

Régression sur données de survie



Révision : 24/09/2023



Résumé.....	1
Entrée des données.....	3
Modèle statistique.....	5
Résumé de l'analyse.....	6
Options d'analyse.....	7
Graphique du modèle ajusté.....	9
Quantiles.....	10
Graphique des quantiles.....	12
Graphique de probabilités des quantiles.....	14
Prévisions.....	16
Graphique des valeurs observées et ajustées.....	17
Graphique de probabilités des résidus.....	17
Résidus non usuels.....	19
Graphiques des résidus.....	20
Matrice des corrélations.....	22
Graphique treillis.....	22
Graphique de la loi des défaillances.....	24
Enregistrement des résultats.....	25
Formules des calculs.....	26

Résumé

La procédure **Régression sur données de survie** est conçue pour ajuster un modèle statistique paramétrique reliant des temps des défaillances à une ou plusieurs variables explicatives. Ces variables explicatives peuvent être quantitatives ou qualitatives. Des modèles du premier ordre et du second ordre peuvent être ajustés, avec ou sans interactions.

La loi des temps des défaillances peut prendre l'une quelconque des formes Weibull, exponentielle, normale, log-normale, logistique, log-logistique ou plus petite valeur extrême. Les temps des défaillances peuvent être censurés ou non.

Les résultats fournis par cette procédure incluent une estimation de la fonction de risque et le calcul de quantiles des temps des défaillances. Des prévisions peuvent être effectuées à partir du modèle ajusté et les résidus non usuels peuvent être détectés.

Exemple de StatFolio : *lifedata reg.sgp*

Données de l'exemple :

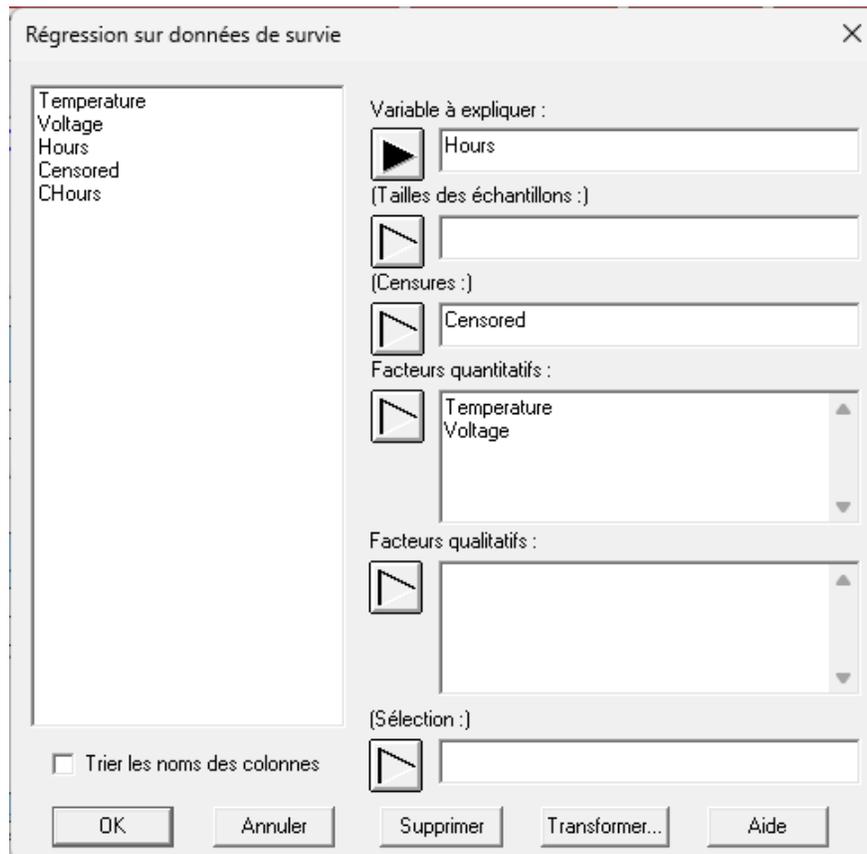
Les données de l'exemple sont contenues dans le StatFolio *lifedata reg.sgp*. Ces données sont constituées des 25 temps des défaillances d'objets soumis à des niveaux variés de deux facteurs d'accélération *temperature* et *voltage*. Les données sont montrées ci-dessous :

<i>Temperature</i>	<i>Voltage</i>	<i>Hours</i>	<i>Censored</i>	<i>CHours</i>
85	6	500	1	>500
85	6	500	1	>500
85	6	500	1	>500
85	6	480	0	480
85	6	475	0	475
85	8	350	0	350
85	8	325	0	325
85	8	315	0	315
85	8	330	0	330
85	8	310	0	310
45	12	500	1	>500
45	12	500	1	>500
45	12	475	0	475
45	12	495	0	495
45	12	450	0	450
65	12	250	0	250
65	12	230	0	230
65	12	245	0	245
65	12	210	0	210
65	12	200	0	200
85	12	60	0	60
85	12	55	0	55
85	12	70	0	70
85	12	65	0	65
85	12	55	0	55

Les données ont été collectées pour *Temperature* = 45, 65 et 85 degrés Celsius et pour *Voltage* = 6, 8 et 12 volts. *Hours* contient le temps jusqu'à la défaillance de l'objet ou 500 si l'objet n'est pas défaillant après 500 heures. *Censored* est défini à 0 pour les temps des défaillances réels ou à 1 si l'objet n'a pas été défaillant après 500 heures. Une colonne additionnelle intitulée *CHours* est également présente. Cette colonne est une colonne de données numériques censurées. Dans cette colonne, les données censurées à droite sont représentées par la notation >500.

Entrée des données

La boîte de dialogue d'entrée des données demande des informations sur les temps des défaillances et les variables explicatives :



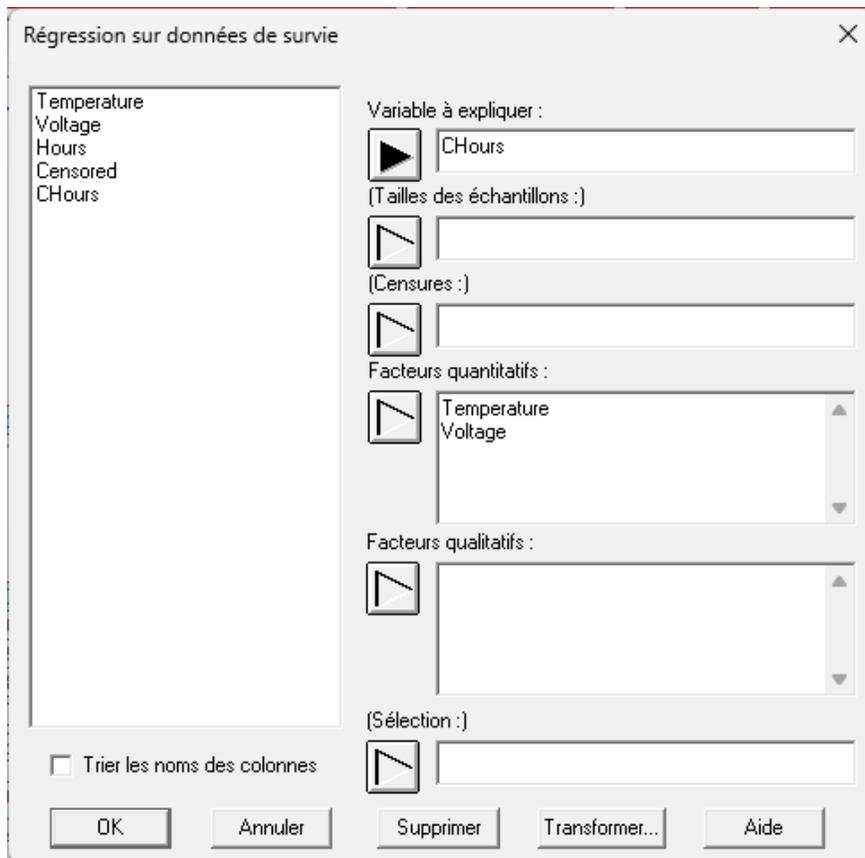
- **Variable à expliquer** : une variable numérique Y contenant les temps des défaillances (pour des données non censurées) ou des temps de censures (pour des données censurées). Elle peut être de deux types :

Une colonne numérique contenant les temps des défaillances. Si des données censurées sont présentes, la colonne des temps des défaillances doit être associée à une seconde colonne contenant les indicateurs des censures.

