimation non linéaire

| N=9 | Response = b0 - b0/(1+(Dose/b2)**b1)<br>Loss: (OBS-PRED)**2<br>Final loss: 20,1880<br>R= 0,99932<br>Variance explained: 99,865% | | |
|---|---|---|---|
| | b0 | b2 | b1 |
| Estimate | 99,54 | 4,80 | 6,76 |
| Std.Err. | 1,57 | 0,05 | 0,43 |
| t(6) | 63,33 | 95,13 | 15,80 |
| -95%CL | 95,69 | 4,68 | 5,71 |
| +95%CL | 103,39 | 4,92 | 7,81 |
| p-value | 0,00 | 0,00 | 0,00 |

| r | Model is:<br>Response = b0 - b0/(1+(Dose/b2)**b1) | | |
|---|---|---|---|
| | Observed | Predicted | Residuals |
| 1 | 0,500 | 0,0025 | 0,4975 |
| 2 | 2,300 | 0,2669 | 2,0331 |
| 3 | 3,400 | 3,9845 | -0,5845 |
| 4 | 24,000 | 22,4756 | 1,5244 |
| 5 | 54,700 | 56,6077 | -1,9077 |
| 6 | 82,100 | 81,5184 | 0,5816 |
| 7 | 94,800 | 92,3411 | 2,4589 |
| 8 | 96,200 | 96,4906 | -0,2906 |
| 9 | 96,400 | 98,1416 | -1,7416 |




## modélisation non linéaire

généralement avec des

études statistiques confirmatoires

considération d'une ou plusieurs

famille de modèles non linéaires

# Modèles non linéaires

logiciel **JMP**

référence: Nonlinear Regression with Built-In Models p.101-122

*JMP® 12 Specialized Models*. Cary, NC: SAS Institute Inc., March 2015

Dans les sciences biologiques et les sciences de la vie, il existe des familles de fonctions de forme connue qui décrivent la relation entre 2 variables. Par exemple, les chercheurs en pharmacologie conduisent des expériences pour comprendre la force de la réponse en fonction de la concentration du médicament. La famille de fonctions logistiques est souvent employée pour l'étude de ce processus. Un autre exemple est l'utilisation des courbes de croissance exponentielle pour prédire le nombre de personnes d'une population en fonction du temps.

## Description générale des fonctions

**Courbes en forme de S croissantes ou décroissantes (sigmoïde)**

- Fonctions logistiques, fonctions Gompertz ; ces fonctions ont une asymptote supérieure et une asymptote inférieure.
- Fonctions logistiques symétriques avec 2, 3 et 4 paramètres.
- Fonctions logistiques avec 5 paramètres, fonctions Gompertz : elles sont non symétriques.
- Fonctions logistique avec 2 paramètres : la réponse doit être entre 0 et 1.

**Courbes exponentielles de croissance et de décroissance**

- Fonctions exponentielles, double exponentielle (bi-exponentielle), fonction mécanique de croissance.
- Fonctions exponentielles avec 2 et 3 paramètres ; la fonction avec 3 paramètres a une asymptote.
- Fonctions bi-exponentielle possède 2 procédés de croissance ou de décroissance.
- Fonctions mécaniques et fonctions exponentielle avec 3 paramètres : ces fonctions sont toujours croissantes mais le taux de croissance diminue et les fonctions ont une asymptote.

**Courbes avec un pic**

- Fonctions gaussiennes, fonctions de Lorentz ; ces fonctions sont croissantes jusqu'à un pic et ensuite elles sont décroissantes.
- La fonction gaussienne est la densité de Gauss (normale) à un facteur d'échelle près.
- La fonction de Lorentz est la densité de Cauchy à un facteur d'échelle près.

**Courbes pharmacocinétique : étude de la concentration d'un médicament dans le corps**

- Fonctions de dose orale avec 1 compartiment ou 2 compartiments (Bolus).
- Fonctions bi-exponentielle avec 4 paramètres.

**Courbes Michaelis-Menten en cinétique biochimique.**

# Modèles non linéaires

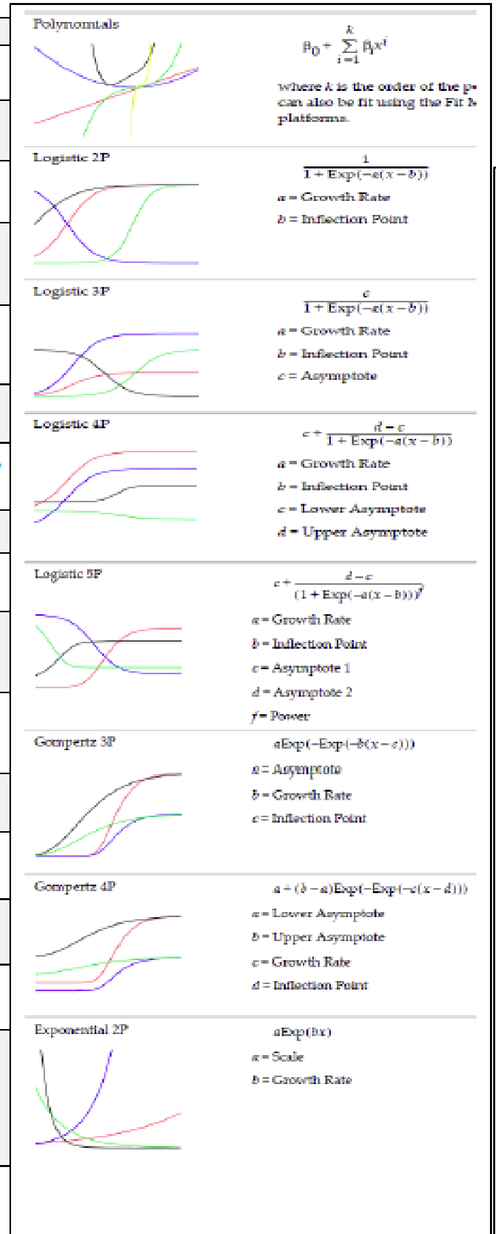
**Courbes sigmoid** : Logistic, Gompertz
**Croissance / décroissance exponentielle**:
Exponentiel, Bi-exponentiel, mechanique
**Modèles sommet** : Gauss, Lorentz
**Modèles pharmacocinétiques** : 1-compartiment dose orale, dose Bolus 2-compartiment Bi-exponentielle 4P
**Michaelis Menten**: cinétique biochimique