Une colonne spéciale numérique censurée contenant à la fois les temps des défaillances et les indicateurs des censures. Si une censure par intervalle est présente, cette option doit être utilisée. Les données des censures par intervalles utilisent une notation telle que [200 ; 250] pour indiquer l'intervalle contenant le temps de défaillance réel.

- **(Censures)** : une colonne optionnelle indiquant si une donnée est censurée ou non. Entrer 0 si la valeur de la variable à expliquer représente un temps de défaillance non censuré. Entrer 1 si la valeur est censurée à droite (le vrai temps de défaillance est supérieur à la valeur entrée) ou -1 si la valeur est censurée à gauche (le vrai temps de défaillance est inférieur à la valeur entrée). Ce champ peut être vide si la variable à expliquer est une colonne numérique censurée. *
- **Facteurs quantitatifs** : colonnes numériques contenant les valeurs des facteurs quantitatifs à inclure dans le modèle.
- **Facteurs qualitatifs** : colonnes numériques ou non numériques contenant les niveaux des facteurs qualitatifs à inclure dans le modèle.
- **Sélection** : sélection d'un sous-ensemble.

Si une colonne numérique censurée est utilisée, la boîte de dialogue d'entrée des données doit être renseignée comme montré ci-dessous :



Modèle statistique

STATGRAPHICS ajuste deux types de modèles paramétriques de régression sur données de survie : modèles de régression position et échelle, modèles de régression log-position et log-échelle.

Modèles position et échelle

Pour ce type de modèle, les quantiles de la loi des temps de survie sont reliés aux variables explicatives par une fonction linéaire de la forme :

$$Y_p = \mu + \Phi^{-1}(p)\sigma = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \Phi^{-1}(p)\sigma \quad (1)$$

où μ est le paramètre de position qui dépend des variables explicatives, σ est le paramètre d'échelle et $\Phi^{-1}(p)$ est l'inverse de la fonction de répartition standardisée des temps des défaillances, c'est-à-dire :

$$F(Y) = \Phi\left(\frac{Y - \mu}{\sigma}\right) \quad (2)$$

Dans un tel modèle, les temps des défaillances sont supposés suivre une loi normale ou une loi logistique ou une loi à plus petite valeur extrême.

Modèles log-position et log-échelle

Pour ce type de modèle, les quantiles de la loi des temps des défaillances sont reliés aux variables explicatives par une fonction log-linéaire de la forme :

$$\log(Y_p) = \mu + \Phi^{-1}(p)\sigma = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \Phi^{-1}(p)\sigma \quad (3)$$

où :

$$F(Y) = \Phi\left(\frac{\log(Y) - \mu}{\sigma}\right) \quad (4)$$

Dans un tel modèle, les temps des défaillances sont supposés suivre une loi log-normale, une loi log-logistique, une loi de Weibull ou une loi exponentielle.

Résumé de l'analyse

Le *Résumé de l'analyse* affiche le modèle estimé et les tests sur les rapports de vraisemblance pour la significativité des coefficients du modèle.

<u>Régression sur données de survie - Hours</u>				
Variable à expliquer: Hours				
Censures : Censored				
Facteurs:				
Temperature (degrees C)				
Voltage				
Nombre de valeurs non censurées: 20				
Nombre de valeurs censurées à droite: 5				
Nombre de valeurs censurées à gauche: 0				
Nombre de valeurs censurées par intervalle : 0				
Modèle estimé de régression - Weibull				
Paramètre	Estimation	Erreur-type	LC inf. à 95,0%	LC sup. à 95,0%
CONSTANTE	13,3606	0,511737	12,3576	14,3636
Temperature	-0,0547279	0,00414181	-0,0628457	-0,0466101
Voltage	-0,370413	0,0228475	-0,415193	-0,325633
SIGMA	0,134798	0,0227323	0,0968579	0,187598
Log de la vraisemblance = -100,554				
Tests sur les rapports de vraisemblance				
Facteur	Khi-carré	Ddl	Proba.	
Temperature	69,5073	1	0,0000	
Voltage	72,421	1	0,0000	

Le tableau inclut les informations suivantes :

- **Résumé des données** : un résumé des données en entrée, incluant les nombre n d'observations utilisées pour ajuster le modèle.
- **Modèle estimé de régression** : estimations des coefficients du modèle de régression, avec erreurs-types et intervalles de confiance approximés.
- **Tests sur les rapports de vraisemblance** : tests effectués pour déterminer si les coefficients sont ou non significativement différents de 0. Les probabilités bilatérales sont affichées. De petites valeurs des probabilités (inférieures à 0,05 pour un niveau de signification de 5%) correspondent à des coefficients statistiquement significatifs.

Le tableau ci-dessus affiche les résultats de l'ajustement d'un modèle du premier ordre pour les données de nos condensateurs, en supposant une loi de Weibull pour les temps des défaillances pour des valeurs données des variables explicatives. Le modèle estimé a pour paramètres :

$$\mu = 13,3606 - 0,0547279 \text{ temperature} - 0,370413 \text{ voltage} \quad (5)$$

$$\sigma = 0,134798 \quad (6)$$

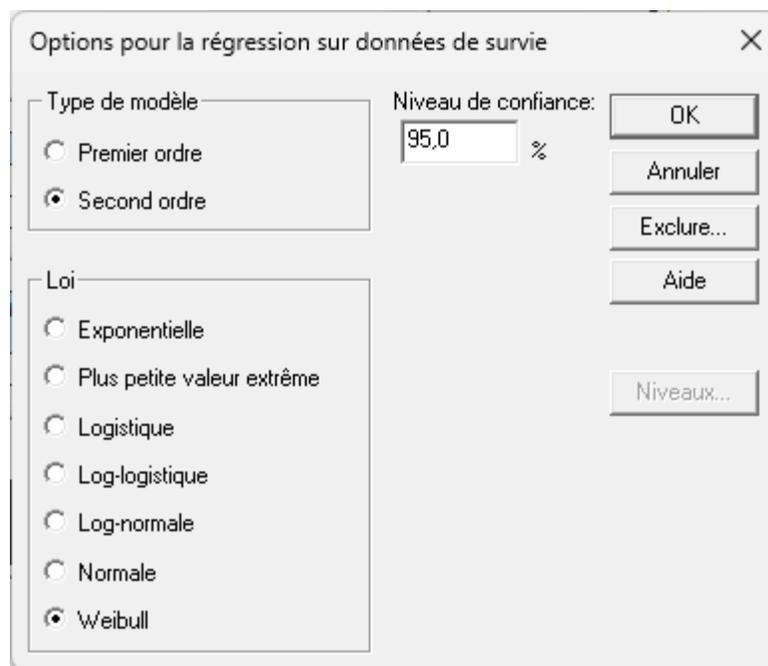
en se basant sur le modèle log-linéaire. L'équation pour le *pième* quantile est :

$$\text{hours}_p = \exp(13,3606 - 0,0547279 \text{ temperature} - 0,370413 \text{ voltage} + 0,134798 \log(-\log(1-p))) \quad (7)$$