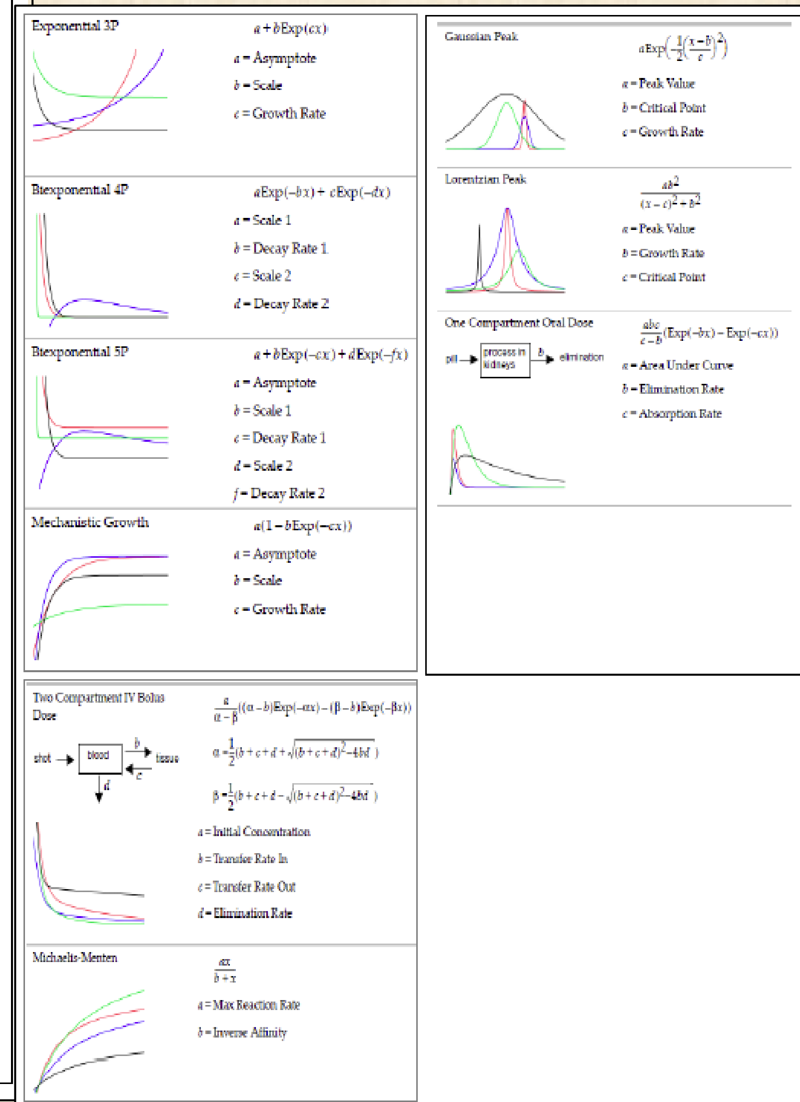
| | | |
|---|---|---|
| Exponentielle 2P | aExp(bx) | a échelle<br>b taux croissance |
| Exponentielle 3P | a + bExp(cx) | a asymptote<br>b échelle<br>c taux croissance |
| Bi Exponentielle 4P | aExp(– bx) + cExp(– dx) | a échelle 1<br>b taux décroissance 1<br>c échelle 2<br>d taux décroissance 2 |
| Bi Exponentielle 5P | a + bExp(– cx) + dExp(– fx) | a asymptote<br>b échelle 1<br>c taux décroissance 1<br>d échelle 2<br>f taux décroissance 2 |
| Croissance mécanique | a(1 – bExp(– cx)) | a asymptote<br>b échelle<br>c taux croissance |
| Sommet Gauss | $a\mathrm{Exp}\left(-\frac{1}{2}\left(\frac{x-b}{c}\right)^2\right)$ | a sommet<br>b point critique<br>c taux croissance |
| Sommet Lorentz | $\frac{ab^2}{(x-c)^2+b^2}$ | a sommet<br>b taux croissance<br>c point critique |

| MODÈLE | FORMULE Y = | PARAMÈTRES |
|---|---|---|
| Polynômes | $\beta_0 + \sum_{i=1}^{k} \beta_i x^i$ | |
| Logistique 2P | $\frac{1}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point d'inflexion |
| Logistique 3P | $\frac{c}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point inflexion<br>c asymptote |
| Logistique 4P | $c+\frac{d-c}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point inflexion<br>c asymptote inférieure<br>d asymptote supérieure |
| Logistique 5P | $c+\frac{d-c}{(1+\mathrm{Exp}(-a(x-b)))^f}$ | a taux croissance<br>b point inflexion<br>c asymptote inférieure<br>d asymptote supérieure<br>f puissance |
| Gompertz 3P | $a\mathrm{Exp}(-\mathrm{Exp}(-b(x-c)))$ | a asymptote<br>b taux croissance<br>c point Inflection |
| Gompertz 4P | $a+(b-a)\mathrm{Exp}(-\mathrm{Exp}(-c(x-d))$ | a asymptote inférieure<br>b asymptote supérieure<br>c taux croissance<br>d pont inflection |
| dose orale 1 compartiment | $\frac{abc}{c-b}(\mathrm{Exp}(-bx)-\mathrm{Exp}(-cx))$ | a aire courbe<br>b taux élimination<br>c taux absorption |
| dose Bolus 2 compartiments | $\frac{a}{\alpha-\beta}((\alpha-b)\mathrm{Exp}(-\alpha x)-(\beta-b)\mathrm{Exp}(-\beta x))$<br>$\alpha=\frac{1}{2}(b+c+d+\sqrt{(b+c+d)^2-4bd})$<br>$\beta=\frac{1}{2}(b+c+d-\sqrt{(b+c+d)^2-4bd})$ | a concentration initiale<br>b taux transfert In<br>c taux transfert Out<br>d taux éimination |
| Michaelis Menten | $\frac{ax}{b+x}$ | a taux max réaction<br>b infinité adverse |

# logiciel JMP

| MODÈLE | FORMULE Y = | |
|---|---|---|
| Polynômes | $\beta_0 + \sum_{i=1}^{k} \beta_i x^i$ | |
| Logistique 2P | $\frac{1}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point d'inflexion |
| Logistique 3P | $\frac{c}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point inflexion<br>c asymptote |
| Logistique 4P | $c+\frac{d-c}{1+\mathrm{Exp}(-a(x-b))}$ | a taux croissance<br>b point inflexion<br>c asymptote inférieure<br>d asymptote supérieure |
| Logistique 5P | $c+\frac{d-c}{(1+\mathrm{Exp}(-a(x-b)))^f}$ | a taux croissance<br>b point inflexion<br>c asymptote inférieure<br>d asymptote supérieure<br>f pulsasance |
| Gompertz 3P | $a\mathrm{Exp}(-\mathrm{Exp}(-b(x-c)))$ | a asymptote<br>b taux croissance<br>c point inflexion |
| Gompertz 4P | $a+(b-a)\mathrm{Exp}(-\mathrm{Exp}(-c(x-d)))$ | a asymptote inférieure<br>b asymptote supérieure<br>c taux croissance<br>d point inflection |
| Exponentielle 2P | aExp(bx) | a échelle<br>b taux croissance |
| Exponentielle 3P | a + bExp(cx) | a asymptote<br>b échelle<br>c taux croissance |
| Bi-Exponentielle 4P | aExp(– bx) + cExp(– dx) | a échelle 1<br>b taux décroissance 1<br>c échelle 2<br>d taux décroissance 2 |
| Bi-Exponentielle 5P | a + bExp(– cx) + dExp(– fx) | a asymptote<br>b échelle 1<br>c taux décroissance 1<br>d échelle 2<br>f taux décroissance 2 |
| Croissance mécanique | a(1 – bExp(– cx)) | a asymptote<br>b échelle<br>c taux croissance |
| Sommet Gauss | $a\mathrm{Exp}\left(-\frac{1}{2}\left(\frac{x-b}{c}\right)^2\right)$ | a sommet<br>b point critique<br>c taux croissance |
| Sommet Lorentz (Cauchy) | $\frac{ab^2}{(x-c)^2+b^2}$ | a sommet<br>b taux croissance<br>c point critique |
| dose orale 1 compartiment | $\frac{abc}{c-b}(\mathrm{Exp}(-bx)-\mathrm{Exp}(-cx))$ | a aire courbe<br>b taux élimination<br>c taux absorption |
| dose Bolus 2 compartiments | $\frac{a}{\alpha-\beta}((\alpha-b)\mathrm{Exp}(-\alpha x)-(\beta-b)\mathrm{Exp}(-\beta x))$<br>$\alpha=\frac{1}{2}(b+c+d+\sqrt{(b+c+d)^2-4bd})$<br>$\beta=\frac{1}{2}(b+c+d-\sqrt{(b+c+d)^2-4bd})$ | a concentration initiale<br>b taux transfert in<br>c taux transfert out<br>d taux élimination |
| Michaelis-Menten | $\frac{ax}{b+x}$ | a taux max réaction<br>b infinité adverse |



# Critères pour la comparaison de modèles

**AIC : Akaike Information Criterion** $AIC = -2 \log(L) + 2k$

L : vraisemblance maximisée k : nombre de paramètres du modèle
La déviance du modèle (- 2 log(L)) est pénalisée par 2 fois le nombre de paramètres k
L'AIC compromis entre le biais (qui diminue avec k ) et la parcimonie
bon modèle : min (décrire les données avec le plus petit nombre de paramètres)

**AICc : AIC corrigé** $AICc = AIC + [\, 2k(k-1) / (n-k+1) \,]$

modification du critère Akaide pour des petits échantillons ($n < 40k$).
On doit avoir $n \geq k + 2$. AICc est préférable à AIC pour de petits échantillons
AICc permet de comparer 2 modèles ou plus.
bon modèle : min

**AICc Weight**

Normalisation des valeurs AICc pour avoir une somme de 1
probabilité que le modèle est le meilleur parmi les modèles ajustés.
bon modèle : le plus près de 1

**BIC : Bayesian Information Criterion** $BIC = -2 * LL + k * \log(n)$

mesure de comparaison entre modèles
bon modèle : min

**SSE :** Sum of Square of Errors = somme différences entre Y observées et Y prédites

**MSE :** Mean Square Error valeur moyenne de SSE bon modèle : min

**RMSE :** Root Mean Square Error racine carrée de MSE bon modèle : min

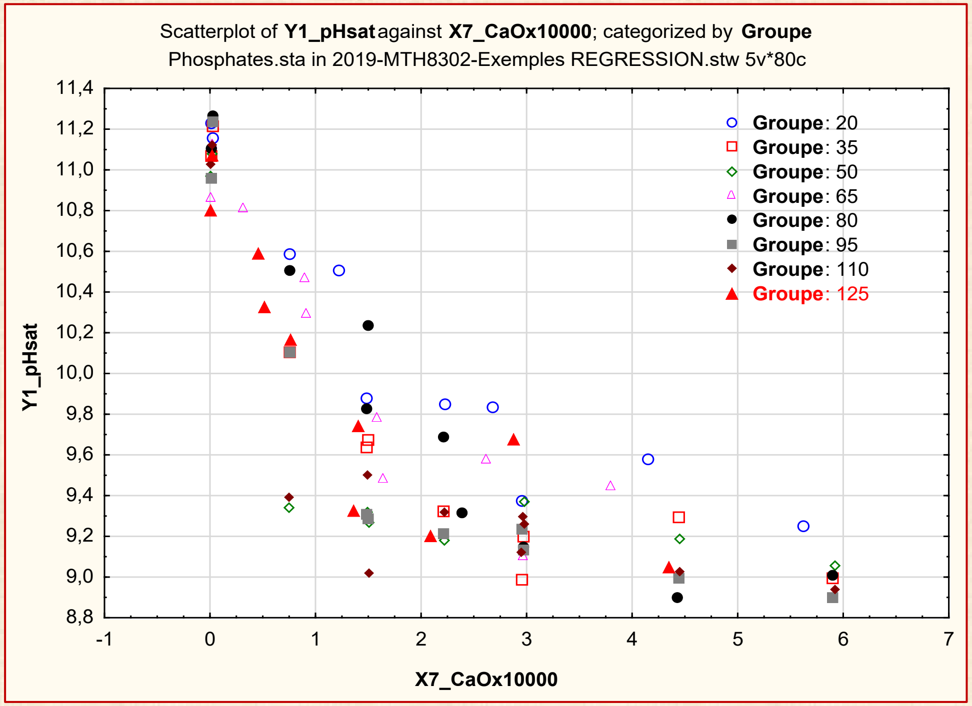
**R-Square :** proportion de la variation de la variable de réponse expliquée par le modèle.
bon modèle : R-Square près de 1

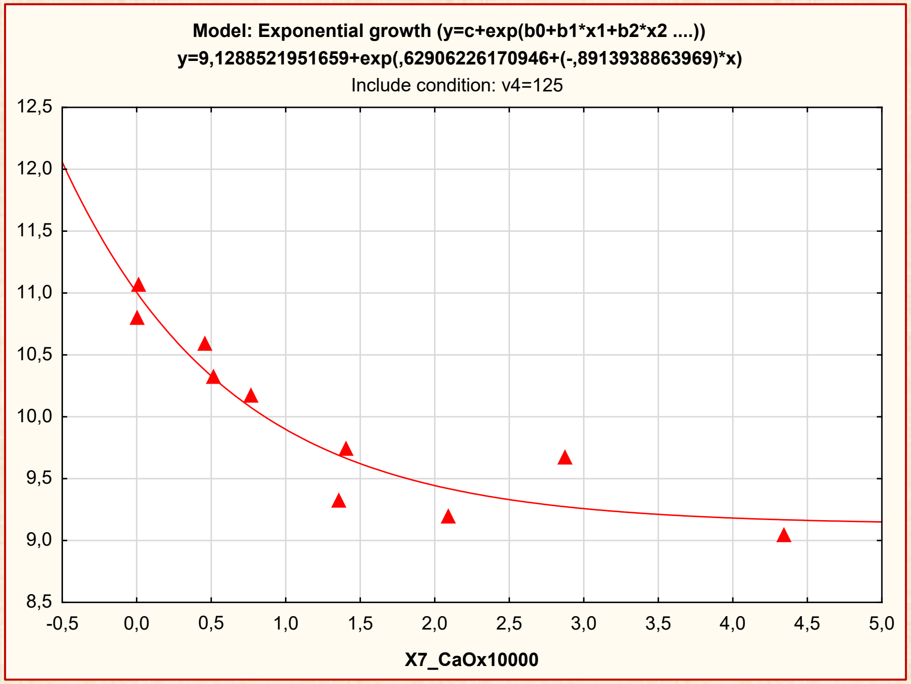
**Remarque** AIC ou AICc préférable à R-square car tient en compte le nombre de paramètres