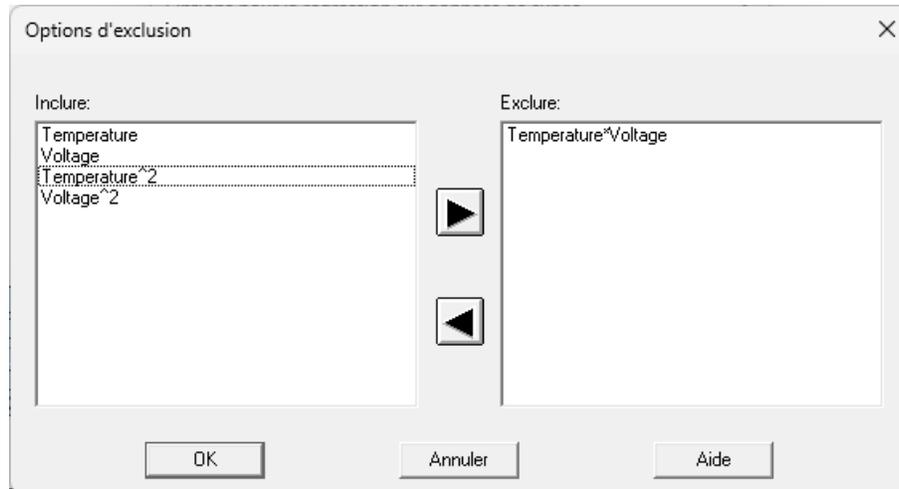
A la fois *voltage* et *temperature* ont un fort effet négatif sur les durées de vie des objets.

Options d'analyse

Le modèle statistique à ajuster est précisé dans les *Options d'analyse* :



- **Type de modèle** : sélectionner *Premier ordre* pour ajuster un modèle utilisant uniquement les effets directs de chaque facteur. Sélectionner *Second ordre* pour inclure les effets quadratiques pour les facteurs quantitatifs et les interactions d'ordre 2 entre toutes les variables.
- **Loi** : la loi supposée des temps des défaillances pour des valeurs données des variables explicatives.
- **Niveau de confiance** : pourcentage de confiance pour l'intervalle d'estimation des coefficients du modèle.
- **Niveaux** : cliquer sur ce bouton pour réordonner les niveaux des facteurs qualitatifs.
- **Exclure** : cliquer sur ce bouton pour exclure des termes spécifiques du modèle. Une boîte de dialogue s'affiche alors.



Double cliquer sur un effet pour le déplacer du champ *Inclure* dans le champ *Exclure* et vice-versa.

Exemple : Ajuster un modèle quadratique

Pour ajuster un modèle quadratique, sélectionner *Second Ordre* dans la boîte de dialogue des *Options d'analyse* puis cliquer sur le bouton *Exclure* pour supprimer le terme croisé *voltage*temperature* qui ne peut pas être estimé car les faibles voltages n'ont été utilisés qu'avec les faibles températures. Les résultats de l'ajustement sont montrés ci-dessous :

Régression sur données de survie - Hours

Variable à expliquer: Hours
 Censures : Censored
 Facteurs:
 Temperature (degrees C)
 Voltage

Nombre de valeurs non censurées: 20
 Nombre de valeurs censurées à droite: 5
 Nombre de valeurs censurées à gauche: 0
 Nombre de valeurs censurées par intervalle : 0

Modèle estimé de régression - Weibull

Paramètre	Estimation	Erreur-type	LC inf. à 95,0%	LC sup. à 95,0%
CONSTANTE	7,98315	0,607048	6,79336	9,17295
Temperature	0,0361367	0,0121787	0,0122668	0,0600066
Voltage	0,190469	0,100988	-0,00746355	0,388402
Temperature^2	-0,000673033	0,0000916152	-0,000852596	-0,00049347
Voltage^2	-0,029916	0,00533424	-0,040371	-0,0194611
SIGMA	0,0632086	0,0115833	0,0441356	0,0905239

Log de la vraisemblance = -87,8563

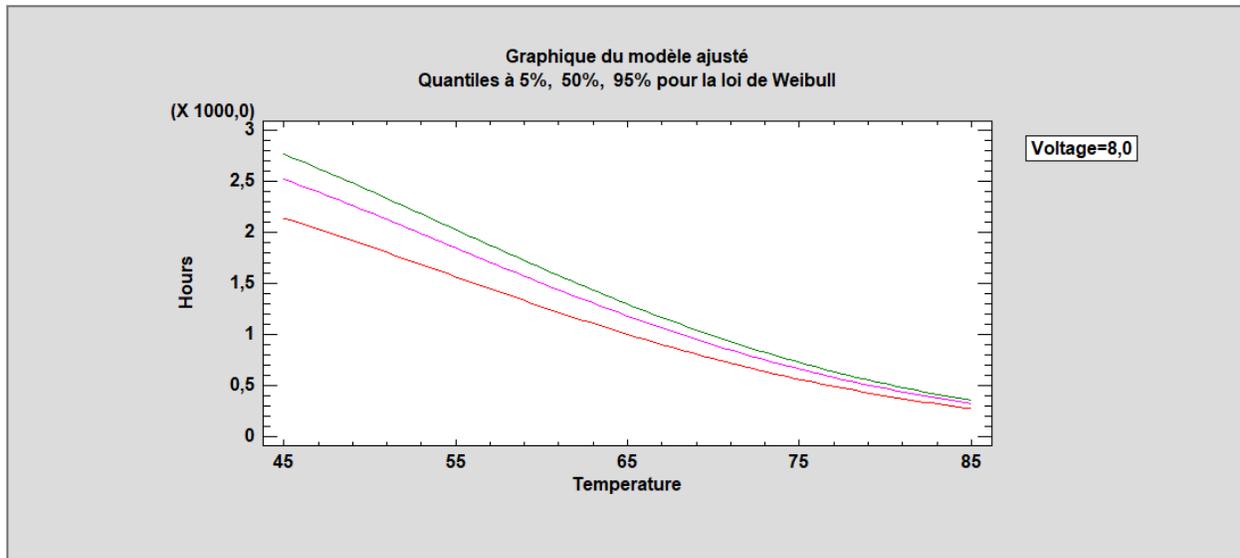
Tests sur les rapports de vraisemblance

Facteur	Khi-carré	Ddl	Proba.
Temperature	6,33446	1	0,0118
Voltage	2,57768	1	0,1084
Temperature^2	21,9476	1	0,0000
Voltage^2	12,6421	1	0,0004

Les tests sur les rapports de vraisemblance pour les termes quadratiques indiquent la présence d'une courbure significative dans les données.

Graphique du modèle ajusté

La fenêtre *Graphique du modèle ajusté* affiche les quantiles en fonction d'une unique variable X, toutes les autres étant maintenues à des valeurs fixes.



Par exemple, le graphique ci-dessus affiche les quantiles à 5%, 50% et 95% en fonction de la *temperature* pour une valeur de *voltage* égale à 8. Les temps des défaillances et la variabilité décroissent lorsque la température augmente.

Options pour la fenêtre

- **Facteur** : sélectionner un facteur à afficher sur l'axe horizontal et indiquer des limites inférieure et supérieure pour le graphique. Pour tous les autres facteurs, préciser les valeurs auxquelles ils sont fixés.
- **Quantiles** : pourcentages désirés des quantiles.
- **Afficher la moyenne** : inclure une ligne pour la moyenne estimée des temps des défaillances.
- **Afficher les résultats à** : si coché, affiche les statistiques estimées pour la valeur indiquée de X.
- **Suivant** ou **Précédent** : à utiliser pour afficher d'autres facteurs lorsqu'il y en a plus de 16.