# Exemple :
## décroissance phosphates

| 1<br>ID | 2<br>X7_CaOx10000 | 3<br>Y1_pHsat | 4<br>Groupe | 5<br>XA_col |
|---|---|---|---|---|
| 1 | 0,018 | 11,16 | 20 | 1 |
| 2 | 1,209 | 10,51 | 20 | 1 |
| 3 | 2,678 | 9,84 | 20 | 1 |
| 4 | 4,145 | 9,58 | 20 | 1 |
| 5 | 5,611 | 9,25 | 20 | 1 |
| 6 | 0,008 | 11,23 | 20 | 2 |
| 7 | 0,748 | 10,59 | 20 | 2 |
| 8 | 1,482 | 9,88 | 20 | 2 |
| 9 | 2,216 | 9,85 | 20 | 2 |
| 10 | 2,949 | 9,38 | 20 | 2 |
| 11 | 0,016 | 11,22 | 35 | 1 |
| 12 | 1,495 | 9,68 | 35 | 1 |
| 13 | 2,962 | 9,20 | 35 | 1 |
| 14 | 4,429 | 9,30 | 35 | 1 |
| 15 | 5,895 | 9,00 | 35 | 1 |
| 16 | 0,006 | 11,07 | 35 | 2 |
| 17 | 0,745 | 10,11 | 35 | 2 |
| 18 | 1,479 | 9,64 | 35 | 2 |
| 19 | 2,213 | 9,33 | 35 | 2 |
| 20 | 2,946 | 8,99 | 35 | 2 |
| 71 | 0,014 | 11,07 | 110 | 1 |
| 72 | 0,515 | 10,33 | 125 | 1 |
| 73 | 1,404 | 9,74 | 125 | 1 |
| 74 | 2,872 | 9,68 | 125 | 1 |
| 75 | 4,343 | 9,05 | 125 | 1 |
| 76 | 0,005 | 10,80 | 125 | 2 |
| 77 | 0,453 | 10,59 | 125 | 2 |
| 78 | 0,770 | 10,17 | 125 | 2 |
| 79 | 1,359 | 9,33 | 125 | 2 |
| 80 | 2,095 | 9,20 | 125 | 2 |

**groupe = 125**



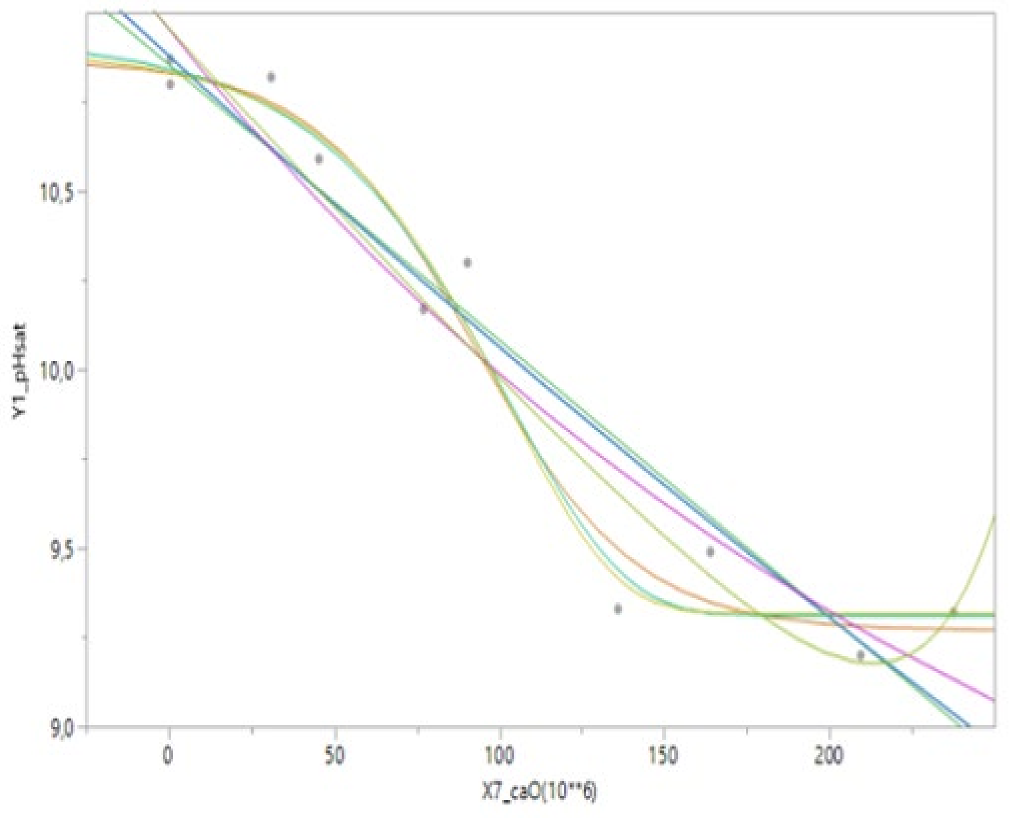



**Exemple : décroissance phosphates**

**calculs avec JMP**

**groupe = 125**

| Modèle | AICc | Poids d'Akaike corrigé (AICc) | 0...1 | BIC | Somme des carrés des écarts (SSE) | Erreur quadratique moyenne (MSE) | RMSE | R carré |
|---|---|---|---|---|---|---|---|---|
| Exponentielle 2 paramètres | 5,6468267 | 0,4565538 | | 2,554582 | 0,3788521 | 0,0473565 | 0,2176155 | 0,9112445 |
| Linéaire | 6,3539489 | 0,3205843 | | 3,2617042 | 0,4066115 | 0,0508264 | 0,2254472 | 0,9047411 |
| Gompertz à 4 paramètres | 9,1808096 | 0,0780005 | | -4,306265 | 0,1203669 | 0,0200611 | 0,1416374 | 0,9718011 |
| Logistique à 4 paramètres | 9,9208518 | 0,0538765 | | -3,566223 | 0,1296124 | 0,0216021 | 0,1469764 | 0,9696351 |
| Exponentielle 3 paramètres | 10,358711 | 0,0432832 | | 3,5690516 | 0,3330639 | 0,0475806 | 0,2181297 | 0,9219715 |
| Logistique à 3 paramètres | 11,648305 | 0,0227137 | | 4,8586452 | 0,3789081 | 0,0541297 | 0,232658 | 0,9112313 |
| Gompertz à 3 paramètres | 11,70708 | 0,0220559 | | 4,9174208 | 0,3811417 | 0,0544488 | 0,2333427 | 0,9107081 |
| Biexponentielle à 4 paramètres | 15,771108 | 0,0028909 | | 2,2840333 | 0,2326591 | 0,0387765 | 0,1969175 | 0,9454938 |
| Logistique à 5 paramètres | 24,26804 | 0,0000413 | | -1,91645 | 0,1214214 | 0,0242843 | 0,1558342 | 0,971554 |




**Meilleur modèle :**

**Exponentiel à 2 paramètres**

| Coefficient | Estimation | Erreur standard |
|---|---|---|
| Échelle | 10,88053 | 0,1125622 |
| Taux de croissance | -0,000782 | 0,0000872 |

# ANALYSE des EXEMPLES avec JMP

| Exemple | data |
|---|---|
| Régression linéaire simple …… | épaisseur |
| Régression logistique ………… | Burns |
| Non linéaire ……………………... | phosphates |

reg simple - JMP Pro

| | id | new2 | AREA | PRICE | new5 | X_PRESSI | Y_EPAIS | ORDRE | RESIDU |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | maison | 134 | 5495 | epaisseur | 3 | 14 | 4 | 1,8 |
| 2 | 2 | | 115 | 4700 | | 3 | 12 | 3 | -0,1999998 |
| 3 | 3 | | 167 | 7500 | | 3 | 13 | 8 | 0,8000002 |
| 4 | 4 | | 185 | 7775 | | 4 | 10 | 2 | -1,1 |
| 5 | 5 | | 84 | 3500 | | 4 | 9 | 12 | -2,1 |
| 6 | 6 | | 98 | 4000 | | 4 | 11 | 14 | -0,1000004 |
| 7 | 7 | | 115 | 4850 | | 5 | 9 | 6 | -1 |
| 8 | 8 | | 185 | 7500 | | 5 | 10 | 15 | 0 |
| 9 | 9 | | 164 | 6900 | | 5 | 11 | 7 | 1 |
| 10 | 10 | | 145 | 5950 | | 6 | 9 | 10 | 0,1000004 |
| 11 | 11 | | 123 | 5010 | | 6 | 8 | 5 | -0,8999996 |
| 12 | 12 | | 128 | 5600 | | 6 | 10 | 1 | 1,1 |
| 13 | 13 | | 167 | 6750 | | 7 | 8 | 9 | 0,1999998 |
| 14 | 14 | | 115 | 5000 | | 7 | 9 | 11 | 1,2 |
| 15 | 15 | | 178 | 7200 | | 7 | 7 | 13 | -0,8000002 |
| 16 | 16 | | 97 | 4290 | | • | • | • | • |
| 17 | 17 | | • | • | | • | • | • | • |
| 18 | 18 | | • | • | | • | • | • | • |
| 19 | 19 | | • | • | | • | • | • | • |