Options pour le graphique du modèle ajusté

	Bas	Haut	Maintenir
<input checked="" type="radio"/> Temperature	45,0	85,0	45,0
<input type="radio"/> Voltage	6,0	12,0	8,0
<input type="radio"/>			

Quantiles

5,0	%
50,0	%
95,0	%
	%
	%

Afficher la moyenne

Afficher les résultats à :

0,0

OK

Annuler

Aide

Suivant

Précédent

Quantiles

La fenêtre *Quantiles* affiche un tableau des quantiles estimés pour des valeurs données des variables explicatives. Elle affiche également le MTTF (temps moyen avant défaillance).

Des intervalles de confiance, basés sur une approximation normale pour de grands échantillons, sont affichés. Par exemple, pour voltage = 8 et température = 45, il est estimé que 50% des objets seront défaillants après approximativement 2521 heures.

L'intervalle de confiance à 95% pour ce quantile à 50% s'étend de 2266 heures à 2805 heures.

A noter qu'il s'agit d'une extrapolation en-dehors de la plage des données mesurées comme cela pourrait être fait lors d'un test de durée de vie accéléré.

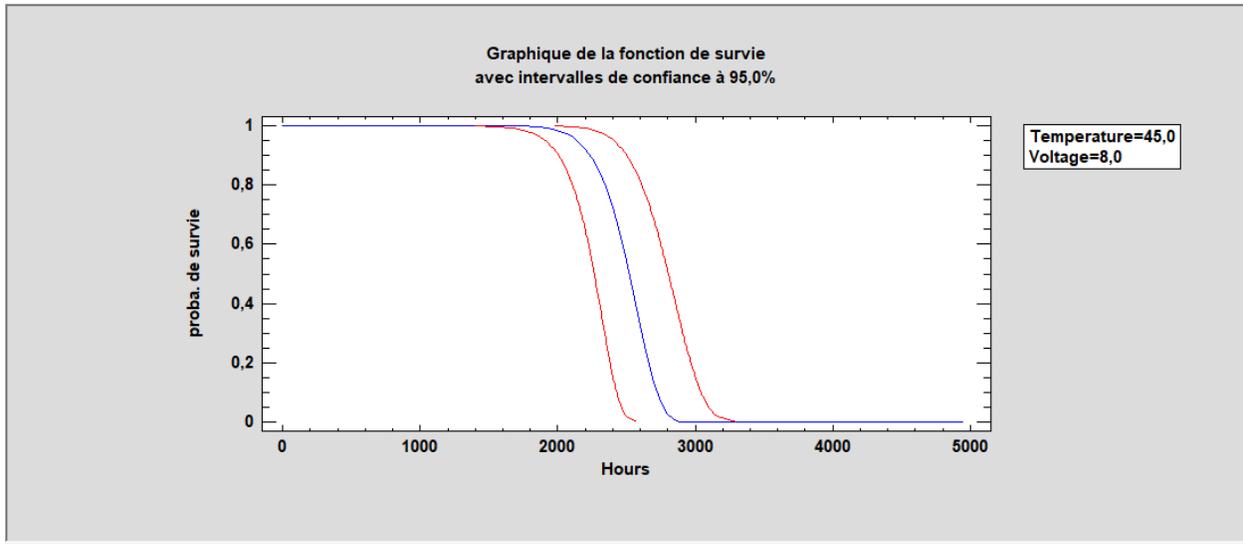
Quantiles des temps de défaillance pour Hours

Temperature=45,0

Voltage=8,0

MTTF: 2495,51

Pourcentage	Quantile	Erreur type	LC inf. à 95,0%	LC sup. à 95,0%
0,1	1667,28	147,51	1401,84	1982,97
0,5	1846,05	137,696	1594,97	2136,66
1,0	1929,04	133,67	1684,06	2209,65
2,0	2016,08	130,148	1776,47	2288,0
3,0	2069,08	128,474	1832,0	2336,85
4,0	2107,74	127,525	1872,04	2373,11
5,0	2138,38	126,954	1903,48	2402,26
6,0	2163,88	126,611	1929,43	2426,82
7,0	2185,8	126,417	1951,55	2448,16
8,0	2205,07	126,326	1970,87	2467,1
9,0	2222,3	126,312	1988,03	2484,19
10,0	2237,92	126,354	2003,48	2499,79
15,0	2300,08	127,067	2064,04	2563,11
20,0	2346,63	128,199	2108,35	2611,84
25,0	2384,62	129,517	2143,81	2652,47
30,0	2417,24	130,939	2173,76	2688,0
35,0	2446,26	132,43	2199,99	2720,09
40,0	2472,75	133,978	2223,62	2749,79
45,0	2497,46	135,582	2245,37	2777,85
50,0	2520,92	137,247	2265,77	2804,8
55,0	2543,57	138,985	2285,24	2831,1
60,0	2565,79	140,813	2304,12	2857,16
65,0	2587,94	142,756	2322,74	2883,43
70,0	2610,45	144,85	2341,44	2910,37
75,0	2633,82	147,151	2360,64	2938,62
80,0	2658,79	149,748	2380,91	2969,1
85,0	2686,57	152,804	2403,17	3003,39
90,0	2719,66	156,666	2429,3	3044,73
91,0	2727,37	157,599	2435,33	3054,43
92,0	2735,61	158,611	2441,75	3064,84
93,0	2744,53	159,722	2448,68	3076,14
94,0	2754,33	160,96	2456,25	3088,58
95,0	2765,28	162,369	2464,67	3102,55
96,0	2777,87	164,017	2474,3	3118,67
97,0	2792,94	166,032	2485,76	3138,07
98,0	2812,32	168,688	2500,39	3163,16
99,0	2841,47	172,816	2522,16	3201,2
99,5	2866,76	176,525	2540,84	3234,49
99,9	2915,23	183,947	2576,11	3299,01



Options pour la fenêtre

Options pour les niveaux des facteurs

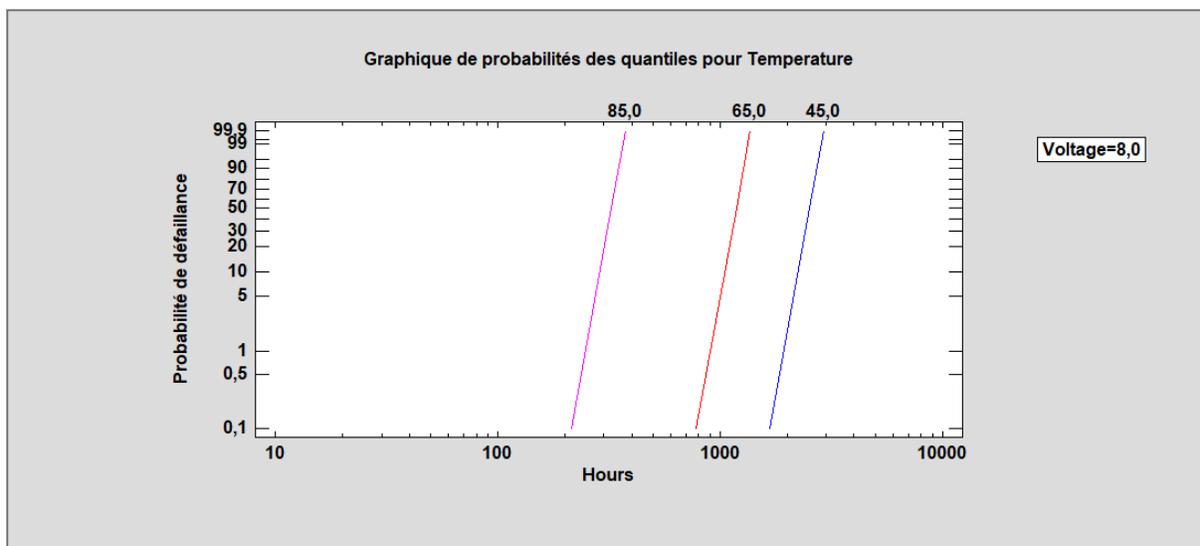
Facteur:	Niveau:	Niveau de confiance:	OK
Temperature	45,0	95,0 %	Annuler
Voltage	8,0		Aide
		Graphique	Suivant
		<input type="radio"/> Répartition	Précédent
		<input checked="" type="radio"/> Survie	

- **Niveau** : valeurs des variables explicatives pour lesquelles les quantiles sont estimés.

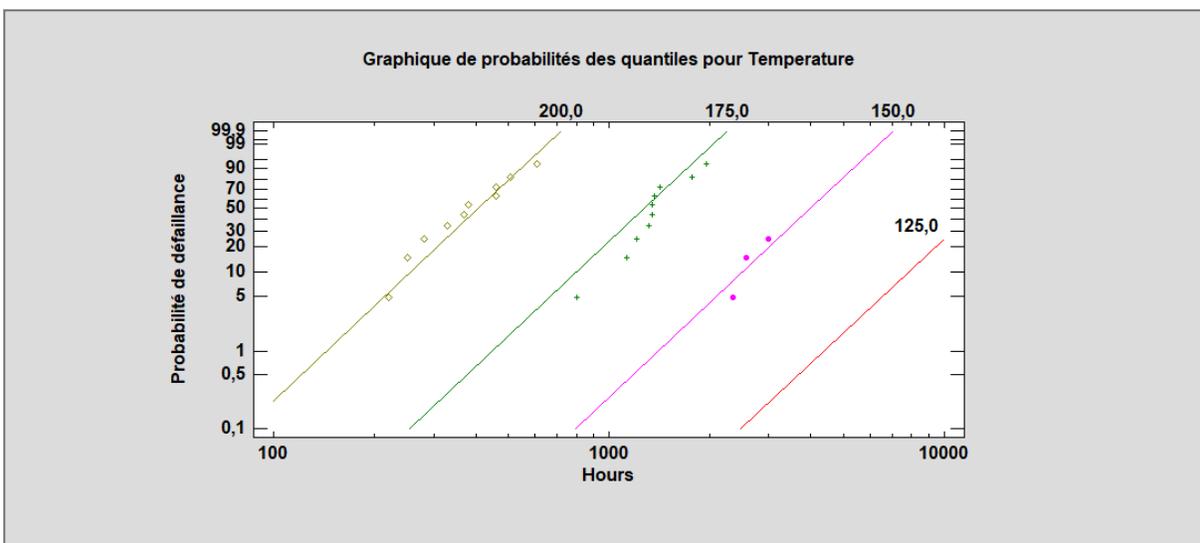
- **Niveau de confiance** : pourcentage de confiance pour l'intervalle d'estimation des quantiles.
- **Graphique** : sélectionner *Répartition* pour afficher les quantiles ou *Survie* pour afficher les probabilités estimées de survie.
- **Suivant** ou **Précédent** : à utiliser pour afficher d'autres facteurs lorsqu'il y en a plus de 16.

Graphique de probabilités des quantiles

Cette fenêtre affiche les quantiles estimés dans un graphique dont l'échelle est telle que la fonction de répartition est une ligne droite. Des lignes sont tracées pour chaque valeur observée de la variable sélectionnée.



Si le modèle contient une unique variable, les valeurs des observations non censurées sont également affichées. Un exemple, pour des données différentes, est montré ci-dessous.



Prévisions

La fenêtre *Prévisions* affiche des prévisions pour le modèle ajusté. Par défaut, le tableau inclut une ligne pour chaque observation dans le tableur pour laquelle toutes les informations pour les variables X sont présentes et l'information pour la variable Y est manquante. Cela permet d'ajouter des lignes dans le tableur qui correspondent à des niveaux pour lesquels vous désirez des prévisions sans affecter le modèle ajusté.

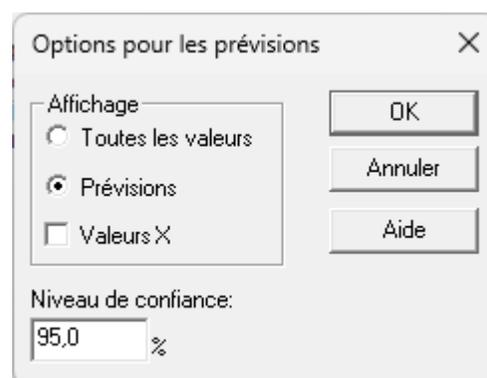
Par exemple, supposons qu'une prévision soit désirée pour un objet sujet à un voltage de 8 et à une température de 45. Dans la ligne 26 du tableur, ajouter ces valeurs tout en laissant la colonne *Hours* vide. Le tableau obtenu est montré ci-dessous.

Ligne	Ajusté	Erreur type	LC inf. à 95,0% pour la valeur ajustée	LC sup. à 95,0% pour la valeur ajustée
26	2580,0	0,0550564	2316,09	2873,98

Les informations suivantes sont présentes dans ce tableau :

- **Ligne** – le numéro de la ligne dans le tableur.
- **Observé** – si présente, la valeur observée Y_i .
- **Ajusté** – la valeur ajustée, donnée par $\hat{\mu}_i$ pour les modèles position-échelle et par $\exp(\hat{\mu}_i)$ pour les modèles log-position et log-échelle.
- **Erreur-type** – l'erreur-type associée à $\hat{\mu}_i$.
- **Intervalles de confiance** – les limites de confiance approximées pour les valeurs ajustées.

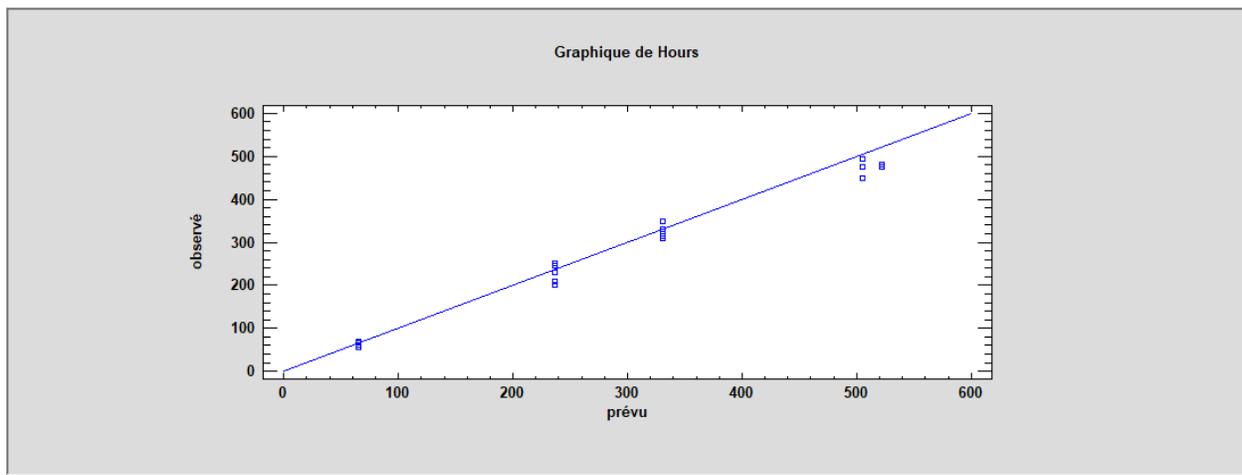
Options pour la fenêtre



- **Affichage** : Toutes les valeurs peuvent être affichées ou uniquement les *Prévisions* (c'est-à-dire les lignes pour lesquelles la valeur de la variable à expliquer est manquante).
- **Valeurs X** : permet de préciser si les valeurs des variables explicatives sont affichées ou non.
- **Niveau de confiance** : pourcentage de confiance pour l'intervalle de confiance des valeurs estimées.