burns - JMP Pro

| | id_patient | Surface | log(Surface+1) | X | 10X | survit? | Y | Yc |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 4,1006532009691 | 1,6293686101503 | 1,6 | 160 | oui | 0 | 1 |
| 2 | 2 | 6,818745162709 | 2,0565240765546 | 2,05 | 205 | oui | 0 | 1 |
| 3 | 3 | 5,4860910214026 | 1,8696600413054 | 1,85 | 185 | oui | 0 | 1 |
| 4 | 4 | 4,0925310011067 | 1,627774956701 | 1,6 | 160 | oui | 0 | 1 |
| 5 | 5 | 6,8411251890701 | 2,0593823430785 | 2,05 | 205 | oui | 0 | 1 |
| 6 | 6 | 6,0806056563879 | 1,9573594487344 | 1,95 | 195 | non | 1 | 0 |
| 7 | 7 | 6,8408159450268 | 2,0593429035677 | 2,05 | 205 | non | 1 | 0 |
| 8 | 8 | 4,767476862857 | 1,7522346994161 | 1,75 | 175 | oui | 0 | 1 |
| 9 | 9 | 7,6260619485925 | 2,1547880798127 | 2,15 | 205 | non | 1 | 0 |
| 10 | 10 | 4,8685594426448 | 1,7696091936104 | 1,75 | 175 | oui | 0 | 1 |
| 11 | 11 | 3,9921514161059 | 1,6078669623591 | 1,6 | 160 | oui | 0 | 1 |
| 12 | 12 | 6,8012507254303 | 2,0542840702551 | 2,05 | 205 | non | 1 | 0 |
| 13 | 13 | 6,8984802474846 | 2,0666703672307 | 2,05 | 205 | oui | 0 | 1 |
| 14 | 14 | 5,4968440465237 | 1,8713165276891 | 1,85 | 185 | oui | 0 | 1 |
| 15 | 15 | 6,0834445538793 | 1,9577603083032 | 1,95 | 195 | non | 1 | 0 |
| 16 | 16 | 9,620563059878 | 2,3627920332292 | 2,35 | 235 | non | 1 | 0 |
| 17 | 17 | 5,4894791057518 | 1,8701822664397 | 1,85 | 185 | oui | 0 | 1 |
| 18 | 18 | 6,1675809577443 | 1,9695682138277 | 1,95 | 195 | oui | 0 | 1 |
| … | | | | | | | | |
| 432 | 432 | 6,1358856048566 | 1,9651363644397 | 1,95 | 195 | non | 1 | 0 |
| 433 | 433 | 6,133507989848 | 1,9648031176181 | 1,95 | 195 | oui | 0 | 1 |
| 434 | 434 | 8,6294283451594 | 2,2648238621747 | 2,25 | 225 | non | 1 | 0 |
| 435 | 435 | 7,6371215838834 | 2,1560693772957 | 2,15 | 215 | oui | 0 | 1 |

Lignes: Toutes les lignes 435, Sélectionnée(s) 0, Exclue(s) 0, Masquée(s) 0, Étiquetée(s) 0

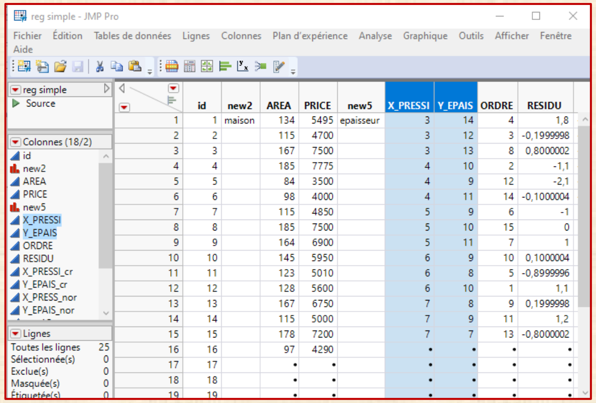
phosphates - JMP Pro

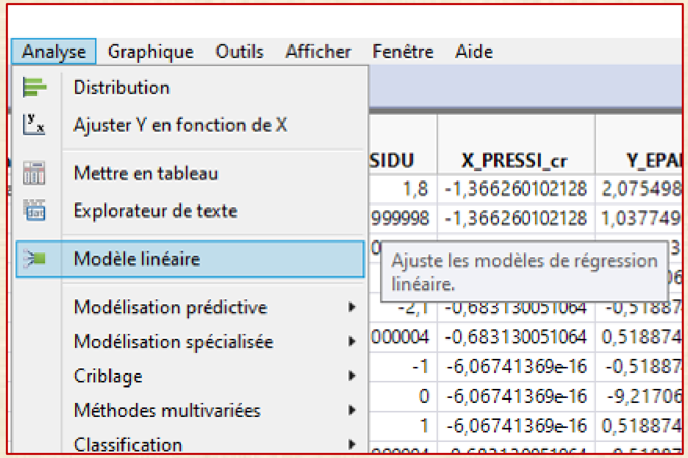
| | ID | X7_CaOx10000 | Y1_pHsat | Groupe | XA_col |
|---|---|---|---|---|---|
| 1 | 1 | 0,018 | 11,16 | 20 | 1 |
| 2 | 2 | 1,209 | 10,51 | 20 | 1 |
| 3 | 3 | 2,678 | 9,84 | 20 | 1 |
| 4 | 4 | 4,145 | 9,58 | 20 | 1 |
| 5 | 5 | 5,611 | 9,25 | 20 | 1 |
| 6 | 6 | 0,008 | 11,23 | 20 | 2 |
| 7 | 7 | 0,748 | 10,59 | 20 | 2 |
| 8 | 8 | 1,482 | 9,88 | 20 | 2 |
| 9 | 9 | 2,216 | 9,85 | 20 | 2 |
| 10 | 10 | 2,949 | 9,38 | 20 | 2 |
| 11 | 11 | 0,016 | 11,22 | 35 | 1 |
| 12 | 12 | 1,495 | 9,68 | 35 | 1 |
| 13 | 13 | 2,962 | 9,2 | 35 | 1 |
| … | | | | | |
| 73 | 73 | 1,404 | 9,74 | 125 | 1 |
| 74 | 74 | 2,872 | 9,68 | 125 | 1 |
| 75 | 75 | 4,343 | 9,05 | 125 | 1 |
| 76 | 76 | 0,005 | 10,8 | 125 | 2 |
| 77 | 77 | 0,453 | 10,59 | 125 | 2 |
| 78 | 78 | 0,77 | 10,17 | 125 | 2 |
| 79 | 79 | 1,359 | 9,33 | 125 | 2 |
| 80 | 80 | 2,095 | 9,2 | 125 | 2 |