Graphique des valeurs observées et ajustées

La fenêtre *Graphique des valeurs observées et ajustées* affiche les temps observés des défaillances Y_i par rapport à $\hat{\mu}_i$ pour les modèles position-échelle et à $\exp(\hat{\mu}_i)$ pour les modèles log-position et log-échelle.



Si le modèle s'ajuste bien, les points doivent être aléatoirement répartis autour de la ligne diagonale.

Graphique de probabilités des résidus

Comme dans toute procédure de régression, il est important de calculer et de visualiser graphiquement les résidus.

La procédure *Régression sur données de survie* affiche trois types différents de résidus :

1. Résidus ordinaires :

$$\text{pour les modèles position-échelle : } r_i = y_i - \hat{\mu}_i \quad (8)$$

$$\text{pour les modèles log-position et log-échelle : } r_i = y_i - \exp(\hat{\mu}_i) \quad (9)$$

2. *Résidus standardisés* :

pour les modèles position-échelle : $e_i = \frac{y_i - \hat{\mu}}{\hat{\sigma}}$ (10)

pour les modèles log-position et log-échelle : $e_i = \exp\left(\frac{\ln(y_i) - \hat{\mu}_i}{\hat{\sigma}}\right)$ (11)

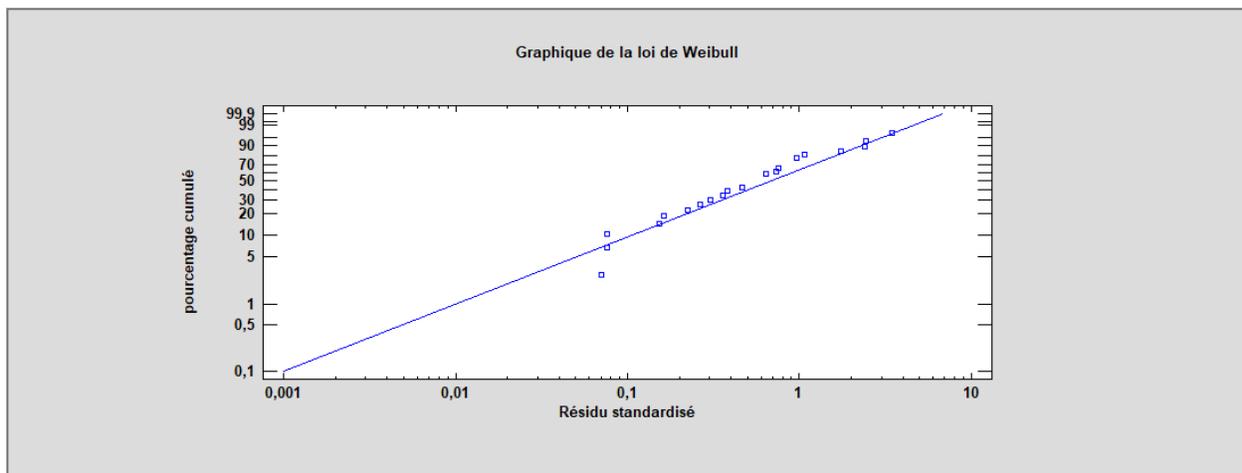
3. *Résidus de Cox-Snell* – un type de résidus de Cox-Snell, compris entre 0 et 1, défini par :

$$\hat{u}_i = \hat{F}(Y_i) \quad (12)$$

qui est la loi estimée des temps cumulés des défaillances évaluée à la valeur observée du temps de défaillance.

Les *résidus ordinaires* mesurent les différences entre les valeurs observées et les valeurs ajustées. Les *résidus standardisés* ont une échelle telle qu'ils doivent suivre une forme standardisée de la loi supposée des temps des défaillances. Les résidus de Cox-Snell peuvent être utiles pour identifier les points extrêmes.

La fenêtre *Graphique de probabilités des résidus* affiche les résidus standardisés dans un graphique conçu pour aider à déterminer si la loi supposée des temps de survie est raisonnable pour les données :



Si la loi supposée est adéquate pour les données, les points doivent s'aligner le long de la ligne diagonale de référence.

Résidus non usuels

La fenêtre *Résidus non usuels* liste toutes les observations qui ont des résidus inhabituellement grands.

Ligne	Y	Y prévu	Résidu	Résidu standardisé	Résidu Cox-Snell
6	350,0	330,71	19,2895	2,45	0,9139
16	250,0	236,56	13,4396	2,40	0,9090
23	70,0	64,7051	5,29494	3,47	0,9689

Le tableau affiche :

- *Ligne* – le numéro de la ligne dans le tableau.
- *Y* – la valeur observée du temps de défaillance (éventuellement censurée).
- *Prévu* – la valeur ajustée, donnée par $\hat{\mu}_i$ pour les modèles position-échelle et $\exp(\hat{\mu}_i)$ pour les modèles log-position et log-échelle.
- *Résidu* – le résidu ordinaire.
- *Résidu standardisé* – le résidu standardisé e_i .
- *Résidu de Cox-Snell* – le résidu de Cox-Snell \hat{u}_i .

Une ligne est affichée dans cette liste si le résidu de Cox-Snell est inférieur à 0,025 ou supérieur à 0,975, c'est-à-dire si le résidu est en-dehors des 95% centraux de la loi estimée des temps des défaillances. Une attention particulière doit donc être apportée à tout résidu en-dehors de l'intervalle :

$$0,00135 \leq \hat{u}_i \leq 0,99865$$

car cela est équivalent à être au-delà de 3 écarts-types si la loi est normale.

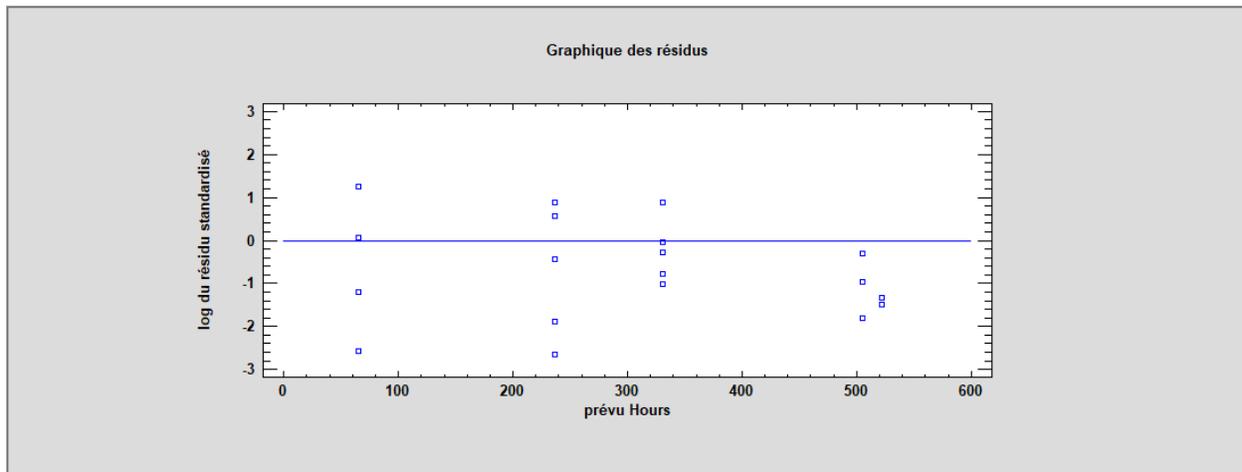
Une ligne est également affichée dans le tableau si sa valeur absolue du résidu standardisé est supérieure à 2.

Graphiques des résidus

Plusieurs types de graphiques des résidus peuvent être créés :

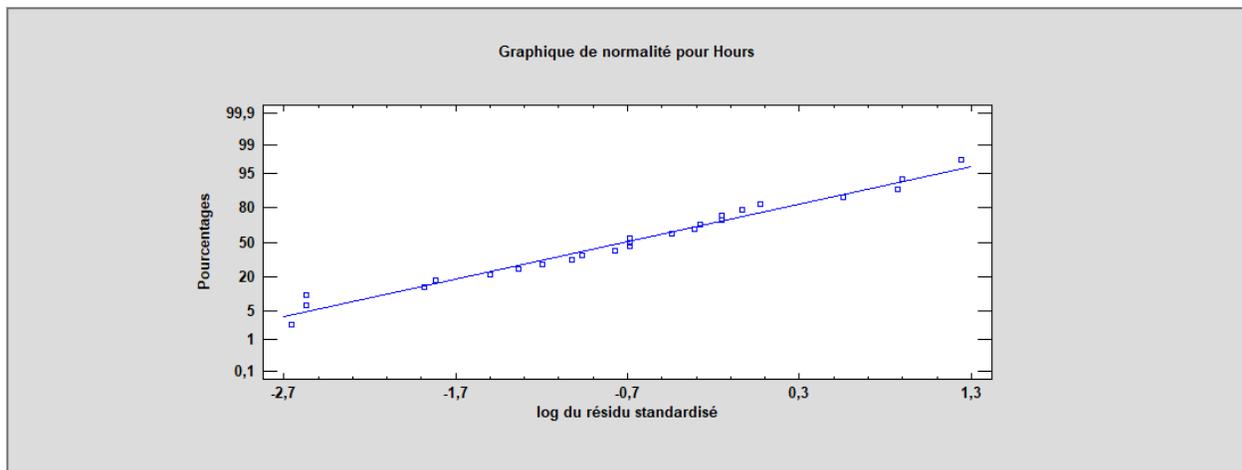
Résidus par rapport aux valeurs prévues

Ce graphique est utile pour aider à visualiser si la variabilité des résidus est constante ou dépend des valeurs prévues.



Graphique de normalité

Ce graphique peut être utilisé pour déterminer si les écarts suivent ou non une loi normale.

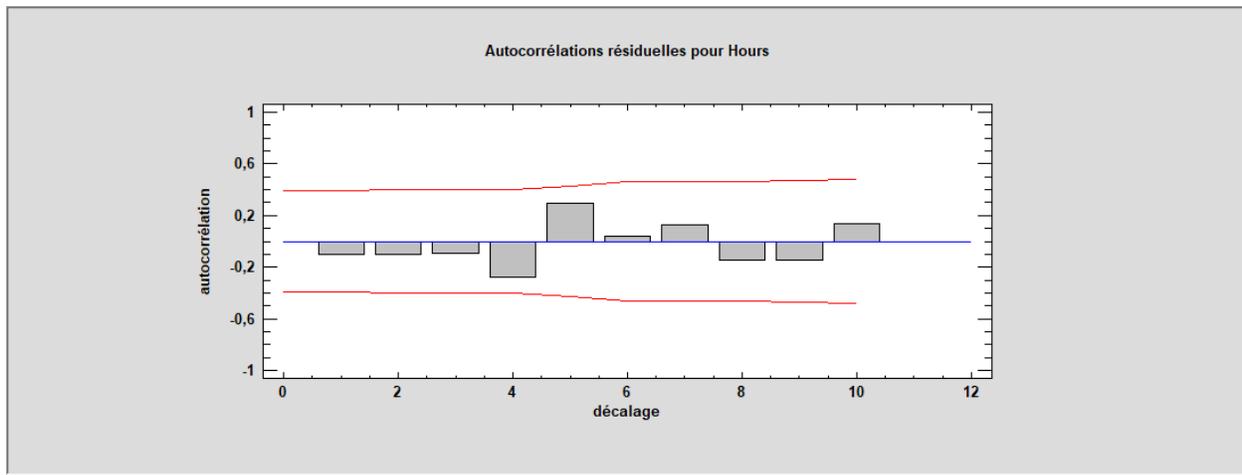


Bien que ce graphique soit affiché dans toutes les procédures de régression, le *Graphique de probabilités des résidus* décrit précédemment est plus utile pour des résidus de données de survie.

Fonction d'autocorrélation

Ce graphique calcule la fonction d'autocorrélation entre les résidus comme une fonction du nombre de lignes entre eux dans le tableur.

Il n'a de sens que si les données ont été collectées dans un ordre séquentiel. Toutes les barres qui s'étendent au-delà des limites des probabilités indiquent des dépendances significatives entre les résidus séparés par les décalages indiqués.



Options pour la fenêtre

Options pour le graphique des résidus ✕

Graphique

Résidus

Résidus standardisés

Orientation

Horizontale

Verticale

OK

Annuler

Aide

Type

Nuage de points

Graphique de normalité

Fonction d'autocorrélation

Ligne ajustée

Aucune

Quartiles

Moindres carrés

Graphique par rapport à:

- Valeurs prévues
- Numéros des observations
- Température
- Voltage

Nombre de décalages:

Niveau de confiance:

- **Graphique** : le type de résidus à afficher.
- **Type** : le type de graphique à créer. Un *Nuage de points* est utilisé pour tester la courbure. Un *Graphique de normalité* est utilisé pour déterminer si les résidus du modèle suivent une loi normale. Une *Fonction d'autocorrélation* est utilisée pour tester la dépendance entre les résidus consécutifs.
- **Graphique par rapport à** : pour un *Nuage de points*, la caractéristique à afficher sur l'axe horizontal.
- **Nombre de décalages** : pour une *Fonction d'autocorrélation*, le nombre maximum de décalages. Pour de petits ensembles de données, le nombre de décalages doit être inférieur à cette valeur.
- **Niveau de confiance** : pour une *Fonction d'autocorrélation*, le niveau utilisé pour calculer les limites des autocorrélations.

Matrice des corrélations

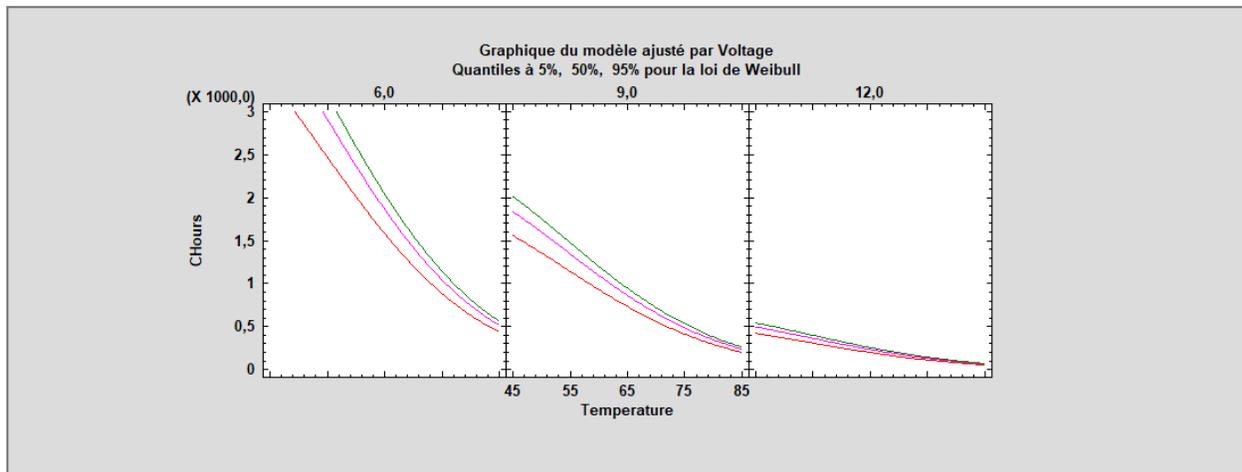
La *Matrice des corrélations* affiche des estimations des corrélations entre les coefficients estimés.