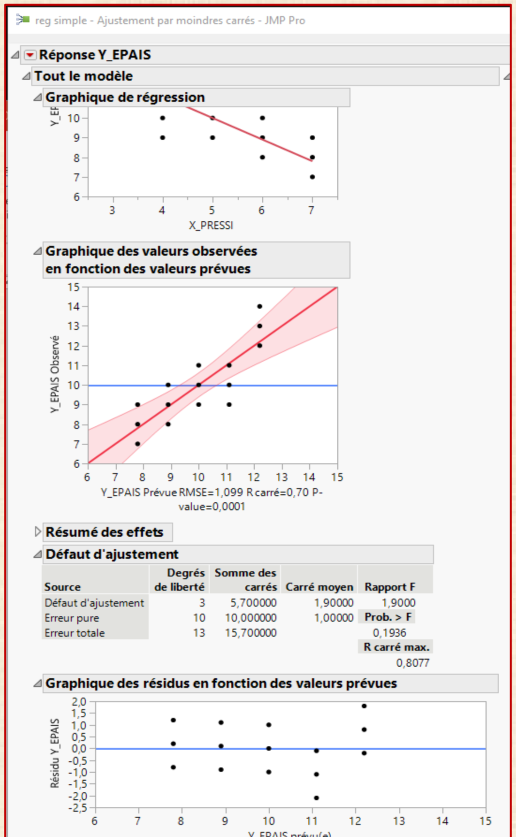
Lignes: Toutes les lignes 80, Sélectionnée(s) 0, Exclue(s) 0, Masquée(s) 0, Étiquetée(s) 0

# ÉPAISSEUR

# ANALYSE des EXEMPLES avec JMP




## Résumé de l'ajustement

| | |
|---|---|
| R carré | 0,698077 |
| R carré ajusté | 0,674852 |
| Racine de l'erreur quadratique moyenne | 1,098951 |
| Moyenne de la réponse | 10 |
| Observations (ou sommes pondérées) | 15 |

## Analyse de la variance

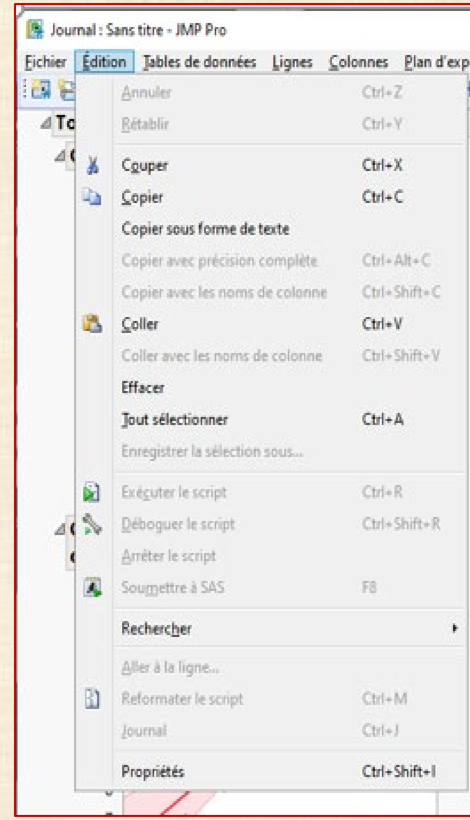
| Source | Degrés de liberté | Somme des carrés | Carré moyen | Rapport F |
|---|---|---|---|---|
| Modèle | 1 | 36,300000 | 36,3000 | 30,0573 |
| Erreur | 13 | 15,700000 | 1,2077 | Prob. > F |
| Total corrigé | 14 | 52,000000 | | 0,0001* |

## Estimations des coefficients

| Terme | Estimation | Erreur standard | t ratio | Prob. > \|t\| |
|---|---|---|---|---|
| Constante | 15,5 | 1,042556 | 14,87 | <,0001* |
| X_PRESSI | -1,1 | 0,20064 | -5,48 | 0,0001* |

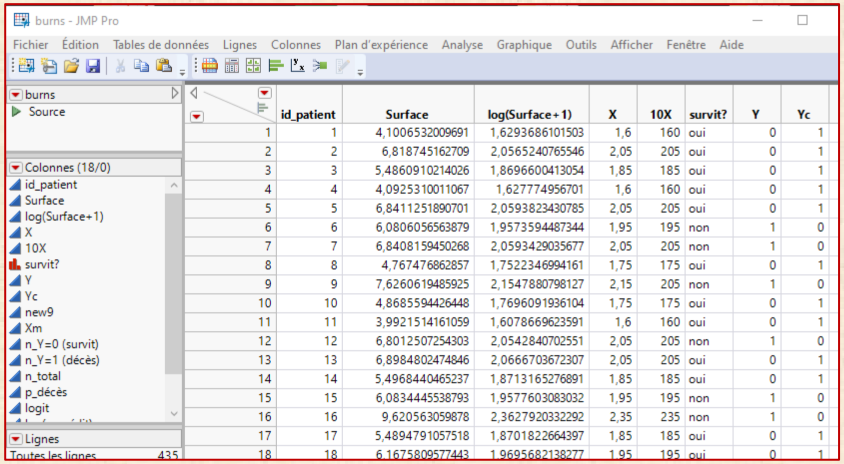
## Tests des effets

| Source | Nombre de coefficients | Degrés de liberté | Somme des carrés | Rapport F | Prob. > F |
|---|---|---|---|---|---|
| X_PRESSI | 1 | 1 | 36,300000 | 30,0573 | 0,0001* |

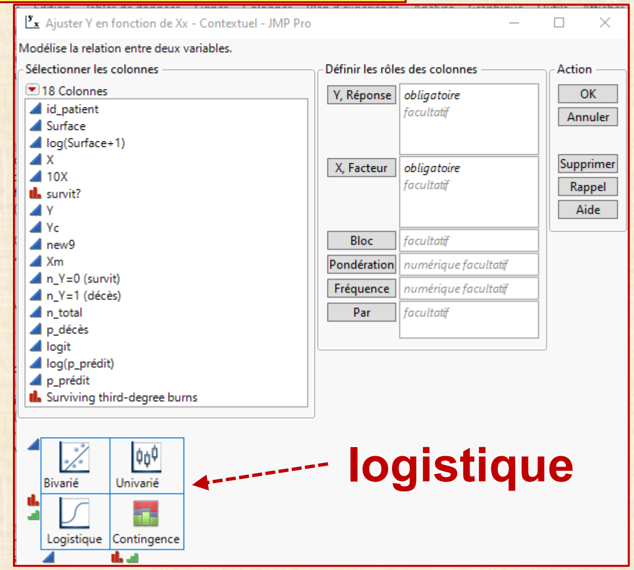


# BURNS

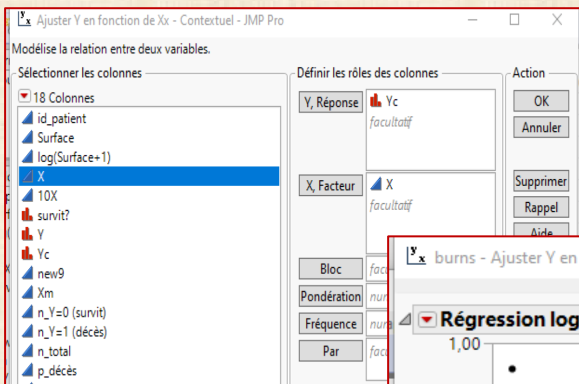
# ANALYSE des EXEMPLES avec JMP




| | id_patient | Surface | log(Surface+1) | X | 10X | survit? | Y | Yc |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 4,1006532009691 | 1,6293686101503 | 1,6 | 160 | oui | 0 | 1 |
| 2 | 2 | 6,818745162709 | 2,0565240765546 | 2,05 | 205 | oui | 0 | 1 |
| 3 | 3 | 5,4860910214026 | 1,8696600413054 | 1,85 | 185 | oui | 0 | 1 |
| 4 | 4 | 4,0925310011067 | 1,627774956701 | 1,6 | 160 | oui | 0 | 1 |
| 5 | 5 | 6,8411251890701 | 2,0593823430785 | 2,05 | 205 | oui | 0 | 1 |
| 6 | 6 | 6,0806056563879 | 1,9573594487344 | 1,95 | 195 | non | 1 | 0 |
| 7 | 7 | 6,8408159450268 | 2,0593429035677 | 2,05 | 205 | non | 1 | 0 |
| 8 | 8 | 4,767476862857 | 1,7522346994161 | 1,75 | 175 | oui | 0 | 1 |
| 9 | 9 | 7,6260619485925 | 2,1547880798127 | 2,15 | 205 | non | 1 | 0 |
| 10 | 10 | 4,8685594426448 | 1,7696091936104 | 1,75 | 175 | oui | 0 | 1 |
| 11 | 11 | 3,9921514161059 | 1,6078669623591 | 1,6 | 160 | oui | 0 | 1 |
| 12 | 12 | 6,8012507254303 | 2,0542840702551 | 2,05 | 205 | non | 1 | 0 |
| 13 | 13 | 6,8984802474846 | 2,0666703672307 | 2,05 | 205 | oui | 0 | 1 |
| 14 | 14 | 5,4968440465237 | 1,8713165276891 | 1,85 | 185 | oui | 0 | 1 |
| 15 | 15 | 6,0834445538793 | 1,9577603083032 | 1,95 | 195 | non | 1 | 0 |
| 16 | 16 | 9,620563059878 | 2,3627920332292 | 2,35 | 235 | non | 1 | 0 |
| 17 | 17 | 5,4894791057518 | 1,8701822664397 | 1,85 | 185 | oui | 0 | 1 |
| 18 | 18 | 6,1675809577443 | 1,9695682138277 | 1,95 | 195 | oui | 0 | 1 |




**logistique**





### Itération(s)

### Tester tout le modèle

| Modèle | -Log-vraisemblance | Degrés de liberté | Khi deux | Prob. > khi deux |
|---|---|---|---|---|
| Différence | 92,89696 | 1 | 185,7939 | <,0001* |
| Complet | 168,00193 | | | |
| Réduit | 260,89889 | | | |

| | |
|---|---|
| R carré (U) | 0,3561 |
| AICc | 340,032 |
| BIC | 348,155 |
| Observations (ou sommes pondérées) | 435 |

### Détails de l'ajustement

### Estimations des coefficients

| Terme | Estimation | Erreur standard | Khi deux | Prob. > khi deux |
|---|---|---|---|---|
| Constante | -22,454537 | 2,2526374 | 99,36 | <,0001* |
| X | 10,5232724 | 1,0752925 | 95,77 | <,0001* |