Matrice des corrélations des coefficients estimés					
	CONSTANTE	Temperature	Voltage	Temperature^2	Voltage^2
CONSTANTE	1,0000	-0,6666	-0,7779	0,6561	0,2440
Temperature	-0,6666	1,0000	0,0729	-0,9953	-0,1157
Voltage	-0,7779	0,0729	1,0000	-0,0794	-0,1798
Temperature^2	0,6561	-0,9953	-0,0794	1,0000	0,0912
Voltage^2	0,2440	-0,1157	-0,1798	0,0912	1,0000

Ce tableau peut être utile pour déterminer si les effets des différentes variables explicatives ont bien été séparés les uns des autres.

Graphique treillis

Pour des modèles contenant deux variables explicatives ou plus, un graphique treillis peut être utilisé pour afficher les quantiles estimés par rapport à deux ou trois variables. Le graphique ci-dessous montre comment les quantiles évoluent en fonction du voltage et de la température.



Options pour la fenêtre

Facteurs pour le graphique treillis

X C1 C2 Bas Haut Par Maintenir

Temperature

Voltage

Facteurs de conditionnement

1 facteur (1 ligne)

1 facteur (2 lignes ou +)

2 facteurs

Quantiles:

%

%

%

%

%

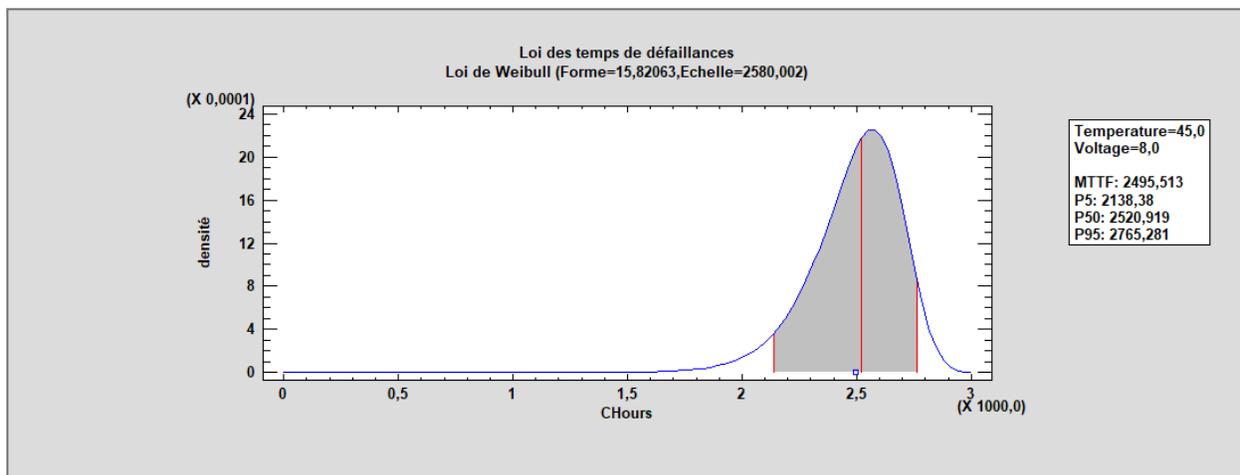
Afficher la moyenne

- **X** : le facteur affiché sur l'axe horizontal du graphique dans chaque segment.
- **C1** : un facteur de conditionnement dont les valeurs définissent les colonnes du graphique treillis.
- **C2** : un facteur de conditionnement dont les valeurs définissent les lignes du graphique treillis.
- **Facteurs de conditionnement** : le nombre de facteurs de conditionnement dans le graphique. Si un unique facteur est sélectionné, plusieurs lignes seront utilisées pour afficher chaque niveau, si souhaité.
- **Bas** : pour un facteur de conditionnement quantitatif, la borne basse de la plage de variation du facteur.

- **Haut** : pour un facteur de conditionnement quantitatif, la borne haute de la plage de variation du facteur.
- **Par** : pour un facteur de conditionnement quantitatif, l'incrément entre les bornes basse et haute.
- **Maintenir** : pour les facteurs non sélectionnés, la valeur à laquelle le facteur doit être maintenu constant.
- **Quantiles** : pourcentages pour les quantiles souhaités.
- **Afficher la moyenne** : inclure une ligne à la moyenne estimée des temps des défaillances.
- **Suivant et Précédent** : boutons à utiliser pour afficher d'autres facteurs lorsqu'il y en a plus de 16.

Graphique de la loi des défaillances

Le graphique de la loi des défaillances peut être affiché pour toute combinaison des facteurs. Un exemple de ce graphique est montré ci-dessous.



Le type de la loi et ses paramètres estimés sont affichés en haut dans le graphique. Le MTTF et les quantiles estimés sont affichés à droite dans le graphique.

Options pour la fenêtre

Niveau : valeurs des facteurs pour lesquelles la loi doit être affichée.

Quantile inférieur : valeur du quantile inférieur à inclure dans le graphique (s'il y en a un).

Quantile supérieur : valeur du quantile supérieur à inclure dans le graphique (s'il y en a un).

Afficher la moyenne : permet d'indiquer si le MTTF doit être affiché ou non.

5. *Résidus* – les n résidus r_i .
6. *Résidus standardisés* – les n résidus standardisés e_i .
7. *Résidus de Cox-Snell* – les n résidus de Cox-Snell \hat{u}_i .
8. *Coefficients* – les coefficients du modèle ajusté.
9. *Pourcentages* – les pourcentages auxquels les quantiles sont calculés.
10. *Quantiles* – les quantiles estimés.
11. *Erreurs-types des quantiles* – les erreurs-types des quantiles estimés.
12. *Limites de confiance inférieures des quantiles* – les limites de confiance inférieures pour les quantiles.
13. *Limites de confiance supérieures des quantiles* – les limites de confiance supérieures pour les quantiles.

Formules des calculs

Lois standardisées

$$\text{Logistique, log-logistique : } \Phi(z) = \exp(z) / [1 + \exp(z)] \quad (13)$$

$$\text{Normale, log-normale : } \Phi(z) = \int_{-\infty}^z (1/\sqrt{2\pi}) \exp(-z^2/2) \quad (14)$$

$$\text{Plus petite valeur extrême, Weibull, exponentielle : } \Phi(z) = 1 - \exp[-\exp(z)] \quad (15)$$

Fonctions de vraisemblance

Soit $\delta_i = 1$ pour un temps exact de défaillance et 0 pour une observation censurée à droite.

Modèles position-échelle :

$$L(\beta, \sigma) = \prod_{i=1}^n \left[\frac{1}{\sigma} \phi\left(\frac{y_i - \mu_i}{\sigma}\right) \right]^{\delta_i} \left[1 - \Phi\left(\frac{y_i - \mu_i}{\sigma}\right) \right]^{1-\delta_i} \quad (16)$$

Modèles log-position et log-échelle :

$$L(\beta, \sigma) = \prod_{i=1}^n \left[\frac{1}{\sigma} \phi\left(\frac{\log(y_i) - \mu_i}{\sigma}\right) \right]^{\delta_i} \left[1 - \Phi\left(\frac{\log(y_i) - \mu_i}{\sigma}\right) \right]^{1-\delta_i} \quad (17)$$

Erreurs-types des coefficients

Ils sont déterminés à partir des dérivées partielles évaluées aux valeurs des estimations du maximum de vraisemblance. Les intervalles de confiance sont basés sur une approximation normale pour de grands échantillons.

Temps moyen des défaillances

<i>Loi</i>	<i>E(Y)</i>
Normale	μ
Log-normale	$\exp(\mu + \sigma^2/2)$
Logistique	μ
Log-logistique	$\exp(\mu)\Gamma(1+\sigma)\Gamma(1-\sigma)$
Plus petite valeur extrême	$\mu-0,5772\sigma$
Weibull	$\exp(\mu)\Gamma(1+\sigma)$
Exponentielle	$\exp(\mu)\Gamma(2)$