Pour les logarithmes des odds de non/oui

### Covariance des estimations

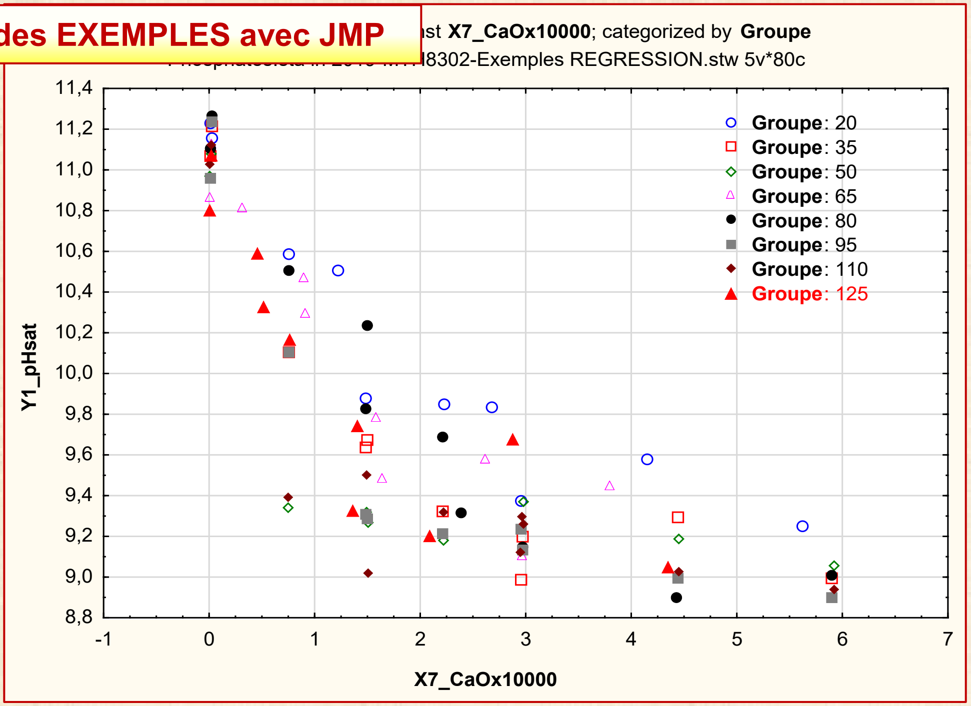
# Exemple :
**décroissance phosphates**

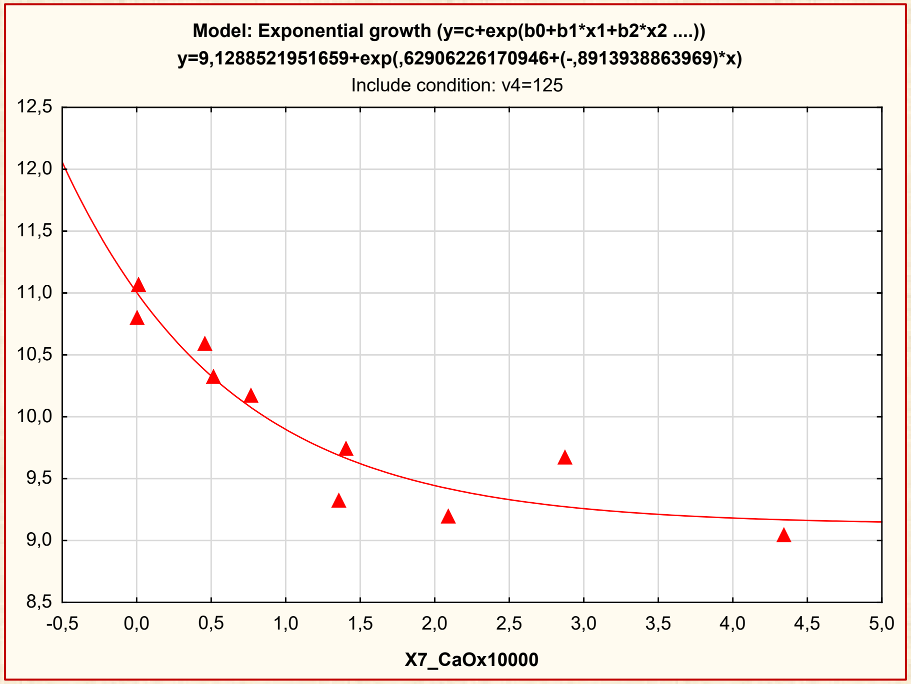
## ANALYSE des EXEMPLES avec JMP

| 1 ID | 2 X7_CaOx10000 | 3 Y1_pHsat | 4 Groupe | 5 XA_col |
|---|---|---|---|---|
| 1 | 0,018 | 11,16 | 20 | 1 |
| 2 | 1,209 | 10,51 | 20 | 1 |
| 3 | 2,678 | 9,84 | 20 | 1 |
| 4 | 4,145 | 9,58 | 20 | 1 |
| 5 | 5,611 | 9,25 | 20 | 1 |
| 6 | 0,008 | 11,23 | 20 | 2 |
| 7 | 0,748 | 10,59 | 20 | 2 |
| 8 | 1,482 | 9,88 | 20 | 2 |
| 9 | 2,216 | 9,85 | 20 | 2 |
| 10 | 2,949 | 9,38 | 20 | 2 |
| 11 | 0,016 | 11,22 | 35 | 1 |
| 12 | 1,495 | 9,68 | 35 | 1 |
| 13 | 2,962 | 9,20 | 35 | 1 |
| 14 | 4,429 | 9,30 | 35 | 1 |
| 15 | 5,895 | 9,00 | 35 | 1 |
| 16 | 0,006 | 11,07 | 35 | 2 |
| 17 | 0,745 | 10,11 | 35 | 2 |
| 18 | 1,479 | 9,64 | 35 | 2 |
| 19 | 2,213 | 9,33 | 35 | 2 |
| 20 | 2,946 | 8,99 | 35 | 2 |

| | | | | |
|---|---|---|---|---|
| 71 | 0,014 | 11,07 | 110 | 1 |
| 72 | 0,515 | 10,33 | 125 | 1 |
| 73 | 1,404 | 9,74 | 125 | 1 |
| 74 | 2,872 | 9,68 | 125 | 1 |
| 75 | 4,343 | 9,05 | 125 | 1 |
| 76 | 0,005 | 10,80 | 125 | 2 |
| 77 | 0,453 | 10,59 | 125 | 2 |
| 78 | 0,770 | 10,17 | 125 | 2 |
| 79 | 1,359 | 9,33 | 125 | 2 |
| 80 | 2,095 | 9,20 | 125 | 2 |

**groupe = 125**



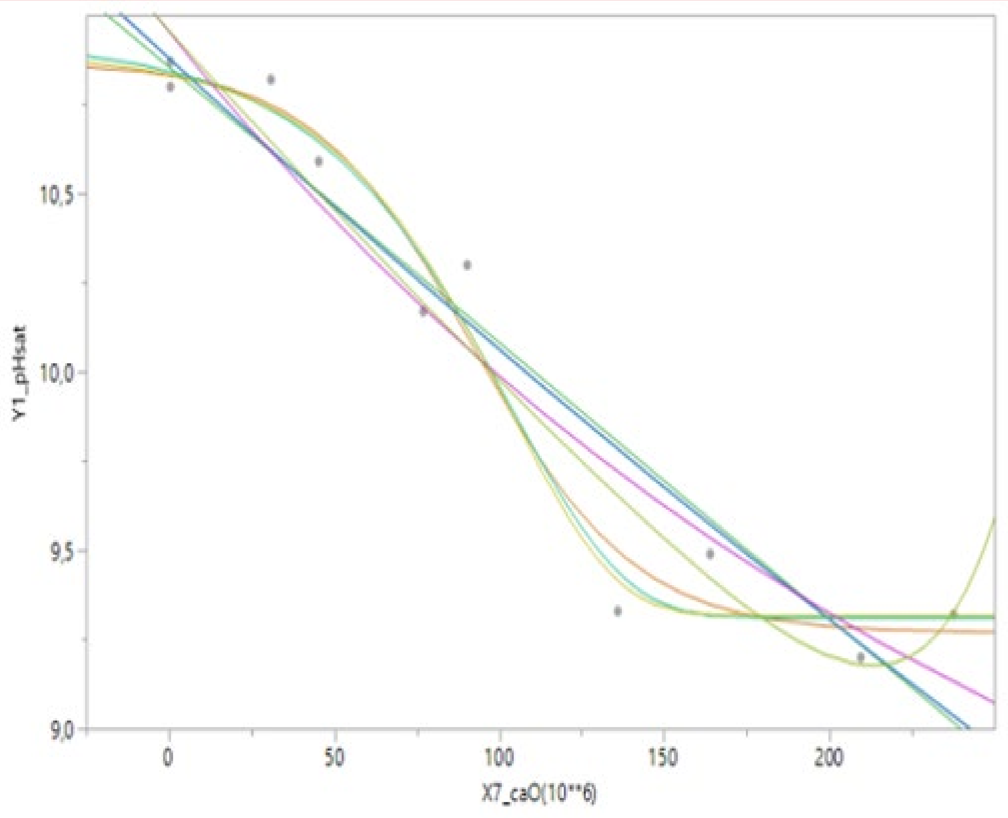



**Exemple : décroissance phosphates**

**calculs avec JMP**

**groupe = 125**

| Modèle | | AICc | Poids d'Akaike corrigé (AICc) | 0...1 | BIC | Somme des carrés des écarts (SSE) | Erreur quadratique moyenne (MSE) | RMSE | R carré |
|---|---|---|---|---|---|---|---|---|---|
| Exponentielle 2 paramètres | — | 5,6468267 | 0,4565538 | | 2,554582 | 0,3788521 | 0,0473565 | 0,2176155 | 0,9112445 |
| Linéaire | — | 6,3539489 | 0,3205843 | | 3,2617042 | 0,4066115 | 0,0508264 | 0,2254472 | 0,9047411 |
| Gompertz à 4 paramètres | — | 9,1808096 | 0,0780005 | | -4,306265 | 0,1203669 | 0,0200611 | 0,1416374 | 0,9718011 |
| Logistique à 4 paramètres | — | 9,9208518 | 0,0538765 | | -3,566223 | 0,1296124 | 0,0216021 | 0,1469764 | 0,9696351 |
| Exponentielle 3 paramètres | — | 10,358711 | 0,0432832 | | 3,5690516 | 0,3330639 | 0,0475806 | 0,2181297 | 0,9219715 |
| Logistique à 3 paramètres | — | 11,648305 | 0,0227137 | | 4,8586452 | 0,3789081 | 0,0541297 | 0,232658 | 0,9112313 |
| Gompertz à 3 paramètres | — | 11,70708 | 0,0220559 | | 4,9174208 | 0,3811417 | 0,0544488 | 0,2333427 | 0,9107081 |
| Biexponentielle à 4 paramètres | — | 15,771108 | 0,0028909 | | 2,2840333 | 0,2326591 | 0,0387765 | 0,1969175 | 0,9454938 |
| Logistique à 5 paramètres | — | 24,26804 | 0,0000413 | | -1,91645 | 0,1214214 | 0,0242843 | 0,1558342 | 0,971554 |




**Meilleur modèle :**

**Exponentiel à 2 paramètres**

| Coefficient | Estimation | Erreur standard |
|---|---|---|
| Échelle | 10,88053 | 0,1125622 |
| Taux de croissance | -0,000782 | 0,0000872 |